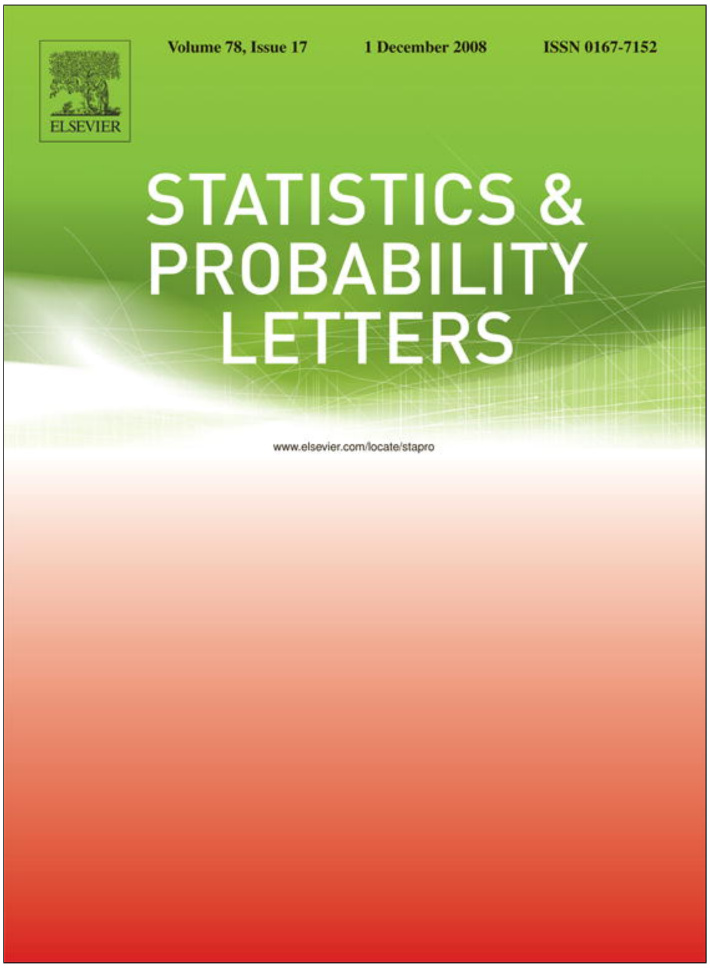
Volume 78, Issue 17
1 December 2008
ISSN 0167-7152
ELSEVIER
STATISTICS &
PROBABILITY
LETTERS
www.elsevier.com/locate/stapro

Statistics and Probability Letters 78 (2008) 3070–3074

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

The multivariate point null testing problem: A Bayesian discussion

Miguel A. Gómez-Villegas*, Beatriz González-Pérez

Departamento de Estadística e Investigación Operativa, Facultad de CC. Matemáticas, Universidad Complutense de Madrid. Madrid, 28040, Spain

ARTICLE INFO

Article history:
Received 15 December 2005
Received in revised form 30 October 2007
Accepted 1 May 2008
Available online 18 May 2008

ABSTRACT

In this paper the problem of testing a multivariate point hypothesis is considered. Of interest is the relationship between the *p*-value and the posterior probability. A Bayesian test for simple $H_0 : \theta = \theta_0$ versus bilateral $H_0 : \theta \neq \theta_0$, with a mixed prior distribution for the parameter θ, is developed. The methodology consists of fixing a sphere of radius δ around θ_0 and assigning a prior mass, π_0, to H_0 by integrating the density $\pi(\theta)$ over this sphere and spreading the remainder, $1 - \pi_0$, over H_1 according to $\pi(\theta)$. A theorem that shows when the frequentist and Bayesian procedures can give rise to the same decision is proved. Then, some examples are revisited.

