TRAITEMENT STATISTIQUE DES DONNÉES EXPÉRIMENTALES

Bruno LECOUTRE

CISIA
ceresta

Bruno LECOUTRE

TRAITEMENT STATISTIQUE DES DONNÉES EXPÉRIMENTALES: DES PRATIQUES TRADITIONNELLES AUX PRATIQUES BAYÉSIENNES

Avec programmes Windows

par Bruno Lecoutre et Jacques Poitevineau

Traitement statistique des données expérimentales :
des pratiques traditionnelles aux pratiques bayésiennes

Avec programmes Windows

Bruno LECOUTRE et Jacques POITEVINEAU

© Copyright 1996, CISIA. Tous droits réservés.

ISBN : 2-906711-10-1

Centre International de Statistique et d'Informatique Appliquées
1, avenue Herbillon, 94160 Saint-Mandé (France)

SOMMAIRE

page

Introduction 7
Chapitre 0 - Quelques éléments de réflexion 13
Chapitre 1 - Du t de Student aux procédures fiducio-bayésiennes 45
Chapitre 2 - De l'ANOVA aux procédures fiducio-bayésiennes 77
Chapitre 3 - Prise en compte d'informations extérieures aux données 101
Chapitre 4 - Analyse spécifique: Plans S<G>*O 121
Chapitre 5 - Analyse spécifique: Illustrations 147
Chapitre 6 - Comparer des moyennes sans supposer l'égalité des variances 193
Chapitre 7 - Compléments sur les procédures fréquentistes 211
Chapitre 8 - Déterminer les effectifs nécessaires 233
Chapitre 9 - Distributions utiles 261
Chapitre 10 - Solutions bayésiennes 295
Installation des programmes 319
LeBayésien 321
LesDistributions 339
Références bibliographiques 343
Table des matières 351
Index 361

INTRODUCTION

L'analyse de variance fait partie des méthodes statistiques les plus utilisées dans les sciences expérimentales. Elle est utilisée pour répondre au problème incontournable de la généralisation des résultats observés sur un échantillon de sujets à une population plus vaste. Cette généralisation repose sur l'usage de procédures d'inférence statistique, et actuellement presque exclusivement sur l'utilisation de "tests de signification" (les fameux *t de Student* et *rapports F*). Le test de signification est devenu progressivement une *norme* indispensable à la communication des résultats expérimentaux. Il relève d'une conception de l'inférence statistique, qui est souvent qualifiée de "classique", mais que nous appellerons ici *traditionnelle* ou *fréquentiste*.

On peut expliquer le succès des tests de signification, à la fois par leur simplicité de mise en œuvre et par leur apparence d'objectivité, qui permet à une large communauté d'admettre la validité des conclusions. En revanche l'expérimentateur n'y trouve pas toujours de réponses précises aux questions qu'il s'était posées en construisant son expérience. La pratique usuelle est d'ailleurs un compromis entre différentes conceptions théoriques, notamment celle de Fisher et celle de Neyman et Pearson; l'existence même de ce compromis est la manifestation des difficultés méthodologiques éprouvées dans la mise en œuvre de l'inférence "classique".

Face à cet usage si répandu, l'inférence statistique *bayésienne* apparaît trop souvent comme un outil très théorique, dont l'intérêt pratique pour l'analyse des données expérimentales reste encore à démontrer. La situation de l'analyse de variance est à cet égard exemplaire: des procédures bayésiennes y ont été développées, mais ne sont généralement pas mentionnées dans les ouvrages de référence, ni incluses dans les logiciels statistiques les plus répandus. Cependant un intérêt sans cesse croissant se porte sur ces procédures, comme le montre la littérature statistique appliquée récente.

Certes les procédures bayésiennes peuvent paraître plus lourdes à mettre en œuvre, en obligeant notamment à expliciter une distribution *a priori*. En revanche elles répondent mieux, et de manière directe, aux questions de l'expérimentateur, en particulier concernant l'importance des effets examinés.

CONTENU DE L'OUVRAGE

L'objectif de cet ouvrage est de proposer à l'utilisateur de l'analyse de variance une approche *pratique*, *réaliste* et *constructive* de l'inférence bayésienne, intégrant les procédures traditionnelles et lui apportant un regard nouveau sur ses données. Il s'agit de fournir des procédures bayésiennes *standard*, aussi objectives et simples à utiliser que tests t ou F familiers, et bien adaptées à l'analyse des plans expérimentaux complexes largement utilisés, tels que les dispositifs avec *mesures répétées* ou croisés (*cross-over*). Ces procédures éclairent les difficultés et les insuffisances des procédures traditionnelles, et renouvellent la méthodologie de l'analyse de variance.

Cette approche s'inscrit dans un cadre général de recherches qui ont été effectuées à Paris dans le *Groupe Mathématiques et Psychologie*, et dont on en trouvera une présentation d'ensemble, en ce qui concerne plus particulièrement l'inférence statistique, dans le livre "*L'inférence statistique dans la démarche du chercheur*" (Rouanet, Lecoutre, Bert, Lecoutre et Bernard, 1991).

Pour garder un volume raisonnable à cet ouvrage, mais aussi pour en rendre le contenu accessible au plus grand nombre de lecteurs, nous avons limité les analyses aux effets unidimensionnels ("sources de variation ou comparaisons de moyennes univariées à un degré de liberté"). Mais des solutions, qui généralisent les procédures présentées ici, existent également pour des effets multidimensionnels, avec prise en compte possible de covariables; elles sont notamment disponibles dans le logiciel **PAC** (Lecoutre et Poitevineau, 1995).

Présentation des procédures à partir de nombreux exemples détaillés

La présentation de toutes les procédures, avec leurs implications méthodologiques, sera effectuée à partir d'exemples réels empruntés à deux domaines d'applications privilégiés, qui sont la *psychologie expérimentale* et les *essais cliniques*. Cette présentation ne fera aucunement appel à des connaissances de spécialistes de ces domaines, et aura donc une portée générale, pouvant être mise à profit par des utilisateurs d'autres domaines.

Programmes informatiques sous Windows

Un programme didactique et convivial, **LeBayésien**, permet la mise en œuvre interactive très simple de toutes les procédures, descriptives et

inférentielles, au fur et à mesure de leur exposé.

Trois autres programmes sont fournis.

• **LesEchantillons** génère des tirages d'échantillons, pour un certains nombre de familles de distributions.

• **LesEffectifs** permet, dans le cadre d'un plan d'expérience donné, de déterminer les effectifs nécessaires.

• **LesDistributions** permet d'obtenir les courbes de densité et les énoncés de probabilité correspondant aux fonctions de répartion directe et inverse de toutes les distributions introduites dans cet ouvrage.

Synthèse des résultats

Même si cet ouvrage poursuit essentiellement un objectif pratique, il nous a paru utile de lui donner un statut autonome, en fournissant également une synthèse technique des résultats utilisés.

PRÉSENTATION DES CHAPITRES

Pour découvrir l'inférence bayésienne: Quelques éléments de réflexion

Le chapitre 0 propose quelques éléments de réflexion sur l'inférence statistique et son utilisation pour l'analyse des données expérimentales. Il fournit une introduction aux procédures d'inférence statistique (pour un non-spécialiste) dans ce domaine, et explicite la conception des méthodes bayésiennes qui nous paraît ici la mieux adaptée.

Pour utiliser les techniques de base: L'inférence fiducio-bayésienne

Les chapitres 1 et 2 présentent sur des exemples les techniques de base pour l'inférence sur une combinaison linéaire de moyennes dans le cadre d'un plan "en groupes indépendants": le "*t* de Student" et l'"ANOVA", et leurs extensions *fiducio-bayésiennes*.

Les autres chapitres peuvent être utilisés de manière relativement indépendante, en fonction des centres d'intérêt du lecteur.

Pour prendre en compte des informations extérieures

Le chapitre 3 illustre la prise en compte d'informations extérieures aux données; il permet de comprendre concrètement comment l'inférence bayésienne permet de combiner des informations.

Pour traiter des plans complexes: L'analyse spécifique

Les chapitres 4 et 5 étendent à des plans d'expérience complexes les techniques de base développées dans les chapitres précédents. Cette extension repose sur le principe raisonnable d'*inférence spécifique*, qui est introduit dans le chapitre 4. De nombreux exemples correspondant à des structures de plans fréquemment utilisées, ou qui posent des problèmes particuliers, sont présentés et discutés dans le chapitre 5.

Pour comparer des moyennes sans supposer l'égalité des variances

Le chapitre 6 apporte des solutions qui permettent d'abandonner les suppositions d'homogénéité des variances des procédures usuelles.

Pour en savoir plus sur les procédures fréquentistes

Le chapitre 7 présente un certain nombre de compléments sur les procédures fréquentistes usuelles, concernant la recherche de conclusions sur l'importance des effets.

Pour choisir les effectifs

Le chapitre 8 développe des solutions générales pour déterminer les effectifs. Ces solutions utilisent les distributions bayésiennes prédictives; elles englobent et étendent les procédures fréquentistes reposant sur la notion de *puissance*.

Pour développer les aspects techniques

Enfin les chapitres 9 et 10 regroupent l'ensemble des résultats utiles pour ceux qui souhaitent développer les aspects techniques des procédures. Un certain nombre de connaissances mathématiques sont ici supposées.

Pour installer et utiliser les programmes Windows

La procédure d'installation et la notice d'utilisation des programmes sont fournies en annexe.

Pour compléter cet ouvrage

Des références bibliographiques complémentaires guident le lecteur, pour découvrir d'autres exemple d'applications, ou pour approfondir des aspects conceptuels ou théoriques des méthodes.

UTILISATION DES PROGRAMMES INFORMATIQUES

L'ouvrage est conçu pour une mise en œuvre interactive de toutes les procédures au fur et à mesure de leur exposé. L'utilisation des programmes est décrite dans des sections délimitées de la manière suivante

LeBayésien

Il suffit d'activer le programme indiqué et de suivre les instructions.

Toutefois l'exposé est rédigé de manière autonome, et il est également possible d'effectuer une lecture indépendante de l'utilisation des programmes, en ignorant les sections correspondantes.

REMARQUES SUR LA TERMINOLOGIE ET LES NOTATIONS

En ce qui concerne la terminologie des distributions bayésiennes, nous préférons, avec de Finetti (1975, page 236), utiliser *initiale* et *finale* plutôt que *a priori* et *a posteriori*.

En ce qui concerne les notations, il est d'usage de distinguer les *variables aléatoires*, souvent désignées par des lettres majuscules, des valeurs prises par ces variables, désignées par des lettres minuscules. En pratique une telle distinction est souvent inutilisable et ne sera pas effectuée en général, dans la mesure où le contexte est suffisamment explicite.

Une simplification qui peut paraître beaucoup plus contestable est celle qui consiste, chaque fois que cela peut se faire sans risque de confusion, à omettre d'indiquer les valeurs qui conditionnent les distributions. Par exemple, il est évidemment essentiel de distinguer la distribution *initiale* sur un paramètre, qui est une distribution *marginale*, de la distribution *finale*, qui est *conditionnelle* aux observations (ce que l'on indique par le symbole "|"). Mais autant cette distinction est indispensable pour un exposé technique, autant elle sera lourde et difficilement utilisable dans la présentation de résultats d'analyses.

Dans le texte les renvois aux chapitres ou sections de chapitre sont indiqués par le signe ↳

CHAPITRE 0

QUELQUES ÉLÉMENTS DE RÉFLEXION

L'objet de ce chapitre préliminaire est d'apporter au lecteur non spécialiste un éclairage pour le guider dans l'utilisation de l'inférence bayésienne (et éventuellement de l'inférence traditionnelle...). L'exposé sera donc nécessairement schématique et incomplet, avec toutes les insuffisances que cela peut comporter s'agissant d'un domaine aussi délicat que celui du raisonnement inductif.

La première partie de ce chapitre est une introduction aux procédures d'inférence statistique pour l'analyse des données expérimentales; elle explicite la conception des méthodes bayésiennes qui nous paraît la mieux adaptée dans ce domaine. La deuxième partie constitue une annexe, qui rappelle quelques notions et résultats généraux.

0.1 De l'inférence traditionnelle à l'inférence bayésienne

Nous illustrerons notre propos à partir du problème relativement simple, mais à valeur d'exemple général, de la comparaison de deux moyennes de groupes indépendants (problème du "t de Student"). Après avoir rappelé le concept fondamental de distribution d'échantillonnage, nous présenterons d'abord le test de signification traditionnel, puis nous introduirons l'inférence bayésienne et montrerons comment elle permet d'évaluer de façon directe et naturelle l'importance réelle d'un effet. Nous mettrons ensuite plus particulièrement en avant la différence fondamentale (et irréductible) entre l'interprétation fréquentiste et l'interprétation bayésienne de l'intervalle de confiance. Nous mentionnerons enfin quelques apports de l'inférence bayésienne.

0.1.1 La distribution d'échantillonnage

Considérons l'exemple suivant emprunté à la pharmacologie. On désire comparer deux traitements A et B, par exemple un anti-hypertenseur et un *placebo*. Pour cela, on réalise un essai clinique dans lequel on observe un échantillon de n patients. Ces patients sont supposés extraits au hasard dans une population de patients hypertendus. Par tirage au sort, n/2 patients reçoivent le traitement A, et les n/2 autres le traitement B. Le critère principal d'évaluation de l'efficacité du traitement est la chute de tension évaluée chez chaque patient comme la différence des tensions systoliques prises avant et après administration du traitement.

Supposons que tous les patients hypertendus de la population soient traités avec le traitement A (respectivement B), et désignons par μ_A (respectivement μ_B) la moyenne réelle inconnue des chutes de tension dans cette population. Alors la différence $\delta=\mu_A-\mu_B$ représente l'effet réel de la comparaison des deux traitements A et B.

■ *Des résultats de base*

Soit D la différence observée entre les deux moyennes des chutes de tension dans chacun des deux groupes de traitement. D est une variable *aléatoire*: si on considère différents échantillons possibles de n patients, D varie d'un échantillon à l'autre. La distribution des valeurs de D, sur l'ensemble de tous les échantillons possibles que l'on peut constituer à partir de la population, est appelée la "distribution d'échantillonnage de D". Cette distribution dépend bien entendu de la différence δ dans la population, mais aussi d'autres caractéristiques, ou paramètres de cette population (tels que variance...).

On montre que la moyenne de la distribution d'échantillonnage de D est égale à δ

$$\text{Moy}(D) = \delta$$

et que sa variance, que nous noterons ε^2, dépend des variances σ_A^2 et σ_B^2 dans la population et de l'effectif n (nous nous en tenons ici à la supposition, faite dans les procédures usuelles, d'une taille de la population infinie)

$$\text{Var}(D) = \varepsilon^2 = 2\frac{\sigma_A^2+\sigma_B^2}{n}$$

L'écart-type ε de la distribution d'échantillonnage traduit l'imprécision expérimentale.

■ *Remarque*

La situation présentée ici est simplifiée. D'une part, si on pouvait répéter la mesure de la tension d'un patient dans des conditions identiques (pour les facteurs contrôlés dans l'expérience), on obtiendrait sans doute un résultat différent dû à la variablité *intra-sujets* et à l'imprécision sur la mesure de la tension ("erreur de mesure"). D'autre part il peut exister d'autres sources d'imprécision liées à des facteurs qui ne sont pas contrôlés. Ceci contribue évidemment à l'imprécision expérimentale ε, mais est sans incidence sur l'exposé des principes d'inférence.

■ *Un concept fondamental*

La distribution d'échantillonnage est l'outil fondamental utilisé par toutes les procédures d'inférence statistique. Si l'imprécision expérimentale ε est très grande (à la limite infinie), aucune inférence n'est possible; si au contraire elle est extrêmement faible (à la limite nulle), nous pouvons pratiquement assimiler l'échantillon à la population, et l'inférence est immédiate. En général on essaye de contrôler la valeur de ε en choisissant un effectif n approprié.

Intuitivement, si nous tirons *au hasard* un échantillon de n patients, la différence entre les moyennes observées ne devrait pas être "trop éloignée" de δ (dans la mesure où l'imprécision expérimentale ε est limitée par le choix de n). Ce raisonnement peut être quantifié, si nous précisons la distribution d'échantillonnage par l'intermédiaire d'un modèle. Nous nous en tiendrons ici encore aux suppositions classiques des procédures usuelles, et nous considérerons que la distribution d'échantillonnage est une distribution *normale*, ce que nous noterons

$$D|\delta,\sigma_A^2,\sigma_B^2 \sim N(\delta,\varepsilon^2)$$

En fait, pour l'exposé des principes d'inférence sur δ, il nous suffira de considérer que l'imprécision expérimentale ε (donc ici la somme des variances $\sigma_A^2+\sigma_B^2$) est connue.

Le passage au cas des variances inconnues pourra être regardé comme une simple modification technique, consistant à utiliser les variances observées et à remplacer la distribution *normale* par celle de *Student*. En réalité la situation est bien entendu plus complexe, surtout si nous ne supposons pas l'égalité des deux variances; et la manière d'"éliminer" les variances "parasites" est elle-même une source de divergence fondamentale entre l'approche traditionnelle et l'approche bayésienne, comme nous le verrons au chapitre ↳6.

La figure 0.1 représente la "courbe de densité" de la distribution d'échantillonnage de D. Si nous tirons un échantillon au hasard, la probabilité que la différence observée pour cet échantillon soit comprise entre a et b est la proportion des échantillons pour lesquels la valeur de D est comprise entre a et b. Elle correspond à l'aire marquée sous la courbe de densité. Bien entendu cette distribution, et donc les probabilités d'échantillonnage, restent hypothétiques, puisque la vraie valeur δ est inconnue; on a en fait une famille de distributions d'échantillonnage, indexée par l'ensemble (infini) des valeurs possibles de δ.

0.1.2 L'inférence statistique traditionnelle: le test de signification

L'ambition (la prétention?) de l'inférence statistique traditionnelle est de fonder ses procédures sur les seules propriétés de la distribution d'échantillonnage. Les probabilités considérées sont uniquement des probabilités d'échantillonnage. Elles s'interprètent donc comme nous l'avons vu comme des proportions, ou des fréquences, d'où l'appellation d'*inférence fréquentiste* qui s'applique à l'inférence traditionnelle (on parle encore de "procédures d'échantillonnage").

La procédure la plus répandue est le test de signification, utilisé pour mettre à l'épreuve l'hypothèse ("nulle") d'absence d'effet du traitement [H_0:δ=0] contre l'hypothèse ("alternative") unilatérale de présence d'un effet [H_1:δ>0]. Ce test repose sur un raisonnement indirect, *hypothético-déductif*: intuitivement, sous l'hypothèse [H_0:δ=0], la différence entre les moyennes observées ne devrait pas être "trop éloignée de 0"; donc, si tel n'est pas le cas, cette hypothèse devient "suspecte".

Si nous avons observé la différence d_{obs} (que nous supposerons positive) sur l'unique échantillon disponible, alors l'aire marquée dans la figure 0.2 représente la proportion p_{sup} des échantillons pour lesquels la valeur de D est supérieure à d_{obs} (c'est-à-dire est plus extrême que la valeur observée), si l'hypothèse d'absence d'effet du traitement est supposée vraie. Formellement

$$p_{sup} = Pr(D>d_{obs} \mid [H_0:\delta=0])$$

On remarquera que nous définissons ici le test unilatéral (*one-sided*) qui est le plus naturel. Mais on utilise souvent le test bilatéral (*two-sided*) qui considère la proportion p des échantillons pour lesquels la valeur de D est supérieure à d_{obs} en valeur absolue (soit ici, puisque la distribution est symétrique, $p=2p_{sup}$ ou $p_{sup}=p/2$). Cette distinction est en fait accessoire, et ce qui suit s'applique de la même manière à p_{sup} ou à p.

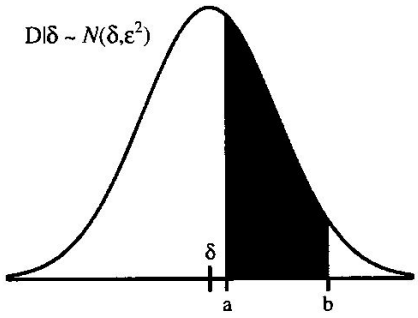


Figure 0.1 - Distribution d'échantillonnage de D

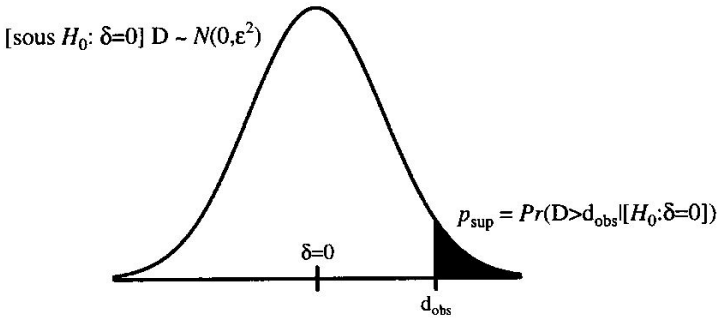


Figure 0.2 - Test de signification: Distribution d'échantillonnage de D sous l'hypothèse nulle [H_0: δ=0]

• Si cette proportion p_{sup} d'échantillons est faible (par exemple plus petite que 5%, comme c'est le cas dans la figure 0.2), cela peut être interprété de deux façons: soit, l'hypothèse $[H_0:\delta=0]$ est vraie et l'événement observé est peu fréquent (moins de 5% des échantillons présentent une valeur plus extrême); soit l'hypothèse $[H_0:\delta=0]$ est fausse. Dans ce cas nous retenons la deuxième interprétation, et nous concluons au "rejet" de l'hypothèse nulle et, par conséquent à l'"acceptation" de l'hypothèse $[H_1:\delta>0]$. Nous disons que le test est *significatif*.

• Si la proportion p_{sup} n'est pas faible (généralement supérieure à 5%), alors l'hypothèse $[H_0:\delta=0]$ ne peut pas être rejetée. Nous disons que le test est *non-significatif*.

Ce test de signification usuel a fait l'objet de nombreuses critiques. Nous en énonçons quelques unes.
1) Il fonde la conclusion de rejet de l'hypothèse nulle sur la probabilité d'événements qui n'ont pas été observés, et qui sont en outre très rares ("les échantillons extrêmes") sous cette hypothèse.
2) Il conduit à une réponse dichotomique, catégorique, alors que la réponse devrait être graduelle.
3) Un résultat significatif met en évidence l'existence d'un effet ($\delta>0$) mais ne donne pas d'indication sur l'importance réelle de cet effet. Le test approprié devrait inclure la borne minimale Δ réellement signifiante, et tester l'hypothèse nulle $[H_0:\delta\leq\Delta]$ contre l'hypothèse $[H_1:\delta>\Delta]$. Mais ceci est trop rarement envisagé dans la pratique.
4) À l'inverse un résultat non-significatif, qui conduit au non rejet de l'hypothèse nulle, n'est en aucune manière une preuve de la validité de cette hypothèse, mais est simplement un *constat d'ignorance*.

Ce même test usuel donne lieu à un certain nombres d'interprétations abusives; mais celles-ci ne sont que le reflet de ses insuffisances. En particulier la valeur p_{sup}, appelée *seuil de signification* (*p-value* en anglais), est souvent interprétée de façon erronée comme "la probabilité de se tromper en rejetant l'hypothèse nulle" alors qu'elle est la probabilité conditionnelle de se tromper en rejetant l'hypothèse nulle SI CETTE HYPOTHÈSE EST VRAIE ("erreur ou risque de première espèce"). La valeur du seuil de signification est également souvent utilisée, de façon non justifiée, comme une indication de l'importance réelle de l'effet.

Par ailleurs, il est encore recommandé de calculer la probabilité, conditionnelle à l'hypothèse d'une valeur $\delta\neq0$, de ne pas rejeter H_0 ("erreur ou risque de deuxième espèce"); mais ceci est rarement fait dans la pratique.

0.1.3 L'inférence bayésienne

■ *La distribution finale*

La distribution bayésienne *finale* ou *a posteriori* (c'est-à-dire après le recueil des données) probabilise les valeurs possibles du paramètre δ (ce qui n'est pas permis par la statistique fréquentiste) et exprime directement l'incertitude sur δ. Deux exemples de cette distribution bayésienne sont montrés dans la figure 0.3: ici par exemple il y a respectivement 72% et 96% de chances pour que la vraie valeur δ appartienne à l'intervalle [5,25]; cette probabilité est donnée par l'aire marquée. Ainsi, pour la différence entre le traitement et le *placebo*, si 5 et 25 sont des valeurs cliniques pertinentes, on obtient directement la probabilité que la vraie valeur δ soit intéressante sur le plan clinique.

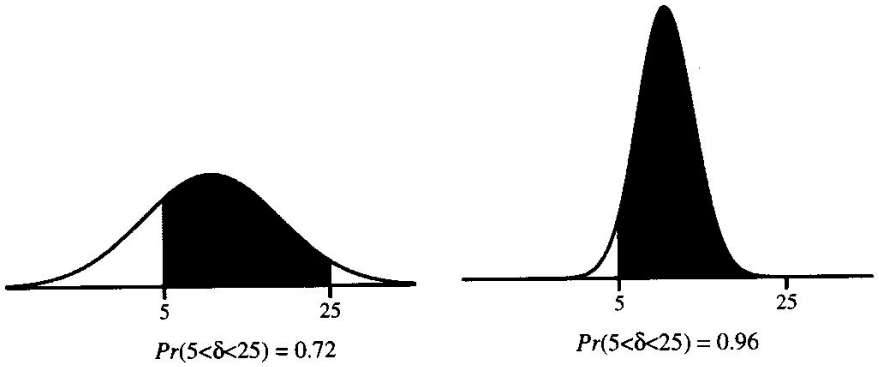


Figure 0.3 - Exemples de deux distributions bayésiennes finales relatives à δ

■ *Comment construire la distribution finale?*

La distribution finale est construite en appliquant le théorème de Bayes qui permet de la dériver des sources d'informations disponibles:
• les données, dont nous supposerons ici qu'elles sont entièrement résumées par la valeur observée d_{obs} (on dit dans ce cas que D est une statistique *exhaustive*);

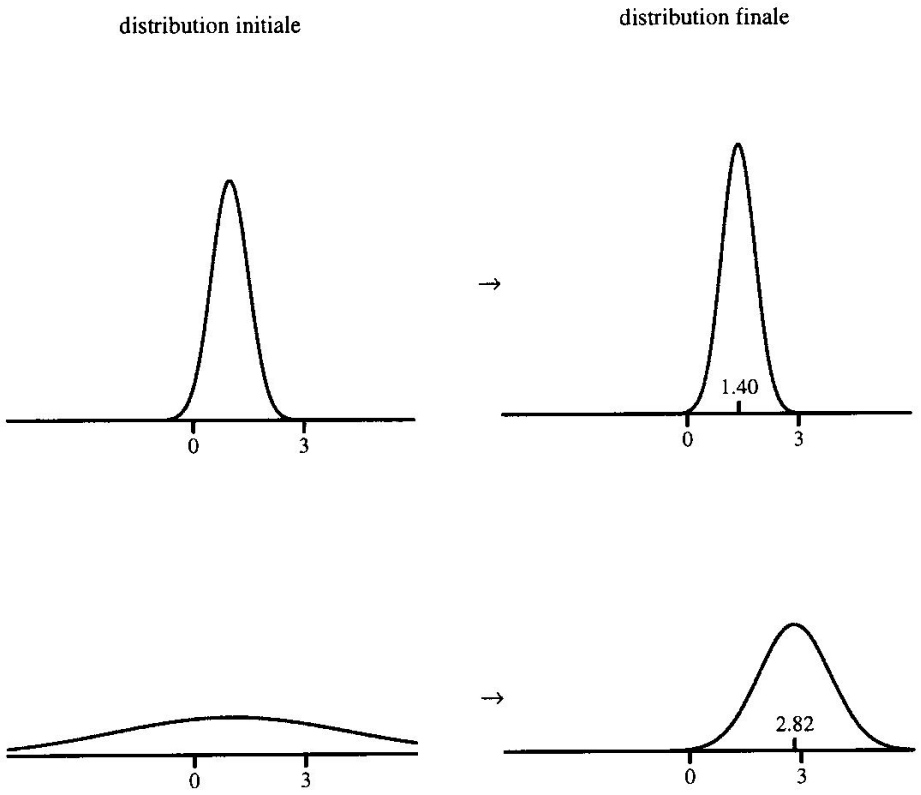


Figure 0.4 - Influence de la distribution initiale sur la distribution finale
Exemple: $d_{obs} = 3$ et $\varepsilon = 1$

• le modèle d'échantillonnage, c'est-à-dire la famille des distributions d'échantillonnage, conditionnellement aux valeurs possibles de δ;
• la distribution *initiale* ou *a priori*, qui traduit l'état d'incertitude sur δ avant le recueil des données.

Plus précisément, nous avons pour chaque valeur possible de δ (rappelons que ε est considéré comme une donnée)

$$p(\delta | D=d_{obs}) = k \times p(D=d_{obs} | \delta) \times p(\delta)$$

où

• $p(\delta | D=d_{obs})$ est la densité de la distribution finale pour la valeur δ considérée;
• $p(D=d_{obs} | \delta)$ est la densité de la distribution d'échantillonnage pour la valeur observée d_{obs} de D conditionnellement à la valeur δ considérée; on notera que $p(D=d_{obs} | \delta)$ est souvent regardée comme fonction de δ (pour la valeur d_{obs} fixée), et qu'on l'appelle alors la *fonction de vraisemblance*;
• $p(\delta)$ est la densité de la distribution initiale pour la valeur δ considérée;
• k est une constante qui assure que la probabilité totale est bien égale à 1.

On voit ainsi comment la distribution bayésienne finale combine l'information initiale avec l'information apportée par les données. Il est clair, comme le montre la figure 0.4, que la distribution finale peut être fortement influencée par la distribution initiale. Nous nous limitons ici à des exemples de distributions initiales *normales* (la distribution finale est encore *normale*); mais on peut utiliser en théorie n'importe quelle distribution.

■ *La distribution prédictive*

En outre nous pouvons calculer la probabilité *marginale* relative à la statistique D. On obtient ainsi une distribution *prédictive* qui tient compte des deux sources d'incertitude, celle liée à l'imprécision expérimentale ε, et celle liée à la valeur de δ exprimée par la distribution initiale. Cette distribution permet, par un jugement probabiliste sur des observations futures, de répondre directement à des questions d'intérêt essentiel pour l'expérimentateur, comme: "quelle est la probabilité d'obtenir tel résultat dans l'expérience envisagée, en fonction des effectifs?".

■ *La conception "subjective"*

Suivant une certaine conception bayésienne *subjective*, les distributions initiales devraient incorporer toutes les connaissances et opinions *a priori*

sur les paramètres de la situation étudiée. La possibilité d'intégrer des informations disponibles avant expérience, notamment des résultats antérieurs, ou encore par exemple l'opinion d'un comité de spécialistes dans le domaine étudié, constitue un apport potentiel considérable.

Par ailleurs cette conception *subjective* se situe généralement dans une perspective *décisionnelle*, dans laquelle l'inférence statistique doit fournir "*un critère d'évaluation, qui décrit les conséquences de chaque décision en fonction des paramètres du modèle*" (Robert, 1992, page 41).

Ces perspectives sont évidemment séduisantes, mais, dans le domaine de l'analyse des données expérimentales, elles impliquent une rupture avec les usages actuels, qui dans la plupart des cas répondent à un besoin d'objectivité. C'est sans doute là une des raisons de la méfiance, voire de l'hostilité, de certains utilisateurs envers la statistique bayésienne.

En tout état de cause, on trouvera dans l'inférence bayésienne la possibilité d'exprimer de manière comparable et cohérente différentes sources d'information, ceci quelle que soit l'attitude qu'on peut avoir quant à l'intérêt de combiner effectivement ces informations.

■ *La conception "non-informative": Procédures bayésiennes standard*

Nous pensons en fait que l'utilisation des procédures bayésiennes ne doit pas impliquer une rupture avec les pratiques fréquentistes actuelles: cette conception est en effet la seule réaliste, face à l'usage si répandu des tests de signification. Pour le moins les procédures bayésiennes devront coexister de nombreuses années avec les procédures fréquentistes.

Les procédures bayésiennes doivent donc intégrer, étendre et raffiner - plutôt que remplacer - les pratiques actuelles. Dans cette optique, nous accorderons une place prépondérante à une autre conception, non moins bayésienne, illustrée notamment par Jeffreys (1961) - pour ne pas remonter à Bayes et surtout Laplace - pour laquelle les distributions initiales peuvent être utilisées pour exprimer un "état d'ignorance" (nous dirions plutôt un état d'*indifférence*) sur les paramètres. De telles distributions initiales ont souvent été appellées "non-informatives"; elles ont donné (et donnent toujours) lieu à d'intéressants débats théoriques. Elles consistent, schématiquement, à ne pas privilégier *a priori* de valeurs particulières du paramètre. Du point de vue de l'utilisateur, ces distributions peuvent donc être considérées comme des distributions "vagues", n'introduisant pas d'information extérieure aux données (ce qui ne veut pas dire que ces informations n'existent pas).

La distribution finale correspondante exprime alors l'apport propre des données; cela lui confère le statut d'objectivité généralement recherché pour la communication des résultats expérimentaux. Les procédures bayésiennes non-informatives ont donc un statut privilégié et sont maintenant reconnues comme un standard. Par exemple, Berger (1985) nous dit:

"We should indeed argue that noninformative prior Bayesian analysis is the single most powerful method of statistical analysis" (page 90);
"At the very least, use of noninformative priors should be recognized as being at least as objective as any other statistical techniques" (page 110).

0.1.4 La méthode fiducio-bayésienne

Il fallait donner un nom à cette conception et aux distributions finales issues de distributions initiales non-informatives. Nous proposons l'appellation

fiducio-bayésienne

En effet, dans les problèmes de comparaisons de moyennes, tels qu'ils se présentent en analyse de variance, cette conception rejoint l'approche fiduciaire (*fiducia* = confiance) développée par Fisher dans les années trente (voir en particulier Fisher, 1990). La motivation de l'approche fiduciaire est d'exprimer l'apport propre aux données ("ce que les données ont à dire") par un jugement probabiliste, en laissant de côté toute information extérieure. On peut ainsi dire que l'inférence fiducio-bayésienne est

bayésienne dans sa *technique* et **fiduciaire** dans sa *motivation*

Il faut cependant savoir que l'approche fiduciaire, en tant que théorie de l'inférence statistique, est considérée comme un *échec* par de nombreux statisticiens. Y faire référence peut donc provoquer des réactions hostiles! Si on souhaite éviter toute polémique, on pourra utiliser l'appellation

bayésienne standard

qui est de plus en plus répandue pour désigner les solutions non-informatives. Mais on perd là l'intérêt de définir clairement une position méthodologique vis-à-vis de l'utilisation de l'inférence statistique, tout en précisant le domaine d'application (les données expérimentales) dans lequel on se situe. En outre, si on ne peut que reconnaître les limites de l'inférence fiduciaire, il n'en reste pas moins que, **dans les situations particulières considérées ici**, sa solution repose sur un argument de symétrie entre les statistiques et les

paramètres si naturel et si manifeste qu'on ne voit pas comment on pourrait admettre une autre solution *standard*.

Dans la pratique, les distributions fiducio-bayésiennes fourniront des probabilités de référence, à "usage public", qui seront un moyen de communication des résultats, à la fois synthétique et objectif.

Ces probabilités pourront parfaitement différer le cas échéant des probabilités "personnelles" qu'on obtiendrait en incorporant des connaissances (voire des préjugés) étrangères aux données traitées. Rien n'empêche d'ailleurs, dans une étape ultérieure, de prolonger les analyses fiducio-bayésiennes, en y incorporant ces connaissances.

■ *Liens avec les procédures fréquentistes traditionnelles...*

Dans la situation considérée ici, la distribution fiducio-bayésienne pour δ est une distribution *normale*, centrée sur la valeur observée d_{obs} et d'écart-type ε. Cette distribution est montrée dans la figure 0.5.

Ce résultat établit le lien conceptuel qui existe dans cette situation entre la procédure bayésienne standard et le test de signification usuel. En effet p_{sup} n'est autre que la probabilité fiducio-bayésienne que la différence δ soit négative (dans le cas d'une différence d_{obs} positive), ceci pour des raisons de symétrie manifestes, qui apparaissent en comparant les figures 0.2 et 0.5. Cette réinterprétation montre que, dans le cadre bayésien, p_{sup} peut être *légitimement* considéré comme la probabilité de se tromper en concluant "$\delta>0$".

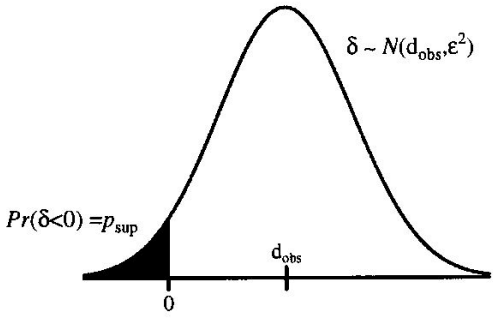


Figure 0.5 - Distribution fiducio-bayésienne pour δ et réinterprétation du seuil unilatéral $p_{sup}=p/2$

Cette réinterprétation assure le lien conceptuel et technique entre l'inférence fiducio-bayésienne et les procédures fréquentistes. Dans les situations de base qui se présentent en analyse de variance, elle fournira à l'utilisateur un passage particulièrement aisé à la méthode fiducio-bayésienne.

Par ailleurs cette réinterprétation éclairera aussi les insuffisances méthodologiques des usages actuels: on voit clairement dans l'exemple précédent que le seuil de signification ne permet une conclusion que sur le signe, et en aucun cas sur l'importance réelle de l'effet.

■ *...et divergences*

Mais, dès que l'on dépasse les situations de base usuelles, on trouve de nombreux cas où l'inférence fiducio-bayésienne et les procédures fréquentistes diffèrent en théorie de manière irréconciable. Un exemple bien connu est celui du problème, dit de Behrens-Fisher, de la comparaison de deux moyennes de groupes indépendants, sous le modèle *normal* avec variances quelconques. L'origine de cette divergence est claire: l'inférence bayésienne est conditionnelle aux variances observées des deux groupes, et l'inférence fréquentiste est conditionnelle aux variances vraies du modèle. La divergence se manifeste dès lors que l'on ne suppose pas l'égalité de ces deux dernières variances.

■ *Les apports méthodologiques de l'inférence fiducio-bayésienne*

Les remarques précédentes pourraient déjà suffire à justifier l'utilisation de la méthode fiducio-bayésienne (et bayésienne en général) en analyse de variance. D'abord elle apporte une interprétation naturelle des procédures, directement en termes de probabilités sur les paramètres. Ensuite elle fournit des solutions dont la justification théorique apparaît beaucoup plus satisfaisante.

Il ne sera pas question ici de revenir sur les fondements de l'inférence bayésienne, ni sur les justifications des distributions non-informatives et les débats qui les ont accompagnées. Nous renvoyons pour cela le lecteur aux ouvrages cités dans la bibliographie. Mais ce qui paraîtra sans doute essentiel à l'utilisateur, ce sont ses apports méthodologiques.

0.1.5 L'inférence bayésienne permet d'évaluer de façon directe l'importance d'un effet

Les deux situations suivantes sont particulièrement exemplaires.

■ *Recherche d'une conclusion d'effet notable (efficacité)*

On compare un anti-hypertenseur A et un *placebo* B et on désire démontrer l'efficacité clinique du traitement. Il s'agit de chercher à conclure que la différence réelle δ entre A et B est supérieure à un écart minimal notable, c'est-à-dire ici signifiant (pertinent) sur le plan clinique. Cet écart minimal peut être défini avant d'avoir observé les données, c'est-à-dire dans le protocole préexpérimental; supposons qu'il soit 5 mm de mercure.

Après le recueil des données, nous déduisons par exemple de la distribution finale que

$$Pr(\delta > 5 \text{ mm Hg}) = 0.82$$

En clair, nous avons une probabilité 0.82 ("82% de chances") que la différence réelle δ soit au moins 5 mm de mercure. Cet énoncé fournit immédiatement une appréciation de l'importance réelle de la différence. Nous pouvons comparer la probabilité obtenue à une *garantie* γ définie avant l'expérience (par exemple γ=0.95). Dans cet exemple, l'expérience n'est pas entièrement concluante (0.82<0.95).

■ *Recherche d'une conclusion d'effet négligeable (équivalence)*

On compare cette fois deux anti-hypertenseurs concurrents A et B et on désire montrer l'équivalence entre les deux traitements. Il s'agit de chercher à conclure que la différence réelle δ entre A et B est inférieure à un écart clinique négligeable (c'est-à-dire ici non signifiant sur le plan clinique), disons ici 3 mm de mercure. Nous calculons alors la probabilité finale

$$Pr(-3 \text{ mm Hg} < \delta < +3 \text{ mm Hg})$$

Si celle-ci est supérieure à une garantie donnée, par exemple 0.95, il y a "beaucoup de chances" pour que la différence réelle soit négligeable et que les deux traitements aient des effets similaires ("équivalents").

0.1.6 Comparaison de l'inférence fréquentiste et de l'inférence bayésienne

■ *Quelle est la différence essentielle entre les deux types de raisonnements?*

Dans l'inférence fréquentiste, seules les valeurs possibles de la grandeur D observable sont probabilisées, et tous les énoncés de probabilité sont conditionnels au paramètre δ. Ainsi le test de signification usuel est basé sur la distribution d'échantillonnage de la statistique de test conditionnellement à une valeur de δ fixée par l'hypothèse nulle.

Dans l'inférence bayésienne, les valeurs possibles du paramètre δ sont également probabilisées, et on obtient directement des énoncés de probabilité relatifs à δ conditionnellement aux données observées.

Cette différence entre les deux types d'inférence est bien illustrée par les interprétations fréquentiste et bayésienne d'un intervalle de confiance (cet intervalle est nommé intervalle de *crédibilité* dans le cadre bayésien).

■ *Interprétations fréquentiste et bayésienne de l'intervalle de confiance*

Dans la construction de l'intervalle de confiance fréquentiste pour le paramètre δ, les bornes de cet intervalle sont des grandeurs aléatoires, qui varient d'un échantillon à un autre. Par exemple, pour trois échantillons différents, les bornes de l'intervalle de confiance seront respectivement

$$[-2.1,+7.4] \qquad [-0.4,+8.9] \qquad [+2.8,+12.7]$$

Dans l'inférence fréquentiste, les bornes particulières observées pour l'échantillon (unique) dont on dispose ne sont interprétables qu'en référence à l'ensemble de tous les intervalles observables. L'interprétation (fréquentiste) correcte de l'intervalle de confiance est alors la suivante: 95% des intervalles calculés sur l'ensemble des échantillons possibles contiennent la vraie valeur δ.

Dans l'inférence fréquentiste, rappelons-le encore, les valeurs possibles du paramètre ne peuvent pas être probabilisées: ainsi si les bornes observées sont [2.1,8.3], l'événement "$2.1<\delta<8.3$" est vrai ou faux, et nous ne pouvons pas lui attribuer de probabilité (sinon 1 ou 0). Il est donc illégitime d'écrire "$Pr(2.1<\delta<8.3)=0.95$". Or l'interprétation communément répandue de l'intervalle de confiance est précisément d'énoncer que "il y a 95 % de chances que le paramètre inconnu δ soit compris entre 2.1 et 8.3". En énonçant cela et en tentant de situer directement le paramètre par rapport à des bornes fixées observées, on fait de l'inférence bayésienne sans le savoir!

Or il se trouve dans le cas présent que la probabilité fiducio-bayésienne (conditionnelle aux données) $Pr(2.1<\delta<8.3)$ est précisément égale à 0.95. Dans le cadre bayésien, l'interprétation précédente devient donc parfaitement légitime. La probabilité 0.95 est conditionnelle aux bornes effectivement observées, considérées cette fois comme fixées.

De plus, dans le cadre bayésien, on peut calculer la probabilité associée à un intervalle de crédibilité dont les bornes sont spécifiées indépendamment des données, tels que $Pr(\delta>+5)$ ou $Pr(-3<\delta<+3)$, et en donner directement une interprétation signifiante (voir les exemples ci-dessus).

0.1.7 En Conclusion

En conclusion, nous esquisserons quelques remarques qui pourront guider et motiver le lecteur pour découvrir l'inférence bayésienne.

- *L'inférence bayésienne n'est pas récente*

L'auto-appellation de "classique" que s'attribue l'inférence fréquentiste ne doit pas faire oublier que la théorie bayésienne est loin d'être récente. Les premiers travaux remontent à Thomas Bayes (1701-1761), mais on doit surtout à Laplace (1749-1827) la démonstration du théorème de Bayes dans un cadre plus général (réalisée, semble-t-il, indépendamment des travaux de Bayes).

- *Elle apparaît supérieure sur le plan théorique*

Les développements théoriques récents établissent la supériorité de l'approche bayésienne (Berger, 1985; Robert, 1992; Bernardo et Smith, 1994...). Ils démontrent notamment que celle-ci recouvre les notions traditionnelles d'optimalité (minimaxité, admissibilité, invariance), et qu'un estimateur optimal du point de vue de la théorie fréquentiste est un estimateur bayésien ou la limite d'estimateurs bayésiens. Enfin, la théorie bayésienne élimine ou surmonte de nombreuses difficultés techniques et de nombreux paradoxes, qui surgissent dans l'approche fréquentiste.

- *Elle est une inférence naturelle*

L'inférence statistique est fondamentalement une démarche d'inversion puisqu'elle vise à remonter des effets aux causes: des données observées aux paramètres. Paradoxalement, ce sont peut être les interprétations "illégitimes" du seuil de signification et de l'intervalle de confiance fréquentiste, en termes de probabilité sur les paramètres, qui rendent ces procédures populaires! En réalité, seule la théorie bayésienne réalise cette inversion de façon légitime et cohérente: partant des données observées et d'un état d'incertitude initial sur le paramètre, le théorème de Bayes permet d'exprimer la nouvelle incertitude sur les valeurs possibles du paramètre. L'inférence bayésienne apporte une analyse quantitative naturelle sur les valeurs possibles du paramètre, qui peut être regardée comme un prolongement direct des procédures descriptives.

- *Elle va devenir de plus en plus facilement utilisable*

Dans les situations courantes, telles que celles rencontrées usuellement en analyse de variance, des procédures bayésiennes standard ont été développées

et peuvent maintenant être mises en œuvre aussi facilement que les tests traditionnels, comme nous le verrons dans la suite de cet ouvrage.

Dans des situations plus complexes, les obstacles nécessités par la réalisation de calculs numériques devraient progressivement disparaître avec le développement d'outils (programmes et ordinateurs de plus en plus sophistiqués) appropriés.

0.2 Quelques notions et résultats préliminaires

Cette partie se veut introductive et intuitive; elle ne prétend pas à la rigueur mathématique. En particulier la notion de distribution d'échantillonnage sera illustrée dans le cas de distributions discrètes (prenant un nombre fini de valeurs) et nous éluderons le problème du passage au cas de distributions continues, qui pourra être regardé simplement comme une situation limite.

0.2.1 Distribution d'échantillonnage dans une population

■ *Population et distribution parentes*

Considérons une *population* d'effectif N, par exemple les N=8 enfants inscrits à un cours de solfège, et une certaine variable numérique, par exemple l'âge (en années entières) de ces enfants

6 7 5 6 7 4 7 6

d'où la *distribution des fréquences* de la variable *âge* dans cette population, figurée ci-après

1/8	1/8	3/8	3/8
4	5	6	7

$moy = \mu = 6 \quad var = \sigma^2 = 1$

Conformément aux usages de l'inférence statistique, nous qualifierons de *parentes* la population et la distribution correspondante, et nous utiliserons des lettres *grecques* pour désigner les caractéristiques de la distribution

parente, notamment
- sa moyenne $\mu = 6$
- son écart-type $\sigma = 1$ et sa variance $\sigma^2 = 1$

■ *Distribution d'échantillonnage de la moyenne*

Supposons que nous choisissions un *échantillon* d'effectif n, par exemple ici que nous choisissions n=2 enfants.

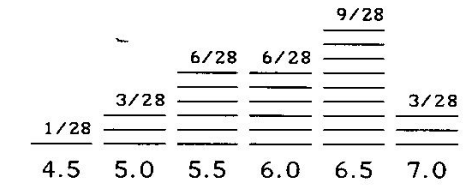
Cet échantillon est ici choisi parmi un ensemble de 28 échantillons possibles, nombre que l'on obtient par la formule

$$\binom{N}{n} = \frac{N!}{(N-n)!n!} = \frac{8!}{(8-2)!2!} = 28$$

Nous pouvons facilement générer les 28 échantillons de taille n=2 et calculer la moyenne m de chacun d'eux.

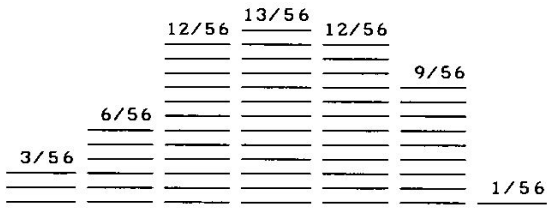
4 5 → 4.5	5 6 → 5.5	6 6 → 6.0	6 7 → 6.5
4 6 → 5.0	5 6 → 5.5	6 7 → 6.5	6 7 → 6.5
4 6 → 5.0	5 6 → 5.5	6 7 → 6.5	6 7 → 6.5
4 6 → 5.0	5 7 → 6.0	6 7 → 6.5	6 7 → 6.5
4 7 → 5.5	5 7 → 6.0	6 6 → 6.0	7 7 → 7.0
4 7 → 5.5	5 7 → 6.0	6 7 → 6.5	7 7 → 7.0
4 7 → 5.5	6 6 → 6.0	6 7 → 6.5	7 7 → 7.0

Pour chaque valeur observable m, nous pouvons dénombrer les échantillons de moyenne m, et par suite calculer la *proportion* de ces échantillons. Nous en déduisons la distribution de ces 28 moyennes en proportions d'échantillons, qui constitue la *distribution d'échantillonnage* de la statistique "moyenne observée" M; elle est figurée ci-après.



n = 2 *moy*(M) = 6 *var*(M) = 0.4286

De la même manière nous pouvons par exemple obtenir la *distribution d'échantillonnage* de M pour des échantillons d'effectifs n=3 (il y a dans ce cas un ensemble de 56 échantillons possibles).



n = 3 *moy*(M) = 6 *var*(M) = 0.2381

La distribution d'échantillonnage présente des propriétés remarquables (que nous énoncerons ici sans démonstration).

1) Sa moyenne, c'est-à-dire la moyenne des moyennes de tous les échantillons, est égale à μ, par exemple pour n=2

$$moy(M) = \frac{1\times4.5 + 3\times5.0 + 6\times5.5 + 6\times6.0 + 9\times6.5 + 3\times7.0}{28} = 6 = \mu$$

2) Sa variance, c'est-à-dire la variance des moyennes de tous les échantillons, que nous noterons ε^2, par exemple pour n=2

$$var(M) = \frac{1\times(4.5-6)^2 + 3\times(5.0-6)^2 + \cdots + (3\times7.0-6)^2}{28} = 0.4286 = \varepsilon^2$$

est égale à

$$\left(1-\frac{n-1}{N-1}\right)\frac{\sigma^2}{n} = \frac{3}{7}\sigma^2$$

Par conséquent, lorsque le rapport n/N (le "taux d'échantillonnage") est faible, on a approximativement

$$\varepsilon^2 \simeq \frac{\sigma^2}{n}$$

et, pour une population de taille arbitrairement grande ("infinie")

$$\varepsilon^2 = \frac{\sigma^2}{n}$$

Donc, l'écart-type ε de la distribution d'échantillonnage de la moyenne décroît avec la taille des échantillons n et, du moment que n est petit par rapport à N, est très peu affecté par la taille de la population.

3) Dans les situations considérées en pratique, la forme de la distribution d'échantillonnage de la moyenne est généralement unimodale et sensiblement symétrique. Il sera alors souvent raisonnable de considérer que cette distribution est approximativement une distribution *normale* de moyenne μ et de variance $\varepsilon^2=\sigma^2/n$; cette approximation sera d'autant meilleure que n est grand. En outre il s'agit d'un résultat *exact* si on suppose que la distribution parente est elle-même *normale*, ce qui est évidemment une idéalisation de la réalité.

Bien entendu nous pouvons procéder de la même manière pour, par exemple la variance, l'écart-type-corrigé, etc., et définir de manière générale la notion de distribution d'échantillonnage d'une statistique (voir plus loin les résultats généraux pour le modèle *normal*).

0.2.2 Test de typicité (inférence ensembliste): Les échantillons extrêmes

De la même manière que nous pouvons situer, pour une variable, un individu par rapport à sa population,

> par exemple ici un enfant de 5 ans est situé dans le quart des enfants les plus jeunes: il y a une proportion de 1+1=2 enfants sur 8 de 5 ans et moins,

nous pouvons situer un échantillon, pour sa moyenne, par rapport à l'ensemble de tous les échantillons de même effectif,

> par exemple un échantillon d'effectif n=2 de moyenne 5 est situé dans les 4/28 d'échantillons "les plus jeunes vis-à-vis de la moyenne": il y a une proportion de 1+3=4 sur 28 échantillons dont la moyenne est inférieure ou égale à 5.

Un échantillon tel qu'il y ait une proportion faible d'échantillons dont la moyenne soit au moins aussi *extrême* (par exemple ici du côté des âges les plus jeunes) peut être jugé *peu typique* de la population pour la moyenne. Ainsi nous pouvons nous donner un critère opérationnel, et décider qu'un échantillon est *atypique* de la population si cette proportion est plus petite qu'un *seuil* fixé α.

Par exemple, pour n=3 et α=0.10, un échantillon de moyenne 5 est *atypique* du côté des âges les plus jeunes: il y a une proportion

$$p_{inf}=3/56=0.054 < \alpha=0.10$$

d'échantillons au moins aussi extrêmes. Mais, pour n=2 et α=0.10, un échantillon de moyenne 5 ne peut pas être déclaré atypique

$$p_{inf}=4/28=0.143 > \alpha=0.10$$

On voit que ce "*test de typicité*" ne suppose aucunement que l'échantillon ait été choisi "au hasard": on a par exemple pu choisir les enfants ayant le moins bien réussi une certaine épreuve...

Mais ce n'est pas pour autant une procédure *descriptive*, dans la mesure où la conclusion, pour une même moyenne donnée, dépend de l'effectif (comme on peut le voir dans l'exemple précédent). Nous dirons qu'il s'agit d'une procédure d'*inférence ensembliste*, l'échantillon étant simplement défini comme une *partie* de la population (*cf.* Rouanet, Bernard et Lecoutre, 1986; Rouanet, Bernard et Le Roux, 1990).

0.2.3 Test de signification - Procédures d'échantillonnage

■ *Échantillonnage au hasard et conversion des proportions en probabilités*

Si l'échantillon a été obtenu par un procédé assimilable à un tirage au sort donnant la même "chance" d'être choisi à tous les échantillons possibles, nous dirons qu'il y a échantillonnage *au hasard* (ou "avec probabilités égales"). Dans ce cas les *proportions* d'échantillons définies précédemment peuvent être converties en *probabilités*: par exemple la proportion des échantillons dont la moyenne est inférieure à 5 devient la probabilité que la moyenne d'un échantillon (tiré au hasard) soit inférieure à 5.

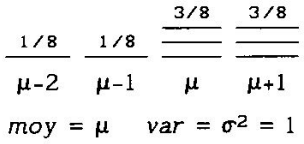
■ *Fonction de vraisemblance*

Dans les procédures usuelles d'inférence sur une moyenne, nous nous interrogeons sur la moyenne μ (*paramètre* inconnu) d'une distribution *parente*, à partir d'un échantillon supposé tiré *au hasard*. La situation *idéale* est celle où μ est la seule caractéristique inconnue de cette distribution.

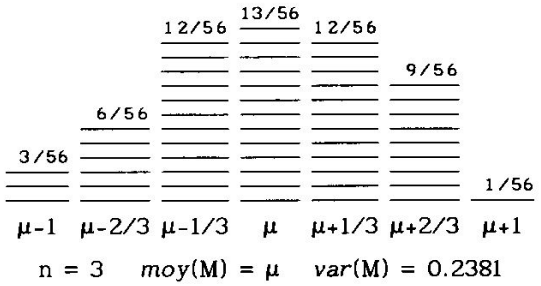
Par exemple, reprenons l'exemple de la variable *âge* dans une population d'effectif 8, avec les valeurs suivantes dépendant du paramètre μ, dont nous supposons pour simplifier qu'il s'agit d'un nombre entier

$$\mu \quad \mu+1 \quad \mu-1 \quad \mu \quad \mu+1 \quad \mu-2 \quad \mu+1 \quad \mu$$

d'où la *distribution en fréquences*



et la *distribution d'échantillonnage* de M pour des échantillons d'effectifs n=3



Supposons que nous ayons observé un échantillon de taille 3 de moyenne m=5. Nous pouvons, pour cette valeur m=5 fixée, calculer la probabilité d'échantillonnage $Pr(M=5|\mu)$, en fonction de l'ensemble des valeurs possibles de μ

	$\mu \leq 3$	$\mu=4$	$\mu=5$	$\mu=6$	$\mu \geq 7$
$Pr(M=5\|\mu)$	0	1/56	13/56	3/56	0

Cette fonction de μ (pour m fixée) est classiquement appelée *fonction de vraisemblance*; clairement il ne s'agit pas d'une distribution de probabilité.

■ *Du test de typicité au test de signification*

Dans la situation précédente, si nous fixons *par hypothèse* une valeur μ_0 ("hypothèse nulle"), nous déterminons une distribution *de référence* (hypothétique) pour laquelle nous pouvons appliquer la procédure du test de typicité.

Si nous posons l'hypothèse que $\mu=7$, nous constatons que cette hypothèse est *strictement incompatible* avec l'observation m=5, puisqu'il est impossible d'observer une moyenne égale à 5 si $\mu=7$ (la plus petite moyenne possible pour un échantillon étant alors 6).

Si nous posons l'hypothèse que μ=6, nous avons exactement la situation considérée précédemment pour le test de typicité, avec en plus la conversion des proportions en probabilités: ainsi, au seuil α=0.10, le résultat atypique (p_{inf}=3/56) s'interprète donc comme la réalisation d'un événement *peu probable* ("rare") *sous l'hypothèse* μ=6. Cette hypothèse est *peu compatible* avec l'observation m=5, ce qui constitue une extension de l'incompatibilité stricte. Par conséquent, au seuil α=0.10 (en fait à tout seuil α inférieur au seuil *observé* p_{inf}), on "rejette l'hypothèse nulle μ=6": le résultat du test est *significatif*.

On voit ici la difficulté conceptuelle soulevée par le test de signification, que nous avons déjà mentionnée précédemment: il fait intervenir dans le calcul du seuil observé la vraisemblance de l'hypothèse nulle, non seulement pour la valeur observée m, mais aussi pour toutes les valeurs (non observées!) plus extrêmes que m sous cette hypothèse.

Une autre difficulté est la nécessité de disposer d'une statistique de test pour situer l'échantillon. Même dans la situation la plus simple d'un seul paramètre, dès lors qu'il n'existe pas de statistique *exhaustive*, c'est-à-dire exprimant toute l'information sur le paramètre contenue dans les données (nous reviendrons plus loin sur cette notion), cette difficulté est considérable. Ainsi, dans l'exemple précédent, l'hypothèse μ=6 est bien évidemment compatible avec une moyenne observée égale à 5; mais si l'échantillon observé est {3,6,6} cette hypothèse est en fait strictement incompatible avec les données (la valeur 3 étant impossible).

0.2.4 Procédures bayésiennes

Dans l'inférence bayésienne, on considère explicitement, pour toutes les valeurs possibles du paramètre μ, l'ensemble de tous les échantillons "semblables" à l'échantillon observé.

■ *Une procédure ensembliste*

Imaginons la situation suivante. Nous disposons de deux *populations* distinctes, par exemple les 8 garçons d'une part et les 6 filles d'autre part inscrits à un cours de solfège. Considérons comme précédemment l'âge (en années entières) de ces enfants, soit

• pour les garçons (nous reprenons les valeurs précédentes)

6 7 5 6 7 4 7 6

• pour les filles

7 5 5 6 7 5

d'où, pour la population des filles, les 15 échantillons de taille n=2 que nous pouvons constituer et les moyennes de chacun d'eux

5 5 → 5.0	5 7 → 6.0	5 6 → 5.5	5 6 → 5.5	6 7 → 6.5
5 5 → 5.0	5 7 → 6.0	5 7 → 6.0	5 7 → 6.0	6 7 → 6.5
5 6 → 5.5	5 5 → 5.0	5 7 → 6.0	5 7 → 6.0	7 7 → 7.0

ainsi que la *distribution d'échantillonnage* de la statistique "moyenne observée" M

3/15	3/15	6/15	2/15	1/15
5.0	5.5	6.0	6.5	7.0

Considérons maintenant tous les échantillons de taille 2 constitués, soit de deux garçons soit de deux filles, dont la moyenne est égale à 5.5. Ces échantillons sont au nombre de 9, respectivement 6 pour les garçons et 3 pour les filles.

Par exemple la proportion de ces échantillons qui proviennent de la population des garçons est égale à

$$\frac{\textit{nombre d'échantillons de 2 garçons de moyenne } 5.5}{\textit{nombre total d'échantillons de moyenne } 5.5} = \frac{6}{6+3} = 0.667$$

	4.5	**5.0**	**5.5**	**6.0**	**6.5**	**7.0**	
garçons	1	3	6	6	9	3	/28
filles	0	3	3	6	2	1	/15
total	1	6	9	12	11	4	/43

■ *La formule de Bayes pour des proportions*

La proportion précédente peut être exprimée en fonction des proportions suivantes

• les proportions d'échantillonnage respectives

$$Pr(\text{M=5.5}|garçons) = 6/28 \quad \text{et} \quad Pr(\text{M=5.5}|filles) = 3/15$$

• les proportions respectives d'échantillons de garçons et de filles pour l'ensemble des 28+15=43 échantillons

$$Pr(garçons) = 28/43 \quad \text{et} \quad Pr(filles) = 15/43$$

à partir desquelles nous obtenons

• les proportions *conjointes*

$$Pr(\text{M=5.5 } et \text{ } garçons) = Pr(\text{M=5.5}|garçons) \times Pr(garçons) = 6/43$$

$$Pr(\text{M=5.5 } et \text{ } filles) = Pr(\text{M=5.5}|filles) \times Pr(filles) = 3/43$$

• et la proportion *marginale* d'échantillons de moyenne 5.5 pour l'ensemble des 43 échantillons

$$Pr(\text{M=5.5}) = Pr(\text{M=5.5 } et \text{ } garçons) + Pr(\text{M=5.5 } et \text{ } filles) = 9/43$$

Ceci conduit à la *formule de Bayes* (pour des proportions)

$$Pr(garçons|\text{M=5.5}) = \frac{Pr(\text{M=5.5}|garçons) \times Pr(garçons)}{Pr(\text{M=5.5})} = \frac{6/43}{9/43} = \frac{6}{9} = 0.667$$

■ *Un calcul probabiliste: La formule de Bayes pour les probabilités*

Plaçons-nous maintenant dans la situation considérée pour le test de signification de l'échantillonnage au hasard dans chaque population, qui permet la conversion des proportions d'échantillonnage en probabilités. On voit que, pour obtenir une inférence probabiliste analogue à la procédure ensembliste précédente, nous devons maintenant compléter ce *modèle* d'échantillonnage, puisque le choix d'un échantillon résulte simultanément du choix d'une population et d'un échantillon dans cette population.

Supposons que l'on ait également choisi au hasard l'une des populations, soit des probabilités marginales

$$Pr(garçons) = Pr(filles) = 1/2$$

qui constituent les *probabilités initiales* de choisir une population, et qui ne sont pas dans ce cas une conversion des proportions d'échantillons.

Si nous avons observé un échantillon de moyenne 5.5, nous en déduisons

• les probabilités *conjointes*

$$Pr(\text{M=5.5 } et \text{ } garçons) = Pr(\text{M=5.5}|garçons)\times Pr(garçons) = \frac{6}{28}\times\frac{1}{2} = \frac{6}{56}$$

$$Pr(\text{M=5.5 } et \text{ } filles) = Pr(\text{M=5.5}|filles)\times Pr(filles) = \frac{3}{15}\times\frac{1}{2} = \frac{3}{30}$$

• la probabilité *marginale*

$$Pr(\text{M=5.5}) = Pr(\text{M=5.5 } et \text{ } garçons) + Pr(\text{M=5.5 } et \text{ } filles) = \frac{6}{56}+\frac{3}{30} = \frac{29}{140}$$

qui constitue la *probabilité prédictive* (avant le tirage) de tirer un échantillon de moyenne 5.5

• la probabilité conditionnelle à la moyenne observée 5.5, obtenue par la *formule de Bayes*

$$Pr(garçons|\text{M=5.5}) = \frac{Pr(\text{M=5.5}|garçons)\times Pr(garçons)}{Pr(\text{M=5.5})} = \frac{6/56}{29/140} = 0.5172$$

Cette *probabilité finale* est bien entendu différente de la proportion précédente. Mais on remarquera qu'elle peut être assimilée à la limite de la fréquence des échantillons issus de la population de garçons que l'on obtiendrait en répétant l'expérience suivante: choisir au hasard l'une des deux populations, puis un échantillon dans cette population, et ne retenir que les échantillons dont la moyenne observée est 5.5.

A l'évidence le calcul précédent est général et peut être effectué, quel que soit le nombre de populations envisagées et quelles que soient les probabilités initiales affectées à ces populations.

En outre nous pouvons l'appliquer pour calculer la probabilité de chaque population, conditionnellement à l'échantillon lui-même, et non seulement à une caractéristique particulière de celui-ci. Ainsi, toujours dans la même situation, si nous avons observé un échantillon (de moyenne 5.5) avec les deux valeurs 5 et 6, nous avons les proportions (donc probabilités) d'échantillonnage respectives

$$Pr(\{5,6\}|garçons) = 3/28 \quad \text{et} \quad Pr(\{5,6\}|filles) = 3/15$$

dont nous déduisons en procédant comme précédemment la probabilité marginale

$$Pr(\{5,6\}) = (3/28)\times(1/2) + (3/15)\times(1/2) = 43/280$$

d'où la probabilité conditionnelle à l'échantillon observé

$$Pr(garçons|\{5,6\}) = \frac{Pr(\{5,6\}|garçons)\times Pr(garçons)}{Pr(\{5,6\})} = \frac{3/56}{43/280} = 0.3488$$

qui diffère ici nettement de la probabilité conditionnelle à la moyenne observée.

■ *Une procédure générale d'inférence statistique*

La procédure précédente prend tout son intérêt quand la probabilité *finale* est relative à une caractéristique de la population qui ne peut pas être déduite des observations (ce qui n'était pas le cas dans l'exemple précédent). C'est typiquement le cas dans la situation de l'inférence sur la moyenne μ, où nous calculons la probabilité (*finale*) relative à μ conditionnellement à l'échantillon observé, à partir des probabilités d'échantillonnage et de la probabilité *initiale* relative à μ.

0.2.5 Quelques résultats de base pour le modèle normal

■ *Distributions d'échantillonnage pour une distribution parente normale*

Pour une distribution parente *normale* $N(\mu,\sigma^2)$, de moyenne μ et de variance σ^2, nous avons pour un échantillon *au hasard* d'effectif n $(x_1, x_2, \ldots x_n)$ les résultats suivants, que nous admettrons.

• La distribution d'échantillonnage de la statistique *moyenne*, qui pour un échantillon donné a la valeur

$$m = \frac{\sum_{i=1}^{n} x_i}{n}$$

est une distribution *normale* de moyenne μ et de variance σ^2/n, ce que nous notons

$$m|\mu,\sigma^2 \sim N(\mu,\frac{\sigma^2}{n})$$

• La distribution d'échantillonnage de la statistique *variance-corrigée*, qui pour un échantillon donné à la valeur

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - m)^2}{n-1}$$

est une distribution, qui ne dépend que de σ^2 (et pas de μ); elle a pour moyenne σ^2 (pour cette raison s^2 est souvent appelée "variance estimée") et pour variance $2\sigma^4/(n-1)$. Nous l'appelons distribution *phi-deux* à q=n-1 degrés de liberté, d'échelle σ^2, et nous notons

$$s^2|\mu,\sigma^2 \sim \sigma^2\varphi_q^2 \qquad [q = n-1]$$

On remarquera que cette définition, bien que très commode en pratique, n'est pas classique: l'usage est de l'appeler distribution *khi-deux* à q degrés de liberté, d'échelle σ^2/n

$$s^2|\mu,\sigma^2 \sim (\sigma^2/n)\chi_q^2$$

Par suite la distribution d'échantillonnage de la statistique *écart-type-corrigé* s est une distribution *phi* (dont la moyenne est proche de σ, d'autant plus que q est grand, mais n'est pas égale à σ)

$$s|\mu,\sigma^2 \sim \sigma\varphi_q$$

- La distribution d'échantillonnage de la statistique

$$\frac{m-\mu}{s/\sqrt{n}}$$

est une distribution *t de Student* à q=n-1 degrés de liberté, qui (miracle!) ne dépend ni de μ ni de σ, ce que nous notons

$$\frac{m-\mu}{s/\sqrt{n}}|\mu,\sigma^2 \sim t_q$$

Cette distribution est symétrique autour de 0; elle a pour moyenne 0 (si q>1) et pour variance q/(q-2) (si q>2).

- La distribution d'échantillonnage de la statistique

$$\frac{m-\mu_0}{s/\sqrt{n}}$$

est une distribution *t de Student non-centrée* (ou *décentrée*), d'indice d'excentricité $(\mu-\mu_0)/(\sigma/\sqrt{n})$, ce que nous notons

$$\frac{m-\mu_0}{s/\sqrt{n}}|\mu,\sigma^2 \sim t'_q\left(\frac{\mu-\mu_0}{\sigma/\sqrt{n}}\right)$$

Cette distribution est asymétrique.

■ *Exhaustivité*

Les statistiques m et s^2 sont conjointement *exhaustives* pour (μ,σ^2). Dans le cadre bayésien, la propriété d'exhaustivité s'exprime de manière particulièrement simple et éclairante: sous le modèle normal, quelle que soit la distribution initiale relative au couple (μ,σ^2), la distribution finale relative à ces paramètres conditionnelle à l'échantillon tout entier coïncide avec celle conditionnelle aux valeurs observées m et s^2. Il sera donc suffisant dans ce cas de considérer cette dernière distribution.

■ *Distributions d'échantillonnage pour deux distributions parentes normales*

Nous avons pour deux échantillons *au hasard* indépendants,

- le premier d'effectif n_1 dans une distribution parente $N(\mu_1,\sigma_1^2)$,
- le second d'effectif n_2 dans une distributions parente $N(\mu_2,\sigma_2^2)$,

les résultats suivants, qui généralisent les précédents.

• La distribution d'échantillonnage d'une combinaison linéaire de moyennes

$$d = c_1m_1 + c_2m_2$$

est une distribution *normale*

$$d|\mu_1,\sigma_1^2,\mu_2,\sigma_2^2 \sim N(\delta,\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}) \qquad \text{avec } \delta = c_1\mu_1 + c_2\mu_2$$

soit, si $\sigma_1^2=\sigma_2^2=\sigma^2$ (*égalité des variances* parentes)

$$d|\mu_1,\mu_2,\sigma^2 \sim N(\delta,b^2\sigma^2) \qquad \text{avec } b^2 = \frac{c_1^2}{n_1} + \frac{c_2^2}{n_2}$$

• La distribution d'échantillonnage de la statistique *variance-corrigée intra-groupe moyenne*, qui est définie comme la moyenne des variances-corrigées, pondérée par les nombres de degrés de liberté

$$s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1-1 + n_2-1}$$

est une distribution *phi-deux*, si et seulement si $\sigma_1^2=\sigma_2^2=\sigma^2$; dans ce cas

$$s^2|\mu_1,\mu_2,\sigma^2 \sim \sigma^2\varphi_q^2 \qquad \text{avec } q = n_1-1 + n_2-1$$

• Si $\sigma_1^2=\sigma_2^2=\sigma^2$ la distribution d'échantillonnage de la statistique

$$\frac{d-\delta}{bs}$$

est une distribution *t de Student* à q degrés de liberté

$$\frac{d-\delta}{bs}|\mu_1,\mu_2,\sigma^2 \sim t_q$$

• Si $\sigma_1^2=\sigma_2^2=\sigma^2$ la distribution d'échantillonnage de la statistique

$$\frac{d-\delta_0}{bs}$$

est une distribution *t de Student non-centrée*

$$\frac{d-\delta_0}{bs}|\mu_1,\mu_2,\sigma^2 \sim t'_q(\frac{\delta-\delta_0}{b\sigma})$$

Les résultats précédents se généralisent au cas de k échantillons indépendants de distributions parentes *normales*.

LesEchantillons

Le programme **LesEchantillons** génère des tirages d'échantillons, et permet de simuler les distributions d'échantillonnage précédentes, ceci non seulement pour des distributions parentes *normales*, mais aussi pour un certains nombres d'autres familles de distributions.

Au lancement du programme, la distribution sélectionnée est la **distribution normale**, de paramètres 0 (**moyenne**) et 1 (**écart-type**). Cliquez sur le bouton **tracer les courbes** pour visualiser cette distribution.

Pour obtenir une simulation de la distribution d'échantillonnage, choisissez:
• l'**effectif** de l'échantillon,
• le **nombre de tirages**,
• la **statistique**,
puis cliquez sur le bouton **tirer des échantillons**.

Bien entendu la simulation sera d'autant meilleure que le nombre de tirages sera élevé; mais ce nombre sera limité par la mémoire disponible et par le temps de calcul qui sera nécessaire (effectuez des essais en augmentant progressivement le nombre de tirages). Vous pouvez apprécier la qualité de la simulation en tirant des échantillons d'effectifs 1, puisque

la distribution d'échantillonnage simulée est alors la distribution parente elle-même.

Pour choisir une autre distribution, cliquez sur la liste déroulante correspondante, et sélectionnez la famille de la distribution dans la liste qui s'affiche. Entrez ensuite les paramètres de la distribution, puis procédez comme précédemment.

Pour étudier le cas de deux distributions parentes, cliquez sur la case à cocher **distribution parente 2** et procédez de la même manière.

Vous pouvez ainsi explorer l'influence sur les distributions d'échantillonnage, à la fois des tailles d'échantillons et des distributions parentes.

■ *Distributions initiales non-informatives*

Le choix des distributions initiales "non-informatives" pour le modèle *normal* est classique. La solution considérée comme standard, et que nous admettrons ici, généralement attribuée à Jeffreys mais déjà donnée par Lhoste (1923), est la suivante: on choisit des distributions *uniformes* indépendantes, respectivement pour la moyenne μ et pour le logarithme de la variance σ^2.

Dans sa généralité, la définition des distributions initiales "non-informatives" est un problème redoutable, surtout si on l'envisage sous l'aspect quasi-philosophique du problème de la "formalisation de l'ignorance". La difficulté du problème est encore accrue par le fait qu'il ne s'agit pas de véritables distributions de probabilité, mais de distributions *impropres*, de "masse infinie". Pour des développements théoriques, nous renvoyons le lecteur, en ce qui concerne plus particulièrement le modèle *normal* à Box et Tiao (1973), et pour les aspects formels les plus généraux notamment à Bernardo et Smith (1994).

Mais, pour les situations particulières qui nous concernent ici, des justifications pratiques apparaîtront sans doute plus convaincantes que des justifications théoriques abstraites. Comme nous le verrons au chapitre 3, le choix d'une distribution initiale "exactement non-informative" n'apparaît pas véritablement crucial, dans la mesure où des distributions "faiblement informatives" conduisent déjà à des distributions finales indiscernables en pratique.

CHAPITRE 1

DU T DE STUDENT AUX PROCÉDURES FIDUCIO-BAYÉSIENNES

Dans ce chapitre, nous illustrerons, à partir d'exemples de base, l'intérêt méthodologique des procédures fiducio-bayésiennes. Nous examinerons également la réinterprétation dans le cadre fiducio-bayésien des procédures fréquentistes usuelles: tests de signification et intervalles de confiance. Cela nous conduira à mettre en avant l'un des aspects essentiels de l'inférence bayésienne, qui est de permettre d'obtenir d'une manière directe et intelligible une conclusion sur *l'importance* des effets, et non seulement sur leur *existence*.

Nous privilégierons deux situations pour leur intérêt méthodologique:

- la situation de recherche d'une conclusion d'effet *notable*, ou au contraire d'effet *non notable*;
- la situation de recherche d'une conclusion d'effet *négligeable*, ou du moins d'effet *limité*.

L'apport des procédures fiducio-bayésiennes à cette problématique de l'effet notable/négligeable a été dégagé notamment dans Lépine et Rouanet (1975), Rouanet, Lépine et Pelnard-Considère (1975), Rouanet, Lépine et Holender (1978), Rouanet et Lecoutre (1983), Lecoutre (1984), Rouanet (1995). Nous proposerons ici des solutions générales.

1.1 Recherche d'une conclusion d'effet notable: Inférence sur une moyenne

1.1.1 L'exemple de Student

Considérons les données ayant servi à Student dans son célèbre article *The probable error of a mean*, paru dans Biometrika en 1908. Elles correspondent à la durée de sommeil supplémentaire (exprimée en heures et fractions décimales

d'heures) pour dix patients auxquels on a administré deux somnifères A et B.

Si nous nous intéressons à la comparaison des deux somnifères, nous pouvons nous poser les questions suivantes (de la plus grossière à la plus fine).

(1) "Est-ce que le gain moyen de sommeil apporté par le somnifère A est *supérieur* à celui apporté par le somnifère B?"

(2) "Si oui, est-ce que le gain moyen de sommeil apporté par le somnifère A est *notablement supérieur* à celui apporté par le somnifère B?"

(3) "Si la réponse est encore oui, est-ce que, dans *la plupart des cas*, le gain de sommeil apporté par le somnifère A est *notablement supérieur* à celui apporté par le somnifère B?"

Pour répondre à ces questions, nous analysons les "effets individuels": pour chaque patient, la différence entre la durée de sommeil supplémentaire obtenue par le somnifère A et celle obtenue par le somnifère B.

patient	somnifère A	somnifère B	effet individuel (différence)
1	+1.9	+0.7	+1.2
2	+0.8	-1.6	+2.4
3	+1.1	-0.2	+1.3
4	+0.1	-1.2	+1.3
5	-0.1	-0.1	0
6	+4.4	+3.4	+1.0
7	+5.5	+3.7	+1.8
8	+1.6	+0.8	+0.8
9	+4.6	0	+4.6
10	+3.4	+2.0	+1.4
moyenne	+2.3300	+0.7500	+1.5800
éty-cor	2.0022	1.7890	1.2300

heures

1.1.2 Analyse descriptive

Une analyse descriptive très simple de ces données apporte les réponses suivantes (pour l'échantillon).

(1) Le gain de sommeil moyen observé est plus grand pour le somnifère A (2.33 heures) que pour le somnifère B (0.75 heure).

(2) La différence entre les gains de sommeil moyens est d=2.33-0.75=+1.58 heure ("effet observé"); cette différence est *notable* (importante); le gain de sommeil apporté par le somnifère A est un peu plus du triple de celui, déjà appréciable, apporté par le somnifère B.

(3) Pour chacun des 10 patients le gain apporté par le somnifère A est supérieur (dans 9 cas) ou égal à celui apporté par le somnifère B. Ceci se traduit par le fait que l'écart-type-corrigé des différences, soit s=1.23 heure, est relativement petit par rapport à la moyenne de ces différences (d=+1.58 heure), le rapport $d_{cal}=d/s$ ("effet moyen calibré") valant +1.58/1.23=+1.28.

1.1.3 Analyse inductive: Techniques de base

Examinons maintenant ces mêmes questions au niveau inductif. Des procédures inférentielles appropriées doivent nous permettre à ce niveau, soit de prolonger les réponses précédentes, soit de suspendre notre jugement par manque d'information. Le problème est ici celui de l'inférence sur la moyenne parente d'un groupe d'observations. Nous adopterons le *modèle normal* usuel: "les observations sont indépendantes et équidistribuées suivant une distribution *normale* de moyenne δ (effet vrai), et d'écart-type σ".

LeBayésien

La fenêtre principale comporte un menu (**mode d'entrée, énoncés**...), et en dessous des boutons (**moyenne G / O**...).

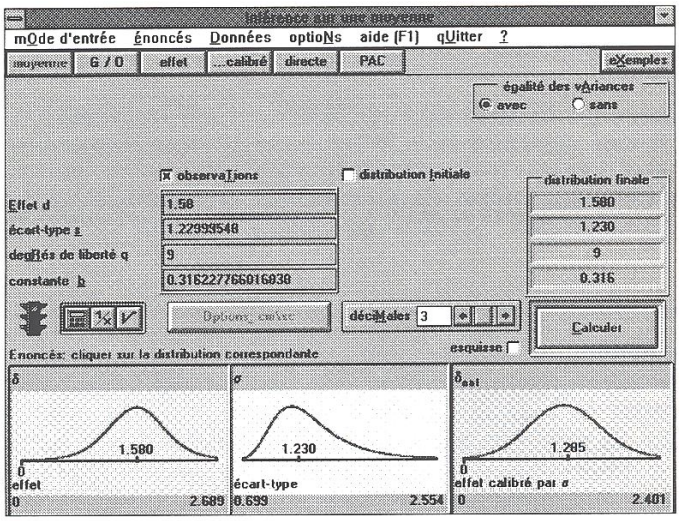
Cliquez sur le bouton **moyenne**, ce qui fait apparaître une fenêtre de saisie des statistiques. Entrez les valeurs appropriées:

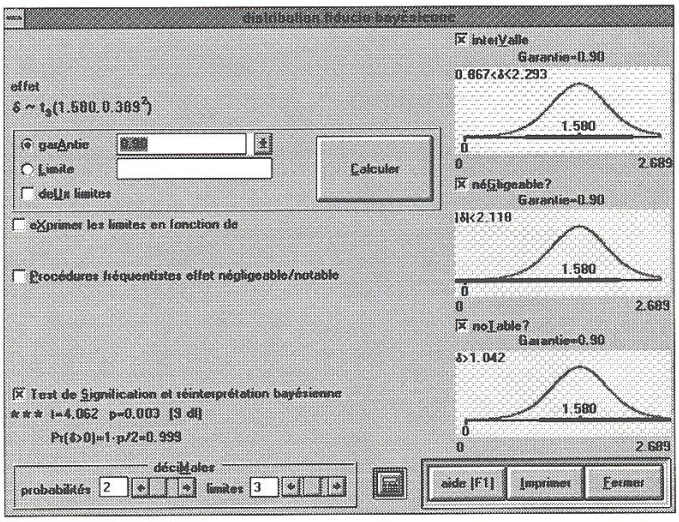
effectif	moyenne	écart-type
10	1.58	1.22999548

puis cliquez sur le bouton **ok**, ce qui vous ramène à la fenêtre principale.

Les courbes des distributions fiducio-bayésiennes relatives
- à l'effet vrai δ (à gauche)
- à l'écart-type σ (au milieu)
- à l'effet calibré $\delta_{cal}=\delta/\sigma$ (à droite)

sont dessinées.


Inférence sur une moyenne
mOde d'entrée énoncés Données optioNs aide (F1) qUitter ?
moyenne G / O effet ...calibré directe PAC eXemples
égalité des vAriances
avec sans
observaTions
distribution Initiale
distribution finale
Effet d
écart-type s
degRés de liberté q
constante b
1.58
1.22999548
9
0.316227766016838
1.580
1.230
9
0.316
Options calcul
déciMales 3
Calculer
Enoncés: cliquer sur la distribution correspondante
esquisse
δ
σ
δ est
1.580
1.230
1.285
effet
écart-type
effet calibré par σ
0
2.689
0.699
2.554
2.401


distribution bayésienne
effet
δ ~ t9(1.580, 0.389²)
garAntie
Limite
deUx limites
Calculer
eXprimer les limites en fonction de
Procédures fréquentistes effet négligeable/notable
Test de Signification et réinterprétation bayésienne
*** t=4.062 p=0.003 [9 dl]
Pr(δ>0)=1-p/2=0.999
déciMales
probabilités 2
limites 3
intervAlle
Garantie=0.90
0.867<δ<2.293
néGligeable?
Garantie=0.90
|δ|<2.110
noTable?
Garantie=0.90
δ>1.042
1.580
0
2.689
aide (F1)
Imprimer
Fermer

(1) La distribution fiducio-bayésienne de la différence δ ("effet vrai") est montrée dans la figure 1.1; elle est centrée sur l'effet observé (+1.58) et sa dispersion traduit l'incertitude expérimentale. Clairement cette distribution nous permet de conclure à l'existence d'un effet moyen (soit δ>0): nous avons une probabilité (une *garantie*) élevée que la différence δ soit supérieure à 0.

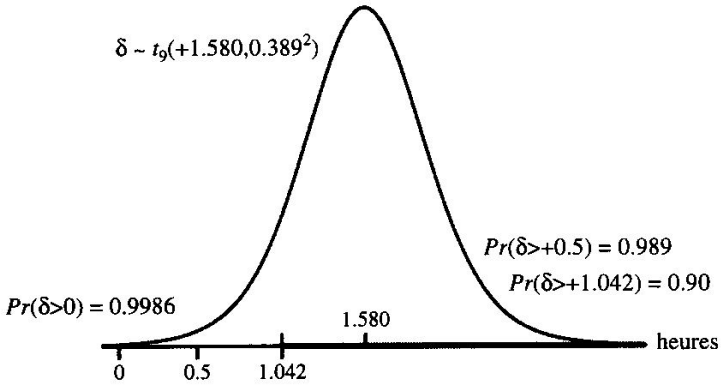


Figure 1.1 - Distribution fiducio-bayésienne pour la différence δ

(2) Pouvons-nous conclure à une différence δ notable?

LeBayésien

Cliquez sur la courbe de gauche relative à δ, ce qui fait apparaître une nouvelle fenêtre, avec trois courbes identiques. Sans rien modifier, cliquez sur le bouton **calculer**. Sur la courbe du bas est figuré l'énoncé fiducio-bayésien de recherche de conclusion d'effet **notable** pour la garantie 0.90: δ>1.042.

□ *Remarque*

Vous pouvez imprimer ou enregistrer cette courbe: voir la notice d'utilisation en annexe de cet ouvrage, ou consulter l'aide du programme.

"δ est supérieure à +1.042 heure avec la garantie fiducio-bayésienne 0.90"; la limite 1.042 correspond bien à une différence importante. Nous pouvons également fixer une limite jugée notable, par exemple une demi-heure.

LeBayésien

Activez l'option **limite** (dans le cadre qui contient le bouton **calculer**), et entrez la valeur 0.5 (que vous pouvez entrer sous la forme 1/2). Entrez 3 pour le nombre de décimales sur les probabilités (en bas à gauche de la fenêtre), puis cliquez sur **calculer**. Vous obtenez (courbe du bas) la garantie 0.989.

"δ est supérieure à une demi-heure avec la garantie fiducio-bayésienne 0.989". Nous pouvons donc considérer que nous avons également une réponse affirmative à la deuxième question.

LeBayésien

Si vous entrez la limite 0 (puis cliquez sur **calculer**), vous trouvez la probabilité 0.9986 (pour 4 décimales) que δ soit supérieur à 0. Remarquez encore que cette probabilité figure dans la rubrique de la fenêtre intitulée **test de signification et réinterprétation bayésienne**; nous y reviendrons plus loin (↳1.2.3).

(3a) Nous pouvons effectuer une inférence relative à l'écart-type vrai σ pour chercher à montrer qu'il est petit par rapport à δ. La distribution fiducio-bayésienne de σ est montrée dans la figure 1.2.

LeBayésien

Cliquez sur le bouton **fermer** pour revenir à la fenêtre principale, puis cliquez sur la courbe du milieu relative à σ, ce qui fait apparaître la fenêtre des énoncés pour l'écart-type. Activez l'option **garantie** (dans le cadre qui contient le bouton **calculer**), et sélectionnez s'il y a lieu la valeur 0.90 dans la liste déroulante, puis cliquez sur le bouton **calculer**. Sur la courbe du haut (intervalle) vous trouvez: $0.897<\sigma<2.024$

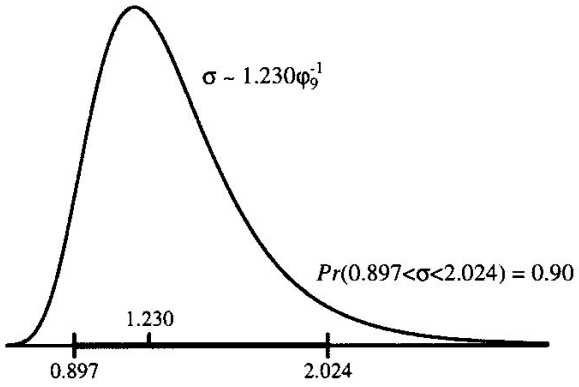


Figure 1.2 - Distribution fiducio-bayésienne pour l'écart-type σ

Mais ici, compte tenu du faible nombre de patients, nous sommes très incertains sur la vraie valeur de l'écart-type: "σ est compris entre 0.897 et 2.024 heures avec la garantie fiducio-bayésienne 0.90".

(3b) Il y a une meilleure manière de procéder, puisque nous pouvons effectuer directement une inférence sur l'effet calibré δ_{cal}, c'est-à-dire le rapport δ/σ. La distribution fiducio-bayésienne de δ_{cal} est montrée dans la figure 1.3.

LeBayésien

Cliquez sur le bouton fermer pour revenir à la fenêtre principale, où vous cliquez maintenant sur la courbe de droite relative à δ_{cal}. Dans la fenêtre des énoncés, cliquez sur le bouton **calculer** (pour la garantie 0.90). Sur la courbe du bas (**notable**), vous trouvez: δ_{cal}>0.696.

"δ_{cal} est supérieur à +0.696 avec la garantie fiducio-bayésienne 0.90". A titre indicatif, pour une distribution *normale* dont la moyenne (positive) vaut 0.70 fois l'écart-type, 75% des valeurs de la variable sont positives et

69% sont supérieures au quart de la moyenne.
Pour la troisième question, la réponse est donc plutôt positive.

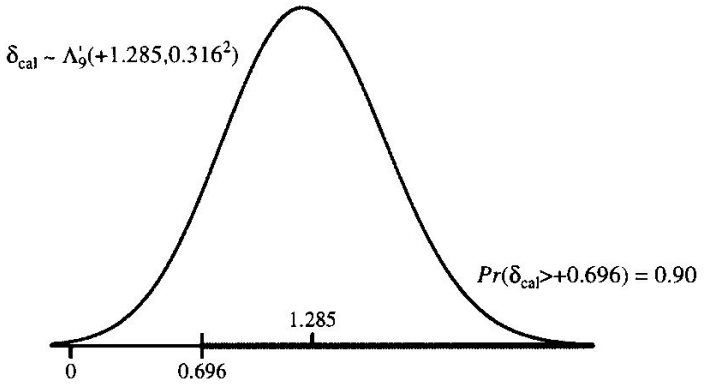


Figure 1.3 - Distribution fiducio-bayésienne pour l'effet calibré δ_{cal}

■ *Remarque: Variabilité inter-sujets et intra-sujets*

La classe de plans d'analyse que nous envisageons dans cet ouvrage ne permet pas (à moins de suppositions peu réalistes) de distinguer la variabilité inter-sujets et la variabilité intra-sujets. Ceci devra être gardé à l'esprit quand nous parlons dans ce contexte de différence ou d'effet *individuel*.

