COMM. STATIST.-THEOR. METH., 14(10), 2437-2446

RECONSIDERATION OF THE F-TEST OF THE ANALYSIS OF VARIANCE: THE SEMI-BAYESIAN SIGNIFICANCE TEST

B. Lecoutre

Université René Descartes
Groupe Mathématiques et Psychologie
12 rue Cujas, 75005 Paris

Key Words and Phrases: significance tests; nuisance parameters; Bayesian predictive distribution; analysis of variance; F and psi-square distributions.

ABSTRACT

The usual F-test of the analysis of variance is reconsidered within the Bayesian framework, in terms of predictive distributions. This leads to the notion of semi-Bayesian significance test, so called because it consists in only probabilizing the space of nuisance parameters, thus bringing a general principle for "eliminating" nuisance parameters, or more exactly incorporating information about these parameters. The approach is shown to extend the F-tests, by allowing the testing of hypotheses of non-zero effects.

Downloaded At: 10:49 19 January 2010

1. INTRODUCTION

It is well known that a given statistical procedure can pertain to different justification and interpretation frameworks. Thus, as has been especially illustrated by Lindley (1965), usual significance tests, such as Student's t or F tests of the analysis of variance, which are traditionally presented in the sampling theory framework, can receive an interpretation in the Bayesian framework, using noninformative priors in the sense of Jeffreys (not to speak of Fisher's fiducial approach).

Such results suggest a continuity, rather than a drastic opposition, between "classical" and Bayesian methods. Yet the reinterpretation of levels of significance in the Bayesian framework is usually performed in terms of posterior distributions, that is distributions on parameters, whereas the same levels in the sampling theory are obtained from sampling distributions of statistics. This paper explores the possibility, in the Bayesian framework, of a more direct reconsideration of usual significance tests, in terms of predictive distributions, that is, distributions concerning statistics. This approach will lead to the notion of *semi-Bayesian significance test*, so called since it consists in only probabilizing the space of nuisance parameters.

The basic idea to replace the unknown nuisance parameters by a probability distribution is certainly not a novel one. Thus, Fisher (1944) gave the following presentation for the elementary Student's t test: "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 $\frac{x}{\sigma}$ 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 $\frac{x}{s}$, and this is not normal. The true value has been divided by a factor $\frac{s}{\sigma}$ which introduces an error. (...) the distribution of $\frac{s}{\sigma}$ 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.".

Curiously enough this presentation, perhaps because expressed in the fiducial framework, seems to have received little attention. However, reconsidered in the Bayesian framework, starting with a prior distribution for σ, such a presentation directly leads to the semi-Bayesian significance test, which brings both conceptually and technically major improvements.

Conceptually, semi-Bayesian significance testing leads to intermediate procedures between the usual significance tests and

Downloaded At: 10:49 19 January 2010

the Bayesian procedures, since it consists in only probabilizing the space of nuisance parameters. Hence it can bring a useful compromise in the situations where one could be reluctant to choose a particular prior distribution for the parameter of interest, but one would not have such a circumspection for the nuisance parameters.

Technically, semi-Bayesian significance testing includes the results of usual significance tests, and extends these results in a substantial way, since it brings a general principle for dealing with the informations available about nuisance parameters.

Furthermore, the procedure can receive a Bayesian justification in terms of partial sufficiency. Thus, in the preceding example, "H-sufficiency" (Hájek, 1965) or "G-sufficiency" (Barnard, 1963) yield the statistic s as containing all the available information about σ in the absence (or ignorance as it is the case in significance testing) of prior knowledge about the true mean: see respectively Basu (1977) and Dawid (1980).

The illustration will be conducted from the familiar example of the univariate analysis of variance (ANOVA), an example of both satisfying complexity and general scope. It will be shown that semi-Bayesian significance testing includes the results of usual F-tests, and furthermore extends these results in problems where the sampling theory approach appears to stumble, such as testing the hypothesis that the variance associated with a fixed-effect has a given non-zero value, or incorporating informations available about the nuisance parameters.

