Journal of Statistical Planning and Inference 79 (1999) 93–105

journal of statistical planning and inference

Two useful distributions for Bayesian predictive procedures under normal models

Bruno Lecoutre *

UPRESA 6085, Analyse et Modèles Stochastiques, C.N.R.S et Université de Rouen, Mathématiques Site Colbert, 76821 Mont-Saint-Aignan Cedex, France

Received 6 October 1997; accepted 2 November 1998

Abstract

The *K-prime* and *K-square* distributions, involved in the Bayesian predictive distributions of standard t and F tests are investigated. They generalize the classical *noncentral* t and *noncentral* F distributions and can receive different characterizations. Their moments and their probability density and distribution functions are made explicit. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Predictive distributions; Normal models; t tests; F ratios; Confidence intervals

1. Introduction

In recent years many authors have stressed the interest of the Bayesian predictive approach for monitoring experiments. Some references are Spiegelhalter and Freedman (1986), Choi and Pepple (1989), Berry (1991), Geisser (1993), Spiegelhalter et al. (1994), Lecoutre et al. (1995), Adcock (1997), Joseph and Bélisle (1997) and Hung et al. (1997). A question of interest has been to determine the predictive distributions of test statistics and p-values or of the limits of confidence intervals. Under standard normal models, assuming a conjugate prior, this leads to define two distributions that we name the *K-prime* and the *K-square* distributions. These distributions can respectively be defined as mixtures of the classical *noncentral* t and *noncentral* F distributions involved in the power of standard tests, or alternatively as mixtures of distributions that are their Bayesian counterparts. Several equivalent characterizations are introduced and point out the links between the different distributions. From these characterizations, the moments and probability distribution functions can be made explicit. The cumulative

* E-mail: bruno.lecoutre@univ-rouen.fr

PII: S0378-3758(98)00231-6

Table 1
Characterizations of univariate distributions

Distribution	Mixture	Particular cases
$t'_r(a)$	$N(\frac{a}{y}, \frac{1}{y^2}) \wedge_{y^2} \frac{\chi^2_r}{r}$	$t'_\infty(a) = N(a,1)$
$\Lambda'_q(a,1)$	$N(ay,1) \wedge_{y^2} \frac{\chi^2_q}{q}$	$\Lambda'_\infty(a,1) = N(a,1)$
$t_p(a,1)$	$N(a, \frac{1}{y^2}) \wedge_{y^2} \frac{\chi^2_p}{p}$	$t_\infty(a,1) = N(a,1)$
$K'_{q,r}(a,1)$	$t'_r(ay) \wedge_{y^2} \frac{\chi^2_q}{q}$	$K'_{\infty,r}(a,1) = t'_r(a)$
$K'_{q,r}(a,1)$	$\Lambda'_q(\frac{a}{y}, \frac{1}{y^2}) \wedge_{y^2} \frac{\chi^2_r}{r}$	$K'_{q,\infty}(a,1) = \Lambda'_q(a,1)$
$K'_{q,r}(a,1)$	$t_{q+r}(ay, \frac{qy^2+r}{q+r}) \wedge_{y^2} F_{q,r}$	$K'_{\infty,\infty}(a,1) = N(a,1)$
$F'_{m,r}(a^2)$	$\frac{1}{y^2}\frac{\chi^2_m(a^2)}{m} \wedge_{y^2} \frac{\chi^2_r}{r}$	$F'_{m,\infty}(a^2) = \frac{\chi^2_m(a^2)}{m}$
$\Lambda^2_{m,q}(a^2)$	$\frac{\chi^2_m(a^2y^2)}{m} \wedge_{y^2} \frac{\chi^2_q}{q}$	$\Lambda^2_{m,\infty}(a^2) = \frac{\chi^2_m(a^2)}{m}$
$\psi^2_{m,p}(a^2)$	$\frac{1}{y^2}\frac{\chi^2_m(a^2y^2)}{m} \wedge_{y^2} \frac{\chi^2_p}{p}$	$\psi^2_{m,\infty}(a^2) = \frac{\chi^2_m(a^2)}{m}$
$K^2_{m,q,r}(a^2)$	$F'_{m,r}(a^2y^2) \wedge_{y^2} \frac{\chi^2_q}{q}$	$K^2_{m,\infty,r}(a^2) = F'_{m,r}(a^2)$
$K^2_{m,q,r}(a^2)$	$\frac{1}{y^2}\Lambda^2_{m,q}(a^2) \wedge_{y^2} \frac{\chi^2_r}{r}$	$K^2_{m,q,\infty}(a^2) = \Lambda^2_{m,q}(a^2)$
$K^2_{m,q,r}(a^2)$	$\frac{qy^2+r}{q+r}\psi^2_{m,q+r}(\frac{q+r}{qy^2+r}a^2y^2) \wedge_{y^2} F_{q,r}$	$K^2_{m,\infty,\infty}(a^2) = \frac{\chi^2_m(a^2)}{m}$

distribution functions can be expressed in terms of infinite series of multiples of incomplete beta function ratios, leading to simple and efficient algorithms for numerical computations. We assume here that the probability density functions and moments of the *chi-square* and *noncentral t* and *F* distributions are known. The necessary results can be found in Rao (1965) and Johnson et al. (1995).