1.1.4 Analyse inductive: Techniques complémentaires

Cette section présente des procédures moins courantes (actuellement...), que nous ne ferons qu'aborder dans cet ouvrage. Sa lecture pourra être omise.

■ *L'inférence prédictive*

Nous pouvons effectuer une inférence *prédictive* relative à une nouvelle observation: que pouvons-nous dire de la différence d_2 que nous observerions entre le somnifère A et le somnifère B pour une observation future (relative à un patient)? Nous pouvons ici énoncer: "cette différence d_2 serait positive

avec la garantie fiducio-bayésienne 0.874; elle serait supérieure à une demi-heure avec la garantie 0.788".

■ *Le paramètre* π_x

Notons π_x la proportion des valeurs de la distribution parente (des effets individuels) qui sont supérieures à x donné. π_x est donc un *paramètre*, qui pour une distribution *normale*, est déterminé par δ et σ. Par exemple, pour x=0.5, δ=+1.58 et σ=1.230, nous avons

$$\pi_{0.5} = Pr(N(+1.58,1.230^2)>0.5) = 0.810$$

Nous pouvons donc obtenir la distribution fiducio-bayésienne relative à ce paramètre et énoncer ici: "$\pi_{0.5}$, la proportion des valeurs de la population parente qui sont supérieures à une demi-heure, est plus grande que 65% avec la garantie fiducio-bayésienne 0.90".

Les procédures précédentes sont une manière directe de répondre à la troisième question. Elles sont reliées par la propriété remarquable suivante: la moyenne de la distribution fiducio-bayésienne de π_x est égale à la probabilité prédictive que d_2 soit supérieure à x (*cf.* Gertsbakh et Winterbottom, 1991).

Ici ces procédures, confirmant la conclusion énoncée précédemment à partir de l'inférence sur l'effet calibré δ_{cal}, apportent une réponse plutôt positive.

1.1.5 Résultats généraux: Techniques de base

(1) La distribution fiducio-bayésienne relative à δ est une distribution *t de Student généralisée* (à ne pas confondre avec la distribution *t non-centrée*).

• Cette distribution est centrée sur la différence observée

$$d = +1.580$$

• L'échelle de la distribution e=0.389 traduit l'imprécision expérimentale. Dans le cas présent de l'inférence sur la moyenne d'un groupe d'observations à partir d'un échantillon d'effectif n, nous avons

$$e = \frac{s}{\sqrt{n}} = \frac{1.230}{\sqrt{10}} = 0.389$$

• Le nombre de degrés de liberté de la distribution est celui associé à l'écart-type-corrigé s, soit ici

$$q = n-1 = 10-1 = 9$$

Nous noterons

$$\delta \sim t_q(d,e^2) \sim t_9(+1.580,0.389^2)$$

Quand l'effectif est élevé, il s'agit approximativement d'une distribution *normale*, de moyenne d et d'écart-type e

$$\delta \simeq N(d,e^2) \quad \text{pour q élevé}$$

□ *Remarques sur l'échelle* e

Pour caractériser la dispersion de la distribution *t de Student*, on utilise l'échelle e et non pas l'écart-type qui présente l'inconvénient de dépendre du nombre de degrés de liberté q. Cet écart-type vaut en effet

$$e\sqrt{q/(q-2)}$$

soit ici 0.441.

Par ailleurs nous suivons l'usage le plus répandu pour la distribution *normale* qui est d'indiquer la variance (soit e^2) plutôt que l'écart-type e. Nous adoptons la même notation pour l'échelle de la distribution *t de Student*.

(2) La distribution fiducio-bayésienne relative à σ est une distribution *phi-inverse*,

- avec également q=9 degrés de liberté,
- d'échelle égale à l'écart-type-corrigé observé s=1.230.

Nous noterons

$$\sigma \sim 1.230\varphi_9^{-1}$$

Cette distribution peut être définie à partie de la distribution usuelle *khi-deux*, puisque

$$\frac{1}{\sigma^2} \sim \frac{1}{qs^2}\chi_q^2 \sim 0.0734\chi_9^2$$

(3) La distribution fiducio-bayésienne relative à δ_{cal} est une distribution que nous appelons *lambda-prime*,

- également à q degrés,
- d'indice d'excentricité d_{cal} = d/s = +1.580/1.230 = +1.285,
- et d'indice d'échelle b = $1/\sqrt{10}$ = 0.316.

Nous noterons

$$\delta_{cal} \sim \Lambda'_q(d_{cal}, b^2) \sim \Lambda'_9(+1.285, 0.316^2)$$

Cette distribution est asymétrique. Sa moyenne est d'autant plus proche de d_{cal} que q est grand. Son écart-type dépend à la fois de d_{cal}, b et q.

Quand l'effectif est élevé, il s'agit approximativement d'une distribution *normale*, de moyenne d_{cal} et d'écart-type b

$$\delta_{cal} \simeq N(d_{cal}, b^2) \quad \text{pour q élevé}$$

Pour δ et σ les courbes, ainsi que les énoncés de probabilité correspondants, peuvent être déduits des densités et des fonctions de répartition des distributions familières, *t de Stu-dent* et *khi-deux*. Seule la distribution relative à δ_{cal} est d'un type non usuel.

LeBayésien

□ *Pour illustrer le rôle de* d, s *et* n

▸ Revenez à la fenêtre principale et cliquez sur le bouton **moyenne**. Entrez l'effectif 5

effectif	moyenne	écart-type
5	1.58	1.22999548

puis cliquez sur le bouton **ok**, ce qui vous ramène à la fenêtre principale. Les distributions sont plus dispersées: il y a moins de données, donc l'incertitude est plus grande. Par exemple nous pouvons encore conclure à l'existence d'une différence moyenne vraie (soit $\delta>0$), mais avec une garantie (0.977) plus faible que précédemment.

▸ Dans la fenêtre principale, cliquez sur le bouton **moyenne**. Rétablissez l'effectif 10, et entrez l'écart-type 4.22999548

effectif	moyenne	écart-type
10	1.58	4.22999548

puis cliquez sur le bouton **ok**. Ici encore les distributions relatives à l'effet et à l'effet calibré sont plus dispersées. Par exemple nous avons seulement la garantie 0.866 que δ soit supérieur à 0, ce qui sera sans

doute jugé insuffisant pour conclure à l'existence d'un effet.

▸ Dans la fenêtre principale, cliquez sur le bouton **moyenne**. Rétablissez l'écart-type 1.22999548, et entrez la moyenne 0.48

effectif	moyenne	écart-type
10	0.48	1.22999548

puis cliquez sur le bouton **ok**. Ici la distribution relative à l'effet δ a la même dispersion que pour données initiales; mais nous avons seulement la garantie 0.876 que δ soit supérieur à 0. La distribution relative à l'effet calibré δ_{cal} est un peu moins dispersée (car d_{cal} est plus petit); mais par exemple la probabilité que δ_{cal} soit supérieur à +0.696 n'est plus que 0.168 (au lieu de 0.90).

1.1.6 Résultats généraux: Techniques complémentaires

La lecture de cette section pourra être omise. Les procédures correspondantes ne sont pas incluses dans la version du programme **LeBayésien** fournie avec cet ouvrage. Les résultats peuvent cependant être obtenus en utilisant le programme **Distributions** (voir le chapitre ↳9).

■ *L'inférence prédictive*

La distribution fiducio-bayésienne prédictive relative à d_2 est, comme celle de δ, une distribution *t de Student généralisée*, centrée sur la différence observée d et avec q degrés de liberté.

L'échelle de la distribution est cette fois égale à $s(b^2+1)^{1/2}$

$$\delta \sim t_q(d, s^2(b^2+1)) \sim t_9(+1.580, 1.290^2)$$

■ *Le paramètre* π_x

D'une manière générale nous avons

$$\pi_x = Pr(N(\delta,\sigma^2)>x) = Pr(N(0,1)<\frac{\delta-x}{\sigma})$$

d'où l'équivalence

$$\pi_x = \kappa \Leftrightarrow \frac{\delta - x}{\sigma} = \bar{z}_\kappa$$

où $\bar{z}_\kappa$ est donné par la fonction de répartition de la distribution *normale*

$$Pr(N(0,1) < \bar{z}_\kappa) = \kappa$$

La fonction de répartition de la distribution fiducio-bayésienne de π_x se déduit directement de celle de $(\delta - x)/\sigma$, puisque, en raison de l'équivalence précédente

$$Pr(\pi_x < \kappa) = Pr(\frac{\delta - x}{\sigma} < \bar{z}_\kappa)$$

La distribution fiducio-bayésienne relative au rapport $(\delta - x)/\sigma$ est, comme celle de δ/σ, une distribution *lambda-prime*

$$\frac{\delta - x}{\sigma} \sim \Lambda'_q(\frac{d-x}{s}, b^2)$$

Nous avons par exemple ici, pour x=0.5

$$\frac{\delta - 0.5}{\sigma} \sim \Lambda'_q(0.878, 0.316^2) \quad \text{d'où } Pr(\frac{\delta - 0.5}{\sigma} < +0.3735) = 0.10$$

$$\text{et par suite } Pr(\pi_{0.5} > 0.65) = 0.90 \quad (0.3735 = \bar{z}_{+0.65})$$

1.1.7 Lien intuitif avec le test de signification

L'*hypothèse* d'un effet nul (δ=0) apparaît ici "peu compatible" avec les données, puisque 0 est situé dans une zone de "valeurs peu probables". Cette hypothèse sera d'autant moins compatible avec les données que 0 sera une valeur plus *extrême* de la distribution, c'est-à-dire que la distance entre 0 et d sera grande. Il est naturel, pour apprécier cette distance, de la rapporter à l'échelle e de la distribution et de considérer le rapport

$$\frac{d-0}{e} = \frac{1.580}{0.389} = +4.062$$

qui n'est autre que la statistique usuelle "*t* de Student" du test de l'hypothèse nulle δ=0.

On voit ici intuitivement comment un résultat du test *significatif* (*t* élevé) se traduit directement dans le cadre fiducio-bayésien.

Par la suite nous utiliserons largement les liens techniques et conceptuels entre l'inférence fiducio-bayésienne et le test de signification. Nous pouvons dès maintenant remarquer que, si nous connaissons l'effet observé d et la valeur t (=d/e) prise par la statistique de test usuelle, il est immédiat de passer du résultat du test à la distribution fiducio-bayésienne relative à δ

$$\delta \sim t_q(d,e^2) \quad \text{avec } e = \frac{d}{t} \quad \text{(pourvu que } d \neq 0)$$

1.1.8 Effet notable/effet négligeable: Vers des solutions générales

Ce premier exemple pourrait laisser croire que l'inférence fiducio-bayésienne ne fait que confirmer les conclusions des analyses usuelles. En particulier on pourrait penser qu'il suffit d'avoir un verdict du test de signification confirmant la conclusion descriptive, c'est-à-dire

- soit un effet observé notable et un test significatif,
- soit un effet observé négligeable et un test non-significatif,

pour pouvoir *ipso facto* prolonger inductivement la conclusion descriptive. Nous verrons qu'il n'en est rien, et que notamment les mises en garde habituelles à l'encontre de l'interprétation d'un résultat non-significatif sont réellement fondées: un tel résultat, compte tenu des effectifs utilisés, est souvent un constat d'ignorance.

Nous ne chercherons pas ici à donner un statut intrinsèque aux notions d'effet notable et d'effet négligeable, l'importance d'un effet ne pouvant être appréciée que dans un certain contexte, relativement à un certain état de connaissance. Il s'agira avant tout ici d'illustrer la mise en œuvre des procédures. Nous avons choisi pour motiver les analyses des exemples réels, mais nous leur garderons un caractère relativement abstrait. Il ne serait guère réaliste, dans ces conditions, de rechercher des critères stricts d'évaluation de l'importance des effets; c'est pourquoi nous procèderons de manière assez arbitraire. Rappelons néanmoins que, de toute évidence, le choix de critères pour juger de l'importance d'un effet n'est pas propre à l'utilisation de procédures inférentielles, mais qu'il se pose dès les analyses descriptives.

Remarquons encore que, même si l'on ne dispose pas de critère absolu, pouvoir apprécier l'importance *relative* des différents effets en jeu dans une expérience donnée sera déjà un apport essentiel.

1.2 Recherche d'une conclusion d'effet notable: Comparaison de deux moyennes de groupes indépendants

1.2.1 Données pharmacologiques: Nouveau médicament/Placebo

Cet exemple est extrait d'une étude clinique pharmacologique. Les données concernent la comparaison d'un nouveau médicament à un *placebo*, administrés tous deux aux mêmes patients en deux "périodes" séparées par un certain délai. Ceux-ci sont répartis en deux groupes de six sujets chacun, suivant la séquence d'administration:

- groupe $_{g1}$: *placebo* puis médicament;
- groupe $_{g2}$: médicament puis *placebo*.

Le tableau suivant fournit, pour chaque sujet $_{s<g>}$ (sujet $_{s}$ affecté au hasard au groupe $_{g}$) la différence $d^{s<g>}$ entre les valeurs observées pour le *placebo* et pour le médicament, pour une certaine variable, que le médicament devrait faire diminuer.

s3 g1	+ 8.1667	
s4 g1	+ 6.8333	
s6 g1	-11.6667	moyenne d^{g1}=+6.8611
s7 g1	+ 6.1667	éty-cor s_{g1}=10.1879
s10 g1	+13.3333	
s12 g1	+18.3333	
s1 g2	+12.6667	
s2 g2	- 3.1667	
s5 g2	+ 5.8333	moyenne d^{g2}=+6.5833
s8 g2	+11.8333	éty-cor s_{g2}= 7.0630
s9 g2	- 0.5000	
s11 g2	+12.8333	

Pour apprécier l'importance des différences observées, il convient de les rapporter aux moyennes

		placebo	médicament	
groupe	$g1$	29.500	22.639	26.069
	$g2$	29.056	22.472	25.764
		29.278	22.556	25.917

Nous adopterons le *modèle équinormal* usuel: "dans chaque groupe g les observations sont indépendantes et équidistribuées $N(\delta_g,\sigma^2)$".

On notera que, compte tenu des faibles effectifs, on pourrait s'interroger ici sur la non-homogénéité éventuelle des variances: nous reviendrons sur ce problème au chapitre ↳6.

Définissons alors la variance-corrigée intragroupe moyenne, qui pour des effectifs égaux est simplement la moyenne arithmétique des variances-corrigées de chaque groupe

$$s^2 = \frac{10.1879^2+7.0630^2}{2} = 8.766^2$$

à laquelle sont associés

$$q = (6-1)+(6-1) = 10 \text{ degrés de liberté}$$

(dans le cas d'effectifs quelconques, on considérerait la moyenne pondérée par les nombres de degrés de liberté associés à chaque groupe).

Nous nous intéressons ici à la moyenne générale vraie

$$\delta = (\delta_{g1}+\delta_{g2})/2 = \tfrac{1}{2}\delta_{g1} + \tfrac{1}{2}\delta_{g2}$$

et le problème est celui de l'inférence sur une combinaison linéaire de moyennes parentes de groupes indépendants.

■ *Remarque: Effet de la comparaison et effet du médicament*

Nous dirons que δ est "l'effet de la comparaison du *placebo* et du nouveau médicament", ou encore, en définissant le facteur *Traitements* comme l'ensemble des deux modalités *placebo* et nouveau traitement, "l'effet global (c'est-à-dire moyenné sur les périodes) du facteur *Traitements*", sous-entendu dans le cadre du plan d'analyse. On évitera d'utiliser des expressions raccourcies telles que "effet du médicament", ou plus généralement "effet du traitement", qui pourraient prêter à confusion, car δ peut bien entendu dépendre d'autres

effets contrôlés dans le plan d'expérience (par exemple ici "l'effet de période" ou "les effets résiduels": voir la discussion sur les plans *cross-over* dans le chapitre §5), ou d'effets non contrôlés. Ainsi nous distinguerons notamment l'effet global du facteur *Traitements* de son effet restreint (ou conditionnel) à la première période.

LeBayésien

Dans la fenêtre principale, cliquez sur le bouton **G / O**, ce qui fait apparaître la fenêtre de saisie des statistiques.

En haut de la fenêtre, sélectionnez les options:

S<G> (groupes indépendants);

2 (nombre de groupes).

Entrez les valeurs appropriées:

effectif	moyenne	écart-type
6	6.86111111	10.18791041
6	6.58333333	7.063010060

Sélectionnez l'une des options **moyenne équipondérée** ou **moyenne pondérée** (ce qui est indifférent ici car les effectifs sont égaux). Utilisez les options (par défaut) **avec égalité des variances** et **spécifique** (celle-ci est indifférente ici, car la comparaison met en jeu tous les groupes).

Puis cliquez sur le bouton **calculer**. Vous obtenez t=2.657 p/2=0.0120...

■ *Test de signification*

L'hypothèse nulle $\delta=0$ peut être testée au moyen du t de Student (avec 5+5=10 degrés de liberté), basé sur la différence moyenne observée $d=(d^{g1}+d^{g2})/2$ =+6.722 et sur l'écart-type-corrigé intragroupe moyen s=8.766.

$$t = \frac{d}{s/\sqrt{12}} = \frac{+6.722}{2.530} = +2.657 \quad [q=10\ dl]$$

Le test est significatif; le seuil observé unilatéral est $p/2=0.012$.

■ *Remarque: la constante* b

Remarquons que, d'une manière générale, le dénominateur de la statistique de test t peut s'écrire sous la forme

$$e = bs$$

où la constante b s'obtient à partir des coefficients v_g affectés aux groupes (ici $v_{g1}=v_{g2}$ et des effectifs f_g des groupes comme la racine carrée de

$$b^2 = \sum_{g \in G} \frac{v_g^2}{f_g}$$

$$\text{Ici } b^2 = \frac{(1/2)^2}{6} + \frac{(1/2)^2}{6} = 1/12$$

■ *Inférence fiducio-bayésienne*

LeBayésien

Cliquez maintenant sur le bouton **ok**, ce qui vous ramène à la fenêtre principale. La situation est la même que dans l'exemple de Student.
Cliquez sur la distribution relative à l'effet δ.

La distribution fiducio-bayésienne de δ est une distribution *t de Student généralisée*, centrée sur la différence observée d=+6.722, d'échelle e=2.530 (dénominateur de la statistique de test *t*), et avec q=10 degrés de liberté

$$\delta \sim t_q(d,e^2) \sim t_{10}(+6.722,2.530^2)$$

La distribution fiducio-bayésienne, montrée dans la figure 1.4, représente l'information sur la différence vraie δ apportée par les données. A partir de cette distribution, nous pouvons d'abord réinterpréter les résultats de l'intervalle de confiance fréquentiste et des tests de signification usuels.

1.2.2 Réinterprétation de l'intervalle de confiance fréquentiste

LeBayésien

Par exemple ici choisissez la **garantie** 0.95 et cliquez sur **calculer**. Vous trouvez dans la courbe du haut pour l'**intervalle**, avec trois décimales: 1.084<δ<12.360.

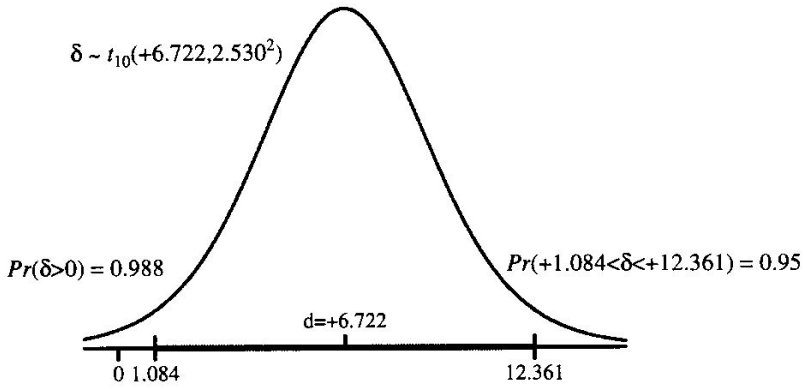


Figure 1.4 - Distribution fiducio-bayésienne pour δ

Nous avons l'énoncé fiducio-bayésien

$$Pr(+1.084<\delta<+12.360) = 0.95$$

L'intervalle de crédibilité [+1.084,+12.360], centré sur la valeur observée d, coïncide avec l'intervalle de confiance fréquentiste usuel. En effet, si t_α désigne la valeur (positive) critique *bilatérale* de la distribution *t de Student* au seuil α telle que

$$Pr(|t_q|>t_\alpha) = \alpha$$

cet intervalle est de la forme

$$[d-et_\alpha\ ,\ d+et_\alpha] = [+1.084,+12.360]$$

(ici, pour α=0.05 et q=10, $t_{0.05}$=2.2281)

L'interprétation fréquentiste est que, si on calculait un très grand nombre d'intervalles de ce type (pour des données chaque fois différentes obtenues sous le même modèle), (1-α)% de ces intervalles contiendraient la vraie valeur δ. La réinterprétation fiducio-bayésienne est immédiate: la probabilité que δ appartienne à cet intervalle, *conditionnellement à l'observation de* d *et de* s, est égale à 1-α

$$Pr(d-et_\alpha < \delta < d+et_\alpha \mid d,s) = 1-\alpha$$

$$\text{soit encore } Pr(|\delta-d| < et_\alpha \mid d,s) = 1-\alpha$$

Si la précision expérimentale est bonne (e est "petit"), nous aurons donc une probabilité fiducio-bayésienne élevée que δ soit "proche" de d.

1.2.3 Réinterprétation du seuil observé du test de signification

Le seuil observé unilatéral (ou orienté) du test t de l'hypothèse nulle usuelle $\delta=0$ ($p/2$) est simplement la probabilité fiducio-bayésienne que l'effet vrai soit de *signe opposé* à l'effet observé. Autrement dit, dans le cas présent où le résultat est significatif, nous avons une probabilité fiducio-bayésienne élevée ($1-p/2$) que l'effet vrai soit de même signe que d, donc positif.

$$Pr(\delta>0) = 1-\frac{p}{2} \text{ (si } d>0) \quad \text{ou} \quad Pr(\delta<0) = 1-\frac{p}{2} \text{ (si } d<0)$$

$$\text{Ici} \quad Pr(\delta>0) = 0.988$$

C'est, on le voit, une manière directe et naturelle d'exprimer le résultat du test t. En cas de résultat significatif comme ici, nous pouvons donc conclure à l'existence de l'effet avec une bonne garantie. En cas de résultat non-significatif, nous jugerons en général que la garantie obtenue n'est pas suffisante (tout au moins si nous adoptons les critères habituels) pour pouvoir conclure sur le sens de l'effet.

Cette réinterprétation du test de signification usuel fait apparaître la pauvreté de l'information qu'il apporte relativement à l'information contenue dans la distribution fiducio-bayésienne. En particulier un résultat significatif nous permet de nous prononcer sur le signe de l'effet, mais en lui-même ne nous autorise en aucun cas à conclure que cet effet est important.

Notons que, pour des raisons de symétrie manifestes (la distribution est centrée sur d et est symétrique), nous avons encore

$$Pr(0<\delta<2d) = 1-p \text{ (si } d>0) \quad \text{ou} \quad Pr(2d<\delta<0) = 1-p \text{ (si } d<0)$$

qui constitue la réinterprétation fiducio-bayésienne du seuil *bilatéral* p.

■ *Le test bilatéral et l'approche bayésienne sont-ils irréconciliables?*

Dans cette réinterprétation, p n'est pas la probabilité que δ soit nul, ou du moins proche de 0 (disons $|\delta|<\alpha$). Certains auteurs (notamment Berger, 1985) y ont vu la manifestation d'une "irréconciabilité" du test bilatéral et de

l'approche bayésienne. Nous ne partageons pas ce point de vue, car ce résultat apparaît correspondre à la réinterprétation intuitive selon laquelle, en cas de résultat significatif, "il y a de grandes chances que δ soit proche de d".

1.2.4 Réinterprétation du seuil fixé du test de signification

LeBayésien

Toujours pour la **garantie** 0.95, vous trouvez dans la courbe du bas (**notable**), avec trois décimales: δ>2.136

Nous avons l'énoncé fiducio-bayésien

$$Pr(\delta>+2.136) = 0.95$$

Pour plus de clarté nous noterons maintenant d_{obs} au lieu de d la différence observée. Considérons $d_{\alpha/2}$ la valeur de l'effet, de même signe que d_{obs} donc ici positive, qui serait juste significative au seuil unilatéral $\alpha/2$.

Par exemple, pour $\alpha/2$=0.05 et q=10, nous avons

$$Pr(t_{10}>+1.8125) = 0.05$$

Par conséquent $d_{0.05}/e$ (valeur que prendrait la statistique de test) est égale à la valeur critique (unilatérale) +1.8125, soit

$$d_{0.05} = 2.530\times1.8125 = +4.586 \quad (e=2.530)$$

La limite +2.136 trouvée pour la garantie 0.95 est égale à la différence:

$$d_{obs}-d_{0.05} = +6.722-4.586 = +2.136$$

D'une manière générale, nous avons l'énoncé fiducio-bayésien

$$Pr(\delta<d_{obs}-d_{\alpha/2}) = \frac{\alpha}{2} \quad \text{soit} \quad Pr(\delta>d_{obs}-d_{\alpha/2}) = 1-\frac{\alpha}{2}$$

qui constitue la réinterprétation fiducio-bayésienne du seuil fixé unilatéral $\alpha/2$.

Par symétrie nous avons encore

$$Pr(\delta<d_{obs}-d_{\alpha/2} \text{ ou } \delta<d_{obs}+d_{\alpha/2}) = \alpha \quad \text{soit} \quad Pr(d_{obs}-d_{\alpha/2}<\delta<d_{obs}+d_{\alpha/2}) = 1-\alpha$$

qui constitue la réinterprétation fiducio-bayésienne du seuil fixé bilatéral α (nous retrouvons encore ici la réinterprétation de l'intervalle de confiance

fréquentiste).

Les limites qui interviennent dans les énoncés précédents sont d_{obs}-$d_{\alpha/2}$ et d_{obs}+$d_{\alpha/2}$. Selon les conventions usuelles le seuil fixé est choisi "petit", de sorte que 1-$\alpha/2$ peut être considéré comme une garantie suffisante pour permettre une conclusion inductive.

■ *Si* d_{obs} *est "nettement notable" et si le résultat est "très significatif"*

Dans la mesure où ces conditions correspondent à une valeur d_{obs}–$d_{\alpha/2}$ notable, une conclusion d'effet notable est souvent raisonnable. Il s'agit manifestement là d'une situation qui met à l'aise l'utilisateur, puisqu'il y a convergence du résultat du test et de la conclusion descriptive. Mais, la conclusion reste tout de même assez impressionniste...

■ *Si* d_{obs} *est notable et si le résultat est non-significatif*

Dans ce cas la conclusion inductive doit être suspendue: on ne peut conclure à l'existence d'un effet et il est évidemment hors de question de conclure à un effet négligeable. Il s'agit là d'une situation souvent embarassante pour l'utilisateur, qui ne peut généraliser la conclusion descriptive: la tentative pour s'en sortir est souvent de proclamer "*il y a une tendance*". Mais cette situation n'est pas en vérité conflictuelle: elle indique simplement que l'information expérimentale est insuffisante pour pouvoir conclure.

■ *Si* d_{obs} *est "nettement négligeable" et si le résultat est "très significatif"*

Dans la mesure où ces conditions correspondent à des valeurs d_{obs}–$d_{\alpha/2}$ et d_{obs}+$d_{\alpha/2}$ négligeables (elles impliquent au moins que d_{obs}–$d_{\alpha/2}$ est "nettement négligeable"), une conclusion d'effet négligeable est généralement raisonnable. Il s'agit là d'une situation qui apparaît souvent conflictuelle, ou pour le moins embarassante, à l'utilisateur. Il n'y a pourtant là aucun paradoxe, sinon le fait que cette situation ne peut se produire que si la précision expérimentale est très bonne, c'est-à-dire si l'échelle e est très petite par rapport à d_{obs} (donc t=d_{obs}/e "très grand"). En pareil cas, selon le test de signification, même une différence observée faible pourra être significative (le test est très *puissant*). Mais la distribution fiducio-bayésienne sera peu dispersée autour de d_{obs}, ce qui montre qu'il s'agit en fait d'une situation privilégiée!

■ *Si* d_{obs} *est "nettement négligeable" et si le résultat est non-significatif*

Ces conditions, à elles seules, impliquent simplement que d_{obs} est inférieur à $d_{\alpha/2}$ mais peuvent correspondre à des valeurs $d_{obs}-d_{\alpha/2}$ et $d_{obs}+d_{\alpha/2}$ négligeables aussi bien qu'à des valeurs non négligeables. Dans cette situation le seuil de signification en lui-même n'apporte pas d'information utile: on ne peut rien conclure. Pourtant cette situation, comme la première, est assez souvent perçue comme favorable par l'utilisateur (en dépit des mises en garde envers les résultats non-significatifs), puisqu'il y a en apparence convergence du résultat du test et de la conclusion descriptive.

On voit ici les difficultés de ce que serait une pratique fiducio-bayésienne naïve consistant à se prononcer intuitivement sur l'importance de l'effet à partir de sa valeur observée et du résultat du test de signification usuel.

1.2.5 Recherche de conclusion d'effet notable

Dans le cadre fiducio-bayésien, nous pouvons poser dans les termes les plus directs la question de la *notabilité* de l'effet. Ici la différence observée entre le *placebo* et le nouveau médicament, d=+6.722, apparaît relativement élevée en regard de la moyenne générale observée 25.92 (dont elle vaut un peu plus du quart). Pour chercher à prolonger cette conclusion descriptive, nous regardons tout naturellement si nous avons une garantie suffisante que δ soit supérieure à une limite notable. Par exemple, nous avons trouvé précédemment

$$Pr(\delta>+2.136) = 0.95$$

LeBayésien

Pour exprimer la limite +2.136 en fonction de la moyenne générale observée:
activez l'option **exprimer les limites en fonction de**,
entrez pour **x** la valeur 25.917,
activez l'option **%**,
cliquez sur **calculer**.
Vous obtenez maintenant: δ>8.24%x

La limite 2.136 représente 8.24% de la moyenne générale observée. Si cette limite est jugée acceptable, nous pouvons conclure à un effet notable (avec la garantie 0.90). Sinon nous constatons que l'information contenue dans les

données ne permet pas de prolonger la conclusion descriptive d'un effet élevé.

Remarquons que nous pensons qu'une probabilité 0.90 constitue souvent un choix raisonnable pour une conclusion sur l'importance de l'effet. En pratique il est en effet tout à fait défendable d'admettre une garantie plus faible que 0.95 pour obtenir une conclusion beaucoup plus intéressante. Nous avons ici

$$Pr(\delta > +3.25) = 0.90$$

■ *Analyses des autres effets: Séquences, Périodes*

Il conviendrait bien entendu ici d'analyser également les effets des autres facteurs. En particulier, l'interprétation du résultat précédent sur la différence entre le *placebo* et le nouveau médicament, moyennée sur les deux périodes, dépendra de l'effet du facteur *Groupes* (ou *Séquences*). Nous y reviendrons au chapitre 5.

Remarquons encore que la comparaison des moyennes relatives à la première et à la seconde administration conduit à considérer la différence

$$(\delta\varepsilon^1 - \delta\varepsilon^2)/2 = \tfrac{1}{2}\delta\varepsilon^1 - \tfrac{1}{2}\delta\varepsilon^2$$

qui représente l'effet du facteur *Périodes*; son analyse s'effectue de la même manière que celle de la comparaison du *placebo* et du nouveau médicament.

1.3 Recherche d'une conclusion d'effet non notable

Dans les deux exemples précédents, il s'agissait de montrer la *supériorité* d'un traitement par rapport à un autre, donc de rechercher une conclusion d'effet *notable*, ce qui supposait d'obtenir un résultat "significatif" (*existence* de l'effet). Nous considérerons maintenant la situation dans laquelle il s'agit de montrer qu'un traitement n'est pas *notablement inférieur* à un autre traitement. Cette situation, si elle est moins usuelle, n'en est pas moins très importante dans la pratique, comme le montrent les deux exemples suivants.

• Le traitement habituel d'une certaine maladie nécessite une hospitalisation de dix jours. On envisage de réduire cette durée à cinq jours: il faut montrer que le traitement "court" est *équivalent* au traitement "long", en ce sens qu'*il n'est pas moins efficace* que ce dernier. On trouvera un exemple d'analyse bayésienne sur des données binaires pour une telle situation dans Lecoutre,

Derzko et Grouin (1995).

• Dans une expérience psychologique de résolution d'un certain type de problèmes sur les représentations probabilistes chez des adultes, Lecoutre et Durand (1988) constatent un biais systématique de réponse, qui apparaît résister à des connaissances de la théorie du calcul des probabilités. Si on compare des étudiants en psychologie à des étudiants en mathématiques, il s'agit ici de montrer que ces derniers *ne réussissent pas notablement mieux ces problèmes*.

1.3.1 Reprenons l'exemple de Student

Pour ne pas introduire ici un nouvel exemple, reprenons celui de Student, en modifiant les données, de telle manière que la différence moyenne observée entre le somnifère A et le somnifère B soit d=+0.28 heure, et que l'écart-type-corrigé s=1.230 soit inchangé.

patient	effet individuel (différence)
1	-0.1
2	+1.1
3	0
4	0
5	-1.3
6	-0.3
7	+0.5
8	-0.5
9	+3.3
10	+0.1
moyenne	+0.2800
éty-cor	1.2300

heures

Supposons que l'objectif soit de montrer que B, qui par exemple provoque moins d'effets secondaires indésirables, "n'est pas moins efficace que A". Nous devons donc nous donner une limite de tolérance: nous admettrons ici qu'un somnifère est effectivement meilleur que l'autre, en ce qui concerne le gain de sommeil, si la différence en sa faveur est d'au moins 1/2 heure.

Descriptivement, le somnifère B est donc pour ce critère "aussi bon" (en moyenne) que le somnifère A: le gain supplémentaire de sommeil observé en faveur de A (d=0.28 heure) est inférieur à 1/2. Il s'agit de chercher à conclure que la différence vraie δ est inférieure à 1/2 heure.

LeBayésien

Cliquez sur le bouton **moyenne**, ce qui fait apparaître la fenêtre de saisie des statistiques. Entrez les valeurs appropriées:

effectif	moyenne	écart-type
10	0.28	1.22999548

puis cliquez sur le bouton **calculer** pour obtenir la statistique de test t. Cliquez ensuite sur le bouton **ok**, ce qui vous ramène à la fenêtre principale, et cliquez sur la courbe relative à δ. Dans la fenêtre des énoncés, activez l'option **limite**, entrez la valeur 0.5 (que vous pouvez entrer sous la forme 1/2), puis cliquez sur **calculer**: vous obtenez (courbe du bas) pour δ>0.50 la garantie 0.293.

Refaites ensuite le même calcul pour un effectif de 80...

Le test t est non-significatif ($p/2$=0.24)

$$t = \frac{0.280}{0.389} = +0.720 \quad [9\ dl]$$

Mais un tel résultat dépend de l'effectif. Si les mêmes valeurs d et s avaient été obtenues, par exemple pour un échantillon de 80 patients, le test serait significatif (en faveur de A) au seuil unilatéral $p/2$=0.023

$$t = \frac{0.280}{0.138} = +2.036 \quad [79\ dl]$$

et nous ne pouvons donc pas l'utiliser comme un argument pour la conclusion recherchée de non notabilité.

1.3.2 Recherche d'une conclusion d'effet non notable

La procédure est en fait la même que pour la recherche d'une conclusion d'effet notable. Nous calculons dans les deux situations la probabilité que la différence δ entre A et B soit supérieure à +0.50.

• Si cette probabilité est élevée, nous pouvons conclure à la supériorité de A pour le critère de notabilité adopté: effet *notable*.
• Si cette probabilité est faible, nous pouvons au contraire conclure à la non supériorité de A: effet *non notable*.

Nous trouvons ici

pour 10 patients: $\delta \sim t_9(+0.28, 0.389^2)$ d'où $Pr(\delta > +0.50) = 0.293$

pour 80 patients: $\delta \sim t_{79}(+0.28, 0.138^2)$ d'où $Pr(\delta > +0.50) = 0.057$

Les deux distributions sont montrées dans la figure 1.5. Dans le cas de 10 patients (test non-significatif), l'information expérimentale est en fait insuffisante pour obtenir une conclusion. Dans le cas de 80 patients, le test significatif permet certes de conclure à une différence en faveur de A, mais nous pouvons montrer que cette différence est en fait limitée, inférieure au critère de tolérance d'une demi-heure ($\delta < +0.50$) avec une garantie 0.943.

L'analyse sur l'effet moyen δ peut être complétée par des analyses sur les effets individuels, de la même manière que pour l'effet notable.

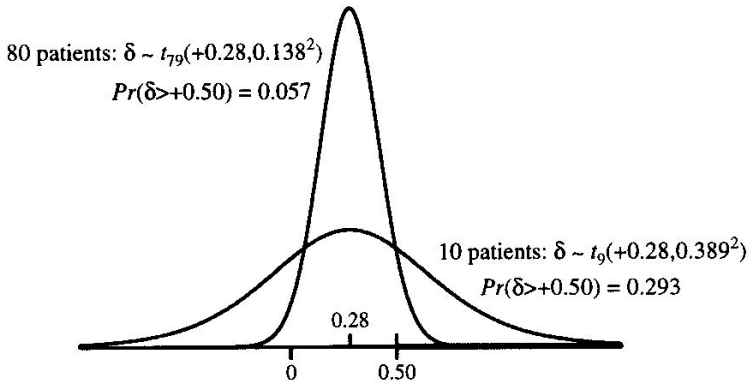


Figure 1.5 - Distributions fiducio-bayésiennes pour δ avec 10 et 80 patients

1.4 Recherche d'une conclusion d'effet négligeable

La recherche d'une conclusion d'effet *négligeable* est un problème incontournable dans la pratique. Elle peut répondre à l'objectif principal de l'expérience, comme dans l'exemple traité ci-après. Mais, même quand cet objectif est par exemple de mettre en évidence un effet *notable* ou *non notable* d'un certain facteur, il faut simultanément dans la quasi totalité des cas montrer que les effets de facteurs secondaires ou les effets d'interaction sont *négligeables* (ou du moins limités). Il s'agit ici de montrer que l'effet δ doit être *faible en valeur absolue*, ce qui distingue donc la notion d'effet "négligeable" de celle d'effet "non notable".

Cette problématique de l'effet négligeable apparaît dans la méthodologie des essais cliniques sous l'appellation d'*études de bioéquivalence* (dont un exemple est donné au chapitre ↳7).

1.4.1 Exemple de la validation d'un modèle additif

Cet exemple relève de la psychologie expérimentale. Il s'agit d'une expérience réalisée par Holender et Bertelson (1975) pour analyser les étapes de traitement de l'information dans une situation de temps de réaction. La tâche du sujet (adulte) est de réagir le plus rapidement possible à la présentation d'un signal. Deux facteurs expérimentaux principaux sont en jeu.

- Le facteur, noté ici A, est la fréquence du signal et comporte deux modalités: il y a un signal "fréquent" a_1 (de fréquence 0.75) et un signal "rare" a_2 (de fréquence 0.25).
- Le facteur, noté ici B, est la durée de la période préparatoire, c'est-à-dire la période qui précède l'apparition du signal et comporte également deux modalités: il y a une période "courte" b_1 (de durée 0.5 seconde) et un période "longue" b_2 (de durée 5 secondes).

Ces deux facteurs ont chacun un effet *notable* sur le temps de réaction: celui-ci est nettement plus petit pour le signal fréquent que pour le signal rare, et pour la période courte que pour la période longue. L'objectif de l'expérience est d'examiner l'hypothèse de l'*additivité* des effets des facteurs A et B, donc l'hypothèse d'un effet d'interaction δ entre les facteurs A et B ("modèle additif") *négligeable*.

Pour chaque sujet, il y a plusieurs essais, dans chacune des quatre conditions expérimentales, définies par les différentes combinaisons des modalités des facteurs A et B, ceci pendant plusieurs sessions d'expérience. A titre illustratif, nous introduisons un facteur supplémentaire (non présent dans

l'expérience originale), en supposant que les sujets sont répartis en trois groupes d'effectifs inégaux, avec respectivement 5, 3 et 4 sujets. Les données de base traitées ici sont les moyennes des essais (temps exprimés en millisecondes), pour chacun des douze sujets, dans chacune des quatre conditions, pour l'une des sessions.

	a1b1	a2b1	a1b2	a2b2	effet individuel (interaction)	
s1 g1	387	435	416	473	+ 9	
s2 g1	321	336	343	368	+10	moyenne d^{g1}=-3.00
s3 g1	333	362	358	390	+ 3	éty-cor d_{g1}=23.6326
s4 g1	344	430	352	393	-45	
s5 g1	368	432	432	504	+ 8	
s6 g2	357	367	394	411	+ 7	
s7 g2	336	346	340	421	+71	moyenne d^{g2}=+23.00
s8 g2	387	454	438	496	- 9	éty-cor s_{g2}=42.3320
s9 g3	345	408	417	479	- 1	
s10 g3	358	389	372	407	+ 4	moyenne d^{g3}=-19.75
s11 g3	317	375	341	393	- 7	éty-cor s_{g3}=37.1068
s12 g3	386	510	464	513	-75	

Nous définirons l'effet d'interaction entre les facteurs A et B comme la "différence des différences": différence entre l'effet de A pour b1 et l'effet de A pour b2 qui, par symétrie, est encore la différence entre l'effet de B pour a1 et l'effet de B pour a2. Puisque chaque sujet est soumis à l'ensemble des conditions, l'effet d'interaction observé peut être calculé pour chaque sujet, et la situation est alors une généralisation immédiate des exemples traités précédemment.

Par exemple, pour le sujet s1, nous obtenons l'effet observé (individuel)

$$d^{s1} = 387 - 435 - 416 + 473 = +9$$

1.4.2 Analyse descriptive

(1) L'effet d'interaction *moyen* observé est très faible, que l'on considère la moyenne *équipondérée* (indiquée par la mention [E])

$$d_{[E]} = \frac{-3.00 + 23.00 - 19.75}{3} = +0.083$$

ou la moyenne *pondérée par les effectifs* (indiquée par la mention [N]

$$d_{[N]} = \frac{5\times(-3.00) + 3\times(+23.00) + 4\times(-19.75)}{5 + 3 + 4} = -2.083$$

des effets individuels. A titre didactique, nous examinerons ici les solutions correspondant à chacun de ces deux choix.

(2) L'écart-type-corrigé *intragroupe* moyen des effets d'interaction individuels s est la moyenne *quadratique* pondérée par les nombres de degrés de liberté des écarts-types de chacun des trois groupes. Nous calculons

$$s^2 = \frac{4\times23.6326^2 + 2\times42.3320^2 + 3\times37.1068^2}{4 + 2 + 3} \quad \text{d'où } s = 33.248$$

La valeur s=33.248 est élevée et traduit la forte variabilité des effets individuels (de -75 *ms* à +71 *ms*). Par suite l'effet calibré moyen est très petit

$$d_{cal[E]} = +0.003 \quad \text{et} \quad d_{cal[N]} = -0.063$$

1.4.3 Analyse inductive

Le problème est encore celui de l'inférence sur une combinaison linéaire de moyennes parentes de groupes indépendants, et les procédures sont les mêmes que dans les exemples précédents. Nous nous intéressons ici à la moyenne générale vraie, suivant le choix de la pondération

$$\delta_{[E]} = (\delta^{g1}+\delta^{g2}+\delta^{g3})/3 = \frac{1}{3}\delta^{g1} + \frac{1}{3}\delta^{g2} + \frac{1}{3}\delta^{g3}$$

$$\delta_{[N]} = (5\times\delta^{g1}+3\times\delta^{g2}+4\times\delta^{g3})/12 = \frac{5}{12}\delta^{g1} + \frac{3}{12}\delta^{g2} + \frac{4}{12}\delta^{g3}$$

LeBayésien

Maintenant que vous avez assimilé la manière d'utiliser le programme, vous trouverez peut être fastidieux de devoir entrer les informations pour chaque exemple. Pour alléger votre tâche (et éviter les risques d'erreurs), tous les exemples de ce manuel sont directement accessibles: dans la fenêtre principale, cliquez sur le bouton **exemples**, ce qui affiche la liste des exemples et sélectionnez **modèle additif**.

Effectuez le calcul pour l'option **moyenne équipondérée**. Vous obtenez d=0.083, s=33.248, etc.

Cliquez ensuite sur **ok**, puis dans la fenêtre principale sur la courbe de l'effet δ. Remarquez que la distribution est très dispersée.

Sélectionnez la garantie 0.90 et cliquez sur **calculer**. Vous obtenez pour la courbe **négligeable**: $|\delta|<17.98$. Pour cette même garantie, vous pouvez obtenir par exemple pour l'écart-type et l'effet calibré

$$\sigma>26.03 \quad \text{et} \quad |\delta_{cal}|<0.49$$

□ *Remarque*

Si dans la fenêtre principale vous cliquez sur l'une des courbes avec le bouton secondaire de la souris, le calcul des énoncés est effectué automatiquement, ceci pour les options (garantie/limite) précédemment utilisées.

Revenez à la fenêtre des statistiques (bouton **G / 0** dans la fenêtre principale) et effectuez les mêmes calculs pour l'option **moyenne pondérée**. Etc.

Nous en déduisons les t de Student correspondants

$$t_{[E]} = \frac{+0.083}{9.809} = +0.008 \quad \text{et} \quad t_{[N]} = \frac{-2.083}{9.598} = +0.217 \quad [9\ dl]$$

à partir des constantes

$$b^2_{[E]} = \frac{(1/3)^2}{5} + \frac{(1/3)^2}{3} + \frac{(1/3)^2}{4} = 47/540$$

$$b^2_{[N]} = \frac{(5/12)^2}{5} + \frac{(3/12)^2}{3} + \frac{(4/12)^2}{4} = 1/12$$

et des échelles

$$e_{[E]} = \sqrt{47/540}\times 33.248 = 9.809 \quad \text{et} \quad e_{[N]} = \sqrt{1/12}\times 33.248 = 9.598$$

Ces tests sont non-significatifs, aux seuils *bilatéraux* observés respectifs p=0.497 et p=0.416. Ce résultat reflète ici, à la fois le fait que l'effet observé est très petit et que l'imprécision expérimentale (traduite par e) est grande: effectif faible et écart-type élevé.

Nous avons les distributions fiducio-bayésiennes (figure 1.6)

$$\delta_{[E]} \sim t_{10}(+0.083, 9.809^2) \quad \text{et} \quad \delta_{[N]} \sim t_{10}(-2.083, 9.598^2)$$

Pouvons nous généraliser la conclusion descriptive d'un effet d'interaction observé faible?

Dans le cadre fiducio-bayésien, la question de la *négligeabilité* de l'effet se pose dans les termes les plus directs: nous regardons si nous avons une

garantie suffisante que l'effet vrai δ soit inférieur en valeur absolue à une limite négligeable. Pour chacune des deux distributions, nous trouvons ici, pour la garantie 0.90, que δ est compris entre -18 et +18

$$Pr(|\delta|<18) = 0.90$$

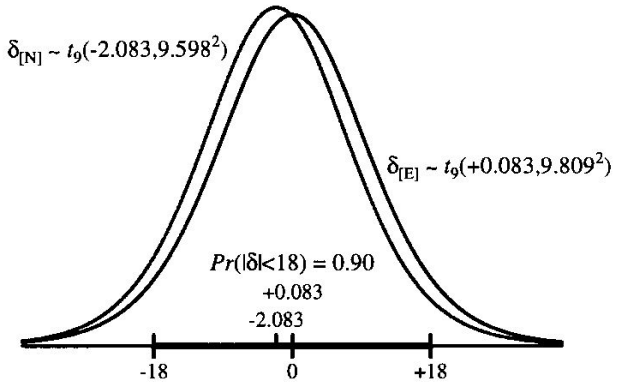


Figure 1.6 - Distributions fiducio-bayésiennes pour l'effet d'interaction

Comparativement aux effets principaux des facteurs A et B, la limite obtenue 18 *ms* peut apparaître relativement faible. Une analyse de ces effets principaux conduit en effet aux énoncés fiducio-bayésiens pour la garantie 0.90

- pour A (fréquence du signal): $\delta_{a2,a1[E]}>+40$ et $\delta_{a2,a1[N]}>+40$
- pour B (durée de la période préparatoire): $\delta_{b2,b1[E]}>+26$ et $\delta_{b2,b1[N]}>+25$

Mais, s'agissant ici du problème de la validation d'un modèle, il s'agit de conclure à un écart tolérable à l'hypothèse stricte d'une interaction nulle, et la limite 18 *ms* apparaîtra alors sans doute trop élevée.

Même si le modèle additif était jugé acceptable "en moyenne", l'écart-type σ apparaîtrait ici trop élevé pour pouvoir étendre cette conclusion au niveau individuel; nous avons

$$Pr(\sigma>26) = 0.90$$

CHAPITRE 2

DE L'ANOVA AUX PROCÉDURES FIDUCIO-BAYÉSIENNES

Dans ce chapitre, nous poursuivrons les illustrations des procédures fiducio-bayésiennes pour les combinaisons linéaires de moyennes de groupes indépendants. Nous expliciterons en particulier les liens avec l'analyse de variance univariée (ANOVA), qui est la méthode traditionnelle pour traiter ces plans. Les lecteurs familiarisés avec cette technique pourront apprécier le fait que les solutions proposées en sont une extension immédiate. Les autres lecteurs verront que la connaissance de l'analyse de variance n'est pas indispensable, puisque l'approche utilisée ici permet d'en retrouver les résultats.

2.1 Combinaisons linéaires de moyennes de groupes indépendants

2.1.1 Exemple de données anthropologiques: Les crânes de Rao

Cet exemple est emprunté à Rao (1965); on dispose de la largeur, en unités arbitraires, de 142 crânes appartenant à trois séries stratigraphiques différentes. Le plan d'analyse est un plan en trois groupes indépendants non équilibré, que nous noterons $S<G_3>$ ("*Sujets emboîtés* dans les *Groupes*"). Il y a respectivement 83, 51 et 8 "sujets" (ici *crânes*) par groupe. On se propose de comparer, d'une part l'ensemble des séries 2 et 3 à la série 1 (comparaison à un degré de liberté notée "g2_g3,g1"), et d'autre part les séries 2 et 3 entre elles (comparaison à un degré de liberté "g2,g3").

Les totaux et moyennes pour chacun des trois groupes sont les suivants

série	total	effectif	moyenne
g1	11277	83	135.867
g2	7049	51	138.216
g3	1102	8	137.750

Descriptivement, il apparaît une différence importante entre la première série et l'ensemble des deux autres séries, tandis que ces deux dernières séries diffèrent relativement peu l'une de l'autre. Plus précisément la différence observée entre la moyenne (pondérée par les effectifs) des séries 2 et 3 et la moyenne de la série 1 vaut

$$d = \frac{7049+1102}{51+8} = \frac{51}{59}\,138.216 + \frac{8}{59}\,137.750 - 135.867 = +2.285$$

tandis que la différence entre les séries 2 et 3 vaut seulement

$$d = 138.216 - 137.750 = +0.466$$

Le tableau d'analyse de variance peut être reconstitué à partir des totaux et de la somme-des-carrés *intragroupe* $sc_{S(G)}$=4378.05 fournie par Rao. On notera que la variance-corrigée intragroupe moyenne s^2 est égale au carré-moyen intragroupe $cm_{S(G)}$=31.497. Ce tableau suppose le *modèle équinormal* usuel: "dans chaque groupe g les observations sont indépendantes et équidistribuées $N(\mu g,\sigma^2)$".

comparaison	*dl*	*sc*	*cm*	*F* [d1]	*p*
inter G	2	181.570	90.785		
g2_g3,g1 [N]	1	180.070	180.070	5.717 [1,139]	0.018
g2,g3	1	1.50	1.50	0.048 [1,139]	0.827
intra S(G)	139	4378.05	31.497		

La mention [N] indique que nous avons utilisé la solution "pondérée par les effectifs"

(certains résultats diffèrent de ceux de Rao qui sont erronés)

■ *Remarque sur la décomposition de la comparaison globale des séries*

la somme-des-carrés *intergroupe* sc_G=181.750 est ici la somme des sommes-des-carrés partielles $sc_{g2_g3,\,g1\,[N]}$=180.070 et $sc_{g2,\,g3}$=1.50; on dira que la décomposition est *orthogonale*. Ce ne serait pas le cas si nous avions considéré la moyenne équipondérée des séries 2 et 3, qui donne (la solution est indiquée

par la mention [E])

$$d = 137.983 - 135.867 = +2.115 \quad \text{et} \quad sc_{g2_g3,\,g1\,[E]} = 92.838$$

Pour d'autres considérations sur cette question de l'orthogonalité, voir le chapitre ↳5.

LeBayésien

Dans la fenêtre principale, cliquez sur le bouton **exemples**, et sélectionnez **crânes de Rao**. La fenêtre de saisie des statistiques est affichée. Remarquez ici que l'écart-type moyen s=5.6122 est utilisé pour chacun des trois groupes (les écarts-types des groupes n'étant pas fournis).

Pour la comparaison "g2_g3,g1", nous nous intéressons à la **combinaison linéaire** des moyennes de coefficients [-1 +51/59 +8/59] qui est préselectionnée.

Cliquez sur le bouton **calculer**. Vous obtenez d=2.285, s=5.612, etc.

En dessous du résultat du test de signification (t=2.391 p/2=0.0091), se trouve une **liste déroulante**. Pour afficher cette liste cliquez sur la **flèche** (à droite) pour la "dérouler", ou sur le **losange** (à gauche) pour l'afficher entièrement dans une fenêtre. Vous trouvez dans cette liste le détail des résultats, avec les notations utilisées ici (d, s, q, b, e), ainsi que les statistiques de l'analyse de variance (**ANOVA**).

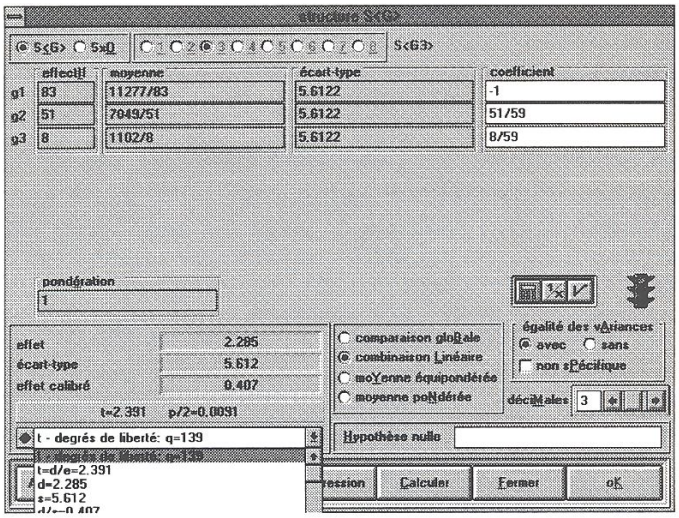
Pour les procédures fiducio-bayésiennes, cliquez sur le bouton **ok**, puis procédez comme dans les exemples précédents...

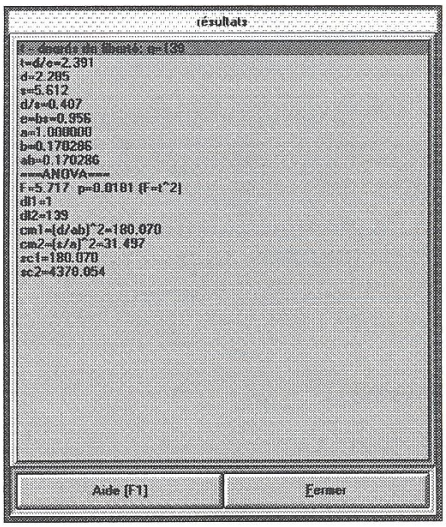
Revenez ensuite à la fenêtre de saisie des statistiques (bouton **G / O** de la fenêtre principale). Entrez maintenant les coefficients [0 +1 -1] pour la comparaison "g2,g3". Notez ici dans le cadre **égalité des variances** l'option **non spécifique**, qui correspond au modèle supposant l'égalité des variances des trois groupes (voir la section suivante).

Cliquez sur le bouton **calculer**...

2.1.2 Tests de signification

Nous pouvons retrouver les rapports F du tableau d'analyse de variance, en généralisant la procédure exposée au chapitre ↳1 dans le cas de un et deux groupes.


structure S<G>
S<G>
effectif
moyenne
écart-type
coefficient
g1 83 11277/83 5.6122 -1
g2 51 7049/51 5.6122 51/59
g3 8 1102/8 5.6122 8/59
pondération
1
effet 2.285
écart-type 5.612
effet calibré 0.407
t=2.391 p/2=0.0091
comparaison globale
combinaison Linéaire
moYenne équipondérée
moyenne poNdérée
égalité des vAriances
avec sans
non sPécifique
décimales 3
t - degrés de liberté: q=139
Hypothèse nulle
t=d/e=2.391
d=2.285
s=5.612
Calculer
Fermer
oK


résultats
t=d/e=2.391
d=2.285
s=5.612
d/s=0.407
e=bs=0.956
a=1.000000
b=0.170286
ab=0.170286
===ANOVA===
F=5.717 p=0.0181 (F=t^2)
dl1=1
dl2=139
cm1=(d/ab)^2=180.070
cm2=(s/a)^2=31.497
sc1=180.070
sc2=4370.054
Aide (F1)
Fermer

■ *Comparaison* "g2_g3,g1"

L'effet vrai de la comparaisons est $\delta = \frac{51}{59}\mu_{g2} + \frac{8}{59}\mu_{g3} - \mu_{g1}$.

Nous avons la statistique de test t de l'hypothèse nulle δ=0

$$t = \frac{d}{e} = \frac{d}{bs} = \frac{+2.285}{0.956} = +2.391 \quad [q=139\ dl]$$

où b est obtenu à partir des coefficients v_g affectés aux groupes et des effectifs

$$b^2 = \sum_{g \in G} \frac{v_g^2}{f_g} = \frac{(-1)^2}{83} + \frac{(51/59)^2}{51} + \frac{(8/59)^2}{8} = \frac{142}{4897} = 0.1703^2$$

Le rapport F est le carré de t

$$F = \frac{cm_{g2_g3,g1}}{cm_{S(G)}} = t^2 = 5.717$$

■ *Comparaison* "g2,g3"

L'effet vrai de la comparaisons est $\delta = \mu_{g2} - \mu_{g3}$

Nous avons les statistiques de test t et F de l'hypothèse nulle δ=0

$$t = \frac{d}{e} = \frac{d}{bs} = \frac{+0.466}{2.134} = +0.218 \quad [q=139\ dl]$$

$$\text{avec } b^2 = \sum_{g \in G} \frac{v_g^2}{f_g} = \frac{0^2}{83} + \frac{(+1)^2}{51} + \frac{(-1)^2}{8} = \frac{59}{408} = 0.3802^2$$

$$F = \frac{cm_{g2,g3}}{cm_{S(G)}} = t^2 = 0.048$$

■ *Remarque sur le choix de* s

La comparaison "g2,g3" ne fait pas intervenir le premier groupe; si nous disposions des écarts-types-corrigés observés dans chacun des groupes, nous aurions une solution *spécifique* consistant à calculer l'écart-type moyen s à partir seulement des groupes g2 et g3. Cette solution aurait l'avantage de reposer sur des conditions de validité moins exigeantes, puisqu'il suffirait de supposer l'égalité des variances de ces deux groupes (voir ↳7.7.2).

2.1.3 Procédures fiducio-bayésiennes

Nous en déduisons immédiatement les distributions fiducio-bayésiennes relatives à δ et à $\delta_{cal}=\delta/\sigma$ pour les comparaisons "g2_g3,g1" et "g2,g3"

"g2_g3,g1" $\quad \delta \sim t_{139}(+2.285, 0.956^2)$ et $\delta_{cal} \sim \Lambda'_{139}(+0.407, 0.170^2)$

"g2,g3" $\quad \delta \sim t_{139}(+0.466, 2.134^2)$ et $\delta_{cal} \sim \Lambda'_{139}(+0.083, 0.380^2)$

Avec la garantie fiducio-bayésienne 0.90: pour la comparaison "g2_g3,g1", l'effet vrai δ dépasse +1.05 (et $\delta_{cal}>+0.19$); tandis que pour la comparaison "g2,g3", δ est inférieur en valeur absolue à 3.62 (et $|\delta_{cal}|<0.64$).

Il est clair que, quel que soit le critère adopté, nous ne pouvons conclure à la fois que l'effet est notable dans le premier cas et négligeable dans le second cas. Ceci veut dire que l'information apportée par les données n'est pas suffisante pour prolonger inductivement les conclusions descriptives (au moins en ce qui concerne la comparaison "g2,g3").

■ *Remarque sur l'utilisation de l'effet calibré*

Pour une comparaison *intra-sujets* (comme dans l'exemple de Student) nous avons utilisé l'effet calibré δ_{cal} pour étudier les effets individuels, ce qui n'est pas pertinent pour une comparaison *inter-sujets* (comme dans l'exemple présent). Mais nous pouvons également utiliser δ_{cal} pour nous prononcer sur l'importance de l'effet moyen (*relativement à* σ); nous y reviendrons en ↳2.2.

2.1.4 Reconstitution du tableau ANOVA

A partir de l'analyse fiducio-bayésienne, nous pouvons reconstituer le tableau ANOVA. Pour chaque comparaison (ou *source de variation*) systématique à un degré de liberté, les statistiques de l'analyse de variance se déduisent des valeurs d, s et e. Les carrés-moyens du numérateur et du dénominateur du rapport F sont respectivement proportionnels à d^2 et s^2.

Par exemple, pour la comparaison "g2_g3,g1", nous avons (les calculs sont effectués avec la précision nécessaire)

$$cm_{g2_g3,g1} = sc_{g2_g3,g1} = \left(\frac{d}{b}\right)^2 = \left(d\,\frac{s}{e}\right)^2 = \left(2.285\,\frac{5.612}{0.956}\right)^2 = 180.07$$

$$cm_{S(G)} = s^2 = 5.612$$

$$F = \left(\frac{d}{e}\right)^2 = \left(\frac{2.285}{0.956}\right)^2 = 5.717$$

Ces formules sont à valeur générale pour un plan en groupes indépendants. Nous verrons qu'elles s'étendent aux plans plus complexes à plusieurs facteurs systématiques, incluant le cas des *cross-over*, *blocs incomplets*, *mesures répétées*... (voir les chapitres §4 et §5).

2.1.5 Du tableau ANOVA à l'analyse fiducio-bayésienne

Inversement, à partir des sommes-des-carrés ou des carrés-moyens, nous pouvons retrouver la valeur absolue de d et s

$$|d| = b\sqrt{sc_{g2_g3,g1}} = b\sqrt{cm_{g2_g3,g1}} \quad \text{et} \quad s = \sqrt{sc_{S(G)}/q} = \sqrt{cm_{S(G)}}$$

et par suite la distribution fiducio-bayésienne, même si nous ne disposons pas des données complètes (et même pas des statistiques descriptives d et s). Un exemple est donné en §7.1.3.

2.1.6 Du test de signification à l'analyse fiducio-bayésienne: Inférence sur δ

Nous pouvons même déduire la distribution fiducio-bayésienne relative à δ de l'effet observé d et de la statistique de test t ou F, puisque

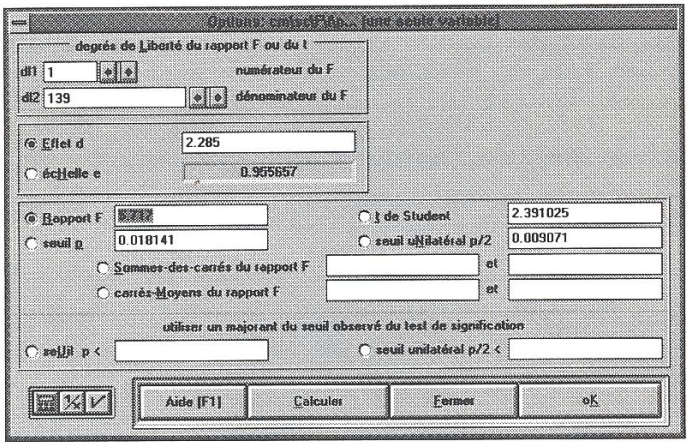
$$\delta \sim t_q(d,e^2) \quad \text{avec } e = \frac{d}{t} = \frac{|d|}{\sqrt{F}} \quad \text{si } d\neq 0 \text{ (donc } t\neq 0 \text{ et } F\neq 0)$$

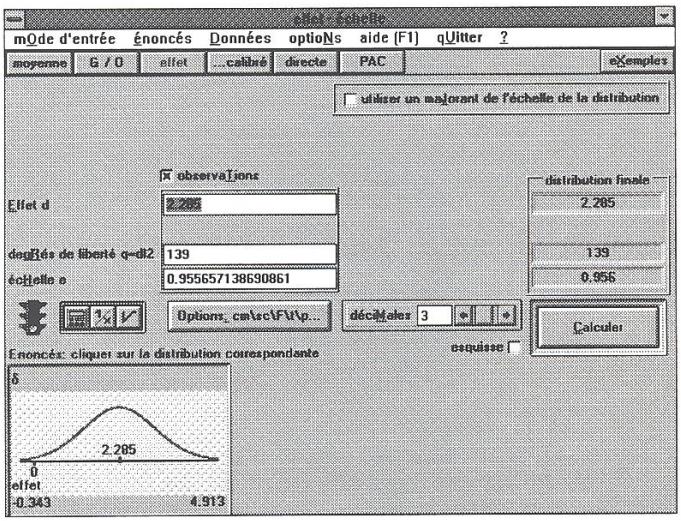
Ceci permet notamment de réanalyser des résultats, à partir des informations trouvées dans un rapport ou dans une publication: il suffit de connaître l'effet observé d et la valeur prise par la statistique de test (t ou $F=t^2$), pourvu que d soit non nul. Par exemple nous pouvons immédiatement en déduire pour δ l'intervalle de crédibilité centré sur d avec la garantie γ (§1.2.2)

$$[d-\frac{d}{t}t_{1-\gamma}\ ,\ d+\frac{d}{t}t_{1-\gamma}] \equiv [d(1-\frac{1}{t})t_{1-\gamma}\ ,\ d(1+\frac{1}{t})t_{1-\gamma}]$$

■ *Attention à la précision*

Il se pose ici le problème de la précision des calculs. On commencera donc dans la mesure du possible par recalculer avec la meilleure précision disponible d (à partir des moyennes ou mieux des totaux) et F (à partir des carrés-moyens, ou mieux des sommes-des-carrés). On fera plus particulièrement attention aux petites valeurs, qui peuvent conduire à des résultats erronés comme le montre l'exemple suivant. Supposons que nous disposions des valeurs d=0.01 et t=0.01, fournies avec deux décimales (ce qui, utilisé tel quel, donnerait


Options: calculs pour une seule variable
degrés de Liberté du rapport F ou du t
dl1 1
numérateur du F
dl2 139
dénominateur du F
Effet d
2.285
échElle e
0.955657
Rapport F
5.212
t de Student
2.391025
seuil p
0.018141
seuil uNilatéral p/2
0.009071
Sommes-des-carrés du rapport F
et
carrés-Moyens du rapport F
et
utiliser un majorant du seuil observé du test de signification
seUil p <
seuil unilatéral p/2 <
Aide [F1]
Calculer
Fermer
oK


effet - échelle
mOde d'entrée énoncés Données optioNs aide (F1) qUitter ?
moyenne G / O effet ...calibré directe PAC
eXemples
utiliser un majorant de l'échelle de la distribution
observaTions
Effet d
2.285
degRés de liberté q=dl2
139
échElle e
0.955657138690861
distribution finale
2.285
139
0.956
Options_cm\src\F\t\p...
déciMales 3
Calculer
esquisse
Enoncés: cliquer sur la distribution correspondante
δ
2.285
0
effet
-0.343
4.913

l'échelle e=d/t=1). En fait, nous avons pour trois décimales

$$0.005 \le d \le 0.015 \quad \text{et} \quad 0.005 \le t \le 0.015$$

et nous pouvons donc simplement dire que l'échelle e de la distribution est comprise entre

$$e = 0.005/0.015 = 1/3 \quad \text{et} \quad e = 0.015/0.005 = 3$$

Si nous n'avons pas d'autre précision, la prudence impose de nous placer dans le plus mauvais cas possible, correspondant à l'incertitude maximale, soit ici e=3, qui constitue un *majorant* de l'échelle de la distribution fiducio-bayésienne.

LeBayésien

Dans la fenêtre principale cliquez sur le bouton **effet**. Ce mode d'entrée permet d'entrer directement la distribution fiducio-bayésienne relative à l'effet δ si ses caractéristiques sont connues, ou encore de la définir à partir des résultats de l'analyse de variance (ou d'un test t). La présentation de la fenêtre est modifiée en conséquence; en particulier il n'y a plus d'écart-type, ni par suite l'option distribution initiale.

Cliquez ensuite sur le bouton **options: cm\sc\F\t\p**, ce qui fait apparaître une nouvelle fenêtre. Entrez l'effet (2.285) de la comparaison "g2_g3,g1" (sélectionnez s'il y a lieu l'option **effet d**), les nombres de degrés de liberté du rapport F (1 et 139), et le rapport F (5.717). Vérifiez que l'option **rapport F** est sélectionnée (c'est l'option par défaut), et cliquez sur le bouton **calculer**. L'échelle $e=|d|/\sqrt{F}$ de la distribution est calculée: vous retrouvez (avec 3 décimales) e=0.956.

Ici on améliorera la précision, en recalculant d à partir des totaux (d=2.28507249...) et en utilisant les sommes-des-carrés. Sélectionnez l'option **sommes-des-carrés du rapport F**, et entrez les valeurs respectives 180.070 et 4378.05, puis cliquez sur le bouton **calculer**. Etc.

Cliquez sur le bouton **OK**, ce qui vous ramène à la fenêtre principale, avec la courbe de la distribution relative à δ, etc.

□ *Remarque*

L'option **échelle e** permet d'effectuer le calcul inverse, donc si on connaît e (et pas d) de recalculer d (ou au moins sa valeur absolue), à partir des résultats de l'analyse de variance: $d=et$ et $|d|=e\sqrt{F}$. Mais cette situation est beaucoup moins usuelle.

2.1.7 Du test de signification à l'analyse fiducio-bayésienne: Inférence sur δ_{cal}

■ *Si l'effet observé* d_{cal} *est connu*

Il est de la même manière immédiat de déduire la distribution fiducio-bayésienne relative à l'effet calibré δ_{cal}, à partir de l'effet observé $d_{cal}=d/s$ et de la statistique de test t ou F, puisque

$$\delta_{cal} \sim \Lambda'_q(d_{cal}, b^2) \quad \text{avec } b = \frac{d_{cal}}{t} = \frac{d_{cal}}{\sqrt{F}} \quad \text{si } d_{cal}\neq 0 \text{ (donc } t\neq 0 \text{ et } F\neq 0)$$

Les mêmes remarques que dans le cas précédent s'appliquent.

LeBayésien

Dans la fenêtre principale cliquez sur le bouton **...calibré**. Ce mode d'entrée permet d'entrer directement la distribution fiducio-bayésienne relative à l'effet calibré δ_{cal} si ses caractéristiques sont connues, ou encore de la définir à partir des résultats de l'analyse de variance (ou d'un test t).

Cliquez ensuite sur le bouton **options: cm\sc\F\t\p**. Sélectionnez s'il y a lieu l'option **effet calibré d**. Entrez l'effet calibré (2.285/5.612 que vous pouvez entrer directement sous forme de rapport) de la comparaison "g2_g3,g1", les nombres de degrés de liberté du rapport F (1 et 139), et le rapport F (5.717). Sélectionnez s'il y a lieu l'option **rapport F**, et cliquez sur le bouton **calculer**. La constante $b=|d_{cal}|/\sqrt{F}$ est calculée: vous retrouvez (avec 3 décimales) e=0.170.

Etc.

■ *Si l'effet observé* d_{cal} *n'est pas connu*

Plus encore: même si nous ne disposons que du tableau ANOVA, ou que de la statistique de test, ou même que du seuil observé, la distribution fiducio-bayésienne relative à δ_{cal} peut être reconstituée, pourvu que les informations sur le plan d'expérience soient suffisantes pour recalculer la constante b. Nous en déduisons en effet l'effet calibré observé

$$d_{cal} = bt \quad \text{ou} \quad |d_{cal}| = b\sqrt{F} \quad \text{si } F\neq 0$$

Si cela est nécessaire, le signe sera généralement indiqué par le contexte. Quand ce n'est pas le cas, il s'agira généralement d'un effet faible, et son

signe ne sera pas nécessaire pour rechercher une conclusion d'effet négligeable.

En pratique la procédure est cependant moins immédiate que dans le cas où l'effet est connu, puisqu'il faut déterminer la constante b.

LeBayésien

Dans la fenêtre principale cliquez sur le bouton **...calibré**, puis sur le bouton **options: cm\sc\F\t\p**. Sélectionnez l'option **constante b**, et entrez la constante b de la comparaison "g2_g3,g1"

$$\sqrt{142/4897} = 0.170286...$$

les nombres de degrés de liberté du rapport F (1 et 139), et le rapport F (5.717). Sélectionnez s'il y a lieu l'option **rapport F**, et cliquez sur le bouton **calculer**. La valeur absolue de l'effet calibré $|d_{cal}|=b/\sqrt{F}$ est calculée: vous retrouvez (avec 3 décimales) 0.407.

Etc.

□ *Remarque*

Le bouton **calculer la constante b** permet de déterminer automatiquement b pour les problèmes de comparaisons correspondants aux plans usuels élémentaires équilibrés S_n, $S_n<G_2>$, $S_n\times O_2$ et $S_n<G_2>\times O_2$ (sur ces plans, voir le chapitre ↳4).

2.1.8 Du seuil de signification à l'analyse fiducio-bayésienne

■ *Si l'effet observé est connu*

Il suffit même, pour obtenir la distribution fiducio-bayésienne relative à δ (*resp.* δ_{cal}), de connaître l'effet observé d (*resp.* d_{cal}) et le seuil de signification observé p (ou $p/2$) du test t ou F (avec le nombre de degrés de liberté), puisque nous pouvons déduire de ce dernier la valeur de la statistique de test.

Par exemple, pour la comparaison "g2_g3,g1", de p=0.018 nous déduisons F=5.731, d'où e=0.954. Le calcul est bien entendu moins précis, et il sera préférable d'utiliser un majorant du seuil (ici p<0.0185): voir ci-après.

Plus encore en effet, il suffit de connaître le seuil observé sous forme d'inégalités. Ceci est typiquement le cas quand on sait simplement que le test

est "significatif au seuil α" (soit $p \leq \alpha$). On en déduit un *majorant* de l'échelle, et les énoncés fiducio-bayésiens seront obtenus, à l'instar du seuil du test, sous forme d'inégalités.

LeBayésien

Revenez à la situation considérée en ↳2.1.6. Dans la fenêtre de définition de l'échelle (avec l'effet 2.28507249 et dl2=139), sélectionnez maintenant l'option **seuil p** et entrez la valeur 0.018. Cliquez sur le bouton **calculer**. Vous obtenez F=5.731 (au lieu de F=5.717 qui correspond en fait à p=0.01814...) et l'échelle 0.954 (au lieu de 0.956).

Sélectionnez maintenant l'option **seuil p <** et entrez la valeur 0.05. Cliquez sur le bouton **calculer**. Vous trouvez que l'échelle est *inférieure* à 1.1557... (pour un F correspondant de 3.909). Cliquez sur le bouton **ok**, ce qui vous ramène à la fenêtre principale, avec la courbe de la distribution relative à δ. Cliquez sur celle-ci pour accéder à la fenêtre des énoncés et effectuez le calcul pour la garantie 0.90. Vous obtenez pour la courbe **notable**: δ>0.797. La garantie associée à cet énoncé est *supérieure* à 0.90.

Par exemple, pour la comparaison "g2_g3,g1", de p<0.05 nous déduisons F>3.909, d'où e>1.156. Pour e=1.156 nous aurions

$$\delta \sim t_{139}(+2.285, 1.156^2) \quad \text{d'où par exemple } Pr(\delta > +0.797) = 0.90$$

Pour e>1.156 nous pouvons donc énoncer que l'effet δ est supérieur à 0.797 avec une garantie *supérieure* à 0.90.

Si nous savons en outre que p>0.01, nous déduisons F<6.821 et e<0.875. Pour e=0.875, nous aurions $Pr(\delta > +0.797)$=0.954, d'où l'encadrement

$$0.90 < Pr(\delta > +0.797) < 0.954$$

■ *Si l'effet observé n'est pas connu*

Dans ce cas, si nous connaissons l'échelle ou la constante b, nous pouvons encore:

• si nous savons que $p \leq \alpha$, déduire un *minorant* de l'effet observé, et par suite rechercher s'il y a lieu une conclusion d'effet notable (mais pas d'effet négligeable);

• si nous savons que $p \geq \alpha$, déduire un *majorant* de l'effet observé, et par suite rechercher s'il y a lieu une conclusion d'effet négligeable (mais pas d'effet

notable).

LeBayésien

Dans la fenêtre principale cliquez sur le bouton **...calibré**, puis sur le bouton **options: cm\sc\F\t\p**. Sélectionnez l'option **constante b**, et entrez la constante b de la comparaison "g2_g3,g1"

$$\sqrt{142/4897} = 0.170286...$$

les nombres de degrés de liberté du rapport F (1 et 139), et le rapport F (5.717).

Sélectionnez l'option **seuil p** et entrez pour p la valeur observée 0.018 (avec 3 décimales): vous obtenez pour $|d_{cal}|$ avec 3 décimales 0.408 (au lieu de 0.407).

Supposons que nous sachions seulement que $p \leq 0.05$. Effectuez le calcul pour la valeur $p=0.05$: vous obtenez 0.337 (soit donc $|d_{cal}| \geq 0.337$). Effectuez l'inférence relative à δ_{cal} pour cette valeur (bouton **ok**: un message vous rappelle que seul le signe de l'effet est déterminé). Etc.

Vous obtenez pour la garantie 0.90 l'énoncé (courbe **notable**): $\delta > 0.116$. La garantie associée à cet énoncé est en fait *supérieure* à 0.90.

Par exemple, pour la comparaison "g2_g3,g1" (b=0.170), de $p<0.05$ nous déduisons $F>3.909$, d'où $|d_{cal}| = b\sqrt{F} \geq 0.337$. Pour $d_{cal}=+0.337$ (si nous savons que l'effet est positif), nous aurions

$$\delta_{cal} \sim \Lambda'_{139}(+0.337, 0.170^2) \quad \text{d'où par exemple } Pr(\delta_{cal} > +0.116) = 0.90$$

Pour $d_{cal} \geq +0.337$, nous pouvons donc énoncer que l'effet δ_{cal} est supérieur à 0.116 avec une garantie *d'au moins* 0.90.

De la même manière, pour la comparaison "g2,g3" (b=0.380), de $p>0.10$ nous déduisons $F<2.742$, d'où $|d_{cal}| = b\sqrt{F} \leq 0.630$. Pour $d_{cal}=0.630$, nous aurions

$$\delta_{cal} \sim \Lambda'_{139}(0.630, 0.380^2) \quad \text{d'où par exemple } Pr(|\delta_{cal}| < 1.118) = 0.90$$

Pour $|d_{cal}| \leq 0.630$, nous pouvons donc énoncer que l'effet $|\delta_{cal}|$ est inférieur en valeur absolue à 1.118 avec une garantie *d'au moins* 0.90.

2.2 PLAN EN GROUPES INDÉPENDANTS AVEC DEUX FACTEURS DE CLASSIFICATION

2.2.1 Exemple d'une enquête psychométrique: Milieu et Pédagogie de l'enseignement des mathématiques

L'exemple suivant nous permettra de montrer que les procédures précédentes s'appliquent directement au cas d'un plan en groupes indépendants non équilibré avec deux (ou plus) facteurs de classification.

■ *Utilisation de l'effet calibré*

Nous illustrerons en outre l'utilisation de l'effet calibré pour apprécier l'importance de l'effet δ, en le rapportant à l'écart-type σ. Cohen (1977) a suggéré les critères suivants pour la valeur absolue d'une différence de moyennes

négligeable (*small*)	$\lvert\delta\rvert \leq 0.20\sigma$	soit $\lvert\delta_{cal}\rvert \leq 0.20$
intermédiaire (*medium*)	$0.20\sigma < \lvert\delta\rvert < 0.80\sigma$	soit $0.20 < \lvert\delta_{cal}\rvert < 0.80$
notable (*large*)	$\lvert\delta\rvert \geq 0.80\sigma$	soit $\lvert\delta_{cal}\rvert \geq 0.80$

Ce critère pourra être étendu à toute combinaison linéaire de moyennes.

D'une certaine manière, il pourrait être tentant d'adopter de telles définitions comme une norme, ce qui éviterait de se poser le problème du choix de critères dans une situation particulière. Elles pourront en tout état de cause fournir une base utile pour comparer des résultats d'expériences différentes. Mais bien entendu d'autres critères peuvent être adoptés (voir par exemple Sechrest et Yeaton, 1982), en fonction du type de comparaison considéré, mais aussi en fonction du domaine étudié et des connaissances disponibles.

Notons encore qu'en pratique on pourra éventuellement se contenter de rapporter l'effet vrai à l'écart-type-corrrigé observé s, et adopter par exemple les critères

négligeable (*small*)	$\lvert\delta\rvert \leq 0.20s$
intermédiaire (*medium*)	$0.20s < \lvert\delta\rvert \leq 0.80s$
notable (*large*)	$\lvert\delta\rvert > 0.80s$

Ceci a l'avantage de conduire à des procédures très simples, puisqu'il suffit de rapporter les limites obtenues pour δ à la valeur de s. En outre, si le nombre de degrés de liberté est élevé il sera pratiquement équivalent de rap-

porter l'effet à σ ou à s (voir plus loin).

■ *Exemple: Plan d'analyse*

Considérons des données extraites d'une enquête psychométrique sur l'enseignement des mathématiques effectuée en France il y a plusieurs années (*cf.* Rouanet, Lépine, Pelnard-Considère, 1975; Rouanet, Lecoutre, Bernard, 1987). Ces données concernent 334 élèves répartis en quatre groupes, définis par le croisement de deux facteurs à deux niveaux chacun:

- Facteur *Milieu*: *favorisé* opposé à *défavorisé*
- Facteur *Pédagogie*: *moderne* opposé à *traditionnelle*

Nous avons donc un plan en quatre groupes indépendants, qui peut être écrit $S<G_4>$ ou mieux $S<M_2 \times P_2>$: "*Sujets emboîtés dans les Groupes* $G_4=M_2 \times P_2$".

Les variables retenues ici pour l'analyse sont les notes numériques aux quatre épreuves suivantes passées en fin d'année scolaire: *calcul Combinatoire* (Comb); *calcul des Probabilités* (Prob); *Logique des Propositions* (LogP); *Mathématiques* (Math).

■ *Exemple: variable Combinatoire*

Considérons par exemple la variable *Combinatoire*. Les statistiques relatives à chacun des quatre groupes sont fournies dans le tableau ci-après.

groupe	Milieu	Pédagogie	effectif	moyenne	écart-type
g1	favorisé	moderne	86	5.6395	1.5935
g2	défavorisé	moderne	86	4.8488	1.6840
g3	favorisé	traditionnelle	80	4.3250	1.7559
g4	défavorisé	traditionnelle	82	3.7439	1.5300

■ *Effets calibrés*

Le tableau suivant fournit pour chacun des quatre groupes les moyennes *réduites*, obtenues en divisant les moyennes brutes par l'écart-type-corrigé intragroupe moyen s=1.642, ainsi que les moyennes marginales équipondérées, pour chaque facteur (les autres procédures de moyennage envisageables sont discutées au chapitre ↳5).

		Milieu favorisé	Milieu défavorisé	
Pédagogie	moderne	$3.434_{/86}$	$2.953_{/86}$	3.193
	traditionnelle	$2.634_{/80}$	$2.280_{/82}$	2.457
		3.034	2.616	(s=1.642)

$$\text{Ici } s^2 = \frac{85\times1.5935^2 + 85\times1.6840^2 + 79\times1.7559^2 + 81\times1.5300^2}{86 + 85 + 79 + 81} = 1.642^2$$

Nous en déduisons les effets observés *calibrés* pour cette variable

- Effet du facteur *Milieu*, défini comme différence des deux moyennes relatives à "favorisé" et "défavorisé"

$$d_{cal} = 3.034-2.616 = +0.418$$

- Effet du facteur *Pédagogie*, défini comme différence des deux moyennes relatives à "moderne" et "traditionnelle"

$$d_{cal} = 3.193-2.457 = +0.737$$

- Effet d'interaction, défini comme différence des différences

$$d_{cal} = (3.434-2.953)-(2.634-2.280) = +0.128$$

LeBayésien

Dans la fenêtre principale, cliquez sur le bouton **exemples**, et sélectionnez **Milieu et pédagogie - Combinatoire**. La fenêtre de saisie des statistiques est affichée.

L'effet calibré du facteur *Milieu* par exemple est la combinaison linéaire des moyennes de groupes

$$d_{cal} = \tfrac{1}{2}\,4.434 - \tfrac{1}{2}\,2.953 + \tfrac{1}{2}\,2.634 - \tfrac{1}{2}\,2.280$$

Les coefficients correspondants sont affichés

groupe		coefficient
g1	...	1/2
g2	...	-1/2
g3	...	1/2
g4	...	-1/2

Cliquez sur le bouton **calculer**, ce qui donne les statistiques:
effet d = 0.686
écart-type s = 1.642
effet calibré d_{cal} = 0.418

L'effet du facteur *Pédagogie* est obtenu de la même manière avec les coefficients [1/2 1/2 -1/2 -1/2]; et l'effet d'interaction est obtenu avec les coefficients [1 -1 -1 1].

2.2.2 Analyses descriptives

En procédant de même pour chacune des autres variables, nous obtenons les 3×4=12 effets calibrés observés suivants

	Comb	Prob	LogP	Math
Milieu	+0.418	+0.514	+0.354	+0.324
Pédagogie	+0.737	+0.229	+0.273	+0.127
Interaction	+0.128	-0.163	-0.007	+0.102

S'agissant d'effets calibrés, leurs grandeurs sont comparables d'une variable à l'autre.

- Pour le facteur *Milieu* tous les effets sont positifs, ce qui veut dire que les enfants de milieu favorisé réussissent mieux que les autres (un résultat guère surprenant!).

- Pour le facteur *Pédagogie* tous les effets sont également positifs, ce qui veut dire que la pédagogie moderne est plus efficace que la pédagogie traditionnelle (un résultat encourageant!).

▸ Enfin deux effets d'interaction sont positifs et deux sont négatifs. Dans le premier cas, cela veut dire que la différence entre enfants favorisés et enfants défavorisés augmente lorsque l'on passe de la pédagogie traditionnelle à la pédagogie moderne (un résultat qui va à l'encontre de l'espoir de réduction des inégalités que l'on aurait pu avoir en introduisant la pédagogie moderne).

Nous devons maintenant nous prononcer sur l'importance de ces effets observés. Supposons pour fixer les idées que nous considérions ici comme notable un effet calibré supérieur à 1/3 et comme négligeable un effet calibré inférieur en valeur absolue à 1/3. En choisissant ces critères, nous recherchons ici avant tout une cohérence interne des conclusions, sans nous référer à des critères extérieurs arbitraires tels que ceux mentionnés précédemment.

Pour simplifier, nous prenons les mêmes critères pour les effets principaux et pour l'effet d'interaction; nous pourrions également prendre des limites différentes pour définir les effets notables et négligeables, ce qui introduirait une région "intermédiaire", sans que cela modifiât le principe des procédures (voir plus loin en ↳2.2.4). Selon ces critères trois variables sur quatre ont un effet descriptivement notable pour le facteur *Milieu*, et seulement une pour le facteur *Pédagogie*, alors que tous les effets d'interaction apparaissent négligeables.

	Comb	Prob	LogP	Math
Milieu	notable	notable	notable	négligeable
Pédagogie	notable	négligeable	négligeable	négligeable
Interaction	négligeable	négligeable	négligeable	négligeable

2.2.3 Analyses inférentielles

Pour chacune des quatre variables, considérons pour le plan en groupes indépendants $S<G_4> = S<M_2 \times P_2>$ le *modèle équinormal* usuel, avec l'écart-type intragroupe σ supposé identique pour les quatre groupes.

Compte tenu des effectifs élevés, le problème de la non homogénéité éventuelle des variances est ici assez peu crucial. Nous reviendrons sur ce problème au chapitre ↳6.

Les analyses inférentielles sont effectuées de la même manière que dans l'exemple précédent et permettent notamment de reconstituer le tableau d'analyse de variance traditionnel, par exemple pour la variable *Combinatoire*

source de variation	*dl*	*sc*	*cm*	*F* [dl]	*p*
inter G	3	163.372	54.457		
Milieu M [E]	1	39.245	39.245	14.551 [1,330]	0.0002
Pédagogie P [E]	1	122.080	122.080	45.264 [1,330]	0.0000
Interaction M.P	1	0.916	0.916	0.340 [1,330]	0.5604
intra S(G)	330	890.032	2.697		

La mention [E] indique que nous avons utilisé la solution "équipondérée"

On remarquera ici que la décomposition n'est pas *orthogonale*: la somme-des-carrés *intergroupe* sc_G=163.372 n'est pas la somme des trois sommes-des-carrés partielles (sur cette question voir ↳5.2)

■ *Résultats des tests de signification*

Les résultats des tests de signification sont fournis en utilisant selon la tradition une grille repère de seuils de signification

*** p<0.001 ** p<0.01 * p<0.05 ns *non-significatif*

	Comb	Prob	LogP	Math
Milieu	***	***	**	**
Pédagogie	***	*	*	ns
Interaction	ns	ns	ns	ns

Compte tenu du nombre élevé de sujets, un effet jugé descriptivement négligeable peut parfaitement ici être significatif. L'interprétation des tests usuels appelle donc plus que jamais de sérieuses réserves. En particulier, pour le facteur *Pédagogie* trois variables sur quatre donnent un résultat significatif, alors que seule la *Combinatoire* a un effet observé notable. Néanmoins les effectifs ne sont pas non plus suffisants pour pouvoir généraliser les conclusions descriptives sans recourir à des procédures inférentielles. C'est ce que montre bien l'analyse fiducio-bayésienne.