Downloaded At: 10:49 19 January 2010

2. A MAJOR DIFFICULTY OF USUAL SIGNIFICANCE TESTING IN THE SAMPLING THEORY

One of the major difficulties encountered in the sampling theory is the "elimination" of nuisance parameters. Even when sufficient statistics are available, one can find many examples where there is difficulty in eliminating nuisance parameters, especially the problem of comparing two means under the normal model with unequal variances or again several problems that arise in analysis of variance. In the case considered here of usual significance testing, this difficulty is solved when an appropriate test statistic can be found, whose sampling distribution under the hypothesis to be tested does not involve any parameter.

Thus the F-tests of univariate analysis of variance, as is well known, are built from two sufficient statistics, respectively denoted here SS_{effect} and SS_{error} (Sums of Squares), whose sampling distributions are independent and such that:

$$SS_{effect} \sim k\sigma^2_{error}\chi^2_m(\sigma^2_{effect}/b\sigma^2_{error})$$

$$SS_{error} \sim k\sigma^2_{error}\chi^2_q$$

where k and b are numerical coefficients.

Then the test of the null hypothesis: $\sigma^2_{effect} = 0$ is obtained from the F-ratio of the mean squares: $F = MS_{effect}/MS_{error}$ $= (SS_{effect}/m)/(SS_{error}/q)$, whose sampling distribution under H_0 is a central F-distribution and therefore does not involve the nuisance parameter σ^2_{error}.

Under H_0: $\sigma^2_{effect} = 0$, $F \sim F_{m,q}$

Unhapilly, the preceding solution is far from being a universal principle and, to consider only this situation, if we want to test another point hypothesis that σ^2_{effect} has a given non-zero value σ^2_0, the nuisance parameter σ^2_{error} does not let itself eliminated.

3. THE PRINCIPLE OF THE SEMI-BAYESIAN SIGNIFICANCE TEST: AN ILLUSTRATION

Let us now formally reconsider the preceding problem, starting with the sampling distribution of the SS_{effect} statistic, which, given σ^2_{error}, is sufficient for σ^2_{effect}. Under H_0: $\sigma^2_{effect} = \sigma^2_0$, this distribution is conditional on σ^2_{error}:

$$SS_{effect}|\sigma^2_{error} \sim k\sigma^2_{error}\chi^2_m(\sigma^2_0/b\sigma^2_{error}) \quad (1)$$

The idea which can easily come to the mind of a Bayesian

Downloaded At: 10:49 19 January 2010

trying to solve this problem is to replace the unknown value σ^2_{error} by a probability distribution for this parameter, the only one to be considered under H_o. A reasonable principle is to assume a *prior* distribution for σ^2_{error} and to consider the *predictive* distribution for SS_{effect}, conditional on the observed value, denoted s^2, taken by the MS_{error} statistic, instead of the sampling distribution in (1). This is the principle of what can be called a *semi-Bayesian significance test*.

As an example, the simplest solution, called here the *standard* solution for σ^2_{error}, is obtained by assuming a prior distribution (often called vague or noninformative), defined by a probability density $p(\sigma^2_{error})$ proportional to $1/\sigma^2_{error}$ (see e.g. Lindley, 1965, or Box & Tiao, 1973). In this important particular case, the conditional predictive distribution for SS_{effect} given s^2 (under H_o) is shown to be:

$$\text{Under } H_o:\ \sigma^2_{effect} = \sigma^2_o,$$

$$SS_{effect} | s^2 \sim s^2 \psi^2_{m,q}(k\sigma^2_o/bs^2) \qquad (2)$$

This result can be obtained here from the distributions:

$$SS_{effect} | \sigma^2_{error}, s^2 \sim k\sigma^2_{error}\chi^2_m(\sigma^2_o/b\sigma^2_{error}) \qquad (1')$$

which is deduced form (1) taking into account the independence of the sampling distributions of SS_{effect} and SS_{error},

$$\text{and } \sigma^2_{error} | s^2 \sim \frac{s^2}{k}(\chi^2_q/q)^{-1} \qquad (3)$$

which is derived from the sampling distribution of SS_{error} and the prior distribution for σ^2_{error} by a usual application of the Bayes' formula.