The main characterizations as mixtures of the univariate distributions involved in this paper are summarized in Table 1.[1]

[1] A Windows™ interactive computer program, "LesDistributions" (B. Lecoutre and J. Poitevineau), that calculates probability statements for all these distributions is available on the Internet at address: `http://epeire.univ-rouen.fr/labos/eris/pac.html` or upon request to the author.

2. The K-prime distribution

2.1. First characterization

The multivariate *K-prime* distribution has been introduced in Lecoutre (1984) with the following characterization. If $\boldsymbol{x}|y^2$ has the multivariate *generalized* t distribution

$$\boldsymbol{x}|y^2 \sim \boldsymbol{t}_{m;q+r}\left(y\boldsymbol{a}, \frac{qy^2+r}{q+r}b^2\boldsymbol{I}_m\right) \quad (y>0, q>0, r>0, m\geqslant 1)$$

and y^2 has the central F distribution $y^2 \sim F_{q,r}$, then $\boldsymbol{x}$ has the m-variate *K-prime* distribution, with q and r degrees of freedom and with parameters $\boldsymbol{a}$ and b^2

$$\boldsymbol{x} \sim \boldsymbol{K}'_{m;q,r}(\boldsymbol{a}, b^2\boldsymbol{I}_m)$$

From this characterization, b appears as a scale factor

$$\boldsymbol{K}'_{m;q,r}(\boldsymbol{a}, b^2\boldsymbol{I}_m) = b\boldsymbol{K}'_{m;q,r}\left(\frac{1}{b}\boldsymbol{a}, \boldsymbol{I}_m\right)$$

hence it will be sufficient to study the case $b^2=1$.

In what follows we consider only the univariate distribution ($m=1$), denoted by $K'_{q,r}(a,b^2)$. The univariate *K-prime* distribution includes as particular cases:

- for $a=0$, the usual t distribution

  $$K'_{q,r}(0,1) = t_r,$$

- for $q=\infty$, the *noncentral* t distribution

  $$K'_{\infty,r}(a,1) = t'_r(a),$$

- for $r=\infty$, the distribution that we name *lambda-prime*, involved in the sampling distribution of confidence limits for μ and in the posterior Bayesian distribution of the ratio μ/σ, under the usual normal model with mean μ and standard deviation σ, using a conjugate prior

  $$K'_{q,\infty}(a,1) = \Lambda'_q(a,1).$$

Note that this last distribution has been used in the fiducial framework by Fisher (1990) (pp. 126–127).

2.2. Equivalent characterizations for the univariate distribution

If one of the three following sets of conditions is satisfied (a, $q>0$ and $r>0$ are real numbers)

(1) $x|y^2 \sim \dfrac{1}{y}\Lambda'_q(a,1)$ $(y>0)$ and $y^2 \sim \dfrac{\chi^2_r}{r}$,

(2) $x = \dfrac{w}{y}$, w and y being independent, with $w \sim \Lambda'_q(a,1)$ and $y \sim \dfrac{\chi^2_r}{r}$,

(3) $x|y^2 \sim t'_r(ay)$ $(y>0)$ and $y^2 \sim \dfrac{\chi^2_q}{q}$

then $x \sim K'_{q,r}(a,1)$.

Proof. The equivalence of these characterizations can be stated as follows. Consider three real variables x, g^2 and h^2, with g^2 and h^2 independent, such that

$$x|g^2,h^2 \sim N\left(a\frac{g}{h},h^2\right), g^2 \sim \frac{\chi^2_q}{q}, h^2 \sim \frac{\chi^2_r}{r} \quad (g>0, h>0)$$

then $x \sim K'_{q,r}(a,1)$ and this result can be obtained from each of the four characterizations.

First these conditions can be rewritten

$$x|u^2,v^2 \sim N\left(av, \frac{qv^2+r}{u^2}\right),$$

where, as a classical result $u^2 = qg^2 + rh^2$ and $v = g/h$ have independent distributions, respectively $u^2 \sim \chi^2_{q+r}$ and $v^2 \sim F_{q,r}$. Then the main characterization applies, since

$$x|v^2 \sim t_{q+r}\left(av, \frac{qv^2+r}{q+r}\right).$$

For the three equivalent characterizations, the result can be respectively deduced from

$$x|h^2 \sim \frac{1}{h}\Lambda'_q(a,1),$$

$$w = \frac{x}{h}|h^2 \sim \Lambda'_q(a,1) \text{ hence } w \sim \Lambda'_q(a,1),$$

$$x|g^2 \sim t'_r(ag),$$

where these distributions are themselves deduced from the main characterization for the particular cases $r=\infty$ and $q=\infty$.