■ *Inférence fiducio-bayésienne sur les effets calibrés*

Pour chaque comparaison examinée et chaque variable dépendante, la distribution fiducio-bayésienne de l'effet calibré vrai $\delta_{cal}=\delta/\sigma$ est obtenue comme dans l'exemple de Student (↳1.1).

Par exemple, pour le facteur *Milieu* et la variable *Combinatoire*

$$\delta_{cal} \sim \Lambda'_q(d_{cal}, b^2) \sim \Lambda'_{330}(+0.418, 0.109^2)$$

Ici b=0.109 est obtenu à partir des coefficients sur les groupes [+1/2 -1/2 +1/2 -1/2] qui définissent l'effet du facteur *Milieu* et des effectifs

$$b^2 = \frac{(+1/2)^2}{86} + \frac{(-1/2)^2}{86} + \frac{(+1/2)^2}{80} + \frac{(-1/2)^2}{82} = 6763/564160 = 0.109^2$$

Nous pouvons de cette manière obtenir chacune des 3×4=12 distributions voulues et chercher à prolonger, s'il y a lieu, la conclusion descriptive en une conclusion inductive

- soit d'effet notable, en calculant la probabilité que δ_{cal} soit supérieur à +1/3, si d_{cal}>+1/3,
- soit d'effet négligeable, en calculant la probabilité que $|\delta_{cal}|$ soit inférieur à 1/3, si $|d_{cal}|$<1/3.

LeBayésien

Dans la fenêtre de saisie des statistiques, revenez aux données de la *Combinatoire* et à l'effet du facteur *Milieu* (**combinaison linéaire** [1/2 -1/2 +1/2 -1/2]). Cliquez sur **ok**, puis dans la fenêtre principale cliquez sur la courbe de droite relative à l'effet calibré.

Activez l'option **limite** et entrez la limite 1/3, puis cliquez sur **calculer**. Sur la courbe **notable**, vous obtenez la garantie 0.776.

Les résultats ainsi obtenus (avec trois décimales) sont les suivants

	effet notable				**effet négligeable**			
	Comb	Prob	LogP	Math	Comb	Prob	LogP	Math
Milieu	0.776	0.948	0.573	/////	/////	/////	/////	0.534
Pédagogie	1.000	/////	/////	/////	/////	0.830	0.710	0.970
Interaction	/////	/////	/////	/////	0.809	0.770	0.872	0.831
	$Pr(\delta_{cal}>+1/3)$				$Pr(\lvert\delta_{cal}\rvert<1/3)$			

Nous remarquerons que dans cet exemple la distribution relative à δ/σ ne diffère pour ainsi dire pas de celle relative à δ/s (s étant donné). Celle-ci est de la forme générale

$$\frac{\delta}{s} \sim t_q(\frac{d}{s}, b^2) \text{ qui est approximativement } N(\frac{d}{s}, b^2) \text{ pour q élevé}$$

Par exemple, pour le facteur *Milieu* et la variable *Combinatoire*, la distribution de δ_{cal} a une moyenne +0.417, un écart-type 0.111 et un troisième moment centré de l'ordre de 10^{-7}. Elle est donc très bien approximée par la distribution symétrique

$$\frac{\delta}{\sigma} \simeq N(+0.417, 0.111^2)$$

qui est très voisine de la distribution

$$\frac{\delta}{s} \sim t_{330}(+0.418, 0.109^2)$$

LeBayésien

Par exemple, toujours pour la variable *Combinatoire* et l'effet du facteur *Milieu*, dans la fenêtre principale cliquez sur la courbe de gauche relative à l'effet. Dans la fenêtre des énoncés activez l'option **exprimer les limites en fonction de**, puis cliquez sur **s=1.642**. Pour la limite 1/3, vous obtenez maintenant sur la courbe **notable** la garantie 0.779.

Nous avons ici

$$Pr(\frac{\delta}{s} > \frac{1}{3}) = Pr(\delta > \frac{1}{3}s) = 0.779$$

En pratique, cela veut dire que si le nombre de degrés de liberté associé à s est élevé, on peut utiliser la distribution $\delta/s|d,s$ comme approximation de la distribution $\delta/\sigma|d,s$. Ainsi ici les probabilités obtenues de cette manière diffèrent des précédentes d'au plus 0.004. Mais, dans le cas général, les deux distributions peuvent largement différer, et la distribution de δ/s ne peut plus alors être utilisée pour approximer celle de δ_{cal}.

■ *Différents types de conclusions*

Illustrons maintenant sur des exemples les différents types de conclusions auxquels nous pouvons aboutir. Les distributions considérées sont figurées dans la figure 2.1.

▸ Conclusion d'effet notable pour la variable *Probabilités* et le facteur *Milieu* – Dans ce cas la majeure partie de la distribution fiducio-bayésienne (approximativement centrée sur +0.51) se trouve dans la région des valeurs notables; nous pouvons donc énoncer une conclusion d'effet notable basée sur l'énoncé

$$Pr(\delta_{cal} > +1/3) = 0.948$$

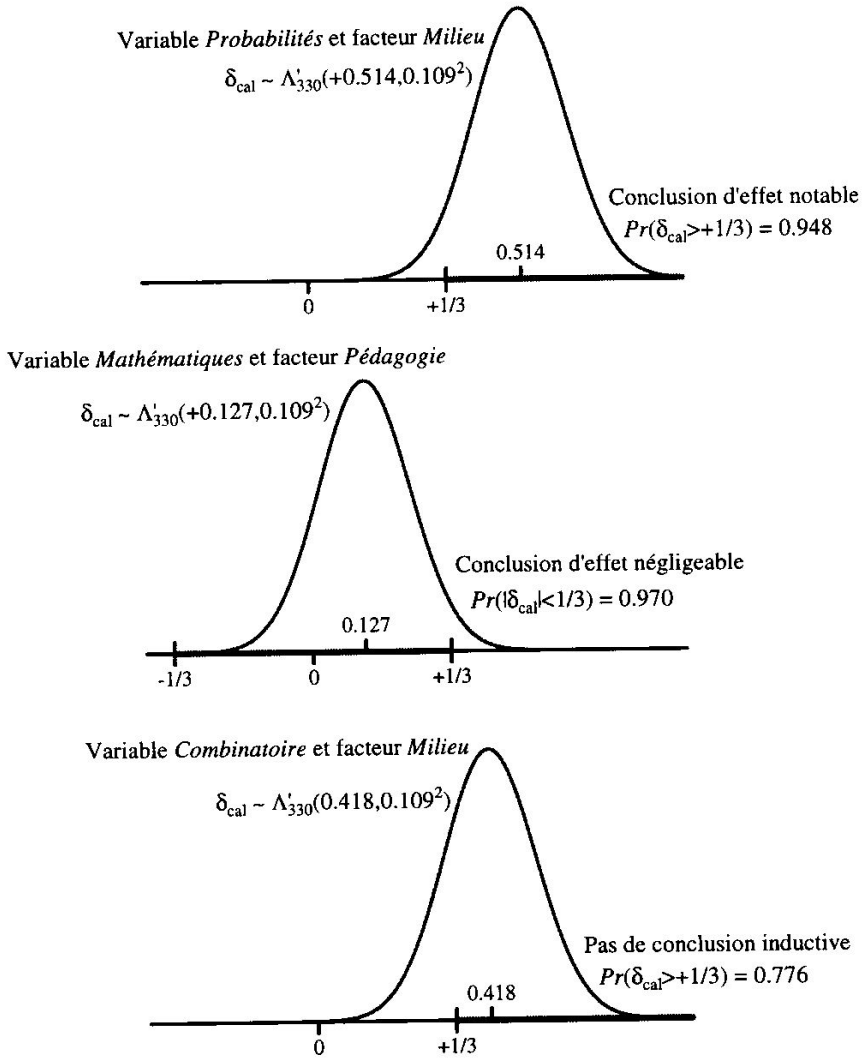


Figure 2.1 - Différents types de conclusions

▸ Conclusion d'effet négligeable pour la variable *Mathématiques* et le facteur *Pédagogie* – Dans ce cas la majeure partie de la distribution fiducio-bayésienne (approximativement centrée sur +0.13) se trouve dans la région des valeurs négligeables; nous pouvons donc énoncer une conclusion d'effet négligeable basée sur l'énoncé

$$Pr(|\delta_{cal}|<1/3) = 0.970$$

▸ Pas de conclusion inductive pour la variable *Combinatoire* et le facteur *Milieu* – Dans ce cas la distribution fiducio-bayésienne est approximativement centrée sur la valeur notable +0.42, mais la probabilité 0.776 associée à la région des valeurs notables (δ_{cal}>+1/3) sera certainement jugée insuffisante pour généraliser la conclusion descriptive.

En résumé, si nous souhaitons une garantie d'au moins 0.90, seuls trois cas conduisent à une conclusion inductive

	effet	variable
effet notable	Milieu	Probabilités
	Pédagogie	Combinatoire
effet négligeable	Pédagogie	Mathématiques

2.2.4 Définition de plusieurs régions d'importance de l'effet

Supposons par exemple maintenant que nous considérions ici comme notable un effet calibré supérieur à 2/3 et comme négligeable un effet calibré inférieur en valeur absolue à 1/3, un effet compris entre ces deux limites étant *intermédiaire* ("ni notable ni négligeable").

Descriptivement, il y a deux effets intermédaires, pour le facteur *Milieu* et les variables *Combinatoire* et *Probabilités*. Pour chercher à prolonger la conclusion descriptive en une conclusion inductive nous calculons dans ce cas la probabilité que δ_{cal} soit compris entre +1/3 et +2/3, soit respectivement

$$Pr(+1/3<\delta_{cal}<+2/3) = 0.764 \ (Combinatoire)$$

$$Pr(+1/3<\delta_{cal}<+2/3) = 0.863 \ (Probabilités)$$

LeBayésien

Considérons encore la variable *Combinatoire* et l'effet du facteur *Milieu*, pour l'effet calibré. Dans la fenêtre des énoncés activez l'option **deux limites**, puis entrez les limites 1/3 et 2/3. Cliquez sur le bouton

calculer; vous obtenez maintenant:
- sur la courbe du haut (**intervalle**) la garantie 0.764 pour l'intervalle [+1/3,+2/3];
- sur la courbe du milieu (**négligeable**) la garantie 0.224 pour la limite 1/3;
- sur la courbe du bas (**notable**) la garantie 0.012 pour la limite +2/3.

Que nous considérions l'effet ou l'effet calibré, il serait en fait immédiat de prendre en compte plusieurs régions d'importance des effets ("plus ou moins négligeable", "plus ou moins notable"...). Mais il sera d'autant plus difficile d'obtenir une conclusion inductive que l'étendue de ces régions sera faible.

CHAPITRE 3

PRISE EN COMPTE D'INFORMATIONS EXTÉRIEURES AUX DONNÉES

Dans ce chapitre, nous illustrerons l'utilisation de distributions initiales de la même famille que la distribution fiducio-bayésienne. Dans ce cas la distribution finale est encore de la même famille, qui est dite "conjuguée naturelle" de la distribution d'échantillonnage. Par conséquent, une fois déterminée la distribution finale, les procédures, notamment en ce qui concerne la recherche d'une conclusion sur l'importance de l'effet, sont techniquement les mêmes que celles que nous avons examinées jusqu'à présent, ce qui conduit à des procédures *standard* et présente donc un intérêt pratique évident.

Le problème général du passage de l'information *a priori* à la distribution initiale est évidemment un problème d'une importance considérable; et il existe une littérature abondante sur le sujet (voir Robert, 1992, chapitre 3). Nous nous contenterons ici de discuter brièvement quelques principes généraux, qui nous paraissent devoir guider le choix et l'utilisation des distributions initiales pour l'analyse des données expérimentales.

Nous examinerons ensuite, à partir d'exemple, l'influence du choix de la distribution initiale sur la distribution finale, en distinguant d'abord l'inférence relative à l'effet δ (pour σ fixé), et celle relative à σ seul.

Nous examinerons une utilisation méthodologique de la distribution initiale, pour examiner la *sensibilité* des conclusions issues d'une expérience vis-à-vis d'informations extérieures.

■ *Remarque*

Nous nous limitons ici à considérer une distribution initiale relative au couple (δ,σ). Dans la situation générale considérée, δ est en fait une combinaison linéaire de moyennes μ_i. On pourrait donc souhaiter choisir une distribution pour l'ensemble de ces moyennes, et éventuellement aussi pour les autres paramètres qui leur sont associés (variances et s'il y a lieu covariances). Mais ceci nécessite de recourir à des procédures multidimensionnelles, ce qui sort du cadre de cet ouvrage. La question qui se pose est en fait de savoir dans quelle mesure l'inférence marginale relative au couple (δ,σ)

pourrait être modifiée: voir la discussion sur l'*analyse spécifique* au chapitre ↳4.

3.1 Choix de la distribution initiale: Différentes attitudes

Comme nous l'avons dit dans l'introduction, on peut adopter vis-à-vis du choix de la distribution initiale différentes attitudes. Nous en considérerons ici trois, qui ne sont ni exhaustives, ni nécessairement exclusives.

■ *Trois attitudes*

• Une première attitude consiste à choisir une distribution initiale, qui soit la distribution fiducio-bayésienne obtenue à partir de données recueillies dans une autre expérience réalisée (ou éventuellement dans plusieurs expériences). Dans ce cas la procédure peut présenter un caractère parfaitement objectif, dans la mesure où les conditions expérimentales sont identiques.

• Une deuxième attitude consiste à choisir une distribution initiale qui soit *équivalente* à la distribution fiducio-bayésienne obtenue à partir de données *fictives* (donc de la même famille que cette distribution). Nous pouvons également avec un tel choix adopter une démarche objective, en cherchant à répondre à des questions du type:
• "quelle information extérieure aux données conduirait à une conclusion différente de celle obtenue à partir de la seule information expérimentale?" (quand une conclusion a été obtenue à partir des données);
• "quelle information extérieure aux données serait nécessaire pour pouvoir atteindre telle conclusion?" (quand cette conclusion n'a pu être obtenue à partir des seules données).

• Une troisième attitude consiste à choisir une distribution initiale qui reflète des connaissances antérieures éventuellement de nature diverses. Cette distribution exprime une information *effective* extérieure aux données, mais qui présente un caractère plus ou moins subjectif. Il pourra notamment s'agir de résultats de *méta-analyses* (voir par exemple DuMouchel, 1989) portant sur plusieurs expériences effectuées dans des conditions différentes, et/ou de l'opinion d'experts du domaine étudié (*cf.*, pour les essais cliniques, Spiegelhalter, Freedman et Parmar, 1994). La distribution finale correspondante sera utile pour prendre une décision; mais elle pourra à l'évidence être plus

ou moins difficile à faire admettre. En fait elle apparaîtra généralement largement *arbitraire* et *artificielle* à un non spécialiste du domaine concerné. C'est pourquoi nous ne donnerons pas ici d'exemples d'une telle utilisation des procédures bayésiennes. Son statut dans la méthodologie de l'analyse des données expérimentales reste à notre avis à clarifier.

D'un point de vue technique, si nous adoptons cette attitude, il apparaît bien sûr très contraignant de nous limiter à une famille particulière de distributions (voir par exemple Press, 1989, chapitre 4). Il conviendrait donc d'envisager pour la distribution initiale d'autres choix que ceux auxquels nous nous limitons ici.

■ *Examiner la sensibilité des conclusions à des informations supplémentaires*

Les procédures fiducio-bayésiennes, en exprimant *l'apport propre des données* fourniront une référence. Nous pouvons alors les compléter par des analyses correspondant à différentes distributions initiales, et à différentes attitudes. Dans la méthodologie expérimentale, un rôle essentiel des distributions initiales sera d'examiner la sensibilité des conclusions obtenues à partir de la distribution fiducio-bayésienne vis-à-vis d'informations supplémentaires. Si ces informations peuvent être exprimées sous forme d'une distribution de la même famille, on appliquera les procédures exposées dans ce chapitre.

■ *Distributions favorables (enthousiastes) et défavorables (sceptiques)*

Il peut alors sembler intéressant de considérer des choix "extrêmes" vis-à-vis de la conclusion recherchée, c'est-à-dire des distributions *a priori favorables* ou au contraire *défavorables*, soit pour reprendre les expressions de Spiegelhalter, Freedman et Parmar (1994) des distributions respectivement "*enthousiastes*" et "*sceptiques*". Mais ici encore le statut méthodologique de ces distributions devra être clarifié. En particulier, nous voyons dans leur utilisation le danger de vouloir réitérer le raisonnement hypothético-déductif du test de signification, en remplaçant l'hypothèse nulle ponctuelle sur le paramètre par une distribution initiale ne faisant qu'exprimer cette hypothèse de manière un peu moins stricte, mais sans réellement traduire une connaissance sur ce paramètre.

C'est pourquoi nous préférons regarder ces distributions comme traduisant une information (donc une connaissance) *potentielle*, notamment celle qui pourrait être fournie par une expérience *future* (*cf.* Lecoutre, 1984, page 266).

3.2 INFLUENCE DE LA DISTRIBUTION INITIALE

3.2.1 Inférence relative à δ conditionnellement à σ

■ *L'exemple de Student*

Reprenons l'exemple de Student" du chapitre ↳1, pour lequel nous avons d=+1.580 et b=1/√10 (soit $1/b^2=n=10$). Supposons que l'écart-type vrai σ soit une donnée (soit "connu"), et prenons-le égal à la valeur observée s=1.230.

La distribution fiducio-bayésienne relative à δ est dans ce cas une distribution *Normale*, de moyenne d et d'échelle (ici écart-type) $e=b\sigma$

$$\delta \mid d,\sigma \sim N(d,b^2\sigma^2) \sim N(+1.58,\sigma^2/10)$$

Celle relative à δ_{cal} est également *Normale* de moyenne d/σ et d'échelle b.

■ *La distribution initiale*

Etant donné σ, considérons pour δ une distribution initiale *Normale*, centrée sur d_0 et d'échelle $e_0=b_0\sigma$

$$\delta \mid \sigma \sim N(d_0,e_0^2) \qquad \text{avec } e_0 = b_0\sigma$$

Ici cette distribution peut être regardée comme la distribution fiducio-bayésienne issue de données, avec

$$n_0=1/b_0^2 \text{ sujets}$$

pour lesquelles on observerait un effet d_0. En pratique, il pourra s'agir de données *réelles* (déjà recueillies), mais aussi de données *fictives*, par exemple de données "potentielles" que pourrait apporter une expérience *future*, ou de données qui apporteraient une "information équivalente" à une information *a priori* éventuellement de caractère subjectif. Il n'y a donc évidemment pas de raison de nous restreindre à des valeurs de n_0 entières ou supérieures à 1.

■ *La distribution finale*

La distribution finale est également une distribution *Normale* (↳10.1.2)

$$\delta \mid d,\sigma \sim N(d_1,e_1^2) \qquad \text{avec } e_1 = b_1\sigma$$

• centrée sur une valeur d_1, intermédiaire entre d_0 et d, qui est simplement la moyenne, pondérée par n_0 et n, de ces deux quantités

$$d_1 = \frac{n_0 d_0 + nd}{n_0 + n} \qquad \left(\text{soit } d_1 = \frac{b_0^2 d + b^2 d_0}{b_0^2 + b^2}\right)$$

• et d'échelle $e_1 = b_1\sigma$ inférieure à la fois à e_0 et à e, la valeur b_1 étant telle que

$$1/b_1^2 = n_1 = n_0 + n \qquad \left(\text{soit } b_1^2 = \frac{b_0^2 b^2}{b_0^2 + b^2}\right)$$

Cette distribution finale est donc la distribution fiducio-bayésienne que l'on obtiendrait pour l'échantillon de $n_1 = n_0 + n$ données, constitué des données de l'échantillon et des données (réelles ou fictives) correspondant à la distribution initiale.

Les valeurs limites $d_1 = d$ et $b_1 = b$ sont obtenues quand $b_0 \to \infty$ (soit $n_0 \to 0$); la distribution finale coïncide alors avec la distribution fiducio-bayésienne.

■ *Exemples de distributions finales*

LeBayésien

Sélectionnez l'exemple **Student**. Dans la fenêtre des statistiques, cliquez sur le bouton **ok**, ce qui vous place dans la fenêtre principale.

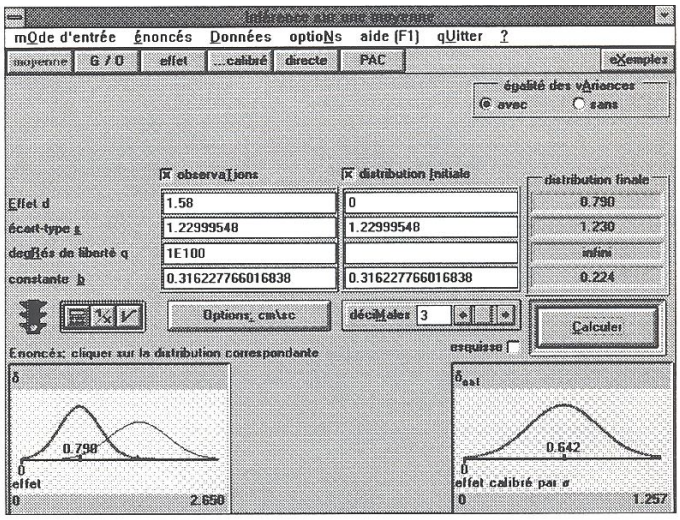
Cliquez sur le **feu rouge**, qui "passe au vert", pour pouvoir modifier les observations. Entrez pour le nombre de degrés de liberté q une valeur très élevée, par exemple 1E100 (10^{100}), ce qui revient à considérer σ comme une donnée (égale à 1.230). Cliquez sur le bouton **calculer** pour obtenir les distributions fiducio-bayésiennes relatives à δ et δ_{cal} (qui sont des distributions *normales*). Il n'y a bien entendu pas de distribution finale relative à σ.

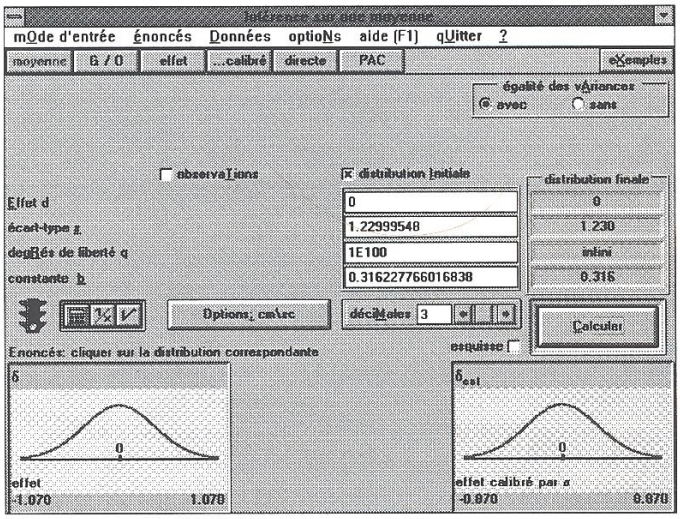
Activez maintenant la case à cocher **distribution initiale**. Considérons pour δ la distribution initiale caractérisée par les valeurs $d_0 = 0$ et $b_0 = 1/\sqrt{10}$, soit $\delta \sim N(0, \sigma/\sqrt{10})$ avec ici $\sigma = 1.230$. Entrez les valeurs correspondantes

effet = 0 et **constante b** = 0.316227766016838 ($1/\sqrt{10}$)

(effacez s'il y a lieu l'**écart-type** et les **degrés de liberté**)

puis cliquez sur le bouton **calculer**. Vous obtenez la distribution finale correspondante. Pour δ la courbe de la distribution fiducio-bayésienne est également tracée sur la même figure, en trait plus fin, ce qui permet d'apprécier l'apport de la distribution initiale.


mQde d'entrée énoncés Données optioNs aide (F1) qUitter ?
moyenne G / O effet ...calibré directe PAC eXemples
égalité des vAriances
avec sans
observaTions
distribution Initiale
distribution finale
Effet d
écart-type s
degRés de liberté q
constante b
1.58
1.22999548
1E100
0.316227766016838
0
1.22999548
0.316227766016838
0.790
1.230
infini
0.224
Options: cm\sec
déciMales 3
Calculer
esquisse
Enoncés: cliquer sur la distribution correspondante
δ
0.790
0
effet
0
2.650
δest
0.642
0
effet calibré par σ
0
1.257


mQde d'entrée énoncés Données optioNs aide (F1) qUitter ?
moyenne G / O effet ...calibré directe PAC eXemples
égalité des vAriances
avec sans
observaTions
distribution Initiale
distribution finale
Effet d
écart-type s
degRés de liberté q
constante b
0
1.22999548
1E100
0.316227766016838
0
1.230
infini
0.316
Options: cm\sec
déciMales 3
Calculer
esquisse
Enoncés: cliquer sur la distribution correspondante
δ
0
effet
-1.070
1.070
δest
0
effet calibré par σ
-0.970
0.970

Remarquez que la valeur de l'écart-type est dans ce cas automatiquement reportée dans la distribution initiale.

Vous pouvez également étudier la distribution initiale seule. Pour cela cliquez sur la case à cocher **observations** (ce qui fait disparaître les données), puis sur le bouton **calculer**.

Vous pouvez obtenir, de la manière habituelle, les énoncés de probabilité associés respectivement à la distribution finale et à la distribution initiale. Par exemple pour la **garantie** 0.95 vous trouvez les intervalles de crédibilité:

$+0.818<\delta<+2.342$ pour la distribution fiducio-bayésienne;
$-0.762<\delta<+0.762$ pour la distribution initiale;
$+0.251<\delta<+1.329$ pour la distribution finale.

Vous pouvez étudier les modifications de la distribution finale obtenues en faisant varier les valeurs de d_0 et/ou b_0. Observez notamment que, quand b_0 augmente, l'apport de la distribution initiale diminue et la distribution finale tend vers la distribution fiducio-bayésienne. Par exemple, avec trois décimales, les deux distributions coïncident pour $d_0=0$ et $b_0 \geq 17.8$.

Nous obtenons par exemple, en fonction de b_0, pour $d_0=0$

b_0	n_0	d_1	b_1	$\gamma=0.95$
0.01	10000	+0.002	0.0100	$-0.023<\delta<+0.026$
0.05	400	+0.039	0.0494	$-0.081<\delta<+0.158$
0.1	100	+0.144	0.0953	$-0.086<\delta<+0.373$
0.2	50	+0.451	0.1690	$+0.044<\delta<+0.859$
0.3162	10	+0.790	0.2236	$+0.251<\delta<+1.329$
0.5	4	+1.129	0.2673	$+0.484<\delta<+1.773$
1	1	+1.436	0.3015	$+0.710<\delta<+2.163$
4	0.0625	+1.570	0.3152	$+0.810<\delta<+2.330$
10	0.01	+1.578	0.3161	$+0.816<\delta<+2.340$
20	0.0025	+1.580	0.3162	$+0.817<\delta<+2.342$
$b_0 \to \infty$	$n_0 \to 0$	+1.580	0.3162	$+0.818<\delta<+2.342$

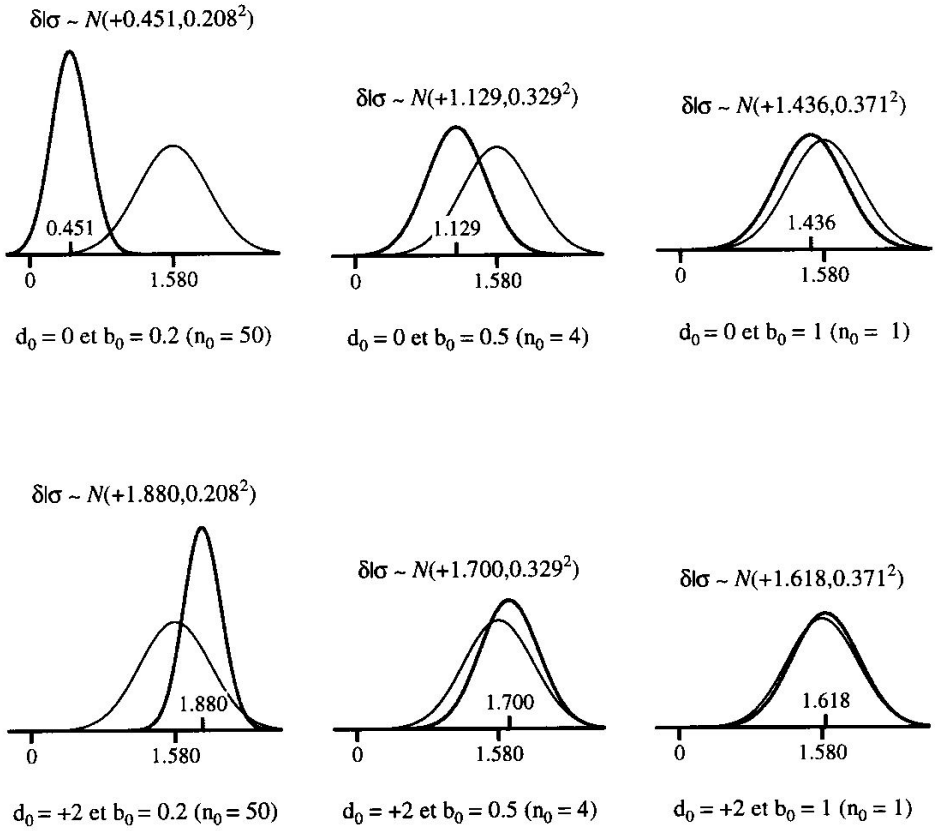


Figure 3.1 - Influence de d_0 et b_0 sur la distribution finale relative à δ conditionnellement à σ=1.230
En trait plus fin, la distribution fiducio-bayésienne: $\delta|\sigma \sim N(+1.580,0.389^2)$

et pour d_0=+2

b_0	n_0	d_1	b_1	γ=0.95
0.01	10000	+2.000	0.0100	+1.975<δ<+2.024
0.05	400	+1.990	0.0494	+1.871<δ<+2.109
0.1	100	+1.962	0.0953	+1.732<δ<+2.192
0.2	25	+1.880	0.1690	+1.473<δ<+2.287
0.3162	10	+1.790	0.2236	+1.251<δ<+2.329
0.5	4	+1.700	0.2673	+1.056<δ<+2.344
1	1	+1.618	0.3015	+0.891<δ<+2.345
4	0.0625	+1.570	0.3152	+0.810<δ<+2.330
10	0.01	+1.580	0.3161	+0.818<δ<+2.342
20	0.0025	+1.580	0.3162	+0.818<δ<+2.342
$b_0 \to \infty$	$n_0 \to 0$	+1.580	0.3162	+0.818<δ<+2.342

Des exemples de ces distributions finales sont présentés dans la figure 3.1. On peut ainsi voir

• d'une part les apports respectifs de distributions *a priori* "sceptiques" (centrées sur 0), et "enthousiastes" (centrées sur deux heures), quant à la supériorité du somnifère B sur le somnifère A;

• d'autre part l'influence de la distribution initiale par rapport aux données, lorsque l'on fait varier l'incertitude *a priori* (exprimée ici par b_0).

On n'oubliera pas ici que les distributions sont conditionnelles à une valeur donnée σ. Cependant les valeurs d_1 et b_1 obtenues ne dépendent pas de σ.

3.2.2 Inférence (marginale) relative à σ

■ *L'exemple de Student*

Rappelons que la distribution fiducio-bayésienne relative à σ est une distribution *phi-inverse* d'échelle s=1.230, à q=10-1 degrés de liberté (↳1.1.4)

$$\sigma \sim s\varphi_q^{-1} \sim 1.230\varphi_9^{-1}$$

■ *La distribution initiale*

Considérons pour le couple (δ,σ) la distribution initiale obtenue en choisissant

• pour δ conditionnellement à σ la distribution précédente

$$\delta \mid \sigma \sim N(d_0, b_0^2\sigma^2) \qquad \text{(et posons encore } n_0 = 1/b_0^2\text{)}$$

• pour σ la distribution marginale *phi-inverse*, centrée sur s_0 et avec q_0 degrés de liberté

$$\sigma \sim s_0\varphi^{-1}_{q_0}$$

Ici la distribution relative à σ peut être regardée comme la distribution fiducio-bayésienne issue de données, avec

$$m_0 = q_0+1 \text{ sujets}$$

pour lesquelles on observerait un écart-type-corrigé s_0.

Il n'y a évidemment pas de raison de nous restreindre au cas particulier où m_0 est égal à n_0, les informations relatives à δ et à σ pouvant être de sources différentes. Il n'y a pas non plus de raison de nous restreindre à des valeurs de m_0 entières ou supérieures à 1.

■ *La distribution finale*

La distribution finale marginale relative à σ est également une distribution *phi-inverse*

$$\sigma \mid d,s \sim s_1\varphi^{-1}_{q_1}$$

• d'échelle s_1, telle que s_1^2 est est une moyenne pondérée de s_0^2, de s^2 et du carré de la différence entre d_0 et d (elle dépend donc du choix de la distribution initiale relative à δ et de la valeur observée d)

$$s_1^2 = \frac{(m_0-1)s_0^2 + (n-1)s^2 + \frac{n_0 n}{n_0+n}(d_0-d)^2}{(m_0-1) + (n-1) + 1} \qquad \text{soit } s_1^2 = \frac{q_0 s_0^2 + qs^2 + \frac{(d_0-d)^2}{b_0^2+b^2}}{q_0 + q + 1}$$

• et avec un nombre de degrés de liberté

$$q_1 = m_0-1+n-1+1 = m_0+n-1 \qquad \text{soit } q_1 = q_0+q+1$$

Dans le cas particulier $m_0=n_0$, la distribution initiale relative au couple (δ,σ) peut être regardée comme la distribution fiducio-bayésienne issue d'un seul échantillon de n_0 sujets. La distribution finale relative à σ est alors la distribution fiducio-bayésienne que l'on obtiendrait pour l'échantillon de $n_1=n_0+n$ données, constitué des données de l'échantillon et des données (réelles ou fictives) correspondant à la distribution initiale.

Les valeurs limites $s_1=s$ et $q_1=q$ sont obtenues quand $b_0 \to \infty$ (soit $n_0 \to 0$) et $q_0 \to -1$ (soit $m_0 \to 0$); la distribution finale coïncide alors avec la distribution fiducio-bayésienne.

■ *La distribution finale conditionnelle à* s

Pour σ, la situation est donc relativement plus complexe, puisque la distribution finale dépend, non seulement de s, q, s_0 et q_0, mais aussi de d, b, d_0 et b_0. Pour mieux apprécier le rôle de ces différentes quantités, il est intéressant de considérer d'abord la distribution finale conditionnelle à s seul, c'est-à-dire ignorant l'information apportée par d et d_0. Il s'agit toujours d'une distribution *phi-inverse*

$$\sigma \,|\, s \sim u_1 \varphi^{-1}_{r_1}$$

où, plus simplement, u_1^2 est la moyenne s_0^2 et de s^2 pondérée par les degrés de liberté

$$u_1^2 = \frac{q_0 s_0^2 + q s^2}{q_0 + q}$$

et r_1 est la somme des nombres de degrés de liberté

$$r_1 = q_0 + q$$

On remarquera qu'il s'agit encore de la distribution finale conditionnelle à d et s (donc à tout l'échantillon) que l'on obtiendrait en choisissant pour δ une distribution initiale indépendante de celle de σ (voir ↳10.1.4).

■ *Exemples de distributions finales pour* σ *conditionnellement à* s

LeBayésien

Sélectionnez à nouveau l'exemple **Student**. Dans la fenêtre des statistiques, cliquez sur le bouton **ok**, ce qui vous place dans la fenêtre principale.

Activez la case à cocher **distribution initiale**. Pour ignorer l'information apportée par d et d_0, il suffit d'effacer la zone de texte de la distribution initiale correspondant à la **constante b** (remarquez que c'est le seul cas où une zone de texte vide n'est pas traitée comme la valeur zéro).

Considérons pour σ la distribution initiale caractérisée par les valeurs s_0=1 et q_0=9. Entrez les valeurs correspondantes

écart-type = 1 et **degrés de liberté** = 9

(effacez s'il y a lieu la **constante b**)

puis cliquez sur le bouton **calculer**. Vous obtenez la distribution finale correspondante. La courbe de la distribution fiducio-bayésienne est également tracée sur la même figure, en trait plus fin, ce qui permet d'apprécier l'apport de la distribution initiale.

Vous pouvez également étudier la distribution initiale seule. Pour cela cliquez sur la case à cocher **observations** (ce qui fait disparaître les données), et cliquez sur **calculer**.

Vous pouvez obtenir, de la manière habituelle, les énoncés de probabilité associés respectivement à la distribution finale et à la distribution initiale. Par exemple pour la **garantie** 0.95 vous trouvez les intervalles de crédibilité:

0.846<σ<2.246 pour la distribution fiducio-bayésienne;
0.688<σ<1.826 pour la distribution initiale;
0.847<σ<1.658 pour la distribution finale.

Vous pouvez étudier les modifications de la distribution finale obtenues en faisant varier les valeurs de s_0 et/ou q_0. Observez notamment que, quand q_0 diminue, l'apport de la distribution initiale diminue et la distribution finale tend vers la distribution fiducio-bayésienne, obtenue pour q_0=0.

Nous obtenons par exemple, en fonction de q_0, pour s_0=1

q_0	u_1	r_1	γ=0.95
5000	1.000	5009	0.981<σ<1.020
1000	1.002	1009	0.960<σ<1.048
500	1.005	509	0.946<σ<1.070
100	1.021	109	0.902<σ<1.177
50	1.038	59	0.880<σ<1.266

q_0	u_1	r_1	γ=0.95
20	1.077	29	0.857<σ<1.447
10	1.115	19	0.848<σ<1.628
5	1.153	14	0.844<σ<1.819
1	1.209	10	0.845<σ<2.122
0	1.230	9	0.846<σ<2.246

Des exemples de ces distributions finales sont présentés dans la figure 3.2.

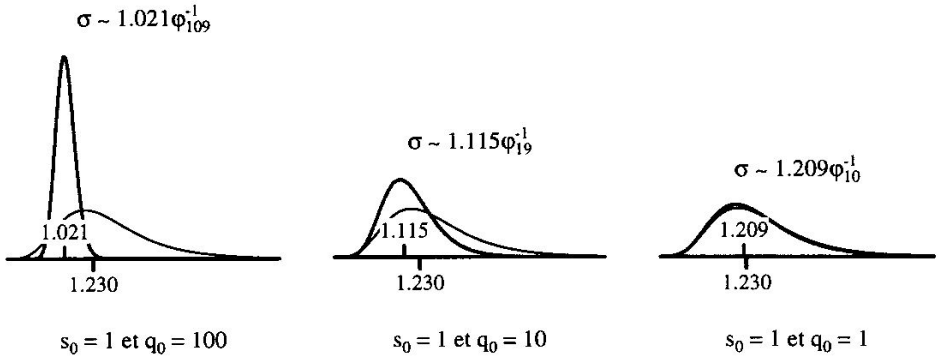


Figure 3.2 - Influence de q_0 sur la distribution finale relative à σ conditionnelle à s ($s_0 = 1$)
En trait plus fin, la distribution fiducio-bayésienne: $\sigma \sim 1.230\varphi_9^{-1}$

■ *Exemples de distributions finales pour σ conditionnellement à* d et s

LeBayésien

Dans la situation précédente, entrez maintenant pour la distribution initiale les valeurs

effet = 0 **écart-type** = 1

degrés de liberté = 9 **constante b** = 0.3162 (1/√10)

puis cliquez sur le bouton **calculer**. Vous obtenez pour la distribution finale une valeur s_1=1.359, qui est supérieure à la fois à s_0 et à s, en raison de la différence importante entre d_0 et d.

Nous reviendrons plus loin sur les distributions finales relatives à δ et δ_{cal}.

Vous pouvez étudier les modifications de la distribution finale de σ obtenues en faisant varier les valeurs de d_0, s_0 q_0, b_0. Observez notamment que, comme on peut s'y attendre:

- quand q_0 augmente (l'information initiale relative à σ est plus précise), l'apport de la différence d_0-d diminue;
- quand b_0 augmente (l'information initiale relative à δ est moins précise), l'apport de la différence d_0-d diminue;

- et bien entendu quand la différence d_0-d diminue, son apport diminue.

Par exemple, pour d_0=0 et s_0=1, nous obtenons les valeurs s_1 en fonction de q_0 et b_0 (avec ici $q_1=q_0+10$)

q_0	b_0=0.01	b_0=0.10	b_0=0.3162	b_0=1	$b_0 \to \infty$
5000	1.003	1.003	1.002	1.001	1.003
1000	1.014	1.013	1.008	1.003	1.002
500	1.028	1.025	1.016	1.006	1.004
100	1.122	1.113	1.071	1.026	1.016
50	1.215	1.199	1.126	1.030	1.030
20	1.397	1.370	1.240	1.094	1.059
10	1.558	1.522	1.343	1.138	1.115
5	1.704	1.660	1.440	1.180	1.153
1	1.896	1.842	1.570	1.239	1.209
0	1.964	1.906	1.615	1.260	1.230

On voit que l'augmentation de s_1 due à la différence d_0-d (élevée ici) peut être très importante.

Bien entendu, pour une différence d_0-d relativement faible, la valeur s_1 varie beaucoup moins, même pour des petites valeurs de q_0, comme on peut le vérifier par exemple pour d_0=+1.5 (et toujours s_0=1)

q_0	b_0=0.01	b_0=0.10	b_0=0.3162	b_0=1	$b_0 \to \infty$
10	1.088	1.088	1.087	1.087	1.087
5	1.116	1.116	1.115	1.114	1.114
1	1.155	1.155	1.154	1.153	1.153
0	1.170	1.169	1.168	1.167	1.167

Des exemples de ces distributions finales sont présentés dans la figure 3.3.

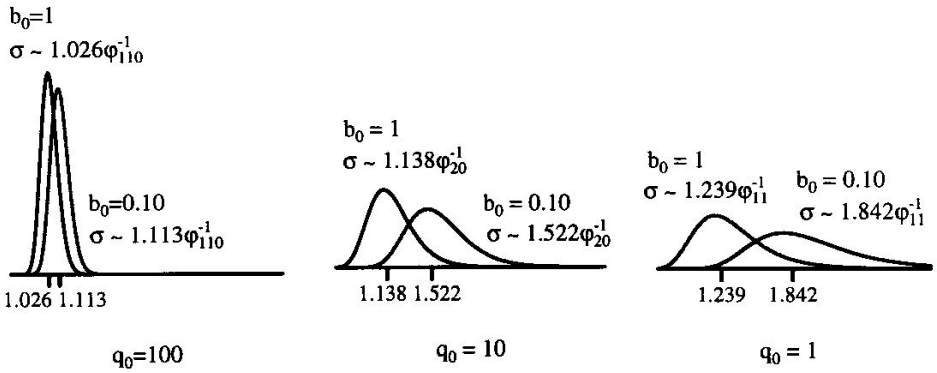


Figure 3.3 - Influence de q_0 et b_0 sur la distribution finale relative à σ conditionnelle à d et s (d_0=0 et s_0=1)

3.2.3 Inférences (marginales) relatives à δ et δ_{cal}

■ *L'exemple de Student*

Rappelons que la distribution fiducio-bayésienne relative à δ est une distribution *t de Student généralisée*, de moyenne d et d'échelle e=b*s*

$$\delta \,|\, d,s \;\sim\; t_q(d,e^2) \;\sim\; t_9(+1.580,0.389^2)$$

Celle relative à $\delta_{cal}=\delta/\sigma$ est une distribution *lambda-prime*, d'indice d'excentricité d/s et d'échelle b

$$\delta_{cal} \,|\, d,s \;\sim\; \Lambda'_q(\tfrac{}{s},b^2) \;\sim\; \Lambda'_9(+1.285,0.316^2)$$

■ *La distribution finale*

Pour la distribution initiale relative au couple (δ,σ) définie précédemment (↳3.1.2), les distributions finales *marginales* de δ et δ_{cal} sont le mélange des distributions conditionnelles à σ, pondéré par la distribution marginale de σ.

Il s'agit encore pour δ d'une distribution *t généralisée*

$$\delta \,|\, d,s \sim t_{q_1}(d_1, e_1^2) \qquad \text{avec } e_1 = b_1 s_1$$

et pour δ_{cal} d'une distribution *lambda-prime*

$$\delta_{cal} \,|\, d,s \sim \Lambda'_{q_1}\left(\frac{d_1}{s_1}, b_1^2\right)$$

où les quantités d_1, s_1, q_1 et b_1 ont été définies précédemment

qui coïncident avec la distribution fiducio-bayésienne quand $b_0 \to +\infty$ et $q_0 \to -1$.

■ *Exemples de distributions finales*

Considérons à titre d'illustration pour d_0 les valeurs

$$d_0 = 0 \qquad d_0 = +1.50 \qquad d_0 = +2$$

et pour s_0 les valeurs

$$s_0 = 1 \qquad s_0 = 2$$

En regard de la conclusion recherchée (on veut montrer la supériorité du somnifère A sur le somnifère B), les valeurs d_0=0 et s_0=2, qui défavorisent l'obtention de cette conclusion, peuvent être regardée comme "sceptiques", et les valeurs d_0=2 et s_0=1, qui la favorisent, comme "enthousiastes".

Nous nous limiterons ici à des choix de b_0 et q_0 correspondant à des effectifs $n_0=1/b_0^2$ et $m_0=q_0+1$ égaux.

LeBayésien

Toujours dans la situation précédente, entrez à nouveau pour la distribution initiale les valeurs

effet = 0 **écart-type** = 1

degrés de liberté = 9 **constante b** = 0.3162 ($1/\sqrt{10}$)

puis cliquez sur le bouton **calculer**. La distribution finale de δ est centrée sur +0.790 (au lieu de +1.58 pour la distribution fiducio-bayésienne) et a une dispersion plus faible.

Vous pouvez étudier les modifications de la distribution finale de δ (et de celle de δ_{cal}) obtenues en faisant varier les valeurs de d_0, s_0 q_0, b_0.

Par exemple, pour les valeurs q_0 et b_0 précédentes:

• pour d_0=0 et s_0=2, vous trouvez toujours pour δ une distribution centrée sur +0.790, mais sa dispersion est maintenant sensiblement du même ordre que celle de la distribution fiducio-bayésienne;

• pour d_0=+1.5 et s_0=1, vous trouvez pour δ une distribution centrée sur +1.540, dont la dispersion est nettement plus faible que celle de la distribution fiducio-bayésienne;
• pour d_0=+1.5 et s_0=2, vous trouvez pour δ une distribution relativement proche de la distribution fiducio-bayésienne.
Etc.

Nous obtenons par exemple les intervalles de crédibilité relatifs à δ, pour la garantie γ=0.95, en fonction de n_0

	d_0=0 et s_0=1	d_0=+1.5 et s_0=1	d_0=+2 et s_0=1
n_0=100	-0.067<δ<+0.354	+1.315<δ<+1.699	+1.768<δ<+2.155
n_0=25	+0.013<δ<+0.890	+1.161<δ<+1.884	+1.153<δ<+2.247
n_0=10	+0.154<δ<+1.426	+1.029<δ<+2.051	+1.270<δ<+2.310
n_0=4	+0.489<δ<+1.768	+1.022<δ<+2.093	+1.157<δ<+2.243
n_0=1	+0.748<δ<+2.125	+0.935<δ<+2.211	+0.977<δ<+2.260
$n_0 \to 0$	+0.867<δ<+2.293	+0.867<δ<+2.293	+0.867<δ<+2.293

	d_0=0 et s_0=2	d_0=+1.5 et s_0=2	d_0=+2 et s_0=2
n_0=100	-0.233<δ<+0.520	+1.141<δ<+1.874	+1.595<δ<+2.329
n_0=25	-0.214<δ<+1.117	+0.906<δ<+2.140	+1.260<δ<+2.500
n_0=10	-0.056<δ<+1.636	+0.783<δ<+2.297	+1.027<δ<+2.553
n_0=4	+0.377<δ<+1.880	+0.893<δ<+2.222	+1.029<δ<+2.371
n_0=1	+0.748<δ<+2.125	+0.935<δ<+2.211	+0.977<δ<+2.260
$n_0 \to 0$	+0.867<δ<+2.293	+0.867<δ<+2.293	+0.867<δ<+2.293

3.3 Examiner la sensibilité des conclusions issues d'une expérience

3.3.1 Mise à l'épreuve d'une conclusion

Dans l'exemple de base de Student, nous avons une probabilité fiducio-bayésienne 0.989 que la différence δ soit supérieure à une demi-heure. Quelle est la sensibilité de cette conclusion? Les résultats de la section précédente

fournissent des indications utiles sur l'influence de la distribution initiale. Considérons plus particulièrement la distribution initiale correspondant aux données potentielles obtenues pour dix patients supplémentaires (soit $q_0=q=9$ et $b_0 = b=1/\sqrt{10}$); nous pouvons examiner dans quelle mesure un résultat supplémentaire "défavorable" (ou "moins favorable") permettrait encore d'obtenir la même conclusion. Le tableau suivant donne la probabilité finale $Pr(\delta>+0.5)$, en fonction de différentes valeurs d_0 et s_0.

	$d_0=0$	$d_0=+0.25$	$d_0=+0.5$	$d_0=+0.75$
$s_0=0.8$	0.835	0.928	0.975	0.993
$s_0=1.0$	0.824	0.917	0.968	0.990
$s_0=1.2$	0.811	0.905	0.960	0.986
$s_0=1.4$	0.798	0.892	0.950	0.980
$s_0=1.6$	0.785	0.877	0.939	0.973
$s_0=1.8$	0.772	0.863	0.926	0.964

■ *Probabilité prédictive*

Ces analyses pourraient être complétées en calculant, à partir des données disponibles (ici dix patients), la probabilité *prédictive* d'observer pour dix patients supplémentaires un couple de valeurs (d_0,s_0) tel que la probabilité finale précédente $Pr(\delta>+0.5)$ soit supérieure à une garantie donnée γ. Cette probabilité prédictive constitue une information utile, pour décider d'arrêter l'expérimentation ou de la poursuivre dans le cas d'une analyse *intermédiaire* portant sur une partie seulement des observations qui avaient été planifiées (voir notamment Choi et Pepple, 1989; Grouin, 1994; Lecoutre, Derzko et Grouin, 1995). On trouvera dans le chapitre 8 une utilisation un peu différente des probabilités prédictives, pour déterminer les effectifs nécessaires.

3.3.2 Evaluer l'information supplémentaire nécessaire pour obtenir une conclusion

Considérons maintenant la situation où les données ne permettent pas d'obtenir la conclusion souhaitée. On pourra ici distinguer le cas où les données sont franchement défavorables à cette conclusion, du cas où elles lui sont plutôt favorables mais où l'information expérimentale n'apparaît pas suffisante à elle seule. Le premier cas nous ramène en fait à la situation envisagée à la section précédente, dans la mesure où nous aurions alors une probabilité

fiducio-bayésienne élevée en faveur d'une conclusion contraire.

Illustrons maintenant le second cas. Reprenons l'exemple de la validation d'un modèle additif (↳1.4). Nous avons une probabilité fiducio-bayésienne 0.90 que l'effet d'interaction δ soit inférieur en valeur absolue à 18 centièmes de seconde. Ce résultat est plutôt en faveur du modèle additif, mais on souhaiterait une limite plus petite que 18. Examinons donc l'influence d'une information supplémentaire "favorable au modèle" sur cet énoncé.

Considérons successivement:

- des distributions initiales pour δ centrées sur 0 (en gardant pour σ la distribution non informative, soit $q_0=0$), obtenues en faisant varier b_0;
- des distributions initiales pour σ d'échelle 20 (en gardant pour δ la distribution non informative, soit $b_{0=+\infty}$), obtenues en faisant varier q_0.

LeBayésien

Sélectionnez l'exemple **modèle additif**. Dans la fenêtre des statistiques, sélectionnez l'option **moyenne pondérée**, puis cliquez sur le bouton **ok**, ce qui vous place dans la fenêtre principale.

Activez la case à cocher **distribution initiale**, et entrez pour la distribution initiale les valeurs

effet = 0 **degrés de liberté** = 0 **constante b** = 0.5

(la valeur de l'écart-type est dans ce cas non pertinente)

puis cliquez sur le bouton **calculer**. Etc.

Pour la garantie 0.90 vous obtenez maintenant $|\delta|<14.53$. Etc.

Entrez maintenant pour la distribution initiale les valeurs

écart-type = 20 **degrés de liberté** = 9 **constante b** = 1E100

(la valeur de l'effet est dans ce cas non pertinente)

Pour la garantie 0.90 vous obtenez maintenant $|\delta|<13.76$. Etc.

Nous pouvons ainsi évaluer l'information supplémentaire qui serait nécessaire pour obtenir une conclusion plus nettement favorable au modèle additif.

■ *Information relative à* δ: $d_0=0$ ($q_0=0$)

Nous obtenons par exemple

b_0	n_0	d_1	e_1	$\gamma=0.90$
0.50	4	-1.562	7.891	$\|\delta\|<14.53$
0.2887	12	-1.042	6.447	$\|\delta\|<11.81$
0.20	25	-0.676	5.195	$\|\delta\|<9.48$
0.10	100	-0.223	2.987	$\|\delta\|<5.43$
0.05	400	-0.061	1.558	$\|\delta\|<2.83$

■ *Information relative à* σ: $s_0=20$ ($b_0=+\infty$ d'où $d_1=d$)

Nous obtenons par exemple

q_0	e_1	$\gamma=0.90$
9	7.709	$\|\delta\|<13.76$
18	7.144	$\|\delta\|<12.63$
27	6.836	$\|\delta\|<12.03$
54	6.409	$\|\delta\|<11.23$
108	6.127	$\|\delta\|<10.72$
$q_0 \to \infty$	5.774	$\|\delta\|<10.10$

■ *Information relative à* δ et à σ

En prenant en compte à la fois des informations relatives à δ et à σ, nous obtenons par exemple pour $d_0=0$, $s_0=20$, $q_0=9$ et $b_0=0.2887$ ($1/\sqrt{12}$)

$$Pr(|\delta|<9.59) = 0.90$$

CHAPITRE 4

ANALYSE SPÉCIFIQUE: PLANS S<G>*O

4.1 INTRODUCTION À L'ANALYSE SPÉCIFIQUE

4.1.1 L'approche du modèle linéaire général et ses difficultés

L'approche traditionnelle en analyse de variance est celle du *Modèle Linéaire Général* (ou modèle de la *régression*), posé sur l'ensemble de toutes les données. Ce modèle met en jeu un certain nombre de conditions de validité ou suppositions (*assumptions*), lesquelles peuvent être classées en quatre catégories principales:

- les conditions d'*indépendance*;
- les conditions de *normalité*;
- s'il y a plus d'un groupe en jeu, les conditions d'*homogénéité* entre les variances et covariances des différents groupes;
- s'il y a pour une variable plusieurs observations pour chaque "sujet" (*mesures répétées*), les conditions de *structure* (*circularité* ou *sphéricité*...) sur les matrices de variances et covariances associées.

Mais, dès que le plan d'analyse est un peu complexe, en particulier en ce qui concerne les mesures répétées, il règne souvent une sorte de flou quand il s'agit d'établir des conditions appropriées pour le modèle général. Une difficulté considérable de l'approche du modèle général est que les suppositions faites doivent assurer la validité des inférences sur *tous* les effets de la décomposition standard du plan d'analyse. La recherche de conditions nécessaires et suffisantes pour réaliser un tel objectif dans un plan complexe constitue sans doute une intéressante ligne de recherche théorique. Pourtant, du point de vue de l'utilisateur, ces conditions apparaissent de plus en plus difficiles à admettre et même à comprendre quand la complexité des plans augmente. Il s'ensuit qu'on en vient rapidement à suspecter *toutes* les inférences qui sont faites, dès lors qu'une condition est en doute.

4.1.2 L'approche de l'analyse spécifique

Il apparaît donc nécessaire d'envisager une autre approche. Or des solutions beaucoup plus simples peuvent être trouvées quand la recherche des conditions de validité est restreinte à un effet particulier. Dans cette ligne, nous en venons à l'idée que, dans ce cas, des conditions globales, établies au niveau d'un modèle général, sont tout simplement *non pertinentes*. L'approche de *l'analyse spécifique* conduit précisément à rechercher et expliciter pour chaque comparaison d'intérêt les conditions de validité minimales.

Etant donné une comparaison à laquelle nous nous intéressons, nous dérivons d'abord, à partir des données de base, un ensemble de *données pertinentes*. Puis un *modèle spécifique* est posé sur ces données pertinentes, et les inférences pour cette comparaison sont faites par l'intermédiaire de ce modèle spécifique.

■ *Des avantages décisifs*

Dans un plan d'expérience de structure complexe, l'inférence spécifique pour une comparaison réduit de manière considérable le nombre de paramètres "parasites". Par suite elle offre la possibilité de s'affranchir aisément des conditions de validité irréalistes sur ces paramètres, en réduisant au minimum ces conditions, voire en les supprimant.

En outre toutes les procédures inférentielles peuvent être obtenues à partir de statistiques descriptives *naturelles*, qui sont facilement interprétables: moyennes, écarts-types, variances et covariances. La *formalisation algébrique* des notions de dérivation et de comparaison permet d'établir le lien avec les statistiques traditionnelles de l'analyse de variance: sommes-des-carrés et/ou carrés-moyens.

■ *Décomposition des sources de variation*

Contrairement à l'approche traditionnelle de l'analyse de variance, l'analyse spécifique ne présuppose pas une décomposition standard des sources de variation, déterminée par la structure du plan. Elle aura donc une beaucoup plus grande souplesse de mise en œuvre, en élargissant à la fois les plans qui peuvent être traités et les effets qui peuvent être analysés dans le cadre de ces plans.

Dans le cadre d'un plan d'analyse, l'utilisateur pourra choisir librement la décomposition des sources de variation systématiques qui est la mieux adaptée à ses *questions*. S'il le souhaite, il pourra tenir compte des mêmes contraintes que celles que l'on impose souvent dans l'approche traditionnelle,

notamment des contraintes d'*orthogonalité*, sans que cela constitue bien entendu une nécessité.

■ *Des classes de plans générales*

La conjugaison de la formalisation algébrique et de l'inférence spécifique permet de traiter des classes de plan très générales, caractérisées en fait uniquement par les relations entre les facteurs aléatoires (ceux qui dans l'inférence feront l'objet de la généralisation). Le même type de solutions s'appliquera, quels que soient les facteurs systématiques et leurs relations.

Nous définirons donc maintenant une structure générale de plans à un facteur aléatoire permettant de formaliser des plans complexes, la classe (*Sujets*, *Groupes*, *Occasions*) ou S<G>*O. Nous nous limiterons ici à l'étude de cette classe de base. Des classes de plans à plusieurs facteurs aléatoires peuvent être définies en combinant deux (ou plus) structures de ce type (un exemple est donné en ↳5.6); mais la complexité des solutions augmente rapidement, en même temps que les conditions de validité peuvent devenir rapidement exorbitantes.

4.2 PLANS S<G>*O

4.2.1 Plans S<G>*O: Définition

■ *Plans* S<G> *et Facteurs inter-sujets*

Le type de plan (*Sujets*, *Groupes*) est celui des groupes *séparés* ou *indépendants*. Chaque "sujet" $_s$ est affecté à un "groupe" $_g$ et un seul; nous dirons que "le facteur *Sujets* S est *emboîté* dans le facteur *Groupes* G", ce que nous noterons S<G>. Pour chaque sujet $_{s<g>}$, il y a exactement une observation (pour la variable analysée).

Les groupes pourront être décrits par des facteurs *inter-sujets*; il s'agira typiquement de facteurs de classification des sujets (âges, catégories socio-professionnelles, antécédents médicaux...), et de facteurs rendant compte de l'affectation des sujets à un ensemble de conditions expérimentales.

■ *Plans* S×O *et Facteurs intra-sujets*

Le type de plan (*Sujets*, *Occasions*) est celui des groupes *appariés* ou plans

à mesures répétées. Pour chaque "sujet" s et chaque "occcasion" o, il y a exactement une observation (pour la variable analysée). Nous dirons que "les facteurs *Sujets* S et *Occasions* O sont *croisés*", ce que nous noterons S×O.

Les occasions pourront être décrites par des facteurs *intra-sujets*. Il s'agira typiquement de facteurs rendant compte de conditions expérimentales auxquelles sont soumis tous les sujets, et de facteurs de localisation des observations effectuées, temporelle (heures, rangs...), spatiale (mains, hémisphères cérébraux...), etc.

■ *Plans* S<G>*O

Le type de plan (*Sujets*, *Groupes*, *Occasions*) combine les deux structures précédentes. Les sujets S sont emboîtés dans les groupes G et, pour chaque groupe $g \in G$, les sujets $S_{<g>}$ affectés à ce groupe sont croisés avec un sous-ensemble O_g des occasions O (soit $S_{<g>} \times O_g$).

La relation entre les groupes et les occasions peut être quelconque (mais pas celle entre les sujets et les occasions); en conséquence nous noterons ce type de plans S<G>*O, réservant la notation S<G>×O au cas où G et O sont croisés (pour chaque groupe O_g=O).

Aucune supposition n'est faite sur la structure des facteurs systématiques G et O, qui peuvent être décrits par des facteurs élémentaires en relation absolument quelconque. Des exemples sont donnés dans le chapitre ↳5.

4.2.2 Analyse spécifique des comparaisons dans un plan S<G>*O: Modèles à effets fixés et modèles mixtes

Dans un plan S<G>*O, G et O sont traités comme des facteurs systématiques, "à effets fixés". Les sources de variation *aléatoires* sont celles faisant intervenir le facteur *Sujets*. On peut en distinguer deux types: d'une part les comparaisons *inter-sujets* (par exemple la comparaison *intra*-G globale notée $\mathbb{S}(G)$) et d'autre part les comparaison d'interaction entre ces dernières comparaisons et les comparaisons sur G*O (par exemple l'interaction $\mathbb{S}(G).\mathbb{O}$). C'est pourquoi dans l'approche traditionnelle de l'analyse de variance, on présente généralement le modèle (général) pour un tel plan comme un modèle "*mixte*". Mais en fait les procédures usuelles sont fondamentalement les mêmes que pour un modèle "*à effets fixés*", c'est-à-dire que pour un plan S<G>; c'est ce que montre bien l'analyse spécifique de ces plans. L'analyse véritable suivant un modèle mixte, c'est-à-dire la prise en compte simultanée effective des deux types de sources de variation précédents, peut se faire dans le cadre de l'analyse spécifique, en considérant une classe de plans plus générale (à deux

facteurs aléatoires). Ceci ne sera cependant pas envisagé dans cet ouvrage.

4.2.3 Comparaisons de type produit sur G*O

Les comparaisons de type produit sur G*O sont celles qui peuvent être engendrées par des formes linéaires produit de formes marginales sur G et sur O (voir les exemples de ce chapitre). Ces comparaisons sont privilégiées car, d'une part comme nous le verrons elles se ramènent à une même procédure générale, et d'autre part les suppositions usuelles d'homogénéité des variances sont en général pertinentes (sinon admissibles) pour ces comparaisons. D'un point de vue pratique, ces comparaisons peuvent en outre être analysées dans un plan partiel S<$\mathcal{G}$>×$\mathcal{O}$, où $\mathcal{G}$ et $\mathcal{O}$ sont des sous-ensembles croisés de G et O. La relation particulière entre G et O n'intervient pas dans l'analyse elle-même.

4.3 Plan du chapitre et exemples de données

4.3.1 Plan du chapitre

Nous développerons dans ce chapitre la mise en œuvre de l'*analyse spécifique* pour les principaux types de comparaisons *à un degré de liberté* d'un plan S<G>*O. Ceci nous permettra dans chaque cas d'expliciter les conditions de validité de chaque inférence effectuée. Nous montrerons ainsi que le tableau usuel d'analyse de variance peut être retrouvé comme un ensemble d'inférences spécifiques. Mais il faut bien insister sur le fait que nous pouvons analyser de la même manière toute comparaison de type produit sur G*O: les exemples présentés sont à valeur générale pour ce type, y compris dans un plan non équilibré.

Puisqu'il s'agit essentiellement ici de présenter les procédures de calcul, nous considérerons un exemple artificiel avec des données simples et en petit nombre. De même nous nous limiterons, en ce qui concerne les procédures inférentielles, à indiquer les statistiques permettant d'appliquer les différentes procédures, sans en fournir d'illustration.

Nous discuterons ensuite les conditions d'application de l'analyse spécifique permettant d'en faire un usage raisonnable.

Interviennent dans ce chapitre les notions formalisées dans le cadre de l'*algèbre linéaire*: dérivation, pondération, forme linéaire, norme, contraste, comparaison, décomposition. Il sera suffisant ici de prendre ces notions dans un sens intuitif.

4.3.2 Exemple de données

Considérons l'exemple typique d'un plan S<G>*O, avec:

- deux *groupes* (g_1 et g_2) de quatre sujets chacun, soit les effectifs $f_{g1}=4$ et $f_{g2}=4$;
- six *occasions* (o_1, o_2... o_6) définies par le croisement de deux facteurs *répétés* A et B avec respectivement trois et deux modalités, soit $O_6=A_3\times B_2$.

Nous supposerons que les six occasions sont affectées d'un poids uniforme (que nous prendrons égal à 1), sans que cela diminue la généralité des procédures.

Nous pouvons écrire ce plan $S_4<G_2>\times A_3\times B_2$. Les données relatives à une certaine variable sont fournies dans le tableau ci-après

groupes		occasions					
		o1 = a1 b1	o2 = a2 b1	o3 = a3 b1	o4 = a1 b2	o5 = a2 b2	o6 = a3 b2
g1	s1	44	48	37	29	22	24
	s2	33	24	24	39	34	28
	s3	16	15	27	25	22	34
	s4	19	25	16	9	8	5
g2	s5	13	6	6	12	14	20
	s6	16	11	4	21	34	26
	s7	15	3	1	24	11	11
	s8	36	31	21	15	8	3

Le tableau suivant fournit les statistiques usuelles de l'analyse de variance, sommes-des-carrés et carrés-moyens, selon la décomposition standard des sources de variation du plan.

Nous allons maintenant voir comment nous pouvons reconstituer ce tableau comme un ensemble d'analyses spécifiques.

Comparaison	*dl*	*sc*	*cm*
G	1	1250.5208	1250.5208
S(G)	6	1786.2920	297.7153
A	2	199.6250	99.8125
G.A	2	52.7917	26.3958
S(G).A	12	569.5833	47.4653
B	1	3.5208	3.5208
G.B	1	150.5208	150.5208
S(G).B	6	1979.7917	329.9653
A.B	2	37.0417	18.5208
G.A.B	2	65.0417	32.5208
S(G).A.B	12	138.5833	11.5486

4.4 Comparaison à un degré de liberté sur les groupes

Comparons globalement les moyennes des deux groupes. Nous noterons cette comparaison G ou "g1,g2".

4.4.1 Protocole dérivé pertinent

Nous commençons par calculer, pour chaque sujet, la moyenne de ses six observations (pour chacune des six occasions), par exemple, pour le sujet s1,

$$\frac{44+48+37+29+22+24}{6} = \frac{1}{6}44 + \frac{1}{6}48 + \frac{1}{6}37 + \frac{1}{6}29 + \frac{1}{6}22 + \frac{1}{6}24 = 34$$

Puisque nous avons supposé un poids uniforme de 1 pour toutes les occasions, nous considérons ici la moyenne équipondérée; nous pouvons bien entendu utiliser la même procédure pour toute autre dérivation.

Les moyennes individuelles ainsi obtenues constituent le "protocole dérivé" pertinent naturel pour la comparaison "g1,g2".

g^1	s^1	34.0000		
	s^2	30.3333	moyenne y^{g1}	25.2917
	s^3	23.1667	éty-cor s_{g1}	8.9612
	s^4	13.6667		
g^2	s^5	11.8333		
	s^6	18.6667	moyenne y^{g2}	15.0833
	s^7	10.8333	éty-cor s_{g2}	4.3515
	s^8	19.0000		

4.4.2 Dérivations sur les occasions et sur les groupes

Le protocole dérivé est obtenu par une *dérivation sur les occasions*, qui est définie par la forme linéaire $\mathbf{w}_O=(w_o)_{o\in O}$ sur O de coefficients

$$\mathbf{w}_O = [1/6 \;\; 1/6 \;\; 1/6 \;\; 1/6 \;\; 1/6 \;\; 1/6]$$

dont la somme est égale à 1: c'est une dérivation "*par moyennage*".

Définissons la constante a>0 telle que

$$a^2 = \sum_{o\in O} w_o^2 = (1/6)^2 + (1/6)^2 + (1/6)^2 + (1/6)^2 + (1/6)^2 + (1/6)^2 = \frac{1}{6}$$

Le protocole dérivé pertinent est muni de la pondération uniforme qui affecte à chaque sujet s le poids

$$\pi_s = 1/a^2 = 6$$

Ici, pour le protocole dérivé, nous nous intéressons à la différence des moyennes des deux groupes. Ceci correspond à une *dérivation sur les groupes*, qui est définie par la forme linéaire $\mathbf{v}_G=(v_g)_{g\in G}$ sur G de coefficients

$$\mathbf{v}_G = [+1 \;\; -1]$$

dont la somme est nulle: c'est une dérivation "*par contraste*"

Définissons la constante b>0 telle que

$$b^2 = \sum_{g\in G} \frac{v_g^2}{f_g} = \frac{(+1)^2}{4} + \frac{(-1)^2}{4} = \frac{1}{2}$$

Finalement la comparaison "g1,g2" est représentée par le contraste sur G*O

	o1	o2	o3	o4	o5	o6	
g1	+1/6	+1/6	+1/6	+1/6	+1/6	+1/6	+1
g2	-1/6	-1/6	-1/6	-1/6	-1/6	-1/6	-1
	1/6	1/6	1/6	1/6	1/6	1/6	

qui est le produit croisé de la forme linéaire $\mathbf{w}_O$ sur O et du contraste $\mathbf{v}_G$ sur G.

■ *Composantes de norme*

Dans la formalisation algébrique, le produit des constantes a et b, ν=ab =1/√12=0.2887, est la *norme* de ce contraste sur G*O; c'est pourquoi a et b sont encore appelées "composantes de norme".

4.4.3 Analyse descriptive

LeBayésien

Dans la fenêtre principale, cliquez sur le bouton **exemples**, et sélectionnez **plan S<G2>×A3×B2**. Les données sont affichées dans l'éditeur de données.

Cliquez sur le bouton **calculer**. Apparaît une fenêtre pour le "choix de la structure dérivée". Sélectionnez **S<G> moyennage sur toutes les occasions** (option par défaut), puis cliquez sur le bouton **ok**. Les statistiques du protocole dérivé pertinent sont calculées. Notez la pondération 6=1/a².

Cliquez sur **combinaison linéaire** et entrez les deux coefficients [1 -1] (ici, puisqu'il n'y a que deux groupes, l'option **comparaison globale** donnerait le même résultat).

	effectif	moyenne	écart-type	coefficient
g1	4	25.2916...	8.9612...	1
g2	4	15.0833...	4.3514...	-1

Puis, sans rien modifier d'autre, cliquez sur le bouton **calculer** pour effectuer l'analyse. Cliquez sur la liste déroulante **t - degrés de liberté: q=6** pour obtenir toutes les quantités mentionnées ci-après. Pour agrandir cette liste, cliquez sur le losange à sa gauche.

Le protocole dérivé pertinent est résumé par les statistiques descriptives suivantes:

• les moyennes y^{g1}=25.2917 et y^{g2}=15.0833, à partir desquelles nous calculons l'effet observé de la comparaison "g1,g2" comme la différence

$$d = \sum v_g y^g = 25.2917-15.0833 = +10.208$$

• les écarts-types-corrigés s_{g1}=8.9612 et s_{g2}=4.3515 (avec chacun 4-1=3 degrés de liberté), à partir desquels nous calculons l'écart-type-corrigé intra-groupe moyen s (auquel sont associés q=3+3=6 degrés de liberté) comme la moyenne quadratique

$$s = ((8.9612^2+4.3515^2)/2)^{1/2} = 7.044 \quad [q = 6\ dl]$$

(dans le cas de groupes non équilibrés d'effectifs f_g, cette moyenne serait pondérée par les nombres de degrés de liberté f_g-1).

Les valeurs d=+10.208 et s=7.044 constituent le résumé descriptif de l'analyse de la comparaison "g1,g2". L'effet *calibré* observé de la comparaison est

$$d_{cal} = \frac{d}{s} = +1.449$$

4.4.4 Procédures d'inférence

Nous sommes ramenés au problème général de l'inférence sur une combinaison linéaire de moyennes de groupes indépendants (ici une différence de deux moyennes), et nous pouvons appliquer les procédures d'inférence développées dans les chapitres précédents (sous le *modèle normal* usuel avec égalité des variances).

Par exemple la distribution fiducio-bayésienne relative à l'effet vrai δ est

$$\delta \sim t_q(d,e^2) \sim t_6(+10.208,4.981^2)$$

$$\text{avec } e = bs = 4.981 \quad (b = 1/\sqrt{2})$$

Cette solution s'étendra de manière évidente à toute dérivation linéaire sur les occasions et à toute combinaison linéaire δ de moyennes des groupes, que le plan soit ou non équilibré.

4.4.5 Rapport F de l'analyse de variance et t de Student

L'analyse spécifique de la comparaison $\mathbb{G}$ correspond à un problème élémentaire d'analyse de variance: il s'agit d'un plan en groupes indépendants ("*one-way*"). L'hypothèse nulle d'absence d'effet du facteur *Groupes* est traditionnellement testée au moyen du rapport F du carré-moyen inter $cm_{\mathbb{G}}$ au carré-moyen intra $cm_{\mathbb{S}(G)}$

$$F = 1250.5208/297.7153 = 4.200 \quad [1 \text{ et } 6\ dl]$$

Ce rapport F est le carré du t de Student

$$t = \frac{d}{e} = +2.049 \quad [6\ dl]$$

car ses carrés-moyens sont respectivement proportionnels à d^2 et s^2 (compte tenu de la pondération du protocole dérivé pertinent)

$$cm_{\mathbb{G}} = d^2/a^2b^2 = 1250.521 \quad (1/a^2b^2 = 12)$$

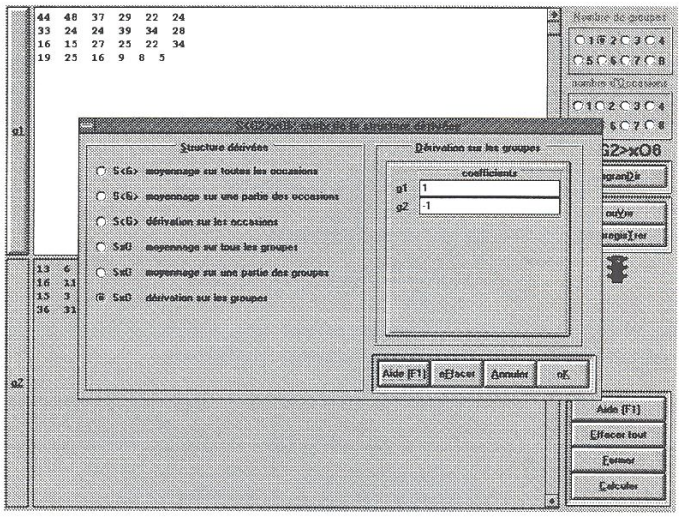
$$cm_{\mathbb{S}(G)} = s^2/a^2 = 297.715 \quad (1/a^2 = 6)$$

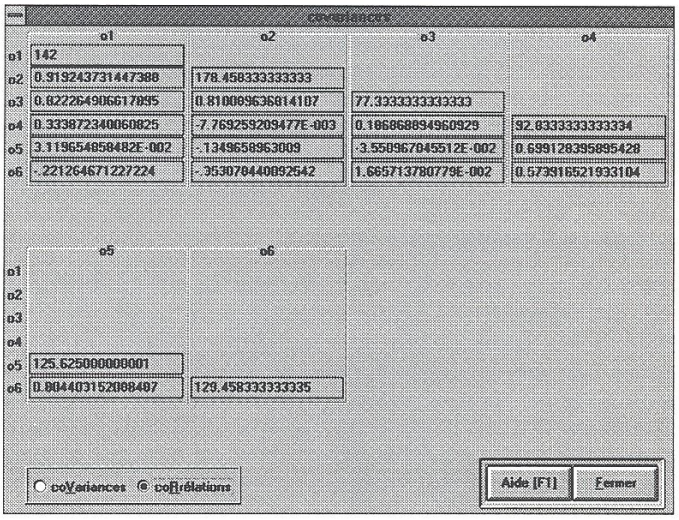
$$F = \frac{d^2/a^2b^2}{s^2/a^2} = \frac{d^2}{b^2s^2} = 4.200 = t^2$$

Nous retrouvons donc comme une inférence spécifique la solution usuelle de l'approche traditionnelle de l'analyse de variance univariée pour le plan à mesures répétées considéré. Mais ici les conditions de validité de cette inférence particulière sont parfaitement définies et faciles à comprendre. Plus simplement encore, nous pouvons tester l'hypothèse nulle d'absence d'effet du facteur *Groupes* au moyen d'un t de Student, avec l'avantage de ne faire intervenir que des statistiques *élémentaires naturelles*: moyennes et écarts-types. Ceci prépare à l'évidence la mise en œuvre des procédures bayésiennes.

■ *Remarque sur les constantes* a *et* b

Seule la constante b intervient en fait dans le calcul de la statistique de test (t ou F). La constante a ne sert que pour le calcul des statistiques de l'analyse de variance, sommes-des-carrés et carrés-moyens. Par conséquent si nous effectuons sur les données dérivées pertinentes une analyse de variance usuelle pour un plan en groupes indépendants, sans tenir compte de cette constante (c'est-à-dire si nous la prenons égale à 1), nous retrouvons les mêmes statistiques de test, mais avec des carrés-moyens seulement proportionnels à leurs valeurs calculées sur les données de base.


44 48 37 29 22 24
33 24 24 39 34 28
16 15 27 25 22 34
19 25 16 9 8 5
Nombre de groupes
Nombre d'occasions
Structure dérivée
S<G> moyennage sur toutes les occasions
S<G> moyennage sur une partie des occasions
S<G> dérivation sur les occasions
SxG moyennage sur tous les groupes
SxG moyennage sur une partie des groupes
SxG dérivation sur les groupes
Dérivation sur les groupes
coefficients
g1 1
g2 -1
Aide [F1]
Effacer tout
Fermer
Calculer


covariances
o1 o2 o3 o4
o1 142
o2 0.919243731447388 178.458333333333
o3 0.822264906617095 0.810089636814107 77.3333333333333
o4 0.333872340060825 -7.769259209477E-003 0.186868894960929 92.8333333333334
o5 3.119654858482E-002 -.1349658963009 -3.550967045512E-002 0.699128395895428
o6 -.221264671227224 -.953078440892542 1.685713780779E-002 0.579916521933104
o5 o6
o1
o2
o3
o4
o5 125.625000000001
o6 0.804403152088407 129.458333333335
coVariances
coCorrélations
Aide [F1]
Fermer

4.4.6 Effectuer d'abord la dérivation sur les groupes

Toutes les statistique globales précédentes peuvent être obtenues en effectuant d'abord la dérivation sur les groupes, puis ensuite celle sur les occasions, ce qui illustre les propriétés des dérivations.

LeBayésien

Dans la situation précédente (↳4.4.4), dans la fenêtre d'entrée des statistiques, cliquez sur le bouton **données**, puis, dans l'éditeur de données, cliquez sur le bouton **calculer**. Dans la fenêtre de "choix de la structure dérivée", sélectionnez maintenant **S×O dérivation sur les groupes**, entrez les deux coefficients [1 -1], puis cliquez sur le bouton **ok**, ce qui vous ramène à la fenêtre des statistiques

Dans la colonne **effectif**, vous obtenez pour chacune des six occasions la même valeur 2, qui correspond en fait ici à un *poids* égal à $1/b^2$ (notez l'analogie avec le poids $1/a^2$ associé précédemment aux données pertinentes obtenues par dérivation sur les occasions).

Dans la colonne **moyenne**, vous obtenez pour chacune des six occasions la combinaison linéaire des moyennes des groupes correspondant à la dérivation sur les groupes (ici la différence entre g_1 et g_2).

Dans la colonne **écart-type**, vous obtenez les écarts-types-corrigés intragroupes moyens associés à chaque occasion, et calculés sur l'ensemble des groupes en jeu.

Dans la situation présente d'un plan (ici dérivé) de structure S×O, interviennent les covariances associées à chaque paire d'occasions. C'est pourquoi apparaît une nouvelle zone de texte **covariances** (avec ici la valeur **142...**). En cliquant sur cette zone, vous faîtes apparaître une fenêtre, avec les variances et covariances associées aux différentes occasions. En cliquant sur l'option **corrélations**, vous obtenez les coefficients de corrélation linéaire entre ces occasions. Fermez cette fenêtre, pour revenir à la fenêtre des statistiques.

Cliquez sur **moyenne équipondérée**, pour obtenir les coefficients [1/6 ... 1/6] définissant la dérivation sur les occasions.

	effectif	moyenne	écart-type	coefficient
o1	2	8	11.9163...	1/6
o2	2	15.25	13.3588...	1/6
	...	...	...	...
o6	2	7.75	11.3779...	1/6

Puis cliquez sur le bouton **calculer** pour effectuer l'analyse, et vérifiez que les résultats sont identiques à ceux obtenus précédemment.

Cette manière de procéder présente un double intérêt dans l'utilisation du programme: obtenir les covariances et les corrélations entre les différentes occasions, et pouvoir étudier aisément la régression polynomiale sur les occasions (voir des exemples au chapitre 5).

4.5 Comparaison à un degré de liberté sur les occasions

Les solutions pour une comparaison à un degré de liberté sur les occasions sont techniquement analogues à celles du cas précédent. Considérons l'exemple de la comparaison B ou "b1,b2".

4.5.1 Analyse spécifique: Dérivation

Dans ce cas, nous calculons pour chaque sujet la différence entre la moyenne de ses trois observations relatives à b1 et la moyenne de ses trois observations relatives à b2, par exemple, pour le sujet s1,

$$\frac{44+48+37}{3} - \frac{29+22+24}{3} = \frac{1}{3}44 + \frac{1}{3}48 + \frac{1}{3}37 - \frac{1}{3}29 - \frac{1}{3}22 - \frac{1}{3}24 = 43-25 = +18$$

Ces différences individuelles constituent le protocole dérivé pertinent naturel pour la comparaison "b1,b2".

La dérivation sur les occasions est définie par le contraste sur O

$$\mathbf{w}_O = [+1/3 \ +1/3 \ +1/3 \ -1/3 \ -1/3 \ -1/3]$$

d'où la constante a>0 telle que

$$a^2 = \sum_{o \in O} w_o^2 = (1/3)^2 + (1/3)^2 + (1/3)^2 + (1/3)^2 + (1/3)^2 + (1/3)^2 = \frac{2}{3}$$

Le protocole dérivé pertinent est muni de la pondération uniforme qui affecte à chaque sujet s le poids $\pi_s = 1/a^2 = 3/2$.

		moyenne pour b1	moyenne pour b2	différence
g1	s1	43.0000	25.0000	+18.0000
	s2	27.0000	33.6667	- 6.6667
	s3	19.3333	27.0000	- 7.6667
	s4	20.0000	22.0000	+12.6667
			moyenne d^{g1}	+4.0833
			éty-cor s_{g1}	13.1779

		moyenne pour b1	moyenne pour b2	différence
g2	s5	8.3333	15.3333	- 7.0000
	s6	10.3333	27.0000	-16.6667
	s7	6.3333	15.3333	- 9.0000
	s8	29.3333	8.6667	+20.6667
			moyenne d^{g2}	-3.0000
			éty-cor s_{g2}	16.3186

Ici, pour le protocole dérivé, nous nous intéressons à la moyenne générale des deux groupes, obtenue par la forme linéaire sur G

$$\mathbf{v}_G = [1/2 \ \ 1/2]$$

d'où la constante b telle que

$$b^2 = \sum_{g \in G} \frac{v_g^2}{f_g} = \frac{(+1/2)^2}{4} + \frac{(-1/2)^2}{4} = \frac{1}{8}$$

Finalement la comparaison "b1,b2" est représentée par le contraste sur G*O

	o1 b1	o2 b1	o3 b1	o4 b2	o5 b2	o6 b2	
g1	+1/6	+1/6	+1/6	-1/6	-1/6	-1/6	1/2
g2	+1/6	+1/6	+1/6	-1/6	-1/6	-1/6	1/2
	+1/3	+1/3	+1/3	-1/3	-1/3	-1/3	

qui est le produit croisé du contraste sur $\mathbf{w}_O$ et de la forme linéaire $\mathbf{v}_G$, et qui a la norme ν=ab=1/√12=0.2887

4.5.2 Analyse descriptive

LeBayésien

Il est supposé que vous venez d'effectuer l'analyse de la comparaison "g1,g2" (sinon placez vous dans la situation correspondante). Dans la fenêtre d'entrée des statistiques, cliquez sur le bouton **données**, puis, dans l'éditeur de données, cliquez sur le bouton **calculer**. Dans la fenêtre de "choix de la structure dérivée", sélectionnez **S<G> dérivation sur les occasions**, entrez les six coefficients [1/3 1/3 1/3 -1/3 -1/3 -1/3], puis cliquez sur le bouton **ok**.

Les statistiques du protocole dérivé pertinent sont calculées. Notez la pondération $1.5=1/a^2$.

Cliquez sur **moyenne équipondérée**.

	effectif	moyenne	écart-type	coefficient
g1	4	4.0833...	13.1779...	1/2
g2	4	-3	16.3185...	1/2

Puis cliquez sur le bouton **calculer** pour effectuer l'analyse.

Le protocole dérivé est résumé par:

• les moyennes des différences à l'intérieur de chaque groupe $d^{g1}=+4.0833$ et $d^{g2}=-3.0000$, à partir desquelles nous calculons l'effet de la comparaison "b1, b2" comme la moyenne

$$d = ((+4.0833)+(-3.0000))/2 = +0.542$$

• les écarts-types-corrigés $s_{g1}=13.1779$ et $s_{g2}=16.3186$, à partir desquels nous définissons l'écart-type intragroupe moyen

$$s = ((13.1779^2+16.3186^2)/2)^{1/2} = 14.832 \quad [q = 3+3 = 6\ dl]$$

Les valeurs d=+0.542 et s=14.832 constituent le résumé descriptif de l'analyse de la comparaison "b1,b2". L'effet calibré est ici

$$d_{cal} = \frac{d}{s} = +0.037$$

4.5.3 Inférence spécifique

Nous retrouvons le problème général de l'inférence sur une combinaison linéaire de moyennes de groupes indépendants. Simplement l'inférence porte maintenant sur la moyenne générale $(\delta\varepsilon^1+\delta\varepsilon^2)/2$ (dans le cas de groupes non équilibrés, on pourrait choisir entre la moyenne équipondérée et une moyenne pondérée en fonction des effectifs).

Nous avons notamment ici, sous le *modèle normal* usuel avec égalité des variances

$$e = bs = 5.244 \quad (b = 1/\sqrt{8})$$

$$t = \frac{d}{e} = +0.103 \quad [6\ dl]$$

En analyse de variance, l'hypothèse nulle d'absence d'effet du facteur B est traditionnellement testée au moyen du rapport F du carré-moyen $cm_{\mathbb{B}}$ associé à la comparaison $\mathbb{B}$ au carré-moyen $cm_{\mathbb{S}(G).\mathbb{B}}$ associé à la comparaison *adjointe* $\mathbb{S}(G).\mathbb{B}$. Compte tenu de la pondération du protocole dérivé pertinent, ces deux carrés-moyens sont respectivement proportionnels à d^2 et s^2

$$cm_{\mathbb{B}} = d^2/a^2b^2 = 3.521 \quad (1/a^2b^2 = 12)$$

$$cm_{\mathbb{S}(G).\mathbb{B}} = s^2/a^2 = 329.965 \quad (1/a^2 = 3/2$$

$$\text{d'où } F = t^2 = cm_{\mathbb{B}}/cm_{\mathbb{S}(G).\mathbb{B}} = 0.011 \quad [1 \text{ et } 6\ dl]$$

4.6 Interaction a un degré de liberté entre les groupes et les occasions

Les solutions pour une comparaison d'interaction entre une comparaison à un degré de liberté sur les groupes et une comparaison à un degré de liberté sur les occasions sont encore techniquement analogues à celles des cas précédent. Considérons, comme exemple d'une comparaison d'interaction entre les groupes et les occasions, la comparaison $\mathbb{G}.\mathbb{B}$ ou "g1,g2.b1,b2".

4.6.1 Analyse spécifique: Dérivations

Les différences individuelles calculées précédemment pour la comparaison "b1,b2" constituent encore le protocole dérivé pertinent naturel. Simplement ici l'inférence porte sur la différence des différences $(\delta\varepsilon^1-\delta\varepsilon^2)/2$, et non

plus sur leur moyenne. Nous avons donc

- le même contraste sur les occasions que pour la comparaison "b1,b2"

$$\mathbf{w}_O = [+1/3 \ +1/3 \ +1/3 \ -1/3 \ -1/3 \ -1/3]$$

- le même contraste sur les groupes que pour la comparaison "g1,g2"

$$\mathbf{v}_G = [+1 \ -1]$$

	o1 b1	o2 b1	o3 b1	o4 b2	o5 b2	o6 b2	
g1	+1/3	+1/3	+1/3	-1/3	-1/3	-1/3	+1
g2	-1/3	-1/3	-1/3	+1/3	+1/3	+1/3	-1
	+1/3	+1/3	+1/3	-1/3	-1/3	-1/3	

Ce contraste produit a la norme

$$\nu = ab = 1/\sqrt{3} = 0.5774$$

avec $a^2 = \Sigma w_o^2 = 2/3$ et $b^2 = \Sigma v_g^2/f_g = 1/2$

4.6.2 Analyse descriptive

LeBayésien

Dans la fenêtre d'entrée des statistiques, dans la même situation que pour la comparaison "b1,b2", cliquez maintenant sur **combinaison linéaire** et entrez les deux coefficients [1 -1].

	effectif	moyenne	écart-type	coefficient
g1	4	4.0833...	13.1779...	1
g2	4	-3	16.3185...	-1

Puis cliquez sur le bouton **calculer** pour effectuer l'analyse.