1. Introduction

Let $f(x_1, \ldots, x_n|\boldsymbol{\theta})$ ($\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$ unknown) be the likelihood of a sample $(X_1, \ldots, X_n)$. For a specified $\boldsymbol{\theta_0} = (\theta_{01}, \ldots, \theta_{0m})$ we want to test

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \quad \text{versus } H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0. \tag{1}$$

As usual, we use the following mixed prior distribution to test (1)

$$\pi^*(\boldsymbol{\theta}) = \pi_0 I_{H_0}(\boldsymbol{\theta}) + (1 - \pi_0)\, \pi(\boldsymbol{\theta})\, I_{H_1}(\boldsymbol{\theta}). \tag{2}$$

For a proper metric d and a value of δ sufficiently small, consider

$$H_{0\delta} : d(\boldsymbol{\theta}_0, \boldsymbol{\theta}) \leq \delta, \quad \text{versus } H_{1\delta} : d(\boldsymbol{\theta}_0, \boldsymbol{\theta}) > \delta.$$

Denote $B(\boldsymbol{\theta}_0, \delta) = \left\{\boldsymbol{\theta} \in \Theta^r, \sum_{i=1}^m (\theta_i - \theta_{0i})^2 \leq \delta^2\right\}$. Gómez-Villegas et al. (2007) proposed to compute π_0 by means of

$$\pi_0 = \int_{B(\boldsymbol{\theta}_0, \delta)} \pi(\boldsymbol{\theta})\, d\boldsymbol{\theta}. \tag{3}$$

In this way, π_0 is the probability assigned to H_0, through $\pi^*(\boldsymbol{\theta})$, and $H_{0\delta}$, through $\pi(\boldsymbol{\theta})$, the Kullback–Leibler discrepancy between $\pi(\theta)$ and $\pi^*(\theta)$ goes to zero as δ goes to zero, and the posterior probability of the null is

$$P_{\pi^*}(\boldsymbol{\theta}_0|x_1, \ldots, x_n) = \left[1 + \frac{1 - \pi_0}{\pi_0} \eta(x_1, \ldots, x_n)\right]^{-1}, \tag{4}$$

$$\eta(x_1, \ldots, x_n) = \frac{\int_\Theta f(x_1, \ldots, x_n|\boldsymbol{\theta})\, \pi(\boldsymbol{\theta})\, d\boldsymbol{\theta}}{f(x_1, \ldots, x_n|\boldsymbol{\theta}_0)}. \tag{5}$$

* Corresponding author.
E-mail addresses: ma.gv@mat.ucm.es (M.A. Gómez-Villegas), beatrizg@mat.ucm.es (B. González-Pérez).

0167-7152/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.spl.2008.05.010

Suppose that H_0 is rejected when $P_{\pi^*}(\boldsymbol{\theta}_0|x_1,\ldots,x_n) \le 1/2$. On the other hand, if $\Lambda(x_1,\ldots,x_n) = \sup_{\boldsymbol{\theta}\in\Theta} f(x_1,\ldots,x_n|\boldsymbol{\theta}) / f(x_1,\ldots,x_n|\boldsymbol{\theta}_0)$ is the test statistic, then the associated p-value,

$$p(x_1,\ldots,x_n) = P(\Lambda(X_1,\ldots,X_n) > \Lambda(x_1,\ldots,x_n)|\boldsymbol{\theta}_0), \tag{6}$$

can be approximated by $P\left(\chi^2_m > 2\ln\Lambda(x_1,\ldots,x_n)\right)$. In this case, H_0 is rejected when $p(x_1,\ldots,x_n) \le p^*$, for a level of significance p^*. The aim is to find conditions for δ in order that, when p^* is fixed, whatever be $(x_1,\ldots,x_n)\in\chi$, the same decision is reached with both methods. In Section 2, a theorem that shows when both approaches are in agreement is proved. In Section 3 some relevant examples are revisited. Conclusions and comments are included in Section 4.

2. Comparison between both approaches

In parametric testing of a simple null hypothesis, as opposed to the one-sided (see Casella and Berger, 1987), Bayesian and frequentist procedures can give rise to different decisions (see Lindley, 1957; Berger and Sellke, 1987; Berger and Delampady, 1987, among others). In most of the Bayesian approaches, the infimum of a Bayesian evidence measure over a wide class of priors is computed and is substantially larger than the p-value. Other important references are Oh and Dasgupta (1999), Gómez-Villegas and Sanz (2000), Sellke et al. (2001), Gómez-Villegas et al. (2002) and Gómez-Villegas and González-Pérez (2005, 2006).

2.1. Preliminaries

In this section we introduce some definitions and results in order to prove a characterization theorem of the agreement between both methods to test (1).