Note that the $\Lambda'_q(a,1)$ distribution can again be characterized as the distribution of $x+ay$, where x and $y>0$ are independent, with $x \sim N(0,1)$ and $y^2 \sim \chi^2_q/q$, hence $x+ay|y \sim N(ay,1)$. This follows for instance as a particular case from the first characterization (for $r=\infty$).

2.3. Properties of the cumulative distribution function

The *cdf* has the following properties:

$$\Pr(K'_{q,r}(-a,1) < -\ell) = \Pr(K'_{q,r}(\ell,1) > a),$$

$$\Pr(K'_{q,r}(a,1) < 0) = \Pr(\Lambda'_q(a,1) < 0) = \Pr(t_q > a),$$

$$\Pr(K'_{r,q}(a,1) < \ell) = \Pr(K'_{q,r}(\ell,1) > a).$$

Proof. The first property is an immediate consequence of the characterizations. The second property comes from the fact that, if $x \sim K'_{q,r}(a,1)$, then under the conditions of the first equivalent characterization, $x/y|y^2 \sim x/y \sim \Lambda'_q(a,1)$. Moreover, since y is a positive variable $\Pr(x<0)=\Pr(x/y<0)$. The third property expresses a remarkable relationship between the *K-prime* distribution with q and r degrees of freedom and the *K-prime* distribution with r and q degrees of freedom. In the particular case $r=\infty$, it gives $\Pr(t'_q(a)<\ell)=\Pr(\Lambda'_q(\ell,1)>a)$. Considering three independent variables x, $y>0$ and $z>0$, such that $x\sim N(0,1)$, $y^2\sim\chi^2_r/r$ and $z^2\sim\chi^2_q/q$, we get

$$x+ay\sim\Lambda'_r(a,1) \text{ and } x-\ell z\sim\Lambda'_r(-\ell,1)$$

hence, from the second of the equivalent characterizations,

$$\frac{x+ay}{z}\sim K'_{r,q}(a,1) \quad \text{and} \quad \frac{x-\ell z}{y}\sim K'_{q,r}(-\ell,1).$$

The result is then deduced from the equivalence of the two conditions $(x+ay)/z<\ell$ and $(x-\ell z)/y<-a$, since $\Pr(K'_{q,r}(-\ell,1)<-a)=\Pr(K'_{q,r}(\ell,1)>a)$.

2.4. Moments

The moments of the $K'_{q,r}(a,1)$ distribution can be directly deduced from its characterizations. In particular,

$$E(x)=ak,$$

$$E(x^2)=\frac{r}{r-2}(a^2+1),$$

$$\mathrm{Var}(x)=\left(\frac{r}{r-2}-k^2\right)a^2+\frac{r}{r-2},$$

$$E(x^3)=\frac{r}{r-3}\left(\frac{q+1}{q}a^3+3a\right)k,$$

$$E(x^4)=\frac{r^2}{(r-2)(r-4)}\left(\frac{q+2}{q}a^4+6a^2+3\right),$$

where

$$k=\sqrt{\frac{r}{q}}\frac{\Gamma(\frac{q+1}{2})\Gamma(\frac{r-1}{2})}{\Gamma(\frac{q}{2})\Gamma(\frac{r}{2})}$$

if respectively $r>1$, $r>2$, $r>2$, $r>3$ and $r>4$.

For instance, from the moments of the conditional distribution $x|y^2\sim t'_r(ay)$, we get

$$E(x)=E(E(x|y^2))=E(ak'y)=ak'E(y),$$

where

$$k'=\sqrt{\frac{r}{2}}\frac{\Gamma(\frac{r-1}{2})}{\Gamma(\frac{r}{2})},$$

$$E(x^2)=E(E(x^2|y^2))=\frac{r}{r-2}(a^2E(y^2)+1)$$

and the results are deduced from the moments of the marginal distribution $y^2\sim\chi^2_q/q$.