Then we get (2) by the definition of the psi-square distribution introduced by Lecoutre (1981, 1984) and Rouanet & Lecoutre (1983). This distribution can be characterized as follows:

if $x^2 | y^2 \sim y^2\chi^2_m(a^2/y^2)$ and $y^2 \sim (\chi^2_q/q)^{-1}$ then x^2 is distributed as ψ^2 with eccentricity a^2 and m and q degrees of freedom; it can be shown that, if $u^2 | y^2 \sim F_{m,q}(a^2/y^2)$ and $y^2 \sim (\chi^2_q/q)^{-1}$, then $u^2 \sim \frac{1}{m}\psi^2_{m,q}(a^2)$. Lastly, if $x^2 \sim \psi^2_{m,q}(a^2)$, the density of x^2 is given by:

Downloaded At: 10:49 19 January 2010

$$f(x^2) = q^{q/2} \frac{1}{\Gamma(\frac{q}{2})} (x^2)^{(m/2)-1} (q+a^2+x^2)^{-(m+q)/2} \times \sum_{j=0}^{+\infty} \frac{1}{j!} \frac{\Gamma(\frac{m+q}{2}+2j)}{\Gamma(\frac{m}{2}+j)} \left(\frac{a^2x^2}{(q+a^2+x^2)^2}\right)^j$$

The predictive $SS_{effect}|s^2$ distribution can be used, in a straightforward way, to proceed to a semi-Bayesian significance test of the null hypothesis $\sigma^2_{effect} = \sigma^2_0$: this test will consist in examining if the data are (in a certain sense) compatible with this predictive distribution.

More precisely, the data will be declared compatible with this predictive distribution ("nonsignificant result") if the observed value of SS_{effect}, denoted SS_{obs}, lies in an expected (rather a not unexpected) region, that is, formally, if, under H_0: $\sigma^2_{effect} = \sigma^2_0$, $P(SS_{effect} > SS_{obs}|s^2)$ is greater than a given α level. On the contrary, the data will be declared non compatible with this predictive distribution ("significant result") if SS_{obs} lies in a unexpected region, that is, if, under H_0, $P(SS_{effect} > SS_{obs}|s^2) < \alpha$.

It is easy to see that, in the particular case $\sigma^2_0 = 0$, where the usual significance test is defined, the results and conclusions of the *standard* semi-Bayesian significance test coincide with the ones of the usual F-test, since one can deduce from (2):

$$\text{Under } H_0\text{: } \sigma^2_{effect} = \sigma^2_0,\ F|s^2 \sim \frac{1}{m}\psi^2_{m,q}(k\sigma^2_0/bs^2) \qquad (4)$$

with, in the particular case $\sigma^2_0 = 0$

$$\text{Under } H_0\text{: } \sigma^2_{effect} = 0,\ F|s^2 \sim \frac{1}{m}\psi^2_{m,q}(0) \sim F_{m,q} \qquad (4')$$

4. FURTHER IMPROVEMENTS OF THE SEMI-BAYESIAN SIGNIFICANCE TEST

In addition to the possibility, illustrated in the preceding Section, of testing any point hypothesis for σ^2_{effect}, the semi-Bayesian significance test allows us to remove many limitations inherent in the usual significance tests. Still for the case of the univariate analysis of variance, two major improvements of the semi-Bayesian significance test will be illustrated hereafter.

Downloaded At: 10:49 19 January 2010

Downloaded At: 10:49 19 January 2010

Incorporating other sources of information. There is obviously no necessity to restrict semi-Bayesian significance testing to the two extreme cases usually distinguished in the sampling theory: σ^2_{error} known vs σ^2_{error} unknown apart from the data. The solution is immediate for the intermediate cases, since one can incorporate some other source of information about σ^2_{error} by choosing an appropriate prior distribution.