Definition 2.1. Let $\pi(\boldsymbol{\theta})$ be the prior about $\boldsymbol{\theta}$. $T = T(X_1,\ldots,X_n)$ is a *sufficient statistic to test* $H_0 : \theta\in\Theta_0$, *versus* $H_1 : \theta\in\Theta_1$, with $\Theta_0\cap\Theta_1 = \phi$ and $\Theta_0\cup\Theta_1 = \Theta$, if $P_\pi(\Theta_0|x_1,\ldots,x_n) = P_\pi(\Theta_0|t)$, when $T(x_1,\ldots,x_n) = t$.

Note that the Bayes factor η given in (5) is a *sufficient statistic to test* the multivariate point null given in (1) when the mixed prior (3) is used. In fact, with this prior choice, if T is a *sufficient statistic to test* (1), then there is a function $g : \mathrm{R}\to\mathrm{R}$ such that $g(T) = \eta$. Moreover, the usual definition of *sufficient statistic* is not equal to this new concept of *sufficient statistic to test*. If T is a *sufficient statistic* then T verifies Definition 2.1. The reciprocal is not true. We introduce the next concept in order to compare two statistics.

Definition 2.2. Let $T_1 = T_1(X_1,\ldots,X_n)$ and $T_2 = T_2(X_1,\ldots,X_n)$ be univariate statistics. T_1 is an *increasing tendency statistic* with respect to T_2 in a value $T_1 = t$ if $\sup_{T_1(x_1,\ldots,x_n)<t} T_2(x_1,\ldots,x_n) \le \inf_{T_1(x_1,\ldots,x_n)\ge t} T_2(x_1,\ldots,x_n)$.

Proposition 2.1. *If* $T_1 = h(T_2)$ *and* $h : \mathrm{R}\to\mathrm{R}$ *is a non-decreasing monotonous function, then* T_1 *is an* increasing tendency statistic *with respect to* T_2 *for any value* $T_1 = t$. *Furthermore, when h is a strictly increasing function, then*

$$\sup_{T_1(x_1,\ldots,x_n)<t} T_2(x_1,\ldots,x_n) = \inf_{T_1(x_1,\ldots,x_n)\ge t} T_2(x_1,\ldots,x_n) = h^{-1}(t).$$

2.2. Agreement between frequentist and Bayesian approaches

Theorem 2.1. *Suppose that we wish to test* (1) *with the prior distribution given in* (2) *and* π_0 *as in* (3). *If* Λ *is an* increasing tendency statistic *with respect to* η *in* $\Lambda = \lambda^*$, *then for* $p^* = P_{\boldsymbol{\theta}_0}\{\Lambda\ge\lambda^*\}$ *there is an interval of values of* $\pi_0 = \pi_0(\delta)$, $I = I(p^*, n) = (\ell_1, \ell_2)$, *where both methods are in agreement.*

Proof. Define $A_\eta(\kappa) = \{(x_1,\ldots,x_n), \eta(x_1,\ldots,x_n) = \kappa\}$. $P_{\pi^*}(\boldsymbol{\theta}_0|\kappa) > 1/2$ is verified when $\pi_0 > \kappa(\kappa+1)^{-1}$. $\pi^0(\kappa) = \kappa(\kappa+1)^{-1}$ is a strictly increasing function. Moreover, if $\lambda_1 < \lambda_2$, then $p(\lambda_1) = P_{\theta_0}\{\Lambda\ge\lambda_1\} \ge P_{\theta_0}\{\Lambda\ge\lambda_2\} = p(\lambda_2)$. If Λ is an *increasing tendency statistic* with respect to η in $\Lambda = \lambda^*$ then $\kappa_* = \sup_{(x_1,\ldots,x_n),p>p^*}\eta = \sup_{\Lambda<\lambda^*}\eta \le \inf_{\Lambda\ge\lambda^*}\eta = \inf_{(x_1,\ldots,x_n),p\le p^*}\eta = \kappa^*$ and $\ell_1 = \pi^0(\kappa_*) = \sup_{(x_1,\ldots,x_n),\,p>p^*}\pi^0(\eta) \le \inf_{(x_1,\ldots,x_n),\,p\le p^*}\pi^0(\eta) = \pi^0(\kappa^*) = \ell_2$ are verified for $p^* = P_{\boldsymbol{\theta}_0}\{\Lambda\ge\lambda^*\}$. Consider $\pi_0\in(\ell_1,\ell_2)$. If $(x_1,\ldots,x_n)\in A_\eta(\kappa)$, $\kappa<\kappa_*\le\kappa^*$, then $\pi_0 > \ell_1 > \kappa(\kappa+1)^{-1}$ and $P_{\boldsymbol{\theta}_0}\{\Lambda\ge\Lambda(x_1,\ldots,x_n)\} > p^*$. On the other hand, if $(x_1,\ldots,x_n)\in A_\eta(\kappa)$, $\kappa\ge\kappa^*\ge\kappa_*$, then $\pi_0 < \ell_2 \le \kappa(\kappa+1)^{-1}$ and $P_{\boldsymbol{\theta}_0}\{\Lambda\ge\Lambda(x_1,\ldots,x_n)\} \le p^*$. □