2.5. Probability density function

If $x \sim K'_{q,r}(a,1)$ its *pdf* is defined for q and r positive real numbers and can be expressed as

$$p(x) = \frac{1}{\sqrt{\pi r}} \frac{1}{\Gamma(\frac{q}{2})\Gamma(\frac{r}{2})} \left(\frac{q}{q+a^2}\right)^{\frac{q}{2}} \left(\frac{r}{r+x^2}\right)^{\frac{r+1}{2}}$$

$$\sum_{j=0}^{+\infty} \frac{1}{j!} \Gamma\left(\frac{q+j}{2}\right) \Gamma\left(\frac{r+j+1}{2}\right) \left(\frac{4}{(q+a^2)(r+x^2)}\right)^{\frac{j}{2}} (ax)^j.$$

Proof. From the characterization as a mixture of *noncentral t* distributions, we get $p(x) = \int p(x|y^2)p(y^2)\,\mathrm{d}y^2$, where

$$p(x|y^2) = \frac{1}{\sqrt{r}} \frac{1}{\Gamma(\frac{r}{2})} \exp\left(-\frac{a^2}{2}y^2\right) \left(\frac{r}{r+x^2}\right)^{\frac{r+1}{2}}$$

$$\times \sum_{j=0}^{+\infty} \frac{1}{j!} \Gamma\left(\frac{r+j+1}{2}\right) (ax)^j \left(\frac{2}{r+x^2}\right)^{\frac{j}{2}} (y^2)^{\frac{j}{2}},$$

$$p(y^2) = \frac{1}{\Gamma(\frac{q}{2})} \left(\frac{q}{2}\right)^{\frac{q}{2}} (y^2)^{\frac{q}{2}-1} \exp\left(-\frac{q}{2}y^2\right).$$

Inverting the summation and the integration, the result is deduced from the gamma integral (see for instance p. 144 of Box and Tiao (1973)),

$$\int_0^{+\infty} (y^2)^{\frac{q+j}{2}-1} \exp\left(-\frac{q+a^2}{2}y^2\right) \mathrm{d}y^2 = \Gamma\left(\frac{q+j}{2}\right) \left(\frac{q+a^2}{2}\right)^{-\frac{q+j}{2}}.$$

2.6. Cumulative distribution function

1. If $a = 0$

$$\Pr(K'_{q,r}(0,1) < \ell) = \Pr(t_r < \ell).$$

2. If $a > 0$ and $\ell < 0$ the *cdf* of the $K'_{q,r}(a,1)$ distribution can be expressed in terms of an infinite series of multiples of incomplete beta function ratios

$$\Pr(K'_{q,r}(a,1) < \ell) = \sum_{j=0}^{+\infty} (-1)^j e_j I_{\ell^2/(r+\ell^2)}\left(\frac{j+1}{2}, \frac{r}{2}\right),$$

where

$$e_j = \frac{1}{2} \frac{\Gamma(\frac{q+j}{2})}{\Gamma(\frac{q}{2})\Gamma(1+\frac{j}{2})} \left(\frac{q}{q+a^2}\right)^{\frac{q}{2}} \left(\frac{a^2}{q+a^2}\right)^{\frac{j}{2}} \quad \text{and} \quad \sum_{j=0}^{+\infty} e_j = 1.$$

This series can be evaluated from recurrence relationships, both for the coefficients e_j and for the incomplete beta function ratios. The algorithm developed for the *psi-square* distribution in Lecoutre et al. (1992) can be extended to the present situation.

Proof. From the *pdf*, we get

$$\Pr(K'_{q,r}(a,1) < \ell) = \sqrt{\pi}\frac{1}{\Gamma(\frac{q}{2})\Gamma(\frac{r}{2})}\left(\frac{q}{q+a^2}\right)^{\frac{q}{2}}\sum_{j=0}^{+\infty}\frac{1}{j!}\Gamma\left(\frac{q+j}{2}\right)\Gamma\left(\frac{r+j+1}{2}\right)$$
$$\times r^{\frac{j+1}{2}}2^j\left(\frac{a^2}{q+a^2}\right)^{\frac{j}{2}}(-1)^j\int_0^\ell x^j\left(1+\frac{x^2}{r}\right)^{-\frac{r+j+1}{2}}\mathrm{d}x,$$

where the integral

$$\int_0^\ell x^j\left(1+\frac{x^2}{r}\right)^{-\frac{r+j+1}{2}}\mathrm{d}x = \frac{1}{2}\frac{\Gamma(\frac{1+j}{2})\Gamma(\frac{r}{2})}{\Gamma(\frac{r+j+1}{2})}r^{-\frac{j+1}{2}}I_{\ell^2/(r+\ell^2)}\left(\frac{j+1}{2},\frac{r}{2}\right)$$

can be deduced from the *pdf* of the *central F* distribution. Then the result is obtained using the duplication formula $2^j\Gamma(\frac{1+j}{2})\Gamma(1+\frac{j}{2}) = \sqrt{\pi}j!$

3. If $a > 0$ and $\ell > 0$ we have

$$\Pr(K'_{q,r}(a,1) < \ell) = \Pr(K'_{q,r}(a,1) < 0) + \Pr(0 < K'_{q,r}(a,1) < \ell)$$
$$= \Pr(t_q > a) + \sum_{j=0}^{+\infty} e_j I_{\ell^2/(r+\ell^2)}\left(\frac{j+1}{2},\frac{r}{2}\right).$$

The first term is given by the *cdf* of the t distribution, and the second term is obtained as in the case $\ell < 0$.