L'effet observé de la comparaison "g1,g2.b1,b2" est défini comme la différence

$$d = +4.0833 - (-3.0000) = +7.083$$

d'où l'effet calibré

$$d_{cal} = \frac{d}{s} = \frac{+7.083}{14.832} = +0.478$$

4.6.3 Inférence spécifique

Nous avons notamment ici sous le *modèle normal* usuel avec égalité des variances

$$e = bs = 10.488 \quad (b = 1/\sqrt{2})$$

$$t = \frac{d}{e} = +0.675 \quad [6\ dl]$$

En analyse de variance, l'hypothèse nulle d'absence d'effet d'interaction entre les facteurs G et B est traditionnellement testée au moyen du rapport F du carré-moyen $cm_{\mathbb{G}.\mathbb{B}}$ associé à la comparaison $\mathbb{G}.\mathbb{B}$ au carré-moyen $cm_{\mathbb{S}(G).\mathbb{B}}$ associé à la comparaison *adjointe* $\mathbb{S}(G).\mathbb{B}$.

$$cm_{\mathbb{B}} = d^2/a^2b^2 = 150.521 \qquad (1/a^2b^2 = 3)$$

$$cm_{\mathbb{S}(G).\mathbb{B}} = s^2/a^2 = 329.965 \qquad (1/a^2 = 3/2$$

$$\text{d'où } F = t^2 = cm_{\mathbb{G}.\mathbb{B}}/cm_{\mathbb{S}(G).\mathbb{B}} = 0.456 \quad [1 \text{ et } 6\ dl]$$

4.7 Vers un usage raisonnable de l'analyse spécifique

4.7.1 Comment justifier une analyse spécifique?

On peut se demander sous quelles conditions l'inférence spécifique relative à un effet coïncide avec l'inférence que l'on pourrait effectuer sur ce même effet dans le cadre d'un modèle général. Il suffira ici de considérer la situation de la comparaison des deux moyennes associées aux occasions dans un plan avec deux occasions o_1 et o_2 (et un seul groupe). δ est la différence des deux moyennes parentes. Cette situation a été discutée par Lindley (1965, page 85) dans les termes suivants, où x_1 et x_2 désignent les variables associées respectivement à chacune des occasions, z est la différence x_1-x_2 et w est la somme x_1+x_2.

"We can write... $p(x_1,x_2)=p(z,w)=p(z)p(w|z)$. Whilst $p(z)$ depends on δ, $p(w|z)$ typically will not, and whatever parameters the latter does depend on, they usually do not occur in $p(z)$ and will, prior to the observations, be independent of those in $p(z)$. Consequently as far as inferences about δ are concerned one may confine attention to $p(z)$, absorbing $p(w|z)$ into the constant of proportionality... and the inference using z only will not lose any information."

Clairement l'argument de Lindley s'applique à l'analyse spécifique de toute comparaison sur les occasions dans un plan S<G>*O. Il suffit de remplacer la différence z par l'effet associé à cette comparaison et w par l'effet associé à sa résiduelle par rapport à la comparaison globale des occasions. La condition est que la distribution initiale relative aux paramètres du modèle général soit telle que le sous-espace (pertinent) correspondant du modèle spécifique soit indépendant du sous-espace résiduel des paramètres. Cette condition apparaîtra raisonnable dans le cadre de l'inférence fiducio-bayésienne. On trouvera une discussion détaillée dans Rouanet et Lecoutre (1983).

On notera que l'intérêt de l'analyse spécifique est souvent implicitement reconnu. Par exemple Hand et Taylor (1987) proposent d'effectuer une dérivation systématique des données pertinentes avant d'utiliser les programmes informatiques standard d'analyse de variance.

4.7.2 Choix méthodologiques liés à l'analyse spécifique

L'analyse spécifique d'une comparaison fait apparaître, dès les procédures descriptives, les choix méthodologiques qui conditionnent les procédures inférentielles.

L'argument précédent s'applique en effet à une comparaison faisant intervenir l'ensemble des groupes. Si tel n'est pas le cas, les conditions de validité portant sur les variances (et éventuellement d'autres paramètres) interviendront et pourront éventuellement occasionner des différences entre l'inférence spécifique et l'inférence générale.

Ainsi, si nous analysons une comparaison *partielle* sur les groupes, par exemple la comparaison "g1,g2" dans un plan avec trois groupes, nous pouvons choisir, pour calculer l'écart-type intragroupe moyen

- soit de ne faire intervenir que les *deux* groupes pertinents;
- soit de considérer l'ensemble des *trois* groupes.

Il correspond à ces deux choix des solutions inférentielles différentes, avec des conditions de validité qui sont faciles à comprendre: il s'agit ici

de supposer l'égalité des variances, soit seulement des deux groupes pertinents, soit de l'ensemble des trois groupes.

Reprenons l'exemple précédent du plan $S_4<G>\times A_3\times B_2$, en ajoutant les données d'un troisième groupe de sujets $g3$. Les données du dérivées du groupe $g3$, obtenues par moyennage sur les six occasions, ont pour moyenne et écart-type-corrigé

$$y^{g3} = 27.5417 \quad \text{et} \quad s_{g3} = 5.9959$$

LeBayésien

Dans la fenêtre principale, cliquez sur le bouton **exemples**, et sélectionnez **plan S<G3>*A3*B2**. Les données sont affichées dans l'éditeur de données.

Cliquez sur le bouton **calculer**. Apparaît une fenêtre pour le "choix de la structure dérivée". Sélectionnez l'option **S<G> moyennage sur toutes les occasions**, puis cliquez sur le bouton **ok**. Les statistiques du protocole dérivé pertinent sont calculées.

Cliquez sur **combinaison linéaire** et entrez les coefficients [1 -1 0].

	effectif	moyenne	écart-type	coefficient
g1	4	25.2916...	8.9612...	1
g2	4	15.0833...	4.3514...	-1
g3	4	27.5416...	5.9959...	0

Effectuez l'analyse, d'une part avec les options **avec égalité des variances** et **spécifique**, et d'autre part avec les options **avec égalité des variances** et **non spécifique** (désactivez la case à cocher **spécifique**).

L'analyse de l'effet de la comparaison "g1,g2" conduit aux indices d'échelle et aux t respectifs:

- pour la solution *spécifique* ne faisant intervenir que les groupes $g1$ et $g2$

$$e = 4.981 \quad t = +2.049 \quad [6\ dl]$$

- pour la solution *non spécifique* faisant intervenir les trois groupes

$$e = 4.747 \quad t = +2.151 \quad [9\ dl]$$

La solution *spécifique* est celle donnée précédemment pour le plan avec deux groupes $S(G_2>\times A3\times B2$ (↳4.4.4).

La solution *non spécifique* est obtenue en calculant l'écart-type intragroupe sur l'ensemble des trois groupes

$$s = ((8.9612^2+4.3515^2+5.9959^2)/3)^{1/2} = 6.713 \quad \text{d'où } e = bs = 4.747$$

Techniquement, le passage à une inférence "moins spécifique" est immédiat: il suffit de remplacer les variances du protocole dérivé pertinent *spécifique* par celles adjointes à une autre comparaison dont la comparaison analysée est une sous-comparaison. Les conditions de validité sont celles relatives à la "sur-comparaison" choisie et sont évidemment plus restrictives.

4.7.3 Quelques lignes directrices pour un usage raisonnable de l'analyse spécifique

Il suffira ici encore de considérer la situation de la comparaison des deux moyennes associées aux occasions dans un plan avec deux occasions $o1$ et $o2$ (et un seul groupe). Le modèle général *binormal* usuel comporte cinq paramètres, les deux moyennes, les deux variances et la covariance; l'inférence effectuée sous ce modèle sur la différence entre les deux moyennes parentes μ^{o1} et μ^{o2} coïncide avec l'inférence spécifique basée sur le protocole dérivé des différences pour chaque sujet.

■ *Contraintes sur les paramètres du modèle général*

Cependant, si certaines contraintes sont introduites sur les cinq paramètres, cette coïncidence peut disparaître. Il convient ici de distinguer le cas où ces contraintes portent sur les moyennes comparées, et le cas où elles portent sur les variances et covariances.

• Comme exemple du premier cas, imaginons que l'on suppose dans le modèle que la différence $\delta=\mu^{o1}-\mu^{o2}$ est liée fonctionnellement à la somme $\mu^{o1}+\mu^{o2}$; alors l'inférence sur δ obtenue par le modèle général peut différer largement de l'inférence spécifique et sera certainement mieux appropriée.

• Comme exemple du second cas, supposons que la covariance parente est nulle; alors l'inférence basée sur le modèle général revient à la comparaison de deux groupes indépendants. Supposons en outre l'égalité des deux variances parentes associées aux deux occasions; le nombre de degrés de liberté sera alors deux fois plus élevé, et (comme cela est bien connu) l'inférence générale, "plus puissante", sera préférable à l'analyse spécifique. Mais ceci nécessite de prendre réellement au sérieux la supposition d'une covariance nulle...

■ *Des procédures de routine raisonnables*

A l'examen l'analyse spécifique apparaît fournir des procédures de routine qui, à la fois présentent dans la plupart des situations courantes un caractère raisonnable, et permettent de discuter de manière précise les conditions de validité de chaque inférence.

Cette discussion pourra déboucher sur deux sortes d'extension de ces procédures de routine.

• D'une part l'approche de l'analyse spécifique donne la possibilité de s'affranchir aisément des contraintes liées aux modèles usuels. En particulier, on pourra utiliser des solutions qui ne supposent plus l'égalité des variances. Par exemple, si la supposition de normalité est suspecte, il sera directement approprié d'envisager une transformation des données dérivées pertinentes. Etc.

• D'autre part, et à l'opposé, on pourra constamment avoir un regard critique sur chaque inférence spécifique effectuée, et si cela paraît approprié recourir à un modèle plus général, qu'un ensemble d'analyses spécifiques soigneusement choisies pourra permettre d'étayer.

4.7.4 Quelle stratégie adopter vis-à-vis des conditions de validité?

Une recommandation préliminaire est de s'interroger, avant de réaliser une expérience, sur les analyses spécifiques qu'elle permettra d'effectuer. Il est souvent possible de choisir un plan qui conduira naturellement à considérer des comparaisons à un degré de liberté, dont les résultats auront une interprétation claire et seront en outre souvent valides sous des conditions assez faibles. En particulier on élude ainsi totalement le problème de la condition de *circularité* (ou *sphéricité*), qui intervient seulement pour les comparaisons à plusieurs degrés de liberté sur les occasions.

Nous passerons maintenant rapidement en revue les différentes catégories de conditions de validité. Bien entendu, il faut être réaliste et considérer que ces différentes conditions (et en tout premier lieu celle de normalité) ne sont jamais entièrement satisfaites et apportent donc essentiellement des approximations commodes. La question réelle est de savoir si les conclusions peuvent être remises en cause quand ces conditions sont manifestement inappropriées.

■ *Les conditions d'indépendance*

On s'assurera chaque fois que cela est possible des conditions d'indépendance par les habituelles précautions expérimentales (ne pas laisser les sujets "copier" les uns sur les autres...). Si ces conditions doivent être assumées par hypothèse (voir la discussion sur l'exemple *apprentissage perceptif incident* en §6.5), on gardera à l'esprit qu'elles ont un caractère crucial et que les procédures peuvent être fortement faussées si elles ne sont pas remplies; un examen particulier s'imposera donc dans ce cas.

■ *Les conditions de normalité*

On sait que les écarts à la normalité les plus fréquemment rencontrés ont peu d'impact sur les distributions des statistiques de test t et F (*cf.* Pearson, 1931; Tan, 1982). Cette propriété s'applique même à une variable dépendante dichotomique dans un plan équilibré (Lunney, 1970).

L'inférence spécifique relative à l'effet de la comparaison ne devrait donc en principe être sérieusement remise en cause par la non-normalité que si les effectifs sont faibles ou fortement inégaux. Mais dans ce dernier cas les distributions fiducio-bayésiennes seront le plus souvent fortement dispersées et conduiront donc à constater l'insuffisance de l'information expérimentale.

La même propriété s'appliquera à l'inférence relative à l'effet calibré. En revanche on se montrera beaucoup plus prudent à l'égard des inférences relatives aux écarts-types, qui peuvent être fortement faussées.

On observera que l'approche de l'analyse spécifique permet facilement de considérer des types de distributions plus généraux, par exemple avec un troisième moment non nul, et de traiter les paramètres supplémentaires dans le cadre bayésien. Néanmoins, pour obtenir des solutions applicables à toute comparaison dans des plans généraux, il faudrait encore examiner de nouvelles conditions d'homogénéité liées à ces paramètres eux-mêmes...

■ *Les conditions d'homogénéité des variances et covariances*

Dans ce cas des solutions de rechange sont disponibles (voir le chapitre §6). En pratique ces solutions, auront le plus souvent un rôle rassurant, car elles sont généralement proches de celles supposant les conditions d'homogénéité. Cependant la non homogénéité ("*heteroscedasticity*") des variances peut dans certains cas avoir des conséquences importantes: *cf.* Brown et Forsythe (1974); Ramsey (1980), Wilcox, Charlin et Thompson (1986). Dans le cas où la solution sans condition d'homogénéité conduit à des résultats sensiblement différents, une attitude prudente impose donc de recourir à la solution la

moins favorable à la conclusion énoncée (qui n'est pas nécessairement celle supposant l'égalité des variances: *cf.* §6.3).

D'une manière générale, il nous semble en tout état de cause raisonnable de considérer les conditions d'homogénéité minimales associées aux données pertinentes: supposer l'égalité des variances de quatre groupes quand la comparaison ne porte que sur deux d'entre eux risque d'aboutir à un résultat illusoire. Ces conditions d'homogénéité minimales constitueront donc le plus souvent un compromis honorable.

■ *Les conditions de structure sur les matrices de variances et covariances*

Comme nous l'avons vu, l'analyse spécifique des comparaisons à un degré de liberté dans un plan *à mesures répétées* ne met en jeu aucune condition particulière de structure sur les matrices de variances et covariances.

4.8 L'ANALYSE DES COMPARAISONS QUI NE SONT PAS DE TYPE PRODUIT

L'analyse spécifique permet de traiter des comparaisons pour lesquelles les suppositions usuelles d'homogénéité des variances sont tout à fait irréalistes. C'est en général le cas pour une comparaison qui n'est pas de type produit sur G*O. Considérons ainsi un exemple simple d'analyse qui pourra être facilement généralisé. Supposons que, dans une certaine expérience, un premier groupe de sujets a été soumis à deux traitements, et un second groupe a été soumis seulement à l'un des traitements (par exemple un contrôle). Le plan est alors une structure incomplète $S<G_2>*O_2$ avec seulement les trois modalités g_1o_1, g_1o_2 et g_2o_1. La comparaison des deux traitements n'est pas une comparaison de type produit.

	o1	o2
g1	t1	t2
g2	t1	

Nous pouvons cependant appliquer très facilement l'approche de l'inférence spécifique. Considérons par exemple la solution équipondérée, nous définissons alors l'effet vrai de la comparaison des deux traitements comme la différence

$$\delta = \mu_{g1o2} - \tfrac{1}{2}(\mu_{g1o1}+\mu_{g2o1})$$

Les données dérivées pertinentes sont définies de la manière suivante, à partir des données de base $x^{s<g>o}$

pour le groupe $g1$ $\quad y^{s<g1>} = x^{s<g1>o2} - \frac{1}{2}x^{s<g1>o1}$

pour le groupe $g2$ $\quad y^{s<g2>} = \frac{1}{2}x^{s<g2>o1}$

Posons le *modèle normal* usuel

$$y^{s<g>} \sim N(\vartheta g, \sigma_g^2) \quad \text{où } \vartheta g1 = \mu g1o2 - \tfrac{1}{2}\mu g1o1 \text{ et } \vartheta g2 = \tfrac{1}{2}\mu g2o1$$

(avec les conditions d'indépendance)

Dans ce cas supposer l'égalité des deux variances σ^2_{g1} et σ^2_{g2} serait tout à fait irréaliste. L'inférence sur le paramètre $\delta = \vartheta g1 - \vartheta g2$ est en fait le problème de la comparaison de deux moyennes sous le modèle *normal*, avec variances quelconques. Nous pouvons donc appliquer les solutions "sans égalité des variances" développées au chapitre ↳6.

■ *Données manquantes*

La procédure précédente peut dans certains cas fournir une solution élégante pour traiter des plans avec données manquantes, sans pour autant bien entendu résoudre les problèmes d'interprétation liés à la cause de ces données manquantes.

Supposons par exemple que chaque sujet ait bien été soumis aux deux traitements; mais que pour certains sujets, on ne dispose des données que d'un seul traitement. Par exemple, en pharmacologie les sujets n'ont eu aucune réaction à l'un des traitements; ou dans une expérience de psychologie où l'on considère les temps de résolution de problèmes, certains sujets ne sont pas parvenus à résoudre l'un des problèmes. On se ramènera alors à la situation précédente, en distinguant trois sous-groupes de sujets:

- ceux qui ont des données pour les deux traitements;
- ceux qui ont des données seulement pour le traitement $t1$;
- ceux qui ont des données seulement pour le traitement $t2$.

En pratique, pour pouvoir utiliser les solutions proposées au chapitre ↳6, il faudra qu'il y ait au moins 6 sujets dans chacun de ces sous-groupes. En théorie, il est cependant toujours possible d'appliquer la procédure, en recourant à un calcul par intégration numérique, cette contrainte étant seulement liée au recours à des approximations faciles à mettre en œuvre.

CHAPITRE 5

ANALYSE SPÉCIFIQUE: ILLUSTRATIONS

Nous avons en fait déjà illustré l'approche de l'analyse spécifique. En particulier les données analysées par Student pour illustrer l'inférence "sur une moyenne" (↳1.1.1) étaient des données dérivées: pour chaque patient, la différence des gains de sommeil procurés par les deux somnifères. Ainsi Student "faisait de l'analyse spécifique"!

Dans ce chapitre, nous envisagerons sans aucune prétention d'exhaustivité différentes situations, correspondant soit à des types de plans usuels, plans *à mesures répétés*, plans *non-équilibrés*, plans *cross-over* notamment, soit à des plans posant des problèmes particuliers par leur complexité.

Comme nous l'avons dit au chapitre ↳4, l'approche de l'analyse spécifique s'appuye sur la formalisation algébrique des notions de *dérivation* et de *comparaison*. Cette formalisation fournit notamment une définition rigoureuse de la notion d'*orthogonalité*, qui intervient dans ce chapitre.

5.1 Exemples de plans à mesures répétées

5.1.1 Une expérience de temps de réaction

Considérons l'expérience suivante de psychologie expérimentale. Le sujet a devant lui un pupitre comportant un nombre égal de voyants et de boutons réponses. Dès que l'un des voyants s'allume, il doit appuyer le plus rapidement possible sur le bouton correspondant. On se propose d'étudier l'évolution du temps de réaction en fonction du nombre de voyants (incertitude du signal).

Considérons les données précédemment analysées par Rouanet, Oléron et Régnier (1966) et par Lecoutre (1981b). Cinq sujets adultes ont effectué chacun 64 essais, dans chacune des quatre conditions expérimentales caractérisées par le nombre de voyants: respectivement un, deux, quatre et huit. Nous adopterons comme plan d'analyse le plan *à mesures répétées* que nous noterons $S_5 \times C_4$

("sujets *croisés* avec les conditions"). Ce plan associe à chaque couple (s,c) la moyenne des 64 temps de réaction recueillis pour ce sujet et cette condition. Les données correspondantes (en centièmes de seconde) sont fournies dans le tableau suivant.

	1 voyant	2 voyants	4 voyants	8 voyants
sujet 1	52.218750	63.812500	75.750000	94.156250
sujet 2	70.203125	76.578125	73.609375	84.687500
sujet 3	38.343750	59.406250	67.046875	80.828125
sujet 4	45.015625	58.640625	64.515625	68.828125
sujet 5	41.703125	57.296875	71.250000	81.093750
moyenne	49.496875	63.146875	70.434375	81.918750

cs

L'analyse de cette expérience a pour objet d'examiner les deux hypothèses suivantes:

- une hypothèse générale - le temps de réaction augmente d'une manière importante lorsque le nombre de signaux augmente;
- une hypothèse spécifique - le temps de réaction est (approximativement) une *fonction linéaire du logarithme binaire* du nombre de signaux possibles.

La figure 5.1 représente les données relatives à chacun des cinq sujets, ainsi que les données de groupe que constituent les moyennes des données individuelles pour chaque condition. Elle fait apparaître la droite ajustée par la méthode des moindres carrés lorsque l'on porte en abscisse le logarithme binaire du nombre de voyants, c'est-à-dire les valeurs 0, 1, 2, 3.

Pour mettre à l'épreuve la première hypothèse, nous analyserons ici la comparaison 𝕃in C (composante *linéaire* de la régression polynomiale). C'est la comparaison -𝕃in C (*résiduelle*) qui permettra de mettre à l'épreuve la deuxième hypothèse. Cette comparaison à deux degrés de liberté se décompose *orthogonalement* comme la somme des deux sous-comparaisons 𝕢ua C (composante *quadratique*) et 𝕔ub C (composante *cubique*); ce sont ces deux sous-comparaisons qui seront analysées ici.

Le tableau suivant fournit le tableau d'analyse de variance univariée usuel, supposant la condition de *circularité* pour les comparaisons à plusieurs degrés de liberté 𝕔 et -𝕃in C (les solutions inférentielles pour ces deux dernières comparaisons ne seront pas examinées ici).

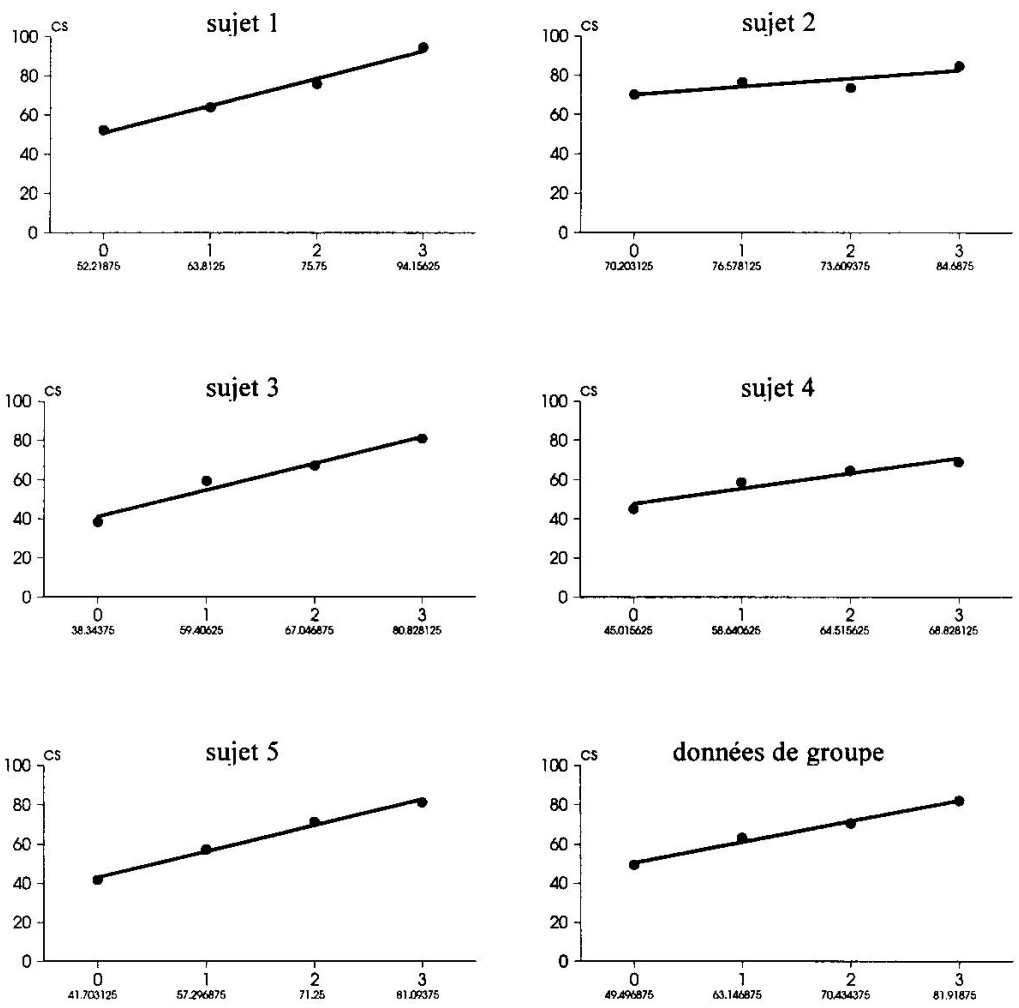
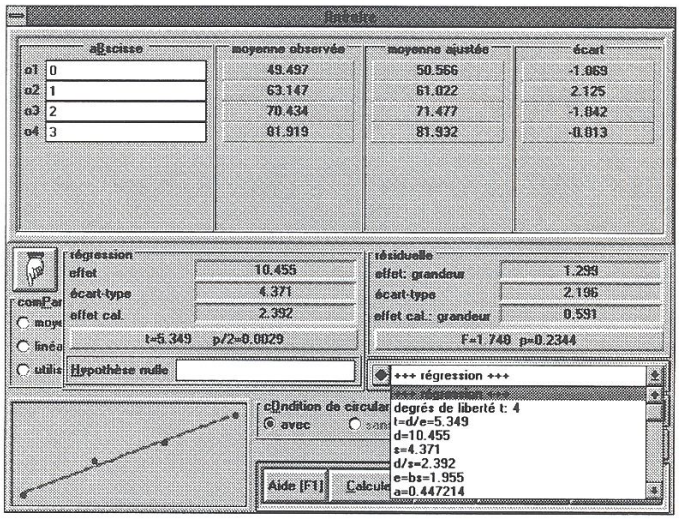
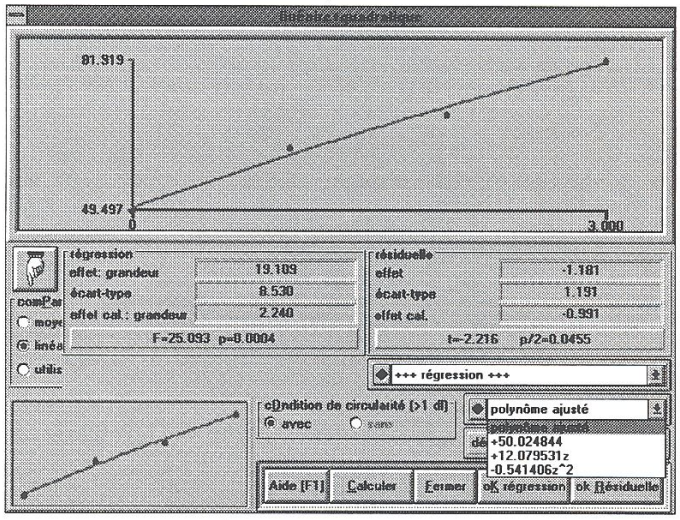


Figure 5.1 - Données individuelles et données de groupe pour l'expérience de temps de réaction avec en abscisse le logarithme binaire du nombre de signaux.


aBscisse
moyenne observée
moyenne ajustée
écart
o1 0
o2 1
o3 2
o4 3
49.497
63.147
70.434
81.919
50.566
61.022
71.477
81.932
-1.069
2.125
-1.042
-0.013
régression
effet 10.455
écart-type 4.371
effet cal. 2.392
t=5.349 p/2=0.0029
résiduelle
effet: grandeur 1.299
écart-type 2.196
effet cal.: grandeur 0.591
F=1.749 p=0.2344
Hypothèse nulle
+++ régression +++
degrés de liberté t: 4
t=d/e=5.349
d=10.455
s=4.371
d/s=2.392
e=bs=1.955
a=0.447214
cOndition de circular
avec
Aide [F1]
Calcule


81.919
49.497
0
3.000
régression
effet: grandeur 19.109
écart-type 8.530
effet cal.: grandeur 2.240
F=25.093 p=0.0004
résiduelle
effet -1.181
écart-type 1.191
effet cal. -0.991
t=-2.216 p/2=0.0455
+++ régression +++
cOndition de circularité (>1 dl)
avec
polynôme ajusté
+50.024844
+12.079531z
-0.541406z^2
Aide [F1]
Calculer
Fermer
ok régression
ok résiduelle

comparaison	*dl*	*sc*	*cm*	*F* [dl]	*p*
C	3	2766.58	922.19	24.10 [3,12]	0.010
S.C	12	459.27	38.27		
Lin C	1	2732.84	2732.84	28.61 [1,4]	0.006
S.Lin C	4	382.09	95.52		
Qua C	1	5.86	5.86	0.43 [1,4]	0.43
S.Qua C	4	54.47	13.62		
Cub C	1	27.88	27.88	4.91 [1,4]	0.09
S.Cub C	4	22.71	5.68		
-Lin C	2	33.74	16.87	1.75 [2,8]	0.23
S.-Lin C	8	77.18	9.65		

le symbole "." désigne une comparaison d'*interaction*

5.1.2 Comparaison Lin C

LeBayésien

Dans la fenêtre principale, cliquez sur le bouton **exemples**, et sélectionnez **temps de réaction**. Les données sont affichées dans l'éditeur de données.

Cliquez sur le bouton **calculer**. Les statistiques sont calculées.

En cliquant sur la zone de texte **covariances** (avec ici la valeur **160.34...**), vous faîtes apparaître une fenêtre, avec les variances et covariances associées aux quatre conditions. En cliquant sur l'option **corrélations**, vous obtenez les coefficients de corrélation linéaire entre ces conditions. Fermez cette fenêtre, pour revenir à la fenêtre des statistiques.

Cliquez maintenant sur le bouton **régression**, ce qui fait apparaître une nouvelle fenêtre. Entrez les **abscisses** [0 1 2 3] égales ici au logarithmes binaires des nombres de signaux (respectivement 1, 2, 4 et 8). Ceci peut être obtenu directement en cliquant sur le bouton **0,1,2...**

Dans le cadre **comparaison**, vérifiez que l'option **linéaire** est sélectionnée, et cliquez sur le bouton **calculer**.

	abscisse	moyenne observée	moyenne ajustée	écart
o1	0	49.147	50.566	-1.069
o2	1	63.147	61.022	2.125
o3	2	70.434	71.477	-1.042
o4	3	81.919	81.932	-0.013

La droite de régression (ajustée par la méthode des moindres carrés) est représentée; cliquez sur la figure pour faire apparaître une figure plus grande (pour la faire disparaître, cliquez à nouveau sur l'une des deux figures).

La liste déroulante **polynôme ajusté** fournit l'équation de cette droite de régression en fonction de l'abscisse z

+50.566250 +10.455313z

d'où les "moyennes ajustées" 50.566 (pour z=0), 61.022 (pour z=1), etc.

Vous obtenez l'analyse de la comparaison 𝕃in C (**régression**) et celle de la comparaison -𝕃in C (**résiduelle**). L'effet de la comparaison 𝕃in C est ici la pente de la droite de régression (+10.455), et le test *t* correspondant vaut +5.349.

La liste déroulante **+++ régression +++** fournit toutes les quantités qui interviennent dans ces analyses.

Cliquez sur le bouton **ok régression**, ce qui vous ramène dans la fenêtre principale, avec les distributions fiducio-bayésiennes relatives l'effet δ (la pente *vraie*), à l'écart-type σ, etc.

Nous pouvons prendre comme indicateur de l'effet associé à l'augmentation du nombre de voyants la pente de la droite de régression. Cette pente est donnée par le contraste *linéaire*

$$\mathbf{w}_0 = [-3/10 \quad -1/10 \quad +1/10 \quad +3/10]$$

$$\text{d'où } a^2 = (-3/10)^2+(-1/10)^2+(1/10)^2+(3/10)^2 = 1/5$$

qui définit la dérivation sur les occasions, et l'analyse spécifique est alors effectuée selon la procédure générale exposée au chapitre ➧4.

Ici il n'y a qu'un groupe, et la dérivation sur les groupes est donc trivialement définie par

$$\mathbf{v}_G = [1] \quad \text{d'où } b^2 = 1^2/5 = 1/5$$

Nous calculons donc la pente pour chaque sujet, et nous retenons ces effets individuels comme données pertinentes pour l'analyse

sujet 1	+13.7750
sujet 2	+ 4.0484
sujet 3	+13.5094
sujet 4	+ 7.7312
sujet 5	+13.2125
moyenne	+10.4553
é ty-cor	4.3709

cs

Il apparaît que l'effet observé est important pour chacun des cinq sujets. L'effet moyen d=+10.46 représente une augmentation de 21% par rapport au temps de réaction de base moyen 49.50 *cs*, observé pour la situation à un voyant (donc sans incertitude).

Nous retrouvons ici le rapport F=28.61 sous la forme du carré de la statistique t de Student du test de l'hypothèse nulle d'une "pente vraie" δ égale à 0, sous le *modèle normal* usuel, de paramètres δ et σ, pour les données pertinentes

$$t = d/e = +5.35 \quad [q=4\ dl] \quad \text{pour } e = bs = 4.3709/\sqrt{5} = 1.955$$

Le résultat est significatif (au seuil unilatéral $p/2$=0.006).

Les carrés-moyens s'obtiennent comme

$$cm_{\text{LIN C}} = d^2/a^2b^2 = 25d^2 = 2732.84 \quad \text{et} \quad cm_{\text{S.LIN C}} = s^2/a^2 = 5s^2 = 95.52$$

■ *Inférence fiducio-bayésienne*

Les distributions fiducio-bayésiennes relatives à δ et σ sont

$$\delta \sim t_q(d,e^2) \sim t_4(+10.455,1.955^2)$$

$$\sigma \sim s\varphi_q^{-1} \sim 4.371\varphi_4^{-1}$$

d'où les limites pour la garantie 0.90

$$Pr(\delta>+7.46) = 0.90 \quad \text{et} \quad Pr(2.84<\sigma<10.37) = 0.90$$

Nous pouvons donc conclure que le temps de réaction moyen augmente d'une manière notable lorsque le nombre de voyants croît.

5.1.3 Comparaisons ℂub C et ℚua C

Nous pouvons ajuster, toujours par la méthode des moindres carrés un polynôme de degré 2, associé aux composantes *linéaire* et *quadratique*. Ces deux composantes engendrent une comparaison à deux degrés de liberté, dont la résiduelle est la comparaison ℂub C.

De la même manière nous pouvons ajuster un polynôme de degré 3 associé aux composantes *linéaire* et *cubique*. Ces deux composantes engendrent une comparaison à deux degrés de liberté, dont la résiduelle est la comparaison ℚua C.

Nous pouvons analyser séparément ces deux comparaisons résiduelles ℂub C et ℚua C, en procédant de la même manière que pour la comparaison 𝕃in C, avec les contrastes respectifs

ℂub C $\quad \mathbf{w}_0 = [+0.1118 \ -0.3354 \ +0.3354 \ -0.1118]$

ℚua C $\quad \mathbf{w}_0 = [-0.2500 \ +0.2500 \ +0.2500 \ -0.2500]$

Ces contrastes sont ici choisis de telle manière que leurs effets observés, respectivement d=-1.181 et d=+0.541, correspondent à l'écart entre les moyennes observées et les moyennes ajustées, pour chacun des deux polynômes précédents. Il s'agit de montrer que chacun des deux effets *vrais* correspondants est négligeable, ou du moins limité.

LeBayésien

Dans l'exemple précédent, après avoir analysé la comparaison 𝕃in C, cliquez sur le bouton **G / 0**, ce qui vous ramène à la fenêtre de la régression polynomiale.

Dans le cadre **comparaison**, sélectionnez maintenant, dans la deuxième ligne, l'option **linéaire+quadratique**, et cliquez sur le bouton **calculer**.

	abscisse	moyenne observée	moyenne ajustée	écart
o1	0	49.147	50.025	-0.528
o2	1	63.147	61.563	1.584
o3	2	70.434	72.018	-1.584
o4	3	81.919	81.391	0.528

Le polynôme de degré 2 correspondant est représenté; son équation est fournie dans la liste déroulante **polynôme ajusté**

$$+50.0248 + 12.0795z - 0.5414z^2$$

d'où les "moyennes ajustées" 50.025 (pour z=0), 61.563 (pour z=1), etc.

Ici nous nous intéressons aux écarts (*résidus*) entre les moyennes observées et les moyennes ajustées. En fait ici, la comparaison résiduelle (représentée par le contraste *cubique*) a un seul degré de liberté, et ces écarts (dont la somme est nulle) sont proportionnels (1.584=3×0.528). L'effet de cette comparaison résiduelle (d=-1.181) est tel que sa valeur absolue est la moyenne quadratique des quatre écarts

$$((0.528^2 + 1.584^2 + 1.584^2 + 0.528^2)/4)^{1/2} = 1.181$$

Il s'agit ici de montrer que le paramètre correspondant δ a une valeur négligeable, ou du moins limité. Cliquez sur le bouton **ok résiduelle**, ce qui vous ramène dans la fenêtre principale, avec les distributions fiducio-bayésiennes relatives l'effet δ (la pente *vraie*), à l'écart-type σ, etc.

Vous obtenez par exemple l'énoncé: $Pr(|\delta|<2.01) = 0.90$

Revenez encore à la la fenêtre de la régression polynomiale, et dans le cadre **comparaison**, sélectionnez maintenant, dans la troisième ligne, l'option **utiliser la liste des composantes**, et cochez les cases 1 et 3. Puis cliquez sur le bouton **calculer**.

	abscisse	moyenne observée	moyenne ajustée	écart
o1	0	49.147	50.038	-0.541
o2	1	63.147	62.605	0.541
o3	2	70.434	69.893	0.541
o4	3	81.919	82.460	-0.541

Le polynôme de degré 3 correspondant est représenté; son équation est fournie dans la liste déroulante **polynôme ajusté**

$$+50.0383 +18.7268z -7.9195z^2 +1.7599z^3$$

d'où les "moyennes ajustées" 50.038 (pour z=0), 62.605 (pour z=1), etc.

La comparaison résiduelle (représentée par le contraste *quadratique*) a un seul degré de liberté, et chacun des quatre écarts est ici égal (en valeur absolue) à son effet (d=+0.541).

Il s'agit ici encore de montrer que le paramètre correspondant δ a une valeur négligeable, ou du moins limité. Cliquez sur le bouton **ok résiduelle**, etc.

Vous obtenez par exemple l'énoncé: $Pr(|\delta|<1.96) = 0.90$

□ *Remarques*

• Si vous analysez directement les composantes *quadratique* et *cubique*, vous obtenez les effets observés respectifs

$$d = -0.5414 \quad \text{et} \quad d = +1.7599$$

qui sont respectivement le coefficients du terme de degré 2 du polynôme ajusté avec les composantes *linéaire* et *cubique*, et le coefficients du terme de degré 3 du polynôme ajusté avec les composantes *linéaire* et *quadratique*.

Les contrastes correspondants sont proportionnels à ceux utilisés pour les résiduelles, lesquels donnent ici une interprétation plus directe vis-à-vis de la question examinée (écart au modèle *logarithmique*).

• Si vous sélectionnez les trois composantes, *linéaire*, *quadratique*, et *cubique*, vous obtenez un polynôme de degré 3, qui s'ajuste "parfaitement" aux moyennes observées.

■ *Inférence fiducio-bayésienne*

Les distributions fiducio-bayésiennes relatives à δ sont respectivement

$$\text{Cub C} \quad \delta \sim t_4(-1.181, 0.533^2)$$

$$\text{Qua C} \quad \delta \sim t_4(+0.541, 0.825^2)$$

d'où les limites pour la garantie 0.90

$$\text{Cub C} \quad Pr(|\delta| < 2.01) = 0.90$$

$$\text{Qua C} \quad Pr(|\delta| < 1.96) = 0.90$$

■ *Comment juger l'acceptabilité du modèle logarithmique?*

Un bon moyen est de procéder à une comparaison avec d'autres modèles. A titre d'exemple considérons le modèle *linéaire* selon lequel le temps de réaction serait une fonction linéaire du nombre de signaux possibles. Cela ne veut pas dire que nous considérons ce modèle comme un concurrent réaliste du modèle logarithmique. Mais, précisément parce qu'il fournit un ajustement médiocre au niveau descriptif, il nous procure un critère de non-acceptabilité. En procédant de la même manière que pour le modèle logarithmique, nous trouvons respectivement les effets observés

Cub C (résiduelle des composantes *linéaire* et *quadratique*) d = -2.185

Qua C (résiduelle des composantes *linéaire* et *cubique*) d = +3.182

Les limites obtenues pour le modèle *logarithmique* sont inférieures à ces effets, et peuvent donc être jugées relativement acceptables; ici notre exigence doit être modeste en raison du faible nombre de sujets...

LeBayésien

Dans l'exemple précédent, effectuez les mêmes analyses avec les abscisses 1, 2, 4 et 8 pour la régression polynomiale.

■ *Remarque sur l'analyse de la comparaison* -Lin C

L'analyse de la comparaison à deux degrés de liberté -Lin C permettraient une inférence simultanée sur les effets des deux sous-comparaisons précédentes. Les données pertinentes pour l'analyse seraient, pour chaque sujet les *deux* effets individuels associés à chacune des deux comparaisons Qua C et Cub C. Pour se ramener à un problème *univarié*, il faudrait supposer que les deux variables correspondantes sont non corrélées, ce qui pourra apparaître peu réaliste. On voit ici les difficultés particulières liées à l'analyse des comparaisons à plusieurs degrés de liberté sur les *occasions*.

5.2 Plans non-équilibrés

Les plans non équilibrés, avec les choix qu'ils posent, sont souvent une source de difficultés pour l'utilisateur. La plupart des logiciels d'analyse de variance donnent en effet plusieurs solutions pour ces plans. Nous chercherons ici à clarifier ce problème, en l'illustrant à partir de l'exemple d'un plan en groupes indépendants à deux facteurs de classification; cet exemple sera facilement généralisé à des plans plus complexes.

Reprenons l'enquête psychométrique sur l'enseignement des mathématiques déjà introduite en §2.2. Rappelons que les éléves étaient répartis en quatre groupes (d'effectifs voisins) définis par le croisement de deux facteurs à deux niveaux chacun

- Facteur *Milieu*: *favorisé* opposé à *défavorisé*
- Facteur *Pédagogie*: *moderne* opposé à *traditionnelle*

Pour les données analysées précédemment, les effectifs étaient élevés et très voisins. A titre didactique, nous considérerons ici des données fictives relatives à seulement 33 élèves, en nous limitant à une seule variable: la note à une épreuve de calcul combinatoire passée en fin d'année scolaire. Ces données sont fournies ci-après.

g1=m1p1	3.9	6.3	5.6	4.3	7.3	5.5	6.4				f_{g1}= 7
g2=m2p1	5.5	3.7	5.5	3.2	4.0	5.0	4.7	5.8	4.4	6.6	f_{g2}=10
g3=m1p2	4.7	6.1	2.7	5.4	3.0	5.8	2.7				f_{g3}= 7
g4=m2p2	3.7	6.2	4.0	2.5	5.7	2.5	2.3	3.7	2.9		f_{g4}= 9

Nous avons les moyennes des quatre groupes de sujets, et les moyennes marginales correspondantes suivant la dérivation par moyennage adoptée: équipondérée, ou pondérée par les effectifs.

équipondération

	m1	m2	
p1	5.6143	4.8400	5.2271
p2	4.3429	3.7222	4.0325
	4.9786	4.2811	

pondération par les effectifs

	m1	m2	
p1	5.6143 (7)	4.8400 (10)	5.1588
p2	4.3429 (7)	3.7222 (9)	3.9937
	4.9786	4.3105	

Considérons à titre d'exemple la comparaison ℙ, ou "p1,p2", qui est représentée par le contraste [+1 -1] sur P.

5.2.1 Solution équipondérée

La solution équipondérée, que nous indiquerons par la mention [E], correspond au contraste sur les groupes

$p1,p2_{[E]}$

	m1	m2	
p1	+1/2	+1/2	+1
p2	-1/2	-1/2	-1
	0	0	

(les marges sont les sommes des coefficients)

dont l'effet observé

$$d = \tfrac{1}{2}\,5.6143 + \tfrac{1}{2}\,4.8400 - \tfrac{1}{2}\,4.3429 - \tfrac{1}{2}\,3.7222 = 5.2271\text{-}4.0325 = +1.1946$$

est la moyenne (équipondérée) des différences respectives à l'intérieur de m1 (d_1=5.6143-4.3429=+1.2714) et à l'intérieur de m2 (d_2=4.8400-3.7222=+1.1178)

$$d = (1.2714+1.1178)/2 = +1.1946$$

5.2.2 Solution pondérée par les effectifs

La solution pondérée par les effectifs, que nous indiquerons par la mention [N], correspond au contraste sur les groupes

$p1,p2_{[N]}$

	m1	m2	
p1	+7/17	+10/17	+1
p2	-7/16	-9/16	-1

dont l'effet observé est

$$d = \frac{7}{17}\,5.6143 + \frac{10}{17}\,4.8400 - \frac{7}{16}\,4.3429 - \frac{9}{16}\,3.7222 = 5.1588\text{-}3.9937 = +1.1651$$

Mais cette différence n'est pas la moyenne des différences respectives à l'intérieur de m1 (d_1=+1.2714) et à l'intérieur de m2 (d_2=+1.1178). Par exemple, si ces deux différences étaient nulles, d ne le serait pas; ces deux différences pourraient même être toutes les deux inférieures ou supérieures à d. Ceci est manifestement un désavantage de la solution pondérée par les effectifs.

5.2.3 Solution orthogonale

Les solutions précédentes ne conduisent pas à une décomposition *orthogonale*. En analyse de variance, cela se traduit par la non-additivité des sommes-de-carrés: la somme-des-carrés associée à la comparaison globale (à trois degrés de liberté) des quatre groupes n'est pas la somme des sommes-des-carrés partielles associées à chacune des comparaisons (à un degré de liberté) $\mathbb{M}$, $\mathbb{P}$ et $\mathbb{M}.\mathbb{P}$.

D'une manière générale la condition d'orthogonalité de deux contrastes $[u_g]_{g\in G}$ et $[v_g]_{g\in G}$ sur G s'écrit (pour des effectifs $[f_g]_{g\in G}$)

$$\sum_{g\in G} u_g v_g / f_g = 0$$

Ainsi, pour la solution équipondérée, la comparaison $\mathbb{P}_{[E]}$ n'est pas orthogonale à la comparaison d'interaction $\mathbb{M}.\mathbb{P}$ (dont l'effet est la "différence des différences de moyennes"), laquelle est définie de manière canonique, c'est-à-dire indépendamment de la dérivation par moyennage.

$\mathbb{P}_{[E]}$

	m1	m2	
p1	+1/2	+1/2	+1
p2	-1/2	-1/2	-1
	0	0	

$\mathbb{M}.\mathbb{P}$

	m1	m2	
p1	+1	-1	0
p2	-1	+1	0
	0	0	

$$\sum_{g\in G}(u_g v_g / f_g) = \frac{+1/2}{7} + \frac{-1/2}{10} + \frac{+1/2}{7} + \frac{-1/2}{9} = \frac{47}{630} \neq 0$$

Pour la solution pondérée par les effectifs, les comparaisons $\mathbb{P}_{[N]}$ et $\mathbb{M}_{[N]}$ ne sont pas orthogonales, mais chacune d'elle est en revanche orthogonale à la comparaison d'interaction $\mathbb{P}.\mathbb{M}$. Enfin nous avons l'orthogonalité des comparaisons $\mathbb{M}$ et $\mathbb{P}$ si nous prenons pour l'une la solution équipondérée et pour l'autre la solution pondérée par les effectifs.

■ *Remarque: Plans orthogonaux*

On remarquera que les comparaisons $\mathbb{P}_{[N]}$ et $\mathbb{M}_{[N]}$ sont orthogonales si les effectifs vérifient une condition particulière, qui s'écrit

$$f_{pm} = f_p f_m / f_{total}$$

où f_{pm} est l'effectif associé au coupe (p,m), f_p et f_m sont les effectifs marginaux associés respectivement à p et m, et f_{total} est l'effectif total.

Dans le cas particulier de dux facteurs à deux modalités, il suffit de vérifier que (par exemple)

$$\frac{f_{p1m1}}{f_{p2m1}} = \frac{f_{p1}}{f_{p2}}$$

Les plans dont les effectifs vérifient cette condition sont appelés *plans orthogonaux*. Ils règlent le problème de l'orthogonalité de la décomposition, à condition toutefois d'adopter la solution pondérée par les effectifs.

■ *Recherche d'une décomposition orthogonale*

Maintenant, si nous recherchons une décomposition orthogonale, nous pouvons prendre pour la comparaison $\mathbb{M}$ la solution pondérée par les effectifs, puisque nous avons l'orthogonalité de $\mathbb{M}_{[N]}$ et $\mathbb{M}.\mathbb{P}$, soit le contraste

$m1,m2_{[N]}$

	m1	m2	
p1	+1/2	-10/19	
p2	+1/2	-9/19	
	+1	-1	

Mais nous devons alors adopter une autre procédure de dérivation par moyennage sur M, de manière à ce que la comparaison $\mathbb{P}$ soit orthogonale à la fois à $\mathbb{M}_{[N]}$ et $\mathbb{M}.\mathbb{P}$.

Nous cherchons par conséquent un contraste sur P×M tel que:
(1) la somme de ses coefficients est égale à +1 à l'intérieur de p1 et à -1 à l'intérieur de p2 (nous retrouvons [+1 -1] comme contraste marginal);
(2) la somme de ses coefficients à l'intérieur de chaque m est nulle (on dira que le contraste est "*intra*-M"), ce qui assure l'orthogonalité avec la comparaison $\mathbb{M}$;
(3) il est orthogonal au contraste d'interaction représentant la comparaison $\mathbb{M}.\mathbb{P}$.

Ce contraste est donc de la forme

	m1	m2	
p1	c	1-c	+1
p2	-c	c-1	-1
	0	0	

et la condition d'orthogonalité avec le contraste d'interaction s'écrit ici

$$\frac{c}{7} + \frac{c-1}{10} + \frac{c}{7} + \frac{c-1}{9} = 0$$

Nous en déduisons la valeur $c = 133/313$, d'où le contraste pour cette solution "orthogonale" que nous indiquerons par la mention [0]

$p1,p2_{[0]}$

	m1	m2	
p1	+133/313	+180/313	+1
p2	-133/313	-180/313	-1
	0	0	

d'effet observé

$$d = \frac{133}{313}5.6143 + \frac{180}{313}4.8400 - \frac{133}{313}4.3429 - \frac{180}{313}3.7222 = +1.1831$$

Nous avons ainsi décomposé le contraste représentant la comparaison $\mathbb{P}_{[E]}$, comme la somme

	m1	m2	
p1	+1/2	+1/2	+1
p2	-1/2	-1/2	-1
	0	0	

=

	m1	m2	
p1	+133/313	+180/313	+1
p2	-133/313	-180/313	-1
	0	0	

+

	m1	m2	
p1	+47/626	-47/626	0
p2	-47/626	+47/626	0
	0	0	

de deux contrastes orthogonaux, le second étant un contraste d'interaction de la comparaison $\mathbb{M}.\mathbb{P}$.

Remarquons encore que cette solution correspond à la dérivation par moyennage sur M définie par les poids

	m1	m2
p1	7	180/19
p2	7	180/19

où 180/19=9.4737 est la moyenne *harmonique* des deux effectifs 10 et 9 associés à la modalité m2

$$\frac{1}{180/19} = \frac{(1/10)+(1/9)}{2}$$

Cette dérivation fournit les moyennes marginales

	m1	m2	
p1	5.6143 (7)	4.8400 (180/19)	5.1690
p2	4.3429 (7)	3.7222 (180/19)	3.9859
	4.9786	4.2811	

et la différence entre les deux moyennes d=5.1690-3.9859=+1.1831 est la moyenne pondérée des différences à l'intérieur de m1 et à l'intérieur de m2

$$+1.1831 = (133\times1.2714+180\times1.1178)/313$$

LeBayésien

Sélectionnez l'exemple **milieu et pédagogie - données de 33 élèves**. Effectuez les calculs pour les différentes combinaisons linéaires définies ci-dessus. Notez la correspondance:

g1=m1p2, g2=m2p1, g3=m1p2, g4=m2p2

Ceci permet notamment de reconstituer le tableau d'analyse de variance de la solution orthogonale.

Effectuez également le calcul pour l'option **comparaison globale**, de manière à obtenir la somme-des-carrés intergroupe totale (les procédures inférentielles pour cette comparaison sortent du cadre de cet ouvrage).

Nous avons pour cette solution *orthogonale* le tableau d'analyse de variance

source de variation	*dl*	*sc*	*cm*	*F* [dl]	*p*
Milieu $\mathbb{M}_{[N]}$	1	3.597	3.597	2.178 [1,29]	0.151
Pédagogie $\mathbb{P}_{[O]}$	1	11.529	11.529	6.979 [1,29]	0.013
Interaction $\mathbb{M}.\mathbb{P}$	1	0.048	0.048	0.029 [1,29]	0.866
Groupes $\mathbb{G}$	3	15.174			
intragroupe $\mathbb{S}(G)$	29	47.905	1.652		

qui se déduit des effets des différents contrastes et des constantes b correspondantes (le carré-moyen étant égal à $(d/b)^2$)

		d	b
[1/2 -10/19 1/2 -9/19]	$\mathbb{M}_{[N]}$	+0.6680	0.352222
[133/313 180/313 -133/313 -180/313]	$\mathbb{P}_{[O]}$	+1.1831	0.348433
[1 -1 -1 1]	$\mathbb{M}.\mathbb{P}$	+0.1537	0.704858

■ *Inconvénients de la solution orthogonale*

Cette solution vérifie par construction la propriété d'orthogonalité des comparaisons $\mathbb{M}_{[N]}$, $\mathbb{P}_{[O]}$ et $\mathbb{M}.\mathbb{P}$; mais elle présente (au moins si l'on s'en tient à la justification d'orthogonalité) deux inconvénients majeurs.

(1) Elle revient à adopter deux procédures différentes pour les comparaisons $\mathbb{M}$ et $\mathbb{P}$. En outre le choix pour l'affectation de ces procédures est arbitraire: nous pouvons aussi bien ici prendre la solution pondérée par les effectifs pour $\mathbb{P}$ et la solution orthogonale pour $\mathbb{M}$, ce qui conduit pour cette dernière comparaison au contraste

$m1,m2_{[O]}$

	m1	m2	
p1	+160/313	-160/313	0
p2	+153/313	-153/313	0
	+1	-1	

d'effet observé

$$d = \frac{160}{313}5.6143 - \frac{160}{313}4.8400 + \frac{153}{313}4.3429 - \frac{153}{313}3.7222 = +0.6992$$

(2) Quand on la généralise à des facteurs à plus de deux modalités, il apparaît que cette procédure dépend du contraste analysé. Ainsi, par exemple si P comporte trois modalités, la moyenne associée à la modalité p1 sera différente, selon que l'on considère le contraste [+1 -1 0] ou le contraste [+1 0 -1]. Par suite il n'apparaît pas possible de définir, comme pour les autres solutions, la comparaison de façon intrinsèque sur P.

5.2.4 Autre point de vue: Modèle sans interaction

Un autre point de vue que l'orthogonalité de la décomposition peut être adopté pour justifier la solution orthogonale: supposer que l'effet théorique de la comparaison d'interaction $\mathbb{M}.\mathbb{P}$ est nul ("modèle sans interaction").

Nous pouvons alors construire des moyennes associées à chacun des quatre groupes telles que l'effet d'interaction observé soit nul (il y a "additivité" des effets des facteurs M et P). Nous donnerons ici le résultat sans justification (laquelle nécessite un détour par la formalisation algébrique). Ces moyennes, auxquelles nous pouvons appliquer les dérivations par moyennage équipondéré et pondéré par les effectifs, sont les suivantes

équipondération

	m1	m2	
p1	5.5701	4.8709	5.2205
p2	4.3870	3.6879	4.0374
	4.9786	4.2794	

pondération par les effectifs

	m1	m2	
p1	5.5701 (7)	4.8709 (10)	5.1588
p2	4.3870 (7)	3.6879 (9)	3.9937
	4.9786	4.3105	

On peut vérifier que l'effet d'interaction ("différence des différences") est nul

$$5.5701 - 4.8709 - 4.3870 + 3.6879 = 0$$

Pour le moyennage pondéré par les effectifs, nous retrouvons les mêmes moyennes marginales que pour le protocole de base, d'où les mêmes effets des comparaisons $\mathbb{M}_{[N]}$ et $\mathbb{P}_{[N]}$. Ceci est normal, puisque ces deux comparaisons sont dans ce cas orthogonales à $\mathbb{M}.\mathbb{P}$.

Pour le moyennage équipondéré, nous obtenons les effets respectifs des comparaisons $\mathbb{M}_{[E]}$ et $\mathbb{P}_{[E]}$

$$\mathbb{M}_{[E]} \rightarrow d = 4.9786-4.2794 = +0.6992$$

$$\mathbb{P}_{[E]} \rightarrow d = 5.2205-4.0374 = +1.1831$$

qui ne sont autres, par construction, que les effets obtenus précédemment dans la solution "orthogonale".

L'argument utilisé ici (hypothèse d'une interaction nulle) nous paraît méthodologiquement plus convaincant que l'argument d'orthogonalité de la décomposition, à condition toutefois que cette hypothèse puisse réellement être prise au sérieux: on ne saurait se contenter, comme cela se pratique trop souvent, d'avancer la justification d'un test *non-significatif*!

5.3 Exemples de plans "cross-over"

5.3.1 Cross-over 2×2

Dans la situation où nous voulons comparer deux traitements t_1 et t_2, les solutions les plus simples consistent:

- soit à administrer ces traitements à deux groupes de sujets différents, d'où un plan $S<G_2>=S<T_2>$;
- soit (si cela est possible) à administrer les deux traitements à chaque sujet, d'où un plan $S\times T_2$.

Cette dernière solution est généralement recommandée: elle conduit en principe à une meilleure précision expérimentale et permet des analyses intra-sujets. Mais elle pose évidemment le problème de l'ordre d'administration des traitements. La procédure la plus courante est le plan "*cross-over*" à deux périodes, dans lequel chacun des traitements est administré une fois à chaque sujet, dans l'ordre t_1 puis t_2 à une partie des sujets (en général la moitié) et dans l'ordre t_2 puis t_1 à l'autre partie. Un exemple typique d'un tel plan a été donné en ↳1.2. avec les données pharmacologiques concernant la comparaison d'un nouveau médicament (t_2) et d'un placebo (t_1). Un tel plan met en jeu trois facteurs systématiques formant un *carré-latin*:

T_2={t1,t2} – *Traitements* (ou *Conditions expérimentales...*)
P_2={p1,p2} – *Périodes* (ou *Rangs* ou *Essais...*)
G_2={s1,s2} – *Groupes* (ou *Séquences* ou *Ordres* d'administration)

■ *Deux représentations*

Nous avons deux représentations possibles d'un tel plan comme un plan S<G>×O, selon que nous considérons comme Occasions les *Traitements* ou les *Périodes*.

	o1 = t1	o2 = t2
g1	p1	p2
g2	p2	p1

	o1 = p1	o2 = p2
g1	t1	t2
g2	t2	t1

Nous adopterons par exemple la représentation comme un plan S<G_2>×P_2. Les effets des différentes comparaisons (à un degré de liberté) sont fournis par les contrastes suivants sur G×P.

■ *Solutions équipondérée et pondérée par les effectifs*

• Solution équipondérée (indiquée par la mention [E])

	p1	p2
g1	+1/2	+1/2
g2	-1/2	-1/2

$\mathbb{G} \equiv \mathbb{P}.\mathbb{T}$

	p1	p2
g1	+1/2	-1/2
g2	+1/2	-1/2

$\mathbb{P}_{[E]} \equiv \mathbb{G}.\mathbb{T}$

	p1	p2
g1	+1/2	-1/2
g2	-1/2	+1/2

$\mathbb{T}_{[E]} \equiv \mathbb{G}.\mathbb{P}$

Dans ce cas chaque comparaison associée à l'un des trois facteurs est confondue avec la comparaison d'interaction associée aux deux autres facteurs (laquelle est définie de manière canonique, c'est-à-dire indépendamment de la dérivation par moyennage): $\mathbb{P}_{[E]} \equiv \mathbb{G}.\mathbb{T}$ et $\mathbb{T}_{[E]} \equiv \mathbb{G}.\mathbb{T}$. Par exemple l'effet de la comparaison "t1,t2" est la moitié de l'effet de la comparaison d'interaction "g1,g2.p1,p2", défini comme la différence des différences.

$\mathbb{G}$ est *orthogonale* (voir ↳5.2.3) à $\mathbb{P}_{[E]}$ et $\mathbb{T}_{[E]}$, mais $\mathbb{P}_{[E]}$ et $\mathbb{T}_{[E]}$ ne sont orthogonales que si les effectifs des deux groupes f_{g1} et f_{g2} sont identiques.

• Solution pondérée par les effectifs (indiquée par la mention [N])

Si les effectifs des deux groupes sont différents, nous pouvons également envisager la solution "pondérée par les effectifs".

exemple: $f_{g1}=5$ et $f_{g2}=3$

	p1	p2
g1	+1/2	+1/2
g2	-1/2	-1/2

$\mathbb{G}\equiv\mathbb{P}.\mathbb{T}$

	p1	p2
g1	+5/8	-5/8
g2	+3/8	-3/8

$\mathbb{P}_{[N]}$

	p1	p2
g1	+5/8	-5/8
g2	-3/8	+3/8

$\mathbb{T}_{[N]}$

Dans ce cas nous avons seulement $\mathbb{P}_{[N]}\oplus\mathbb{T}_{[N]}\equiv\mathbb{G}.\mathbb{T}\oplus\mathbb{G}.\mathbb{P}$. Par exemple le contraste d'interaction représentant la comparaison $\mathbb{G}.\mathbb{P}$ est obtenu comme la combinaison linéaire

+1	-1	
-1	+1	

$\mathbb{G}.\mathbb{P}$

$= -(8/15)$

+5/8	-5/8
+3/8	-3/8

$\mathbb{P}_{[N]}$

$+ (32/15)$

+5/8	-5/8
-3/8	+3/8

$\mathbb{T}_{[N]}$

$\mathbb{G}$ est orthogonale à $\mathbb{P}_{[N]}$ et $\mathbb{T}_{[N]}$, mais $\mathbb{P}_{[N]}$ et $\mathbb{T}_{[N]}$ ne sont pas orthogonales si les effectifs des deux groupes sont différents.

• Si l'on veut une décomposition orthogonale, il faut prendre la solution équipondérée pour l'une des comparaisons $\mathbb{P}$ ou $\mathbb{T}$ et la solution pondérée par les effectifs pour l'autre. Par exemple

	p1	p2
g1	+1/2	+1/2
g2	-1/2	-1/2

$\mathbb{G}\equiv\mathbb{P}.\mathbb{T}$

	p1	p2
g1	+5/8	-5/8
g2	+3/8	-3/8

$\mathbb{P}_{[N]}$

	p1	p2
g1	+1/2	-1/2
g2	-1/2	+1/2

$\mathbb{T}_{[E]}\equiv\mathbb{G}.\mathbb{P}$

■ *Modélisation et effet résiduel*

Pour représenter les mécanismes en jeux dans un plan *cross-over*, on pose souvent un modèle du type suivant sur les moyennes vraies: la moyenne μ^{gpt} (du groupe g, à la période p, au cours de laquelle a été administré le traitement t) est la somme des composantes ("fixées") suivantes:

- la moyenne générale μ
- une quantité τ^t dépendant du traitement t;
- une quantité π^p dépendant de la période;
- une quantité $\kappa^{t'}$ dépendant du traitement t' administré à la période précédente (qui est par conséquent nulle pour la première période).

Cette dernière quantité $\kappa^{t'}$ correspond à un effet *résiduel* (ici d'ordre 1) du traitement t', souvent appelé effet "*carry-over*". Les quantités τ^t et π^p sont souvent appelées "effet traitement" et "effet période"; on les distinguera des effets des facteurs *Traitements* et *Périodes* (voir les considérations sur l'emploi du mot effet en ↳1.2.1).

Ce modèle peut être rendu plus complexe, par exemple en supposant que les composantes τ et/ou les composantes κ dépendent de la période, ou encore dans la généralisation à des plans comportant plus de deux périodes que les composantes κ dépendent également des traitements administrés aux cours des deux (ou plus) périodes précédentes (effets résiduels d'ordre 2, 3...).

En nous en tenant au modèle précédent (*"carry-over simple"*), nous obtenons pour le plan *cross-over* 2×2

	p1	p2
g1	$\mu^{g1p1t1} = \mu+\tau^{t1}+\pi^{p1}$	$\mu^{g1p2t2} = \mu+\tau^{t2}+\pi^{p2}+\kappa^{t1}$
g2	$\mu^{g2p1t2} = \mu+\tau^{t2}+\pi^{p1}$	$\mu^{g2p2t1} = \mu+\tau^{t1}+\pi^{p2}+\kappa^{t2}$

Pour la solution équipondérée les effets vrais des trois contrastes sont par conséquent les suivants:

$$\text{"g1,g2"} \rightarrow \delta = (\kappa^{t1}-\kappa^{t2})/2$$
$$\text{"p1,p2"} \rightarrow \delta = \pi^{p1}-\pi^{p2}-(\kappa^{t1}+\kappa^{t2})/2$$
$$\text{"t1,t2"} \rightarrow \delta = \tau^{t1}-\tau^{t2}+(\kappa^{t1}-\kappa^{t2})/2$$

d'où des difficultés d'interprétation si les traitements ont des effets résiduels différents ($\kappa^{t1}\neq\kappa^{t2}$).

Un autre exemple d'un tel plan est donné en ↳7.1.3.

■ *Généralisations*

La situation de base précédente a donné lieu à toutes sortes de généralisations (voir Jones et Kenward, 1989; Senn, 1993).

On peut bien entendu introduire des facteurs supplémentaires. Ainsi, à la structure précédente, on peut ajouter des facteurs:
• inter-sujets, par exemple des facteurs de classification;
• intra-sujets, typiquement à l'intérieur de chacune des deux périodes on effectuera un certain nombre d'observations successives pour le traitement correspondant.

Un exemple d'un tel plan est fourni en ↳5.3.4.

On peut également utiliser plusieurs traitements et/ou plusieurs périodes. Des exemples sont donnés ci-après.

5.3.2 Cross-over: Exemple avec deux traitements et trois périodes

Chacun des deux traitements peut être administré plusieurs fois à chaque sujet. Par exemple, en utilisant trois périodes, on peut avoir

	o1 = p1	o2 = p2	o3 = p3
g1	t1	t2	t2
g2	t2	t1	t1

■ *Modélisation et effets résiduels*

Si nous considérons le même modèle que pour le *cross-over* 2×2 (*carry-over* simple), nous avons

	p1	p2	p3
g1	$\mu^{g1p1t1} = \mu+\tau^{t1}+\pi^{p1}$	$\mu^{g1p2t2} = \mu+\tau^{t2}+\pi^{p2}+\kappa^{t1}$	$\mu^{g1p3t2} = \mu+\tau^{t2}+\pi^{p3}+\kappa^{t2}$
g2	$\mu^{g2p1t2} = \mu+\tau^{t2}+\pi^{p1}$	$\mu^{g2p2t1} = \mu+\tau^{t1}+\pi^{p2}+\kappa^{t2}$	$\mu^{g2p3t1} = \mu+\tau^{t1}+\pi^{p3}+\kappa^{t1}$

■ *Décomposition*

Dans cette situation la comparaison des deux traitements $\mathbb{T}$, que l'on prenne la solution équipondérée ou la solution pondérée par les effectifs, n'est pas orthogonale à la comparaison des deux groupes $\mathbb{G}$. On le vérifie par exemple dans le cas équilibré

	p1	p2	p3
g1	+1/3	+1/3	+1/3
g2	-1/3	-1/3	-1/3

$\mathbb{G}$

	p1	p2	p3
g1	+1/3	-1/3	-1/3
g2	-1/3	+1/3	+1/3

$\mathbb{T}_{[E]}$

Ceci nous amène à considérer, pour l'une de ces deux comparaisons, le contraste correspondant à la solution équipondérée (indiqué par la mention [E]), et pour l'autre le contraste *orthogonal* (indiqué par la mention [O]), soit l'un des deux choix

	p1	p2	p3	
g1	+1/3	+1/3	+1/3	+1
g2	-1/3	-1/3	-1/3	-1
	1/3	1/3	1/3	

$\mathbb{G}_{[E]}$

$\delta = (\tau^{t2}-\tau^{t1})/3$

et

	p1	p2	p3	
g1	+1/2	-1/4	-1/4	+1/2
g2	-1/2	+1/4	+1/4	-1/2
	+1	-1/2	-1/2	

$\mathbb{T}_{[O]}$

$\delta = \tau^{t1}-\tau^{t2}$

ou

	p1	p2	p3	
g1	+1/2	+1/4	+1/4	+1
g2	-1/2	-1/4	-1/4	-1
	1/2	1/4	1/4	

$\mathbb{G}_{[O]}$

$\delta = 0$

et

	p1	p2	p3	
g1	+1/3	-1/3	-1/3	+1/3
g2	-1/3	+1/3	+1/3	-1/3
	+1	-1	-1	

$\mathbb{T}_{[E]}$

$\delta = \tau^{t1}-\tau^{t2}$

(ici les marges correspondent aux formes linéaires marginales v_g et w_0)

La combinaison linéaire correspondante δ des moyennes μ_{gpt}, sous le modèle précédent, est indiquée dans chaque cas. On voit ici l'intérêt du plan avec trois périodes:

- pour la comparaison $\mathbb{T}$, quelle que soit la solution retenue, δ ne dépend pas des effets résiduels κ^{t1}, et est bien la différence $\tau^{t1}-\tau^{t2}$;
- pour la comparaison $\mathbb{G}$, δ est proportionnelle à la différence $\tau_{t2}-\tau_{t1}$ pour la solution équipondérée, et est nulle pour la solution orthogonale, ce qui permet une vérification de la plausibilité du modèle.

A titre didactique, nous envisagerons ici les deux solutions précédentes.

On remarquera encore la comparaison $\mathbb{T}_{[O]}$ est confondue avec l'interaction de la comparaison $\mathbb{G}$ et de la comparaison "p1,p2_p3"$_{[E]}$ qui oppose la première période à l'ensemble des deux autres périodes.

Le plan d'expérience suggère d'ailleurs de décomposer la comparaison globale des périodes (à deux degrés de liberté) comme la somme de "p1,p2_p3" et de "p2,p3". Cependant, dans un plan non équilibré la comparaison "p1,p2_p3"$_{[E]}$ (qui est confondue avec la comparaison d'interaction $\mathbb{G}.\mathbb{T}$) n'est pas orthogonale à la comparaison $\mathbb{T}$. Si nous voulons une décomposition orthogonale, il faudra donc, comme dans le *cross-over* 2×2, retenir pour les périodes la solution pondérée par les effectifs, soit les deux sous-comparaisons "p1,p2_p3"$_{[N]}$ et "p2,p3"$_{[N]}$. En notant les effectifs des groupes f_{g1} et f_{g2} et $s=f_{g1}+f_{g2}$, nous considérerons donc les deux contrastes

	p1	p2	p3	
g1	$+f_{g1}/s$	$-f_{g1}/2s$	$-f_{g1}/2s$	f_{g1}/s
g2	$+f_{g2}/s$	$-f_{g2}/2s$	$-f_{g2}/2s$	f_{g2}/s
	+1	-1/2	-1/2	

$$p1,p2_p3_{[N]}$$

$$\delta = \pi^{p1} - \tfrac{1}{2}(\pi^{p2} + \pi^{p3}) + \frac{f_{g1} - f_{g2}}{s}(\tau^{t1} - \tau^{t2}) - \tfrac{1}{2}(\kappa^{t1} - \kappa^{t2})$$

	p1	p2	p3	
g1	0	$+f_{g1}/s$	$-f_{g1}/s$	f_{g1}/s
g2	0	$+f_{g2}/s$	$-f_{g2}/s$	f_{g2}/s
	0	+1	-1	

$$p2,p3_{[N]}$$

$$\delta = \pi^{p2} - \pi^{p3} + \frac{f_{g1} - f_{g2}}{s}(\kappa^{t1} - \kappa^{t2})$$

Remarquons encore que les deux comparaisons d'interaction "𝔾.p2,p3" et "𝕋.p2,p3" (qui sont définies de manière canonique) coïncident et sont confondues avec la comparaison "g1p2_g2p3,g1p3_g2p2"$_{[E]}$, qui est représentée par le contraste

	p1	p2	p3	
g1	0	+1/2	-1/2	+1/2
g2	0	-1/2	+1/2	-1/2
		+1	-1	

$$g1p2_g2p3,g1p3_g2p2_{[E]} \equiv \mathbb{G}.p2,p3 \equiv \mathbb{T}.p2,p3$$

$$\delta = \kappa^{t1} - \kappa^{t2}$$

tel que δ est la différence des effets résiduels

■ *Exemple numérique*

Considérons l'exemple de données fourni par Jones et Kenward (1989, page 159), concernant la comparaison de deux traitements de la tension artérielle. La variable analysée est la pression artérielle systolique (en mm de mercure); les moyennes observées y_g sont les suivantes

	p1	p2	p3	
g1	t1 157.0909	t2 151.3636	t2 145.9091	f_{g1}=22
g2	t2 147.1111	t1 150.5556	t1 150.4444	f_{g2}=27

Nous en déduisons les effets d, les constantes a et b, et les sommes-des-carrés $sc=(d/ab)^2$ des contrastes de la décomposition précédente. Par exemple, pour la comparaison $p2,p3_{[N]}$, nous avons

$$d = \frac{22}{49}(151.3636-145.9091) + \frac{27}{49}(150.5556-150.4444) = +2.5102$$

$$a^2 = (+1)^2 + (-1)^2 = 2 = 1.414214^2$$

$$b^2 = \frac{(22/49)^2}{22} + \frac{(27/49)^2}{27} = 1/49 = 0.142857^2$$

Que nous prenions pour $\mathbb{G}$ et $\mathbb{T}$ l'une ou l'autre des deux solutions envisagées, cette décomposition est *orthogonale*: le total des cinq sommes-des-carrés est égal à la somme-des-carrés associée à la comparaison globale des six moyennes (1740.50).

	d	a	b	*sc*
$g1,g2_{[E]}$	+2.0842	0.577350	0.287213	157.97
$t1,t2_{[O]}$	+5.9217	1.224745	0.143607	1133.59
$g1,g2_{[O]}$	+4.0581	0.612372	0.287213	532.35
$t1,t2_{[E]}$	+4.5690	1.732051	0.095738	759.20
$p1,p2_p3_{[N]}$	+1.9286	1.224745	0.142857	121.50
$p2,p3_{[N]}$	+2.5102	1.414214	0.142857	154.38
$g1p2,g2p3,g1p3_g2p2_{[E]}$	+2.6717	1.414214	0.143607	173.06

Nous retrouvons les résultats fournis par Jones et Kenward dans leur tableau 4.11 (page 159), avec la correspondance suivante

Periods → $p1,p2_p3_{[N]} + p2,p3_{[N]}$

Direct treatments → $t1,t2_{[O]}$

Carry-over → $g1p2_g2p3,g1p3_g2p2_{[E]}$

On remarquera que l'effet de la comparaison "$g1,g2"_{[E]}$ est inclus dans leur source de variation "*Between subjects*".

A partir des données complètes, les statistiques de test *t*, et par suite les énoncés fiducio-bayésiens, sont obtenus suivant la méthode générale exposée au chapitre §4.

	d	s	t=d/bs	dl	p	*énoncé fiducio-bayésien*
$g1,g2_{[E]}$	+2.084	14.705	+0.493	[47]	0.624	$Pr(\|\delta\|<7.87)=0.90$
$t1,t2_{[O]}$	+5.922	16.363	+2.520	[47]	0.015	$Pr(\ \delta\ >2.87)=0.90$
$g1,g2_{[O]}$	+4.058	14.610	+0.967	[47]	0.338	$Pr(\|\delta\|<9.54)=0.90$
$t1,t2_{[E]}$	+4.569	27.824	+1.715	[47]	0.092	$Pr(\ \delta\ >1.11)=0.90$
$p1,p2_p3_{[N]}$	+1.929	16.363	+0.825	[47]	0.414	$Pr(\|\delta\|<5.00)=0.90$
$p2,p3_{[N]}$	+2.510	18.222	+0.964	[47]	0.340	$Pr(\|\delta\|<5.91)=0.90$
$g1p2,g2p3,g1p3_g2p2_{[E]}$	+2.672	18.222	+1.021	[47]	0.312	$Pr(\|\delta\|<6.09)=0.90$

LeBayésien

Dans la fenêtre principale, cliquez sur le bouton **exemples**, et sélectionnez **cross-over deux traitements et trois périodes**. Les données sont affichées dans l'éditeur de données.

Considérons par exemple l'analyse de la différence des effets résiduels: comparaison $g1p2_g2p3,g1p3_g2p2_{[E]}$.

Cliquez sur le bouton **calculer**, ce qui fait apparaître la fenêtre pour le "choix de la structure dérivée". Cliquez sur **dérivation sur les occasions** et entrez les trois coefficients appropriés [0 1 -1], puis cliquez sur le bouton **ok**. Les statistiques du protocole dérivé pertinent sont calculées.

Cliquez sur **combinaison linéaire** et entrez les coefficients [1/2 -1/2].

	effectif	moyenne	écart-type	coefficient
g1	22	5.4545...	18.5054...	+1/2
g2	27	0.1111...	17.9900...	-1/2

Puis cliquez sur le bouton **calculer** pour effectuer l'analyse: t=1.021, etc.

Cliquez sur le bouton **ok**, puis dans la fenêtre principale cliquez sur la distribution de l'effet. Dans la fenêtre des énoncés, vous obtenez par exemple pour la garantie 0.90:

$$-1.72<\delta<7.06 \quad \text{et} \quad |\delta|<6.09$$

L'analyse de la comparaison $p2,p3_{[N]}$ est obtenue pour les mêmes données dérivées, avec les coefficients [22/49 27/49].

Etc.

L'analyse spécifique des comparaisons "t1,t2"$_{[O]}$ et "g1p2_g2p3,g1p3_g2p2"$_{[E]}$ correspond à la procédure qui est exposée par Jones et Kenward, sous le titre de *"a simple and robust analysis for two-group dual designs"* (page 160). La seule différence est que leurs contrastes sont de signes opposés et ont des coefficients divisés par 2.

Mais les auteurs s'arrêtent au test t et en concluent par exemple que: *"clearly, there is no evidence (p=0.32) of a difference in carry-over effects"*. Ceci ne signifie évidemment pas que l'on a démontré l'absence de différence ou même l'existence d'une différence négligeable: nous obtenons pour la comparaison g1p2_g2p3,g1p3_g2p2$_{[E]}$ l'énoncé fiducio-bayésien

$$Pr(|\delta|<6.09) = 0.90$$

La limite 6.09 est supérieure à la différence observée entre les deux traitements (5.92 pour la solution orthogonale et 4.57 pour la solution équipondérée); elle ne peut donc en aucun cas être considérée comme relativement négligeable: nous obtenons par exemple pour la comparaison t1,t2$_{[O]}$ un test significatif au seuil unilatéral $p/2$=0.008 et

$$Pr(\delta>2.87) = 0.90$$

5.3.3 Cross-over: Autres exemples de généralisations

Nous donnerons encore les exemples de généralisations suivants, bien que ceux-ci fassent intervenir des comparaisons à plusieurs degrés de liberté qui sortent du cadre de cet ouvrage.

■ *Deux traitements et quatre périodes*

Par exemple, en utilisant quatre périodes, on peut avoir

	o1 = p1	o2 = p2	o3 = p3	o4 = p4
g1	t1	t1	t2	t2
g2	t2	t2	t1	t1

Une généralisation directe de situation avec deux périodes est obtenue en introduisant le facteur *auxiliaire* à deux modalités B_2={b1,b2} (pour B*locs*) qui oppose les deux premières périodes aux deux dernières. La comparaison à trois degrés de liberté $\mathbb{P}$ est alors décomposée orthogonalement comme la somme des comparaisons $\mathbb{B}$ [à 1 *dl*] *et* $\mathbb{P}$(B) [à 2 *dl*].

En particulier si les effectifs des deux groupes sont identiques, nous avons la décomposition orthogonale de $\mathbb{G}*\mathbb{O}$ suivant les cinq comparaisons (de forme produit) suivantes

$$\begin{array}{rcll} \mathbb{G} & \equiv & \mathbb{B}.\mathbb{T} & [1\ dl] \\ \mathbb{B} & \equiv & \mathbb{G}.\mathbb{T} & [1\ dl] \\ \mathbb{T} & \equiv & \mathbb{B}.\mathbb{G} & [1\ dl] \\ \mathbb{P}(B) & & & [2\ dl] \\ \mathbb{P}.\mathbb{G}(B) & \equiv & \mathbb{P}.\mathbb{T}(B) & [2\ dl] \end{array}$$

■ *Trois traitements et trois périodes*

Par exemple, avec trois traitements et trois périodes, nous pouvons avoir

	o1 o2 o3 = = = p1 p2 p3		
g1	t1	t2	t3
g2	t2	t3	t1
g3	t3	t1	t2

	o1 o2 o3 = = = t1 t2 t3		
g1	p1	p2	p3
g2	p3	p1	p2
g3	p2	p3	p1

Dans ce cas les comparaisons $\mathbb{G}$, $\mathbb{P}$ et $\mathbb{T}$ ont chacune deux degrés de liberté. Leur somme n'épuise pas par conséquent les 8 degrés de liberté de la comparaison $\mathbb{G}*\mathbb{O}$. En particulier, si les effectifs des trois groupes sont identiques, ces trois comparaisons sont orthogonales et chacune d'elle est une sous-comparaison de l'interaction des deux autres. La comparaison résiduelle de leur somme par rapport à $\mathbb{G}*\mathbb{O}$ a deux degrés de liberté; elle est l'intersection des interactions $\mathbb{G}.\mathbb{P}$, $\mathbb{G}.\mathbb{T}$ et $\mathbb{P}.\mathbb{T}$. Nous avons donc la décomposition orthogonale pour le cas équilibré

$$\begin{array}{ll} \mathbb{G}\subset\mathbb{P}.\mathbb{T} & [2\ dl] \\ \mathbb{P}\subset\mathbb{G}.\mathbb{T} & [2\ dl] \\ \mathbb{T}\subset\mathbb{P}.\mathbb{T} & [2\ dl] \\ \mathbb{G}.\mathbb{P}\cap\mathbb{G}.\mathbb{T}\cap\mathbb{P}.\mathbb{T} & [2\ dl] \end{array}$$

On remarquera que dans ce cas chacune des comparaisons $\mathbb{P}$ et $\mathbb{T}$ n'est de type produit sur G*O que dans la représentation où l'on prend pour les occasions respectivement les périodes et les traitements.

On peut également construire une décomposition orthogonale pour le cas non équilibré ou prendre en compte les effets résiduels, en procédant de la même manière que dans le cas précédent. On notera cependant que dans le cas général, même pour un plan équilibré, les comparaisons ainsi obtenues pourront ne pas être de type produit et que, par conséquent, les solution usuelles ("avec égalité des variances") seront peu raisonnables: *cf.* Senn, 1993, page 246.

L'analyse spécifique permet d'expliciter les conditions de validité en jeu et de recourir aux solutions "sans égalité de variance", en utilisant les procédures présentées au chapitre ↳6.

5.3.4 Exemple: La conservation des morceaux de viande

Il s'agit d'un exemple correspondant à l'utilisation de quatre traitements et de quatre périodes dans un contexte différent.

Dans une étude sur la conservation des morceaux de viande de bovins, on s'intéresse à quatre conditions expérimentales, définies par le croisement de deux facteurs principaux, *Température* et *Ventilation*, à deux modalités chacun. Chaque bovin est divisé en quatre quartiers, suivant le croisement des deux facteurs à deux modalités, que nous appellerons *Côté* (c1: droit ou c2: gauche) et *Moitié* (m1: avant ou m2: arrière). Les modalités des facteurs *Température* (t1 et t2) et *Ventilation* (v1 et v2) sont affectées aux quatre quartiers d'un même bovin, suivant quatre arrangements différents, ce qui conduit à définir quatre *groupes* de bovins, correspondant chacun à un arrangement différent. Nous avons ici les deux représentations possibles suivantes comme un plan S<G>*O, selon que nous considérons comme *Occasions* les *Quartiers* ou les *Conditions expérimentales*.

	o1 = t1v1	o2 = t2v1	o3 = t1v2	o4 = t2v2
g1=x1y1	c1m1	c2m1	c1m2	c2m1
g2=x2y1	c2m1	c1m1	c2m1	c1m2
g3=x1y2	c1m2	c2m1	c1m1	c2m1
g4=x2y2	c2m1	c1m2	c2m1	c1m1

	o1 = c1m1	o2 = c2m1	o3 = c1m2	o4 = c2m1
g1=x1y1	t1v1	t2v1	t1v2	t2v2
g2=x2y1	t2v1	t1v1	t2v2	t1v2
g3=x1y2	t1v2	t2v2	t1v1	t2v1
g4=x2y2	t2v2	t1v2	t2v1	t1v1

Les trois facteurs à quatre modalités *Groupes* (G_4), *Conditions expérimentales* ($T_2 \times V_2$) et *Quartiers* ($C_2 \times M_2$) constituent un *carré-latin*: ils sont croisés deux à deux, mais pas dans leur ensemble. Plus précisément, il apparaît ici que chacun des quatre plans partiels

$S<\{g1,g2\}>\times\{o1,o2\}$ $S<\{g1,g2\}>\times\{o3,o4\}$
$S<\{g3,g4\}>\times\{o1,o2\}$ $S<\{g3,g4\}>\times\{o3,o4\}$

a la même structure que le *cross-over* 2×2.

Ceci motive l'introduction, pour décrire les groupes, de deux facteurs auxiliaires croisés à deux modalités chacun:

- X_2 qui indique à quel côté est affectée chaque modalité du facteur *Température*

x1 → t1 à droite et t2 à gauche
x2 → t2 à droite et t1 à gauche

- Y_2 qui indique à quelle moitié est affectée chaque modalité du facteur *Ventilation*

y1 → v1 à l'avant et v2 à l'arrière
y2 → v2 à l'avant et v1 à l'arrière

Grâce à l'introduction des facteurs auxiliaires X et Y dont l'interprétation est claire, les 15 degrés de liberté de la comparaison globale G*O sont entièrement décomposés en autant de comparaisons à un degré de liberté, ce qui permet une analyse très fine de ces effets. Les effets principaux des six facteurs T_2, V_2, C_2, M_2, X_2 et Y_2, ainsi que toutes leurs interactions, sont des comparaisons à un degré de liberté, de type produit sur G*O, quelle que soit la représentation du plan adoptée, et peuvent donc être facilement analysées. Bien entendu la structure particulière du plan entraîne des *confusions* entre les comparaisons. On peut vérifier que les décompositions canoniques de chacun des deux croisements $X_2 \times Y_2 \times T_2 \times V_2$ et $X_2 \times Y_2 \times C_2 \times M_2$ coïncident, selon les équivalences suivantes

$X_2 \times Y_2 \times T_2 \times V_2$		$X_2 \times Y_2 \times C_2 \times M_2$
X	≡	X
Y	≡	Y
X.Y	≡	X.Y
T	≡	X.C
X.T	≡	C
Y.T	≡	X.Y.C
X.Y.T	≡	Y.C
V	≡	Y.M
X.V	≡	X.Y.M
Y.V	≡	M
X.Y.V	≡	X.M
T.V	≡	X.Y.C.M
X.T.V	≡	Y.C.M
Y.T.V	≡	X.C.M
X.Y.T.V	≡	C.M

La décomposition est en outre orthogonale si les effectifs des quatre groupes sont identiques.

5.4 MÉMORISATION DE RÉCITS

5.4.1 Présentation générale et objectifs de l'expérience

Cette expérience de psychologie porte sur la mémorisation de récits par des enfants. Précédemment on avait pu mettre en évidence le fait que le rappel d'une information varie fortement en fonction de son importance relative dans le récit. Si cet effet est présent à tous les âges, on constate cependant une coupure radicale dans la performance entre 8 et 9 ans: par exemple, dans une expérience de rappel immédiat, après une lecture d'un récit, des groupes âgés de 7 et 8 ans rappelaient en moyenne 20% et 22% des propositions, alors que ceux âgés de 9 et 10 ans en rappelaient 41% et 40%. L'expérience rapportée ici (Denhière et Lecoutre, 1983; Lecoutre, Rouanet et Denhière, 1988) se proposait de mettre à l'épreuve les modèles susceptibles de rendre compte de ces phénomènes dans une situation de *reconnaissance* de récits. Dans ce cas le sujet, après la lecture d'un récit, doit dire si les propositions qu'on lui soumet appartiennent ou non au récit (et non essayer de retrouver ces propositions comme dans l'expérience de rappel).

Une telle problématique amène inéluctablement la prise en compte simultanée d'un grand nombre de facteurs, d'où un plan d'expérience très complexe:

- reconnaissance immédiate ou différée (après un certain délai), d'où récits et énoncés différents;
- niveaux d'âges;
- facteurs inhérents à la structure du matériel que l'on s'efforce de contrôler, origine des énoncés soumis à la reconnaissance (originaux ou étrangers au récit), proximité des énoncés étrangers avec les énoncés originaux sur le plan sémantique, importance des énoncés dans le récit;
- enfin facteurs relevant des habituelles précautions méthodologiques, contrebalancement de l'ordre de passation des récits par exemple.

180 enfants de trois niveaux d'âge, 7, 8 et 10 ans, ont participé à cette expérience. Trois récits intitulés *Géant 1*, *Géant 2* et *Ourson* étaient utilisés. Chaque sujet a été soumis à deux épreuves de *reconnaissance*: reconnaissance immédiate d'un premier récit aussitôt après sa lecture, et reconnaissance différée d'un second récit une semaine après sa lecture. Chaque sujet a été affecté à l'un des six arrangements, obtenus à partir des trois combinaisons possibles pour choisir les deux récits parmi les trois utilisés et des deux ordres d'administration possibles pour chacune de ces combinaisons.

L'épreuve de reconnaissance consistait (pour un récit donné) en la présentation de 40 énoncés, dont le sujet devait dire s'ils appartenaient ou non au

récit. Ces énoncés se composaient de 20 énoncés originaux et 20 énoncés distracteurs dérivés, dont 10 sémantiquement proches et 10 sémantiquement lointains (il y avait également 10 énoncés étrangers qui, n'ayant donné lieu à pratiquement aucune erreur, n'ont pas été retenus dans l'analyse). Les énoncés originaux se répartissaient en deux catégories: 10 d'entre eux exprimaient la *macrostructure* du récit, c'est-à-dire l'information la plus importante, celle nécessaire pour construire un résumé du récit. Pour chacun des trois récits, les énoncés distracteurs proches et lointains étaient encore divisés en deux catégories à partir du numéro d'ordre des énoncés originaux dont ils étaient dérivés. Les 10 énoncés proches impairs étaient proposés à la moitié des sujets, et les énoncés complémentaires à l'autre moitié. Dans chaque cas l'ensemble des 40 énoncés étaient divisés en 5 blocs de 8 énoncés (4 originaux + 4 distracteurs: 2 proches et 2 lointains). Par permutation circulaire à partir d'un ordre de succession tiré au hasard, on obtenait 5 ordres de présentation différents.

5.4.2 Plan d'analyse

Le plan d'expérience est particulièrement complexe et combine les structures de *blocs incomplets* et de *mesures répétées*. La relation entre les *Sujets* et les *Récits* n'est ni un emboîtement ni un croisement (chaque sujet étant soumis à deux des trois récits). Les données peuvent néanmoins être analysées suivant un plan *Sujets<Groupes>×Occasions*.