Corollary 2.1. *Let* $\Lambda = h(\eta)$ *with* $h : \mathrm{R}\to\mathrm{R}$ *be a non-decreasing monotonous function. If* $(x_1,\ldots,x_n)\in A_\eta(\kappa)$ *is observed, the same decision is reached with the posterior probability* $P_{\pi^*}(\boldsymbol{\theta}_0|\kappa) = [1+(1-\pi_0(\delta))\kappa/\pi_0(\delta)]^{-1}$ *and* δ *such that* $\pi_0(\delta)\in(\ell_1,\ell_2)$,

$$\ell_1 = \ell_1\left(p^*, n\right) = \sup\nolimits_{(x_1,\ldots,x_n),p>p^*}\eta(\eta+1)^{-1}, \tag{7}$$

$$\ell_2 = \ell_2\left(p^*, n\right) = \inf_{(x_1,\ldots,x_n),p\le p^*}\eta(\eta+1)^{-1}, \tag{8}$$

Table 1
Values of $\delta^*(p^*, n, m) = \delta\sqrt{n}/\sigma^2$ where the agreement is achieved over unimodal and symmetric prior distribution class

δ^*	$m = 2$	$m = 5$	$m = 10$	$m = 20$
$p^* = 0.5$	2	2.78	3.64	4.88
$p^* = 0.1$	4.47	4.53	5.08	6.12
$p^* = 0.05$	6.32	5.44	5.7	6.6
$p^* = 0.01$	14.14	8.13	7.28	7.7
$p^* = 0.001$	44.72	13.95	10.02	9.34

and $p^* = P(\Lambda \geq \lambda^*)$ *to quantify the p-value,* $p(h(\kappa)) = P_{\theta_0}\{\Lambda \geq h(\kappa)\}$*. Moreover, when h is strictly increasing continuous, then* $\eta = h^{-1}(\Lambda)$ *and*

$$\ell_1 = \ell_2 = \pi_0(\delta, p^*, n) = h^{-1}(\lambda^*)\left[h^{-1}(\lambda^*) + 1\right]^{-1}. \tag{9}$$

The proof of Corollary 2.1 is easy by using Theorem 2.1 and Proposition 2.1. Another immediate outcome is that, when the sample size n and the level of significance p^* are fixed, the verification of $\ell_1 \leq \ell_2$, defined in (7) and (8), is sufficient for the existence of agreement between the p-value and the posterior probability. Moreover, when $\Lambda = h(\eta)$ and $h : \mathrm{R} \to \mathrm{R}$ is non-decreasing monotonous, $\ell_1 \leq \ell_2$ is always true whatever be the value of p^*, and if h is strictly increasing continuous then $\ell_1 = \ell_2$ and the agreement is obtained for (9).

3. Applications

3.1. Lindley's paradox

Let $(x_1, \ldots, x_m)$ be a sample from a univariate normal of mean θ and known variance σ^2 and the prior mass of $H_0 : \theta = \theta_0$, be c. Distribute the remainder $1 - c$ uniformly over some interval I containing θ_0. $\bar{x}$, the arithmetic mean, and a minimal sufficient statistic, is well within I for n sufficiently enough. Then $\bar{x} - \theta_0$ tends to zero as n increases and the posterior probability that $\theta = \theta_0$ is