As an illustration, let us assume a prior distribution on σ^2_{error} such that

$$\sigma^2_{error} \sim \frac{s'^2}{k}(\chi^2_{q'}/q')^{-1}$$

(where the numerical coefficient k is introduced only for reasons of convenience)

Hence we get the semi-Bayesian significance test of the null hypothesis $\sigma^2_{effect} = \sigma^2_0$, given by the following distribution

Under H_0: $\sigma^2_{effect} = \sigma^2_0$,

$$SS_{effect} \mid s^2 \sim w^2 \psi^2_{m,q+q'}(k\sigma^2_0/bw^2) \qquad (5)$$

$$\text{with } w^2 = \frac{qs^2+q's'^2}{q+q'}$$

This result can be obtained, as for the standard solution, the distribution in (3) being remplaced by

$$\sigma^2_{error} \mid s^2 \sim \frac{w^2}{k}(\chi^2_{q+q'}/(q+q'))^{-1} \qquad (6)$$

It will be noted that, when $q' \to 0$, one finds again, as a limiting case, the standard semi-Bayesian significance test given by (2).

Random effect models. An other example of improvement brought by semi-Bayesian significance testing pertains to the case where two (or more) nuisance parameters are functionally related. Such a situation typically occurs in "*random effect models*": in these models the variance σ^2_{error} is divided into two or more additive components. As an illustration, let us assume that, in addition of the two SS_{effect} and SS_{error} statistics, we have an additional statistic, let us say SS'_{error}, whose the sampling distribution, independent of the preceding ones, involves only one component, let us say η^2_{error}, of the two components of the total variance σ^2_{error}. More precisely, we have the independent sampling distributions:

$$SS_{effect} \sim k\sigma^2_{error}\chi^2_m(\sigma^2_{effect}/b\sigma^2_{error})$$
$$SS_{error} \sim k\sigma^2_{error}\chi^2_q \qquad (7)$$
$$SS'_{error} \sim k\eta^2_{error}\chi^2_p$$

with the constraint $\sigma^2_{error} > \eta^2_{error}$

The three statistics SS_{effect}, SS_{error} and SS'_{error} are assumed jointly sufficient for the three parameters σ^2_{effect}, σ^2_{error}, η^2_{error}.

In the sampling theory, this situation is embarassing, since it is quite difficult to take into account the information brought by the SS'_{error} statistic, for lack to be able to explicitely integrate the constraint $\sigma^2_{error} > \eta^2_{error}$, which is here an information about the parameters related to the very structure of the problem.

For the semi-Bayesian significance test, such a situation does not generate new conceptual difficulties; simply the procedures will be technically more complex.

As an illustration, let us consider the standard solution corresponding to the prior probability density $p(\sigma^2_{error},\eta^2_{error})$ proportional to $1/\sigma^2_{error}\eta^2_{error}$, with the constraint $\sigma^2_{error} > \eta^2_{error}$. Then the standard semi-Bayesian significance test of the null hypothesis H_o: $\sigma^2_{effect} = \sigma^2_o$ is given by the predictive distribution for SS_{effect}, conditional on s^2 and u^2, the observed value of $MS'_{error} = SS'_{error}/q'$. This distribution can be characterized by its density:

Under H_o: $\sigma^2_{effect} = \sigma^2_o$,

$$p(SS_{effect}|s^2,u^2) = \int_0^1 p(SS_{effect}|\varepsilon,s^2,u^2)p(\varepsilon|s^2,u^2)d\varepsilon \qquad (8)$$

$$\text{with } SS_{effect}|\varepsilon,s^2,u^2 \sim \frac{pu^2+qs^2\varepsilon}{(p+q)\varepsilon}\,\psi^2_{m,p+q}\left(\frac{k\sigma^2_o(p+q)\varepsilon}{b(pu^2+qs^2\varepsilon)}\right) \qquad (8.1)$$

$$\text{and } p(\varepsilon|s^2,u^2) = \frac{1}{P(F_{q,p} < \frac{s^2}{u^2})}\,p(\frac{u^2}{s^2}F_{q,p} = \varepsilon) \qquad (0<\varepsilon<1) \qquad (8.2)$$