- *Groupes*

Les groupes sont définis par le croisement de quatre facteurs élémentaires: *Niveaux d'âges* à 3 modalités noté A_3, *Combinaison des récits* à 3 modalités noté C_3; *Ordre d'administration* (pour chaque combinaison donnée de deux récits) à 2 modalités noté X_2; le facteur, que nous appellerons *Pair-Impair*, relatif au mode de dérivation des énoncés distracteurs, à 2 modalités noté Z_2.

Les groupes sont donc un facteur *complet* $G_{36}=A_3\times C_3\times X_2\times Z_2$. On notera que la facteur "Ordre de présentation des blocs" à cinq modalités n'a pas été retenu parce qu'il se trouve dans chaque groupe *confondu* avec le facteur Sujets.

- *Occasions*

Pour chaque sujet, les 80 observations retenues pour l'analyse constituent autant d'*occasions*, qui peuvent d'abord être décrites comme le croisement des deux facteurs *Délai de reconnaissance* à 2 modalités (immédiate et différée) et

Items à reconnaître à 40 modalités: $O_{80}=D_2\times I_{40}$. A l'intérieur de chaque groupe, la combinaison particulière des facteurs C, O et Z détermine pour chaque délai les énoncés (et le récit) auxquels correspondent les items à reconnaître.

Les 40 items se répartissent en 8 catégories. D'abord chaque item est constitué, soit par un énoncé original, soit par un énoncé distracteur: facteur *Origine des items* à 2 modalités noté O_2. Ensuite à l'intérieur de chaque origine, chacun des items est soit proche soit lointain d'un autre item appartenant à l'autre origine: facteur *Proximité des items* à 2 modalités noté P_2. Enfin chaque item est un énoncé qui appartient ou qui n'appartient pas à la macrostructure du récit correspondant (ou dérivé d'un tel énoncé): facteur *Appartenance à la macrostructure* à deux modalités noté M_2.

Les facteurs O, P et M étant croisés les occasions peuvent s'écrire sous la forme $D_2\times I<O_2\times P_2\times M_2>$ (l'emboîtement n'étant pas équilibré). On remarquera que n'ont été retenus, ni le facteur "Blocs d'énoncés", ni la relation d'appariement entre énoncés originaux et énoncés distracteurs introduite par le mode de dérivation de ces derniers. Mais ceci est sans conséquence sur les analyses effectuées.

■ *Plan Sujets<Groupes>×Occasions*

Finalement le plan d'analyse peut être écrit

$$S_5<A_3\times C_3\times X_2\times Z_2>\times D_2\times I<O_2\times P_2\times M_2>$$

La correspondance entre les modalités des facteurs de ce plan et les récits, qui permet d'effectuer dans le cadre de ce plan les analyses spécifiques relatives aux facteurs *Récits* et *Enoncés* est la suivante.

Combinaison \ Ordre	Délai de reconnaissance d1 (immédiate)	d2 (différée)
c1 x1	Géant 1	Géant 2
c2 x1	Ourson	Géant 1
c3 x1	Géant 2	Ourson
c1 x2	Géant 2	Géant 1
c2 x2	Géant 1	Ourson
c3 x2	Ourson	Géant 2

Par exemple, la comparaison globale des trois récits en reconnaissance immédiate ("r1,r2,r3/d1") équivaut à la comparaison

"c1x1_c2x2,c3x1_c1x2,c2x1_c3x2"

■ *Planification des analyses spécifiques*

Sans qu'il soit nécessaire ici de connaître les théories psychologiques qui permettront l'interprétation des résultats, il est clair que les questions essentielles qui se posent à propos de ces données concernent, d'une part l'importance des différences entre les enfants d'âges différents (comparaison 𝔸) et d'autre part l'effet du facteur *Appartenance la macrostructure du récit* (comparaison 𝕄). Il convient en outre de s'assurer que les effets de ces comparaisons ne sont pas sensiblement affectés par les autres facteurs (en particulier par le délai de reconnaissance).

Etant donné la relation particulière entre les sujets et le facteur *Récit*, il est apparu judicieux de comparer les récits, non pas sur l'ensemble des occasions, mais d'une part en reconnaissance immédiate et d'autre part en reconnaissance différée (comparaisons ℝ/d1 et ℝ/d2). De même les comparaisons relatives aux facteurs Z (*Pair-Impair*), P (*Proximité des items*) et M (*Appartenance à la macrostructure*) sont considérées conditionnellement à l'origine des items (elles ont peu d'intérêt globalement). Par suite les analyses spécifiques peuvent être organisées autour des principales comparaisons (avec leur nombres de degrés de liberté)

𝔸 [2]	ℤ/o2 [1]	𝔻 [1]	𝕆 [1]
ℝ/d1 [2]	ℝ/d2 [2]	𝕏(R)/d1 [3]	𝕏(R)/d2 [3]
ℙ/o2 [1]	𝕄/o1 [1]		

Compte tenu des objectifs de l'expérience, la comparaison des âges est en outre décomposée en "a2,a1" et "a3,a2".

Des analyses plus fines pourront être obtenues, en considérant les comparaisons d'interaction (notamment 𝔸.𝔻), ou en conditionnant les comparaisons précédentes. Notamment les comparaisons 𝔸 et 𝔻 pourront être examinées conditionnellement à l'origine des items, et la comparaison 𝕄 conditionnellement à l'âge ou au délai. Les effets du facteur *Items "à l'intérieur* de O×P×M" peuvent enfin être analysés pour chaque récit et chaque délai.

5.4.3 Principaux résultats

L'espace d'observation est ici un ensemble binaire: pour chaque item soumis à reconnaissance, nous avons codé 1 une réponse correcte (acceptation pour un énoncé original et rejet pour un énoncé distracteur) et 0 une réponse incorrecte. Nous pouvons néanmoins utiliser sans trop d'inquiétudes les procédures

d'analyse spécifique pour les comparaisons de moyennes, en raison des effectifs très élevés à partir desquels sont calculés les fréquences en jeu. Toutes les comparaisons envisagées sont facilement analysées selon ces procédures.

Une synthèse des principaux résultats est fournie dans le tableau suivant.

comparaison*	effet	e	F dl	énoncé fiducio-bayésien
z1, z2/o2	d=+0.004	0.015	0.07 [1,144]	Pr(|δ|<0.025) =0.90
a2, a1	d=+0.029	0.011	7.72 [1,96]	Pr(0.012<δ<0.047)=0.90
a3, a2	d=+0.033	0.010	10.91 [1,96]	Pr(0.016<δ<0.049)=0.90
d1, d2	d=+0.059	0.007	77.26 [1,144]	Pr(0.047<δ<0.070)=0.90
o2, o1	d=+0.002	0.015	0.02 [1,144]	Pr(|δ|<0.024) =0.90
p2, p1/o1	d=+0.002	0.010	0.05 [1,144]	Pr(|δ|>0.018) =0.90
p2, p1/o2	d=+0.418	0.012	1249.39 [1,144]	Pr(δ >0.403) =0.90
m2, m1/o1	d=+0.020	0.010	3.69 [1,144]	Pr(|δ|<0.033) =0.90 Pr(δ <0) =0.03

*Dans la formule désignant la comparaison, les modalités sont ordonnées selon l'ordre décroissant des fréquences de réponses correctes observées.

LeBayésien

Supposons que nous ne disposions ici (comme dans la plupart des publications) que de l'effet observé et du rapport F. Nous pourrions retrouver la distribution fiducio-bayésienne à partir de ces deux quantités, selon la procédure illustrée au chapitre 4. Nous examinerons ici les précautions à prendre pour tenir compte de la précision (ou plutôt de l'imprécision) avec laquelle sont données les valeurs.

Dans la fenêtre principale, cliquez sur le bouton **effet**, puis sur le bouton **options: cm\sc\F\t\p**, ce qui fait apparaître une nouvelle fenêtre.

Par exemple pour la comparaison z2,z2/o2, l'effet donné avec 3 décimales (d=0.004) est en fait compris entre 0.0035 et 0.0045. La prudence nous impose de considérer le cas le plus défavorable à la conclusion recherchée (ici d'effet négligeable), soit 0.0045. Le rapport F donné avec 2 décimales (F=0.07) est compris entre 0.065 et 0.075; nous devons considérer la plus petite de ces deux valeurs, qui correspond à la plus grande imprécision expérimentale. Entrez donc l'effet 0.0045 (sélectionnez s'il y a lieu l'option **effet d**), les nombres de degrés de liberté du rapport F (1 et

144), et le rapport F (0.065). Vérifiez que l'option **rapport F** est sélectionnée, et cliquez sur le bouton **calculer**.

L'échelle $e=|d|/\sqrt{F}$ de la distribution est calculée: vous obtenez ici e=0.01765... qui est un *majorant* de la valeur réelle (qui est avec 3 décimales 0.015).

Cliquez sur le bouton **OK**, ce qui vous ramène à la fenêtre principale, avec la courbe de la distribution relative à δ. Cliquez sur la case à cocher **utiliser un majorant de l'échelle de la distribution**, puis procédez de la manière habituelle pour obtenir les énoncés. Vous obtenez par exemple les énoncés:

pour la garantie 0.90: $Pr(|\delta|<0.030) > 0.90$
pour la limite 0.25: $Pr(|\delta|<0.025) > 0.828$

En utilisant directement les valeurs données (d=0.004 et F=0.07), vous auriez obtenu:

pour la garantie 0.90: $Pr(|\delta|<0.026) = 0.90$
pour la limite 0.25: $Pr(|\delta|<0.025) = 0.888$

et en utilisant les valeurs les plus favorables à la conclusion (d=0.0035 et F=0.075):

pour la garantie 0.90: $Pr(|\delta|<0.022) > 0.90$
pour la limite 0.25: $Pr(|\delta|<0.025) > 0.939$

Les autres comparaisons sont analysées de manière analogue.

Bien entendu, si on ne disposait que de l'effet et du rapport F, la prudence imposerait de considérer les énoncés les moins favorables aux conclusions, ce qui ne modifierait pas sensiblement celles-ci.

Nous ferons ci-après quelques commentaires généraux.

■ *Facteur "Pair-Impair"*

La comparaison "z1,z2/o2" permet de vérifier que le facteur *Pair-Impair*, comme on pouvait l'espérer a un effet négligeable pour les items distracteurs (les seuls modifiés par ce facteur). Par référence cela permet de fixer un critère *interne* prenant en compte la précision expérimentale pour apprécier dans le contexte de cette expérience l'importance des autres effets. On observe effectivement des fréquences de rejet correct des distracteurs très voisines pour les deux modalités: 0.695 et 0.691. Avec une garantie fiducio-bayésienne 0.90, la différence vraie est inférieure en valeur absolue à 0.025. Nous considérerons par conséquent qu'un effet inférieur à cette limite peut être tenu pour négligeable. On remarquera que, rapporté au nombre total

d'items soumis à reconnaissance (40), la valeur 0.025 correspond à une différence d'exactement un item.

■ *Niveaux d'âge*

Les fréquences de réponses correctes observées s'ordonnent comme les niveaux d'âges, soit 0.661 (7 ans), 0.691 (8 ans) et 0.723 (10 ans). Toutefois les différences, sans être négligeables apparaissent relativement limitées. Le résultat très significatif des tests de signification pour chacune des deux comparaisons "a2,a1" et "a3,a2" indique dans chaque cas que l'existence des différences est bien établie. Mais les procédures fiducio-bayésiennes permettent également de conclure inférentiellement que ces différences sont relativement limitées, notamment entre les enfants de 8 ans et ceux de 10 ans, sans commune mesure avec celles constatées dans les épreuves de rappel. Cette dernière conclusion peut en outre être établie aussi bien en reconnaissance immédiate qu'en reconnaissance différée: l'effet de l'interaction "a3,a2.d2,d1" (non présenté dans le tableau précédent) est négligeable.

Une analyse plus fine montre que, descriptivement, la différence de performance entre les enfants de 8 ans et ceux de 10 ans peut être attribuée au rejet des énoncés distracteurs, et non à la reconnaissance des énoncés originaux. Pour ces derniers la fréquence de réponses correctes est même supérieure pour les enfants de 8 ans: 0.710 contre 0.693. Au contraire la fréquence de rejet correcte est assez nettement supérieure pour les enfants les plus âgés: 0.672 (8 ans) contre 0.754 (10 ans). Les énoncés fiducio-bayésiens correspondants sont respectivement

$$Pr(|\delta|<0.032) = 0.90 \quad \text{(énoncés originaux)}$$
$$Pr(\delta>+0.065) = 0.90 \quad \text{(distracteurs)}$$

On peut ici être plus exigeant et souhaiter un énoncé conjoint. Prenant en compte le fait que ces deux distributions ne sont pas indépendantes, nous pouvons énoncer ici que nous avons une probabilité d'au moins

$$1-(1-0.90)-(1-0.90) = 0.80$$

pour que, simultanément, la différence soit inférieure à 0.032 dans le premier cas et supérieure à +0.065 dans le second

$$Pr(|\delta_1|<0.032) \textit{ et } \delta_2>+0.065)) > 0.80$$

Si la garantie 0.80 ne nous paraît pas suffisante, nous pouvons choisir par exemple une probabilité 0.95 pour chacun des énoncés pris séparément, d'où

$$Pr(|\delta_1|<0.037) \textit{ et } \delta_2>+0.060)) > 0.90$$

■ *Délai de reconnaissance*

Globalement la fréquence de réponses correctes observées diminue avec le délai: 0.663 en reconnaissance différée contre 0.721 en reconnaissance immédiate. Ici encore l'effet est bien établi, mais d'importance limitée.

■ *Origine et Proximité des items*

Globalement, les fréquences de réponses correctes observées sont du même ordre pour les items originaux (acceptation de l'item) et les items distracteurs (rejet de l'item): 0.691 et 0.693. Mais on constate une hétérogénéité considérable, parmi les distracteurs entre les items "proches" et les items "lointains", les fréquences de rejet correct étant respectivement 0.484 et 0.902. Les conclusions descriptives, très nettes, sont bien entendu prolongées par les analyses inférentielles.

■ *Appartenance à la macrostructure*

Globalement les énoncés originaux importants du récit (appartenant à la macrostructure) ne sont pas mieux reconnus que les autres, les fréquences d'acceptation observées étant respectivement 0.681 et 0.701. Inférentiellement on peut conclure que la différence est faible, mais pas qu'elle est négligeable pour le critère adopté (on obtient la limite 0.033). En tout état de cause on a une probabilité fiducio-bayésienne égale seulement à 0.03 que les énoncés importants soient mieux reconnus (soit une différence de signe opposé à celle observée). Cette conclusion est la même quand on conditionne par le niveau d'âge, par l'origine des énoncés, ou par le délai de reconnaissance.

5.5 APPRENTISSSAGE PERCEPTIF INCIDENT

5.5.1 Présentation générale et objectifs de l'expérience

Dans cette expérience d'apprentissage (voir Lecoutre et Lecoutre, 1979), la tâche du sujet consiste à relier par un trait de crayon les 26 lettres de l'alphabet disposées au hasard sur une feuille (suivant une disposition constante). Les sujets sont répartis en deux groupes. Le premier groupe effectue d'abord la tâche; puis le second groupe l'effectue avec un apprentissage perceptif *incident*, qui consiste à observer un sujet en train d'effectuer la

tâche (sans que les sujets aient été avertis qu'ils auraient à effectuer la tâche ultérieurement). Pour cela les sujets sont d'abord répartis au hasard par *paires*. Dans chaque paire, l'un des sujets est choisi au hasard comme "sujet de l'expérience" et l'autre est chargé du rôle d'"expérimentateur". Chaque expérimentateur fait exécuter à son sujet six *essais* successifs de la tâche, en le surveillant attentivement pour vérifier qu'il ne commet pas d'erreur et enregistre les temps d'exécution. Quand le sujet a terminé, on inverse les rôles: sans en avoir été averti auparavant, l'"expérimentateur" doit à son tour effectuer six essais.

5.5.2 Plan d'analyse

Cette expérience, réalisée avec 40 sujets met en jeu les quatre facteurs élémentaires suivants:

S_{40} – *Sujets*
P_{20} – *Paires de sujets* à 15 modalités
E_6 – *Essais* à 6 modalités
C_2 – *Conditions de passation* à 2 modalités

Chaque sujet est affecté à exactement une paire et une condition, et S est croisé avec E. Le plan d'expérience peut donc s'écrire

$$S_1 \langle P_{20} \times C_2 \rangle \times E_6$$

Il y a un seul sujet à l'intérieur de chaque paire et de chaque condition. En laissant de côté le facteur E, cette situation est classique: on compare deux conditions avec deux fois k sujets groupés en k paires et on retient généralement l'une des deux solutions suivantes.

• La première solution est d'adopter le plan $P \times C_2$, ce qui conduit à l'analyse de deux groupes *appariés*, et tient compte fort justement de l'appariement des sujets.

• La seconde solution est d'adopter le plan $S\langle C_2 \rangle$, ce qui conduit à traiter les données sans tenir compte de cet appariement (plan en groupes *indépendants*). Or ce dernier plan est en général non approprié, puisqu'il exige des conditions de validité supplémentaires, notamment celle d'une corrélation nulle entre les deux variables associées aux deux conditions. Ici notamment cette condition est peu réaliste puisque un expérimentateur ayant observé un sujet lent aura eu un temps d'apprentissage incident supérieur à celui d'un expérimentateur ayant observé un sujet rapide.

Cette discussion conduit à retenir le plan d'analyse complet du type *Sujets* ×*Occasions*, avec $O_{12}=C_2\times E_6$, soit le plan

$$P_{20}\times C_2\times E_6$$

La possibilité de tenir compte des deux facteurs aléatoires P et S dans l'analyse ne sera pas examinée ici.

■ *Analyses spécifiques*

La décomposition canonique du croisement $C_2\times E_6$ (les comparaisons **C**, **E** et **C.E**) est ici de peu d'intérêt, car elle ne s'accorde pas aux objectifs de l'expérience qui peuvent être résumés au moyen des trois questions suivantes.
(1) "Est-ce que la répétition de la tâche permet d'obtenir une diminution du temps d'exécution pour les sujets de la première condition?".
(2) "Est-ce que l'apprentissage perceptif préalable procure aux sujets de la deuxième condition un gain initial (c'est-à-dire au premier essai) dans le temps d'exécution?".
(3) "Est-ce que ce gain éventuel tend à diminuer avec la répétition de la tâche?".

Des réponses directement appropriées à ces trois questions peuvent être obtenues à partir des analyse spécifiques des comparaisons respectives

(1) **L**in E/c1 (2) c2,c1/e1 (3) c2,c1.**L**in E

On pourrait également à titre complémentaire analyser la résiduelle -**L**in E/c1 et son interaction avec **C** (ou les composantes de degrés plus élevés **Q**ua E/c1, **C**ub E/c1, etc.).

■ *Résultats*

La figure 5.2 présente les moyennes observées pour chaque condition et chaque essai (temps exprimés en secondes). Le tableau suivant et la figure 5.3 fournissent une synthèse des principaux résultats.

comparaison	effet	e	*t* [dl]	*p*/2	énoncé fiducio-bayésien
Lin E/c1	-11.081	1.669	-6.639 [19]	0.000	*Pr*(δ <-8.86)=0.90
c2, c1/e1	-17.900	8.549	-2.094 [19]	0.025	*Pr*(δ <-6.55)=0.90
c2, c1.**L**in E	+ 0.811	1.680	+0.483 [19]	0.317	*Pr*(|δ|< 3.20)=0.90

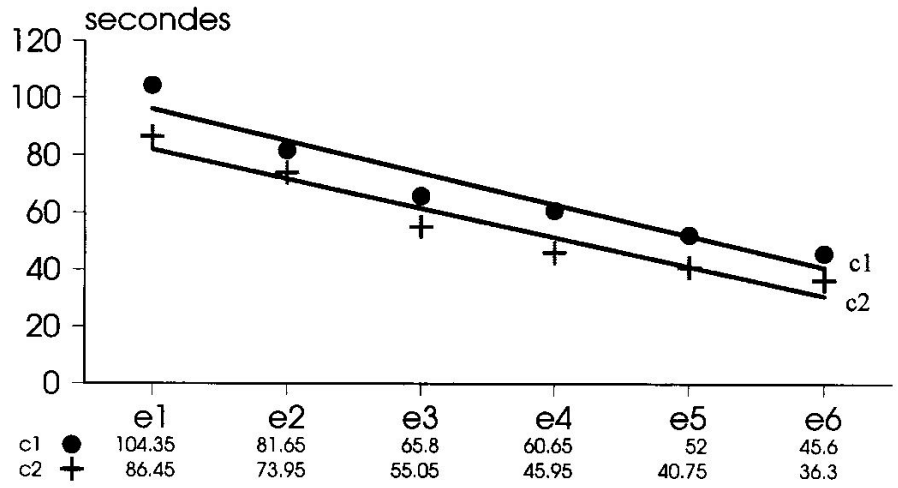


Figure 5.2 - Expérience apprentissage perceptif incident: Moyennes observées pour chaque essai avec les droites de régression ajustées dans chaque condition

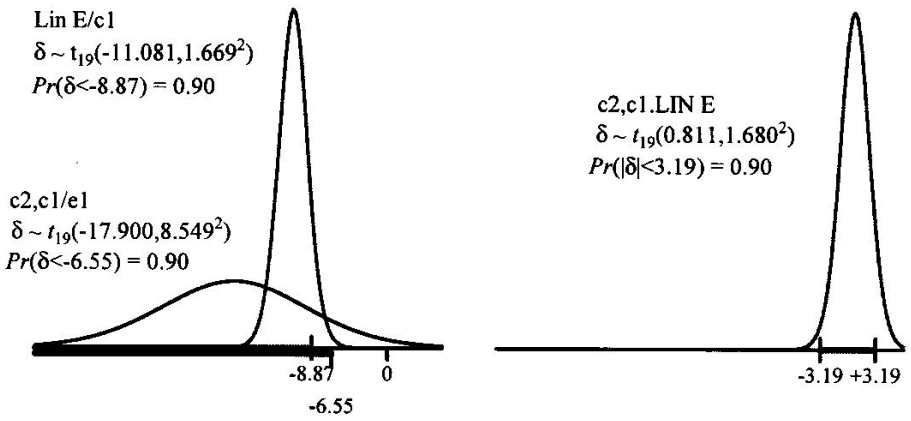


Figure 5.3 - Distributions fiducio-bayésiennes des effets

LeBayésien

Le plan comporte 12 *occasions*, ce qui dépasse en principe les limites (8 *occasions*) du programme. Cependant les comparaisons "c2,c1/e1" et "Lin E/c1" ne font intervenir respectivement que 2 et 6 *occasions*. Nous illustrerons ici l'utilisation d'un fichier ASCII standard de données (cette procédure permet donc d'importer des fichiers créés par ailleurs).

Dans la fenêtre principale, cliquez sur le bouton **exemples**, et sélectionnez **apprentissage perceptif incident**. Apparaît une fenêtre comportant dans la partie supérieure une zone de texte où est affiché le fichier des données. Celui-ci comporte, pour chacune des 20 *paires de sujets*, 13 *objets* séparés par des espaces, soit successivement:

- le numéro du sujet (objet 1),
- les observations correspondant aux 6 essais de la première condition (objets 2 à 7),
- les observations correspondant aux 6 essais de la deuxième condition (objets 8 à 13).

Ici le nombre d'objets affiché (13, qui est le nombre d'objets trouvés sur la première ligne du fichier) est le bon.

Cliquez sur le bouton **choisir les objets**, ce qui fait apparaître deux listes. Chacune d'elle comporte la liste des 13 objets, numérotés de ¤1 à ¤13; pour faciliter l'identification des objets, il est associé à chacun d'eux la valeur correspondante pour la première paire de sujets. La liste de gauche permet de sélectionner (s'il y a lieu) les objets fournissant les codes des modalités des facteurs décrivant les groupes (ou "GFacteurs") et n'est pas utilisée ici. La liste de droite permet de sélectionner les *occasions* (8 au maximum).

La comparaison "c2,c1/e1", qui met en jeu les objets 2 et 8: dans la liste **occasions**, cliquez successivement sur **¤2 = 85** puis sur **¤8 = 65**, pour sélectionner ces deux objets (pour supprimer un objet sélectionné par erreur, cliquez à nouveau sur cet objet). Au dessus de la liste doit maintenant être affiché: ¤2 ¤8.

Cliquez sur le bouton **ok**, ce qui fait apparaître l'éditeur de données, avec un plan $S_{20} \times O_2$ correspondant aux deux *occasions* sélectionnées. Cliquez sur le bouton **Calculer**. Les statistiques sont affichées:

	effectif	moyenne	écart-type
o1	20	104.35	35.7642...
o2	20	86.45	25.2388...

Suivez la démarche habituelle pour l'analyse: sélectionnez l'option **combinaison linéaire** et entrez les coefficients -1 (pour o1) et 1 (pour o2), puis cliquez sur le bouton **calculer**, etc.

Pour la comparaison "Lin E/c1", cliquez à nouveau sur le bouton **exemples** de la fenêtre principale, et sélectionnez maintenant les objets ¤2, ¤3, ¤4, ¤5, ¤6 et ¤7, etc. Dans la fenêtre des statistiques, cliquez sur le bouton **régression**, puis procédez comme dans l'exemple temps de réaction en §6.2 (avec ici les abscisses [1, 2,... 6].

Seule la comparaison "c2,c1.Lin E" ne peut pas être traitée directement: il faudrait créer d'abord un fichier de données dérivées qui comporterait, pour chaque paire de sujets, la différence entre la condition 2 et la condition 1 pour chacun des essais; on procéderait ensuite comme dans le cas de la comparaison "Lin E/c1".

Des réponses précises aux questions posées peuvent ainsi être obtenues.

(1) Nous pouvons conclure à une diminution notable du temps d'exécution pour la première condition, puisque nous avons la garantie fiducio-bayésienne 0.90 que le gain moyen par essai (fourni par la pente de la droite de régression) est supérieur à 8.86 secondes.

(2) Nous pouvons conclure que l'apprentissage incident procure un gain initial notable, puisque nous avons la garantie fiducio-bayésienne 0.90 que ce gain soit supérieure à 6.55 secondes (ce qui représente un peu plus de la moitié du gain moyen observé par essai dans la première condition).

(3) Ici nous avons la garantie 0.90 que la différence des deux pentes des droites de régression dans chacune des deux conditions soit inférieure en valeur absolue à 3.20 secondes. Cette limite peut difficilement être tenue pour négligeable. La précision expérimentale est insuffisante pour obtenir une conclusion satisfaisante.

5.6 Enquête pédagogique: Questionnaire mathématique

5.6.1 Description du plan d'enquête

Dans cette enquête pédagogique réalisée en Suisse (Tourneur et Cardinet, 1981), 80 *élèves*, répartis en 40 *groupes* de deux élèves chacun, ont été soumis à un questionnaire mathématique comportant 60 questions. Chaque groupe est caractérisé à partir des trois facteurs (inter-sujets):

- *Classes* à 20 modalités - il y a 4 élèves par classe;
- *Formes du questionnaire* à 5 modalités - chaque classe a été soumise à l'une des cinq formes différentes comportant 12 questions chacune;
- *Ages* à 2 modalités - dans chaque classe, deux sous-groupes d'élèves sont distingués en fonction de leur âge.

En résumé les *Elèves* (E_{80}) sont emboîtés dans les *Classes* (C_{20}) - elles-mêmes emboîtées dans les *Formes* du questionnaire (F_5) - et dans les *Ages* (A_2), les facteurs C et A étant croisés, ce que nous pouvons écrire $E_2<C_4<F_5>\times A_2>$.

A l'intérieur de chaque forme du questionnaire, les 12 questions concernent quatre *Domaines* des mathématiques (avec 3 questions par domaine): ensembles-relations, numération, opérations, découverte de l'espace (soit le facteur D_4), d'où l'emboîtement $Q_3<F_5\times D_4>$. Chaque élève, étant affecté à une forme du questionnaire, n'a donc été soumis qu'à 12 questions, à raison de 3 par domaine, soit 12 *occasions*.

Les facteurs E et D sont donc croisés, mais pas les facteurs E et Q, ce qui correspond à une structure de *blocs incomplets*). Ceci motive l'introduction d'un facteur *auxiliaire*, que nous appellerons *Réalisations*, à 3 modalités (R_3), qui exprime le fait que chaque élève est soumis à trois "réalisations" du facteur *Questions* dans chaque domaine. Les facteurs E, D et R sont croisés et le plan peut finalement être écrit

$$E_2<G_{40}>\times O_{12} \equiv E_2<C_4<F_5>\times A_2>\times D_4\times R_3$$

où l'on retrouve le facteur *Questions* sous la forme $Q_{60}\equiv F_5\times D_4\times R_3$.

Ici les facteurs *Elèves* et *Classes* peuvent tous les deux être considérés comme *aléatoires*, les classes ayant été choisies "au hasard" parmi les classes de même niveau des écoles suisses. Nous avons donc ici un exemple de plan à deux facteurs aléatoires. Les procédures développées dans cet ouvrage ne permettant de traiter qu'un seul facteur aléatoire à la fois, on pourra procéder à deux analyses selon des plans de structure $S<G>\times O$.

- La première analyse sera effectuée pour le plan caractérisé par

$$S\equiv E \qquad G_{40}\equiv C_4<F_5>\times A_2> \qquad O_{12}\equiv D_4\times R_3$$

où C est traité comme un facteur systématique; dans ce cas les conclusions inférentielles ne portent que sur les 20 classes de l'expérience - il n'y a pas de *généralisation* pour ce facteur.

• La seconde analyse sera effectuée pour le plan caractérisé par

$$S \equiv C \qquad G_5 \equiv F_5 \qquad O_{24} \equiv A_2 \times D_4 \times R_3$$

où on associe à chaque classe et chaque occasion la moyenne des deux élèves de la classe de l'âge considéré; dans ce cas on ne tient pas compte de la variabilité *intra-classes*.

On trouvera un exemple de procédures permettant de prendre en compte simultanément les deux facteurs aléatoires E et C dans Rouanet, Lépine et Holender (1978).

5.6.2 Analyses spécifiques

Le statut méthodologique du facteur *Réalisations* peut être pris explicitement en compte dans les analyses spécifiques: nous ne nous intéressons pas à l'effet global de R, mais à l'effet de R à l'intérieur de F (puisque d'une forme à l'autre les réalisations correspondent à des questions différentes). Les analyses spécifiques peuvent alors être organisées autour des cinq comparaisons suivantes

$$\mathbb{D}\ [2] \qquad \mathbb{F}\ [4] \qquad \mathbb{R}(F)\ [10] \qquad \mathbb{D}.\mathbb{F}\ [12] \qquad \mathbb{D}.\mathbb{R}(F)\ [30]$$

qui pourront ensuite être redécomposées en comparaisons partielles.

CHAPITRE 6

COMPARER DES MOYENNES SANS SUPPOSER L'ÉGALITÉ DES VARIANCES

Nous considérerons dans ce chapitre la généralisation du problème (dit de Behrens-Fisher) de la comparaison de deux moyennes de deux groupes d'observations indépendants sous le *modèle normal avec variances quelconques*. Nous envisagerons l'inférence sur une combinaison linéaire des moyennes, d'abord de deux groupes, puis de plusieurs groupes. Nous définirons des solutions fiducio-bayésiennes, abandonnant les suppositions d'homogénéité des variances des procédures usuelles de l'analyse de variance univariée. Ces solutions s'étendront de manière évidente aux plans de structure S<G>*O, grâce à l'approche de l'analyse spécifique développée aux chapitres §4 et §5.

Nous serons conduit à adopter sur le test de signification un autre point de vue que celui adopté dans le cadre fréquentiste, en ce qui concerne "l'élimination des paramètres parasites". Nous verrons que ce point de vue, que nous appelons *semi-bayésien*, rejoint en fait la conception de R.A. Fisher du test de Student.

6.1 Introduction aux procédures semi-bayésiennes

6.1.1 Le problème de Behrens-Fisher

Renoncer aux suppositions d'homogénéité des variances est une source d'embarras pour l'approche fréquentiste. Nous illustrerons ce problème à partir de l'exemple pharmacologique de la comparaison d'un nouveau médicament et d'un *placebo* introduit précédemment en §1.2. Il s'agit, rappelons-le, de l'inférence sur la moyenne générale de deux groupes indépendants de six patients chacun, avec la moyenne observée

$$d_{obs} = \tfrac{1}{2}\, d\varepsilon^1 + \tfrac{1}{2}\, d\varepsilon^2 = +6.722$$

et les écarts-types-corrigés observés respectifs

$$s_{g1} = 10.188 \text{ et } s_{g2} = 7.063$$

Nous pouvons dans une telle situation nous interroger sur la condition d'homogénéité des variances. Procéder à un "test préliminaire" de cette condition ne saurait être décisif, en raison des petites tailles d'échantillons qui correspondent à ces variances (sans parler ici des difficultés d'interprétation d'un résultat non-significatif...). Différentes solutions ont été considérées du point de vue de l'inférence fréquentiste pour traiter le problème de l'hétérogénéité des variances (ou plus généralement dans le cas multivarié des matrices de variances et covariances), en particulier par Satterhwaite (1946), Welch (1947, 1951), James (1951), Brown et Forsythe (1974), Nel et Van der Merwe (1986). Mais généralement la distribution d'échantillonnage des statistiques de test dépend des paramètres parasites. Par suite cette distribution ne peut pas être connue exactement (sinon asymptotiquement); elle engendre donc une solution qui ne peut pas être entièrement satisfaisante.

6.1.2 Test conditionnel au rapport des écarts-types

Compte tenu du fait qu'il n'y a que deux écarts-types en jeu, nous pouvons ici considérer leur rapport

$$\eta = \sigma_{g1}/\sigma_{g2}$$

Les observations du second groupe sont un échantillon d'une distribution de moyenne δ_{g2} et d'écart-type $\sigma_{g2}=\eta\times\sigma_{g1}$. Si la valeur de η est supposée connue, nous pouvons nous ramener à la solution usuelle avec égalité des variances, en divisant les observations du second groupe par η. En effet nous obtenons alors pour ce groupe un échantillon d'une distribution de moyenne δ_{g2}/η et d'écart-type $\sigma_{g2}/\eta=\sigma_{g1}$. Le paramètre $\delta=\frac{1}{2}\delta_{g1}+\frac{1}{2}\delta_{g2}$ est maintenant la combinaison linéaire des moyennes des deux groupes de coefficients $\frac{1}{2}$ et $\frac{1}{2}\eta$

$$\delta = \tfrac{1}{2}\,\delta_{g1} + \tfrac{1}{2}\,\eta\,\frac{\delta_{g2}}{\eta}$$

D'une manière générale nous nous intéressons à un paramètre δ de la forme

$$\delta = v_{g1}\delta_{g1} + v_{g2}\delta_{g2} = v_{g1}\delta_{g1} + (\eta v_{g2})\frac{\delta_{g2}}{\eta}$$

avec les coefficients v_{g1} et v_{g2} et les effectifs des groupes f_{g1} et f_{g2}. Après transformation des observations du second groupe comme ci-dessus, la statistique du test de l'hypothèse nulle $[H_0{:}\delta=0]$, *conditionnellement à* η, est d'observations calculée comme dans le cas "avec égalité des variances", en remplaçant simplement s_{g2} par s_{g2}/η et v_{g2} par $\eta\times v_{g2}$. En notant $t(\eta)$ cette statistique du test conditionnel à η, nous avons

$$t(\eta) = \frac{d}{b(\eta)s(\eta)} \quad [q = f_{g1}+f_{g2}-2\ dl]$$

$$\text{avec } b^2(\eta) = \frac{v_{g1}^2}{f_{g1}} + \frac{(\eta v_{g2})^2}{f_{g2}}$$

$$s^2(\eta) = \frac{(f_{g1}-1)s_{g1}^2 + (f_{g2}-1)(s_{g2}/\eta)^2}{f_{g1}-1 + f_{g2}-1}$$

Ceci inclut bien entendu, pour η=1, le cas particulier d'homogénéité des variances.

LeBayésien

Dans la fenêtre principale, cliquez sur le bouton **exemples**, et sélectionnez **médicament/placebo**. Pour pouvoir modifier les statistiques, cliquez sur le **feu rouge** qui "passe au vert" (n'effacez pas les données). Supposons par exemple η=0.7, qui est une valeur proche du rapport des écarts-types observés s_{g2}/s_{g1}=0.6933.

Cliquez sur **combinaison linéaire** et entrez les valeurs appropriées:

	effectif	moyenne	écart-type	coefficients
g1	6	6.86110950	10.18791045	1/2
g2	6	6.58333267/0.7	7.06300965/0.7	0.7/2

Puis cliquez sur le bouton **calculer**. Vous obtenez t = 2.661.

Nous pouvons vérifier que, pour les valeurs de η qui apparaissent "vraisemblables" au vu des données, la statistique $t\eta$) du test conditionnel varie relativement peu. Nous trouvons par exemple les valeurs

$$t(0.5) = 2.392 \quad t(0.6) = 2.565 \quad t(0.7) = 2.661$$

$$t(0.8) = 2.698 \quad t(0.9) = 2.692 \quad t(1.0) = 2.657$$

Nous pouvons vérifier que $t(0.6933)=t(1)=2.657$. En effet, dans le cas d'effectifs égaux nous avons $t(s_{g2}/s_{g1})=t(1)$, $t(1)$ étant la solution usuelle "avec égalité des variances". Ceci est évidemment une propriété importante du cas équilibré.

Nous retrouvons ici la propriété bien connue de robustesse du t de Student vis-à-vis de l'inégalité des variances, pourvu que les effectifs des groupes ne soient pas trop petits ou trop différents.

6.1.3 Le test de signification semi-bayésien: Principe général

C'est notre incertitude sur la vraie valeur de η qui est en jeu. Le fait que nous sommes enclins à penser que η doit être "relativement proche" du rapport observé 0.6933 est un argument important. Mais on voit mal comment il pourrait être explicité dans le cadre de l'approche fréquentiste.

■ *Test semi-bayésien*

L'idée qui vient à l'esprit d'un bayésien essayant de résoudre ce problème est de "remplacer" les écarts-types inconnus par une distribution de probabilité. Un principe raisonnable est de choisir une distribution initiale pour les écarts-types, et de considérer pour la statistique d
• la distribution bayésienne *prédictive*, sous H_0, conditionnelle aux écarts-types-corrigés intragroupes observés s_{g1} et s_{g2},
• au lieu de la distribution d'échantillonnage (conditionnelle à σ_{g1} et σ_{g2}).
C'est ce principe qui a été introduit dans Lecoutre (1985) sous le nom de *test de signification semi-bayésien.*

■ *Résultat significatif/non-significatif*

La distribution prédictive conditionnelle $d|(s_g)$ peut être utilisée d'une manière directe, pour procéder à un test de signification semi-bayésien de l'hypothèse nulle δ=0: ce test consiste à examiner si les données sont (dans un certain sens) compatibles avec cette distribution prédictive. Plus précisément, les données seront déclarées incompatibles avec cette distribution prédictive si la valeur observée d_{obs} est situé dans une zone inattendue ("résultat *significatif*").
Par exemple, pour le test unilatéral le résultat est significatif si pour un seuil fixé α ("petit")

$$[\text{sous } H_0]\ Pr(d>d_{obs}|(s_g)) \leq \alpha$$

Au contraire, les données seront déclarées compatibles avec cette distribution prédictive si d_{obs} est située dans une zone attendue, ou plutôt non inattendue, ("résultat *non-significatif*"), donc si

$$[\text{sous } H_0]\ Pr(d>d_{obs}|(s_g)) > \alpha$$

■ *Solution fiducio-bayésienne*

Le test de signification semi-bayésien nous permet de prendre explicitement en considération notre incertitude sur η. En particulier la solution standard, ou *fiducio-bayésienne*, utilise les distributions fiducio-bayésiennes, supposées indépendantes, relatives aux écarts-types σ_{g1} et σ_{g2}.

La distribution fiducio-bayésienne relative au paramètre dérivé η, montrée dans la figure 6.1, ne dépend en fait que du rapport s_{g2}/s_{g1}. Elle est proportionnelle à la racine-carrée d'une distribution F à 5 et 5 degrés de liberté, ce que nous noterons

$$\eta|s_{g1},s_{g2} \sim (s_{g2}/s_{g1})F^{1/2}_{f_{g_1}-1,f_{g2}-1} \sim 0.6933F^{1/2}_{5,5}$$

d'où par exemple $Pr(\eta>0.308) = 0.95$ et $Pr(\eta<1.558) = 0.95$

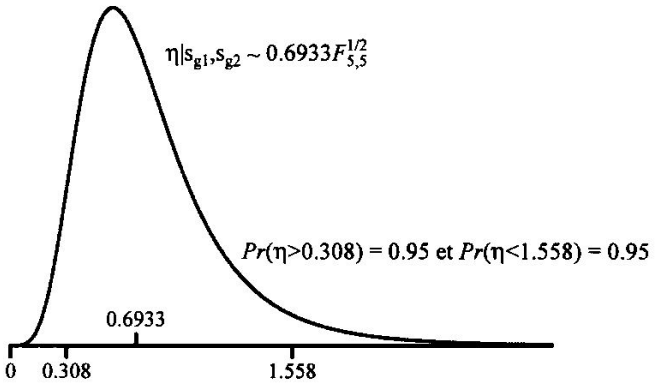


Figure 6.1 - Distribution fiducio-bayésienne relative au paramètre η

Le test de signification (semi-fiducio-bayésien) correspondant peut être regardé comme le mélange des tests conditionnels à η, pondérés par cette distribution.

6.1.4 Mise en œuvre pratique

■ *Comparaison du médicament et du placebo*

La distribution prédictive semi-fiducio-bayésienne de d, construite suivant le principe précédent, peut être approchée par une distribution t de *Student généralisée* ayant les même quatre premiers moments.

$$[\text{sous } H_0{:}\delta{=}0]\ \mathrm{d}\,|\,s_{\varepsilon 1}, s_{\varepsilon 2} \simeq t_{5.78}(0, 2.642^2)$$

LeBayésien

Revenez aux statistiques de base de l'exemple **médicament/placebo**, avec les coefficients [1/2 1/2] (option **moyenne équipondérée**). Cliquez sur l'option **sans égalité des variances**, puis sur le bouton **calculer**. Vous obtenez t=2.544 (avec 5.78 degrés de liberté).

Ainsi en pratique il sera possible de procéder au test fiducio-bayésien au moyen d'une distribution t, comme dans la procédure supposant l'égalité des variances. Le test semi-fiducio-bayésien conduit simplement à remplacer la constante b=0.2887 (1/√12) par $\tilde{b}$=0.3014 et le nombre de degrés de liberté q=10 par $\tilde{q}$=5.78, d'où l'*échelle*

$$\tilde{e} = \tilde{b}s = 0.2887 \times 8.766 = 2.642 \quad \text{(au lieu de } e = 2.530)$$

et la statistique de test

$$\tilde{t} = \frac{d}{\tilde{e}} = \frac{+6.722}{2.642} = +2.544 \quad [5.78\ dl]$$

Comme sous la supposition d'égalité des variances, le résultat est significatif, le seuil observé unilatéral passant cependant de $p/2$=0.012 à $p/2$=0.023.

■ *Effet du facteur Périodes*

Nous pouvons de la même manière tester l'hypothèse d'un effet du facteur *Périodes* nul, soit $\frac{1}{2}\delta_{\varepsilon 1} - \frac{1}{2}\delta_{\varepsilon 2} = 0$ (↳1.2.5).

LeBayésien

Pour les statistiques de base de l'exemple **médicament/placebo**, entrez les coefficients [1/2 -1/2] (option **combinaison linéaire**). Effectuez le calcul pour chacune des deux options **avec** et **sans égalité des variances.**

Le test semi-fiducio-bayésien ne supposant pas l'égalité des deux variances est construit de la même manière que précédemment

$$\tilde{t} = \frac{\frac{1}{2}d\varepsilon^1 - \frac{1}{2}d\varepsilon^2}{\tilde{e}} = \frac{+0.139}{2.642} = +0.053 \quad [5.78\ dl]$$

et donne un résultat très voisin de celui de la solution "avec égalité des variances" (t=0.055, 10 dl), le seuil observé unilatéral étant $p/2$=0.480 au lieu $p/2$=0.479.

Dans la section ↳6.4.1 sont explicitées les formules correspondant à ce cas particulier de deux groupes indépendants avec effectifs égaux. On remarquera que l'on retrouve ici la solution dite de Behrens-Fisher pour la comparaison de deux groupes indépendants avec variances quelconques (Fisher, 1935). Mais dans le cadre bayésien cette solution est habituellement justifiée en probabilisant également le paramètre d'intérêt δ (par exemple Box et Tiao, 1973).

6.1.5 Quelques commentaires généraux

L'idée de base de remplacer les paramètres "parasites" inconnus par une distribution de probabilité n'est certainement pas une idée nouvelle. Ainsi Fisher donnait dans son ouvrage *Statistical methods for research workers* la présentation suivante du test élémentaire du t de Student (Fisher, 1990, page 118):

"If x (for example the mean of a sample) is a value normally distributed about zero, and σ is its true standard error, then the probability that x/σ exceeds any specified value may be obtained from the appropriate table of the normal distribution; but if we do not know σ, but in its place have s, an estimate of the value of σ, the distribution required will be that of x/s, and this is not normal. The true value has been divided by a factor s/σ which introduces an error. (...) the distribution of s/σ is calculable, and although σ is unknown, we can use in its place the fiducial distribution of σ given s to find the probability of x exceeding a given multiple of s".

Assez curieusement cette présentation, peut-être parce qu'elle a été exprimée dans le cadre fiduciaire, ne semble pas avoir beaucoup retenu l'attention. Pourtant, reconsidérée dans le cadre bayésien, partant d'une distribution initiale pour σ, une telle présentation conduit directement au test de signification semi-bayésien.

L'approche semi-bayésienne nous amène à renoncer à une interprétation fréquentiste stricte en termes de distribution d'échantillonnage. En échange elle fournit un apport conceptuel et technique majeur.

■ *Conceptuellement*

Le test de signification semi-bayésien conduit à des procédures intermédiaires entre les tests de signification usuels et les procédures bayésiennes, puisqu'il consiste à probabiliser seulement l'espace des paramètres "parasites". Il préserve ainsi le principe fondamental de l'inférence fréquentiste pour le paramètre principal δ, et peut donc être considéré comme un amendement technique des tests usuels.

Dans le cadre bayésien, le statut du test semi-bayésien, défini par une distribution prédictive, est parfaitement clair. En outre, la procédure peut recevoir une justification en termes d'exhaustivité partielle. Ainsi, dans l'exemple précédent du test élémentaire du t de Student, la "G-exhaustivité" (Barnard, 1963) ou la "H-exhaustivité" (Hájek, 1965) font apparaître que la statistique s contient toute l'information disponible sur σ en l'absence (ou l'ignorance comme cela est le cas dans le test usuel) de connaissance *a priori* sur la moyenne vraie: voir respectivement Dawid (1980) et Basu (1977).

■ *Techniquement*

Le test de signification semi-bayésien englobe les résultats des tests de signification usuels et les étend de manière substantielle, puisqu'il apporte un principe général, facile à justifier, pour traiter l'information disponible sur les paramètres parasites. Clairement l'approche semi-bayésienne peut être appliquée à de nombreux problèmes qui se posent en analyse de variance, et apparaît particulièrement fructueuse pour éliminer les restrictions de validité arbitraires liées aux paramètres variances.

■ *Choix d'une distribution initiale*

La solution semi-fiducio-bayésienne apparaît privilégiée, dans la mesure où, dans la situation de base, avec égalité des variances, elle coïncide avec la solution fréquentiste usuelle.

Mais il n'y a à l'évidence aucune raison de restreindre le test de signification semi-bayésien à l'un des deux cas extrêmes habituellement distingués dans l'approche fréquentiste: variances connues opposé à variances inconnues en dehors des données. Nous avons une solution immédiate pour les cas intermédiaires, puisque nous pouvons incorporer d'autres sources d'information sur les variances en choisissant une distribution initiale appropriée (même si nous n'illustrons pas cette possibilité dans cet ouvrage).

6.1.6 Du test de signification semi-fiducio-bayésien à l'inférence fiducio-bayésienne

Si nous admettons le point de vue semi-bayésien, nous avons déjà résolu une grande partie du problème pour passer à une inférence bayésienne complète: le test semi-fiducio-bayésien jouera le même rôle d'intermédiaire technique et conceptuel que le test usuel dans la situation avec égalité des variances.

LeBayésien

Revenez aux statistiques de base de l'exemple **médicament/placebo**, avec les coefficients [1/2 1/2] (option **moyenne équipondérée**).

Cliquez sur l'option **sans égalité des variances**, puis sur le bouton **ok**, ce qui vous amène à la fenêtre principale. Il n'y a dans ce cas que la courbe relative à l'effet δ. Cliquez sur cette courbe, puis dans la fenêtre des énoncés, effectuez le calcul des limites pour la garantie 0.90. Vous obtenez par exemple dans la courbe du bas (**notable**): δ>2.90.

Ainsi, dans la comparaison du médicament et du *placebo* sans supposer l'égalité des variances, nous pouvons déduire immédiatement l'approximation de la distribution fiducio-bayésienne relative à δ des résultats précédents: il s'agit d'une distribution t généralisée, centrée sur d_{obs}=+6.722, d'échelle $\tilde{e}$=2.642, et avec $\tilde{q}$=5.78 degrés de liberté (les quatre premiers moments sont ceux de la distribution exacte)

$$\delta \simeq t_{\tilde{q}}(d,\tilde{e}^2) \simeq t_{5.78}(+6.722,2.642^2)$$

Cette distribution, montrée dans la figure 6.2, est seulement un peu plus dispersée que celle obtenue sous la supposition d'égalité des variances.

Nous pouvons réinterpréter le seuil observé du test semi-fiducio-bayésien de la même manière que pour le test usuel

$$Pr(\delta>0) = 1-p/2 = 0.977$$

Pour une recherche de conclusion d'effet notable, nous trouvons

$$Pr(\delta>+2.90) = 0.90$$

au lieu de $Pr(\delta>+3.25)=0.90$ pour la solution avec égalité des variances.

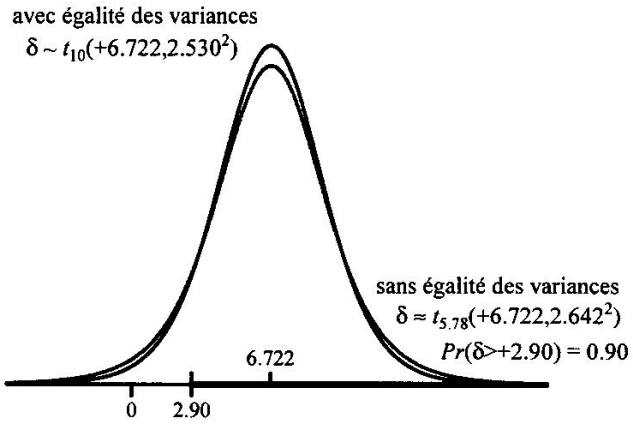


Figure 6.2 - Comparaison médicament/*placebo*: Distributions fiducio-bayésiennes relatives à δ

6.2 Combinaisons linéaires de moyennes de groupes indépendants

6.2.1 Données psychométriques: Milieu et Pédagogie de l'enseignement des mathématiques

Reprenons l'enquête "*Milieu et Pédagogie de l'enseignement des mathématiques*", considérée aux chapitres §2 et §5. Rappelons que les éléves étaient répartis en quatre groupes (d'effectifs voisins) définis par le croisement de deux facteurs à deux niveaux chacun

- Facteur *Milieu*: *favorisé* opposé à *défavorisé*
- Facteur *Pédagogie*: *moderne* opposé à *traditionnelle*

Pour les données complètes analysées au chapitre ↳2, compte tenu des effectifs élevés, le problème de la non homogénéité des variances n'était guère crucial. Nous considérerons donc ici les données relatives à 33 élèves présentées au chapitre ↳5. Pour ces données, nous avons le tableau d'analyse de variance (pour la solution *équipondérée*)

source de variation		*dl*	*sc*	*cm*	*F* [dl]	*p*
Milieu	$\mathbb{M}_{[E]}$	1	3.916	3.916	2.371 [1,29]	0.134
Pédagogie	$\mathbb{P}_{[E]}$	1	11.490	11.490	6.955 [1,29]	0.013
Interaction	$\mathbb{M}.\mathbb{P}$	1	0.048	0.048	0.029 [1,29]	0.866
intra	$\mathbb{S}(G)$	29	47.905	1.652		

Intéressons-nous à l'effet du facteur *Pédagogie*, qui est le seul résultat significatif en supposant l'égalité des variances. L'effet observé est défini par la différence entre les deux moyennes

pédagogie moderne (5.164+4.840)/2 = 5.227
pédagogie traditionnelle (4.343+3.722)/2 = 4.033

(nous prenons ici les moyennes équipondérées, sans que cela affecte la généralité de la procédure)

soit d = 5.227 - 4.033 = +1.195

La différence parente correspondante δ est la combinaison linéaire $\Sigma_g v_g \mu_g$ des moyennes des quatre groupes, définie par les coefficients $v_{g1}=+1/2$, $v_{g2}=+1/2$, $v_{g3}=-1/2$, $v_{g4}=-1/2$.

LeBayésien

Dans la fenêtre principale, cliquez sur le bouton **exemples**, et sélectionnez **milieu et pédagogie - données de 33 élève**. Effectuez les calculs pour chacune des deux options, **avec** et **sans égalité des variances**...

6.2.2 En supposant l'égalité des variances

Si nous supposons l'égalité des variances des quatre groupes, nous utilisons la statistique variance-corrigée intragroupe moyenne s^2

$$s^2 = \Sigma(f_g-1)s_g^2 / \Sigma(f_g-1) = 1.285^2$$

à laquelle sont associés q=6+9+6+8=29 degrés de liberté. s^2 est encore ici le carré-moyen intragroupe $cm_{S(G)}$.

Nous avons l'échelle

$$e = bs = \sqrt{313/2520} \times 1.285 = 0.453$$

$$\text{avec } b^2 = \sum_{g\in G} \frac{v_g^2}{f_g} = \frac{1}{28} + \frac{1}{40} + \frac{1}{28} + \frac{1}{36} = \frac{313}{2520} = 0.3524^2$$

d'où la statistique de test

$$t = \frac{d}{e} = \frac{+1.195}{0.453} = +2.637 \quad [29\ dl]$$

qui est la racine carrée du rapport F

$$F = t^2 = cm_P/cm_{S(G)} = 6.955 \quad [1 \text{ et } 29\ dl]$$

et la distribution fiducio-bayésienne relative à δ

$$\delta \sim t_{29}(+1.195, 0.453^2)$$

6.2.3 En ne supposant pas l'égalité des variances

Nous pouvons dans ce cas généraliser la méthode illustrée précédemment dans le cas particulier de deux variances. Les calculs sont détaillés dans la section ↳6.4.2, et conduisent pour le test de signification au t de Student "approché"

$$\tilde{t} = \frac{d}{\tilde{e}} = \frac{+1.195}{0.505} = +2.637 \quad [11.88\ dl]$$

Cette solution revient à remplacer la constante b=0.3524 par $\tilde{b}$=0.3932

$$\tilde{e} = \tilde{b}s = 0.3931 \times 1.285 = 0.505$$

et le nombre de degrés de liberté q=29 par $\tilde{q}$=11.88

Nous pouvons donc également définir le rapport F "approché"

$$\tilde{F} = \tilde{t}^2 = cm_P/1.481 = 5.588 \quad [1 \text{ et } 11.88\ dl]$$

obtenu en remplaçant le carré-moyen $cm_{S(G)}$=1.652 par la valeur 1.481.

Nous en déduisons l'approximation de la distribution fiducio-bayésienne relative à δ

$$\delta \simeq t_{\tilde{q}}(d,\tilde{e}^2) \simeq t_{11.88}(+1.195,0.505^2)$$

Les deux résultats sont significatifs: le seuil observé unilatéral est $p/2$ =0.007 pour la solution avec égalité des variances, et $p/2$=0.018 pour la solution sans égalité des variances. Nous obtenons les énoncés fiducio-bayésiens de recherche d'une conclusion d'effet notable respectifs

$$Pr(\delta>+0.60) = 0.90 \text{ et } Pr(\delta>+0.51) = 0.90$$

■ *Remarque: effets calibrés*

Dans ce cas nous pourrions encore chercher une inférence sur un effet calibré par un "écart-type moyen" σ. Malheureusement, que nous prenions pour σ^2 la moyenne équipondérée ou une moyenne pondérée par les effectifs ou par les nombres de degrés de liberté des variances des différents groupes, σ^2 n'est pas proportionnelle à la variance de la distribution d'échantillonnage de d. La distribution relative à δ/σ devient par conséquent beaucoup plus complexe. On pourra cependant toujours dans cette situation rapporter δ à l'écart-type-corrigé intragroupe moyen s.

6.3 Un résultat plus significatif en ne supposant pas l'égalité des variances

On pourrait croire que la solution qui ne suppose pas l'égalité des variances est "pénalisante", c'est-à-dire fournit systématiquement un résultat moins significatif, ou de manière plus éclairante une distribution fiducio-bayésienne relative à δ plus dispersée. Ceci est vrai dans le cas où l'on compare deux parties d'effectifs égaux. Mais en dehors de ce cas, le contraire peut parfaitement se produire, comme le montre l'exemple suivant de la comparaison de deux groupes indépendants d'effectifs respectifs 15 et 6. Nous avons, pour chacun de ces deux groupes, les moyennes et écarts-types-corrigés respectifs

	effectif	moyenne	écart-type
g1	15	22.2667	10.8659
g2	6	14.5000	1.0488

LeBayésien

Dans la fenêtre principale, cliquez sur le bouton **exemples**, et sélectionnez **exemple plus significatif "sans égalité des variances"**. Effectuez les calculs pour chacune des deux options, **avec** et **sans égalité des variances**...

■ *En supposant l'égalité des variances*

L'écart-type intragroupe moyen

$$s = \left(\frac{14\times10.8659^2 + 5\times1.0488^2}{14 + 5}\right)^{1/2} = 9.343$$

accorde un poids supérieur au groupe $g1$, et est par suite proche de s_{g1} qui est le plus grand des deux écarts-types. Il en résulte que le test donné par

$$t = \frac{d}{e} = \frac{+7.767}{4.513} = +1.721 \quad [19\ dl]$$

$$\text{avec } b^2 = \frac{(+1)^2}{15} + \frac{(-1)^2}{6} = \frac{7}{30} = 0.4830^2$$

$$\text{d'où } e = bs = 0.4830\times9.343 = 4.513$$

est non-significatif aux seuils usuels ($p/2$=0.051).

La distribution fiducio-bayésienne de δ est

$$\delta \sim t_{19}(+7.767, 4.513^2)$$

■ *En ne supposant pas l'égalité des variances*

La solution revient à remplacer e=4.513 par $\tilde{e}$=2.861 et q=19 par $\tilde{q}$=14.56. Cette fois le test est significatif: $\tilde{t}$=d/$\tilde{e}$=+2.715 [14.56 *dl*] d'où $p/2$=0.008.

Cette solution équivaut encore à remplacer s=9.343 par $\tilde{s}$=5.923, qui est en fait voisin de la moyenne arithmétique équipondérée des écarts-types de chacun des deux groupes, (10.8659+1.0488)/2=5.957

$$\tilde{e} = b\tilde{s} = 0.4830\times5.923 = 2.861$$

La distribution fiducio-bayésienne de δ est sensiblement moins dispersée que dans le cas précédent

$$\delta \simeq t_{14.56}(7.767, 2.861^2)$$

■ *L'influence des effectifs*

LeBayésien

Dans l'exemple précédent modifiez les effectifs 15 et 6 des groupes (cliquez au préalable sur le **feu rouge**), successivement en 6 et 15, 15 et 15, 6 et 6. Effectuez dans chaque cas les calculs pour chacune des deux options, **avec** et **sans égalité des variances**...

Comme on pouvait s'y attendre, si on permute les effectifs, le résultat est moins significatif "sans égalité des variances"; nous trouvons respectivement

t = +2.848 [19 dl] $p/2$=0.005 et $\tilde{t}$ = +1.748 [5.01 dl] $p/2$=0.070

Par contre, si les deux effectifs étaient égaux à 15, les deux résultats seraient très proches

t = +2.756 [29 dl] $p/2$=0.005 et $\tilde{t}$ = +2.753 [14.19 dl] $p/2$=0.007

Et même pour deux effectifs égaux à 6, les deux résultats seraient encore étonnamment proches

t = +1.743 [10 dl] $p/2$=0.056 et $\tilde{t}$ = +1.741 [5.02 dl] $p/2$=0.071

6.4 Formules de calcul des approximations

6.4.1 Comparaison de deux moyennes de groupes indépendants d'effectifs égaux

Notons $\bar{f}$ l'effectif commun des groupes

et $F^{12} = s^2_{g1}/s^2_{g2}$ le rapport des deux variances observées

$$\tilde{q} = \bar{f}-1+\frac{2}{F_{12}+\frac{1}{F_{12}}}(\bar{f}-5) \quad \text{et} \quad \tilde{c} = \left(1+\frac{2}{F_{12}+\frac{1}{F_{12}}}\frac{\bar{f}-5}{\bar{f}-3}\right)\Big/\left(1+\frac{2}{F_{12}+\frac{1}{F_{12}}}\frac{\bar{f}-5}{\bar{f}-1}\right)$$

- pour l'inférence sur la moyenne générale
  ou sur la demi-différence des moyennes

$$\tilde{e}^2 = \frac{s^2}{2\bar{f}}\tilde{c}$$

- pour l'inférence sur la différence des moyennes

$$\tilde{e}^2 = \frac{2s^2}{\bar{f}}\tilde{c}$$

On remarquera les propriétés suivantes:

$$\tilde{e}^2 \to e^2 \text{ quand } F_{12} \to 0 \text{ ou } F_{12} \to \infty$$

$$\tilde{e}^2 \to \frac{\bar{f}-4}{\bar{f}-3}e^2 \text{ quand } F_{12} = 1$$

où e est le dénominateur du test t
pour la solution "avec égalité des variances"

■ *Application numérique*

Dans l'exemple traité en ↳6.1.4, nous avons

$$\bar{f} = 6 \text{ et } F^{12} = 2.0806 \text{ d'où } \tilde{q} = 5.781 \text{ et } \tilde{c} = 1.090$$

- pour l'inférence sur la moyenne générale $\tilde{e}^2 = 2.642^2$
- pour l'inférence sur la différence des moyennes $\tilde{e}^2 = 5.284^2$

6.4.2 Inférence sur une combinaison linéaire de moyennes de groupes indépendants d'effectifs quelconques

Les effectifs des groupes sont f_g d'où $q_g = f_g - 1$

Les coefficients de la combinaison linéaire sont v_g

$$\text{Posons } c_g^2 = \frac{v_g^2}{f_g}$$

$$M = \sum_{g \in G} c_g^2 \frac{q_g}{q_g - 2} s_g^2 \quad \text{et} \quad V = \sum_{g \in G} c_g^4 \frac{2q_g^2}{(q_g-2)^2(q_g-4)} s_g^4$$

▸ L'approximation est caractérisée par

$$\tilde{q} = 4 + 2\frac{M^2}{V}$$

$$\tilde{e}^2 = \frac{\tilde{q}-2}{\tilde{q}} M^2$$

■ *Application numérique*

Dans l'exemple traité en §6.2.3, nous avons

$$v_{g1} = v_{g2} = v_{g3} = v_{g4} = \tfrac{1}{2}$$

$$f_{g1} = 7 \quad f_{g2} = 10 \quad f_{g3} = 7 \quad f_{g4} = 9$$

$$c_{g1}^2 = 1/28 \quad c_{g2}^2 = 1/40 \quad c_{g3}^2 = 1/28 \quad c_{g4}^2 = 1/36$$

$$s_{g1}^2 = 1.435 \quad s_{g2}^2 = 1.092 \quad s_{g3}^2 = 2.276 \quad s_{g4}^2 = 1.977$$

$$M = \frac{2.152}{28} + \frac{1.403}{40} + \frac{3.414}{28} + \frac{2.636}{36} = 0.307$$

$$V = \frac{4.632}{28^2} + \frac{0.788}{40^2} + \frac{11.657}{28^2} + \frac{3.474}{36^2} = 0.024$$

▸ D'où $\tilde{q} = 11.88$ et $\tilde{e}^2 = 0.505^2$

CHAPITRE 7

COMPLÉMENTS SUR LES PROCÉDURES FRÉQUENTISTES

Dans le cas de recherche d'une conclusion sur l'importance de l'effet δ, il s'agit, dans le cadre fréquentiste, de tester une hypothèse composée. Des solutions sont disponibles, et seront illustrées dans la première partie de ce chapitre. Elles peuvent aussi bien s'appliquer à l'effet δ qu'à l'effet calibré δ/σ. Dans le cas de l'effet *notable*, les procédures fréquentistes et fiducio-bayésiennes conduisent à la même conclusion (à la différence d'interprétation près); en revanche, le cas de l'effet *négligeable* apparaît comme un problème crucial dans le choix entre les deux types de procédures.

Dans la deuxième partie de ce chapitre, nous verrons que le test d'une hypothèse composée soulève des difficultés dans le cadre de l'inférence d'échantillonnage. Les principes qui distinguent les différentes procédures concernent le paramètre δ, et il suffira donc ici de les discuter dans le cas où la variance σ^2 est connue. Nous considérerons successivement le problème de recherche d'une conclusion d'effet négligeable, puis celui de recherche d'une conclusion d'effet non négligeable, et nous discuterons dans chaque cas deux procédures d'échantillonnage envisageables.

7.1 ILLUSTRATIONS

7.1.1 Effet notable

■ *Test unilatéral de notabilité*

Si nous voulons montrer par un test de signification (au seuil α) que la différence δ est supérieure à α>0 (ou inférieure à α<0), nous devons tester l'hypothèse nulle *composée* (que nous cherchons à rejeter) $\delta<\alpha$ (*resp.* $\delta>x$). La solution consiste à tester l'hypothèse nulle ponctuelle $\delta=\alpha$ (soit $\delta-\alpha=0$) au

seuil *unilatéral* α. Il suffit pour cela de remplacer d-0 par d-α, d'où pour les notations générales utilisées dans les chapitres précédents la statistique de test

$$t^{[\alpha]} = \frac{d-\alpha}{e} \quad \text{avec } e = bs$$

et le seuil observé correspondant

$$p[H_0:\delta<\alpha] = Pr(t_q>t^{[\alpha]}) \quad \text{ou} \quad p[H_0:\delta>\alpha] = Pr(t_q<t^{[\alpha]})$$

■ *L'exemple de Student*

Reprenons l'exemple de Student introduit au chapitre ↳1, pour lequel nous avons d=+1.580, e=0.389 et q=9, et considérons l'hypothèse nulle $\delta<1/2$, d'où la statistique de test

$$t^{[1/2]} = \frac{1.580-0.500}{0.389} = +2.777$$

LeBayésien

Sélectionnez l'exemple **Student** (voir le chapitre ↳1) et placez vous dans la fenêtre des énoncés pour la distribution fiducio-bayésienne de δ.

Activez la case à cocher **procédures fréquentistes effet négligeable/notable**, ainsi que s'il y a lieu l'option **limite**. Entrez la limite 1/2 (pas de deuxième limite) et cliquez sur le bouton **calculer**. Dans le cadre des procédures fréquentistes, vous obtenez (avec 3 décimales pour les probabilités) les deux dernières lignes

test: notable? $H_0:\delta<0.500$
t=2.777 p=0.011

(les deux premières lignes fournissent un test dit de "non négligeabilité": voir ↳7.1.2).

Remarquez pour la courbe du bas (effet notable) la garantie fiducio-bayésienne associée à $\delta>0.500$: 0.989 (=1-0.011).

Nous pouvons rejeter l'hypothèse nulle au seuil observé $p[H_0:\delta<1/2]=0.011$.

■ *Du test usuel au test de notabilité*

D'une manière générale, il est immédiat de procéder au test de notabilité de l'effet, si nous connaissons d et la valeur t_{obs} (ou F_{obs}) observée pour la

statistique de test usuelle de l'hypothèse nulle H_0:δ=0, puisque

$$t^{[\alpha]} = \frac{d-\alpha}{d} t_{obs} = \frac{d-\alpha}{|d|}\sqrt{F_{obs}} \quad \text{(pourvu que } d\neq 0)$$

Ainsi, dans l'exemple de Student d=+1.580, α=1/2 et t_{obs}=+4.062, d'où

$$t^{[1/2]} = \frac{+1.080}{1.580} 4.062 = +2.777$$

■ *Réinterprétation fiducio-bayésienne*

Nous avons ici une généralisation de la réinterprétation fiducio-bayésienne du seuil unilatéral du test usuel considérée en §1.2.3 (qui correspond au cas particulier α=0)

$$Pr(\delta>\alpha) = 1-p[H_0:\delta<\alpha] \quad \text{ou} \quad Pr(\delta<\alpha) = 1-p[H_0:\delta>\alpha]$$

$$\text{ici } Pr(\delta>+0.5) = 1-0.011 = 0.989$$

■ *Intervalle de confiance (effet notable)*

Nous pouvons construire un intervalle (ou plutôt un demi-intervalle) de confiance 1-α pour δ comme l'ensemble des α_0 tels que l'hypothèse nulle $\delta<\alpha_0$ (ou $\delta>\alpha_0$) est rejetée au seuil α. L'intervalle $[\alpha,+\infty[$ (ou $]-\infty,\alpha]$) ainsi obtenu coïncide avec l'intervalle de crédibilité fiducio-bayésien donné par l'énoncé

$$Pr(\delta>\alpha) = 1-\alpha \quad (resp.\ Pr(\delta<\alpha) = 1-\alpha)$$

LeBayésien

Activez l'option **garantie** et sélectionnez la garantie 0.95, puis cliquez sur **calculer**. Dans le cadre des procédures fréquentistes, la dernière ligne donne l'intervalle de confiance 0.95

pour δ: notable? [0.867,infini[

Effectuez maintenant le calcul pour la **limite** 0.867: vous obtenez le test correspondant

test: notable? H_0:δ<0.867

t=1.833 p=0.05

Nous obtenons ici l'intervalle de confiance 0.95 [+0.867,+∞[, où +0.867 est la valeur "juste significative" au seuil 1-0.95=0.05, donc telle que

$$p[H_0:\delta<+0.867] = 0.05$$

7.1.2 Effet non négligeable

■ *Test bilatéral de non négligeabilité*

Un autre test envisageable concerne l'hypothèse nulle composée $|\delta|<\alpha$ (avec $\alpha\geq 0$). Il s'agit dans ce cas de chercher à rejeter l'hypothèse d'un effet négligeable, c'est pourquoi ce test sera appelé "test de non négligeabilité". Ce test, qui est la généralisation du test *bilatéral* usuel, est moins intéressant que le précédent, puisqu'il ne permet pas de se prononcer sur le sens de l'effet; mais c'est lui qui se prêtera aux extensions multidimensionnelles.

La solution consiste (nous supposons bien entendu $|d|>\alpha$) à tester chacune des hypothèses nulles ponctuelles $\delta=\alpha$ et $\delta=-\alpha$ au seuil *bilatéral* α, et à déclarer le résultat significatif (au seuil α) si chacun des ces deux tests est significatif. En pratique, il suffit donc de considérer la statistique de test (avec les notations précédentes)

$$t_{min}^{[\alpha]} = min(|t^{[\alpha]}|, |t^{[-\alpha]}|) = \frac{||d|-\alpha|}{e}$$

$$= \frac{||d|-\alpha|}{d} t_{obs}| = \frac{||d|-\alpha|}{|d|}\sqrt{F_{obs}} \quad (\text{si } d\neq 0)$$

et le seuil observé correspondant

$$p[H_0: |\delta|<\alpha] = Pr(|t_q|>t_{min}^{[\alpha]})$$

■ *L'exemple de Student*

Considérons l'hypothèse nulle $|\delta|<1/2$.

LeBayésien

Activez l'option **limite** et entrez la valeur 1/2, puis cliquez sur **calculer**. Dans le cadre des procédures fréquentistes, vous obtenez (avec 3 décimales pour les probabilités) les deux premières lignes

test: non négligeable? $H_0: |\delta|<0.500$
t=2.777 p=0.022

La statistique de test vaut

$$t_{min}^{[1/2]} = \frac{|1.580-0.500|}{0.389} = 2.777$$

et nous pouvons rejeter l'hypothèse nulle au seuil observé

$$p[H_0:|\delta|<1/2] = 0.022 = 2p[H_0:\delta<1/2]$$

■ *Réinterprétation fiducio-bayésienne*

Nous avons une généralisation immédiate de la réinterprétation fiducio-bayésienne du seuil bilatéral du test usuel considérée en ↳1.2.3 (qui correspond au cas particulier $\alpha=0$)

$$Pr(\alpha<\delta<2d-\alpha) = 1-p[H_0:|\delta|<\alpha] \quad (\text{si } d>0)$$
$$Pr(2d-\alpha<\delta<-\alpha) = 1-p[H_0:|\delta|<\alpha] \quad (\text{si } d<0)$$

LeBayésien

Activez l'option **deux limites**, et entrez la deuxième limite 2.66 (=2×1.580-0.5), puis cliquez sur **calculer**. Vous obtenez dans la courbe du haut la garantie fiducio-bayésienne 0.978 pour l'intervalle [0.5,2.66].

Ici $Pr(0.5<\delta<2.56) = 1-0.022 = 0.978$

Du point de vue fiducio-bayésien, cet énoncé permet une conclusion d'effet notable ($\delta>1/2$), mais il exclut les valeurs supérieures à +2.56.

■ *Caractérisation équivalente à partir de l'intervalle de confiance usuel*

Comme le montre la réinterprétation précédente, ce test revient à calculer l'intervalle de confiance $1-\alpha$ usuel pour δ (centré sur la valeur observée d), et à rejeter l'hypothèse nulle $|\delta|<\alpha$ au seuil α si cet intervalle n'a aucun point commun avec l'intervalle $[-\alpha,\alpha]$.

■ *Intervalle de confiance pour $|\delta|$ (effet non négligeable)*

Nous pouvons construire un intervalle de confiance $1-\alpha$ pour $|\delta|$ comme l'ensemble des $\alpha_0>0$ tels que l'hypothèse nulle $|\delta|<\alpha_0$ est rejetée au seuil α.

En raison des considérations précédentes, si les deux limites de l'intervalle de confiance $1-\alpha$ usuel pour δ sont de même signe, la borne α de l'intervalle $[\alpha,+\infty[$ ainsi obtenue est la plus petite en valeur absolue de ces deux limites. Du point de vue fiducio-bayésien, la probabilité associée à cet intervalle est supérieure à $1-\alpha$. Nous obtenons ainsi ici pour $|\delta|$ l'intervalle de confiance 0.95: $[0.700,+\infty[$; et la probabilité fiducio-bayésienne correspondante

$$Pr(|\delta|>0.700) = 0.9751$$

est voisine de $Pr(\delta>+0.700)=0.975$.

Si les deux limites de l'intervalle de confiance usuel sont de signes différents, l'intervalle de confiance pour $|\delta|$ est l'ensemble $\mathbb{R}_+$ tout entier (une solution "fâcheuse" pour un intervalle de confiance $1-\alpha$!). Dans ce cas aucune valeur α_0 inférieure à $|d|$ n'est significative au seuil α pour le test de non négligeabilité.

LeBayésien

Activez l'option **garantie** et sélectionnez la garantie 0.95, puis cliquez sur **calculer**. Dans le cadre des procédures fréquentistes, vous obtenez sur la troisième ligne l'intervalle de confiance 0.95 pour $|\delta|$

pour δ: non négligeable? [0.700,infini[

et dans la courbe du haut l'intervalle de crédibilité fiducio-bayésien [0.700,2.460] (qui coïncide avec l'intervalle de confiance usuel pour δ).

Effectuez le calcul pour la **limite** 0.700: vous obtenez sur la courbe du milieu (négligeable) la garantie 0.0249 (avec 4 décimales), soit

$$Pr(|\delta|>0.700) = 1-0.0249 = 0.9751$$

Effectuez maintenant le calcul pour la **garantie** 0.999: vous obtenez cette fois l'intervalle de confiance

pour δ: non négligeable? [0,infini[

et, dans la courbe du haut, l'intervalle de crédibilité fiducio-bayésien [-0.280,3.440], dont les deux bornes sont de signes différents.

7.1.3 Effet négligeable ("équivalence")

Si nous voulons montrer par un test de signification que la différence δ est inférieure en valeur absolue à α, nous devons tester l'hypothèse nulle *composée* (que nous cherchons à rejeter) $|\delta|>\alpha$.

■ *Procédure des "deux seuils unilatéraux"*

La solution consiste (nous supposons bien entendu $|d|<\alpha$) à tester chacune des hypothèses nulles ponctuelles $\delta=\alpha$ et $\delta=-\alpha$ au seuil *unilatéral* α, et à déclarer le résultat significatif (au seuil α) si chacun de ces deux tests est significatif. En pratique, il suffit donc de considérer la même statistique de test que pour le test de non négligeabilité (↳7.1.2)

$$t^{[\alpha]}_{min} = min(|t^{[\alpha]}|, |t^{[-\alpha]}|) = \frac{||d|-\alpha|}{e}$$

$$= \frac{||d|-\alpha|}{d} t_{obs} = \frac{||d|-\alpha|}{|d|}\sqrt{F_{obs}} \quad \text{si } d \neq 0)$$

mais avec ici le seuil observé

$$p[H_0: |\delta| > \alpha] = Pr(t_q > t^{[\alpha]}_{min})$$

■ *Exemple: Une étude de bioéquivalence*

Considérons les données d'une étude typique de bioéquivalence fournies par Steinjans et Diletti (1983). Il s'agit d'un plan *cross-over* 2×2 (*cf* §5.3.1), avec 12 patients. Le traitement t_1 est comparé au traitement *de référence* t_2; 6 patients sont affectés à la séquence t_1-t_2 (groupe g_1) et 6 à la séquence t_2-t_1 (groupe g_2). Une telle étude a pour objectif de montrer que deux traitements ont la même activité pharmacologique, en comparant l'évolution de leurs concentrations plasmatiques. La variable la plus souvent considérée est le logarithme de l'aire sous la courbe de ces concentrations.

Il est habituel de définir l'intervalle de bioéquivalence "à p%" comme l'intervalle centré sur 0 pour la différence des logarithmes

$$\left[log(\frac{100-p}{100}) , log(\frac{100}{100-p}\right]$$

d'où par exemple l'intervalle de bioéquivalence à 10%, que nous considérerons ici

$$[log(0.90), log(1/0.90)] = [-0.105361, +0.105361]$$

Le tableau suivant donne pour chaque sujet, les valeurs observées pour t_1 et t_2, leurs logarithmes, et la différence de ces logarithmes, qui constituent ici les données *pertinentes*.

Les auteurs fournissent également un tableau d'analyse de variance, avec pour la comparaison des deux traitements les deux carrés-moyens respectifs pour le rapport F

numérateur	0.0101701	[1 *dl*]
dénominateur	0.00652858	[10 *dl*]

sujet	groupe	t1	t2	t1:log	t2:log	différence
1	g1	135.7	136.0	4.910447	4.912655	-0.002208
2	g1	155.3	152.6	5.045359	5.027820	+0.017539
5	g1	139.2	115.7	4.935912	4.751001	+0.184911
6	g1	91.7	72.0	4.518522	4.276666	+0.241856
8	g1	133.2	151.1	4.891852	5.017942	-0.126090
10	g1	150.3	156.1	5.012633	5.050497	-0.037864
3	g2	148.9	123.1	5.003275	4.812997	+0.190278
4	g2	81.2	77.0	4.396915	4.343805	+0.053110
7	g2	118.7	116.4	4.776599	4.757033	+0.019567
9	g2	115.6	118.9	4.750136	4.778283	-0.028147
11	g2	223.9	222.4	5.411200	5.404478	+0.006722
12	g2	154.1	158.1	5.037602	5.063228	-0.025626

LeBayésien

Nous illustrerons ici le mode d'entrée à partir seulement des carrés-moyens et des caractéristiques du plan, qui permet de retrouver la valeur absolue de la différence d (voir §2.1.5). Dans la fenêtre principale, cliquez sur le bouton **directe**, puis sur le bouton **options: cm\sc**.

Dans la fenêtre des options, entrez le nombre de degrés de liberté du dénominateur du F (10). Activez l'option **constante b**; puis, dans le cadre **plans de base équilibrés...**, entrez l'**effectif par groupe** (6). Cliquez sur le bouton **calculer b et a** et sélectionnez l'option **S<G2>*O2: comparaison O** (qui correspond au plan et à la comparaison analysée ici). Les constantes b=0.2866... ($1/\sqrt{12}$) et a=1.4142... ($1/\sqrt{2}$) sont affichées.

Sélectionnez l'option **carrés moyens du rapport F** et entrez les valeurs

0.0101701 et 0.00652858

puis cliquez sur le bouton **calculer**. Le rapport F (1.557...), ainsi que les statistiques descriptives d=0.041... et s=0.114... sont affichées.

Cliquez sur le bouton **ok** (un message vous rappelle que seule la valeur absolue de d est déterminée), puis dans la fenêtre principale cliquez sur la distribution de l'effet.

Dans la fenêtre des énoncés, sélectionnez l'option **procédures fréquentistes effet négligeable/notable** et effectuez le calcul pour la limite 0.105361. Vous obtenez (avec 3 décimales pour les probabilités) les deux premières lignes

test: négligeable? H_0: $|\delta|>0.105361$
t=1.946 p=0.040

Sur la courbe du bas (notable), remarquez la garantie fiducio-bayésienne associée à $\delta>0.105361$: 0.040.

□ *Remarque*

Vous pouvez également obtenir ces résultats à partir des données qui sont fournies. Dans la fenêtre principale, cliquez sur le bouton **exemples**, et sélectionnez **une étude de bioéquivalence**. Les données sont affichées dans l'éditeur de données. Cliquez sur le bouton **calculer**, ce qui fait apparaître la fenêtre pour le "choix de la structure dérivée". Cliquez sur **dérivation sur les occasions** et entrez les coefficients appropriés pour obtenir la différence des deux traitements

[1 -1]

puis cliquez sur le bouton **ok**. Les statistiques du protocole dérivé pertinent sont calculées.

Cliquez sur l'option **moyenne équipondérée**. Puis cliquez sur le bouton **calculer** pour effectuer l'analyse: d=0.041, t=1.248, etc.

Nous pouvons utiliser le rapport F usuel

$$F = \frac{0.0101701}{0.00652858} = 1.5578$$

On remarquera que la situation est celle considérée au chapitre 1 pour la comparaison du médicament et du *placebo*, et que ce rapport F est le carré de la statistique du test t de l'hypothèse nulle $\delta=0$ obtenue par la procédure décrite en 1.2.1.

La valeur absolue de la différence moyenne observée est ici $|d|=0.0412$. Il est immédiat d'en déduire le test de *négligeabilité* associé à l'hypothèse nulle $|\delta|>0.105361$; la statistique de test est

$$t_{\min}^{[0.105361]} = \frac{|0.0412-0.1054|}{0.0412}\sqrt{1.5578} = 1.946$$

et nous pouvons rejeter cette hypothèse nulle, avec le seuil observé

$$p[H_0: |\delta|>0.105361] = Pr(t_{10}>1.946) = 0.040$$

■ *Réinterprétation fiducio-bayésienne*

Nous avons une généralisation de la réinterprétation fiducio-bayésienne du seuil unilatéral du test usuel

$$Pr(\delta<\alpha) = 1-p[H_0:|\delta|>\alpha] \quad (\text{si } d>0)$$
$$Pr(\delta>-\alpha) = 1-p[H_0:|\delta|>\alpha] \quad (\text{si } d<0)$$

Ici $Pr(\delta<+0.105361) = 1-0.040 = 0.960$

Du point de vue bayésien, cet énoncé ne constitue pas une conclusion d'effet négligeable, puisqu'il inclut les valeurs de δ inférieures à -0.105361.

■ *Caractérisation équivalente à partir de l'intervalle de confiance usuel*

Le test précédent revient à considérer l'intervalle de confiance $1-2\alpha$ (et non $1-\alpha$!) usuel pour δ (centré sur la valeur observée d) et à rejeter l'hypothèse nulle $|\delta|>\alpha$ au seuil α si cet intervalle est contenu dans l'intervalle $[-\alpha,\alpha]$.

■ *Intervalle de confiance pour $|\delta|$ (effet négligeable)*

Nous pouvons construire un intervalle de confiance $1-\alpha$ pour $|\delta|$ comme l'ensemble des $\alpha_0>0$ tels que l'hypothèse nulle $|\delta|>\alpha_0$ est rejetée au seuil α. En raison des considérations précédentes, la borne α de l'intervalle $[0,\alpha]$ ainsi obtenue est la plus petite en valeur absolue des limites de l'intervalle de confiance $1-2\alpha$ (et non $1-\alpha$) usuel pour δ. Du point de vue fiducio-bayésien, la probabilité associée à cet intervalle est comprise entre $1-2\alpha$ et $1-\alpha$. Elle est d'autant plus proche de $1-2\alpha$ que d est proche de 0.

LeBayésien

Dans la situation précédente, activez l'option **garantie** et sélectionnez la garantie 0.90 (soit 1-2×0.05), puis cliquez sur **calculer**. Vous obtenez, dans la courbe du haut, l'intervalle de crédibilité fiducio-bayésien (centré sur d) [-0.0186,0.1010] (avec 4 décimales).

Activez l'option **garantie** et sélectionnez la garantie 0.95, puis cliquez sur **calculer**. Dans le cadre des procédures fréquentistes, vous obtenez sur la deuxième ligne l'intervalle de confiance 0.95 pour $|\delta|$

pour δ: négligeable? [0,0.1010]

Effectuez maintenant le calcul pour cette limite 0.1010; vous obtenez dans la courbe du milieu (négligeable) pour $|\delta|<0.1010$ la garantie 0.949.

Nous obtenons ainsi ici l'intervalle de confiance 0.95 pour $|\delta|$: [0,0.1010] et la probabilité fiducio-bayésienne correspondante $Pr(|\delta|<0.1010)$ est égale à 0.949.

7.1.4 Effet calibré notable, non négligeable, négligeable

Les différentes procédures fréquentistes précédentes peuvent également être appliquées à l'effet calibré $\delta_{cal}=\delta/\sigma$. Nous illustrerons les cas de l'effet notable et de l'effet négligeable.

■ *Test de l'hypothèse nulle ponctuelle* $\delta_{cal}=\alpha$

L'hypothèse nulle ponctuelle $H_0:\delta_{cal}=\alpha$ peut être testée avec la statistique du test usuel $t=d/bs=d_{cal}/b$, dont la distribution d'échantillonnage sous $H_0:\delta_{cal}=\alpha$ est une distribution *t non-centrée* à q degrés de liberté, d'indice d'excentricité α/b (↳10.2.1)

$$[\text{sous } H_0:\delta_{cal}=\alpha]\ t \sim t'_q(\frac{\alpha}{b})$$

■ *Effet calibré notable*

Il s'agit de tester l'hypothèse nulle *composée* (que nous cherchons à rejeter) $\delta_{cal}<\alpha$ (*resp.* $\delta_{cal}>x$). La solution consiste à tester l'hypothèse nulle ponctuelle $\delta_{cal}=\alpha$ au seuil *unilatéral* α, d'où le seuil observé correspondant

$$p[H_0:\delta_{cal}<\alpha] = Pr(t'_q(\frac{\alpha}{b})>t_{obs}) \quad \text{ou} \quad p[H_0:\delta_{cal}>\alpha] = Pr(t'_q(\frac{\alpha}{b})<t_{obs})$$

Il est immédiat de procéder à ce test de notabilité si nous connaissons les valeurs observées d_{cal} et t_{obs}, puisque l'indice d'excentricité α/b s'en déduit

$$\frac{\alpha}{b} = \frac{\alpha}{d_{cal}} t_{obs} \quad \text{(pourvu que } d_{cal}\neq 0)$$

Pour le cas particulier $\alpha=0$, nous retrouvons bien entendu le test usuel, $\delta_{cal}=0$ étant équivalent à $\delta=0$.

LeBayésien

Sélectionnez à nouveau l'exemple **Student** et placez vous dans la fenêtre des énoncés pour la distribution fiducio-bayésienne de δ_{cal}.

Sélectionnez l'option **procédures fréquentistes effet négligeable/notable** (case à cocher), et effectuez le calcul pour la limite 1/2 (pas de deuxième limite). Dans le cadre des procédures fréquentistes, vous obtenez

(avec 3 décimales pour les probabilités) les deux dernières lignes

test: notable? H_0:δ_{cal}<0.500
t'=4.062 p=0.040

(la notation t' rappelle que la distribution utilisée est celle du t non centrée).

Remarquez pour la courbe du bas (effet notable) la garantie fiducio-bayésienne associée à δ_{cal}>0.500: 0.960 (=1-0.040).

Effectuez maintenant le calcul pour la **garantie** 0.95. Dans le cadre des procédures fréquentistes, la dernière ligne donne l'intervalle de confiance 0.95

pour δ_{cal}: notable? [0.545,infini[

qui coïncide avec l'intervalle de crédibilité fiducio-bayésien pour l'effet notable (δ_{cal}>+0.545) associé à cette même garantie.

Reprenons encore l'exemple de Student, pour lequel nous avons d_{cal}=+1.285, t_{obs}=+4.062, q=9, et considérons l'hypothèse nulle δ_{cal}<1/2; nous utilisons la distribution t *non-centrée* $t'_9(+1.581)$ où $+1.581=(\alpha/d_{cal})t_{obs}$, d'où

$$p[H_0:\delta_{cal}<1/2] = Pr\{t'_9(+1.581)>+4.062\} = 0.040$$

La réinterprétation fiducio-bayésienne est donnée par l'énoncé

$$Pr(\delta_{cal}>\alpha) = 1-p[H_0:\delta_{cal}<\alpha] \quad \text{ou} \quad Pr(\delta_{cal}<\alpha) = 1-p[H_0:\delta_{cal}>\alpha]$$

$$\text{Ici } Pr(\delta_{cal}>+0.5) = 1-0.040 = 0.960$$

Nous pouvons construire un intervalle de confiance 1-α pour δ_{cal} comme l'ensemble des α_0 tels que l'hypothèse nulle $\delta_{cal}<\alpha_0$ (ou $\delta_{cal}>\alpha_0$) est rejetée au seuil α. L'intervalle $[\alpha,+\infty[$ (ou $]-\infty,\alpha]$) ainsi obtenu coïncide encore avec l'intervalle de crédibilité fiducio-bayésien donné par l'énoncé

$$Pr(\delta_{cal}>\alpha) = 1-\alpha \quad (resp.\ Pr(\delta_{cal}<\alpha) = 1-\alpha)$$

■ *Effet calibré négligeable*

Nous supposons pour ce test que $|d_{cal}|<\alpha$ (α>0). Le test de l'hypothèse nulle composée $|\delta_{cal}|>\alpha$ est significatif au seuil α si chacun des deux tests des hypothèses ponctuelles $\delta_{cal}=\alpha$ et $\delta_{cal}=-\alpha$ est significatif au seuil *unilatéral* α. En pratique, il suffit donc de considérer la valeur absolue $|t_{obs}|$ et de calculer le seuil observé

$$[H_0:|\delta_{cal}|>\alpha] = Pr(t'_q(\frac{\alpha}{b})<|t_{obs}|)$$

Reprenons l'exemple de l'étude de bioéquivalence décrit en ↳7.1.3. Considérons maintenant l'hypothèse nulle $|\delta_{cal}|>1/2$. Nous avons ici $|d|=0.0412$ et $s=0.1143$, d'où $|d_{cal}|=0.3603$, $\alpha=1/2$ et $|t_{obs}|=\sqrt{F}=1.2481$.