4. If $a < 0$ we can use

$$\Pr(K'_{q,r}(a,1) < \ell) = \Pr(K'_{q,r}(-a,1) > -\ell) = 1 - \Pr(K'_{q,r}(-a,1) < -\ell).$$

3. The K-square distribution

3.1. Characterization

The *K-square* distribution has been introduced in Lecoutre (1984) with the following characterization. If $\boldsymbol{x}$ has the multivariate *K-prime* distribution $\boldsymbol{K}'_{m;q,r}(\boldsymbol{a}, \boldsymbol{I}_m)$, then $\boldsymbol{x}'\boldsymbol{x}/m$ has the *K-square* distribution with m, q and r degrees of freedom and parameter $a^2 = \boldsymbol{a}'\boldsymbol{a}$

$$\frac{\boldsymbol{x}'\boldsymbol{x}}{m} \sim K^2_{m,q,r}(a^2).$$

The *K-square* distribution includes as particular cases

- for $a = 0$, the usual F distribution

$$K^2_{m,q,r}(0) = F_{m,r},$$

- for $q = \infty$, the *noncentral F* distribution

$$K^2_{m,\infty,r}(a^2) = F'_{m,r}(a^2),$$

- for $r=\infty$, the distribution that we name *lambda-square*, involved in Bayesian analysis of variance procedures

$$K^2_{m,q,\infty}(a^2)=\Lambda^2_{m,q}(a^2).$$

This last distribution has been used by Geisser (1965) and Rouanet and Lecoutre (1983). It has been considered by Schervish (1992, 1995) under the name of *alternate chi-square* distribution (with a different scale factor).

3.2. Equivalent characterizations

If one of the four following sets of conditions is satisfied (a, $q>0$ and $r>0$ are real numbers)

$$\text{(1)}\quad x^2|y^2 \sim \frac{qy^2+r}{q+r}\psi^2_{m,q+r}\left(\frac{q+r}{qy^2+r}a^2y^2\right) \text{ and } y^2\sim F_{q,r},$$

$$\text{(2)}\quad x^2|y^2 \sim \frac{1}{y^2}\Lambda^2_{m,q}(a^2) \text{ and } y^2\sim\frac{\chi^2_r}{r},$$

$$\text{(3)}\quad x^2=\frac{w^2}{y^2}, w^2 \text{ and } y^2 \text{ being independent, with } w^2\sim\Lambda^2_{m,q}(a^2) \text{ and } y^2\sim\frac{\chi^2_r}{r},$$

$$\text{(4)}\quad x^2|y^2\sim F'_{m,r}(a^2y^2) \text{ and } y^2\sim\frac{\chi^2_q}{q}$$

then $x^2\sim K^2_{m,q,r}(a^2)$.

The *psi-square* distribution has been introduced in Lecoutre (1981); see also Rouanet and Lecoutre (1983) and Lecoutre (1985). It has been considered by Schervish (1992, 1995) under the name of *alternate F* distribution. If $\boldsymbol{x}$ has the multivariate *generalized t* distribution $\boldsymbol{t}_{m;q}(\boldsymbol{a},\boldsymbol{I}_m)$, then $\boldsymbol{x}'\boldsymbol{x}/m$ has the *psi-square* distribution with m and q degrees of freedom and parameter $a^2=\boldsymbol{a}'\boldsymbol{a}$: $\boldsymbol{x}'\boldsymbol{x}/m\sim\psi^2_{m,q}(a^2)$ with the particular case $\psi^2_{m,q}(0)=F_{m,q}$.

Proof. The equivalence of these characterizations can be stated as follows. Considering three variables x, g^2 and h^2, with g^2 and h^2 independent, such that

$$\boldsymbol{x}|g^2,\qquad h^2\sim\boldsymbol{N}_m\left(\frac{g}{h}\boldsymbol{a},\frac{1}{h^2}\boldsymbol{I}_m\right),\qquad g^2\sim\frac{\chi^2_q}{q},\qquad h^2\sim\frac{\chi^2_r}{r}\quad(g>0,h>0)$$

then $\boldsymbol{x}'\boldsymbol{x}/m\sim K^2_{m,q,r}(a^2)$ and this result can be obtained from each of the four characterizations.