Downloaded At: 10:49 19 January 2010

This result can be obtained from the distributions:

$$SS_{effect} | \sigma^2_{error}, \eta^2_{error}, s^2, u^2 \sim k\sigma^2_{error}\chi^2_m(\sigma^2_0/b\sigma^2_{error}) \quad (7')$$

which is deduced from (7) taking into account the independence of the sampling distributions,

$$\sigma^2_{error} | s^2, u^2 \sim \frac{s^2}{k}(\chi^2_q/q)^{-1}$$

and

$$\eta^2_{error} | s^2, u^2 \sim \frac{u^2}{k}(\chi^2_p/p)^{-1}$$

(with the constraint $\sigma^2_{error} > \eta^2_{error}$, the two distributions, apart from the constraint being independent)

which are derived by a usual application of the Bayes' formula.

Then, after the change of variables $\varepsilon = \eta^2_{error}/\sigma^2_{error}$, one obtains the distribution in (8.2) and

$$\sigma^2_{error} | \varepsilon, s^2, u^2 \sim \frac{pu^2+qs^2\varepsilon}{k(p+q)\varepsilon}(\chi^2_{p+q}/(p+q))^{-1} \qquad (0<\varepsilon<1)$$

and, consequently, the distribution in (8.1).

It will be noted that the distribution in (8.1) is a central *F* if $\sigma^2_0 = 0$, and that it reduces to (2) for $p = 0$.

Here again an other distribution that the noninformative one could be considered for σ^2_{error} and η^2_{error}.

BIBLIOGRAPHY

Barnard, G.A. (1963), Some logical aspects of the fiducial argument, *J. Roy. Statist. Soc. B*, *25*, 111-114

Basu, D. (1977), On the elimination of nuisance parameters, *J. Amer. Statist. Ass.*, *72*, 355-366.

Box, G.E.P. & Tiao, G.C. (1973), *Bayesian Inference in Statistical Analysis*, Reading: Addison-Wesley.

Dawid, A.P. (1980), A bayesian look at nuisance parameters, *in* J.M. Bernardo, M.H. De Groot; D.V. Lindley & A.F.M. Smith (Eds.), *Bayesian Statistics*, Valencia: University Press, 167-184.

Fisher, R.A. (1944), *Statistical Methods for Research Workers*, Edinburgh and London: Oliver and Boyd, ninth edition.

Hájek, J. (1965), On basic concepts of statistics. *Fith Berkeley Symposium on Mathematical Statistic and Probability*, *1*, 139-162.

Downloaded At: 10:49 19 January 2010

Lecoutre, B. (1981), Extensions de l'analyse de la variance: l'analyse bayésienne des comparaisons, *Mathématiques et Sciences Humaines*, *75*, 49-69.

Lecoutre, B. (1984), *L'Analyse Bayésienne des Comparaisons*, Lille: Presses Universitaires de Lille.

Lindley, D.V. (1965), *Introduction to Probability and Statistics from a Bayesian Viewpoint, Part 2 - Inference*, Cambridge: Cambridge University Press.

Rouanet, H. & Lecoutre, B. (1983), Specific inference in ANOVA: from significance tests to Bayesian procedures, *British J. of Math. and Stat. Psychol.*, *36*, 252-268.

Received by Editorial Board member, July, 1984; Revised April, 1985.

Recommended by Shelemyahu Zacks, State University of New York at Binghamton, NY

Refereed by C. Pereira, University of Sao Paulo, Sao Paulo, Brazil.

Downloaded At: 10:49 19 January 2010