LeBayésien

Sélectionnez à nouveau l'exemple **une étude de bioéquivalence** et procédez comme précédemment: dans l'éditeur de données, cliquez sur le bouton **calculer**, puis entrez les coefficients [1 -1] de la **dérivation sur les occasions**, et cliquez sur le bouton **ok**; dans la fenêtre des statistiques, pour l'option **moyenne équipondérée**, cliquez sur le bouton **ok**

Placez-vous maintenant dans la fenêtre des énoncés pour la distribution fiducio-bayésienne de l'effet calibré δ_{cal}.

Sélectionnez l'option **procédures fréquentistes effet négligeable/notable**, et effectuez le calcul pour la limite 1/2. Dans le cadre des procédures fréquentistes, vous obtenez (avec 4 décimales pour les limites et pour les probabilités) les deux premières lignes

test: négligeable? H_0: $|\delta_{cal}|>0.5000$
$t'=1.248$ $p=0.3097$

Remarquez sur la courbe du bas (effet notable) la garantie fiducio-bayésienne associée à $\delta_{cal}>0.5000$: 0.3097.

Activez l'option **garantie** et sélectionnez la garantie 0.95, puis cliquez sur **calculer**. Dans le cadre des procédures fréquentistes, vous obtenez sur la deuxième ligne l'intervalle de confiance 0.95 pour $|\delta_{cal}|$

pour $|\delta_{cal}|$: négligeable? [0,0.8443]

Remarquez sur la courbe du milieu (effet négligeable) la garantie fiducio-bayésienne associée à $|\delta_{cal}|<0.8444$: 0.9500.

Nous utilisons la distribution *t non-centrée* $t'_{10}(+1.7321)$ où

$$+1.7321 = (\alpha/b)t_{obs} = (0.5/0.3603)\times 1.2481$$

d'où $p[H_0:\delta_{cal}<1/2] = Pr(t'_{10}(+1.7321)>+1.2481) = 0.310$

Nous pouvons construire un intervalle de confiance $1-\alpha$ pour $|\delta_{cal}|$ comme l'ensemble des α_0 tels que l'hypothèse nulle $|\delta_{cal}|<\alpha_0$ est rejetée au seuil α. De même que pour $|\delta|$, la probabilité fiducio-bayésienne associée à l'intervalle $[\alpha,+\infty]$ obtenu est inférieure à $1-\alpha$. Nous obtenons ainsi ici l'intervalle de confiance 0.95 pour $|\delta_{cal}|$ [0,0.8443], qui dans ce cas coïncide presque avec l'intervalle de crédibilité fiducio-bayésien: $Pr(|\delta_{cal}|<0.8444=0.9500$.

7.2 QUELQUES CONSIDÉRATIONS FORMELLES SUR LE TEST DE NÉGLIGEABILITÉ

Dans cette section, nous supposons, sans perte de généralité sur les principes des procédures discutées, que la variance σ^2 est donnée. Nous définirons le test de *négligeabilité* de l'effet de la manière suivante: étant donné un réel positif α, il s'agit de tester l'hypothèse nulle $|\delta|>\alpha$ (que l'on cherche à rejeter) contre $|\delta|\leq\alpha$ (conclusion de négligeabilité).

Rappelons que, dans la situation générale que nous considérons, la distribution d'échantillonnage de la statistique d (exhaustive pour δ) est $N(\delta, b^2\sigma^2)$.

7.2.1 Première procédure

Une solution directe consiste à utiliser la statistique de test $\ell=|d|$ (ou de manière équivalente $|d|/b\sigma$). La région critique de rejet pour $|d|$ de l'hypothèse nulle $|\delta|>\alpha$ (donc de conclusion d'effet négligeable) au seuil α est l'intervalle $[0,\ell_\alpha]$ tel que

$$Pr(|d|\leq\ell_\alpha \,|\, [H_0: |\delta|=\alpha]) = \alpha$$

donc encore tel que, écrit symboliquement

$$Pr(|N(\alpha, b^2\sigma^2)|\leq\ell_\alpha) = \alpha$$

■ *Seuil de signification observé*

Le seuil de signification observé p est le plus petit seuil auquel on peut rejeter l'hypothèse nulle, donc conclure à un effet négligeable. A la valeur observée $|d_{obs}|$ correspond le seuil

$$p = Pr(|d|\leq|d_{obs}| \;|\; [H_0: |\delta|=\alpha]) = Pr(|N(\alpha, b^2\sigma^2)|\leq\ell_{obs})$$

On peut vérifier que $p=|p^{[\alpha]}-p^{[-\alpha]}|$, où $p^{[\alpha]}$ et $p^{[-\alpha]}$ désignent respectivement les seuils de signification unilatéraux observés associés aux tests usuels de chacune des deux hypothèses ponctuelles $\delta=\alpha$ et $\delta=-\alpha$. La procédure équivaut donc à conclure à un effet négligeable ($|\delta|\leq\alpha$) si

$$|p^{[\alpha]}-p^{[-\alpha]}| \leq \alpha$$

■ *Des propriétés indésirables*

Malheureusement, ce test présente des propriétés indésirables.

(1) Quand la valeur observée d_{obs} est nulle, le seuil de signification observé p est toujours nul. Donc dans ce cas on conclut toujours avec ce test à un

effet négligeable à n'importe quel seuil α, quels que soient α et $b\sigma$.

(2) Pour une valeur donnée $|d_{obs}|$, la valeur critique ℓ_α varie de façon non monotone en fonction de $b\sigma$. Par exemple, dans le cas de l'inférence sur une moyenne ($b^2=1/n$), pour $\alpha=\sigma/4$ la valeur critique de rejet $\ell_{0.05}$ au seuil $\alpha=0.05$ est donnée en fonction de n par $Pr(|N(1/4,1/n)|\leq\ell_{0.05}/\sigma)=0.05$. Nous avons par exemple les valeurs

n	$\to 0$	4	16	36	49	100	$\to \infty$
$\ell_{0.05}$	$\to \infty$	0.036σ	0.026σ	0.032σ	0.040σ	0.086σ	$\to 0.25\sigma$

■ *Intervalle de confiance pour $|\delta|$ (effet négligeable)*

Nous pouvons construire un intervalle de confiance $1-\alpha$ pour $|\delta|$ comme l'ensemble des α_0 tels que l'hypothèse nulle $|\delta|>\alpha_0$ est rejetée au seuil α (on conclut à un effet négligeable). Nous retrouvons bien entendu la même absence de monotonie que pour le test. Par exemple, dans le cas de l'inférence sur une moyenne, pour la moyenne observée $|d_{obs}|=\sigma/30$ nous obtenons les intervalles de confiance 0.95 pour $|\delta|$

n=4 $[0,0.175\sigma]$
n=16 $[0,0.308\sigma]$ (plus large que pour n=4...)
n=36 $[0,0.225\sigma]$
n=100 $[0,0.185\sigma]$
etc.

Cette procédure a été proposée notamment par Hodges et Lehmann (1954), Bondy (1969), Anderson et Hauck (1983), Hauck et Anderson (1984), Rocke (1984). Ses propriétés indésirables ont été discutées en particulier par Selwyn, Hall et Dempster (1985), Rocke (1985), Schuirmann (1987).

7.2.2 Seconde procédure

La seconde procédure qui vient naturellement à l'esprit est de conclure à un effet négligeable ($|\delta|\leq\alpha$) si *toutes* les hypothèses nulles ponctuelles, $[H_0:\delta=\delta_0]$ avec $|\delta_0|>\alpha$), sont rejetées. Il suffit donc que les deux hypothèses $[H_0:\delta=\alpha]$ et $[H_0:\delta=-\alpha]$ soient rejetées. Une question essentielle est de savoir à quel seuil doivent être effectués les tests de ces deux hypothèses.

■ *Détermination de la zone de rejet*

En effet, pour effectuer le test de négligeabilité au seuil α, nous déterminons comme dans la première procédure un intervalle de rejet $[0,\ell_\alpha]$ pour $|d|$. Or nous ne pouvons pas utiliser la probabilité d'échantillonnage associée à cet intervalle, puisque c'est le seuil de signification observé associé à cette valeur ℓ_α dans la première procédure. Donc cela nous ramènerait aux propriétés indésirables précédentes. L'aspect embarassant de cette seconde procédure est que nous devons donc utiliser une majoration de cette probabilité. Celle-ci est inférieure à

$$Pr(d \leq \ell_\alpha \mid [H_0:\delta=\alpha]) = Pr(d \geq -\ell_\alpha \mid [H_0:\delta=-\alpha])$$

(voir figure 7.1) et tend vers cette valeur (pour α fixé) lorsque $b\sigma$ tend vers 0 (on ne peut donc pas trouver de "meilleur" majorant). Pour effectuer le test de négligeabilité au seuil α nous déterminons donc ℓ_α de sorte que

$$Pr(d \leq \ell_\alpha \mid [H_0:\delta=\alpha]) = Pr(d \geq -\ell_\alpha \mid [H_0:\delta=-\alpha]) = Pr(N(\alpha, b^2\sigma^2) \leq \ell_\alpha) = \alpha$$

Cette procédure revient donc à choisir pour le test de chacune des deux hypothèses $[H_0:\delta=\alpha]$ et $[H_0:\delta=-\alpha]$ le seuil *unilatéral* α (ou le seuil bilatéral 2α). Remarquons que, si nous avons utilisé ici la symétrie de la distribution d'échantillonnage, la procédure s'applique encore dans le cas d'une distribution non symétrique.

Manifestement l'intervalle ainsi déterminé est vide si $Pr(d \leq 0 \mid [H_0:\delta=\alpha]) \geq \alpha$. Ceci permet d'éviter les propriétés indésirables de la solution précédente, mais conduit en contrepartie à une région de rejet plus petite. Par exemple dans le cas de l'inférence sur une moyenne ($b^2=1/n$), au seuil $\alpha=0.05$ pour $\alpha=\sigma/4$ nous avons les régions de conclusion d'effet négligeable respectives

	procédure 1	procédure 2
n=4	$\lvert d_{obs}\rvert \leq 0.036\sigma$	ø
n=16	$\lvert d_{obs}\rvert \leq 0.026\sigma$	ø
n=36	$\lvert d_{obs}\rvert \leq 0.032\sigma$	ø
n=49	$\lvert d_{obs}\rvert \leq 0.040\sigma$	$\lvert d_{obs}\rvert \leq 0.015\sigma$
n=64	$\lvert d_{obs}\rvert \leq 0.053\sigma$	$\lvert d_{obs}\rvert \leq 0.044\sigma$
n=81	$\lvert d_{obs}\rvert \leq 0.069\sigma$	$\lvert d_{obs}\rvert \leq 0.067\sigma$
n=100	$\lvert d_{obs}\rvert \leq 0.086\sigma$	$\lvert d_{obs}\rvert \leq 0.086\sigma$

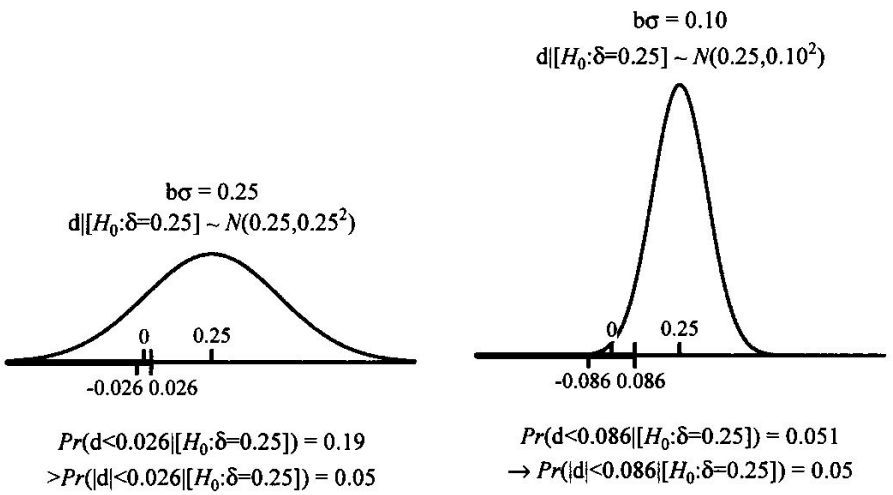


Figure 7.1 - Probabilités associées à la zone de rejet pour les deux procédures pour x = 0.25 et α = 0.05, en fonction de bσ

■ *Seuil de signification observé*

Le seuil de signification observé p est ici déterminé par

$$p = Pr(N(\alpha, b^2\sigma^2) \geq d_{obs}) \text{ si } d_{obs} \geq 0 \quad \text{ou} \quad Pr(N(-\alpha, b^2\sigma^2) \geq d_{obs}) \text{ si } d_{obs} \leq 0$$

et la procédure équivaut donc à conclure à un effet négligeable si *chacun* des deux seuils de signification unilatéraux $p^{[\alpha]}$ et $p^{[-\alpha]}$ définis précédemment est inférieur ou égal à α (au lieu de $|p^{[\alpha]} - p^{[-\alpha]}| \leq \alpha$ pour la première procédure).

■ *Caractérisation équivalente à partir de l'intervalle de confiance usuel*

Cette procédure a été discutée sous l'appellation de "*two one-sided tests procedure*", notamment par Schuirmann (1987). Elle revient encore à calculer l'intervalle de confiance 1-2α usuel pour δ, centré sur la valeur observée

d_{obs} et à conclure à un effet négligeable si cet intervalle est contenu dans l'intervalle $[-\alpha,\alpha]$. Sous cette forme, elle a notamment été proposée par Westlake (1981) et justifiée par Deheuvels (1984). Cette remarque conduit à discréditer l'utilisation d'un intervalle de confiance $1-\alpha$ usuel (suggérée par exemple par Kirkwood, 1981).

■ *Intervalle de confiance pour $|\delta|$ (effet négligeable)*

L'intervalle de confiance $1-\alpha$ pour $|\delta|$, construit comme l'ensemble des α_0 tels que l'hypothèse nulle $|\delta|>\alpha_0$ est rejetée au seuil α, est bien entendu plus large que pour la première procédure. En raison de la remarque précédente, la borne supérieure de cet intervalle est la plus grande en valeur absolue des deux bornes de l'intervalle de confiance $1-2\alpha$ usuel, soit par conséquent

$$|d_{obs}| + b\sigma z_{2\alpha}$$

où $z_{2\alpha}$ désigne la valeur (positive) de la distribution *normale* telle que

$$Pr(|N(0,1)|>z_{2\alpha}) = 2\alpha$$

Dans le même exemple que précédemment ($|d_{obs}|=\sigma/30$ et $b=1/\sqrt{n}$) nous obtenons ainsi les intervalles de confiance 0.95 pour $|\delta|$ ($z_{0.10}=1.645$)

	procédure 1	procédure 2
n=4	$[0,0.175\sigma]$	$[0,0.856\sigma]$
n=16	$[0,0.308\sigma]$	$[0,0.445\sigma]$
n=36	$[0,0.225\sigma]$	$[0,0.308\sigma]$
n=100	$[0,0.185\sigma]$	$[0,0.198\sigma]$

Par conséquent, si $|\delta|$ est plus petit que $b\sigma z_{2\alpha}$, quelle que soit la valeur observée d_{obs}, l'intervalle de confiance précédent contiendra *toujours* $|\delta|$. En particulier, si la limite de négligeabilité α est inférieure à $bz_{2\alpha}$, alors "si l'hypothèse d'un effet négligeable ($|\delta|<\alpha$) est vraie, alors l'intervalle de confiance contient la vraie valeur $|\delta|$ avec une probabilité égale à 1". On voit clairement dans cette situation la différence avec l'inférence bayésienne: conditionnellement à d_{obs}, sauf si on le suppose dans la distribution initiale, la probabilité que $|\delta|$ appartienne à cet intervalle de confiance ne saurait être égale à 1!

■ *Intervalles fréquentistes et intervalles fiducio-bayésiens*

En retour cet intervalle de confiance 1-α pour $|\delta|$ est toujours plus court que l'intervalle de négligeabilité fiducio-bayésien de garantie 1-α, ceci d'autant plus que la valeur observée d_{obs} est proche de 0. Tout se passe comme si dans ce cas nous favorisions *a priori* les valeurs de δ proches de 0: ainsi, si nous observons $d_{obs}=0$, nous considérerons comme intervalle de confiance 1-α pour δ:

• l'intervalle $[-b\sigma z_{\alpha}, +b\sigma z_{\alpha}]$ si nous utilisons l'intervalle usuel;

• l'intervalle $[-b\sigma z_{2\alpha}, +b\sigma z_{2\alpha}]$ (plus court) si nous utilisons "l'intervalle d'effet négligeable".

Ceci est cependant compatible avec la logique de l'inférence fréquentiste, puisque celle-ci n'est pas conditionnelle aux observations. Mais, pour être cohérent nous devons nous en tenir à une application rigide de la procédure: si nous avons prévu (avant l'expérience) une recherche d'effet négligeable, nous devons nous y tenir et considérer l'intervalle correspondant, même si nous observons une valeur d_{obs} notable (qui donnera évidemment pour $|\delta|$ un intervalle très large).

Au contraire l'approche bayésienne, conditionnelle aux observations, est beaucoup plus souple et permet la recherche de la conclusion adaptée à la situation, donc suivant le cas l'utilisation du type d'intervalle approprié.

7.2.3 Remarque: Intervalle de confiance symétrique de Westlake

Westlake (1976) a proposé une autre procédure consistant à modifier l'intervalle de confiance usuel (centré sur l'effet observé) de manière à obtenir un intervalle symétrique autour de 0. L'intervalle ainsi obtenu coïncide en fait avec l'intervalle fiducio-bayésien (mais la solution apparaît artificielle dans le cadre fréquentiste).

7.3 Recherche d'une conclusion d'effet non négligeable

Ne pas pouvoir accepter une conclusion d'effet négligeable ne signifie pas que les données permettent pour autant de rejeter l'hypothèse d'un effet négligeable, ce qui nécessite une autre procédure. Nous allons voir que les procédures d'échantillonnage envisagées précédemment peuvent être adaptées à ce problème, l'hypothèse nulle que l'on cherche à rejeter étant maintenant $|\delta| \leq \alpha$.

Néanmoins une difficulté est de savoir à quelle conclusion la procédure permet réellement d'aboutir. En effet une conclusion telle que $|\delta|$ est supérieur à α est rarement satisfaisante, puisqu'elle ne nous renseigne même pas sur le signe de δ.

7.3.1 Première procédure

La première procédure conduit à rejeter l'hypothèse nulle $|\delta| \leq \alpha$ si $|d_{obs}|$ est situé dans l'intervalle $]\ell_\alpha, +\infty[$ $(\ell_\alpha > 0)$ tel que

$$Pr(|d| > \ell_\alpha | [H_0 : |\delta| = \alpha]) = \alpha$$

donc encore tel que, écrit symboliquement

$$Pr(|N(\alpha, b^2\sigma^2)| > \ell_\alpha) = \alpha$$

Il est facile de vérifier que le seuil de signification observé est dans ce cas la somme $p^{[\alpha]} + p^{[-\alpha]}$ des seuils unilatéraux observés des tests usuels de chacune des deux hypothèses ponctuelles $\delta = \alpha$ et $\delta = -\alpha$. En conséquence ce test ne présente pas les propriétés indésirables précédentes (§6.2.1). En particulier il n'est pas possible de conclure à un effet notable si on ne peut pas rejeter l'hypothèse nulle usuelle $|\delta| = 0$ (propriété hautement souhaitable!).

7.3.2 Seconde procédure

La seconde procédure conduit à rejeter l'hypothèse nulle de négligeabilité $|\delta| \leq \alpha$ au seuil α si toutes les hypothèses ponctuelles $[H_0 : \delta = \delta_0]$ où $\delta_0 \leq \alpha$ sont rejetées à un seuil $h(\alpha)$. La probabilité de la zone de rejet $]\ell_\alpha, +\infty[$ ainsi déterminée est le seuil observé associé à cette valeur ℓ_α dans la première procédure.

Dans ce cas encore nous devons utiliser une majoration de cette probabilité. On peut vérifier que celle-ci est inférieure à α et tend vers α lorsque $b\sigma$ tend vers $+\infty$. Par conséquent nous obtenons simplement $h(\alpha) = \alpha$.

■ *Caractérisation équivalente à partir de l'intervalle de confiance usuel*

Par conséquent la procédure revient à calculer l'intervalle de confiance $1-\alpha$ usuel pour δ (centré sur d_{obs}) et à rejeter l'hypothèse nulle d'un effet négligeable si cet intervalle n'a aucun point commun avec l'intervalle $[-\alpha, \alpha]$ (Deheuvels, 1984). Par conséquent elle revient encore à rejeter cette hypothèse si chacun des deux seuils de signification (unilatéraux) observés $p^{[\alpha]}$ et

$p^{[-x]}$ est inférieur à $\alpha/2$ (au lieu de $p^{[x]}+p^{[-x]}<\alpha$ pour la première procédure). En d'autres termes le seuil de signification observé est le plus grand des seuils *bilatéraux* $2p^{[x]}$ et $2p^{[-x]}$.

CHAPITRE 8
DÉTERMINER LES EFFECTIFS NÉCESSAIRES

Nous étudierons dans ce chapitre des solutions générales pour déterminer les effectifs. Ces solutions utilisent les distributions bayésiennes prédictives; elles englobent et étendent les procédures fréquentistes reposant sur la notion de *puissance*.

Typiquement, dans le cadre d'un plan d'expérience fixé, on se donne une conclusion inductive que l'on souhaite obtenir pour un certain effet, et on cherche l'effectif minimum nécessaire pour "avoir des bonnes chances" d'atteindre cette conclusion. Quelle que soit la procédure utilisée (fréquentiste ou bayésienne), il est bien entendu nécessaire de partir d'informations initiales sur les paramètres. Dans le cadre fréquentiste on doit se contenter de raisonner conditionnellement à des valeurs fixées des paramètres. Le cadre bayésien apporte un avantage immédiat considérable, en permettant de formaliser les informations disponibles, que celles-ci soient de nature "objective" (résultats préliminaires par exemple), ou qu'elles ne correspondent qu'à des *attentes* plus ou moins étayées. Le problème du choix des effectifs est manifestement une situation privilégiée pour utiliser une distribution *a priori* "informative". Cependant, compte tenu du caractère plus ou moins subjectif de cette distribution initiale, on pourra souvent admettre de l'utiliser dans la détermination des effectifs, mais pas dans l'analyse de l'expérience elle-même, pour laquelle on privilégiera plutôt les solutions fiducio-bayésiennes (voire fréquentistes). C'est en tout cas à cette situation que nous nous limiterons ici.

Nous considérerons ici le problème général de l'inférence sur une combinaison linéaire $\delta=\Sigma_g v_g \mu^g$ de moyennes μ^g de groupes indépendants, sous le *modèle normal équivarié*. Les solutions s'étendent aux plans de structure S<G>*O, grâce à l'approche de l'analyse spécifique développée aux chapitres ↳4 et ↳5.

On cherche le plus souvent un effectif identique, que nous noterons n, pour les différents groupes. Nous nous limiterons à ce cas, mais le problème peut

encore être aisément résolu si on cherche des effectifs proportionnels (par exemple, pour deux groupes $f_{g2}=2f_{g1}$).

Les notations générales seront celles introduites dans les chapitres précédents. En particulier nous définirons donc la constante b telle que

$$b^2 = \frac{1}{n}\Sigma_g v_g^2 \quad \text{soit} \quad n = \frac{1}{b^2}\Sigma_g v_g^2$$

Nous envisagerons d'abord des procédures relatives à l'effet calibré δ_{cal}, qui sont d'une certaine manière plus simple puisqu'il suffit de disposer d'une information initiale sur le rapport δ/σ, puis les procédures relatives à l'effet δ. Dans les deux cas nous considérerons les problèmes de recherche de conclusions d'effet notable et d'effet négligeable, par les procédures fiduciobayésiennes et par les procédures fréquentistes.

■ *Remarque sur la précision des calculs*

Les procédures consisteront à déterminer un effectif tel qu'une certaine probabilité (prédictive) soit supérieure à une borne fixée. Le résultat pourra donc dépendre de la précision des calculs, et on pourra dans certains cas obtenir un effectif légèrement différent, en fonction du nombre de décimales utilisé pour calculer la probabilité.

8.1 Effet calibré notable

8.1.1 Le problème

■ *Exemple*

Considérons l'exemple fourni par Schwartz, Flamant et Lellouch (1981, page 97). Il s'agit de savoir si la modification de formule d'un antalgique augmente le temps de rémission de la douleur. On se propose de réaliser un essai expérimental, avec deux groupes indépendants de n patients chacun.

Supposons que l'objectif de cet essai est de montrer que la différence δ des temps de rémission entre la nouvelle formule et l'ancienne est supérieure à la valeur α avec une garantie γ d'au moins 0.975. Ceci inclut le cas particulier $\alpha=0$ (le seul considéré par les auteurs), où il s'agit simplement de montrer

l'existence de la différence.

■ *Le problème général*

Nous supposerons toujours ici qu'il s'agit de montrer que $\delta_{cal}=\delta/\sigma$ est supérieur à α (le cas $\delta_{cal}<\alpha$ étant équivalent par symétrie). Dans la situation générale considérée dans cet ouvrage, les données qui seront recueillies doivent donc être telles que la distribution fiducio-bayésienne

$$\delta_{cal}\,|\,d,s \sim \Lambda'_q(d/s,b^2)$$

conduise à l'énoncé

$$Pr(\delta_{cal}>\alpha)=Pr(\Lambda'_q(d/s,b^2)>\alpha) \geq \gamma$$

La question posée est: quelle est, avant l'expérience, la probabilité (*prédictive*) P d'obtenir un tel énoncé?

Cet énoncé sera satisfait par l'ensemble des valeurs d/s tel que

$$\frac{d}{s} > d_{cal\backslash min}$$

où $d_{cal\backslash min}$ est la valeur telle que

$$Pr(\Lambda'_q(d_{cal\backslash min},b^2)>\alpha) = \gamma$$

soit encore, en raison de la relation remarquable entre les distributions *lambda-prime* et *t non-centrée* (↳9.4.5), telle que

$$Pr(t'_q(\alpha,b^2)<d_{cal\backslash min}) = \gamma$$

La probabilité P cherchée est donc la probabilité prédictive d'observer un rapport d/s supérieur à $d_{cal\backslash min}$

$$Pr(d/s>d_{cal\backslash min}) = P$$

Formulé en ces termes, le problème fait intervenir six quantités:
- la limite α;
- la garantie γ;
- la valeur b (dépendant des effectifs);
- la vraie valeur du paramètre δ_{cal};
- la valeur minimum $d_{cal\backslash min}$ pour obtenir un énoncé fixé;
- la probabilité prédictive P d'obtenir un énoncé fixé.

Ces six quantités sont liées par deux égalités; par conséquent, si nous fixons quatre d'entre elles, les deux autres sont déterminées. Nous pouvons donc envisager en fait quinze questions! Le plus souvent dans la pratique, l'objectif à atteindre (c'est-à-dire α et γ) est fixé, et on n'a pas de raison de fixer la valeur minimum à observer $d_{cal\backslash min}$; En ce cas, il ne reste que trois questions possibles:

- étant donné b et δ_{cal}, trouver $d_{cal\backslash min}$ et la probabilité prédictive P;
- étant donné δ_{cal} et P, trouver $d_{cal\backslash min}$ et b, c'est-à-dire l'effectif minimum tel que la probabilité prédictive d'obtenir l'énoncé souhaité soit supérieure à P;
- étant donné b et P, trouver la plus petite des valeurs δ_{cal} telles que la probabilité prédictive d'obtenir l'énoncé souhaité soit supérieure à P.

Nous nous limiterons ici à ces trois questions, mais nous pourrions également chercher par exemple, étant donné γ, b, δ_{cal} et P, la limite α qui peut être obtenue sous ces conditions, etc.

8.1.2 Simplification du problème: Paramètres fixés et q traité comme infini

■ *Si les valeurs des paramètres sont fixées: Probabilité prédictive et puissance*

Remarquons que si les valeurs des paramètres δ et σ sont fixées, la distribution prédictive relative à d/s est la distribution d'échantillonnage (*t non-centrée*)

$$\frac{d}{s}\Big|\,\delta,\sigma \sim t'_q\left(\frac{\delta}{\sigma}, b^2\right)$$

Nous retrouvons donc dans ce cas les procédures fréquentistes usuelles pour déterminer les effectifs (même si celles-ci se limitent généralement au cas particulier α=0). Dans le cadre fréquentiste, la probabilité prédictive P est la probabilité de rejeter l'hypothèse nulle d'une différence égale à α si la vraie différence vaut δ/σ: P est par définition la *puissance* du test.

Nous verrons plus loin comment prendre en compte dans le cadre bayésien l'incertitude sur les valeurs de δ et σ.

■ q *traité comme infini*

Nous pouvons encore simplifier le problème en supposant que le nombre de degrés de liberté q est élevé (assimilable à une valeur infinie). Les distributions en jeu sont alors des distributions *normales*, et le problème est défini par les deux égalités

$$Pr(N(d_{cal\backslash min}, b^2) > \alpha) = Pr(N(0,1) < \frac{d_{cal\backslash min} - \alpha}{b}) = \gamma$$

$$Pr(N(\delta_{cal}, b^2) > d_{cal\backslash min}) = Pr(N(0,1) < \frac{\delta_{cal} - d_{cal\backslash min}}{b}) = P$$

Il est dans ce cas très facile de déduire deux quelconques des six quantités en jeu à partir des quatre autres. Des formules explicites peuvent être établies, permettant d'obtenir le résultat cherché à l'aide d'une table des valeurs critiques de la distribution *normale*.

■ δ_{cal} *et* n *sont fixés*

$d_{cal\backslash min}$ est d'abord déterminé par

$$d_{cal\backslash min} = \alpha + b\bar{z}^{\gamma}$$

où $\bar{z}^{\gamma}$ est donné par la fonction de répartition de la distribution *normale*

$$Pr(N(0,1) < \bar{z}^{\gamma}) = \gamma$$

$$\text{d'où } P = Pr(N(0,1) < \frac{\delta_{cal} - \alpha - b\bar{z}^{\gamma}}{b})$$

Dans l'exemple précédent du temps de rémission de la douleur, les auteurs considèrent la situation

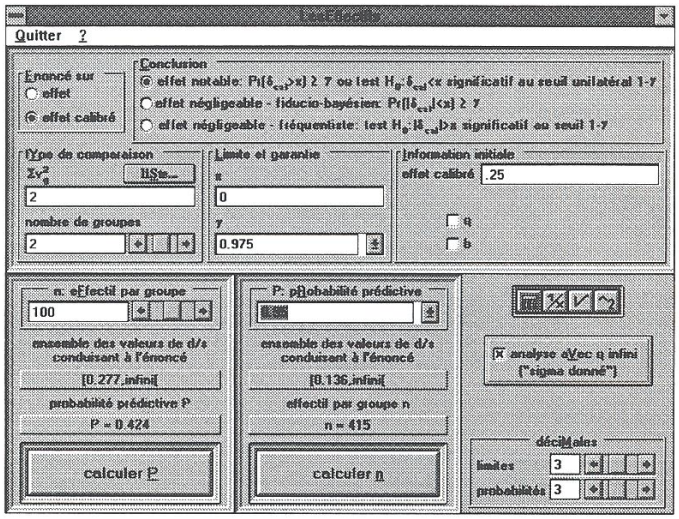
$$\alpha = 0 \qquad \gamma = 0.975 \quad (\bar{z}^{0.975} = +1.9600)$$

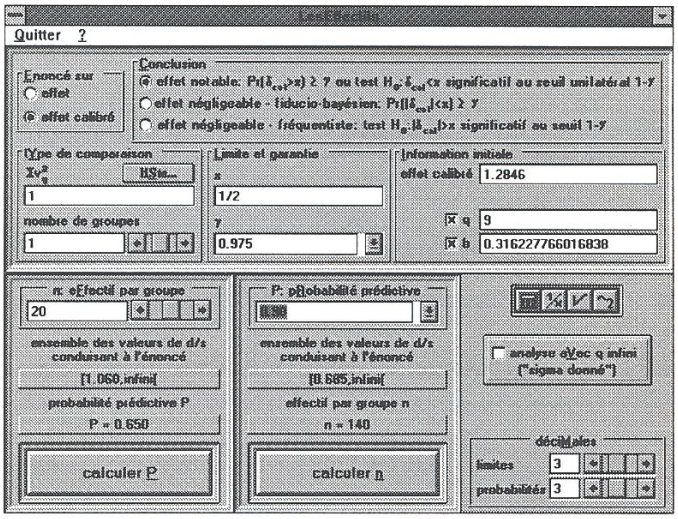
$$\delta = +0.25 \text{ heure et } \sigma = 1, \text{ soit } \delta_{cal} = +0.25$$

Pour le plan considéré, nous avons $\Sigma_g v_g^2 = (+1)^2 + (-1)^2 = 2$

LesEffectifs

Sélectionnez options d'énoncé **"effet calibré"** et de conclusion **"notable"** (ce sont les options par défaut). Entrez les valeurs appropriées:


Quitter ?
Enoncé sur
effet
effet calibré
Conclusion
effet notable: Pr[δ_cal>x) ≥ γ ou test H0:δ_cal<x significatif au seuil unilatéral 1-γ
effet négligeable - fiducio-bayésien: Pr[|δ_cal|<x] ≥ γ
effet négligeable - fréquentiste: test H0:|δ_cal|>x significatif au seuil 1-γ
Type de comparaison
Σv_q^2
Hsie...
2
nombre de groupes
2
Limite et garantie
x
0
γ
0.975
Information initiale
effet calibré .25
q
b
n: eEffectif par groupe
100
ensemble des valeurs de d/s conduisant à l'énoncé
[0.277,infini[
probabilité prédictive P
P = 0.424
calculer P
P: pRobabilité prédictive
0.65
ensemble des valeurs de d/s conduisant à l'énoncé
[0.136,infini[
effectif par groupe n
n = 415
calculer n
analyse aVec q infini ("sigma donné")
décimales
limites 3
probabilités 3


Quitter ?
Enoncé sur
effet
effet calibré
Conclusion
effet notable: Pr[δ_cal>x) ≥ γ ou test H0:δ_cal<x significatif au seuil unilatéral 1-γ
effet négligeable - fiducio-bayésien: Pr[|δ_cal|<x] ≥ γ
effet négligeable - fréquentiste: test H0:|δ_cal|>x significatif au seuil 1-γ
Type de comparaison
Σv_q^2
Hsie...
1
nombre de groupes
1
Limite et garantie
x
1/2
γ
0.975
Information initiale
effet calibré 1.2846
q 9
b 0.316227766016838
n: eEffectif par groupe
20
ensemble des valeurs de d/s conduisant à l'énoncé
[1.060,infini[
probabilité prédictive P
P = 0.650
calculer P
P: pRobabilité prédictive
0.50
ensemble des valeurs de d/s conduisant à l'énoncé
[0.585,infini[
effectif par groupe n
n = 140
calculer n
analyse aVec q infini ("sigma donné")
décimales
limites 3
probabilités 3

$\Sigma_g v_g^2 = 2$ nombre de groupes = 2 $\alpha = 0$ $\gamma = 0.975$

Pour l'**information initiale**, entrez 0.25 pour l'**effet calibré** et désactivez s'il y a lieu les cases à cocher correspondant aux autres informations initiales (écart-type, q et b). Activez l'option **analyse avec q infini**.

Dans le cadre contenant le bouton **calculer P**, entrez l'**effectif par groupe** 100, et cliquez sur ce bouton. Vous obtenez l'ensemble des valeurs de d_{cal} conduisant à l'énoncé $Pr(\delta>0)=0.975$: [0.277,infini[; la probabilité prédictive P de cet ensemble est 0.424.

□ *Remarque*

Le bouton **liste...** permet de déterminer automatiquement $\Sigma_g v_g^2$ et le nombre de groupes pour les problèmes de comparaisons correspondants aux plans usuels élémentaires S, $S<G_2>$, $S\times O_2$ et $S<G_2>\times O_2$ (sur ces plans, voir le chapitre ↳4): ici, sélectionnez **S<G2>: Comparaison G**.

Par exemple, pour n=100, soit b=0.1414, nous obtenons

$$d_{cal\backslash min} = +0.277 \quad \text{et} \quad P = 0.424$$

■ *δ et P sont fixés*

C'est la situation qui donne l'effectif. $d_{cal\backslash min}=\alpha+b\bar{z}^{\gamma}$ est obtenu comme dans le cas précédent, puis b est déterminé par

$$\frac{\delta_{cal}-\alpha-b\bar{z}^{\gamma}}{b} = \bar{z}^{P} \quad \text{soit} \quad b = \frac{\delta-\alpha}{\bar{z}^{\gamma}+\bar{z}^{P}}$$

LesEffectifs

Dans la situation précédente, dans le cadre contenant le bouton **calculer n**, entrez la **probabilité prédictive** 0.95 et cliquez sur ce bouton. Vous obtenez l'effectif n=416 (calcul effectué avec 4 décimales pour les probabilités), et l'ensemble correspondant des valeurs de d_{cal} conduisant à l'énoncé $Pr(\delta>0)=0.975$: [0.136,infini[(la probabilité prédictive P de cet ensemble est 0.95).

Recommencez ce calcul pour $\alpha=0.10$; vous obtenez n=1155.

□ *Remarque*

l'effectif obtenu dépend de la précision demandée sur la probabilité. Avec 4 décimales on obtient P=0.9496 pour n=416. Mais avec 3 décimales, on obtient P=0.950 pour n=415 qui est donc l'effectif donné dans ce cas.

Dans l'exemple précédent, pour

$$\delta = +0.25 \qquad \alpha = 0 \qquad \gamma = 0.975 \qquad P = 0.95 \ (\bar{z}^{0.95}=+1.6449)$$

qui est la situation considérée par les auteurs, nous obtenons

$$b = \frac{0.25}{3.6049} = 0.06935 \quad \text{d'où } n = \frac{2}{b^2} = 415.83$$

et nous retrouvons donc l'effectif n=416 calculé par les auteurs.

Pour α=0.10 et les mêmes valeurs pour les autres quantités, il faudrait un effectif théorique

$$n = 1155.09 \quad \text{correspondant à } b = \frac{0.15}{3.6049} = 0.04161$$

■ n *et* P *sont fixés*

La valeur δ_{cal} cherchée est déterminée, de la même manière que b dans le cas précédent, par

$$\frac{\delta_{cal}-\alpha-b\bar{z}^{\gamma}}{b} = \bar{z}^{P} \quad \text{soit} \quad \delta_{cal} = \alpha + (\bar{z}^{\gamma}+\bar{z}^{P})b$$

Ainsi, dans l'exemple précédent (α=0.10 heure), l'effectif nécessaire sera sans doute jugé trop élevé. Si les contraintes de la situation imposent par exemple un effectif maximum de 200 patients par groupe, nous pouvons nous demander quelle doit être la valeur minimum de δ_{cal} pour avoir une probabilité suffisante d'obtenir l'énoncé voulu avec n=200 (soit b=0.1).

LesEffectifs

L'option de calcul de δ_{cal} n'est pas disponible dans cette version limitée. Vous pouvez cependant obtenir la valeur voulue en procédant par approximations successives pour la valeur initiale **effet calibré**.

Dans la situation précédente (α=0.10), entrez l'**effectif par groupe** 200, et effectuez le calcul de P pour les valeurs successives de l'effet calibré: 0.4, 0.5, 0.45, etc., jusqu'à obtenir la valeur 0.95. 0.46 donne la probabilité prédictive 0.950 (avec 3 décimales).

Pour α=0, γ=0.975, P=0.95, nous trouvons

$$\delta_{cal} = +0.4605$$

8.1.3 Paramètres fixés et q quelconque

Le problème se résoud de la même manière que dans le cas précédent, à ceci près que la distribution *normale* est remplacée par la distribution *lambda-prime*, ou de manière équivalente par la distribution *t non-centrée*, compte tenu de la propriété remarquable qui relie ces deux distributions (↳9.4.5). Le problème est maintenant défini par les deux égalités

$$Pr(\Lambda'_q(d_{cal\backslash min}, b^2) > \alpha) = Pr(t'_q(\alpha, b^2) < d_{cal\backslash min}) = \gamma$$

$$Pr(t'_q(\delta_{cal}, b^2) > d_{cal\backslash min}) = Pr(\Lambda'_q(d_{cal\backslash min}, b^2) < \delta_{cal}) = P$$

On remarquera que, dans ce cas, la valeur fixée de δ_{cal} est seulement utilisée pour calculer la probabilité prédictive; mais contrairement au cas précédent, on ne suppose plus que l'analyse des données sera effectuée en supposant la variance connue (q infini).

■ *Une difficulté pratique supplémentaire*

La difficulté supplémentaire dans l'application des procédures provient de ce qu'il faut maintenant spécifier le lien entre le nombre de degrés de liberté q et l'effectif n, qui dépend du plan d'expérience. En conséquence les distributions en jeu seront différentes pour chaque valeur de b. De plus on ne peut pas établir des tables pour toutes les valeurs des paramètres de ces distributions. Il faudra donc recourir à un programme informatique.

■ *Calcul de* P

LesEffectifs

Revenez à la situation considérée précédemment:

$\Sigma_g v_g^2 = 2$ nombre de groupes = 2 $\alpha = 0$ $\gamma = 0.975$

information initiale: **effet calibré** = 0.25

Désactivez maintenant l'option **analyse avec q infini**, entrez l'**effectif par groupe** 100, et cliquez sur le bouton **calculer P**. Vous obtenez maintenant l'ensemble des valeurs de d_{cal}: [0.279,infini[, et la probabilité prédictive P de cet ensemble: 0.420. Etc.

Nous pouvons ainsi construire des tables donnant P en fonction des valeurs α, γ, δ_{cal} et n. Nous obtenons par exemple, en considérant α=0 et aussi α=0.10 heure, pour γ=0.975 et pour les deux valeurs δ_{cal}=+0.25 et δ_{cal}=+0.35 (on remarquera que, quand n tend vers l'infini, P ne dépend en fait que de la différence $\delta_{cal}-\alpha$) les résultats suivants.

δ_{cal}=+0.25

	α=0	α=0.10
n=100	P=0.420	P=0.182
n=150	P=0.579	P=0.252
n=200	P=0.703	P=0.320
n=250	P=0.797	P=0.386
n=300	P=0.864	P=0.448
n=350	P=0.910	P=0.507
n=400	P=0.942	P=0.561
n=450	P=0.963	P=0.611
n=500	P=0.977	P=0.657
n=550	P=0.985	P=0.699
n=600	P=0.991	P=0.736

δ_{cal}=+0.35

	α=0	α=0.10
n=100	P=0.692	P=0.417
n=150	P=0.856	P=0.576
n=200	P=0.937	P=0.701
n=250	P=0.974	P=0.794
n=300	P=0.990	P=0.862
n=350	P=0.996	P=0.909
n=400	P=0.999	P=0.941
n=450	P=0.999	P=0.962
n=500	P=1.000	P=0.976
n=550	P=1.000	P=0.985
n=600	P=1.000	P=0.991

On peut également tracer des courbes donnant P en fonction de l'une des valeurs (les autres étant fixées).

■ *Quand les effectifs sont élevés*

Bien entendu, quand n tend vers l'infini, nous retrouvons les solutions précédentes. Dans l'exemple précédent, qui nécessitait un effectif élevé, cette procédure apporte donc peu de modification. Ainsi, pour δ_{cal}=+0.25, α=0 et γ=0.975, nous obtenons pour n=416 (qui était l'effectif trouvé précédemment avec 4 décimales pour la probabilité), soit b=0.06938 et q=415+415=830

$$d_{cal\backslash min} = +0.1363 \quad \text{et} \quad P = Pr(\Lambda'_{830}(+0.1363, 0.06938^2) < +0.25) = 0.9496$$

$$\text{au lieu de } P = Pr(N(+0.1359, 0.06938^2) < +0.25) = 0.9501$$

L'effectif nécessaire est maintenant n=417.

8.1.4 Un autre exemple mettant en jeu un effectif plus faible

Reprenons encore l'exemple de Student (↳1.1), et supposons que les données soient celles d'une étude préliminaire, dont nous utilisons les résultats pour construire un essai démonstratif. Nous prendrons comme valeur initiale (fixée) pour δ_{cal} la moyenne de sa distribution fiducio-bayésienne, soit +1.2494.

Supposons que l'objectif soit de montrer que δ_{cal} est supérieur à α=+1/2 avec une garantie d'au moins 0.975. Nous avons ici b^2=1/n et q=n-1.

■ δ_{cal} *et* n *sont fixés*

LesEffectifs

Entrez maintenant les valeurs:

$\Sigma_g v_g^2 = 1$ nombre de groupes = 1 α = 1/2 γ = 0.975

information initiale: **effet calibré** = 1.2494

Entrez l'**effectif par groupe** 20, et effectuez le calcul pour l'option **avec q infini** et sans cette option. Vous obtenez respectivement les probabilités prédictives P: 0.918 et 0.765.

Par exemple, pour n=20, nous obtenons

$$d_{cal\backslash min} = +1.060 \quad \text{et} \quad P = Pr(\Lambda'_{19}(+1.060,0.2236)<+1.2494) = 0.765$$

au lieu de, pour la solution précédente (q infini)

$$d_{cal\backslash min} = +0.938 \quad \text{et} \quad P = Pr\{N(+0.938,0.2236)<+1.2494\} = 0.918$$

■ δ_{cal} *et* P *sont fixés*

LesEffectifs

Dans la situation précédente, effectuez le calcul de n pour la **probabilité prédictive** 0.90; vous obtenez respectivement n=19 et n=28 pour l'option **avec q infini** et sans cette option.

Pour trouver n de manière à avoir une probabilité P égale à 0.90, il faut procéder par ajustements successifs. Nous obtenons ainsi P=0.895 pour n=27 et P=0.907 pour n=28 qui est donc l'effectif cherché. Pour la solution avec q traité comme infini on trouve n=19.

■ n *et* P *sont fixés*

Demandons-nous quelle doit être la valeur minimum de δ_{cal} pour avoir une probabilité P=0.90 d'obtenir l'énoncé voulu avec n=10 (b=0.3162). Nous trouvons dans ce cas

$$d_{cal\backslash min} = +1.410 \quad \text{et} \quad \delta_{cal} = +1.902$$

au lieu de, pour la solution avec q traité comme infini

$$d_{cal\backslash min} = +1.120 \quad \text{et} \quad \delta_{cal} = +1.525$$

8.1.5 Prendre en compte l'incertitude sur les paramètres

Le cadre bayésien permet de prendre en compte l'incertitude sur les vraies valeurs des paramètres δ et σ. Il suffit de se donner une distribution initiale, et de considérer la distribution *prédictive* relative au couple (d,s) pour

cette distribution initiale, au lieu de la distribution d'échantillonnage. Si nous choisissons une distribution conjuguée, caractérisée par

$$\delta|\sigma \sim N(d_0,(b_0^2\sigma^2) \quad \text{et} \quad \sigma \sim s_0\varphi_{q0}^{-1}$$

la distribution prédictive relative à d/s est une nouvelle distribution, que nous appelons *K-prime*, à q_0 et q degrés de liberté

$$\frac{d}{s} \sim K'_{q_0,q}\left(\frac{d_0}{s_0},b_0^2+b^2\right)$$

Puisque l'information initiale n'est pas utilisée pour analyser l'expérience future, $d_{cal\backslash min}$ est bien entendu inchangé, et P est alors défini par l'égalité

$$Pr\left(K'_{q_0,q}\left(\frac{d_0}{s_0},b_0^2+b^2\right)>d_{cal\backslash min}\right) = Pr\left(K'_{q,q_0}\left(d_{cal\backslash min},b_0^2+b^2\right)<\frac{d_0}{s_0}\right) = P$$

qui ne dépend que du rapport d_0/s_0.

Quand $q_0\to\infty$ (soit σ fixé), nous obtenons respectivement comme cas limites les distributions *t non-centrée*, pour la distribution *K-prime* à $q_0\to\infty$ et q degrés de liberté, et *lambda-prime*, pour la distribution *K-prime* à q et $q_0\to\infty$ degrés de liberté. En choisissant en outre $b_0=0$ (soit δ fixé), nous retrouvons donc la solution précédente (δ_{cal} fixé).

Au changement de distribution près, les procédures sont identiques à celles du cas précédent. Reprenons le même exemple et utilisons les valeurs initiales

$$d_0/s_0 = 1.580/1.230 = 1.2846 \qquad q_0 = 9$$

■ δ_{cal} *et* n *sont fixés*

LesEffectifs

Dans la situation précédente, entrez maintenant les informations initiales (activez les cases à cocher correspondantes):

effet calibré = 1.2846 q = 9 b = $1/\sqrt{10}$ = 0.316227766

Effectuez le calcul de la probabilité prédictive P pour les effectifs n=20 et n=200. Vous obtenez respectivement P=0.650 et P=0.913.

Pour n=20 (rappelons que δ=+1.58 et α=+0.50) nous obtenons seulement

$$P = Pr(K'_{19,9}(+1.060,0.387^2)<+1.2846) = 0.650$$

et pour n=200 ($d_{cal\backslash min}$=+0.653)

$$P = Pr(K'_{19,9}(+0.653,0.324^2)<+1.2846) = 0.913$$

■ δ_{cal} *et* P *fixés*

LesEffectifs

Dans la situation précédente, effectuez le calcul de n pour la **probabilité prédictive** 0.90; vous obtenez n=140.

Nous avons ici un exemple de situation où la probabilité prédictive augmente lentement quand n croît. Nous obtenons ainsi

n=75	n=100	n=125	n=130	n=135	n=140	n=145	n=150
P=0.856	P=0.882	P=0.894	P=0.896	P=0.898	P=0.900	P=0.901	P=0.903

Pour avoir une probabilité prédictive 0.90 d'obtenir l'énoncé voulu, il faut donc un effectif de l'ordre de 140, mais un effectif de 75 donne déjà une probabilité P=0.856.

■ n *et* P *sont fixés*

Avec n=10, il faudrait en fait une valeur initiale d_0/s_0=+2.842 pour avoir une probabilité prédictive 0.90 d'obtenir l'énoncé voulu.

■ *Remarque: Cas* b_0=0

En prenant b_0=0, la probabilité prédictive P calculée est la probabilité *conditionnelle* d'obtenir l'énoncé voulu si δ est égal à d_0. Seule intervient ici l'incertitude sur σ. Dans l'exemple précédent, nous trouvons ainsi l'effectif 51 pour avoir une probabilité prédictive 0.90 d'obtenir l'énoncé voulu si δ=+1.58 (et $\sigma \sim 1.230\varphi_9^{-1}$).

8.2 Effet calibré négligeable

8.2.1 Calcul pour la procédure fiducio-bayésienne

■ *Le problème général*

Dans la situation générale considérée, les données qui seront recueillies doivent donc être telles que la distribution fiducio-bayésienne

$$\delta_{cal} \,|\, d,s \sim \Lambda'_q(d/s, b^2)$$

conduise à l'énoncé

$$Pr(|\delta_{cal}| < \alpha) = Pr(|\Lambda'_q(d/s, b^2)| < \alpha) \geq \gamma$$

Cet énoncé sera satisfait par l'ensemble des valeurs d/s tel que

$$\frac{|d|}{s} < d_{cal\backslash max}$$

où $d_{cal\backslash max}$ est la valeur positive telle que

$$Pr(|\Lambda'_q(d_{cal\backslash max}, b^2)| < \alpha) = \gamma$$

soit encore, en raison de la relation remarquable entre les distributions *lambda-prime* et *t non-centrée*, telle que

$$Pr(t'_q(\alpha, b^2) > d_{cal\backslash max}) - Pr(t'_q(\alpha, b^2) < -d_{cal\backslash max}) = \gamma$$

La probabilité P cherchée est donc la probabilité prédictive d'observer un rapport d/s inférieur en valeur absolue à $d_{cal\backslash max}$

$$Pr(|d/s| < d_{cal\backslash max}) = P$$

et est donc donnée, comme précédemment, par une distribution *t non-centrée* (si $q_0 \to \infty$) ou *K-prime* (si $q \to \infty$), avec dans le cas le plus général

$$Pr\left(\left|K'_{q_0,q}\left(\frac{d_0}{s_0}, b_0^2 + b^2\right)\right| < d_{cal\backslash max}\right) = P$$

Les procédures sont analogues à celles de l'effet calibré notable.

8.2.2 Illustrations

■ *Exemple*

On se propose de réaliser une expérience selon un plan *cross-over* $S<G_2>\times T_2$ pour démontrer la bioéquivalence de deux traitements (voir l'exemple traité en ↳7.1.3). Nous supposerons que l'objectif est de montrer que, pour la comparaison T, la différence $\delta_{cal}=\delta/\sigma$ est inférieure en valeur absolue à 1/3 avec une garantie γ d'au moins 0.95. Nous avons ici deux groupes et

$$\Sigma_g v_g^2 = (\tfrac{1}{2})^2 + (\tfrac{1}{2})^2 = 1/2$$

■ *Calcul conditionnellement à* $\delta_{cal}=0$

Dans la situation de recherche d'une conclusion d'effet négligeable, la valeur 0 est une valeur de référence privilégiée; elle correspond au cas "idéal" d'équivalence *stricte*. Calculer l'effectif conditionnellement à un effet vrai $\delta_{cal}=0$ revient à se placer sous la condition la plus favorable pour obtenir l'énoncé voulu.

LesEffectifs

Sélectionnez s'il y a lieu les options d'énoncé "**effet calibré**" et de conclusion "**négligeable fiducio-bayésien**". Entrez les valeurs appropriées:

$\Sigma_g v_g^2 = 1/2$ nombre de groupes = 2 $\alpha = 1/3$ $\gamma = 0.95$

information initiale: **effet calibré** = 0

Effectuez le calcul de l'effectif n pour les probabilités prédictives 0.90, 0.95 et 0.975 (désactivez s'il y lieu les cases à cocher des informations initiales et l'option **analyse avec q infini**). Vous obtenez respectivement les effectifs n: 50, 60 et 69 (avec 3 décimales pour la probabilité).

Nous pouvons obtenir l'effectif par groupe nécessaire, en fonction de la probabilité prédictive P voulue, conditionnellement à $\delta_{cal}=0$, mais aussi conditionnellement à d'autres valeurs moins favorables à la conclusion d'équivalence (les calculs sont effectués avec trois décimales pour la probabilité).

	P=0.90	P=0.95	P=0.975
δ_{cal}=0	n=50	n=60	n=69
δ_{cal}=0.05	n=56	n=70	n=83
δ_{cal}=0.10	n=81	n=102	n=122
δ_{cal}=0.15	n=131	n=166	n=199
δ_{cal}=0.20	n=249	n=315	n=379

■ *Prise en compte de l'incertitude sur* δ_{cal}

Pour tenir compte de notre incertitude sur la vraie valeur de δ_{cal}, nous pouvons considérer une distribution initiale pour ce paramètre. Par exemple, choisissons la distribution $N(0,0.051^2)$, qui donne une probabilité *a priori* 0.95 que δ_{cal} soit inférieur en valeur absolue à 0.10. Rappelons que cette distribution initiale est utilisée seulement pour calculer la probabilité prédictive (et pas pour l'analyse de l'expérience).

LesEffectifs

Dans la situation précédente, entrez les informations initiales (pas d'information pour q_0 pris ici infini):

effet calibré = 0 b = 0.051

Effectuez le calcul de l'effectif n pour les probabilités prédictives 0.90, 0.95 et 0.975. Vous obtenez respectivement les effectifs n: 57, 71 et 86.

Pour une distribution initiale pour δ_{cal} de la forme $N(0,b_0^2)$, nous pouvons obtenir l'effectif par groupe n nécessaire, en fonction de b_0 (les calculs sont effectués avec trois décimales pour la probabilité).

	probabilité a priori	P=0.90	P=0.95	P=0.975
b_0=0.0608	$Pr(\|\delta_{cal}\|<0.10)=0.90$	n=60	n=77	n=96
b_0=0.0510	$Pr(\|\delta_{cal}\|<0.10)=0.95$	n=57	n=71	n=86
b_0=0.0388	$Pr(\|\delta_{cal}\|<0.10)=0.99$	n=54	n=66	n=79
$b_0 \to 0$		n=50	n=60	n=69

8.2.3 Calcul pour la procédure fréquentiste "des deux seuils unilatéraux"

■ *Le problème général*

Pour cette procédure de test de l'hypothèse nulle $|\delta_{cal}|>\alpha$ au seuil α, les données qui seront recueillies, doivent être telles que (*cf.* ↳7.1.4)

$$Pr(t'_q(\alpha/b)<|d|/bs) \leq \alpha \quad \text{soit} \quad Pr(t'_q(\alpha,b^2)>|d|/s) \geq 1-\alpha$$

Cet énoncé sera satisfait par l'ensemble des valeurs d/s tel que

$$\frac{|d|}{s} < d_{cal\backslash max}$$

où $d_{cal\backslash max}$ est la valeur positive telle que

$$Pr(t'_q(\alpha,b^2)>d_{cal\backslash max}) = Pr(\Lambda'_q(d_{cal\backslash max},b^2)<\alpha) = 1-\alpha$$

que l'on rapprochera de la condition trouvée précédemment pour la solution fiducio-bayésienne

$$Pr(t'_q(\alpha,b^2)>d_{cal\backslash max}) - Pr(t'_q(\alpha,b^2)<-d_{cal\backslash max}) = \gamma$$

Il apparaît que, si l'on prend $\alpha=1-\gamma$, les valeurs trouvées pour $d_{cal\backslash max}$ dans les deux cas seront le plus souvent très voisines dans la pratique, la probabilité $Pr(t'_q(\alpha,b^2)<-d_{cal\backslash max})$ étant généralement très petite.

■ *Illustration*

LesEffectifs

Dans la situation précédente, sélectionnez l'option de conclusion "**négligeable fréquentiste**", et effectuez les mêmes calculs. Vous obtenez les mêmes effectifs.

Ainsi, dans l'exemple précédent, nous trouvons dans tous les cas les mêmes effectifs pour les deux procédures.

On pourrait s'étonner de cette similitude, alors que la solution fréquentiste apparaît plus favorable à l'obtention d'une conclusion d'effet négligeable. En fait, les deux procédures divergent essentiellement pour les valeurs observées de d proches de 0, mais coïncident lorsque la valeur de d devient élevée. C'est pourquoi la plus grande des valeurs permettant de conclure à un effet négligeable est sensiblement la même pour les deux procédures.

8.3 Effet (non calibré) notable

8.3.1 Solution générale

■ *Le problème*

La situation, et les procédures, sont analogues à celles de l'effet calibré notable. Ici les données qui seront recueillies doivent être telles que la distribution fiducio-bayésienne

$$\delta \mid d,s \sim t_q(d,b^2s^2)$$

conduise à l'énoncé

$$Pr(\delta>\alpha)=Pr(t_q(d,b^2s^2)>\alpha) \geq \gamma$$

Rappelons que la procédure fréquentiste de test unilatéral de l'hypothèse nulle $H_0:\delta<\alpha$ au seuil $1-\gamma$ conduit à la même condition (↳7.1.1).

Cet énoncé sera satisfait par l'ensemble des couples (d,s) tel que

$$d-bs\bar{t}_q^{\gamma} > \alpha \quad \text{soit encore} \quad \frac{d-\alpha}{bs} < \bar{t}_q^{\gamma}$$

où $\bar{t}_q^{\gamma}$ est donné par la fonction de répartition de la distribution t de Student

$$Pr(t_q<\bar{t}_q^{\gamma}) = \gamma$$

Ici, il n'y a pas de limite inférieure pour d autre que α: pour toute valeur $d>\alpha$ (et quel que soit b), l'inégalité précédente peut être satisfaite pour certaines valeurs de s.

■ *Solution*

La probabilité P cherchée est donc la probabilité prédictive d'observer un tel couple. Pour une distribution initiale caractérisée, comme précédemment, par d_0, b_0, s_0 et q_0, nous pouvons utiliser

• la distribution prédictive de la limite de crédibilité

$$d-bs\bar{t}_q^{\gamma} \sim d_0+K'_{q,q_0}(-bs_0\bar{t}_q^{\gamma},(b_0^2+b^2)s_0^2)$$

$$\text{d'où } Pr(d-bs\bar{t}_q^{\gamma}>\alpha) = Pr\left(K'_{q,q_0}(bs_0\bar{t}_q^{\gamma},(b_0^2+b^2)s_0^2)<d_0-\alpha\right)$$

$$= Pr\left(K'_{q,q_0}(\bar{t}_q^{\gamma},\frac{b_0^2+b^2}{b^2})<\frac{d_0-\alpha}{bs_0}\right) = P$$

• la distribution prédictive de la statistique de test (qui conduit bien entendu à des énoncés équivalents)

$$\frac{d-\alpha}{bs} \sim K'_{q_0,q}\left(\frac{d_0-\alpha}{bs_0},\frac{b_0^2+b^2}{b^2}\right)$$

$$\text{d'où } Pr\left(\frac{d-\alpha}{bs}>\bar{t}_q^{\gamma}\right) = Pr\left(K'_{q_0,q}\left(\frac{d_0-\alpha}{bs_0},\frac{b_0^2+b^2}{b^2}\right)>\bar{t}_q^{\gamma}\right) = P$$

On remarquera que la probabilité P ne dépend en fait de d_0 et s_0 que par l'intermédiaire de $(d_0-\alpha)/s_0$, au lieu de d_0/s_0 dans le cas de l'effet calibré. Bien entendu, pour $\alpha=0$, il est équivalent d'utiliser l'effet calibré ou l'effet.

Quand $q_0\to\infty$ (soit $\sigma\equiv s_0$ fixé) et $b_0=0$ (soit $\delta\equiv d_0$ fixé), nous retrouvons respectivement comme cas limites les distributions d'échantillonnage

- *lambda-prime* pour d-bst_q^{γ};
- *t non-centrée* pour (d-α)/bs.

Rappelons que, dans le cadre fréquentiste, la probabilité P est dans ce cas la *puissance* du test, c'est-à-dire la probabilité de rejeter l'hypothèse nulle d'une différence égale à α si la vraie différence vaut δ.

8.3.2 Illustrations

■ *Calcul de* n *pour* δ *et* σ fixés

Reprenons à nouveau l'exemple de Student. Nous prendrons comme valeurs initiales (fixées), pour δ la différence moyenne observée d_{obs}=+1.580, et pour σ la moyenne de sa distribution fiducio-bayésienne, soit 1.346. Pour nous placer dans une situation comparable à celle considérée précédemment pour l'effet calibré en ↳8.1.4 (δ>σ/2), supposons que l'objectif soit de montrer que la différence δ est supérieure à α=1.346/2=+0.673 avec une garantie d'au moins 0.975.

LesEffectifs

Entrez les valeurs:

$\Sigma_g v_g^2$ = 1 nombre de groupes = 1 α = 0.673 γ = 0.975

information initiale: **effet** = 1.580 **écart-type** = 1.346

Désactivez s'il y a lieu les différentes options (q, b et calcul avec q infini). Sélectionnez les options d'énoncé "**effet**" et de conclusion "**notable**", et effectuez le calcul de n pour la probabilité prédictive P=0.90. Vous obtenez l'effectif n=26.

Nous obtenons ainsi P=0.898 pour n=25 et P=0.910 pour n=26, qui est donc l'effectif cherché.

■ *Prise en compte de l'incertitude sur δ et σ*

Les informations initiales sont

$$d_0 = +1.580 \qquad b_0 = 1/\sqrt{10} \qquad s_0 = 1.230 \qquad q_0 = 9$$

LesEffectifs

Dans la situation précédente, entrez maintenant les informations initiales

effet = 1.580 écart-type = 1.230

$q = 9$ $b = 1/\sqrt{10} = 0.316227766$

Pour la probabilité P=0.90, vous obtenez l'effectif n=73 (en effectuant le calcul avec 3 décimales pour les probabilités).

Nous trouvons ici P=0.900 pour n=73, qui est donc l'effectif cherché.

8.4 Effet (non calibré) négligeable

8.4.1 Calcul pour la procédure fiducio-bayésienne

Ici les données qui seront recueillies doivent être telles que la distribution fiducio-bayésienne

$$\delta \mid d,s \sim t_q(d,b^2s^2)$$

conduise à l'énoncé

$$Pr(|\delta|<\alpha)=Pr(|t_q(d,b^2s^2)|<\alpha|) \geq \gamma$$

Le problème de la détermination de l'ensemble des couples (d,s) satisfaisant l'égalité $Pr(|t_q(d,b^2s^2)|<\alpha|)=\gamma$, et de sa probabilité prédictive, est plus complexe que dans les cas précédents. La méthode est la suivante.

(1) Conditionnellement à s, cet énoncé sera satisfait par l'ensemble des valeurs d tel que

$$|d| < d_{max/s}$$

où $d_{max/s}$ est la valeur positive, obtenue par approximations successives, telle que

$$Pr(|t_q(d_{max/s}, b^2s^2)|<\alpha) = \gamma$$

On en déduit la probabilité prédictive *conditionnelle* à s

$$Pr(|d|<d_{max/s}|s) = Pr\left(|t_{q_0+q}(d_0,(b_0^2+b^2)\frac{q_0s_0^2+qs^2}{q_0+q})|<d_{max/s}\right) = P_s$$

(2) La probabilité marginale cherchée P est obtenue par intégration numérique, à partir de la distribution prédictive marginale relative à s

$$s \sim s_0 F^{1/2}_{q,q_0}$$

Le calcul est donc relativement long. En fait, compte tenu de la propriété mise en évidence dans le cas de l'effet calibré (↳8.2.3) de la similitude avec la procédure fréquentiste "des deux seuils unilatéraux", nous pourrons dans la pratique utiliser les résultats obtenus pour cette procédure, exposés ci-après.

8.4.2 Calcul pour la procédure fréquentiste "des deux seuils unilatéraux"

■ *Le problème général*

Pour cette procédure de test de l'hypothèse nulle $|\delta|>\alpha$ au seuil α (voir ↳7.1.3), les données qui seront recueillies, doivent être telles que l'on ait simultanément

$$d-bs\bar{t}_q^{\gamma} > -\alpha \quad \text{et} \quad d+bs\bar{t}_q^{\gamma} < \alpha$$

$$\text{soit } |d|+bs\bar{t}_q^{\gamma} < \alpha \quad \text{ou encore} \quad \frac{|d|-\alpha}{bs} < -\bar{t}_q^{\gamma}$$

où $\bar{t}_q^{\gamma}$ est donné par la fonction de répartition de la distribution t de Student

$$Pr(t_q<\bar{t}_q^{\gamma}) = \gamma$$

Contrairement au cas de la solution fiducio-bayésienne, l'ensemble des couples (d,s) est ici déterminé de manière explicite. Il s'agit donc de calculer la probabilité prédictive

$$Pr\{|d|+bs\bar{t}_q^{\gamma}<\alpha\} = Pr\left(\frac{|d|-\alpha}{bs}<-\bar{t}_q^{\gamma}\right) = P$$

Cette probabilité, *conditionnellement* à s est donnée par

$$Pr\left(\left|t_{q_0+q}\left(d_0,(b_0^2+b^2)\frac{q_0s_0^2+qs^2}{q_0+q}\right)\right|<\alpha-bs\bar{t}_q^{\gamma}\right) = P_s$$

avec dans le cas particulier $q_0=\infty$

$$Pr\left(\left|N\left(d_0,(b_0^2+b^2)s_0^2\right)\right|<\alpha-bs\bar{t}_q^{\gamma}\right) = P_s$$

La probabilité marginale P est obtenue par intégration numérique, à partir de la distribution prédictive marginale relative à s

$$s \sim s_0(F_{q,q_0})^{1/2}$$

avec pour $q_0=\infty$, $s \sim s_0\varphi_q^{-1}$ (distribution d'échantillonnage)

■ *Remarque*

Dans le cas particulier $b_0=0$ et $q_0=\infty$ (soit $\delta\equiv d_0$ et $\sigma\equiv s_0$ fixés), Owen (1965) a traité ce problème (*cf.* Phillips, 1990) comme celui (équivalent) du calcul de la probabilité conjointe

$$Pr(d-bs\bar{t}_q^{\gamma}>-\alpha \text{ et } d+bs\bar{t}_q^{\gamma}<\alpha)$$

qui met en jeu une distribution du *t non-centrée* bivariée, et a donné une méthode de calcul numérique relativement complexe.

8.4.3 Illustrations

■ *Exemple*

Reprenons la situation du plan *cross-over* considérée précédemment (↳8.2.1) pour l'effet calibré. Pour nous placer dans une situation comparable à celle envisagée dans ce dernier cas, nous supposerons que l'objectif est de montrer que la différence δ est inférieure en valeur absolue à 1/3 avec une garantie γ d'au moins 0.95, pour un écart-type donné égal à 1.

LesEffectifs

Sélectionnez s'il y a lieu les options d'énoncé **"effet"** et de conclusion **"négligeable fréquentiste"**. Entrez les valeurs appropriées:

$\Sigma_g v_g^2 = 1/2$ nombre de groupes = 2 $\alpha = 1/3$ $\gamma = 0.95$

information initiale: effet = 0 écart-type = 1

Effectuez le calcul de l'effectif n pour les probabilités prédictives 0.90, 0.95 et 0.975 (désactivez s'il y lieu les cases à cocher des informations initiales et l'option **analyse avec q infini**). Vous obtenez les effectifs respectifs: 50, 60 et 69 (avec 3 décimales pour la probabilité).

Puis entrez l'informations initiale supplémentaire:

$b = 0.051$

Effectuez le calcul de l'effectif n pour les probabilités prédictives 0.90, 0.95 et 0.975. Vous obtenez respectivement les effectifs n: 57, 71 et 86.

Nous pouvons obtenir l'effectif par groupe nécessaire, en fonction de la probabilité prédictive P voulue, conditionnellement à diférentes valeurs de δ (les calculs sont effectués avec trois décimales pour la probabilité).

	P=0.90	P=0.95	P=0.975
δ=0	n=50	n=60	n=69
δ=0.05	n=56	n=69	n=82
δ=0.10	n=80	n=100	n=120
δ=0.15	n=128	n=162	n=194
δ=0.20	n=242	n=305	n=366

Pour une distribution initiale pour δ de la forme $N(0, b_0^2\sigma^2)$ (avec ici σ=1 donné), nous obtenons l'effectif par groupe n nécessaire, en fonction de b_0.

	probabilité a priori	P=0.90	P=0.95	P=0.975
b_0=0.0608	$Pr(\|\delta\|<0.10)=0.90$	n=60	n=76	n=94
b_0=0.0510	$Pr(\|\delta\|<0.10)=0.95$	n=57	n=71	n=85
b_0=0.0388	$Pr(\|\delta\|<0.10)=0.99$	n=54	n=66	n=78
$b_0 \to 0$		n=50	n=60	n=69

On remarquera que ces effectifs sont très voisins de ceux obtenus précédemment pour l'effet calibré.

■ *Autre exemple (Phillips, 1990)*

Phillips (1990, page 142) fournit un tableau d'effectifs, correspondant à la situation (avec les définitions utilisées ici):

$\Sigma_g v_g^2 = 2$ nombre de groupes = 2 $\alpha = 2$ $\gamma = 0.95$ $q_0 = \infty$ $b_0 = 0$

pour différentes valeurs de σ, d_0 et P. Nous retrouvons les mêmes effectifs, à une unité près (les calculs sont effectués avec cinq décimales pour la probabilité).

σ=1	P=0.7	P=0.8	P=0.9
d_0=0	n=5	n=6	n=7
d_0=0.5	n=6	n=7	n=9
d_0=1	n=11	n=14	n=18
d_0=1.5	n=39	n=51	n=70

σ=2	P=0.7	P=0.8	P=0.9
d_0=0	n=16	n=18	n=23
d_0=0.5	n=19	n=24	n=32
d_0=1	n=39	n=51	n=70
d_0=1.5	n=152	n=199	n=275

σ=3	P=0.7	P=0.8	P=0.9
d_0=0	n=34	n=40	n=50
d_0=0.5	n=41	n=51	n=70
d_0=1	n=86	n=112	n=155
d_0=1.5	n=340	n=446	n=618

On remarquera que, si l'on adopte le critère de l'effet calibré, les valeurs considérées ici peuvent difficilement être admises comme permettant une conclusion d'effet négligeable. Il n'est donc pas surprenant de trouver que des effectifs relativement faibles sont suffisants.

■ *Un exemple avec information initiale sur* σ

Considérons encore l'exemple d'une étude de bioéquivalence suivant un plan *cross-over*. Considérons l'intervalle de bioéquivalence "à p%" (voir ↳7.1.3)

$$[log(\frac{100\text{-p}}{100}) , log(\frac{100}{100\text{-p}})]$$

Supposons qu'une expérience préliminaire ait fourni pour σ la distribution fiducio-bayésienne

$$\sigma \sim 0.7145454\varphi_{16}^{-1}$$

Considérons pour δ la situation la plus favorable $\delta\equiv 0$ (soit $b_0=0$), et cherchons l'effectif nécessaire pour obtenir une conclusion d'équivalence avec la garantie γ=0.95 pour une probabilité prédictive P=0.90.

LesEffectifs

Sélectionnez s'il y a lieu les options d'énoncé "**effet**" et de conclusion "**test... (effet négligeable - fréquentiste)**". Entrez les valeurs appropriées:

$\Sigma_g v_g^2 = 1/2$ nombre de groupes = 2 γ = 0.95

information initiale:

effet = 0 écart-type = 0.7145454 q = 16

Entrez la limite $\alpha=log(1.25)=0.223144$, et effectuez le calcul de l'effectif n pour la probabilité prédictive 0.90 (désactivez s'il y lieu les cases à cocher des informations initiales et l'option **analyse avec q infini**): vous obtenez l'effectif 61 (avec cinq décimales pour la probabilité). Supprimez ensuite le nombre de degrés de liberté q=16 (désactivez la case à cocher correspondante), et refaites le calcul: vous obtenez cette fois l'effectif 57.

On obtient ici les effectifs nécessaires, en fonction de l'intervalle d'équivalence. Les calculs sont également effectués pour la valeur fixée σ=0.7145454 (soit q_0 infini).

intervalle d'équivalence	q_0=16	q_0=∞
10% α=log(100/90)=0.1054	n=306	n=250
15% α=log(100/85)=0.1625	n=129	n=106
20% α=log(100/80)=0.2231	n=69	n=57
25% α=log(100/75)=0.2877	n=42	n=35
30% α=log(100/70)=0.3567	n=28	n=23

CHAPITRE 9
DISTRIBUTIONS UTILES

Nous présentons dans ce chapitre les distributions qui interviennent dans les procédures inférentielles développées dans cet ouvrage. Nous présenterons d'abord les "distributions de base" *gamma*, et une généralisation que nous appellons *gamma-puissance*, *khi-deux* (et distributions dérivées), *normale*, *bêta* et *F* centrée, puis les distributions de Student, *t généralisée* et *t non-centrée* (ou *décentrée*). Nous étudierons ensuite la distribution, que nous dénommons *lambda-prime*, qui intervient dans les inférences bayésiennes relatives aux "effets calibrés". Cette distribution a notamment été utilisée par Fisher, dans "*Statistical Methods and Scientific Inference*" (Fisher, 1990, pages 126-127), mais n'a guère fait l'objet d'études systématiques en dépit de son intérêt. Enfin nous introduirons la distribution, que nous dénommons *K-prime*, qui intervient dans les inférences bayésiennes prédictives, et qui est une généralisation des distributions *t non-centrée* et *lambda-prime*.

Toutes ces distributions sont disponibles dans le programme **LesDistributions** (voir l'annexe sur ce programme).

■ *Définitions générales et notations*

Si X est une variable (aléatoire) continue réelle admettant une *densité* (de probabilité) par rapport à la mesure de Lebesgue sur $\mathbb{R}$, nous noterons $p(x)$ cette *densité* au point x, et $F(x)$ la fonction de répartition $F(x)=Pr(X<x)$.

Le *moment* d'ordre k de X est la *moyenne* (ou *espérance*) de la variable X^k

$$moy(X^k) = \int x^k p(x)dx$$

Le *moment centré* d'ordre k de X, noté $\mu_k(X)$, est la *moyenne* (ou *espérance*) de la variable $(X-moy(X))^k$

$$\mu_k(X) = \int (X-moy(X))^k p(x)dx$$

En particulier les moments centrés d'ordre 2 (ou *variance*), 3 et 4 se déduisent des formules classiques

$$Var(X) = moy(X^2) - (moy(X))^2$$

$$\mu_3(X) = moy(X^3) - 3moy(X^2)moy(X) + 2(moy(X))^3$$

$$\mu_4(X) = moy(X^4) - 4moy(X^3)moy(X) + 6moy(X^2)(moy(X))^2 - 3(moy(X))^4$$

On définit encore

- le coefficient d'asymétrie (en anglais *skewness*) $\frac{\mu_3(X)}{(\mu_2(X))^{3/2}}$
- le coefficient d'aplatissement (en anglais *kurtosis*) $\frac{\mu_4(X)}{(\mu_2(X))^2}$

Un outil essentiel est fourni par le mélange de distributions conditionnelles. Si, conditionnellement à la valeur y de la variable Y, la variable X a la densité $p(x|y)$, alors

$$p(x) = \int p(x|y)p(y)dy$$

$$moy(X^k) = moy(moy(X^k|y)) = \int \left(\int x^k p(x|y)p(y)dy\right)dx$$

Dans ce qui suit, nous ne distinguerons pas dans les notations les variables et leurs valeurs, notées de la même manière en minuscules.

9.1 DISTRIBUTIONS DE BASE

9.1.1 Distribution gamma

■ *Définition et densité*

x a la distribution *gamma* de paramètre a (a>0), d'échelle b (b>0), et de position c, ce que nous noterons

$$x \sim c+bG(a)$$

si x a la densité

$$p(x) = \frac{1}{\Gamma(a)} b^{-a} (x-c)^{a-1} exp\left(-\frac{(x-c)}{b}\right)$$

b et c sont respectivement des paramètres d'échelle et de position puisque

$$\frac{x-c}{b} \sim G(a)$$

où G(a) est la distribution *gamma* à un seul paramètre (soit b=1 et c=0).

A cette densité sont liés les définitions et résultats suivants.

■ *Fonction gamma*

La fonction *gamma* $\Gamma(a)$ est définie par l'intégrale (a>0)

$$\Gamma(a) = \int_0^1 u^{a-1} exp(-u)du$$

Elle vérifie la relation

$$\Gamma(a+1) = a\Gamma(a)$$

On en déduit, pour k entier

$$\Gamma(k) = (k-1)(k-2)...1 = (k-1)!$$

$$\Gamma(k+\tfrac{1}{2}) = (k-\tfrac{1}{2})(k-\tfrac{3}{2})...\tfrac{1}{2}\sqrt{\pi}$$

et les valeurs particulières

$$\Gamma(1) = \Gamma(0) = 1 \text{ et } \Gamma(\tfrac{1}{2}) = \sqrt{\pi}$$

ainsi que la formule dite de *duplication*

$$2^{2a-1}\,\Gamma(a)\,\Gamma(a+\tfrac{1}{2}) = \sqrt{\pi}\,\Gamma(2a)$$

■ *Fonction de répartition et fonction gamma incomplète*

La fonction *gamma incomplète* $I\Gamma_x(a)$ est définie par la fonction de répartition de la distribution G(a)

$$I\Gamma_x(a) = F(x) = \frac{1}{\Gamma(a)} \int_0^x u^{a-1} exp(-u)du$$

On fera attention au fait que la fonction *gamma incomplète* est définie par certains auteurs comme l'intégrale $\int_0^x u^{a-1} exp(-u)du$.

9.1.2 Distribution gamma-puissance

■ *Définition et densité*

Il est commode d'introduire la distribution suivante, que nous appellerons "*gamma-puissance*", qui est la distribution de la variable $x=c+y^r$ (où r est un réel non nul), avec $y \sim bG(a)$. Nous noterons

$$x \sim c+b^rG^r(a)$$

La densité de x est déduite de celle de la distribution gamma par changement de variable

$$p(x) = \frac{1}{|r|}\frac{1}{\Gamma(a)} b^{-a} (x-c)^{(a/r)-1} exp\left(-\frac{(x-c)^{1/r}}{b}\right)$$

■ *Cas particuliers*

Cette distribution inclut comme cas particulier:
• la distribution *gamma* (r=1), qui elle-même inclut la distribution *exponentielle* (pour a=1) et la distribution de *Erlang* (pour b=1);
• les distributions *khi-deux* (centrée) et dérivées qui seront définies en ↳9.1.3, qui sont en fait les seuls cas utilisés dans cet ouvrage;
• ainsi que d'autres distributions plus ou moins classiques, par exemple la distribution de *Weibull* de paramètres g et h qui est (en adoptant la définition la plus usuelle des paramètres) la distribution $gG^{1/h}(1)$ (soit a=1, r=1/h et $b=g^h$).

■ *Fonction de répartition*

La fonction de répartition est donnée par la fonction *gamma incomplète*

$$F(x) = I\Gamma_{(x-c)^{1/r}/b}(a) \qquad \text{si } r>0$$

$$= 1 - I\Gamma_{(x-c)^{1/r}/b}(a) \qquad \text{si } r<0$$

■ *Intégrales utiles*

De la densité précédente, puisque $\int_0^\infty p(x)dx=1$, nous déduisons (dans le cas particulier c=0) l'intégrale

$$\int_0^{+\infty} u^{(a/r)-1} exp\left(-\frac{u^{1/r}}{b}\right)du = |r|\,\Gamma(a)\,b^a$$

et, plus généralement

$$\int_0^x u^{(a/r)-1} exp\left(-\frac{u^{1/r}}{b}\right) du = |r|\,\Gamma(a)\,b^a\,I\Gamma_{x^{1/r}/b}(a) \qquad \text{si } r>0$$

$$= |r|\,\Gamma(a)\,b^a\,(1-I\Gamma_{x^{1/r}/b}(a)) \qquad \text{si } r<0$$

■ *Moments de* $x \sim b^r G^r(a)$

Dans le cas particulier c=0

$$moy(x^k) = b^{kr} \frac{\Gamma(a+kr)}{\Gamma(a)}$$

formule définie pour tout réel k tel que a+kr>0

Ce résultat se déduit de l'intégrale précédente

$$moy(x^k) = \int_0^{+\infty} x^k p(x)dx = \frac{1}{|r|} \frac{1}{\Gamma(a)} b^{-a} \int_0^{+\infty} x^{((a+kr)/r)-1} exp\left(-\frac{x^{1/r}}{b}\right)dx$$

■ *Moments de* $x \sim c+bG(a))$

D'où en particulier pour r=1 et c quelconque

$$moy(x) = ab + c \qquad var(x) = ab^2 \qquad \mu_3(x) = 2ab^3$$

9.1.3 Distributions khi-deux (centrée) et dérivées

Toutes les distributions présentées dans cette section sont des cas particuliers de la distribution *gamma-puissance* $b^r G^r(a)$ (avec c=0). Elles sont introduites pour dénommer de façon commode les distributions relatives aux écarts-types et variances.

■ *Distribution khi-deux*

La distribution classique *khi-deux* (centrée), à q degrés de liberté (où q est un nombre réel positif), d'indice d'échelle e^2, que nous noterons

$$e^2\chi_q^2$$

est la distribution $2e^2G(\frac{1}{2}q)$, soit r=1, $b=2e^2$, $a=\frac{1}{2}q$. Inversement la distribution bG(a) est la distribution $\frac{1}{2}\chi_{2a}$.

■ *Distributions dérivées*

Il est commode d'introduire les distributions dérivées, qui sont caractérisées par les valeurs suivantes (avec e>0 et q>0)

distribution	*notation*	r	b	a
khi-deux	$e^2\chi_q^2$	1	$2e^2$	$\frac{1}{2}q$
phi-deux	$e^2\varphi_q^2$	1	$2e^2/q$	$\frac{1}{2}q$
khi	$e\chi_q$	$\frac{1}{2}$	$2e^2$	$\frac{1}{2}q$
phi	$e\varphi_q$	$\frac{1}{2}$	$2e^2/q$	$\frac{1}{2}q$
khi-deux inverse	$e^2\chi_q^{-2}$	-1	$2/e^2$	$\frac{1}{2}q$
phi-deux inverse	$e^2\varphi_q^{-2}$	-1	$2/qe^2$	$\frac{1}{2}q$
khi inverse	$e\chi_q^{-1}$	$-\frac{1}{2}$	$2/e^2$	$\frac{1}{2}q$
phi inverse	$e\varphi_q^{-1}$	$-\frac{1}{2}$	$2/qe^2$	$\frac{1}{2}q$

Par conséquent, si $x \sim e^2\chi_q^2$ (e>0) alors

$x \sim e^2\chi_q^2$	$\frac{x}{q} \sim e^2\varphi_q^2$
$\frac{1}{x} \sim (1/e^2)\chi_q^{-2}$	$\frac{q}{x} \sim \frac{1}{e^2}\varphi_q^{-2}$
$x^{1/2} \sim e\chi_q$	$\left(\frac{x}{q}\right)^{1/2} \sim e\varphi_q$
$\left(\frac{1}{x}\right)^{1/2} \sim \frac{1}{e}\chi_q^{-1}$	$\left(\frac{q}{x}\right)^{1/2} \sim \frac{1}{e}\varphi_q^{-1}$

La densité, les moments et la fonction de répartition de chacune de ces distributions sont donnés par les résultats correspondants de la distribution $b^rG^r(a)$. Nous expliciterons les formules pour les distributions les plus utilisées dans cet ouvrage.

■ *Densités*

$e^2\varphi_q^2$	$p(x) = \frac{1}{\Gamma(\frac{1}{2}q)} (\frac{q}{2e^2})^{q/2} x^{q/2-1} exp(-\frac{q}{2e^2}x)$
$e\varphi_q$	$p(x) = \frac{2}{\Gamma(\frac{1}{2}q)} (\frac{q}{2e^2})^{q/2} x^{q-1} exp(-\frac{q}{2e^2}x^2)$
$e^2\varphi_q^{-2}$	$p(x) = \frac{1}{\Gamma(\frac{1}{2}q)} (\frac{qe^2}{2})^{q/2} x^{-(q/2+1)} exp(-\frac{qe^2}{2x})$
$e\varphi_q^{-1}$	$p(x) = \frac{2}{\Gamma(\frac{1}{2}q)} (\frac{qe^2}{2})^{q/2} x^{-(q+1)} exp(-\frac{qe^2}{2x^2})$

■ *Moyenne et variance*

	moyenne	*variance*
$e^2\varphi_q^2$	e^2	$\frac{2}{q}e^4$
$e\varphi_q$	$(\frac{2}{q})^{1/2} \frac{\Gamma(\frac{1}{2}(q+1))}{\Gamma(\frac{1}{2}q)} e$	e^2 - (*moyenne*)^2
$e^2\varphi_q^{-2}$	$\frac{q}{q-2}e^2$ [q>2]	$\frac{2q^2}{(q-2)^2(q-4)}e^4$ [q>4]
$e\varphi_q^{-1}$	$(\frac{q}{2})^{1/2} \frac{\Gamma(\frac{1}{2}(q-1))}{\Gamma(\frac{1}{2}q)} e$ [q>1]	$\frac{q}{q-2}e^2$ - (*moyenne*)^2 [q>2]

■ *Mode*

$e^2\varphi_q^2$	$(q-2)e^2$ [q≥2]
$e^2\varphi_q^{-2}$	$\frac{q}{q+2}e^2$
$e\varphi_q^{-1}$	$(\frac{q}{q+1})^{1/2} e$

■ *Fonction de répartition*

$e^2\varphi_q^2$	$F(x) = I\Gamma_{qx/2e^2}(\frac{1}{2}q)$
$e^2\varphi_q^{-2}$	$F(x) = 1-I\Gamma_{qe^2/2x}(\frac{1}{2}q)$
$e\varphi_q^{-1}$	$F(x) = 1-I\Gamma_{qe^2/2x^2}(\frac{1}{2}q)$

9.1.4 Distribution normale

■ *Notation et densité*

x a la distribution *normale* de centre (position) a et d'échelle e (e>0)

$$x \sim N(a,e^2)$$

si x a la densité

$$p(x) = \frac{1}{\sqrt{2\pi}}\frac{1}{e}\,exp\left(-\tfrac{1}{2}\left(\frac{x-a}{e}\right)^2\right)$$

■ *Propriétés*

Si $z \sim N(0,1)$, on vérifie que

$$x=a+ez \sim N(a,e^2) \quad \text{et} \quad z^2 \sim 2G(\tfrac{1}{2}) \sim \chi_1^2$$

■ *Moments de* $z \sim N(0,1)$ pour k>0 entier

si k impair $moy(z^k) = 0$

si k pair $moy(z^k) = (k-1)(k-3)\ldots 1$

En effet si k est impair

$$moy(z^k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} z^k\, exp(-\tfrac{1}{2}z^2)dz = 0$$

et si k est pair $moy(z^k)$ est le moment d'ordre k/2 de la distribution $2G(\frac{1}{2})$.

■ *Moments de* $x \sim N(a,e^2)$

En particulier

$$moy(x) = a \qquad moy(x^2) = a^2 + e^2$$

$$moy(x^3) = a^3 + 3ae^2 \qquad moy(x^4) = a^4 + 6a^2e^2 + 3e^4$$

d'où les premiers moments centrés

$$var(x) = e^2 \qquad \mu_3(x) = 0 \qquad \mu_4(x)$$

Ces résultats se déduisent des moments de $z \sim N(0,1)$ en posant $x=a+ez$

$$moy(x) = a + e\,moy(z) = a$$

$$moy(x^2) = a^2 + (2ae)moy(z) + e^2moy(z^2) = a^2 + e^2$$

etc.

■ *Fonction de répartition*

La fonction de répartition peut être obtenue à partir de la fonction *gamma incomplète*

$$F(x) = \tfrac{1}{2}(1-I\Gamma_u(\tfrac{1}{2})) \quad \text{si } x \geq a$$

$$F(x) = \tfrac{1}{2}(1+I\Gamma_u(\tfrac{1}{2})) \quad \text{si } x \leq a$$

$$\text{avec } u = \tfrac{1}{2}\left(\frac{x-a}{e}\right)^2$$

9.1.5 Distribution bêta

La distribution *bêta* n'est introduite ici que comme intermédiaire pour caractériser la distribution *F centrée*; elle sera sonc présentée brièvement.