Generalizing the demonstration given for the *K-prime* distribution, we get

$$\boldsymbol{x}|u^2,v^2\sim\boldsymbol{N}_m\left(v\boldsymbol{a},\frac{qv^2+r}{u^2}\boldsymbol{I}_m\right) \text{ where } u^2=qg^2+rh^2 \text{ and } v=\frac{g}{h}$$

hence for the main characterization

$$\boldsymbol{x}|v^2 \sim \boldsymbol{t}_{m;q+r}\left(v\boldsymbol{a}, \frac{qv^2+r}{q+r}\boldsymbol{I}_m\right) \quad \text{and} \quad \boldsymbol{x} \sim \boldsymbol{K}'_{m,q,r}(\boldsymbol{a}, \boldsymbol{I}_m).$$

For the other characterizations, the result can be respectively deduced from

$$\frac{\boldsymbol{x}'\boldsymbol{x}}{m}|v^2 \sim \frac{qv^2+r}{q+r}\psi^2_{m,q}\left(\frac{q+r}{qv^2+r}\boldsymbol{a}'\boldsymbol{a}v^2\right),$$

$$\frac{\boldsymbol{x}'\boldsymbol{x}}{m}|h^2 \sim \frac{1}{h^2}\Lambda^2_{m,q}(\boldsymbol{a}'\boldsymbol{a}),$$

$$w^2 = \frac{\boldsymbol{x}'\boldsymbol{x}}{m}h^2|h^2 \sim \Lambda^2_{m,q}(\boldsymbol{a}'\boldsymbol{a}) \text{ hence } w^2 \sim \Lambda^2_{m,q}(\boldsymbol{a}'\boldsymbol{a}),$$

$$\frac{\boldsymbol{x}'\boldsymbol{x}}{m}|g^2 \sim F'_{m,q}(\boldsymbol{a}'\boldsymbol{a}g^2),$$

where these distributions are themselves deduced from the main characterization for the particular cases $r=\infty$ and $q=\infty$.

3.3. Moments

If $x^2 \sim K^2_{m,q,r}(a^2)$ then its moments can be directly deduced from the characterizations of the distribution. In particular,

$$E(x^2) = \frac{r}{m(r-2)}(a^2+m),$$

$$E(x^4) = \frac{r^2}{m^2(r-2)(r-4)}\left(m(m+2)+2(m+2)a^2+\frac{q+2}{q}a^4\right),$$

$$\mathrm{Var}(x^2) = \frac{2r^2}{m^2(r-2)(r-4)}\left(\frac{q+r-2}{q}a^4+2(m+r-2)a^2+m(m+r-2)\right),$$

$$E(x^6) = \frac{r^3}{m^3(r-2)(r-4)(r-6)}$$
$$\left(m(m+2)(m+4)+3(m+2)(m+4)a^2+3(m+4)\frac{q+1}{q}a^4 + \frac{(q+2)(q+4)}{q^2}a^6\right)$$

if respectively $r>2$, $r>4$, $r>4$ and $r>6$.

For instance, from the moments of the conditional distribution $x^2|y^2 \sim F'_{m,r}(a^2y^2)$, we get

$$E(x^2) = E(E(x^2|y^2)) = \frac{r}{m(r-2)}(a^2E(y^2)+m),$$

$$E(x^4) = E(E(x^4|y^2))\frac{r^2}{m^2(r-2)(r-4)}(m(m+2)+2(m+2)a^2E(y^2)+a^4E(y^4))$$

and the results are deduced from the moments of the marginal distribution $y^2 \sim \chi^2_q/q$.

3.4. Probability density function

If $x^2 \sim K^2_{m,q,r}(a^2)$ its *pdf* is defined for m, q and r positive real numbers and can be expressed as

$$p(x^2) = mq^{\frac{q}{2}} \frac{1}{\Gamma(\frac{q}{2})\Gamma(\frac{r}{2})} \sum_{j=0}^{+\infty} \frac{1}{j!} \frac{\Gamma(\frac{q}{2}+j)\Gamma(\frac{m+r}{2}+j)}{\Gamma(\frac{m}{2}+j)} (a^2)^j (q+a^2)^{-\frac{q}{2}-j}$$
$$(mx^2)^{\frac{m}{2}+j-1}(r+mx^2)^{-\frac{m+r}{2}-j}.$$

Proof. From the characterization as a mixture of *noncentral F* distributions, we get $p(x^2) = \int p(x^2|y^2)p(y^2)\mathrm{d}y^2$ where

$$p(x^2|y) = mr^{\frac{r}{2}} \frac{1}{\Gamma(\frac{r}{2})} \sum_{j=0}^{+\infty} \frac{1}{j!} \frac{\Gamma(\frac{m+r}{2}+j)}{\Gamma(\frac{m}{2}+j)}$$
$$\times \left(\frac{a^2}{2}\right)^j (y^2)^j (mx^2)^{\frac{m}{2}+j-1}(r+mx^2)^{-\frac{m+r}{2}-j},$$

$$p(y^2) = \frac{1}{\Gamma(\frac{q}{2})} \left(\frac{q}{2}\right)^{\frac{q}{2}} (y^2)^{\frac{q}{2}-1} \exp\left(-\frac{q}{2}y^2\right).$$