$$\bar{c} = cK^{-1}\exp[-n(\bar{x} - \theta_0)^2/(2\sigma^2)], \tag{10}$$

$K = c\exp[-n(\bar{x} - \theta_0)^2/(2\sigma^2)] + (1 - c)\int_I \exp[-n(\bar{x} - \theta)^2/(2\sigma^2)]\mathrm{d}\theta$. By the assumption about $\bar{x}$ and I, the last integral can be evaluated as $\sigma\sqrt{2\pi/n}$. By using the usual test $\bar{x} = \theta_0 + \lambda_\alpha\sigma\sqrt{n}$ [$\Phi(\lambda_\alpha) = 1 - \alpha/2$, ($\Phi$ being the normal standard distribution function)] is significant at the α% level and

$$\bar{c} = c\exp(-\lambda_\alpha^2/2)\left[c\exp(-\lambda_\alpha^2/2) + (1 - c)\sigma\sqrt{2\pi/n}\right]^{-1}. \tag{11}$$

From (11) we see that $\bar{c} \to 1$ when $n \to \infty$ and whatever be c, it can be found $n = n(c, \alpha)$ such that "*$\bar{x}$ is significantly different from θ_0 at the α% level*" and "*the posterior probability that $\theta = \theta_0$ is* $100(1-\alpha)$%". This is the paradox. Both statements are in direct conflict. However, Lindley (1957) remarks also that if $A = c\exp(-\lambda_\alpha^2/2)[(1 - c)\sqrt{2\pi}]^{-1}$, then $\bar{c} = A(A + \sigma/\sqrt{n})^{-1}$ and $\bar{c} \to 0$ as $\sigma/\sqrt{n} \to \infty$. Therefore, in a small experiment, significance at 5% may give good reasons to doubt the null hypothesis. On the other hand when n and α are fixed, there is a value of $c = c(n, \alpha)$, such that both approaches are in agreement. Observe that the posterior probability (10) may be written $\bar{c} = \left[1 + (1 - c)h^{-1}[T(\bar{x}, \theta_0)]/c\right]^{-1}$, $h^{-1}(u) = \sigma\sqrt{2\pi/n}\exp(u/2)$, $T(\bar{x}, \theta_0) = n(\bar{x} - \theta_0)^2/\sigma^2$. Lindley's argument to show a paradox consists of finding a value of $n = n(c, \alpha)$ such that both decisions differ radically. However, by applying Corollary 2.1 both agree with each other for $c = h^{-1}(\chi^2_{1,\alpha})\left[h^{-1}(\chi^2_{1,\alpha}) + 1\right]^{-1} = \left[1 + n^{1/2}(2\pi\sigma^2)^{-1/2}\exp(-\chi^2_{1,\alpha}/2)\right]^{-1}$. Note that if $\alpha = 0.05$, then $c = 1/2$ when $n/\sigma^2 \approx 300$. In general, for c and α fixed, $n/\sigma^2 = 2\pi c^{-2}(1 - c)^2\exp(\chi^2_{1,\alpha})$.

3.2. Lowers bounds for unimodal and symmetric priors

Let $\mathbf{X} = (X_1, \ldots, X_m)' \sim N_m(\boldsymbol{\theta}, \sigma^2\mathbf{I})$ with σ^2 known, $\mathbf{I}$ the identity matrix $m \times m$ and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)'$ unknown. For testing (1) with a sample of size n the usual test statistic is $T(\overline{\mathbf{X}}, \boldsymbol{\theta}_0) = n\sigma^{-2}(\overline{\mathbf{X}} - \boldsymbol{\theta}_0)'(\overline{\mathbf{X}} - \boldsymbol{\theta}_0)$, with $\overline{\mathbf{X}} = (\overline{X}_1, \ldots, \overline{X}_m)'$, and the p-value of the observed data, $\bar{\mathbf{x}} = (\bar{\mathbf{x}}_1, \ldots, \bar{\mathbf{x}}_m)'$, is $p(\bar{\mathbf{x}}) \approx P\left\{\chi^2_m \geq T(\bar{\mathbf{x}}, \boldsymbol{\theta}_0)\right\}$. If $\pi^*(\boldsymbol{\theta})$ is the mixed prior given in (2) with π_0 computed by (3), the infimum of the posterior probability of the point null when $\pi(\boldsymbol{\theta}) \in Q_{US}$, Q_{US} being the unimodal and symmetric priors about $\boldsymbol{\theta}_0$, is