■ *Définition et densité*

x a la distribution *bêta* de paramètres a et b (où a et b sont des réels positifs)

$$x \sim \beta(a,b)$$

si x a la densité

$$p(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\, x^{a-1}\,(1-x)^{b-1}$$

Si $y/c \sim \beta(a,b)$ $(c>0)$, alors nous dirons que *y* a la distribution *bêta* de paramètres a et b et d'échelle c, ce que nous noterons

$$y \sim c\beta(a,b)$$

y a la densité

$$p(y) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} c^{-a} y^{a-1} \left(1-\frac{y}{c}\right)^{b-1} = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} c^{-(a-b+1)} y^{a-1} (c-y)^{b-1}$$

A la densité de la fonction *bêta* sont liés les définitions et résultats suivants.

■ *Cas particulier: Distribution uniforme*

La distribution $c\beta(1,1)$ est la distribution *uniforme* sur l'intervalle [0,c].

■ *Fonction bêta*

La fonction *bêta* B(a,b) est définie par l'intégrale

$$B(a,b) = \int_0^1 u^{a-1}(1-u)^{b-1}du = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

■ *Moments de* $x \sim c\beta(a,b)$

$$moy(x^k) = \frac{\Gamma(a+k)\Gamma(a+b)}{\Gamma(a)\Gamma(a+b+k)} c^k = \frac{B(a+k,b)}{B(a,b)} c^k$$

formule définie pour tout réel k tel que a+k>0

Ce résultat se déduit de l'intégrale précédente (pour c=1)

$$moy(x^k) = \int_0^{+\infty} x^k p(x)dx = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^{+\infty} x^{a+k-1} (1-x)^{b-1}dx$$

D'où en particulier

$$moy(x) = \frac{a}{a+b} c \qquad var(x) = \frac{ab}{(a+b)^2(a+b+1)} c^2$$

■ *Fonction de répartition et fonction bêta incomplète*

La fonction *bêta incomplète* $IB_x(a)$ est définie par la fonction de répartition de la distribution *bêta*

$$IB_x(a,b) = F(x) = \frac{1}{B(a,b)} \int_0^x u^{a-1}(1-u)^{b-1}du$$

On fera attention au fait que la fonction *bêta incomplète* est définie par certains auteurs comme l'intégrale $\int_0^x u^{a-1}(1-u)^{b-1}du$.

9.1.6 Distribution F (centrée)

La distribution classique *F* (*centrée*) de Fisher-Snedecor peut être considérée dans le cadre fréquentiste comme un cas particulier de la distribution *F non-centrée* (ou *décentrée*), et dans le cadre bayésien comme un cas particulier de la distribution *psi-deux* (*cf.* Lecoutre, 1981a; Schervish, 1992; Lecoutre, Guigues et Poitevineau, 1992). Nous énoncerons ici les résultats sans démonstration.

■ *Caractérisation*

La distribution *F centrée* peut être caractérisée à partir de la distribution *bêta*. Si $x \sim \beta(a,b)$ alors la variable

$$F = e^2 \frac{bx}{a(1-x)}$$

a la distribution *F*, à L et q degrés de liberté (où L et q sont donc des réels postitifs), d'échelle e^2

$$F \sim e^2 F_{L,q}$$

Inversement, si $F \sim e^2 F_{L,q}$ alors

$$\frac{LF}{qe^2+LF} \sim \beta(\tfrac{1}{2}L, \tfrac{1}{2}q)$$

■ *Densité*

$$p(F) = \frac{\Gamma(\frac{1}{2}(L+q))}{\Gamma(\frac{1}{2}L)\Gamma(\frac{1}{2}q)} L^{L/2} (qe^2)^{q/2} F^{L/2-1} (qe^2+LF)^{-(L+q)/2}$$

$$= \frac{1}{B(\frac{1}{2}L, \frac{1}{2}q)} \left(\frac{L}{q}\right)^{L/2} \left(\frac{1}{e^2}\right)^{L/2} F^{L/2-1} \left(1+\frac{LF}{qe^2}\right)^{-(L+q)/2}$$

■ *Fonction de répartition*

La fonction de répartition est donnée par la fonction *bêta incomplète*

$$F(u) = Pr(e^2 F_{L,q} < u) = IB_{Lu/(Lu+qe^2)}(\tfrac{1}{2}L, \tfrac{1}{2}q)$$

■ *Moments*

$$moy(F^h) = \frac{\Gamma(\frac{1}{2}L+h)\Gamma(\frac{1}{2}q-h)}{\Gamma(\frac{1}{2}L)\Gamma(\frac{1}{2}q)} \left(\frac{qe^2}{L}\right)^h$$

formule définie pour tout réel h tel que $-L/2<h<q/2$

$$moy(F) = \frac{q}{q-2} e^2 \quad [q>2]$$

$$var(F) = \frac{2q^2(L+q-2)}{L(q-2)^2(q-4)} e^4 \quad [q>4]$$

$$\mu_3(F) = \frac{8q^3(2L+q-2)(L+q-2)}{L^2(q-2)^3(q-4)(q-6)} e^6 \quad [q>6]$$

■ *Distribution $eF_q^{1/2}$ et intégrale utile*

Si $F \sim e^2 F_{L,q}$ alors la variable $x=F^{1/2}$ a pour densité

$$p(x) = 2\frac{\Gamma(\frac{1}{2}(L+q))}{\Gamma(\frac{1}{2}L)\Gamma(\frac{1}{2}q)} \left(\frac{L}{q}\right)^{L/2} \left(\frac{1}{e^2}\right)^{L/2} x^{L-1} \left(1+\frac{Lx^2}{qe^2}\right)^{-(L+q)/2}$$

et nous noterons

$$x \sim eF_q^{1/2}$$

Nous en déduisons l'intégrale

$$\int_0^x u^{a-1}(1+u^2/c)^{-(a+b)/2}du = \tfrac{1}{2}B(\tfrac{1}{2}a,\tfrac{1}{2}b)\, c^{-a/2}\, IB_{x^2/(c+x^2)}(\tfrac{1}{2}a,\tfrac{1}{2}b)$$

9.2 DISTRIBUTION DE STUDENT GÉNÉRALISÉE

9.2.1 Caractérisations

■ *Première caractérisation*

Si $t|y^2 \sim N(a,e^2y^2)$ et $y^2 \sim \varphi_q^{-2}$ ($e>0$, $y>0$ et $q>0$) alors t a la distribution *t généralisée* à q degrés de liberté, de centre (position) a et d'échelle e. Nous noterons

$$t \sim t_q(a,e^2)$$

Pour a=0 et e=1, nous retrouvons la distribution t_q élémentaire usuelle; nous avons symboliquement

$$t_q(a,e^2) \equiv a+et_q$$

Cette distribution ne coïncide pas avec la distribution "*t non-centrée*" qui intervient en inférence fréquentiste (notamment à propos de la puissance du test de Student) que dans le cas particulier a=0 (voir ↳9.3).

■ *Deuxième caractérisation*

Si $x \sim N(0,1)$ et $y \sim \varphi_q$, x et y étant indépendantes, alors la variable t=a+e(x/y) a la distribution *t généralisée* à q degrés de liberté

$$t \sim t_q(a,e^2)$$

■ *Equivalence*

On vérifie l'équivalence de ces deux caractérisations. Sous les hypothèses de la première caractérisation la distribution

$$\frac{t-a}{ey} \mid y^2 \sim N(0,1) \quad [y>0]$$

ne dépend plus de y^2. En définissant z=(t-a)/ey et u=1/y, nous avons par conséquent les deux distributions indépendantes

$$z \sim N(0,1) \text{ et } u \sim \varphi_q$$

d'où, en utilisant la seconde caractérisation $a+(ez/u)=t \sim t_q(a,e^2)$

9.2.2 Densité

$$p(t) = \frac{1}{\sqrt{\pi q}} \frac{\Gamma(\frac{1}{2}(q+1))}{\Gamma(\frac{1}{2}q)} \frac{1}{e} \left(1+\frac{(t-a)^2}{qe^2}\right)^{-\frac{1}{2}(q+1)}$$

■ *Preuve*

Nous donnons la démonstration pour le cas particulier e=1. Le cas général s'en déduit en effectuant le changement de variable t→et.

En utilisant la première caractérisation

$$p(t \mid y^2) = \frac{1}{\sqrt{2\pi}} (y^2)^{-1/2} exp\left(-\tfrac{1}{2}\frac{(t-a)^2}{y^2}\right)$$

$$p(y^2) = \frac{1}{\Gamma(\frac{1}{2}q)} \left(\frac{q}{2}\right)^{q/2} (y^2)^{-(q/2+1)} exp(-\tfrac{1}{2}\frac{q}{y^2})$$

$$p(t) = \int p(t|y^2)p(y^2) = \frac{1}{\sqrt{\pi}}\frac{1}{\Gamma(\frac{1}{2}q)}\,2^{-(q+1)/2}\,(q^2)^{q/2} \times\int_0^{+\infty}(y^2)^{-\frac{1}{2}(q+1)-1}\,exp\left(-\tfrac{1}{2}(q^2+(t-a)^2)/y^2\right)dy^2$$

et en appliquant l'intégrale donnée en ↳9.1.2 avec

$$r = -1 \qquad a = \tfrac{1}{2}(q+1) \qquad 1/b = \tfrac{1}{2}(q+(t-a)^2)$$

il vient pour l'intégrale précédente

$$|r|\,\Gamma(a)\,b^a = \Gamma(\tfrac{1}{2}(q+1))\,(\frac{q}{2})^{-\frac{1}{2}(q+1)}\,\left(1+\frac{(t-a)^2}{q}\right)^{-\frac{1}{2}(q+1)}$$

d'où le résultat.

9.2.3 Moments

$$moy(t) = a$$

$$moy(t^2) = a^2 + \frac{q}{q-2}e^2 \quad [q>2]$$

$$moy(t^3) = a^3 + 3\frac{q}{q-2}ae^2 \quad [q>2]$$

$$moy(t^4) = a^4 + 6\frac{q}{q-2}a^2e^2 + 3\frac{q^2}{(q-2)(q-4)}e^4 \quad [q>4]$$

Ces moments se déduisent directement des moments correspondants de la distribution conditionnelle $t|y^2 \sim N(a,e^2y^2)$ et des moments de la distribution marginale $y^2 \sim \varphi_q^{-2}$ (soit $(2/q)\Gamma^{-1}(\frac{1}{2}q)$)

$$moy(t) = moy(moy(t|y^2)) = moy(a) = a$$

$$moy(t^2) = moy(moy(t^2|y^2)) = a^2 + e^2moy(y^2) = a^2 + \frac{q}{q-2}e^2 \quad [q>2]$$

etc.

Nous en déduisons les premiers moments centrés

$$var(t) = \frac{q}{q-2}e^2 \quad [q>2] \qquad \mu_3(t) = 0 \qquad \mu_4(t) = 3\frac{q^2}{(q-2)(q-4)}e^4$$

9.2.4 Fonction de répartition

■ *Distribution t_q et distribution $F_{1,q}$*

Si $t \sim t_q$ alors $t^2 \sim F_{1,q}$

■ *Fonction de répartition*

En conséquence du résultat précédent, la fonction de répartition de la distribution $t_q(a,e^2)$ peut être obtenue à partir de la fonction de répartition de la distribution F, et par suite de la fonction *bêta incomplète*

$$F(x) = \tfrac{1}{2}(1 - signe(x-a)IB_u(\tfrac{1}{2},\tfrac{1}{2}q)) \quad \text{avec } u = \frac{(x-a)^2}{(x-a)^2+qe^2}$$

9.3 Distribution de Student non-centrée

9.3.1 Caractérisations

■ *Première caractérisation*

si $t|y^2 \sim N(ay,b^2y^2)$ $(y>0,\ b>0)$ et $y^2 \sim \varphi_q^{-2}$ alors t a la distribution *t non-centrée* (ou *décentrée*) à q degrés de liberté, d'indice d'excentricité a et d'échelle b. Nous noterons

$$t \sim t'_q(a,b^2)$$

Pour a=0, nous retrouvons la distribution t_q élémentaire usuelle; nous avons symboliquement

$$t'_q(0,b^2) \equiv t_q(0,b^2) \equiv bt_q$$

■ *Deuxième caractérisation*

Si $z \sim N(a,b^2)$ et $u \sim \varphi_q$ (soit $y^2=1/u^2 \sim \varphi_q^{-2}$), z et u étant indépendantes, alors la variable t=z/u a la distribution *t non-centrée* à q degrés de liberté

$$t \sim t'_q(a,b^2)$$

■ *Equivalence*

On vérifie l'équivalence de ces deux caractérisations. Sous les hypothèses de la première caractérisation la distribution

$$\frac{t}{by}\Big| y^2 \sim N(a,1) \quad [y>0]$$

ne dépend plus de y^2. En définissant $z=t/by$ et $u=1/y^2$, nous avons par conséquent les deux distributions indépendantes

$$z \sim N(a,1) \text{ et } u \sim \varphi_q$$

d'où, en utilisant la seconde caractérisation $z/u=t \sim t'_q(a,b^2)$

■ *Cas particulier* b=1*: Notation* $t'_q(a)$

L'usage est de noter plus simplement $t'_q(a)$ pour $t'_q(a,1)$.

9.3.2 Propriétés

Des caractérisations précédentes et des propriétés de la distribution *normale*, nous déduisons immédiatement les propriétés suivantes, écrites symboliquement

$$t'_q(a,b^2) \equiv bt'_q(a/b,1)$$

$$t'_q(0,b^2) \equiv t(0,b^2)$$

$$t'_\infty(a,b^2) \equiv N(a,b^2)$$

$$Pr(t'_q(-x,b^2)<-a) = Pr(t'_q(x,b^2)>a)$$

$$Pr(t'_q(a,b^2)<0) = Pr(N(0,1)>\frac{a}{b})$$

■ *Preuve de la dernière propriété*

Pour $t \sim t'_q(a,b^2)$, nous avons sous les hypothèses de la première caractérisation

$t/y|y \sim N(a,b^2)$ qui ne dépend pas de y

Puisque y est une variable positive $Pr(t<0)=Pr(t/y<0)$, d'où le résultat.

9.3.3 Densité

$$p(t) = \frac{1}{\sqrt{\pi q}} \frac{1}{\Gamma(\frac{1}{2}q)} b^q \exp(-\tfrac{1}{2}a^2) \left(\frac{q}{qb^2+t^2}\right)^{(q+1)/2} \times \sum_{j=0}^{+\infty} \frac{1}{j!} \Gamma(\tfrac{1}{2}(q+j+1)) \left(\frac{2}{qb^2+t^2}\right)^{j/2} \left(a\frac{t}{b}\right)^j$$

■ *Preuve*

Nous donnons la démonstration pour le cas particulier b=1. Le cas général s'en déduit en effectuant le changement de variable t→bt et en changeant a en a/b.

En utilisant la première caractérisation

$$p(t\,|\,y^2) = \frac{1}{\sqrt{2\pi}} (y^2)^{-1/2} \exp(-\tfrac{1}{2}a^2) \exp\left(-\tfrac{1}{2}\frac{t^2}{y^2}\right) \exp\left(a\frac{t}{y}\right)$$

$$p(y^2) = \frac{1}{\Gamma(\frac{1}{2}q)} \left(\frac{q}{2}\right)^{q/2} (y^2)^{-(q/2)+1} \exp(-\tfrac{1}{2}\frac{q}{y^2})$$

$$p(t) = \int p(t\,|\,y^2)p(y^2)dy^2$$

en développant en série

$$\exp(a\frac{t}{y}) = \sum_{j=0}^{+\infty} \frac{1}{j!}(at)^j (y^2)^{-j/2}$$

et en intervertissant la sommation et l'intégration, nous obtenons

$$p(t) = \frac{1}{\sqrt{\pi}} \frac{1}{\Gamma(\frac{1}{2}q)} 2^{-(q+1)/2} q^{q/2} \exp(-\tfrac{1}{2}a^2) \times \sum_{j=0}^{+\infty} \frac{1}{j!} a^j t^j \int_0^{+\infty} (y^2)^{-\frac{1}{2}(q+j+1)-1} \exp\left(-\tfrac{1}{2}(q+t^2)/y^2\right) dy^2$$

En appliquant l'intégrale donnée en ↳9.1.2 avec

$$r = -1 \qquad a = \tfrac{1}{2}(q+j+1) \qquad 1/b = \tfrac{1}{2}(q+t^2)$$

il vient pour l'intégrale ci-dessus

$$|r|\,\Gamma(a)\,b^a = \Gamma(\tfrac{1}{2}(q+j+1)) \left(\frac{q+t^2}{2}\right)^{-(q+1)/2} \left(\frac{q+t^2}{2}\right)^{-j/2}$$

d'où le résultat.

9.3.4 Moments

$$moy(t) = a\hbar \quad [q>1]$$

$$moy(t^2) = \frac{q}{q-2}(a^2+b^2) \quad [q>2]$$

$$var(t) = \left(\frac{q}{q-2} - \hbar^2\right)a^2 + \frac{q}{q-2}b^2 \quad [q>2]$$

$$moy(t^3) = \frac{q}{q-3}(a^3+3ab^2)\hbar \quad [q>3]$$

$$moy(t^4) = \frac{q^2}{(q-2)(q-4)}(a^4+6a^2b^2+3b^4) \quad [q>4]$$

$$\text{avec } \hbar = \frac{\Gamma(\frac{1}{2}(q-1))}{\Gamma(\frac{1}{2}q)}(\tfrac{1}{2}q)^{1/2}$$

■ *Preuve*

Ces moments se déduisent directement des moments correspondants de la distribution conditionnelle $t|y^2 \sim N(ay, b^2y^2)$ et des moments de la distribution marginale $y^2 \sim \varphi_q^{-2}$ (soit $(2/q)\Gamma^{-1}(\frac{1}{2}q)$)

$$moy(t) = moy(moy(t|y^2)) = moy(ay) = a\,moy(y)$$

$$moy(t^2) = moy(moy(t^2|y^2)) = (a^2+b^2)moy(y^2) = (a^2+b^2)\frac{q}{q-2} \quad [q>2]$$

etc.

9.3.5 Fonction de répartition

Le $j^{ème}$ terme en t de la densité est de la forme $t^j(qb^2+t^2)^{-(q+j+1)/2}$ et correspond, au signe près, à une distribution *F*. En conséquence la fonction de répartition de la distribution *t non-centrée* peut être écrite comme une somme (infinie) de fonctions *bêta incomplètes*. Nous n'expliciterons pas ici ce résultat, car cette fonction de répartition peut également être obtenue à partir de celle de la distribution *lambda-prime*, introduite dans la section suivante.

9.4 DISTRIBUTION LAMBDA-PRIME

9.4.1 Caractérisations

■ *Première caractérisation*

Si $x|y^2 \sim N(a/y,b^2)$ $(y>0,\ b>0)$ et $y^2 \sim \varphi_q^{-2}$ (où a, b>0 et q>0 sont des réels), alors x a la distribution *lambda-prime* à q degrés de liberté, d'indice d'excentricité a et d'échelle b. Nous noterons

$$x \sim \Lambda'_q(a,b^2)$$

■ *Deuxième caractérisation*

Si $z \sim N(0,1)$ et $u \sim \varphi_q$ (soit $y^2=1/u^2 \sim \varphi_q^{-2}$), z et u étant indépendantes, alors la variable x=bz+au a la distribution *lambda-prime* à q degrés de liberté, d'indice d'excentricité a et d'échelle b.

En effet, sous ces hypothèses, nous avons $x|y \sim N(a/y,b^2)$, d'où le résultat en appliquant la première caractérisation.

9.4.2 Propriétés

Des caractérisations précédentes et des propriétés de la distribution *normale*, nous déduisons immédiatement les propriétés suivantes, écrites symboliquement.

$$\Lambda'_q(a,b^2) \equiv b\Lambda'_q(a/b,1)$$

$$\Lambda'_q(0,b^2) \equiv N(0,b^2)$$

$$\Lambda'_\infty(a,b^2) \equiv N(a,b^2)$$

$$\text{Si } x|y \sim N(ay,b^2) \text{ et } y^2 \sim c\varphi_q^2 \text{ alors } x \sim \Lambda'_q(ac,b^2)$$

$$Pr(\Lambda'_q(-x,b^2)<-a) = Pr(\Lambda'_q(x,b^2)>a)$$

$$Pr(\Lambda'_q(a,b^2)<0) = Pr(t_q>\frac{a}{b})$$

■ *Preuve de la dernière propriété*

Pour $x \sim \Lambda'_q(a,b^2)$, nous avons sous les hypothèses de la première caractérisation

$$xy \mid y \sim N(a,b^2y^2) \text{ et } y^2 \sim \varphi_q^{-2} \text{ donc } xy \sim t_q(a,b^2)$$

Puisque y est une variable positive $Pr(x<0)=Pr(xy<0)$, d'où le résultat.

9.4.3 Densité

$$p(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\Gamma(\frac{1}{2}q)} b^{q-1} \left(\frac{q}{qb^2+a^2}\right)^{q/2} exp(-\tfrac{1}{2}\frac{x^2}{b^2}) \times \sum_{j=0}^{+\infty} \frac{1}{j!} \Gamma(\tfrac{1}{2}(q+j)) \left(\frac{2}{qb^2+a^2}\right)^{j/2} \left(a\frac{x}{b}\right)^j$$

■ *Preuve*

Nous donnons la démonstration pour le cas particulier b=1. Le cas général s'en déduit en effectuant le changement de variable t→bt et en changeant a en a/b.

En utilisant la première caractérisation

$$p(x \mid y^2) = \frac{1}{\sqrt{2\pi}} exp(-\tfrac{1}{2}x^2) \, exp\left(-\tfrac{1}{2}\frac{a^2}{y^2}\right) exp\left(a\frac{x}{y}\right)$$

$$p(y^2) = \frac{1}{\Gamma(\frac{1}{2}q)} \left(\frac{q}{2}\right)^{q/2} (y^2)^{-(q/2+1)} exp(-\tfrac{1}{2}\frac{q}{y^2})$$

$$p(x) = \int p(x \mid y^2) p(y^2) dy^2$$

en développant en série

$$exp(a\frac{x}{y}) = \sum_{j=0}^{+\infty} \frac{1}{j!} (ax)^j (y^2)^{-j/2}$$

et en intervertissant la sommation et l'intégration, nous obtenons

$$p(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\Gamma(\frac{1}{2}q)} exp(-\tfrac{1}{2}x^2) \left(\frac{q}{2}\right)^{q/2} \times \sum_{j=0}^{+\infty} \frac{1}{j!} (ax)^j \int_0^{+\infty} (y^2)^{-\frac{1}{2}(q+j)-1} exp\left(-\tfrac{1}{2}(q+a^2)/y^2\right) dy^2$$

En appliquant l'intégrale donnée en §9.1.2 avec

$$r = -1 \qquad a = \tfrac{1}{2}(q+j) \qquad 1/b = \tfrac{1}{2}(q+a^2)$$

il vient pour l'intégrale ci-dessus

$$|r|\,\Gamma(a)\,b^a = \Gamma(\tfrac{1}{2}(q+j))\left(\frac{q+a^2}{2}\right)^{-(q+j)/2}$$

d'où le résultat.

9.4.4 Moments

$$moy(x) = ak$$

$$moy(x^2) = a^2 + b^2$$

$$var(x) = a^2(1-k^2) + b^2$$

$$moy(x^3) = \left(\frac{q+1}{q}a^3 + 3ab^2\right)k$$

$$\mu_3(x) = \left(2k^2 - \frac{2q-1}{q}\right)a^3k$$

$$moy(x^4) = \frac{q+2}{q}a^4 + 6a^2b^2 + 3b^4$$

$$\text{avec } k = \frac{\Gamma(\frac{1}{2}(q+1))}{\Gamma(\frac{1}{2}q)}\left(\frac{2}{q}\right)^{1/2}$$

■ *Preuve*

Ces moments se déduisent directement des moments correspondants de la distribution conditionnelle $x|y^2 \sim N(a/y, b^2)$ et des moments de la distribution marginale $y^2 \sim \varphi_q^{-2}$ (soit $(2/q)\Gamma^{-1}(\frac{1}{2}q)$)

$$moy(x) = moy(moy(x|y^2)) = moy(a/y) = a\,moy(y^{-1})$$

$$moy(x^2) = moy(moy(x^2|y^2)) = a^2 moy(y^{-2}) + b^2 = a^2 + b^2$$

etc.

9.4.5 Distribution t non-centrée et distribution lambda-prime

Les fonctions de répartition des distributions *t non-centrée* et *lambda-prime* sont liées par la relation remarquable (écrite symboliquement)

$$Pr(t'_q(a,b^2)<x) = Pr(\Lambda'_q(x,b^2)>a)$$

■ *Preuve*

En utilisant la deuxième caractérisation de la distribution *t non-centrée*, la condition "$t'_q(a,b^2)<x$" s'écrit

$$\frac{a+bz}{y} < x \quad \text{où } z \sim N(0,1) \text{ et } y \sim \varphi_q \text{ (z et y indépendantes)}$$

soit encore

$$bz-xy < -a$$

En utilisant la deuxième caractérisation de la distribution *lambda-prime*,

$$bz-xy \sim \Lambda'_q(-x,b^2)$$

d'où le résultat puisque $Pr(\Lambda'_q(-x,b^2)<-a) = Pr(\Lambda'_q(x,b^2)>a)$.

9.4.6 Fonction de répartition

Nous étudions la fonction de répartition pour b=1. Pour b quelconque on utilisera (écrit symboliquement)

$$Pr(\Lambda'_q(a,b^2)<x) = Pr(\Lambda'_q(a/b,1)<x/b)$$

- Si a=0 on utilisera

$$Pr(\Lambda'_q(0,1)<x) = Pr(N(0,1)<x)$$

- Si a>0 et x<0

$$F(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\Gamma(\frac{1}{2}q)} \left(\frac{q}{q+a^2}\right)^{q/2}$$
$$\times \sum_{j=0}^{+\infty} \frac{1}{j!} \Gamma(\tfrac{1}{2}(q+j))\, 2^{j/2} \left(\frac{a^2}{q+a^2}\right)^{j/2} (-1)^j \int_0^x u^j exp(-\tfrac{1}{2}u^2)du$$

En appliquant l'intégrale donnée en §9.1.2 avec

$$r = \tfrac{1}{2} \qquad a = \tfrac{1}{2}(1+j) \qquad b = \tfrac{1}{2}$$
$$\int_0^x u^j exp(-\tfrac{1}{2}u^2)du = \tfrac{1}{2}\Gamma(\tfrac{1}{2}(1+j))\, 2^{(j+1)/2}\, I\Gamma_{x^2/2}(\tfrac{1}{2}(j+1))$$

et en utilisant la formule de *duplication* (§9.1.1)

$$2^j \Gamma(\tfrac{1}{2}(1+j))\, \Gamma(1+\tfrac{1}{2}j) = \sqrt{\pi}\, j!$$

nous obtenons

$$F(x) = \sum_{j=0}^{+\infty} (-1)^j e_j \, I\Gamma_{x^2/2}(\tfrac{1}{2}(j+1))$$

$$\text{avec } e_j = \frac{1}{2} \frac{\Gamma(\frac{1}{2}(q+j))}{\Gamma(\frac{1}{2}q)\Gamma(1+\frac{1}{2}j)} \left(\frac{q}{q+a^2}\right)^{1/2} \left(\frac{a^2}{q+a^2}\right)^{j/2} \qquad (\sum_{j=0}^{+\infty} e_j = 1)$$

- Si a>0 et x>0, nous avons

$$F(x) = Pr(\Lambda'_q(a,1)<0) + Pr(0<\Lambda'_q(a,1)<x)$$

Le premier terme est donné par la fonction de répartition de la distribution t usuelle, en utilisant la propriété énoncée en ↳9.4.2

$$Pr(\Lambda'_q(a,1)<0) = Pr(t_q>a)$$

Le second terme est donné par

$$\frac{1}{\sqrt{2\pi}} \frac{1}{\Gamma(\frac{1}{2}q)} \left(\frac{q}{q+a^2}\right)^{q/2} \sum_{j=0}^{+\infty} \frac{1}{j!} \Gamma(\tfrac{1}{2}(q+j)) \, 2^{j/2} \left(\frac{a^2}{q+a^2}\right)^{j/2} \int_0^x u^j exp(-\tfrac{1}{2}u^2) du$$

soit en procédant comme dans le cas précédent

$$Pr(0<\Lambda'_q(a,1)<x) = \sum_{j=0}^{+\infty} e_j \, I\Gamma_{x^2/2}(\tfrac{1}{2}(j+1))$$

- Si a<0 on utilisera

$$Pr(\Lambda'_q(a,1)<x) = Pr(\Lambda'_q(-a,1)>-x) = 1 - Pr(\Lambda'_q(-a,1)<-x)$$

■ *Remarque*

Les séries précédentes peuvent être calculées par une double récurrence sur les coefficients e_j et sur les termes de la fonction *gamma incomplète*, en utilisant un algorithme analogue à celui développé pour la distribution *psi-deux* (qui intervient pour les effets multidimensionnels) dans Lecoutre, Guigues et Poitevineau (1992).

9.4.7 Distribution lambda-prime et distribution d'échantillonnage des limites de confiance fréquentistes

La distribution d'échantillonnage des limites de confiance fréquentistes pour le paramètre δ de la structure "*normale – phi-deux*" (voir le chapitre ↳10) met en jeu la distribution *lambda-prime*. En effet pour cette structure

nous avons les distributions d'échantillonnage indépendantes

$$d|\delta,\sigma^2 \sim N(\delta,b^2\sigma^2) \quad \text{et} \quad s^2|\delta,\sigma^2 \sim \sigma^2\varphi_q^2$$

soit

$$\frac{d-\delta}{b\sigma}|\delta,\sigma^2 \sim N(0,1) \quad \text{et} \quad \frac{s^2}{\sigma^2}|\delta,\sigma^2 \sim \varphi_q^2$$

Nous en déduisons, en appliquant la deuxième caractérisation, que pour toute constante c

$$d+cs = \delta+b\sigma\frac{d-\delta}{b\sigma}+c\sigma\frac{s}{\sigma} \;|\delta,\sigma^2 \sim \delta+\Lambda_q'(c\sigma,b^2\sigma^2)$$

Nous en déduisons la propriété fondamentale suivante de la distribution d'échantillonnage. La probabilité que la variable $d+cs$ soit inférieure à la valeur δ ne dépend ni de δ ni de σ, mais seulement du rapport c/b. Plus précisément nous avons, en appliquant la propriété donnée en §9.4.2

$$Pr(d+cs<\delta|\delta,\sigma^2) = Pr(\Lambda_q'(c\sigma,b^2\sigma^2)<0) = Pr(t_q>\frac{c}{b})$$

Nous en déduisons, en notant t_α la valeur critique positive (*bilatérale*) telle que $Pr(|t_q|>t_\alpha)=\alpha$ soit $Pr(t_q>t_\alpha)=\alpha/2$

$$Pr(d+cs<\delta|\delta,\sigma^2) = \alpha/2 \Leftrightarrow c = bt_\alpha$$

et, en raison de la symétrie de la distribution t

$$Pr(d+cs>\delta|\delta,\sigma^2) = \alpha/2 \Leftrightarrow c = -bt_\alpha$$

Nous retrouvons ainsi les statistiques "limites de confiance $1-\alpha$ fréquentistes usuelles pour le paramètre δ"

$$\textit{limite inférieure} \quad l_{1-\alpha} = d-bst_\alpha$$

$$\textit{limite supérieure} \quad l^{1-\alpha} = d+bst_\alpha$$

avec les distributions et les probabilités d'échantillonnage

$$l_{1-\alpha}|\delta,\sigma^2 \sim \delta-b\sigma\Lambda_q'(t_\alpha,1) \quad \text{et} \quad l^{1-\alpha}|\delta,\sigma^2 \sim \delta+b\sigma\Lambda_q'(t_\alpha,1)$$

$$Pr(l_{1-\alpha}>\delta|\delta,\sigma^2) = \alpha/2 \quad \text{et} \quad Pr(l^{1-\alpha}<\delta|\delta,\sigma^2) = \alpha/2$$

9.5 Distribution K-prime

9.5.1 Caractérisations

Si l'un des quatre ensembles de conditions suivant est rempli (a, b>0 et q>0 et r>0 sont des réels, et y>0):

- $x|y^2 \sim \Lambda'_q(ay, b^2y^2)$ et $y^2 \sim \varphi_r^{-2}$
- $x=w/u$ avec $w \sim \Lambda'_q(a,b^2)$ et $u \sim \varphi_r$ $(1/u^2 \sim \varphi_r^{-2})$, w et u étant indépendantes
- $x|y^2 \sim t'_r(a/y,b^2)$ et $y^2 \sim \varphi_q^{-2}$
- $x|y^2 \sim t_{q+r}(ay, b^2\frac{qy^2+r}{q+r})$ et $y^2 \sim F_{q,r}$

alors x a la distribution *K-prime* à q et r degrés de liberté, d'indice d'excentricité a et d'échelle b. Nous noterons

$$x \sim K'_{q,r}(a,b^2)$$

■ *Equivalence des caractérisations*

Considérons trois variables x, g^2 et h^2 telles que (g>0 h>0)

$$x|g^2,h^2 \sim N(ah/g, b^2h^2)$$

$$g^2 \sim \varphi_q^{-2} \quad \text{et} \quad h^2 \sim \varphi_r^{-2} \quad \text{avec } g^2 \text{ et } h^2 \text{ indépendantes}$$

alors $x \sim K'_{q,r}(a,b^2)$ et ce résultat s'obtient en appliquant chacune des quatre caractérisations. Pour les trois premières, il se déduit respectivement de

$$x|h^2 \sim \Lambda'_q(ah, b^2h^2)$$

d'où encore $w=\frac{x}{h}|h^2 \sim \Lambda'_q(a,b^2)$ qui ne dépend plus de h^2

$$x|g^2 \sim t'_q(a/g, b^2)$$

Pour la quatrième, en posant

$$u = \frac{h}{g} \quad \text{et} \quad v^2 = \frac{q+r}{(q/g^2)+(r/h^2)}$$

nous obtenons

$$x|u^2,v^2 \sim x|g^2,h^2 \sim N(au, b^2\frac{qu^2+r}{q+r}v^2)$$

et, en appliquant le lemme donné en §9.6.1 pour $q/g^2 \sim \chi^2_q$ et $r/h^2 \sim \chi^2_r$

$$u^2 \sim F_{q,r} \quad \text{et} \quad v^2 \sim \varphi_{q+r}^{-2} \quad \text{avec } u^2 \text{ et } v^2 \text{ indépendantes}$$

d'où

$$x|u^2 \sim t_{q+r}(au, b^2\frac{qu^2+r}{q+r})$$

9.5.2 Propriétés

Des caractérisations précédentes et des propriétés des distributions *lambda-prime*, *t non-centrée* et *t généralisée*, nous déduisons immédiatement les propriétés suivantes, écrites symboliquement

$$K'_{q,r}(a,b^2) = bK'_{q,r}(a/b,1)$$

$$K'_{q,r}(0,b^2) = t_r(0,b^2)$$

$$K'_{q,\infty}(a,b^2) = \Lambda'_q(a,b^2)$$

$$K'_{\infty,r}(a,b^2) = t'_r(a,b^2)$$

$$Pr(K'_{q,r}(-x,b^2)<-a) = Pr(K'_{q,r}(x,b^2)>a)$$

$$Pr(K'_{q,r}(a,b^2)<0) = Pr(\Lambda'_q(a,b^2)<0) = Pr(t_q>\frac{a}{b})$$

■ *Preuve de la dernière propriété*

Pour $x \sim K'_{q,r}(a,b^2)$, nous avons sous les hypothèses de la première caractérisation

$x/y|y^2 \sim \Lambda'_q(a,b^2)$ qui ne dépend pas de y

Puisque y est une variable positive $Pr(x<0)=Pr(x/y<0)$, d'où le résultat.

9.5.3 Relation remarquable

Les fonctions de répartition des distributions *K-prime* à q et r degrés de liberté et *K-prime* à r et q degrés de liberté sont liées par la relation remarquable (écrite symboliquement)

$$Pr(K'_{r,q}(a,b^2)<x) = Pr(K'_{q,r}(x,b^2)>a)$$

qui généralise la relation donnée en ↳9.4.5

$$Pr(t'_q(a,b^2)<x) = Pr(\Lambda'_q(x,b^2)>a)$$

obtenue pour $r=\infty$.

■ *Preuve*

Considérons trois variables indépendantes z, g et h telles que

$$z \sim N(0,1) \qquad g^2 \sim \varphi^2_r \qquad h^2 \sim \varphi^2_q$$

Nous en déduisons, en appliquant les caractérisations des distributions *lambda-prime* et *K-prime*

$$bz+ag \sim \Lambda'_r(a,b^2) \quad \text{d'où} \quad \frac{bz+ag}{h} \sim K'_{r,q}(a,b^2)$$

$$bz-xh \sim \Lambda'_r(-x,b^2) \quad \text{d'où} \quad \frac{bz-xh}{g} \sim K'_{q,r}(-x,b^2)$$

Le résultat se déduit de l'équivalence des deux conditions

$$\frac{bz+ag}{h} < x \quad \text{et} \quad \frac{bz-xh}{g_1} < -a$$

puisque $Pr(K'_{q,r}(-x,b^2)<-a) = Pr(K'_{q,r}(x,b^2)>a)$.

9.5.4 Densité

$$p(x) = \frac{1}{\sqrt{\pi r}} \frac{1}{\Gamma(\frac{1}{2}q)\Gamma(\frac{1}{2}r)} b^{q+r} \left(\frac{q}{qb^2+a^2}\right)^{q/2} \left(\frac{r}{rb^2+x^2}\right)^{(r+1)/2}$$

$$\times \sum_{j=0}^{+\infty} \frac{1}{j!} \Gamma(\tfrac{1}{2}(q+j))\, \Gamma(\tfrac{1}{2}(r+j+1)) \left(\frac{4}{(qb^2+a^2)(rb^2+x^2)}\right)^{j/2} \left(a\frac{x}{b}\right)^j$$

■ *Preuve*

Nous donnons la démonstration pour le cas particulier b=1. Le cas général s'en déduit en effectuant le changement de variable t→bt et en changeant a en a/b.

En utilisant la caractérisation à partir de la distribution *t non-centrée*

$$p(x|y^2) = \frac{1}{\sqrt{\pi r}} \frac{1}{\Gamma(\frac{1}{2}r)} exp(-\tfrac{1}{2}a^2y^{-2}) \left(\frac{r}{r+x^2}\right)^{(r+1)/2}$$
$$\times \sum_{j=0}^{+\infty} \frac{1}{j!} \Gamma(\tfrac{1}{2}(r+j+1)) \, a^j x^j \left(\frac{2}{r+x^2}\right)^{j/2} (y^2)^{-j/2}$$

$$p(y^2) = \frac{1}{\Gamma(\frac{1}{2}q)} \left(\frac{q}{2}\right)^{q/2} (y^2)^{-(q/2+1)} exp(-\tfrac{1}{2}\frac{q}{y^2})$$

$$p(x) = \int p(x|y^2)p(y^2)dy^2$$

et, en intervertissant la sommation et l'intégration, nous obtenons

$$p(x) = \frac{1}{\sqrt{\pi r}} \frac{1}{\Gamma(\frac{1}{2}q)\Gamma(\frac{1}{2}r)} \left(\frac{q}{2}\right)^{q/2} \left(\frac{r}{r+x^2}\right)^{(r+1)/2}$$
$$\times \sum_{j=0}^{+\infty} \frac{1}{j!} \Gamma(\tfrac{1}{2}(r+j+1)) \, a^j x^j \left(\frac{2}{(r+x^2)}\right)^{j/2} \int_0^{+\infty} (y^2)^{-\frac{1}{2}(q+j)-1} exp(-\tfrac{1}{2}(q+a^2)/y^2)dy^2$$

L'intégrale est la même que celle qui intervenait pour la densité de la distribution *lambda-prime* en §9.4.3, d'où le résultat.

9.5.5 Moments

$$moy(x) = ak \quad [q>1]$$

$$moy(x^2) = \frac{r}{r-2}(a^2+b^2) \quad [r>2]$$

$$var(x) = (\frac{r}{r-2} - k^2)a^2 + \frac{r}{r-2} b^2 \quad [r>2]$$

$$moy(x^3) = \frac{r}{r-3}(\frac{q+1}{q} a^3 + 3ab^2)k \quad [r>3]$$

$$moy(x^4) = \frac{r^2}{(r-2)(r-4)}(\frac{q+2}{q} a^4 + 6a^2b^2 + 3b^4) \quad [r>4]$$

$$\text{avec } k = \frac{\Gamma((q+1)/2)\Gamma((r-1)/2)}{\Gamma(q/2)\Gamma(r/2)} \left(\frac{r}{q}\right)^{1/2}$$

■ *Preuve*

Ces moments se déduisent directement à partir de l'une des caractérisations, par exemple des moments correspondants de la distribution conditionnelle $x|y^2 \sim t'_r(a/y,b^2)$ et des moments de la distribution marginale $y^2 \sim e^2\varphi_q^{-2}$

$$moy(x) = moy(moy(x|y^2)) = moy(aq/y) = aq\, moy(y^{-1})$$

$$\text{avec } q = \frac{\Gamma(\frac{1}{2}(r-1))}{\Gamma(\frac{1}{2}r)}(\tfrac{1}{2}r)^{1/2}$$

$$moy(x^2) = moy(moy(x^2|y^2)) = \frac{r}{r-2}(a^2 moy(y^{-2}) + b^2) \quad [r>2]$$

etc.

9.5.6 Fonction de répartition

Nous étudions la fonction de répartition pour b=1. Pour b quelconque on utilisera (écrit symboliquement)

$$Pr(K'_{q,r}(a,b^2)<x) = Pr(K'_{q,r}(a/b,1)<x/b)$$

- Si a=0 on utilisera

$$Pr(K'_{q,r}(0,1)<x) = Pr(t_r(0,1)<x)$$

- Si a>0 et x<0

$$F(x) = \frac{1}{\sqrt{\pi}}\frac{1}{\Gamma(\frac{1}{2}q)\Gamma(\frac{1}{2}r)}\left(\frac{q}{q+a^2}\right)^{q/2}$$
$$\times \sum_{j=0}^{+\infty}\frac{1}{j!}\Gamma(\tfrac{1}{2}(q+j))\Gamma(\tfrac{1}{2}(r+j+1))\, r^{(j+1/2)}\, 2^j \left(\frac{a^2}{q+a^2}\right)^{j/2}(-1)^j$$
$$\times \int_0^x u^j(1+u^2/r)^{-(r+j+1)/2}\,du$$

En appliquant l'intégrale donnée en §9.1.6 avec

$$c = r \qquad a = j+1 \qquad b = r$$

$$\int_0^x u^j(1+u^2/r)^{-(r+j+1)/2}\,du$$
$$= \tfrac{1}{2}\frac{\Gamma(\frac{1}{2}(1+j))\Gamma(\frac{1}{2}r)}{\Gamma(\frac{1}{2}(r+j+1))}\, r^{-(j+1)/2}\, IB_{x^2/(r+x^2)}(\tfrac{1}{2}(j+1),\tfrac{1}{2}r)$$

et en utilisant la formule de *duplication* (↳9.1.1)

$$2^j\,\Gamma(\tfrac{1}{2}(1+j))\,\Gamma(1+\tfrac{1}{2}j) = \sqrt{\pi}\;j!$$

nous obtenons

$$F(x) = \sum_{j=0}^{+\infty}(-1)^j\,e_j\,IB_{x^2/(r+x^2)}(\tfrac{1}{2}(j+1),\tfrac{1}{2}r)$$

avec les mêmes coefficients e_j que pour la distribution *lambda-prime*

$$e_j = \frac{1}{2}\,\frac{\Gamma(\frac{1}{2}(q+j))}{\Gamma(\frac{1}{2}q)\Gamma(1+\frac{1}{2}j)}\left(\frac{q}{q+a^2}\right)^{q/2}\left(\frac{a^2}{q+a^2}\right)^{j/2} \qquad \left(\sum_{j=0}^{+\infty}e_j = 1\right)$$

- Si a>0 et x>0, nous avons

$$F(x) = Pr\{K'_{q,r}(a,1)<0\} + Pr\{0<K'_{q,r}(a,1)<x\}$$

Le premier terme est donné par la fonction de répartition de la distribution t usuelle, en utilisant la propriété énoncée en ↳9.5.2

$$Pr\{K'_{q,r}(a,1)<0\} = Pr(t_q>a)$$

Le second terme est donné par

$$\frac{1}{\sqrt{\pi}}\,\frac{1}{\Gamma(\frac{1}{2}q)\Gamma(\frac{1}{2}r)}\left(\frac{q}{q+a^2}\right)^{q/2}$$

$$\times\sum_{j=0}^{+\infty}\frac{1}{j!}\,\Gamma(\tfrac{1}{2}(q+j))\,\Gamma(\tfrac{1}{2}(r+j+1))\,r^{(j+1)/2}\,2^j\left(\frac{a^2}{q+a^2}\right)^{j/2}$$

$$\times\int_0^x u^j(1+u^2/r)^{-(r+j+1)/2}\,du$$

soit en procédant comme dans le cas précédent

$$Pr\{0<K'_{q,r}(a,1)<x\} = \sum_{j=0}^{+\infty}e_j\,IB_{x^2/(r+x^2)}(\tfrac{1}{2}(j+1),\tfrac{1}{2}r)$$

- Si a<0 on utilisera

$$Pr\{K'_{q,r}(a,1)<x\} = Pr\{K'_{q,r}(-a,1)>-x\} = 1-Pr\{K'_{q,r}(-a,1)<-x\}$$

■ *Remarque*

Comme pour la distribution *lambda-prime*, les séries précédentes peuvent être calculées par une double récurrence sur les coefficients e_j et sur les termes

de la fonction *béta incomplète*.

9.6 Divers résultats utiles

9.6.1 Un lemme sur les distributions χ^2 et F

Si les variables x et y sont indépendantes, de distributions respectives

$$x \sim \chi^2_p \quad \text{et} \quad y \sim \chi^2_q$$

alors, les variables x+y et x/y sont également indépendantes et ont les distributions respectives

$$x+y \sim \chi^2_{p+q} \quad \text{et} \quad \frac{x}{y} \sim \frac{p}{q} F_{p,q}$$

■ *Preuve*

Nous avons la densité conjointe

$$p(x,y) = p(x)p(y) \propto x^{p/2-1}\, y^{-q/2-1}\, exp(-\tfrac{1}{2}(x+y))$$

Effectuons le changement de variable u=x+y, v=x/y, de Jacobien $u/(1+v)^2$

$$p(u,v) \propto u^{(p+q)/2-1}\, exp(-\tfrac{1}{2}u)\, v^{p/2-1}\, (1+v)^{-(p+q)/2}$$

Nous identifions dans l'expression précédente:

- le terme en u de la densité d'une variable distribuée χ^2_{p+q}
- le terme en v de la densité d'une variable distribuée $(p/q)F_{p,q}$, d'où le résultat.

9.6.2 Approximations d'une distribution de moyenne M et de variance V

■ *Par une distribution* $c\varphi_q^{-2}$

Cette approximation est utilisée pour la distribution finale d'une combinaison linéaire de variances dans les solutions "sans égalité des variances" développées au chapitre ↳6.

Nous devons avoir (symboliquement)

$$moy(c\varphi_q^{-2}) = c\frac{q}{q-2} = M \quad \text{et} \quad var(c\varphi_q^{-2}) = c^2\frac{2q^2}{(q-2)^2(q-4)} = V$$

dont nous déduisons

$$q = 4 + 2\frac{M^2}{V} \quad \text{et} \quad c = \frac{q-2}{q}M$$

■ *Par une distribution* $cF_{p,q}$ *(p fixé)*

Nous devons avoir (symboliquement)

$$moy(cF_{p,q}) = c\frac{q}{q-2} = M \quad \text{et} \quad var(cF_{p,q}) = c^2\frac{2q^2(p+q-2)}{p(q-2)^2(q-4)} = V$$

dont nous déduisons

$$q = 2\frac{2pV+(p-2)M^2}{pV-2M^2} \quad \text{et} \quad c = \frac{q-2}{q}M$$

avec la condition $pV > 2M^2$

9.6.3 Approximation d'une distribution par une distribution c+bΓ(a)

Cette approximation est utilisable pour toute distribution d'une variable x de moyenne M, variance V et troisième moment centré non nul W. Nous devons donc avoir

$$moy(x) = ab + c = M \qquad var(x) = ab^2 = V \qquad \mu_3(x) = 2ab^3 = W$$

dont nous en déduisons

$$a = 4\frac{V^3}{W^2} \qquad b = \tfrac{1}{2}\frac{W}{V} \qquad c = M - 2\frac{V^2}{W}$$

Nous avons donc l'approximation de la fonction de répartition $F(x)$ par

$$I\Gamma_{(x-c)/b}(a)$$

9.6.4 Un résultat sur les distributions conditionnelles

Nous énoncerons sans démonstration le lemme suivant, qui est classique dans le cas d'une distribution normale (dans ce cas la distribution conditionnelle est elle-même normale).

Si la variable vectorielle bidimensionnelle [x,y]' a une distribution

• de moyenne [a,b]'

• de matrice de variances et covariances définie positive $\begin{bmatrix} e_{11} & e_{12} \\ e_{12} & e_{22} \end{bmatrix}$

alors nous avons pour la distribution conditionnelle $y|x$

• la moyenne $cx + d$ avec $c = e_{12}/e_{11}$ et $d = b - ca$

• la variance $e_{22} - c^2 e_{11}$

CHAPITRE 10
SOLUTIONS BAYÉSIENNES

Les procédures bayésiennes mises en œuvre dans cet ouvrage sont une généralisation directe des solutions bayésiennes de base pour l'inférence sur la moyenne d'une distribution normale. Nous supposons ici, sans démonstration, l'existence de statistiques *exhaustives* (voir l'annexe de l'introduction), dont les distributions d'échantillonnage caractérisent deux structures statistiques remarquables, que nous appellerons "*normale - phi-deux*" et "*normale - (phi-deux)*s". L'étude de ces deux structures fournit les solutions, respectivement "avec égalité des variances" et "sans égalité des variances".

Ces solutions mettent en jeu, outre les distributions classiques utilisées dans les tests de signification, les distributions *t généralisée*, *lambda-prime*, et *K-prime* (cette dernière intervenant dans l'inférence prédictive). Toutes ces distributions sont présentées dans le chapitre ↳9.

10.1 Structure "normale - phi-deux": Solutions générales

10.1.1 Définition

La structure statistique "*normale - phi-deux*" est définie par la donnée d'un paramètre numérique $\delta \in \mathbb{R}$ et d'un paramètre numérique positif $\sigma^2 \in \mathbb{R}_+$.

■ *Distributions d'échantillonnage*

Sous un certain modèle d'échantillonnage, les statistiques d et s^2 sont conjointement exhaustives pour (δ,σ^2) et ont les distributions d'échantillonnage indépendantes

$$\left\| \begin{array}{l} d|\delta,\sigma^2 \sim N(\delta, b^2\sigma^2) \\ s^2|\delta,\sigma^2 \sim \sigma^2\varphi_q^2 \quad (\text{soit } \frac{\sigma^2}{q}\chi_q^2) \end{array} \right.$$

Les distributions d'échantillonnage de d et s^2 ont les densités respectives

$$p(d|\delta,\sigma^2) = \left(\frac{1}{2\pi b^2\sigma^2}\right)^{1/2} exp\left(-\tfrac{1}{2}\frac{(d-\delta)^2}{b^2\sigma^2}\right)$$

$$p(s^2|\delta,\sigma^2) = \frac{1}{\Gamma(q/2)}\left(\frac{2\sigma^2}{q}\right)^{-q/2}(s^2)^{q/2-1}\, exp\left(-\tfrac{1}{2}q\frac{s^2}{\sigma^2}\right)$$

10.1.2 Inférence relative à δ conditionnellement à σ^2 et à d

Supposons que σ^2 est une donnée, et cherchons la distribution bayésienne finale relative au paramètre δ conditionnellement à d (et à σ^2 que nous omettons pour simplifier les notations). Pour une distribution initiale pour δ de densité $p(\delta)$, cette distribution, donnée par la formule de Bayes, a une densité

$$p(\delta|d) = \frac{p(d,\delta)}{p(d)} = \frac{p(d|\delta)p(\delta)}{\int p(d|\delta)p(\delta)d\delta}$$

Une fois fixée la densité initiale $p(\delta)$:
- $p(d,\delta)$ est la densité de la distribution *conjointe* du couple (d,δ);
- $p(d)$ est la densité de la distribution *prédictive* de d.

Comme fonction de δ, la densité finale $p(\delta|d)$ est donc proportionnelle à la densité conjointe $p(d,\delta)=p(d|\delta)p(\delta)$, ce que nous indiquons avec le symbole $\propto$

$$p(\delta|d) \propto exp\left(-\tfrac{1}{2}\frac{(\delta-d)^2}{b^2\sigma^2}\right)p(\delta)$$

Cette distribution bayésienne sur le paramètre δ exprime l'information apportée sur δ par d, donc par l'échantillon (conditionnellement à σ^2), en prenant éventuellement en compte une information extérieure dans la distribution initiale.

■ *Solution standard ou fiducio-bayésienne*

Choisissons pour δ la distribution initiale non-informative usuelle uniforme pour δ, c'est-à-dire de densité $p(\delta)$ constante sur un intervalle arbitrairement grand, que nous faisons tendre vers $[-\infty,+\infty]$. Ce choix revient à écarter (ou ignorer) toute information extérieure aux données.

Pour cette distribution initiale nous avons

$$p(\delta \mid d) \propto exp\left(-\tfrac{1}{2}\frac{(\delta-d)^2}{b^2\sigma^2}\right)$$

où nous identifions le terme en δ de la densité d'une distribution $N(d, b^2\sigma^2)$, d'où le résultat

$$\delta \mid d \sim N(d, b^2\sigma^2)$$

On observera ici que ce choix fait de la densité conjointe une fonction symétrique de δ et de $-d$.

On remarquera encore que, dans cette solution, on passe de la distribution d'échantillonnage

$$d \mid \delta \sim N(\delta, b^2\sigma^2) \quad \text{soit} \quad d-\delta \mid \delta \sim N(0, b^2\sigma^2)$$

(où d est la variable et la valeur δ est fixée)

à la distribution finale

$$\delta \mid d \sim N(d, b^2\sigma^2) \quad \text{soit} \quad \delta-d \mid d \sim N(0, b^2\sigma^2)$$

(où δ est la variable et la valeur d est fixée)

Ce passage, qui explicite le rôle symétrique joué par δ et $-d$ (compte tenu de la symétrie de la distribution), correspond à un *pivot* intuitif et peut en ce sens être justifié par l'argument *fiduciaire* de Fisher (*cf.* Fisher, 1990). Nous appellerons cette solution la solution bayésienne *standard* ou encore la solution *fiducio-bayésienne* (sur cette appellation, voir l'introduction).

■ *Distribution initiale conjuguée naturelle*

Si nous choisissons pour δ la distribution initiale ("informative") de la même famille que la distribution fiducio-bayésienne (distribution *conjuguée naturelle*)

$$\delta \sim N(d_0, b_0^2\sigma^2)$$

où b_0 est une constante positive (rappel: σ^2 est une donnée)

la densité finale devient proportionnelle à

$$exp\left(-\tfrac{1}{2}\frac{(\delta-\delta)^2}{b^2\sigma^2}\right) exp\left(-\tfrac{1}{2}\frac{(\delta-\delta_0)^2}{b_0^2\sigma^2}\right)$$

donc encore, après quelques calculs élémentaires, à

$$exp\left(-\tfrac{1}{2}\frac{(\delta-d_1)^2}{b_1^2\sigma^2}\right) \qquad \text{avec } d_1 = \frac{b_0^2 d+b^2 d_0}{b_0^2+b^2} \text{ et } b_1^2 = \frac{b_0^2 b^2}{b_0^2+b^2}$$

d'où le résultat

$$\delta \mid d \sim N(d_1, b_1^2\sigma^2)$$

qui donne comme cas limite, quand $b_0 \to \infty$, la solution fiducio-bayésienne.

10.1.3 Inférence relative à σ^2 conditionnellement à s^2

Cherchons la distribution bayésienne finale relative au paramètre σ^2 conditionnellement à s^2.

Pour une distribution initiale de densité $p(\sigma^2)$, cette distribution finale, donnée par la formule de Bayes, a une densité

$$p(\sigma^2 \mid s^2) = \frac{p(s^2,\sigma^2)}{p(s^2)} = \frac{p(s^2 \mid \sigma^2)\,p(\sigma^2)}{\int p(s^2 \mid \sigma^2)p(\sigma^2)d\sigma^2}$$

Une fois fixée la densité initiale $p(\sigma^2)$:
- $p(s^2,\sigma^2)$ est la densité de la distribution *conjointe* du couple (s^2,σ^2);
- $p(d)$ est la densité de la distribution *prédictive* de s^2.

Comme fonction de σ^2, la densité finale $p(\sigma^2 \mid s^2)$ est donc proportionnelle à la densité conjointe $p(s^2,\sigma^2)=p(s^2 \mid \sigma^2)p(\sigma^2)$

$$p(\sigma^2 \mid s^2) \propto (\sigma^2)^{-q/2} exp(-\tfrac{1}{2}q\frac{s^2}{\sigma^2})p(\sigma^2)$$

Cette distribution bayésienne sur le paramètre σ^2 exprime l'information apportée par s^2 sur σ^2, en prenant éventuellement en compte une information extérieure dans la distribution initiale. On remarquera qu'en général, comme nous le verrons plus loin, d apporte une information sur σ^2, et par conséquent cette distribution diffère de la distribution finale conditionnelle à s^2 et à d (donc à l'échantillon).

■ *Solution standard ou fiducio-bayésienne*

Choisissons pour σ^2 la distribution initiale non-informative usuelle uniforme pour $log(\sigma^2)$, c'est-à-dire de densité $p(\sigma^2)$ proportionnelle à $1/\sigma^2$ sur un intervalle arbitrairement grand (que l'on fait tendre vers $]0,+\infty[$). Ceci revient à écarter (ou ignorer) toute information extérieure aux données.

On observera que ce choix fait de la densité conjointe de la statistique s^2 et du paramètre $\sigma^{-2}=1/\sigma^2$ une fonction symétrique de ces deux quantités. Nous avons en effet $p(s^2 \mid \sigma^{-2})=p(s^2 \mid \sigma^2)$ et $p(\sigma^{-2})\propto 1/\sigma^{-2}$, d'où

$$p(s^2,\sigma^{-2}) \propto (s^2)^{q/2-1} (\sigma^{-2})^{q/2-1} exp(-\tfrac{1}{2}qs^2\sigma^{-2})$$

Il résulte de ce choix

$$p(\sigma^2|s^2) \propto (\sigma^2)^{-(q/2+1)} exp(-\tfrac{1}{2}q\frac{s^2}{\sigma^2})$$

où nous identifions le terme en σ^2 de la densité d'une distribution *phi-deux inverse* d'échelle s^2, d'où le résultat

$$\sigma^2|s^2 \sim s^2\varphi_q^{-2}$$

On remarquera que, dans cette solution nous passons de la distribution d'échantillonnage

$$\frac{s^2}{\sigma^2}|\sigma^2 \sim \varphi_q^2$$

(où s^2 est la variable et la valeur σ^2 est fixée)

à la distribution finale

$$\frac{s^2}{\sigma^2}|s^2 \sim \varphi_q^2$$

(où σ^2 est la variable et la valeur s^2 est fixée)

Ce passage, qui explicite le rôle symétrique joué par s^2 et σ^{-2}, peut être justifié par l'argument *fiduciaire* du *pivot* de Fisher (*cf.* Fisher, 1990). Nous appellerons cette solution la solution bayésienne *standard* ou encore la solution *fiducio-bayésienne* (sur cette appellation, voir l'introduction).

■ *Distribution initiale conjuguée naturelle*

Si nous choisissons pour σ^2 la distribution initiale ("informative") de la même famille que la distribution fiducio-bayésienne (distribution *conjuguée naturelle*)

$$\sigma^2 \sim s_0^2\varphi_{q_0}^{-2}$$

soit encore pour σ^{-2} la distribution

$$\sigma^{-2} \sim (1/s_0^2)\varphi_{q_0}^2$$

la densité finale devient proportionnelle à

$$(\sigma^2)^{-(q_0+q)/2-1} exp(-\tfrac{1}{2}\frac{q_0s_0^2+qs^2}{\sigma^2})$$

d'où le résultat

$$\sigma^2 | s^2 \sim \left(\frac{q_0 s_0^2 + q s^2}{q_0 + q}\right) \varphi^{-2}_{q_0+q}$$

qui, quand $q_0 \to 0$, donne encore, comme cas limite, la solution fiducio-bayésienne.

10.1.4 Inférence conjointe relative à δ et σ^2

Cherchons maintenant la distribution bayésienne finale relative au couple (δ,σ^2), conditionnellement à d et s^2, qui sont des statistiques conjointement exhaustives pour (δ,σ^2). Considérons une distribution initiale pour (δ,σ^2) de densité $p(\delta,\sigma^2)$.

■ *Cas de distributions initiales indépendantes pour δ et σ^2*

• Si les distributions initiales marginales relatives à δ et σ^2 sont indépendantes, soit $p(\delta,\sigma^2)=p(\delta)p(\sigma^2)$ et $p(\delta|\sigma^2)=p(\delta)p(\sigma^2)$, alors la distribution finale relative à δ conditionnellement à d et σ^2 s'obtient comme dans le cas où σ^2 est une donnée.

En effet, remarquons d'abord que, conditionnellement à d et σ^2 (mais pas à s^2), la densité finale $p(\delta|d,\sigma^2)=p(\delta,d|\sigma^2)/p(d|\sigma^2)$ est proportionnelle à la densité conjointe (conditionnelle à σ^2)

$$p(d,\delta|\sigma^2) = p(d|\delta,\sigma^2)p(\delta|\sigma^2) = p(d|\delta,\sigma^2)p(\delta)$$

Et cette densité finale n'est autre que celle de la distribution finale pour δ conditionnellement à d, s^2 et σ^2. En effet, en raison de l'indépendance des distributions initiales, nous avons

$$s^2|\delta,\sigma^2 \sim s^2|\delta,\sigma^2,d$$

et encore

$$s^2|\delta,\sigma^2,d \sim s^2|\sigma^2,d \quad (\text{car } s^2|\delta,\sigma^2,d \sim \sigma^2\varphi^2_q)$$

donc s^2 et δ sont indépendantes conditionnellement à σ^2 et d, et finalement

$$\delta|\sigma^2,d \sim \delta|\sigma^2,d,s^2$$

Nous pouvons donc dire que dans ce cas, conditionnellement à σ^2, s^2 "n'apporte pas d'information" sur δ. En revanche la distribution finale relative à σ^2 dépendra en général de d.

• Si les distributions initiales marginales relatives à δ et σ^2 sont indépendantes, et si en outre la distribution finale relative à δ conditionnellement à σ^2 est la distribution fiducio-bayésienne $N(d,b^2\sigma^2)$, alors la distribution

finale relative à σ^2 conditionnellement à d (mais pas à s^2) coïncide avec la distribution initiale pour ce paramètre. Dans cette situation d "n'apporte pas d'information" sur σ^2.

En effet, dans ce cas la densité finale conjointe, conditionnelle à d, $p(\delta,\sigma^2|d)$, est selon la formule de Bayes proportionnelle à

$$p(d|\delta,\sigma^2)p(\delta)p(\sigma^2)$$

où $p(d|\delta,\sigma^2)$ est la densité de la distribution $N(\delta,b^2\sigma^2)$.

Nous avons par ailleurs

$$p(\delta,\sigma^2|d) = p(\delta|\sigma^2,d)p(\sigma^2|d)$$

où, d'après le résultat précédent, $p(\delta|\sigma^2,d)$ est égale à $p(\delta|\sigma^2,d,s^2)$, donc est ici la densité de la distribution $N(d,b^2\sigma^2)$.

Par suite les termes en σ^2 des quantités $p(d|\delta,\sigma^2)$ et $p(\delta|\sigma^2,d)$ sont égaux; il en résulte que $p(\sigma^2|d)=p(\sigma^2)$.

Utilisant encore l'indépendance des distributions d'échantillonnage de d et s^2, nous en déduisons que la distribution finale $\sigma^2|d,s^2$ ne dépend pas de d. Elle se déduit donc directement de la distribution d'échantillonnage de s^2 et de la distribution initiale de σ^2, en appliquant les résultats concernant l'inférence sur σ^2 avec δ donnée.

■ *Caractérisation de la distribution fiducio-bayésienne*

Pour la solution *standard*, ou *fiducio-bayésienne*, nous supposons l'indépendance des distributions initiales relatives à δ et σ^2. En appliquant les résultats de la section précédente, nous considérons pour δ, conditionnellement à σ^2, la distribution fiducio-bayésienne (correspondant au cas σ^2 donnée)

$$\boxed{\delta|\sigma^2,d,s^2 \sim \delta|\sigma^2,d \sim N(d,b^2\sigma^2)}$$

et pour σ^2 la distribution fiducio-bayésienne (correspondant au cas δ donnée)

$$\boxed{\sigma^2|d,s^2 \sim \sigma^2|s^2 \sim s^2\varphi_q^{-2}}$$

■ *Distribution initiale conjuguée naturelle*

Pour une distribution initiale de densité quelconque $p(\delta,\sigma^2)$, la densité de la distribution finale est, à une constante de proportionnalité près, égale au

produit de cette densité initiale par la densité d'échantillonnage conjointe des statistiques d et s^2

$$p(\delta,\sigma^2|d,s^2) \propto p(d,s^2|\delta,\sigma^2)p(\delta,\sigma^2)$$
$$\propto (\sigma^2)^{-(q+1)/2} exp\left(-\tfrac{1}{2}\left(\frac{(\delta-d)^2}{b^2}+qs^2\right)\frac{1}{\sigma^2}\right)p(\delta,\sigma^2)$$

Considérons maintenant le cas particulier où la distribution initiale est de la même famille que la distribution fiducio-bayésienne (distribution *conjuguée naturelle*), c'est-à-dire est caractérisée par

$$\delta|\sigma^2 \sim N(d_0, b_0^2\sigma^2) \quad \text{et} \quad \sigma^2 \sim s_0^2\varphi_{q_0}^{-2}$$

Dans ce cas, la distribution finale est encore de la même famille. Elle est caractérisée par

$$\delta|\sigma^2,d,s^2 \sim N(d_1, b^2\sigma^2)$$
$$\sigma^2|d,s^2 \sim s_1^2\varphi_{q_1}^{-2}$$

$$d_1 = \frac{b_0^2 d + b^2 d_0}{b_0^2+b^2}$$

$$b_1^2 = \frac{b_0^2 b^2}{b_0^2+b^2}$$

$$s_1^2 = \frac{q_0 s_0^2 + qs^2 + \frac{(d_0-d)^2}{b_0^2+b^2}}{q_0+q+1}$$

$$q_1 = q_0+q+1$$

■ *Preuve*

$$p(\delta,\sigma^2) = p(\delta|\sigma^2)p(\sigma^2) \propto (\sigma^2)^{-1/2} exp\left(-\tfrac{1}{2}\frac{(\delta-d_0)^2}{b^2\sigma^2}\right)(\sigma^2)^{-(q_0+1)/2} exp\left(-\tfrac{1}{2}\frac{q_0 s_0^2}{\sigma^2}\right)$$

A partir de l'égalité

$$\frac{(\delta-d_0)^2}{b_0^2}+\frac{(\delta-d)^2}{b^2}+q_0s_0^2+qs^2=\frac{b_0^2+b^2}{b_0^2b^2}\left(\delta-\frac{b_0^2d+b^2d_0}{b_0^2+b^2}\right)^2+\frac{(d_0-d)^2}{b_0^2+b^2}+q_0s_0^2+qs^2$$

$$=\frac{1}{b_1^2}(\delta-d_1)^2+q_1s_1^2$$

la densité finale peut alors s'écrire sous la forme

$$p(\delta,\sigma^2|d,s^2)\propto(\sigma^2)^{-1/2}\,exp\left(-\tfrac{1}{2}\frac{(\delta-d_1)^2}{b_1^2\sigma^2}\right)(\sigma^2)^{-(q_1+1)/2}\,exp\left(-\tfrac{1}{2}\frac{q_1s_1^2}{\sigma^2}\right)$$

où nous identifions:

• dans la première ligne le terme en δ et σ^2 de la densité de la distribution $N(d_1,b_1^2\sigma^2)$;

• dans la seconde ligne le terme en σ^2 de la densité de la distribution $s_1^2\varphi_{q_1}^{-2}$.

■ *Cas particuliers:* $b_0^2\to+\infty$ *et* $q_0\to 0$

• Quand $b_0^2\to+\infty$ nous obtenons en passant à la limite

$$d_1=d\quad\text{et}\quad b_1^2=b^2$$

qui sont inchangés par rapport à la distribution fiducio-bayésienne

$$s_1^2=\frac{q_0s_0^2+qs^2}{q_0+q+1}\ \text{et}\ q_1=q_0+q+1$$

• Quand $q_0\to+\infty$ nous obtenons en passant à la limite

$$d_1\quad\text{et}\quad b_1^2$$

qui sont inchangés par rapport au cas général

$$s_1^2=\frac{qs^2+(d_0-d)^2/(b_0^2+b^2)}{q+1}\quad\text{et}\quad q_1=q+1$$

• Quand simultanément $b_0^2\to+\infty$ et $q_0\to 0$ nous obtenons en passant à la limite la distribution finale caractérisée par

$$\delta|\sigma^2,d,s^2\sim N(d,b^2\sigma^2)\quad\text{et}\quad\sigma^2|d,s^2\sim\frac{q}{q+1}s^2\varphi_{q+1}^{-2}$$

d'où la distribution marginale pour δ

$$\delta|\sigma^2,d,s^2\sim t_{q+1}(d,b^2\frac{q}{q+1}s^2)$$

La distribution relative à σ^2 diffère de la distribution fiducio-bayésienne par l'introduction d'un degré de liberté supplémentaire, qui provient du fait que nous avons introduit ici une non-indépendance des distributions initiales marginales relatives à δ et σ^2.

• Techniquement, la distribution fiducio-bayésienne relative au couple (δ,σ^2) peut être retrouvée pour $b_0^2 \to +\infty$ et $q_0 \to -1$.

On remarquera que dans le cas de l'inférence sur une moyenne, si nous regardons la distribution initiale comme la distribution fiducio-bayésienne associée à un échantillon de taille n_0, nous avons $q_0=n_0-1$, et par suite $q_0 \to -1$ correspond à $n_0 \to 0$.

10.2 STRUCTURE "NORMALE - PHI-DEUX": PROCÉDURES POUR LES PARAMÈTRES

10.2.1 Test de signification fréquentiste

■ *Test fréquentiste usuel de l'hypothèse nulle* $\delta=\delta_0$

Le test de signification (fréquentiste) usuel de l'hypothèse nulle $[H_0:\delta=\delta_0]$ repose sur la distribution d'échantillonnage de la statistique $t=(d-\delta_0)/bs$ qui est une distribution *t usuelle* à q degrés de liberté, qui ne dépend pas de σ^2.

$$[\text{sous } H_0:\delta=\delta_0]\ \ t=\frac{d-\delta_0}{bs}\Big|\sigma^2 \sim t_q$$

En effet nous avons les deux distributions indépendantes

$$\frac{d-\delta_0}{b\sigma}\Big|[\delta=\delta_0],\sigma^2 \sim N(0,1) \quad \text{et} \quad \frac{s}{\sigma}\Big|\delta=\delta_0,\sigma^2 \sim \varphi_q$$

et le résultat découle de la deuxième caractérisation de la distribution *t généralisée* en §9.2.1.

Sous cette même hypothèse nulle, $F=t^2$ a une distribution *F* à 1 et q degrés de liberté

$$[\text{sous } H_0:\delta=\delta_0]\ \ F=t^2\Big|\sigma^2 \sim F_{m,q}$$

■ *Distribution d'échantillonnage de t sous une autre hypothèse*

Sous une autre hypothèse $[H_1:\delta=\delta_1]$, la distribution d'échantillonnage de la

statistique $t=(d-\delta_0)/bs$ est une distribution *t non-centrée* à q degrés de liberté, d'indice d'excentricité $(\delta_1-\delta_0)/b\sigma$, qui dépend donc de σ.

$$[\text{sous } H_1{:}\delta=\delta_1]\ t=\frac{d-\delta_0}{bs}\,|\,\sigma^2 \sim t'_q\left(\frac{\delta_1-\delta_0}{b\sigma}\right)$$

En effet nous avons les deux distributions indépendantes

$$\frac{d-\delta_0}{b\sigma}\,|\,[\delta=\delta_1],\sigma^2 \sim N\left(\frac{\delta_1-\delta_0}{b\sigma},1\right) \quad \text{et} \quad \frac{s}{\sigma}\,|\,\delta=\delta_1,\sigma^2 \sim \varphi_q$$

et le résultat découle de la deuxième caractérisation de la distribution *t non-centrée* en §9.3.1.

10.2.2 Test de signification semi-bayésien

■ *Test semi-fiducio-bayésien de l'hypothèse nulle* $\delta=\delta_0$

Adoptons maintenant pour le test de signification le point de vue semi-bayésien exposé aux chapitre §6. Le test de signification semi-bayésien standard (ou *semi-fiducio-bayésien*) de l'hypothèse nulle $[H_0{:}\delta=\delta_0]$ repose sur la distribution prédictive relative à d sous H_0, conditionnellement à la valeur observée s^2. Cette distribution se déduit de la distribution d'échantillonnage relative à d

$$[\text{sous } H_0{:}\delta=\delta_0]\ d\,|\,\sigma^2 \sim N(\delta_0,b^2\sigma^2)$$

et de la distribution fiducio-bayésienne relative à σ^2

$$\sigma^2\,|\,s^2 \sim s^2\varphi_q^{-2}$$

Compte tenu de l'indépendance des distributions d'échantillonnage de d et s^2, nous avons en effet

$$[\text{sous } H_0{:}\delta=\delta_0]\ d\,|\,s^2,\sigma^2 \sim d\,|\,\sigma^2 \sim N(\delta_0,b^2\sigma^2)$$

et par suite la distribution prédictive relative à d est une distribution *t généralisée* à q degrés de liberté, selon la première caractérisation de la distribution *t généralisée* en §9.2.1.

$$[\text{sous } H_0{:}\delta=\delta_0]\ d\,|\,s^2 \sim t_q(\delta_0,b^2s^2)$$

$$\text{soit } t=\frac{d-\delta_0}{bs}\,|\,s^2 \sim t_q$$

Nous pouvons donc tester directement H_0 à partir de la distribution prédictive de d, et le résultat coïncide avec le test fréquentiste.

■ *Distribution prédictive de t sous une autre hypothèse*

Sous une autre hypothèse [H_1:$\delta=\delta_1$], on montre de la même manière que la distribution prédictive de la statistique $t=(d-\delta_0)/bs$ est une distribution *t généralisée* à q degrés de liberté, de centre $(\delta_1-\delta_0)/bs$

$$[\text{sous } H_1:\delta=\delta_1]\ \ t=\frac{d-\delta_0}{bs}\,|\,s^2 \sim t_q\left(\frac{\delta_1-\delta_0}{bs},1\right) \quad (t \textit{ généralisée})$$

qu'on opposera à la distribution d'échantillonnage

$$[\text{sous } H_1:\delta=\delta_1]\ \ t=\frac{d-\delta_0}{bs}\,|\,\sigma^2 \sim t'_q\left(\frac{\delta_1-\delta_0}{b\sigma},1\right) \quad (t \textit{ non-centrée})$$

10.2.3 Procédures fiducio-bayésiennes

Les distributions fiducio-bayésiennes marginales et celles relatives à des paramètres dérivés se déduisent immédiatement de la distribution conjointe de (δ,σ^2).

■ *Distribution relative à* δ

La distribution fiducio-bayésienne (marginale) relative au paramètre δ est une distribution *t généralisée* à q degrés de liberté, de centre (moyenne) d et d'échelle e=bs (↳9.2)

$$\boxed{\delta\,|\,d,s^2 \sim t_q(d,e^2) \quad \text{avec } e=bs}$$

$$\text{d'où } (\delta-d)/e\,|\,d,s^2 \sim t_q$$

■ *Distribution relative au rapport* $(\delta-h)/\sigma$ (h constante réelle)

Il s'agit d'une distribution *lambda-prime*, à q degrés de liberté (↳9.4)

$$\boxed{\frac{\delta-h}{\sigma}\,|\,d,s^2 \sim \Lambda'_q\left(\frac{d-h}{s},b^2\right)}$$

et en particulier

$$\boxed{\frac{\delta}{\sigma}\,|\,d,s^2 \sim \Lambda'_q\left(\frac{d}{s},b^2\right)}$$

10.2.4 Liens entre l'inférence fiducio-bayésienne et le test de signification

■ *Lien technique*

En adoptant le point de vue semi-fiducio-bayésien, le lien entre l'inférence fiducio-bayésienne et le test de signification usuel de l'hypothèse nulle $\delta=0$ se déduit immédiatement de l'analogie entre:

• la distribution prédictive relative à d sous $[H_0:\delta=0]$,

$$d\,|\,[H_0:\delta=0],s^2 \sim t_q(0,e^2)$$

• et la distribution fiducio-bayésienne relative à la différence δ-d (conditionnelle à la valeur observée d)

$$\delta-d\,|\,d,s^2 \sim t_q(0,e^2)$$

■ *Réinterprétation fiducio-bayésienne du test de signification*

Notons $t_{obs}=d/e=d/bs$ la valeur prise par la statistique de test. Nous avons la réinterprétation fiducio-bayésienne du seuil de signification unilatéral (ou orienté) observé $p/2$ du test t, déduite de la distribution de δ qui s'écrit alors, pourvu que $d\neq 0$

$$\delta\,|\,d,s^2 \sim t_q(d,(d/t_{obs})^2)$$

$$\text{si } d>0,\ Pr(\delta>0\,|\,d,s^2 = Pr(t_q>t_{obs}) = p/2$$

$$\text{si } d<0,\ Pr(\delta<0\,|\,d,s^2 = Pr(t_q>t_{obs}) = p/2$$

10.2.5 Procédures bayésiennes (distribution initiale conjuguée naturelle)

Dans le cas d'une distribution initiale de la même famille que la distribution fiducio-bayésienne, les distributions finales marginales et celles relatives à des paramètres dérivés s'obtiennent comme les distributions fiducio-bayésiennes, en remplaçant simplement d, s, b et q par les valeurs correspondantes d_1, s_1, b_1 et q_1 définies précédemment en ↳10.1.4.

$$\boxed{\delta\,|\,d,s^2 \sim t_{q_1}(d_1,e_1^2) \quad \text{avec } e_1=b_1s_1}$$

$$\boxed{\frac{\delta-h}{\sigma}\,|\,d,s^2 \sim \Lambda'_{q_1}\left(\frac{d_1-h}{s_1}, b_1^2\right)}$$

10.3 Structure "normale - phi-deux": Procédures prédictives

10.3.1 Procédures bayésiennes prédictives

Considérons une distribution initiale conjuguée naturelle caractérisée par

$$\delta\,|\,\sigma^2 \sim N(d_0, b_0^2\sigma^2) \quad \text{et} \quad \sigma^2 \sim s_0^2\varphi^{-2}_{q_0}$$

et explicitons la distribution *prédictive* correspondante pour la structure "*normale - phi-deux*"

$$\left\|\begin{array}{l} d\,|\,\delta,\sigma^2 \sim N(\delta, b^2\sigma^2) \\ s^2\,|\,\delta,\sigma^2 \sim \sigma^2\varphi^2_q \quad (\text{soit } \sigma^2\chi^2_q/q) \\ d \perp\!\!\!\perp s^2\,|\,\delta,\sigma^2 \text{ (indépendance)} \end{array}\right.$$

10.3.2 Distributions prédictives marginales

■ *Distribution relative à* d

Il s'agit d'une distribution t de Student *généralisée*

$$\boxed{d \sim t_{q_0}(d_0, (b_0^2+b^2)s_0^2)}$$

En effet, des distributions

$$d\,|\,\delta,\sigma^2 \sim N(\delta, b_0^2\sigma^2) \quad \text{et} \quad \delta\,|\,\sigma^2 \sim N(d_0, b_0^2\sigma^2)$$

nous déduisons la distribution conditionnelle à σ^2 (ceci est une conséquence du lemme donné en ↳9.6.4)

$$d\,|\,\sigma^2 \sim N(d_0, (b_0^2+b^2))\sigma^2)$$

d'où le résultat énoncé puisque

$$\sigma^2 \sim s_0^2 \varphi_{q_0}^{-2}$$

■ *Distribution relative à* s^2

Il s'agit d'une distribution F

$$\boxed{s^2 \sim s^2 F_{q,q_0}}$$

Ceci se déduit des distributions

$$s^2|\delta,\sigma^2 \sim s^2|\sigma^2 \sim \sigma^2 \varphi_q^2 \quad \text{et} \quad \sigma^2 \sim s_0^2 \varphi_{q_0}^{-2}$$

10.3.3 Distribution prédictive relative à d conditionnellement à s^2

Il s'agit d'une distribution t *généralisée*

$$\boxed{d|s^2 \sim t_{q_0+q}\left(d, (b_0^2+b^2)\frac{q_0 s_0^2+q s^2}{q_0+q}\right)}$$

■ *Preuve*

La distribution finale de σ^2 conditionnellement à s^2 est (↳10.2.3)

$$\sigma^2|s^2 \sim \frac{q_0 s_0^2+q s^2}{q_0+q} \varphi_{q_0+q}^{-2}$$

Le résultat énoncé se déduit de cette dernière distribution et de la distribution de d conditionnelle à σ^2

$$d|\sigma^2 \sim d|s^2,\sigma^2 \sim N(d_0,(b_0^2+b^2)\sigma^2)$$

10.3.4 Distributions prédictives relatives à l'effet calibré

■ *Distribution relative au rapport* d/σ

Il s'agit d'une distribution *lambda-prime* (↳9.4)

$$\boxed{\frac{d}{\sigma} \sim \Lambda'_{q_0}\left(\frac{d_0}{s_0}, b_0^2+b^2\right)}$$

Ceci se déduit des distributions

$$\frac{d}{\delta}|\sigma^2 \sim N\left(\frac{d_0}{\sigma}, b_0^2+b^2\right) \quad \text{et} \quad \sigma^2 \sim s_0^2\varphi_{q_0}^{-2}$$

■ *Distribution relative au rapport* d/s

Il s'agit d'une distribution *K-prime*, à q_1 et q_2 degrés de liberté (↳9.5)

$$\boxed{\frac{d}{s} \sim K'_{q_0,q}\left(\frac{d}{s}, b_0^2+b^2\right)}$$

Ceci se déduit des distributions

$$\frac{d}{s}|s^2 \sim t_{q_0+q}\left(\frac{d_0}{s_0}, (b_0^2+b^2)\frac{q_0(s_0^2/s^2)+q}{q_0+q}\right) \quad \text{et} \quad \frac{s_0^2}{s^2} \sim F_{q_0,q}$$

10.3.5 Distribution prédictive relative à la statistique de test

Nous en déduisons que la distribution prédictive relative à la statistique de test "*t* de Student" *t*=d/bs est une distribution *K-prime*, qui ne dépend que de $t_0=d_0/b_0s_0$

$$\boxed{t=\frac{d}{bs} \sim K'_{q_0,q}\left(\frac{b}{b_0}t_0, \frac{b_0^2+b^2}{b_0^2}\right) \quad \text{avec } t_0 = \frac{d_0}{b_0 s_0}}$$

10.3.6 Distributions prédictives relatives aux limites de crédibilité

Considérons l'intervalle de crédibilité fiducio-bayésien centré sur d pour le paramètre δ obtenu pour la garantie γ. Rappelons que les limites de cet intervalle coïncident avec les limites de confiance γ fréquentistes

$$\textit{limite inférieure} \quad l_\gamma = d+bst_{1-\gamma}$$

$$\textit{limite supérieure} \quad l^\gamma = d-bst_{1-\gamma}$$

où $t_{1-\gamma}$ (positif) est par définition tel que

$$Pr(|t_q|>t_{1-\gamma}) = 1-\gamma \quad \text{soit} \quad Pr(t_q>t_{1-\gamma}) = \tfrac{1}{2}(1-\gamma)$$

Les distributions prédictives relatives à ces limites se déduisent des distributions conditionnelles à s^2. Par exemple, pour la limite inférieure

$$l_\gamma=d-bt_{1-\gamma}\,|\,s^2 \sim d_0 + t_{q_0+q}\left(-bst_{1-\gamma},(b_0^2+b^2)\frac{q_0s_0^2+q^2s^2}{q_0+q}\right)$$

d'où

$$l_\gamma \sim d_0 + K'_{q_0,q}(-bs_0t_{1-\gamma},(b_0^2+b^2)s_0^2)$$

et de même

$$l^\gamma \sim d_0 + K'_{q_0,q}(bs_0t_{1-\gamma},(b_0^2+b^2)s_0^2)$$

■ *Limites de crédibilité tenant compte de la distribution initiale*

On peut également considérer les distributions prédictives relatives aux limites de crédibilité obtenues en tenant compte de la distribution initiale, c'est-à-dire, $d_1+b_1s_1t_{1-\gamma}$ et $d_1-b_1s_1t_{1-\gamma}$ (avec les définitions données en ↳10.1.4 et pour q_1 degrés de liberté). Dans ce cas les distributions sont malheureusement beaucoup plus complexes, car s_1 dépend de d_0 et d. Cependant des approximations très satisfaisantes peuvent être obtenues (*cf.* Grouin, 1994).

10.3.7 Remarque: quand $b^2 \to 0$ et $q \to +\infty$

Quand $b^2 \to 0$ et $q \to +\infty$ – ce qui en pratique correspondrait à la réalisation d'une expérience avec un très grand nombre d'observations – les distributions prédictives coïncident à la limite avec la distribution initiale relative aux paramètres, puisque

$$d|s^2,\sigma^2 \simeq N(d_{01},b_0^2\sigma^2) \quad \text{et} \quad s^2 \simeq s_0^2\varphi_{q_0}^{-2}$$

ce qui est analogue à

$$\delta|\sigma^2 \sim N(d_0,b_0^2\sigma^2) \quad \text{et} \quad \sigma^2 \sim s_0^2\varphi_{q_0}^{-2}$$

Ce résultat est conceptuellement important, puisqu'il permet d'assimiler les paramètres avec les valeurs des statistiques correspondantes obtenues pour un très grand nombre d'observations.

10.4 Structure "normale – (phi-deux)s"

10.4.1 Définition

La structure statistique *"normale - (phi-deux)s"*, qui intervient dans les solutions "sans égalité des variances", est définie par la donnée d'un paramètre numérique $\delta\in\mathbb{R}$ et des paramètres numériques positif $\sigma_g^2\in\mathbb{R}_+$ $(g\in G)$.

■ *Distributions d'échantillonnage*

Sous un certain modèle d'échantillonnage, les statistiques d et s_g^2 sont conjointement exhaustives pour les paramètres de la structure et ont les distributions d'échantillonnage indépendantes

$$\left\| \begin{array}{l} d|\delta,(\sigma_g^2) \sim N(\delta,\varepsilon^2) \quad \text{avec } \varepsilon^2 = \sum_{g\in G} c_g^2\sigma_g^2 \quad (c_g^2\in\mathbb{R}_+) \\ s_g^2|\delta,(\sigma_g^2) \sim \sigma_g^2\varphi_{q_g}^2 \end{array} \right.$$

10.4.2 Distribution fiducio-bayésienne pour le paramètre ε^2

■ *Solution fiducio-bayésienne*

Nous appliquons simplement pour chaque paramètre σ_g^2 la solution fiducio-bayésienne dérivée en ↳10.1.4. D'où, pour chaque que σ_g^2

$$\sigma_g^2|s_g^2 \sim s_g^2\varphi_{q_g}^{-2}$$

Puisque nous recherchons une solution traduisant un état initial d'indifférence, il apparaît approprié de supposer l'indépendance des distributions initiales relatives à l'ensemble des variances. En raison de l'indépendance des distributions d'échantillonnage, les distributions finales sont aussi indépendantes avec, en outre

$$\sigma_g^2 | (s_g^2)_{g\in G} \sim \sigma_g^2 | s_g^2 \sim s_g^2 \varphi_{q_g}^{-2}$$

La distribution finale relative au paramètre $\varepsilon^2 = \Sigma_g c_g^2 \sigma_g^2$ est déduite, par changement de variable et intégration, de la densité conjointe

$$p(\sigma_1^2, \sigma_2^2 \ldots | (s_g^2)) = p(\sigma_1^2 | s_1^2) p(\sigma_2^2 | s_2^2) \ldots$$

Les calculs doivent être effectués par intégration numérique. Mais en pratique de très bonnes approximations peuvent être obtenues.

■ *Moments*

Les moments de la distribution fiducio-bayésienne relative au paramètre ε^2 s'obtiennent de manière immédiate, en raison de l'indépendance des distributions marginales des σ_g^2.

Nous avons en particulier la moyenne et la variance

$$moy(\varepsilon^2 | (s_g^2)) = \Sigma_g c_g^2 s_g^2 \, moy(\varphi_{q_g}^{-2}) = \sum_{g\in G} c_g^2 \frac{q_g}{q_g - 2} s_g^2 \quad [q_g > 2]$$

$$var(\varepsilon^2 | (s_g^2)) = \Sigma_g c_g^4 s_g^4 \, var(\varphi_{q_g}^{-2}) = \sum_{g\in G} c_g^4 \frac{2q_g^2}{(q_g - 2)^2 (q_g - 4)} s_g^4 \quad [q_g > 4]$$

■ *Approximation de la distribution de ε^2*

Nous pouvons approcher la distribution fiducio-bayésienne de ε^2 par une distribution *phi-deux inverse* ayant même moyenne et même variance en utilisant le résultat général donné en ↳9.6.2

$$\varepsilon^2 \simeq \tilde{e}^2 \varphi_{\tilde{q}}^{-2} \quad \text{avec } \tilde{q} = 4 + 2\frac{(moy(\varepsilon^2))^2}{var(\varepsilon^2)} \quad \text{et} \quad \tilde{e}^2 = \frac{\tilde{q} - 2}{\tilde{q}} moy(\varepsilon^2)$$

10.4.3 Test semi-fiducio-bayésien

Selon la procédure définie dans le cas (particulier) de la structure "*normale - phi-deux*" (↳10.2.2), nous pouvons tester l'hypothèse nulle $[H_0 : \delta = \delta_0]$

directement à partir de la distribution prédictive de d. La densité de cette dernière est dérivée à partir de la distribution fiducio-bayésienne de ε^2.