Inverting the summation and the integration, the result is deduced from the gamma integral (see, for instance, of p. 144 Box and Tiao, 1973)

$$\int_0^{+\infty} (y^2)^{\frac{q}{2}+j-1} \exp\left(\frac{q}{2}y^2\right) \mathrm{d}y^2 = \Gamma\left(\frac{q}{2}+j\right)\left(\frac{q}{2}\right)^{-\frac{q+j}{2}}.$$

3.5. Cumulative distribution function

From the density, the *cdf* can be expressed in terms of an infinite series of multiples of incomplete beta function ratios

$$\Pr(K^2_{m,q,r}(a^2) < \ell) = \sum_{j=0}^{+\infty} f_j I_{m\ell/(r+m\ell)}\left(\frac{m}{2}+j, \frac{r}{2}\right) \quad (\ell > 0)$$

where

$$f_j = \frac{1}{j!} \frac{\Gamma(\frac{q}{2}+j)}{\Gamma(\frac{q}{2})} \left(\frac{q}{q+a^2}\right)^{\frac{q}{2}} \left(\frac{a^2}{q+a^2}\right)^j \quad \text{and} \quad \sum_{j=0}^{+\infty} f_j = 1.$$

This series can be evaluated from recurrence relationships, both for the coefficients f_j and for the incomplete beta functions. The algorithm developed for the *psi-square* distribution (involving the same f_j) in Lecoutre et al. (1992) can be extended to the present situation.

4. Predictive distributions

4.1. Predictive distributions for Student's t statistics

The Student's t statistics are defined as the ratio $t = d/bs$ of two independent variables d and s with respective sampling distributions

$$d|\delta,\sigma^2 \sim N(\delta, b^2\sigma^2)(b>0) \text{ and } s^2|\delta,\sigma^2 \sim \sigma^2\frac{\chi_r^2}{r}.$$

Then if (δ,σ^2) has the conjugate prior distribution characterized by

$$\delta|\sigma^2 \sim N(d_0, b_0^2\sigma^2)(b_0>0) \text{ and } \sigma^2 \sim s_0^2\left(\frac{\chi_{q_0}^2}{q_0}\right)^{-1},$$

t has the marginal predictive distribution

$$t = \frac{d}{bs} \sim K'_{q_0,r}\left(\frac{d_0}{bs_0}, \frac{b_0^2+b^2}{b^2}\right).$$

Proof. First it can be deduced that

$$d|\sigma^2 \sim N(d_0, (b_0^2+b^2)\sigma^2)$$

hence noting that d and s^2 are independent conditionally on σ^2

$$d|s^2,\sigma^2 \sim N(d_0, (b_0^2+b^2)\sigma^2).$$

Then the conditions can be rewritten

$$t = \frac{d}{bs}|g^2, \qquad h^2 \sim N\left(\frac{d_0}{bs_0}\frac{g}{h}, \frac{b_0^2+b^2}{b^2}\frac{1}{h^2}\right) \quad (g>0, h>0),$$

$$g^2 = \frac{s_0^2}{\sigma^2} \sim \frac{\chi_{q_0}^2}{q_0} \quad h^2 = \frac{s^2}{\sigma^2} \sim \frac{\chi_r^2}{r}$$

and the result is obtained as in Section 2.2.

As a consequence the predictive distribution of t can be directly deduced from

$$t|\zeta \sim t'_r\left(\frac{\zeta}{b}\right) \quad \text{and} \quad \zeta \sim \Lambda'_{q_0}(z_0, b_0^2) \quad \text{where } \zeta = \frac{\delta}{\sigma} \quad \text{and } z_0 = \frac{d_0}{s_0}.$$

4.2. Predictive distributions for Student's confidence limits

Confidence limits associated with Student's t tests are of the general form $d+cs$ and have a predictive distribution given by

$$d + cs - d_0 \sim K'_{r,q_0}\left(cs_0, (b_0^2+b^2)s_0^2\right).$$

Proof. From $d|s^2,\sigma^2 \sim N((b_0^2+b^2)\sigma^2)$ (see previous Section), it follows that

$$d + cs - d_0|s^2,\sigma^2 \sim N((b_0^2+b^2)\sigma^2)$$

and the result is obtained applying the equivalent characterizations, with the conditional on σ^2 distribution

$$d + cs - d_0|\sigma^2 \sim \Lambda'_r(c\sigma, (b_0^2 + b^2)\sigma^2).$$

4.3. Predictive distributions for F ratios

The results obtained for the Student's t statistics can be immediately generalized, replacing the univariate normal distributions by m-variate distributions, i.e.,

$$\boldsymbol{D}|\boldsymbol{\Lambda}, \sigma^2 \sim \boldsymbol{N}_m(\boldsymbol{\Lambda}, b^2\sigma^2\boldsymbol{I}_m) \quad \text{and} \quad \boldsymbol{\Lambda}|\sigma^2 \sim \boldsymbol{N}_m(\boldsymbol{D}_0, b_0^2\sigma^2\boldsymbol{I}_m) \quad (b > 0, b_0 > 0).$$