$$\inf_{\pi \in Q_{US}} P(H_0|\bar{\mathbf{x}}) = \left[1 + 2^{m/2}\delta^{*-m}\Gamma(m/2 + 1)\exp[T(\bar{\mathbf{x}}, \boldsymbol{\theta}_0)/2]\right]^{-1},$$

where $\delta^* = \delta\sqrt{n}/\sigma^2$ (see Gómez-Villegas et al., 2007). Let t^* be such that $P\left\{\chi^2_m \geq t^*\right\} = p^*$. By applying Corollary 2.1 both methods always agree with each other when $\delta^* = \left[2^{m/2}\Gamma(m/2 + 1)\exp(t^*/2)\right]^{1/m}$. Table 1 shows the values of δ^* obtained

Table 2
Agreement between the p-value and the infimum of the posterior probability when $t = T(\bar{\mathbf{x}}, \boldsymbol{\theta}_0) = 7$

$t = 7$	$m = 2$	$m = 5$	$m = 10$	$m = 20$
$p(t)$	0.0302	0.22064	0.72544	0.9967
$\inf_{\pi \in Q_{us}} P(H_0 \mid t = 7, \delta^*)$	$m = 2$	$m = 5$	$m = 10$	$m = 20$
$p^* = 0.5$	0.0787	0.23133	0.76331	0.99791
$p^* = 0.1$	0.2992	0.77584	0.98894	0.99998
$p^* = 0.05$	0.46066	0.89648	0.99651	0.9999995
$p^* = 0.01$	0.81027	0.98473	0.9997	0.9999998
$p^* = 0.001$	0.97712	0.99897	0.999988	0.9999999

Table 3
Values of $\delta^*(p^*, n, m) = \delta\sqrt{n}/\sigma^2$ where the agreement is achieved over scale mixture of normal prior class

δ^*	$m = 5$	$m = 10$	$m = 15$	$m = 20$
$p^* = 0.5$	2.38	2.72	2.95	3.13
$p^* = 0.1$	5.14	5.77	6.21	6.55
$p^* = 0.05$	7.15	7.96	8.52	8.97
$p^* = 0.01$	15.51	17.01	18.05	18.89
$p^* = 0.001$	48.02	50.63	54.3	56.48

for different values of m and p^* and both methods are compared when $t = T(\bar{\mathbf{x}}, \boldsymbol{\theta}_0) = 7$. We compute the p-values, and the infimum of the posterior probabilities over the values of δ^* given in Table 1, when $m = 2$, $m = 5$, $m = 10$ and $m = 20$. For instance, when $m = 2$, the p-value is $p(7) = 0.0302$. Therefore, a frequentist statistician who uses $p^* = 0.05$, rejects H_0, whereas with $p^* = 0.01$, accepts H_0. The same result is obtained by a Bayesian with $\delta^* = 6.32$ and $\delta^* = 14.14$. In this case, $\inf_{\pi \in Q_{us}} P(H_0 \mid t = 7, \delta^* = 6.32) = 0.46066$ and $\inf_{\pi \in Q_{us}} P(H_0 \mid t = 7, \delta^* = 14.14) = 0.81027$.

3.3. Lowers bounds for scale mixture of normals

In the same context of Section 3.2, we want to test (1) with $\boldsymbol{\theta}_0 = (0, \ldots, 0)'$ and a sample of size n. The usual test statistic is $T(\overline{\mathbf{X}}) = n\sigma^{-2}\overline{\mathbf{X}}'\overline{\mathbf{X}}$. If $m > 2$ the infimum of the posterior probability of H_0 when $\pi(\boldsymbol{\theta}) \in Q_N$, $Q_N = \left\{\pi\left(\theta \mid v^2\right) \approx N_m\left(\mathbf{0}, v^2\mathbf{I}\right), \pi\left(v^2\right) \text{ non-decreasing on } (0,\infty)\right\}$, is

$$\inf_{\pi \in Q_N} P(H_0 \mid t) = \left[1 + \delta^{*-2} F_{m-2}(t) / f_m(t)\right]^{-1},$$