■ *Approximation*

Puisque la distribution fiducio-bayésienne relative à ε^2 peut être approximée par une distribution *phi-deux inverse*, la distribution prédictive de la statistique d (sous l'hypothèse nulle H_0) est approximée par une distribution *t généralisée*

$$[\text{sous } H_0{:}\delta=\delta_0]\ \ d\,|\,(s_g^2) \simeq t_{\tilde{q}}(\delta_0,\tilde{e}^2)$$

Quel que soit le nombre de variances en jeu, nous pouvons donc utiliser des procédures analogues aux procédures familières de la structure "*normale - phi-deux*" (*t* de Student et rapport *F*), en utilisant les approximations

$$[\text{sous } H_0{:}\delta=\delta_0] \quad \tilde{t}=\frac{d-\delta_0}{\tilde{e}}\,|\,(s_g^2) \sim t_{\tilde{q}} \quad \text{et} \quad \tilde{F}=\tilde{t}^2\,|\,(s_g^2) \sim F_{1,\tilde{q}}$$

10.4.4 Procédures fiducio-bayésiennes

La distribution fiducio-bayésienne relative à δ conditionnellement à ε^2 est obtenue de la même manière que dans le cas de la structure "*normale - phi-deux*"

$$\delta\,|\,d,(s_g^2),(\sigma_g^2) \sim \delta\,|\,d,(\sigma_g^2) \sim \delta\,|\,d,\varepsilon^2 \sim N(d,\varepsilon^2)$$

La distribution marginale relative à δ se déduit de cette distribution et de la distribution marginale relative à ε^2.

$$\varepsilon^2\,|\,d,(s_g^2) \sim \varepsilon^2\,|\,(s_g^2)$$

■ *Approximation de la distribution fiducio-bayésienne de* δ

La distribution marginale relative à δ sera donc également approximée par une distribution *t généralisée*

$$\delta\,|\,d,(s_g^2) \simeq t_{\tilde{q}}(d,\tilde{e}^2)$$

■ *Remarque*

Dans le cas particulier de deux variances, avec $c_{g1}^2=1/f_{g1}$ et $c_{g2}^2=1/f_{g2}$, nous retrouvons l'approximation proposée par Patil (1965) pour la distribution de Behrens-Fisher.

10.5 Compléments: Le paradoxe de marginalisation

■ *Un paradoxe...*

La distribution fiducio-bayésienne relative au paramètre dérivé $\delta_{cal}=\delta/\sigma$ ne dépend des statistiques de d et de s que par l'intermédiaire de leur rapport $d_{cal}=d/s$, dont la distribution d'échantillonnage

$$d_{cal}\,|\,\delta,\sigma \sim t'_q(\delta_{cal},b^2)$$

ne dépend elle-même que de δ_{cal}.

On pourrait donc penser pouvoir obtenir cette distribution fiducio-bayésienne directement à partir d'une distribution initiale relative à δ_{cal} (voir cependant les réserves que nous faisons envers cette manière de procéder dans l'introduction). Or une telle distribution initiale n'existe pas: c'est le "*paradoxe de marginalisation*", dénoncé par Stone et Dawid (1972).

■ *...qui n'est cependant pas si dramatique en pratique*

Il faut cependant relativiser ici la portée de ce paradoxe, puisque pour une densité initiale $p(\delta_{cal})$ constante, nous avons

$$\delta_{cal}\,|\,d_{cal} \sim \Lambda'_{q+1}(\tfrac{q+1}{q}d_{cal},b^2)$$

qui, sauf pour les petites valeurs de q (mais celles-ci conduiront généralement à un constat d'ignorance...), différera souvent peu en pratique de la distribution fiducio-bayésienne

$$\delta_{cal}\,|\,d,s \sim \Lambda'_q(d_{cal},b^2)$$

En effet, de la densité de la distribution d'échantillonnage (↳9.3.3)

$$p(d_{cal}\,|\,\delta_{cal}) \propto exp(-\tfrac{1}{2}\tfrac{\delta_{cal}^2}{b^2})\left(\frac{q}{qb^2+d_{cal}^2}\right)^{(q+1)/2} \times \sum_{j=0}^{+\infty}\frac{1}{j!}\,\Gamma(\tfrac{1}{2}(q+j+1))\left(\frac{2}{qb^2+d_{cal}^2}\right)^{j/2}\left(\delta_{cal}\frac{d_{cal}}{b}\right)^j$$

nous déduisons pour $p(\delta_{cal}) \propto 1$

$$p(\delta_{cal} \mid d_{cal}) \propto exp(-\tfrac{1}{2}\frac{\delta_{cal}^2}{b^2}) \times \sum_{j=0}^{+\infty} \frac{1}{j!} \Gamma(\tfrac{1}{2}(q+1+j)) \left(\frac{2}{(q+1)b^2 + \frac{q+1}{q}d_{cal}^2}\right)^{j/2} \left(\frac{\frac{q+1}{q}d_{cal}}{b}\delta_{cal}\right)^{j}$$

soit le résultat (*cf.* ↳9.4.3).

■ *Une distribution non-informative pour* δ_{cal}*?*

Bernardo et Smith (1994, page 330), suivant des arguments qui dépassent le cadre de cet ouvrage, proposent d'utiliser pour δ_{cal} (si c'est le "paramètre d'intérêt") la distribution initiale de densité

$$p(\delta_{cal}) \propto (2+\delta_{cal}^2)^{-1/2}$$

soit une distribution uniforme pour $Argsh(\delta_{cal}/\sqrt{2})$, qui "favorise" les valeurs proches de 0.

Cette solution présente cependant un certain nombre d'inconvénients, qui nous conduisent à ne pas la retenir.

• La probabilité finale que δ_{cal} soit positif diffère de la probabilité que δ soit positif.

• Dans le cas de recherche d'une conclusion d'effet notable pour δ_{cal}, on perd la correspondance avec les procédures fréquentistes (↳7.1.4).

• Surtout on peut être très réticent à l'idée de combiner deux valeurs de l'effet calibré. Si nous regroupons en une seule population deux sous-populations d'effets calibrés connus respectifs δ_1/σ_1 et δ_2/σ_2, la moyenne de ces deux valeurs peut être sensiblement différente de l'effet calibré δ/σ obtenu à partir des valeurs δ et σ calculées sur l'ensemble de la population. La même remarque s'applique à la valeur absolue de la moyenne et plus généralement à une moyenne quadratique de moyennes, à un coefficient de corrélation, etc. En conséquence, il ne nous apparaît donc pas recommandable de considérer directement une distribution initiale relative à un tel paramètre; que celle-ci soit informative ou non-informative ne change rien au problème.

ANNEXE

Installation et notice d'utilisation des programmes Windows

LeBayésien
LesDistributions
LesEchantillons
LesEffectifs

✍ Le fichier "LISEZ.MOI" de la disquette contient s'il y a lieu les compléments et modifications apportés au présent document

✍ Les programmes **LesEchantillons** et **LesEffectifs** sont des versions limitées, qui ne comportent pas de système d'aide. Pour leur utilisation, consultez respectivement les chapitres ☞0 (section 0.2.5) et ☞8.

IBM est une marque déposée de International Business Machines Corporation.
Windows est une marque de Microsoft Corporation.
PAC est une marque déposée.

INSTALLATION DES PROGRAMMES

CONFIGURATION NÉCESSAIRE

❶ Un ordinateur compatible IBM doté d'un microprocesseur 386 ou supérieur, avec Windows version 3.1 ou ultérieure (y compris Windows 95); ou un système compatible.

❷ Au moins quatre méga-octets de mémoire (huit méga-octets étant fortement conseillés).

❸ Un disque dur avec quatre méga-octets d'espace libre pour l'installation des programmes.

INSTALLATION

Faites une copie de sauvegarde

Avant toute chose, faites une copie de sauvegarde des disquettes d'origine, et continuez avec les copies.

✧ Procédure d'installation

L'installation se fait sous Windows, de préférence au début d'une session. Certains programmes résidents peuvent empêcher l'installation; cela vous sera signalé, et vous devrez alors les décharger, et relancer la procédure d'installation.

Le logiciel est installé dans le répertoire que vous spécifiez lors de l'installation ("**C:\BAYES**" par défaut). Ce répertoire contient un sous-répertoire appelé "**GUIDE**" dans lequel vous trouverez les fichiers d'exemples.

Bibliothèques exécutables

Des fichiers de "bibliothèques exécutables" (extensions ".DLL" et ".OCX" notamment) sont installés dans le sous-répertoire SYSTEM du répertoire d'installation de Windows. Ces fichiers peuvent être utilisés simultanément par différents programmes.
Vous en trouverez la liste dans le fichier **SETUP.LST** de la disquette 1.
Si vous désinstallez le logiciel, vous devez vous assurer que ces fichiers ne sont pas utilisés par d'autres programmes avant de les supprimer.

✧ Exécutez la procédure d'installation (sous Windows)

❶ Insérez la disquette 1 dans le lecteur de disquettes.

❷ Exécutez le programme **SETUP.EXE** de la disquette: par exemple, utilisez le sous-menu **exécuter** de la fonction **fichier** du gestionnaire de programme Windows (ou de la fonction **démarrer** de Windows 95).

❸ Suivez ensuite les instructions affichées à l'écran.

EN CAS DE PROBLÈME

Contactez le C.I.S.I.A.

☎ (33-1) 43 74 20 20 Fax (33-1) 43 74 17 29

LeBayésien

LeBayésien est un programme didactique et convivial, qui vous permet d'effectuer toutes les analyses descriptives et inférentielles présentées dans cet ouvrage. Il s'agit d'une version limitée aux comparaisons à un degré de liberté, sur une seule variable, sans covariable dans un plan S<G>*O, c'est-à-dire les situations traitées ici.

> ***Remarque***
> Le "module graphique bayésien" du logiciel **PAC** (version Windows), qui lui est en partie commun, ne comporte pas ces restrictions.

Prise en mains

Les exemples d'utilisation traités dans les chapitres ☞1 à ☞7 permettent une prise en mains progressive du programme; ils sont en principe autonomes et ne nécessitent pas la lecture préalable de la présente notice. Celle-ci consiste en une présentation générale des principales fonctions, et en un certains nombre de compléments, notamment en ce qui concerne:

- la gestion des fichiers;
- l'utilisation de l'aide;
- l'édition des données;
- l'utilisation des exemples de sorties du logiciel **PAC**;
- l'utilisation de la calculatrice;
- l'impression des courbes de distributions.

Un certain nombres de fonctions, qui sont décrites dans les exemples d'utilisation, ne sont pas reprises ici; il s'agit notamment de:

- l'utilisation de la "fenêtre des statistiques" (voir notamment les chapitres ☞1, ☞2 et ☞4, et pour les solutions "sans égalité des variances" ☞6);
- l'utilisation de la "fenêtre des énoncés" (voir notamment les chapitres ☞1, ☞2, et pour les "procédures fréquentistes effet négligeable/notable" ☞6);
- le choix de la distribution initiale (chapitre ☞3);
- l'utilisation de la "fenêtre de la régression polynomiale" (☞5.1).

GÉNÉRALITÉS SUR L'UTILISATION

Utilisation de Windows

Le lecteur est supposé familiarisé avec l'utilisation de logiciels sous Windows.

Usage de la souris

- La mise en œuvre est décrite en supposant l'usage de la souris; mais la plupart des fonctions sont accessibles à partir du clavier, soit par menu, soit par des touches de raccourcis (qui sont indiquées dans les sous-menus), soit en maintenant enfoncé la touche **Alt** et en tapant la lettre soulignée à l'écran.
- "***cliquez***" sans autre précision signifie "***cliquez une seule fois avec le bouton principal de la souris***".

Pratique

D'une manière générale, dans les zones de texte, les valeurs numériques peuvent être entrées directement sous forme de rapport (par exemple "1/10" ou "2617.53/9") ou de produit (par exemple "27.751*0.73"). On ne peut utiliser qu'un seul symbole, soit "/" soit "*" à la fois.

Feu vert/feu rouge

Pour éviter les risques de modifications accidentelles, les données qui proviennent d'un fichier, ou qui sont le résultat d'un calcul, sont "protégées". Ceci est signalé par un "*feu rouge*" (en outre les zones de textes correspondantes sont grises). Si vous souhaitez modifier ces données, cliquez sur le feu rouge, qui "passe au vert"; vous pouvez également effectuer l'opération inverse.

Attention

Si vous vous retrouvez par mégarde dans une fenêtre qui ne correspond pas à ce qui est décrit, revenez en arrière en cliquant suivant les cas sur le bouton **fermer**, ou **annuler**...

Le menu principal et la barre de boutons

Le menu principal

Le menu principal comporte les fonctions suivantes:

mode d'entrée énoncés données options aide [F1] quitter ?

La barre de boutons

La barre de boutons en dessous du menu principal comporte, au lancement du programme, les fonctions (de gauche à droite):

moyenne G / O effet ...calibré directe PAC exemples

qui permettent un accès direct aux fonctions les plus utilisées.

Les fichiers

Tous les fichiers sont des fichiers ASCII (sauf les fichiers "Bitmap"). Pour simplifier leur gestion, les noms de chacun des types de fichiers reçoivent des extensions particulières.

Fichiers de données - Extension: .DBT et .DAT

LeBayésien comporte un éditeur de données. Les fichiers de données sont enregistrés sous un format propre au programme et ont obligatoirement l'extension ".DBT". Il est également possible de créer et d'utiliser des fichiers de données au format de **PAC** (un format standard utilisable par la plupart des logiciels).

Fichiers de commandes du logiciel PAC Windows - Extensions: .PAC et .XPW

Quand vous enregistrez un fichier de données, vous pouvez créer un fichier de commandes pour **PAC**, en mode "normal" (extension ".PAC") et/ou en mode "Express" (extension ".XPW")

Fichiers de résultats du logiciel PAC Windows - Extensions: .OUT et .BAY

Il s'agit des fichiers de résultats créés par **PAC**. Ils ont obligatoirement l'extension ".OUT".

Les fichiers d'extension ".BAY" créés par **PAC** contiennent les informations nécessaires pour les distributions bayésiennes. Ils peuvent être utilisés par **LeBayésien**; vous n'avez pas à vous préoccuper du contenu de ces fichiers.

PAC associe automatiquement un fichier de ce type à chaque fichier de sorties, qui a le même nom avec l'extension [.OUT].

Fichiers de sauvegarde des courbes bayésiennes - Extension: .DIS

Ces fichiers contiennent les informations nécessaires pour construire les courbes des distributions bayésiennes. Ils ont obligatoirement l'extension ".DIS". Vous n'avez pas à vous préoccuper du contenu de ces fichiers.

Fichiers de sauvegarde au format "Bitmap" - Extension: .BMP

Ces fichiers contiennent les courbes des distributions bayésiennes en mode graphique "Bitmap". Ils reçoivent obligatoirement l'extension ".BMP".

Fichiers de sauvegarde des densités (mode Texte) - Extension: .TXT

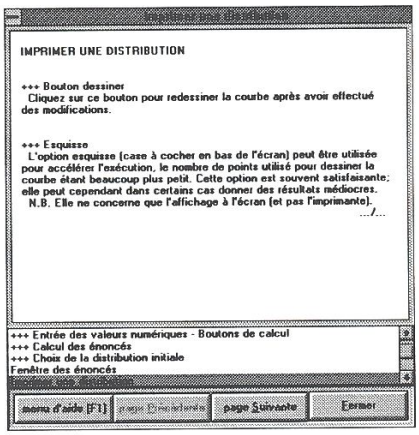
Ces fichiers contiennent les coordonnées (abscisse et densité) des distributions bayésiennes en mode texte. Ils reçoivent obligatoirement l'extension ".TXT".

Le fichier de configuration .CFG

Un fichier de configuration appelé "WBAYES.CFG" est créé au premier lancement du programme, et est mis à jour automatiquement à chaque utilisation. Il conserve notamment les informations sur les répertoires et les fichiers actifs. Des répertoires différents peuvent être utilisés pour les différents types de fichiers.

Utilisation de l'aide

LeBayésien comporte un système d'aide détaillé, accessible suivant les cas, soit par menu, soit par un bouton **aide**, et toujours par la touche **F1** du clavier.



Choisir une section

Dans la fenêtre d'aide, cliquez sur le bouton **menu d'aide** (ou appuyez sur la touche **F1**) pour accéder aux différentes sections qui s'affichent alors: Nouveautés, Les fonction générales, etc.

Se déplacer dans une section

Chaque section comporte un certain nombre de "pages". Vous pouvez vous déplacer d'une page à l'autre en cliquant sur les boutons **page suivante** et **page précédente** correspondants, ou en utilisant les touches **page bas** et **page haut** du clavier.

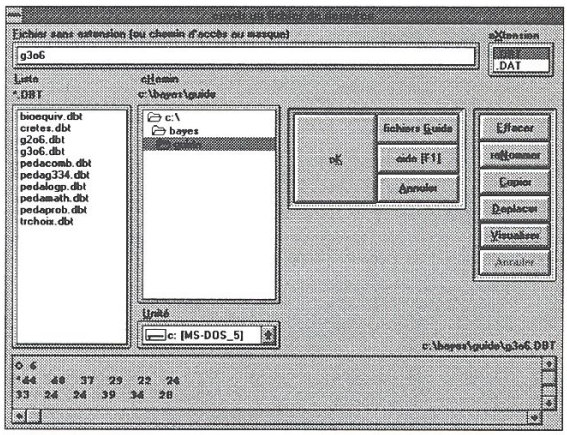
Chaque section est subdivisée en rubriques, qui peuvent être sélectionnées dans la liste qui est affichée en dessous du texte d'aide.

Gestion des fichiers: La fenêtre de dialogue "fichiers"

Choisir un nom de fichier

Quand vous devez choisir un nom de fichier, apparaît une fenêtre de dialogue spéciale (excepté pour les fichiers [*.*] qui font appel à une fenêtre standard de Windows). Pour faire apparaître cette fenêtre, activez par exemple l'option **ouvrir** dans le sous-menu de la fonction **données** du menu principal.

Des listes vous permettent de sélectionner: l'unité, le chemin et le fichier (s'il existe déjà). La sélection est rendue active par un double-clic ou en cliquant sur le bouton **ok**.



✧ ***Fichiers guide***

Le bouton **fichiers guide** permet de sélectionner directement le répertoire où sont installés les fichiers exemples fournis avec le programme (sous-répertoire "Guide" du répertoire d'installation).

✧ ***Zone de texte: Fichier, chemin, masque***

Vous pouvez également entrer dans la zone de texte:

- un chemin d'accès;
- un nom de fichier (sans extension, celle-ci étant déterminée automatiquement), avec ou sans son chemin d'accès;
- un masque de recherche (avec les signes "?" et "*") suivant les règles du DOS.

Validez par la touche **Entrée** du clavier, ou en cliquant sur le bouton **ok**.

Création d'un fichier de "secours"

Quand vous enregistrez un fichier, si un fichier ayant le même nom existe déjà celui-ci est automatiquement renommé "SECOURS", avec la même extension (par exemple: "SECOURS.DBT").

Option de remplacement automatique

Par défaut, il vous est demandé de confirmer le remplacement d'un fichier existant. Vous pouvez supprimer cette demande de confirmation, en activant la case à cocher **remplacer automatiquement les fichiers** de la fenêtre de dialogue.

Opérations sur les fichiers

La fenêtre de dialogue permet en outre d'effectuer un certain nombre d'opérations sur les fichiers.

- ***Effacer***
  Sélectionnez le fichier et cliquez sur bouton **effacer**; vous devez alors confirmer l'effacement.

- ***Renommer, copier, déplacer***
  Sélectionnez le fichier et cliquez sur le bouton approprié: **renommer/copier/déplacer**. La fenêtre change d'aspect; choisissez alors un nouveau nom de fichier avec éventuellement un nouveau chemin d'accès et cliquez à nouveau sur le bouton concerné (ou sur le bouton **annuler** pour annuler l'opération).

- ***Visualiser***
  Sélectionnez le fichier et cliquez sur le bouton **visualiser**; le contenu du fichier (éventuellement tronqué) est affiché en bas de la fenêtre.

- ***Imprimer***
  Sélectionnez le fichier et cliquez sur le bouton **imprimer**; une fenêtre de dialogue vous permet de choisir les options "marges haut et gauche" et "police et taille des caractères". Si texte est trop large pour être imprimé sans troncature cela vous est signalé par un message, et vous devez réduire la marge gauche et/ou la taille des caractères.

Attention

Pour l'impression de certains fichiers, il est nécessaire de choisir une fonte non proportionnelle (telle que "Courier New" ou "Courier").

Accès rapide aux derniers fichiers utilisés

Le nom du dernier fichier utilisé pour chaque fonction est affiché dans le sous-menu correspondant du menu général, ce qui permet un accès rapide à ces fichiers.

LES PRINCIPALES FONCTIONS

La fonction données

Sous-menu:
éditer ouvrir un fichier convertir

L'éditeur de données permet d'éditer des données correspondant à un plan de structure S<G>×O (voir le chapitre ☞4), avec au plus 8 *groupes* et 8 *occasions*.

Editer
Pour créer un nouveau fichier de données, ou éditer un fichier précédemment ouvert.

Ouvrir un fichier
Pour éditer un fichier existant (un fichier dont le nom a l'extension ".DBT" ou ".DAT"). ; Sélectionnez l'extension et choisissez le fichier (voir la rubrique "Choisir un nom de fichier"). Quand vous ouvrez un fichier au format **PAC** (extension ".DAT"), une fenêtre apparaît, dans laquelle vous devez indiquer (voir un exemple en ☞5.5.2):
- le nombre d'*objets* contenus dans le fichier (pour chaque sujet);
- les objets utilisés comme G-*facteurs* (facteurs structurant les *groupes*);
- les objets utilisés comme *occasions*.

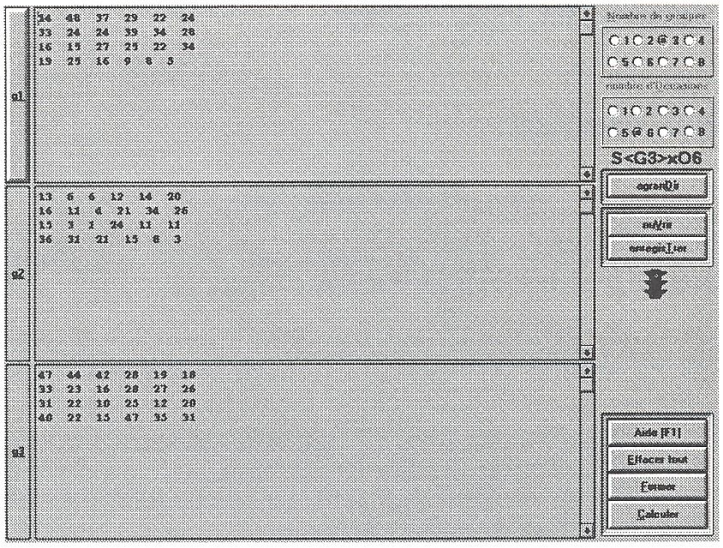
L'éditeur de données
Sélectionnez le nombre de groupes et le nombre d'occasions (passez au **feu vert** s'il y a lieu). L'éditeur comporte une zone de texte par groupe. Par défaut toutes les zones de textes sont affichées simultanément à l'écran; pour afficher une seule zone à la fois, cliquez sur le bouton **agrandir** (qui devient **réduire** pour réaliser l'opération inverse).

Pour chaque groupe entrez successivement, pour chaque sujet, les données relatives à chacune des occasions, séparées par au moins un *espace*. Le passage à la ligne est équivalent à un *espace*. Il n'est pas possible d'utiliser de tabulations.

Attention
Il ne doit pas y avoir de données manquantes. Le nombre d'occasions doit être le même pour chaque sujet et chaque groupe.

Bouton enregistrer
Pour enregistrer le fichier en cours d'édition. Choisissez le fichier (voir la rubrique "Choisir un nom de fichier"); des options permettent de créer des fichiers (données et commandes) au format du logiciel **PAC**.

Bouton fermer
Pour quitter l'éditeur de données (aucun calcul n'est effectué).



✧ ***Bouton calculer***

Si le plan est de structure S<G> ou S×O, les statistiques descriptives (moyennes, écarts-types, covariances) sont calculées et affichées dans la "fenêtre des statistiques". S'il y a plus d'un *groupe* et plus d'une occasion, une fenêtre permet de choisir une structure dérivée: voir notamment les différents exemples du chapitre ☞4.

Convertir

Pour convertir un fichier de données existant de type ASCII (d'extension quelconque) au format standard du logiciel **PAC**. Les caractères particuliers éventuellement utilisés dans le fichier comme séparateurs des données (tabulation et point-virgule) sont automatiquement remplacés par des espaces. Cette fonction permet également de supprimer s'il y a lieu: les caractères non numériques et les lignes d'en-tête. Si la virgule est utilisée dans le fichier comme séparateur décimal, elle permet de la remplacer par le point.

Choisissez le fichier (voir la rubrique "Choisir un nom de fichier"). Apparaît une fenêtre, qui affiche le contenu du fichier et propose un certain nombre d'options.

✧ ***Virgule***

Sélectionnez l'une des options: **convertir en point** (si la virgule est utilisée comme séparateur décimal); **convertir en espace** (si la virgule est utilisée comme séparateur des données); **ne pas convertir**.

✧ ***Caractère non numérique***
Sélectionnez l'une des options: **supprimer**; **convertir en espace**; **convertir en zéro**; **ne pas convertir**.

✧ ***Supprimer lignes d'en-tête***
Indiquez le nombre de lignes d'en-tête à supprimer.

✧ ***Bouton convertir***
Après avoir sélectionné les options, cliquez sur le bouton **convertir**. Le résultat de la conversion est affiché (mais n'est pas enregistré: voir ci-après le bouton **enregistrer**).

✧ ***Annuler convertir***
Après conversion, le bouton devient **annuler convertir**; cliquez dessus si vous souhaitez rétablir le contenu initial du fichier.

✧ ***Bouton enregistrer***
Cliquez sur le bouton **enregistrer** pour enregistrer le résultat de la conversion dans un fichier, qui aura obligatoirement l'extension ".DAT" (voir la rubrique "Choisir un nom de fichier").

Remarque
Vous pouvez utilisez cette fonction pour créer un fichier d'extension ".DAT", si le fichier de données n'a pas cette extension. Dans ce cas utilisez seulement la fonction **enregistrer** sans convertir le fichier.

La fonction mode d'entrée

Sous-menu:
inférence sur une moyenne plan S<G> ou S×O
effet - échelle effet calibré - constante b entrée directe de la distribution
ouvrir un fichier PAC ouvrir un fichier de courbes

Inférence sur une moyenne et plan S<G> ou S×O (boutons moyenne et G / O)
Vous entrez les statistiques descriptives pertinentes dans la "fenêtre des statistiques" qui est affichée (les données complètes ne sont donc pas nécessaires); pour son utilisation, voir les différents exemples donnés dans les chapitres ☞1 à ☞7.

Effet - échelle, effet calibré - constante b et entrée directe de la distribution (boutons effet, ...calibré et directe)
Vous entrez directement les caractéristiques des distributions fiducio-bayésiennes (il est possible d'utiliser des résultats d'analyse de variance ou de tests de signification) relatives, suivant le cas, à l'effet (voir ☞2.1.6), à l'effet calibré (voir ☞2.1.7) ou à l'effet et l'écart-type (voir ☞7.1.3).

Ouvrir un fichier PAC (bouton PAC)

Pour utiliser un fichier de résultats créé par le logiciel **PAC**. Choisissez le fichier (d'extension ".BAY") (voir la rubrique "Choisir un nom de fichier"). Des exemples sont fournis avec le programme.

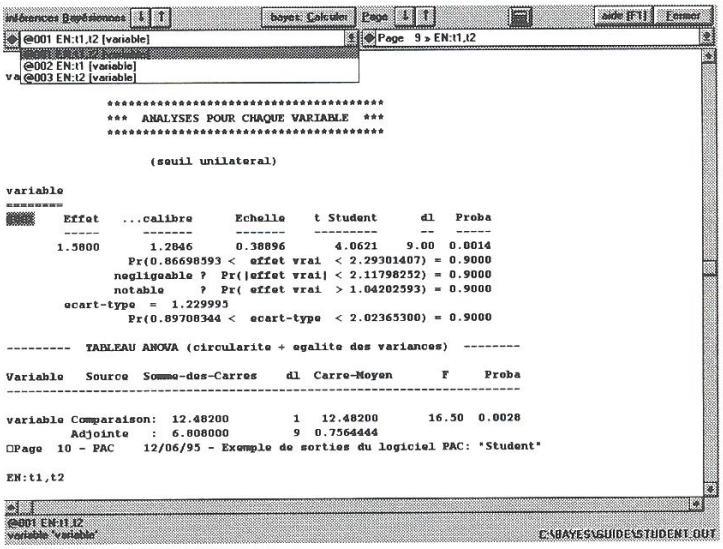
Visualisation des sorties

Un fichier des sorties produites par **PAC** (d'extension ".OUT") est en principe associé au fichier choisi (si toutefois ce fichier n'est plus disponible, vous ne pouvez plus visualiser les sorties, mais cela n'empêche pas d'effectuer les analyses bayésiennes).

La liste des **inférences bayésiennes** apparaît en haut et à gauche de l'écran dans une autre liste déroulante; ces inférences sont numérotées "@001", "@002", etc. Un descriptif de l'inférence sélectionnée est affiché en bas de l'écran.

Les sorties sont paginées; la liste des **pages** (avec un descriptif) apparaît en haut et à droite de l'écran dans une liste déroulante. Les pages sont numérotées "Page 1", "Page 2", etc.

Vous pouvez sélectionner une page ou une inférence bayésienne, en cliquant sur son intitulé dans la liste correspondante. Vous pouvez également parcourir l'ensemble des pages ou des inférences bayésiennes, en cliquant sur les boutons ↓ et ↑ correspondants.



Effectuer une inférence bayésienne

Sélectionnez l'inférence bayésienne voulue dans la liste déroulante correspondante des sorties (en haut à gauche de l'écran), puis cliquez sur le bouton **bayes: calculer**, ce qui

vous ramène à la fenêtre principale Vous pouvez également choisir directement une inférence bayésienne en "double-cliquant" sur la ligne correspondante (commençant par @) dans la sortie de **PAC** si celle-ci est disponible.
Dans la fenêtre principale, un nouveau bouton, **distribution PAC**, apparaît: il vous ramène aux sorties de **PAC**.

Ouvrir un fichier de courbes
Pour imprimer une courbe qui a été précédemment enregistrée. Choisissez le fichier (d'extension ".DIS") (voir la rubrique "Choisir un nom de fichier").
Apparaît une fenêtre avec la liste des courbes contenues dans le fichier. Pour la suite des opérations, voir la rubrique "Imprimer une distribution".

La fonction énoncés

Pour calculer des énoncés relatifs à une distribution bayésienne (après calcul); il suffit en fait de cliquer sur la distribution correspondante. La "fenêtre des énoncés" est alors affichée. Pour la suite des opérations, voir les exemples traités dans les chapitres ☞1 à ☞3 et ☞5 à ☞7

La fonction options

Sous-menu:
pas de distribution initiale feu vert/rouge
esquisse (courbes) calculatrice

Pas de distribution initiale
Pour cacher l'option de choix d'une distribution initiale.

Feu vert/rouge
Pour pouvoir modifier les valeurs qui sont le résultat d'un calcul.

Esquisse (courbes)
Pour accélérer l'exécution, le nombre de points utilisé pour dessiner la courbe étant limité. Cette option est souvent satisfaisante; elle peut cependant dans certains cas donner des résultats médiocres.

Calculatrice
La calculatrice est accessible suivant les cas, soit par menu, soit par son icône, et toujours par la touche **F9** du clavier. Elle fonctionne en notations *algébriques* et en notations *polonaises inverses*. Le mode actif au lancement de l'application est celui de la dernière utilisation.

✧ ***Icône***

Quand l'icône de la calculatrice est visible, pour effectuer un calcul sur une valeur contenue dans une zone de texte (l'option "feu vert" doit être active), cliquez sur l'icône, puis, sans relâcher le bouton de la souris, faites glisser l'image au dessus de la zone de texte et relâchez le bouton de la souris. La valeur voulue de cette zone s'affichera dans la calculatrice et sera remplacée par le résultat obtenu.

Les bouton √ et **1/x** s'utilisent de la même manière pour obtenir la racine carrée ou l'inverse.

Vous pouvez entrer les nombres et les opérateurs de base +, -, *, /, = (↑ ou **Entrée** en notations polonaises inverses) au clavier ou à la souris. Les autres opérateurs utilisent la souris.

En notations polonaises inverses, la pile a dix registres; ses deux premiers registres sont visibles.

✧ ***Mémoires A et B***

Les deux mémoires appelées A et B permettent d'effectuer des sommes (boutons **A+x** et **B+x**), des soustractions (boutons **A-x** et **B-x**) et des sommes de carrés (boutons **A+x^2** et **B+x^2**). Leur contenu est visible. Il est mis à zéro par les boutons **cl A** et **cl B**, et il est rappelé et affiché par les boutons **Rcl A** et **Rcl B**.

✧ ***Mémoires 0 à 9***

Pour stocker la valeur **x** dans l'une de ces mémoires, cliquez sur le bouton **Sto** puis, sans relacher le bouton de la souris, faites glisser le rectangle au dessus de l'un des boutons **0**, **1**,... Les boutons **Sto+** et **Sto-** s'utilisent de la même manière et permettent respectivement d'ajouter la valeur **x** au contenu de la mémoire et de la retrancher. Pour rappeler le contenu de la mémoire, procédez de la même manière avec le bouton **Rcl**.

La fonction quitter

Pour quitter **LeBayésien**. Apparaît une fenêtre de dialogue qui vous demande de confirmer.

La fonction ?

Pour afficher le numéro de version et le *copyright*.

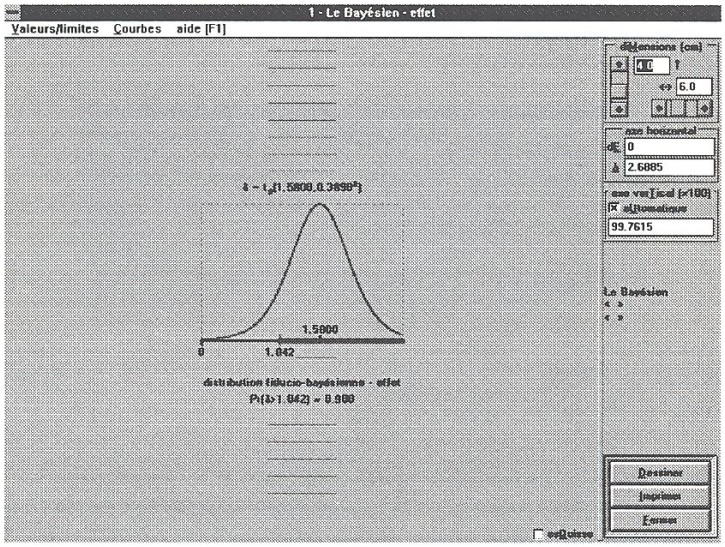
Le bouton exemples

Pour avoir un accès direct aux différents exemples traités dans l'ouvrage, ainsi qu'aux exemples de sorties du logiciel **PAC** fournis.

IMPRIMER UNE DISTRIBUTION

Procédure

1. Dans la "fenêtre des énoncés", cliquez sur la courbe qui vous intéresse. Celle-ci apparaît alors en plein écran.

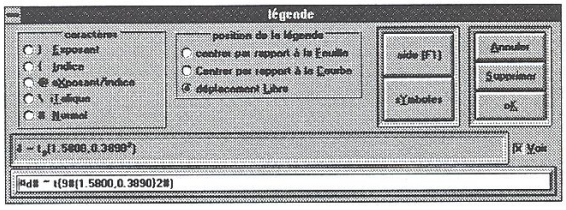


- Dans le cadre en haut et à droite de l'écran, choisissez les dimensions, hauteur (↑) et largeur (↔) de la courbe imprimée en centimètres (les dimensions à l'écran leur sont proportionnelles).
- Dans le cadre **axe horizontal**, choisissez l'étendue des valeurs de la variable sur l'axe horizontal (**de... à**).
- Dans le cadre **axe vertical**, si l'option **automatique** est active la valeur maximale de cet axe (donc son échelle) est déterminée automatiquement (quelle que soit la valeur fournie). La valeur affichée est multipliée par 100.
- Dans le cadre **axe vertical**, si l'option **automatique** est désactivée, c'est la valeur fournie qui est utilisée comme valeur maximale; ceci est utile pour mettre plusieurs courbes à la même échelle, ou pour modifier l'aspect de la courbe de distributions dont la densité prend des valeurs infinies, mais ne devrait pas être utilisé en dehors de ces deux cas.

- L'option **esquisse** (case à cocher en bas de l'écran) peut être utilisée pour accélérer l'exécution, le nombre de points utilisé pour dessiner la courbe étant limité. Cette option est souvent satisfaisante; elle peut cependant dans certains cas donner des résultats médiocres. Elle ne concerne que l'affichage à l'écran (et pas l'imprimante).

2. Modifiez les valeurs précédentes et cliquez sur le bouton **dessiner** pour obtenir les changements.

3. Vous pouvez éditer, modifier et déplacer les légendes. Pour cela cliquez sur la légende, ce qui fait apparaître une fenêtre d'édition.



- Vous pouvez inclure des lettres en italiques, des lettres grecques, des symboles, des indices et des exposants, et même des combinaisons d'indices et d'exposants. Pour cela des caractères *réservés* (non imprimés) sont utilisés comme marqueurs et agissent sur le texte qui suit.

}	petits caractères en exposant - Exemple "}2" donne "^{2}"
{	petits caractères en indice - Exemple "{23" donne "$_{23}$"
@	petits caractères en exposant, utilisé pour obtenir simultanément des exposants et des indices (si { est utilisé ensuite) - Exemple "@2{23" donne "$^{2}_{23}$"
\	caractères en italique - Exemple "\Student" donne "*Student*"
#	retour aux caractères normaux
¤	utilisation de la police "*symbol*" pour le caractère suivant (et pour celui-là seul) - Exemple "¤s=" donne "σ="

Vous pouvez taper directement ces caractères réservés au clavier ou cliquer sur l'option appropriée pour obtenir leur affichage à l'endroit du curseur. L'affichage d'un symbole (ou d'une lettre grecque) peut être obtenu en cliquant sur le bouton **symboles** et en sélectionnant le symbole voulu dans la liste qui apparaît (pour faire disparaître cette liste sans sélection, cliquez à nouveau sur **symboles**).

Remarque

Si l'affichage n'est pas assez rapide, vous pouvez cliquer sur **voir** pour supprimer l'affichage de la légende telle qu'elle apparaîtra à l'écran.

• Vous pouvez également choisir l'option d'alignement de la légende: centrée sur la courbe, centrée sur la feuille imprimée, ou placée librement.

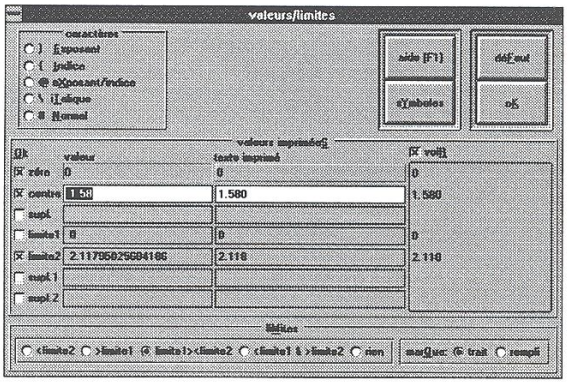
• Une fois effectuées les modifications souhaitées, cliquez sur le bouton **ok**. Positionnez ensuite la légende en la déplaçant à l'aide de la souris puis cliquez pour la "déposer". Dans l'option "déplacement libre" elle sera placée à l'endroit choisi, et dans les options "centrée" elle sera centrée automatiquement (la différence entre les deux options de centrage n'apparaît qu'à l'impression).

✧ ***Déplacer une légende***
Vous pouvez obtenir le déplacement seul de la légende en cliquant avec le bouton secondaire de la souris et en maintenant ce bouton enfoncé jusqu'à la position voulue. Pour obtenir le déplacement libre de la légende (si ce n'est pas déjà l'option active), enfoncez l'une des touches **Majuscule**, **Ctrl**, ou **Alt** du clavier et cliquez avec le bouton secondaire en maintenant cette touche enfoncée.

✧ ***Supprimer une légende***
Vous pouvez supprimer une légende en la déplaçant jusqu'à la partie droite de l'écran, ou au contraire ajouter une légende située dans cette partie droite en cliquant dessus.

✧ ***Ajouter une légende***
Vous pouvez ajouter une légende non utilisée (située dans la partie droite) en l'éditant ou en la déplaçant (les légendes "vides" sont indiquées par les signes "« »".

4. La fonction **valeurs/limites** du menu fait apparaître une fenêtre d'édition, qui permet les opérations suivantes:



- modifier les valeurs des limites;
- éditer les textes qui leur sont associés (avec les mêmes fonctions d'édition que pour les légendes);

- ajouter des points supplémentaires;
- choisir la manière de figurer les limites: par un trait (option par défaut), ou en remplissant la courbe.

Après modification, cliquez sur **ok** pour fermer la fenêtre, puis sur le bouton **dessiner** pour obtenir les changements.

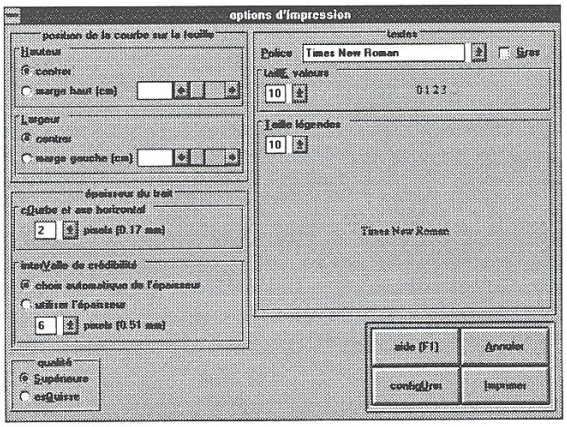
Attention

Quand l'étendue des valeurs de la variable sur l'axe horizontal est modifiée, les limites et valeurs qui peuvent être affichées sont changées et sont automatiquement remises à jour. Il est donc préférable de fixer d'abord cette étendue.

5. Pour imprimer cliquez sur le bouton **imprimer**. Apparaît une fenêtre d'options. Vous pouvez choisir:

- l'alignement de la courbe sur la feuille, centrée en hauteur et largeur (options par défaut) ou positionnée à l'emplacement de votre choix;
- l'épaisseur du trait du dessin pour la courbe et l'axe horizontal;
- l'épaisseur du trait du dessin qui représente l'intervalle de crédibilité;
- la police et la taille des caractères pour les textes, qui peuvent également être imprimés en gras;
- la qualité du dessin; l'option esquisse ne devrait cependant être utile que si vous avez un matériel "lent".

Le bouton **configurer** donne accès aux options de configuration de Windows.



Cliquez sur **imprimer** pour lancer l'impression.

Mettre des courbes à la même échelle

Pour mettre deux ou plusieurs courbes à la même échelle.

1. Le rapport entre la largeur de la courbe imprimée et l'étendue de l'axe horizontal doit être constant:
   - par exemple utilisez la même largeur et la même étendue ("**de** -2 **à** 2" pour la courbe 1, "**de** -1 **à** 3" pour la courbe 2, etc.);
   - si par exemple l'étendue est "**de** 0 **à** 3" pour la courbe 1, "**de** -2 **à** 4 pour la courbe 2, la largeur imprimée doit être double pour la courbe 2.
2. Utilisez la même hauteur pour toutes les courbes et choisissez la valeur maximale de l'axe vertical commune à toutes les courbes (au moins égale au plus grand des modes). Désactivez l'option **automatique**.

Enregistrement/lecture des courbes

Quand vous fermez la fenêtre plein écran, la courbe affichée est automatiquement sauvegardée dans un fichier temporaire ("TEMPO.DIS"), avec ses caractéristiques et ses légendes. Le sous-menu de la fonction **courbes** de cette fenêtre permet les opérations suivantes sur les courbes.

Choisir une courbe

Apparaît une fenêtre avec la liste des courbes précédemment dessinées (au cours de la même session). Sont affichées les légendes, l'étendue des valeurs de la variable sur l'axe horizontal (**de... à**); et les dimensions de la courbe.

- **Bouton *dessiner***
  Pour dessiner une courbe (en remplacement de la courbe actuelle), sélectionnez cette courbe dans la liste et cliquez sur **dessiner**.
- **Bouton *ajouter***
  Pour ajouter une courbe à la courbe actuelle, sélectionnez cette courbe dans la liste et cliquez sur **ajouter**. Les deux courbes sont mises automatiquement à la même échelle et dessinées simultanément.
  Si deux courbes sont déjà dessinées, la courbe sélectionnée remplace la seconde
- ***Options***
  - Garder l'axe horizontal actuel.
  - Garder les dimensions actuelles.

Enregistrer la courbe

La courbe affichée (éventuellement les deux courbes) est ajoutée au fichier temporaire, avec ses caractéristiques et ses légendes. En fait la courbe affichée est automatiquement sauvegardée

chaque fois qu'une nouvelle courbe est dessinée. Cette fonction n'est donc utile que si vous souhaitez conserver une courbe avant d'en modifier les dimensions, l'étendue, les légendes...

Copier dans le presse-papiers

La courbe (éventuellement les deux courbes) est copiée en mode "Bitmap" dans le "presse-papiers" (sans les légendes). Vous pouvez alors la "coller" dans une application graphique.

Créer un fichier "Bitmap"

Choisissez un nom de fichier (d'extension ".BMP") (voir la rubrique "Choisir un nom de fichier"). La courbe affichée (éventuellement les deux courbes) est copiée en mode "Bitmap" dans ce fichier (sans les légendes).

Créer un fichier texte des densités

Choisissez un nom de fichier (d'extension ".TXT") (voir la rubrique "Choisir un nom de fichier"). Les coordonnées de la courbe affichée (éventuellement les deux courbes) sont enregistrées en mode "texte" dans ce fichier. Chaque ligne du fichier contient l'abscisse, puis l'ordonnée de la courbe (ou les ordonnées de chacune des deux courbes s'il y a lieu).

Choisir dans un fichier

Choisissez un fichier de sauvegarde existant (d'extension ".DIS") (voir la rubrique "Choisir un nom de fichier"). La procédure est ensuite la même que pour la fonction **choisir une courbe**.

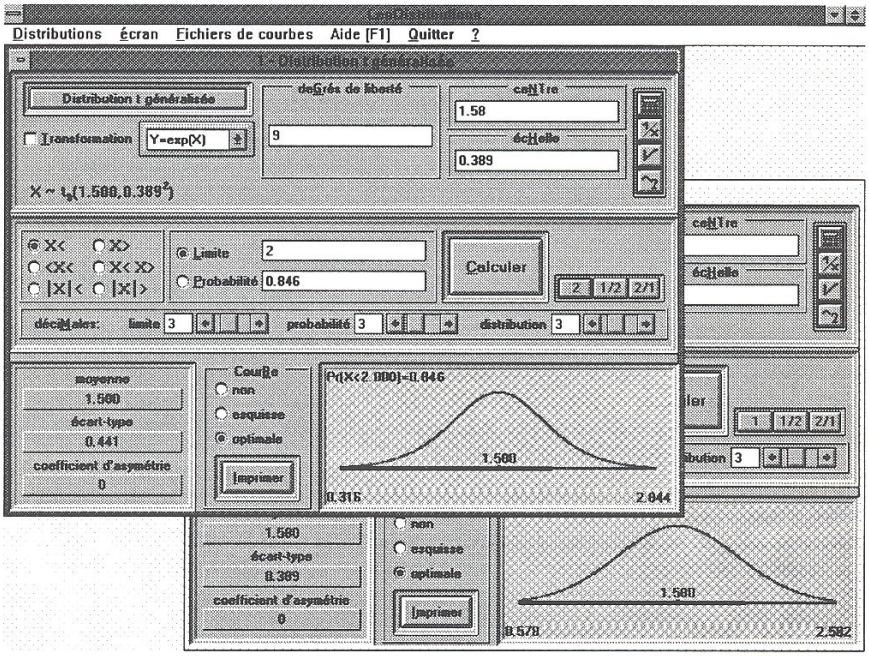
Enregistrer tout dans un fichier

Choisissez un nom de fichier (d'extension ".DIS") (voir la rubrique "Choisir un nom de fichier"), dans lequel toutes les courbes sauvegardées dans la session seront enregistrées.

Supprimer toutes les courbes en mémoire

Supprime toutes les courbes actuellement en mémoire et recommence la numérotation des courbes à 1.

LesDistributions



Le programme **LesDistributions** permet d'obtenir, à la fois les courbes de densité et les énoncés de probabilité correspondant aux fonctions de répartition directe et inverse, de toutes les distributions introduites dans cet ouvrage (voir le chapitre ☞9).

Les principes généraux d'utilisation sont les mêmes que ceux du programme **LeBayésien**. En particulier le module d'impression des courbes de distributions est identique.

Nous donnerons simplement ici un exemple de prise en mains.

Prise en mains

Au lancement du programme, la distribution sélectionnée est la **distribution normale**. Pour choisir une autre distribution, cliquez sur le bouton correspondant ou sur la fonction **distributions** du menu, et sélectionnez la distribution dans la liste qui s'affiche.

Par exemple, pour obtenir la distribution $X \sim t_9(+1.580,0.389^2)$, sélectionnez **t...** puis dans la sous-liste **t généralisée**. Entrez le centre (1.580), l'échelle (0.389), et les degrés de liberté (9), puis cliquez sur le bouton calculer: la courbe est dessinée.

Enoncés

Pour obtenir un énoncé relatif à cette distribution, sélectionnez:

- le type d'énoncé voulu ("**X<**" ou "**X>**"...);
- l'option **limite** pour calculer la probabilité correspondant à une limite (ou deux limites) donnée;
- l'option **probabilité** pour calculer la limite correspondant à une probabilité donnée.

Exemples

Calcul de probabilité

Sélectionnez "**X<**" et l'option **limite** et entrez la limite 2, puis cliquez sur calculer: vous obtenez Pr(X<2.000)=0.846.

Sélectionnez "**<X<**" et entrez les deux limites 2 et 2.5, puis cliquez sur calculer: vous obtenez Pr(2.000X<2.500)=0.133.

Calcul de limite

Sélectionnez "**|X|<**" et l'option **probabilité** et entrez la probabilité 0.975, puis cliquez sur calculer: vous obtenez Pr(|X|<2.460)=0.975.

Transformations

Un certain nombre de transformations sont disponibles (suivant la distribution). Par exemple, dans la situation précédente, cliquez sur la case à cocher **Transformation**, puis sur le bouton calculer (sans rien changer d'autre): vous obtenez la distribution de la variable $Y=exp(X)$ et l'énoncé Pr(|Y|<11.705)=0.975 (11.705=*exp*(2.460)).

Etude simultanée de deux distributions

Vous pouvez étudier simultanément deux distributions. Pour cela, cliquez sur le bouton "**2**" (à droite du bouton **calculer**). Les boutons "**1**", "**2**", "**1/2**" et "**2/1**" (ou la fonction **écran** du menu) permettent de contrôler l'affichage des fenêtres.

Impression des courbes

Cliquez sur le bouton **imprimer** (et non sur la courbe de la distribution).

Remarque

Bien que les fichiers de sauvegarde des courbes des distributions aient la même extension ".DIS" et le même format que dans le programme **LeBayésien**, les courbes qu'ils contiennent ne sont pas nécessairement utilisables par le programme qui ne les a pas créées.

RÉFÉRENCES BIBLIOGRAPHIQUES

RÉFÉRENCES CITÉES DANS L'OUVRAGE

Les nombres entre crochets indiquent les pages où la référence est citée

Anderson S., Hauck W.W. (1983) – A new procedure for testing equivalence in comparative bioavailability and other clinical trials. *Communications in Statistics, Theory and Methods* **12**, 2663-2692. [225]

Barnard G.A. (1963) – Some logical aspects of the fiducial argument. *Journal of the Royal Statistical Society B* **25**, 111-114. [200]

Basu D. (1977) – On the elimination of nuisance parameters. *Journal of the American Statistical Association* **72**,355-366. [200]

Berger G.O. (1985) – *Statistical Decision Theory and Bayesian Analysis*. New York: Springer Verlag. [23, 28, 64]

Bernardo J.M., Smith A.F.M. (1994) – *Bayesian Theory*. Chichester: Wiley. [28, 43, 316]

Bondy W.A. (1969) – A test of an experimental hypothesis of negligible difference between means. *The American Statistician*, 28-30. [225]

Box G.E.P., Tiao G.C. (1973) – *Bayesian Inference In Statistical Analysis*. Reading, Massachusetts: Addison Wesley. [43, 199]

Brown M., Forsythe A. (1974) – The small sample behavior of some statistics which test the equality of several means. *Technometrics* **16**, 129-132. [144, 194]

Choi S.C., Pepple P.A. (1989) – Monitoring clinical trials based on predictive probability of significance. *Biometrics* **45**, 317-323. [118]

Cohen J. (1977) – *Statistical power analysis for the behavioral sciences*. New York: Academic Press. [90]

Dawid A.P. (1980) – A Bayesian look at nuisance parameters. *In* Bernardo J.M., De Groot M.H., Lindley D.V., Smith A.F.M. (Eds.), *Bayesian Statistics*, Valencia: University Press, 167-184. [200]

de Finetti B. (1975) – *Theory of Probability (volume 2)*. New York: Wiley. [11]

Deheuvels P. (1984) – How to analyze bio-equivalence studies? The right use of confidence intervals. *Journal of Organizational Behaviour and Statistics* 1, 1-15. [228, 230]

DuMouchel W. (1989) – Bayesian metaanalysis. *In* Berry D.A. (Ed.), *Statistical Methodology in the Pharmaceutical Science*. New York: Marcel Dekker, 509-529. [102]

Fisher R.A. (1935) – The fiducial argument in statistical inference. *Annals of Eugenics*, 6, 391-398. [199]

Fisher R.A. (1990) – *Statistical Methods, Experimental Design and Scientific Inference* (rééditions). Oxford: Oxford University Press. [23, 199, 261, 297, 299]

Gertsbakh I., Winterbottom A. (1991) – Point and interval estimation of normal tail probabilities. *Communications in Statistics - Theory and Methods* 20, 1497-1514. [53]

Grouin J.-M. (1994) – *Procédures bayésiennes prédictives pour les essais expérimentaux*. Thèse de doctorat de Mathématiques Appliquées. Paris: Université René Descartes. [311]

Hájek J. (1965) – On basic concepts of statistics. *Fith Berkeley Symposium on Mathematical Statistic and Probability*, Berkeley: University of California Press 1, 139-162. [200]

Hand D., Taylor C. (1987) – *Multivariate Analysis of Variances and Repeated Measures: A practical Approach for Behavioural Scientists*. London and New York: Chapman and Hall. [140]

Hauck W.W., Anderson S. (1984) – A new statistical procedure for testing equivalence in two-group comparative bioavailability trials. *Journal of Pharmacokinetics and Biopharmaceutics* 12, 83-91. [225]

Hodges J.L., Lehmann E.L. (1954) – Testing the approximate validity of statistical hypotheses. *Journal of the Royal Statistical Society B* 16, 261-268. [225]

Holender D., Bertelson P. (1975) – Selective preparation and time uncertainty. *Acta Psychologica* 39,193-203. [72]

James G.S. (1951) – The comparison of several groups of observations when the ratios of the population variances are unknown. *Biometrika* **38**, 324-329. [194]

Jeffreys H. (1961) – *Theory of Probability* (third edition). Oxford University Press. [22]

Kirkwood T.B.L. (1981) – Bioequivalence testing: a need to rethink (reader reaction). *Biometrics* **37**, 589-591. [228]

Lecoutre B. (1981a) – Extensions de l'analyse de la variance: L'analyse bayésienne des comparaisons. *Mathématiques et Sciences Humaines* **75**, 49-69. [271]

Lecoutre B. (1981b) – Procédures fiducio-bayésiennes pour l'investigation des mécanismes individuels en psychologie. *L'Année Psychologique* **81**, 453-464. [147]

Lecoutre B. (1984) – *L'Analyse Bayésienne des Comparaisons*. Lille: Presses Universitaires de Lille. [45, 103]

Lecoutre B. (1985) – Reconsideration of the F test of the analysis of variance: The semi-Bayesian significance test. *Communications in Statistics - Theory and Methods* **14**, 2437-2446. [196]

Lecoutre B., Derzko G., Grouin J.-M. (1995) – Bayesian predictive approach for inference about proportions. *Statistics in Medicine*, **14**, 1057-1063. [68, 118]

Lecoutre B., Guigues J.-L., Poitevineau J. (1992) – Distribution of quadratic forms of multivariate Student variables. *Applied Statistics*, **41**,.617-627. [271, 283]

Lecoutre B., Poitevineau J. (1995) – PAC: *Programme d'Analyse des Comparaisons*, version 1.5. Logiciel édité et distribué par: CISIA, 1 avenue Herbillon, 94160 Saint Mandé (France). [8]

Lecoutre M.-P., Durand J.-L. (1988) – Jugements probabilistes et modèles cognitifs: étude d'une situtation aléatoire. *Educational Studies in Mathematics* **19**, 357-368. [69]

Lépine D., Rouanet H. (1975) – Introduction aux méthodes fiduciaires: Inférence sur un contraste entre moyennes. *Cahiers de Psychologie* **18**, 193-218. [45]

Lhoste E. (1923) – Le calcul des probabilités appliqué à l'artillerie. *Revue d'Artillerie* **91**. [43]

Lindley D.V. (1965) – *Introduction to Probability and Statistics from a Bayesian Viewpoint (deux volumes)*. Cambridge University Press. [139]

Lunney G.H. (1970) – Using analysis of variance with a dichotomous dependent variable: An empirical study. *Journal of Educational Measurement* **7**, 263-269. [144]

Nel D.G., Van der Merwe C.A.(1986) – A solution to the multivariate Behrens-Fisher problem. *Communications in Statistics: Theory and Methods* **15**, 3485-3487. [194]

Owen (1965) – A special case of bivariate non-central t distribution. *Biometrika* **52**, 437-446. [256]

Patil V.H. (1965) – Approximations to the Behrens-Fisher distribution. *Biometrika* **52**, 267-271. [314]

Pearson E.S. (1931) – The analysis of variance in cases of non-normal variation. *Biometrika* **21**, 259-286. [144]

Phillips K.F. (1990) – Power of the two one-sided tests procedure in bioequivalence. *Journal of Pharmacokinetics and Biopharmaceutics* **18**, 137-144. [256, 258]

Press S.J. (1989) – *Bayesian Statistics; Principles, Models and Applications.* New York: Wiley. [103]

Ramsey P. (1980) – Exact type I error rates for robustness of Student's t-test with unequal variances. *Journal of Educational Statistics* **5**, 337-349. [144]

Rao C.R. (1965) – *Linear Statistical Inference and Its Applications.* New York: John Wiley & Sons. [77]

Robert C. (1992) – *L'analyse statistique bayésienne.* Paris: Economica. [22, 28, 101]

Rocke D.M. (1984) – On testing for bioequivalence. *Biometrics* **40**, 225-230. [225]

Rocke D.M. (1985) – Reply to correspondence. *Biometrics* **41**, 563. [225]

Rouanet H. (1995) – Assessing the importance of the effects with Bayesian methods. *Psychological Bulletin*, sous presse. [45]

Rouanet H., Bernard J.-M., Lecoutre B. (1986) – Nonprobabilistic statistical inference: A set-theoretic approach. *The American Statistician* **40**, 60-65. [33]

Rouanet H., Bernard J.-M., Le Roux B. (1990) – *Statistiques en sciences humaines: Analyse inductive des données.* Paris: Dunod. [33]

Rouanet H., Lecoutre B. (1983) – Specific inference in ANOVA: From significance tests to Bayesian procedures. *British Journal of Mathematical and Statistical Psychology* **36**, 252-268. [45, 140]

Rouanet H., Lecoutre B., Bernard J.-M. (1987) – L'inférence fiducio-bayésienne comme méthode d'analyse de données: Un exemple d'application à des données psychométriques. *Statistique et Analyse des Données* **11**, 58-74. [33, 91]

Rouanet H., Lecoutre M.-P., Bert M.-C., Lecoutre B., Bernard J.-M. (1991) – *L'inférence statistique dans la démarche du chercheur*. Berne: Lang. [8]

Rouanet H., Lépine D., Holender D. (1978) – Model acceptability and the use of Bayes-fiducial procedures for validating models. *In* J. Requin (Ed.), *Attention and Performance VII*, Hillsdale, NJ: Erlbaum, 687-701. [45]

Rouanet H., Lépine D., Pelnard-Considère J. (1975) – Bayes-fiducial procedures as practical substitutes for misplaced significance testing: An application to educational data. *In* D.N.M. de Gruijter and L.J.Th. Van der Kamp (Eds), *Advances in Psychological and Educational Measurement*, New York: John Wiley & Sons, 33-50. [45, 91]

Rouanet H., Oléron G., Régnier J. (1966) – Analyse de variance pour données appareillées et modèles de dépendance: Illustration d'une démarche. *L'Année Psychologique* **66**, 131-165. [147]

Satterthwaite F.E. (1946) – An approximation distribution of estimates of variance components. *Biometrics Bulletin* **2**, 110-114. [194]

Schervish M.J. (1992) – Bayesian analysis of linear models. *In* Bernardo J.M., Berger J.O., Dawid A.P., Smith A.F.M. (Eds.), *Bayesian Statistics IV*, Oxford: University Press, 419-434. [271]

Schuirmann D.J. (1987) – A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics* **15**, 657-680. [225, 227]

Schwartz D., Flamant R., Lellouch J. (1981) – *L'essai thérapeutique chez l'homme*. Paris: Flammarion. [234]

Sechrest L., Yeaton W.H. (1982) – Magnitudes of experimental effects in social science research. *Evaluation Review* **6**, 579-600. [90]

Selwyn M.R., Hall N.R., Dempster A.P. (1985) – Letter to the editor. *Biometrics* **41**, 561. [225]

Spiegelhalter D.J., Freedman L.S., Parmar M.K.B. (1994) – Bayesian approaches to randomized trials. *Journal of the Royal Society A* **157**, 357-416. [102, 103]

Steinjans V.W., Diletti E. (1983) – Statistical analysis of bioavailability studies: parametric and non parametric confidence intervals. *European Journal of Clinical Pharmacology* **24**, 127-136. [217]

Stone M., Dawid A.P. (1972) – Bayesian implications of improper Bayesian inference in routine statistical problems. *Biometrika* **59**, 369-373. [315]

Student (1908) – The probable error of a mean, *Biometrika* **6**, 1-25. [45]

Tan W. (1982) – Sampling distributions and robustness of t, f and variance-ratio in two samples and ANOVA models with respect to departure from normality. *Communications in Statistics: Theory and Methods* **11**, 2485-2511. [144]

Welch B.L. (1947) – The generalization of Student's problem when several different population variances are involved. *Biometrika* **34**, 28-35. [194]

Welch B.L. (1951) – On the comparison of several mean values: An alternative approach. *Biometrika* **38**, 330-336. [194]

Westlake W.J. (1976) – Symmetrical confidence intervals in analysis of comparative bioavailability trials. *Biometrics* **32**, 741-744. [229]

Westlake W.J. (1981) – Response to bioequivalence testing: A need to rethink (reader reaction response). *Biometrics* **37**, 591-593. [228]

Wilcox R., Charlin V., Thompson K. (1986) – New monte carlo results on the robustness of the ANOVA F, W and F* statistics. *Communications in Statistics: Simulation and Computation* **15**, 933-934. [144]

RÉFÉRENCES COMPLÉMENTAIRES

"Mises en garde" à l'égard des tests de signification

Dans ces articles méthodologiques récents, le lecteur pourra trouver de nombreuses références concernant les débats sur l'usage des tests de signification dans la recherche expérimentale.

Cohen J. (1994) – The earth is round ($p<.05$). *American Psychologist* **12**, 997-1003.

Falk R. (1995), Greenbaum C.W. – Significance tests die hard: The amazing persistence of a probabilistic misconception. *Theory and Psychology* **5**, 75-98.

Freeman P.R. (1993) – The role of *p*-values in analysing trial results. *Statistics in Medicine* **12**, 1443-1452.

Ouvrages techniques d'introduction aux méthodes bayésiennes

Les ouvrages suivants fournissent une introduction générale relativement élémentaire aux méthodes bayésiennes.

• *Ouvrage écrit pour des statisticiens*

Lee P. (1989) – *Bayesian Statistics: An introduction.* Oxford: Oxford University Press.

• *Ouvrages écrits pour des utilisateurs*

Novick M.R., Jackson P.H. (1974) – *Statistical Methods for Educational and Psychological Research.* New York: McGraw-Hill.

Phillips L.D. (1973) – *Bayesian Statistics for Social Scientists.* London: Nelson.

Application des méthodes fiducio-bayésiennes

Ces références complètent celles données dans cet ouvrage, et fournissent des considérations méthodologiques sur l'utilisation des procédures fiducio-bayésiennes, avec des exemples d'applications à des données concrètes.

Ciancia F., Maitte M., Honoré J., Lecoutre B., Coquery J.-M. (1988) – Orientation of attention and sensory gatting: An evoked potential and RT study in cat. *Experimental Neurology* **100**, 274-287.

Hoc J.-M. (1983) – *L'analyse planifiée des données en psychologie.* Paris: Presses Universitaires de France.

Lecoutre B. (1985) – How to derive Bayes-fiducial conclusions from usual significance tests. *Cahiers de Psychologie Cognitive* **5**, 553-563.

Lecoutre B. (1987) – Les procédures fiducio-bayésiennes et semi-bayésiennes dans les études de généralisabilité. *Informatique et Sciences Humaines* **72-73**, 133-156.

Lecoutre B., Lecoutre M.-P. (1979) – A propos d'une expérience d'apprentissage perceptif incident: Quelques aspects de la démarche d'analyse des données et méthodes fiduciaires. *Psychologie Française* **24**, 269-276.

Lecoutre B., Rouanet H., Denhière G. (1988) – L'inférence statistique comme instrument de validation de modèles. *In* Caverni J.-P., Bastien C., Mendelsohn P., Tiberghien G. (Eds), *Psychologie cognitive: Modèles et méthodes*, Grenoble: Presses Universitaires de Grenoble, 383-394.

Articles méthodologiques sur les applications des procédures bayésiennes aux essais cliniques

De nombreux articles méthodologiques récents sont consacrés aux applications des procédures bayésiennes aux essais cliniques.

Berry D.A. (1991) – Experimental design for drug development: A Bayesian approach. *Journal of Pharmaceutical Statistics* **1**, 81-101.

Berry D.A. (1993) – A case for Bayesianism in clinical trials. *Statistics in Medicine* **12**, 1377-1393.

Racine A., Grieve A.P., Flühler H. (1986) – Bayesian methods in practice: Experiences in the pharmaceutical industry. *Applied Statistics* **35**, 93-150.

Spiegelhalter D.J., Freedman L.S. (1988) – Bayesian approaches to clinical trials. *In* J.M. Bernardo, M.H. DeGroot, D.V. Lindley, A.F.M. Smith (Eds.), *Bayesian Statistics III*, Oxford: University Press, 453-477.

Spiegelhalter D.J., Freedman L.S., Parmar M.K.B. (1993) – Applying Bayesian ideas in drug developement and clinical trials. *Statistics in Medicine* **12**, 1501-1511.

TABLE DES MATIÈRES

page

5 **Sommaire**

7 **Introduction**
8 Contenu de l'ouvrage
9 Présentation des chapitres
11 Utilisation des programmes informatiques
11 Remarques sur la terminologie et les notations

13 **Chapitre 0 - Quelques éléments de réflexion**

13 **0.1 De l'inférence traditionnelle à l'inférence bayésienne**
14 0.1.1 La distribution d'échantillonnage
16 0.1.2 L'inférence statistique traditionnelle: le test de signification
19 0.1.3 L'inférence bayésienne
23 0.1.4 La méthode fiducio-bayésienne
25 0.1.5 L'inférence bayésienne permet d'évaluer de façon directe l'importance d'un effet
26 0.1.6 Comparaison de l'inférence fréquentiste et de l'inférence bayésienne
28 0.1.7 En Conclusion

29 **0.2 Quelques notions et résultats préliminaires**
29 0.2.1 Distribution d'échantillonnage dans une population
32 0.2.2 Test de typicité (inférence ensembliste): Les échantillons extrêmes
33 0.2.3 Test de signification - Procédures d'échantillonnage
35 0.2.4 Procédures bayésiennes
39 0.2.5 Quelques résultats de base pour le modèle normal

45 Chapitre 1 - Du t de Student aux procédures fiducio-bayésiennes

45 1.1 Recherche d'une conclusion d'effet notable: Inférence sur une moyenne
45 1.1.1 L'exemple de Student
46 1.1.2 Analyse descriptive
47 1.1.3 Analyse inductive: Techniques de base
52 1.1.4 Analyse inductive: Techniques complémentaires
53 1.1.5 Résultats généraux: Techniques de base
56 1.1.6 Résultats généraux: Techniques complémentaires
57 1.1.7 Lien intuitif avec le test de signification
57 1.1.8 Effet notable/effet négligeable: Vers des solutions générales

59 1.2 Recherche d'une conclusion d'effet notable: Comparaison de deux moyennes de groupes indépendants
59 1.2.1 Données pharmacologiques: Nouveau médicament/Placebo
62 1.2.2 Réinterprétation de l'intervalle de confiance fréquentiste
64 1.2.3 Réinterprétation du seuil observé du test de signification
65 1.2.4 Réinterprétation du seuil fixé du test de signification
67 1.2.5 Recherche de conclusion d'effet notable

68 1.3 Recherche d'une conclusion d'effet non notable
69 1.3.1 Reprenons l'exemple de Student
70 1.3.2 Recherche d'une conclusion d'effet non notable

72 1.4 Recherche d'une conclusion d'effet négligeable
72 1.4.1 Exemple de la validation d'un modèle additif
73 1.4.2 Analyse descriptive
74 1.4.3 Analyse inductive

77 Chapitre 2 - De l'ANOVA aux procédures fiducio-bayésiennes

77 2.1 Combinaisons linéaires de moyennes de groupes indépendants
77 2.1.1 Exemple de données anthropologiques: Les crânes de Rao
79 2.1.2 Tests de signification
82 2.1.3 Procédures fiducio-bayésiennes
82 2.1.4 Reconstitution du tableau ANOVA
83 2.1.5 Du tableau ANOVA à l'analyse fiducio-bayésienne
83 2.1.6 Du test de signification à l'analyse fiducio-bayésienne: Inférence sur δ
86 2.1.7 Du test de signification à l'analyse fiducio-bayésienne: Inférence sur δ_{cal}

87 2.1.8 Du seuil de signification à l'analyse fiducio-bayésienne

90 2.2 Plan en groupes indépendants avec deux facteurs de classification
90 2.2.1 Exemple d'une enquête psychométrique:
Milieu et Pédagogie de l'enseignement des mathématiques
93 2.2.2 Analyses descriptives
94 2.2.3 Analyses inférentielles
99 2.2.4 Définition de plusieurs régions d'importance de l'effet

101 Chapitre 3 - Prise en compte d'informations extérieures aux données

102 3.1 Choix de la distribution initiale: Différentes attitudes

104 3.2 Influence de la distribution initiale
104 3.2.1 Inférence relative à δ conditionnellement à σ
109 3.2.2 Inférence (marginale) relative à σ
115 3.2.3 Inférences (marginales) relatives à δ et δ_{cal}

117 3.3 Examiner la sensibilité des conclusions issues d'une expérience
117 3.3.1 Mise à l'épreuve d'une conclusion
118 3.3.2 Evaluer l'information supplémentaire nécessaire pour obtenir une conclusion

121 Chapitre 4 - Analyse spécifique: Plans S<G>*O

121 4.1 Introduction à l'analyse spécifique
121 4.1.1 L'approche du modèle linéaire général et ses difficultés
122 4.1.2 L'approche de l'analyse spécifique

123 4.2 Plans S<G>*O
123 4.2.1 Plans S<G>*O: Définition
124 4.2.2 Analyse spécifique des comparaisons dans un plan S<G>*O:
Modèles à effets fixés et modèles mixtes
125 4.2.3 Comparaisons de type produit sur G*O

125 4.3 Plan du chapitre et exemples de données
125 4 3.1 Plan du chapitre
126 4.3.2 Exemple de données

127 4.4 Comparaison à un degré de liberté sur les groupes
127 4.4.1 Protocole dérivé pertinent
128 4.4.2 Dérivations sur les occasions et sur les groupes

129 4.4.3 Analyse descriptive
130 4.4.4 Procédures d'inférence
131 4.4.5 Rapport F de l'analyse de variance et t de Student
133 4.4.6 Effectuer d'abord la dérivation sur les groupes

134 4.5 Comparaison à un degré de liberté sur les occasions
134 4.5.1 Analyse spécifique: Dérivation
136 4.5.2 Analyse descriptive
137 4.5.3 Inférence spécifique

137 4.6 Interaction à un degré de liberté entre les groupes et les occasions
137 4.6.1 Analyse spécifique: Dérivations
138 4.6.2 Analyse descriptive
139 4.6.3 Inférence spécifique

139 4.7 Vers un usage raisonnable de l'analyse spécifique
139 4.7.1 Comment justifier une analyse spécifique?
140 4.7.2 Choix méthodologiques liés à l'analyse spécifique
142 4.7.3 Quelques lignes directrices pour un usage raisonnable de l'analyse spécifique
143 4.7.4 Quelle stratégie adopter vis-à-vis des conditions de validité?

145 4.8 L'analyse des comparaisons qui ne sont pas de type produit

147 Chapitre 5 - Analyse spécifique: Illustrations

147 5.1 Exemples de plans à mesures répétées
147 5.1.1 Une expérience de temps de réaction
151 5.1.2 Comparaison Lin C
154 5.1.3 Comparaisons Cub C et Qua C

157 5.2 Plans non-équilibrés
158 5.2.1 Solution équipondérée
159 5.2.2 Solution pondérée par les effectifs
159 5.2.3 Solution orthogonale
164 5.2.4 Autre point de vue: Modèle sans interaction

165 5.3 Exemples de plans "cross-over"
165 5.3.1 Cross-over 2×2
169 5.3.2 Cross-over: Exemple avec deux traitements et trois périodes
174 5.3.3 Cross-over: Autres exemples de généralisations
176 5.3.4 Exemple: La conservation des morceaux de viande

178 5.4 Mémorisation de récits
178 5.4.1 Présentation générale et objectifs de l'expérience

179 5.4.2 Plan d'analyse
181 5.4.3 Principaux résultats

185 5.5 Apprentisssage perceptif incident
185 5.5.1 Présentation générale et objectifs de l'expérience
186 5.5.2 Plan d'analyse

191 5.6 Enquête pédagogique: Questionnaire mathématique
191 5.6.1 Description du plan d'enquête
192 5.6.2 Analyses spécifiques

193 Chapitre 6 - Comparer des moyennes sans supposer l'égalité des variances

193 6.1 Introduction aux procédures semi-bayésiennes
193 6.1.1 Le problème de Behrens-Fisher
194 6.1.2 Test conditionnel au rapport des écarts-types
196 6.1.3 Le test de signification semi-bayésien: Principe général
198 6.1.4 Mise en œuvre pratique
199 6.1.5 Quelques commentaires généraux
201 6.1.6 Du test de signification semi-fiducio-bayésien à l'inférence fiducio-bayésienne

202 6.2 Combinaisons linéaires de moyennes de groupes indépendants
202 6.2.1 Données psychométriques: Milieu et Pédagogie de l'enseignement des mathématiques
203 6.2.2 En supposant l'égalité des variances
204 6.2.3 En ne supposant pas l'égalité des variances

205 6.3 Un résultat plus significatif en ne supposant pas l'égalité des variances

207 6.4 Formules de calcul des approximations
207 6.4.1 Comparaison de deux moyennes de groupes indépendants d'effectifs égaux
208 6.4.2 Inférence sur une combinaison linéaire de moyennes de groupes indépendants d'effectifs quelconques

211 Chapitre 7 - Compléments sur les procédures fréquentistes

211 7.1 Illustrations
211 7.1.1 Effet notable
214 7.1.2 Effet non négligeable
216 7.1.3 Effet négligeable ("équivalence")
221 7.1.4 Effet calibré notable, non négligeable, négligeable

224 7.2 Quelques considérations formelles sur le test de négligeabilité
224 7.2.1 Première procédure
225 7.2.2 Seconde procédure
229 7.2.3 Remarque: Intervalle de confiance symétrique de Westlake

229 7.3 Recherche d'une conclusion d'effet non négligeable
230 7.3.1 Première procédure
230 7.3.2 Seconde procédure

233 Chapitre 8 - Déterminer les effectifs nécessaires

234 8.1 Effet calibré notable
234 8.1.1 Le problème
236 8.1.2 Simplification du problème: Paramètres fixés et q traité comme infini
241 8.1.3 Paramètres fixés et q quelconque
243 8.1.4 Un autre exemple mettant en jeu un effectif plus faible
244 8.1.5 Prendre en compte l'incertitude sur les paramètres

247 8.2 Effet calibré négligeable
247 8.2.1 Calcul pour la procédure fiducio-bayésienne
248 8.2.2 Illustration
250 8.2.3 Calcul pour la procédure fréquentiste "des deux seuils unilatéraux"

251 8.3 Effet (non calibré) notable
251 8.3.1 Solution générale
253 8.3.2 Illustrations

254 8.4 Effet (non calibré) négligeable
254 8.4.1 Calcul pour la procédure fiducio-bayésienne
255 8.4.2 Calcul pour la procédure fréquentiste "des deux seuils unilatéraux"
256 8.4.3 Illustrations

261 Chapitre 9 - Distributions utiles

262 9.1 Distributions de base
262 9.1.1 Distribution gamma
263 9.1.2 Distribution gamma-puissance
265 9.1.3 Distributions khi-deux (centrée) et dérivées
268 9.1.4 Distribution normale
269 9.1.5 Distribution bêta
271 9.1.6 Distribution F (centrée)

272 9.2 Distribution de Student généralisée
272 9.2.1 Caractérisations
273 9.2.2 Densité
274 9.2.3 Moments
275 9.2.4 Fonction de répartition

275 9.3 Distribution de Student non-centrée
275 9.3.1 Caractérisations
276 9.3.2 Propriétés
277 9.3.3 Densité
278 9.3.4 Moments
278 9.3.5 Fonction de répartition

279 9.4 Distribution lambda-prime
279 9.4.1 Caractérisations
279 9.4.2 Propriétés
280 9.4.3 Densité
281 9.4.4 Moments
281 9.4.5 Distribution t non-centrée et distribution lambda-prime
282 9.4.6 Fonction de répartition
283 9.4.7 Distribution lambda-prime et distribution d'échantillonnage des limites de confiance fréquentistes

285 9.5 Distribution K-prime
285 9.5.1 Caractérisations
286 9.5.2 Propriétés
286 9.5.3 Relation remarquable
287 9.5.4 Densité
288 9.5.5 Moments
288 9.5.6 Fonction de répartition

291 9.6 Divers résultats utiles
291 9.6.1 Un lemme sur les distributions χ^2 et F
291 9.6.2 Approximations d'une distribution de moyenne M et de variance V
292 9.6.3 Approximation d'une distribution par une distribution $cb\Gamma(a)$
292 9.6.4 Un résultat sur les distributions conditionnelles

295 Chapitre 10 - Solutions bayésiennes

295 10.1 Structure "normale - phi-deux": Solutions générales
295 10.1.1 Définition
296 10.1.2 Inférence relative à δ conditionnellement à σ^2 et à d

298 10.1.3 Inférence relative à σ^2 conditionnellement à s2
300 10.1.4 Inférence conjointe relative à δ et σ^2

304 10.2 Structure "normale - phi-deux": Procédures pour les paramètres
304 10.2.1 Test de signification fréquentiste
305 10.2.2 Test de signification semi-bayésien
306 10.2.3 Procédures fiducio-bayésiennes
307 10.2.4 Liens entre l'inférence fiducio-bayésienne et le test de signification
307 10.2.5 Procédures bayésiennes (distribution initiale conjuguée naturelle)

308 10.3 Structure "normale - phi-deux": Procédures prédictives
308 10.3.1 Procédures bayésiennes prédictives
308 10.3.2 Distributions prédictives marginales
309 10.3.3 Distribution prédictive relative à d conditionnellement à s^2
310 10.3.4 Distributions prédictives relatives à l'effet calibré
310 10.3.5 Distribution prédictive relative à la statistique de test
311 10.3.6 Distributions prédictives relatives aux limites de crédibilité
311 10.3.7 Remarque: quand $b^2 \rightarrow 0$ et $q \rightarrow +\infty$

312 10.4 Structure "normale - (phi-deux)s"
312 10.4.1 Définition
312 10.4.2 Distribution fiducio-bayésienne pour le paramètre ε^2
313 10.4.3 Test semi-fiducio-bayésien
314 10.4.4 Procédures fiducio-bayésiennes

315 10.5 Compléments: Le paradoxe de marginalisation

317 Annexe: Installation et notice d'utilisation des programmes Windows

319 Installation des programmes

319 Configuration nécessaire

319 Installation
319 Faites une copie de sauvegarde
319 Procédure d'installation
320 Exécutez la procédure d'installation
320 Anomalies de fonctionnement
320 En cas de problème

321 **LeBayésien**

322 Généralités sur l'utilisation
322 Le menu principal et la barre de boutons
323 Les fichiers
324 Utilisation de l'aide
325 Gestion des fichiers: La fenêtre de dialogue "fichiers"

327 Les principales fonctions
327 La fonction données
329 La fonction mode d'entrée
331 La fonction énoncés
331 La fonction options
332 La fonction quitter
332 La fonction ?
332 Le bouton exemples

333 Imprimer une distribution
333 Procédure
337 Mettre des courbes à la même échelle
337 Enregistrement/lecture des courbes

339 **LesDistributions**
340 Prise en mains

343 **Références bibliographiques**

343 Références citées dans l'ouvrage

348 Références complémentaires

351 **Table des matières**

361 **Index**

Dans le texte les renvois aux chapitres ou sections de chapitre sont indiqués par le signe ↳

INDEX

Analyse de variance 121-122-126-131
Analyse spécifique 122,139
Assumptions 121
Asymétrie (coefficient d') 262

Bayes 28
  approche bayésienne 19
  formule de 21-37
Bayésiennes: voir procédures
Behrens-Fisher (problème) 193
Bioéquivalence 217

Carrés-moyens 78-131
Combinaison linéaire de moyennes 41
  inférence sur une... 74-77-130
Comparaisons de type non produit 146
Comparaisons de type produit 124
Composantes de norme 129
Conditions de validité 121-140-193
Constante a 128-131
Constante b 62-128-131
Contraste 128

Degrés de liberté q 40-53-60-130
Dérivation 127-128
  par moyennage, par contraste 128
Distribution
  approximations par les moments 291
  bêta 269
  exponentielle 264
  F 9-11
  gamma 262
  gamma-puissance 263
  K-prime 1-25
  khi, khi-inverse 266
  khi-deux 40-265
  khi-deux inverse 266
  lambda-prime 54-279
  normale 15-39-268
  phi, phi-deux 40-266
  phi-deux inverse 266
  phi-inverse 54-266
  t de Student (élémentaire) 40, 279
  t de Student généralisée 53-273
  t de Student non-centrée 40-221-235-275
  uniforme 270
  Weibull 264
Distribution conguguée naturelle 101-297
Distribution d'échantillonnage 14
  dans une population 29
  sous le modèle normal 40
  des limites de confiance fréquentistes 283
Distribution finale (*a posteriori*) 11-19-21
Distribution initiale (*a priori*) 11-19-21
  choix (attitudes) 102
  enthousiaste/sceptique 103-109-116
  influence de la distribution initiale 20-104
  subjective 21
Distribution initiale non-informative 22
  pour le modèle normal 43-296-298-316
Distribution parente 29
Distribution prédictive 21-56-244
Données
  apprentissage perceptif incident 185
  conservation des morceaux de viande 176
  crânes de Rao 77

cross-over 2 traitements 3 périodes 171
étude de bioéquivalence 217
exemple plus significatif
sans égalité des variances 205
médicament/*placebo* 59-193
mémorisation de récits 178
milieu et pédagogie (334 élèves) 89
milieu et pédagogie (33 élèves) 153-208
modèle additif 72
plan S<G2>*A2*B2 126
plan S<G3>*A2*B2 141
questionnaire mathématique 191
rémission de la douleur 234
Student 45-104-109-115-117-212-214-221-243
temps de réaction 147
Droite de régression 148-189

Ecart-type σ
inférence fiducio-bayésienne sur σ 50-54
Ecart-type-corrigé 40
intragroupe moyen s 61-74-130
Echantillonnage, voir distribution, modèle
Echelle e 53-54-62-83-87-306
Effet 14-48-81-122
observé d 147
inférence fiducio-bayésienne sur δ 50-53
Effet calibré 47-82-86-90
inférence fiducio-bayésienne sur δ_{cal} 51-54
tests négligeabilité/notabilité 221
Effet de la comparaison 60
Effet négligeable/notable 26-58
tests négligeabilité/notabilité 212-214-216
Effet résiduel (*carry-over*) 167
Effets individuels 46-53
Egalité des variances 15-25-41-81-193
Enthousiaste: voir distribution initiale
Exhaustivité 41

Facteur(s) 123
aléatoires/systématiques 123-124-187-191
croisés 124
emboîté dans un autre facteur 123
inter-sujets, intra-sujets 123
f_g (effectifs des groupes) 62
Fiduciaire 23-201-297-299
Fiducio-bayésien 23
voir distribution, écart-type, effet, effet calibré, procédures
Forme linéaire 128

Garantie 26
Groupes 123

Hypothèse nulle 16-18
composée 211

Importance des effets 18-25
voir effet négligeable/notable
Imprécision/précision expérimentale 15-43-64-66-75
Inférence bayésienne 19
Inférence ensembliste (test de typicité) 32
Inférence fréquentiste 7-16
et bayésienne 27
Inférence spécifique 122
et inférence générale 139
Intervalle de bioéquivalence 217
Intervalle de confiance 27-62-65-283
voir négligeable/notable
réinterprétation fiducio-bayésienne 27-62
Intervalle de crédibilité (bayésien) 27

Laplace 28

Marginalisation (paradoxe) 315
Modèle d'échantillonnage 15-21
Modèle normal 39-47-77-95-137-142
Modèle sans interaction 164
Modèle spécifique 122-140
Modèles à effets fixes et mixtes 124
Moments 261
Moyenne(s) 39
comparaison de deux moyennes 59
équipondérée/pondérée 73-78-91-158

inférence sur une moyenne 45

Négligeable/notable: voir effet, test
Non-informative
voir distribution initiale
conception 22
Norme d'une forme linéaire 129
Notations (remarques sur) 12-261

Occasions 123
Orthogonale (décomposition) 78-95-159
Orthogonalité de deux contrastes 160

PAC
Programme d'Analyse des Comparaisons 8
utilisation des fichiers 330
Paramètres 14
Plans d'analyse 121
à mesures répétées 123-147
à blocs incomplets 179-191
cross-over 165
en groupes indépendants 77-83-90-123
non équilibrés 157
orthogonaux 161
Sujets<Groupes>*Occasions 124
S<G>, S*O, S<G>*O 123-124
Poids, pondération 126-128
Précision/imprécision expérimentale 15-43-64-66-75
Prédictive(s):
voir distribution, probabilité, procédures
Probabilité
d'échantillonnage 16
initiale/finale 21-38
prédictive 21-38-118-235
prédictive et puissance 236
Procédures (solutions techniques)
bayésiennes 302-307
fiducio-bayésiennes 306-312
prédictives 308
sans égalité des variances 207
semi-bayésiennes 196-305-313
test fréquentiste 304

Programmes informatiques 8-11-317
LeBayésien 47-321
LesDistributions 339
LesEchantillons 42-317
LesEffectifs 237-317
PAC (utilisation des fichiers) 330
Protocole dérivé pertinent 127
Puissance 236

Rapport F et t de Student 81-131
Régression polynomiale 148

Sceptique: voir distribution initiale
Semi-bayésien(nes):
voir procédures, test de signification
Seuil de signification 18
fixé 32
réinterprétation fiducio-bayésienne 24-64-307
seuil-repère 95
unilatéral/bilatéral 16-64-65
Seuils unilatéraux (procédure des deux) 216
S<G>, S*O, S<G>*O 123-124
Significatif/non-significatif 18-196
Somme-des-carrés 78-83
Spécifique: voir analyse, inférence, modèle

Test
de négligeabilité 216-224
de non négligeabilité 214-230
de notabilité 212
Test de signification 16
voir procédures
de typicité (ensembliste) 32
rapport F 81-131
semi-bayésien 196
t de Student 57
unilatéral/bilatéral 16-64-65

Variabilité intra/inter-sujets 15-52
Variance-corrigée 40
intragroupe moyenne s^2 41-60-78
Variances (égalité des) 15-25-41-81-193

v_g (dérivation sur les groupes) 62-128-135
Vraisemblance (fonction de) 21-33

w_o (dérivation sur les occasions) 128-134

Imprimerie LAUNAY
45, rue Linné 75005 Paris (France)
Dépôt légal : juin 1996 - Numéro 58
ISBN 2-906711-10-1

Cet ouvrage propose à l'utilisateur de l'analyse de variance une approche pratique, réaliste et constructive de l'inférence statistique, qui lui apporte un regard nouveau sur ses données. Les procédures bayésiennes standard sont aussi objectives et simples à utiliser que les procédures traditionnelles (tests de signification t ou F familiers, intervalles de confiance). Intégrant ces dernières, elles en éclairent les difficultés et les insuffisances, et renouvellement en profondeur la méthodologie du traitement statistique des données expérimentales.

Des réponses concrètes sont apportées à des questions essentielles dans la pratique.

- Comment interpréter correctement les procédures d'inférence statistique ?
- Comment juger de l'importance d'un effet
- "significativité clinique (psychologique ...)" et "significativité statistique" ?
- peut-on "prouver l'hypothèse nulle" d'absence d'effet quand c'est l'hypothèse de recherche ?
- Comment apprécier "ce que les données ont à dire" et examiner dans quelle mesure des informations supplémentaires remettraient en cause les conclusions ?
- Comment analyser les plans expérimentaux complexes largement utilisés, tels que les dispositifs avec mesures répétées ou croisés (cross-over) ?
- Comment comparer des moyennes sans supposer l'égalité des variances ?
- Comment déterminer les effectifs nécessaires pour "avoir de bonnes chances" d'obtenir une conclusion donnée ?

La présentation des méthodes est effectuée à partir d'exemples réels. Les programmes informatiques sous Windows, didactiques et conviviaux, permettent la mise en oeuvre interactive très simple de toutes les procédures au fur et à mesure de leur exposé. L'ouvrage présente ainsi une conception originale, qui en fait pour le plus grand nombre de lecteurs un outil précieux, utilisable aussi bien pour une initiation au traitement des données expérimentales que pour des applications sophistiquées.

Bruno LECOUTRE, Docteur en psychologie et en Mathématiques, est directeur de Recherche au C.N.R.S. et consultant dans l'industrie pharmaceutique. Il est (avec Jacques POLTEVINEAU) l'auteur du logiciel PAC édité par le CISIA.

CISIA•CERESTA Editeur
1 Avenue Herbillon 94160 Saint-Mandé (France)
ISBN 2-906711-10-1