It follows that $\boldsymbol{D}/bs$ has the marginal predictive distribution

$$\frac{\boldsymbol{D}}{bs} \sim \boldsymbol{K}'_{m;q_0,r}\left(\frac{\boldsymbol{D}_0}{bs_0}, \frac{b_0^2 + b^2}{b^2}\boldsymbol{I}_m\right),$$

and consequently

$$\frac{\boldsymbol{D}'\boldsymbol{D}}{mb^2s^2} \sim \frac{b_0^2 + b^2}{b^2} K^2_{m,q_0,r}\left(\frac{\boldsymbol{D}'_0\boldsymbol{D}_0}{(b_0^2 + b^2)s_0^2}\right)$$

with the conditional distribution

$$\frac{\boldsymbol{D}'\boldsymbol{D}}{mb^2s^2}|\boldsymbol{\Lambda}, \sigma^2 \sim F'_{m,r}\left(\frac{\boldsymbol{\Lambda}'\boldsymbol{\Lambda}}{b^2\sigma^2}\right).$$

As a consequence, defining $\zeta^2 = \boldsymbol{\Lambda}'\boldsymbol{\Lambda}/m\sigma^2$ and $z_0^2 = \boldsymbol{D}'_0\boldsymbol{D}_0/ms_0^2$, the predictive distribution of F ratios can be directly deduced from

$$F|\zeta^2 \sim F'_{m,r}\left(\frac{m\zeta^2}{b^2}\right) \quad \text{and} \quad \zeta^2 \sim b_0^2\Lambda^2_{m,q_0}\left(\frac{mz_0^2}{b_0^2}\right)$$

which implies

$$F \sim \frac{b_0^2 + b^2}{b^2} K^2_{m,q_0,r}\left(\frac{mz_0^2}{b_0^2 + b^2}\right).$$

References

Adcock, C.J., 1997. Sample size determination: a review. Appl. Statist. 46, 261–283.

Berry, D.A., 1991. Experimental design for drug development: a Bayesian approach. J. Biopharm. Statist. 1, 81–102.

Box, G.E.P., Tiao, G.C., 1973. Bayesian Inference in Statistical Analysis. Addison-Wesley, Reading, MA.

Choi, S.C., Pepple, P.A., 1989. Monitoring clinical trials based on predictive probability of significance. Biometrics 45, 317–323.

Fisher, R.A., 1990. Statistical Methods, Experimental Design and Scientific Inference, Reedition. Oxford University Press, Oxford.

Geisser, S., 1965. Bayesian estimation in multivariate analysis. Ann. Math. Statist. 36, 150–159.

Geisser, S., 1993. Predictive Inference: An Introduction. Chapman & Hall, New York.

Hung, H.M.J., O'Neill, R.T., Baver, P. Köhne, K., 1997. The behavior of the P-Value when the alternative hypothesis is true. Biometrics 53, 11–22.

Johnson, N.L., Kotz, S., Balakrishnan, 1995. Distributions in Statistics, Continuous Univariate Distributions, Wiley Series in Probability and Mathematical Statistics, vol. 2, 2nd ed. Wiley, New York.

Joseph, L., Bélisle, P., 1997. Bayesian sample size determination for normal means and differences between normal means. Appl. Statist. 46, 209–226.

Lecoutre, B., 1981. Extensions de l'analyse de la variance: L'analyse bayésienne des comparaisons. Math. Sci. Hum. 75, 49–69.

Lecoutre, B., 1984. L'Analyse Bayésienne des Comparaisons. Presses Universitaires de Lille, Lille, France.

Lecoutre, B., 1985. Reconsideration of the F test of the analysis of variance: the semi-Bayesian significance test. Comm. Statist. Theory Meth. 14, 2437–2446.

Lecoutre, B., Dzerko, G., Grouin, J.-M., 1995. Bayesian predictive approach for inference about proportions. Statist. Med. 14, 1057–1063.

Lecoutre, B., Guigues, J.-L., Poitevineau, J., 1992. Distribution of quadratic forms of multivariate Student variables. Appl. Statist. 41, 617–627.

Rao, C.R., 1965. Linear Statistical Inference and its Applications. Wiley, New York.

Rouanet, H., Lecoutre, B., 1983. Specific inference in ANOVA: from significance tests to Bayesian procedures. Br. J. Math. Statist. Psych. 36, 252–268.

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, Oxford, pp. 419–434.

Schervish, M.J., 1995. Theory of Statistics. Springer, New York.

Spiegelhalter, D.J., Freedman, L.S., 1986. A predictive approach to selecting the size of a clinical trial. Statist. Med. 5, 421–433.

Spiegelhalter, D.J., Freedman, Parmar, M.K.B., 1994. Bayesian approaches to randomized trials. J.R. Statist. Soc. A 157 357–416.