$t = n\bar{\mathbf{x}}'\bar{\mathbf{x}}/\sigma^2$, $\delta^* = \delta\sqrt{n}/\sigma^2$, F_{m-2} is the distribution function of a χ^2_{m-2} and f_m is the density function of a χ^2_m (see Gómez-Villegas et al., 2007). Let t^* be such that $P\left\{\chi^2_m \geq t^*\right\} = p^*$. By applying Corollary 2.2 both methods always agree with each other when $\delta^*(p^*, n, m) = \left[F_{m-2}(t^*)/f_m(t^*)\right]^{1/2}$. Table 2 shows the values of δ^* computed for different values of m and p^* and both methods are compared when $t = T(\bar{\mathbf{x}}) = 20$. The p-values and the infimums on the values of δ^* of Table 3 are computed. The p-value is $p(20) = 0.02925$ when $m = 10$. Therefore, a frequentist statistician who uses $p^* = 0.05$, rejects H_0, whereas with $p^* = 0.01$, accepts H_0. The same result is reached by a Bayesian by using, respectively, the values $\delta^* = 7.96$ and $\delta^* = 17.01$ given in Table 2. In this case $\inf_{\pi \in Q_{us}} P(H_0 \mid t = 20, \delta^* = 7.96) = 0.3772$ and $\inf_{\pi \in Q_{us}} P(H_0 \mid t = 7, \delta^* = 14.14) = 0.7343$.

4. Conclusions and comments

An important conclusion is that p-values and posterior probabilities can be reconciled in the multivariate point null testing problem. The methodology consists of assigning a prior mass to $\boldsymbol{\theta}_0$ computed by the probability of a sphere of radius δ centered at $\boldsymbol{\theta}_0$ assigned by a density $\pi(\theta)$. The p-value and the posterior probability of the null for the mixed prior $\pi^*(\boldsymbol{\theta}) = \pi_0 I_{H_0}(\boldsymbol{\theta}) + (1 - \pi_0)\, \pi(\boldsymbol{\theta}) I_{H_1}(\boldsymbol{\theta})$ are computed. This procedure allows one to prove a theorem that shows when and how both approaches are in agreement. The analyzed examples show that such agreement is always possible when $\delta^* = \delta^*(p^*, n, m) = \delta\sqrt{n}/\sigma^2$. This is due to the fact that the infimum of the posterior probability over the prior classes used depends on the usual test statistic through an increasing function. When this dependence is not possible, the agreement is in terms of a sufficient condition.

References

Berger, J.O., Delampady, M., 1987. Testing precise hypotheses, (with discussion). Statist. Sci. 2 (3), 317–352.

Berger, J.O., Sellke, T., 1987. Testing a point null hypothesis: The irreconcilability of p-values and evidence, (with discussion). J. Amer. Statist. Assoc. 82, 112–139.

Casella, G., Berger, R.L., 1987. Reconciling bayesian and frequentist evidence in the one-sided testing problem, (with discussion). J. Amer. Statist. Assoc. 82, 106–135.

Gómez-Villegas, M.A., González-Pérez, B., 2005. Bayesian analysis of contingency tables. Commun. Stat.-Theory Methods 34 (8), 1743–1754.
Gómez-Villegas, M.A., González-Pérez, B., 2006. A condition to obtain the same decision in the homogeneity testing problem from the frequentist and Bayesian point of view. Commun. Stat.-Theory Methods 35, 2211–2222.
Gómez-Villegas, M.A., Maín, P., Sanz, L., 2002. A suitable Bayesian approach in testing point null hypothesis: Some examples revisited. Commun. Stat.-Theory Methods 31 (2), 201–217.
Gómez-Villegas, M.A., Maín, P., Sanz, L., 2007. Contraste de hipótesis nula puntual multivariante para variables normales correladas. Actas del XXX Congreso Nacional de Estadística e Investigación Operativa y IV Jornadas de Estadística Pública 2007.
Gómez-Villegas, M.A., Sanz, L., 2000. ε-contaminated priors in testing point null hypothesis: A procedure to determine the prior probability. Statist. Probab. Lett. 47, 53–60.
Lindley, D.V., 1957. A statistical paradox. Biometrika 44, 187–192.
Oh, H.S., Dasgupta, A., 1999. Comparison of the p-value and posterior probability. J. Stat. Plan. Inference 76, 93–107.
Sellke, T., Bayarri, M.J., Berger, J.O., 2001. Calibration of p-values for testing point null hypotheses. Amer. Stat. 55, 62–71.