Unimodal contaminations in testing point null hypothesis

Miguel A. Gómez–Villegas[1] and Luis Sanz

Departamento de Estadística e Investigación Operativa I,

Universidad Complutense de Madrid. 28040 Madrid, Spain.

Abstract

The problem of testing a point null hypothesis from the Bayesian perspective is considered. The uncertainties are modelled through use of ε–contamination class with the class of contaminations including: *i*) All unimodal distributions and *ii*) All unimodal and symmetric distributions. Over these classes, the infimum of the posterior probability of the point null hypothesis is computed and compared with the p–value and a better approach than the one known is obtained.

Key Words: ε–contaminated class, point null hypothesis, posterior probability, p–values.

AMS subject classification: 62F15, 62A15

Resumen

Contaminaciones unimodales en el contraste de hipótesis nula puntual

Se considera el problema del contraste de hipótesis nula puntual desde el punto de vista Bayesiano. La incertidumbre se modeliza mediante el uso de la clase de las distribuciones ε–contaminadas, cuando la clase de las contaminaciones incluye: *i*) todas las distribuciones unimodales y *ii*) todas las distribuciones unimodales y simétricas. Se calcula el ínfimo de las probabilidades a posteriori de la hipótesis nula

[1]Correspondence to: Departamento de Estadística e Investigación Operativa I, Facultad de Ciencias Matemáticas, Universidad Complutense de Madrid, 28040 Madrid, Spain, E-mail:ma.gv@mat.ucm.es

puntual sobre estas clases y se compara con el p–valor, obteniéndose una aproximación aceptable entre ambos valores.

1. Introduction

1.1. The choice of the prior

A random variable, X, having density $f(x-\theta)$ is observed, being θ an unknown real parameter. To perform a Bayesian analysis concerning the parameter θ, it is necessary to express the prior beliefs about θ through a prior distribution of probability. Usually, the prior information can not be exactly quantified in terms of a single prior distribution. Perhaps, after an elicitation process, we can conclude that $\pi_0(\theta)$ represents our prior beliefs, but it looks reasonable that any prior not too far from $\pi_0(\theta)$ would also be a good approximation to our prior beliefs.

This is the reason why often a class of prior distributions is used instead of a concrete prior distribution. In this paper, following the reasoning above, we will use the ε–contamination class given by

$$\Gamma = \{\pi = (1-\epsilon)\pi_0 + \epsilon q,\ q \in Q\} \tag{1}$$

where π_0 is the prior that one would use in a Bayesian analysis with only one prior distribution. The value of ε, with $0 \leq \epsilon \leq 1$, represents the amount of error that we want to introduce in π_0. And Q is the class of probability distributions that contaminates $\pi_0(\theta)$.

Gómez–Villegas and Sanz (2000) use the class (1) to compare p–values and posterior probabilities in the point null testing problem being Q the class of all probability distributions and they conclude that both values can match. The class of all probability distributions is attractive because it is easy to work with and it contains any prior near to π_0 but, on the other hand, this class contains many unreasonable distributions that are too far from π_0.

If $\pi_0(\theta)$, the base prior distribution, is unimodal it looks reasonable the choice of Q as the class of all unimodal distributions (with the same mode as π_0) or the class of all symmetric and unimodal distributions, specially if π_0 is also unimodal and symmetric. See Berger and Berliner (1986), Berger (1985), Sivaganesan and Berger (1989) and Berger (1994) for further information about the choice of Q.

We start, in Section 1, with the problem. Then, in 1.2, we introduce the procedure to make up the mixed prior distribution and in 1.3 a justification for this construction is provided. Section 2 compares the p–value with the infimum of the posterior probability when Q is the class of all unimodal distributions. In Section 3 the comparison is done when Q is the class of all unimodal and symmetric distributions. Finally, Section 4 contains some additional comments.

1.2. The problem

We consider the parametric point null testing problem

$$H_0^* : \theta = \theta_0 \quad \text{versus} \quad H_1^* : \theta \neq \theta_0, \tag{2}$$

based on observing a random variable, X, with density $f(x|\theta)$, $\theta \in \Re$, continuous in θ_0. We suppose, as usually, that the probability of $\theta = \theta_0$ is $p > 0$, in such a way that the prior information is given by a mixed distribution assigning mass p to the null hypothesis and spreading the remainder, $1 - p$, according to a density $\pi(\theta) \in \Gamma$ over $\theta \neq \theta_0$. However there is no rule to fix the value of p –usually $p = 0.5$–.

In many practical situations, it is not usual to test (2). We propose to replace (2) by the more realistic precise hypothesis

$$H_0 : \theta \in I_b \quad \text{versus} \quad H_1 : \theta \in I_b^c, \tag{3}$$

where $I_b = (\theta_0 - b, \theta_0 + b)$ and b is suitable "small" so that any value of $\theta \in I_b$ can be considered indistinguishable from θ_0. Examples of this replacement can be seen in Berger (1985), Berger and Delampady (1987) and Lee (1989) among others.

In the classical approach, (2) can be changed by (3) when the p–value in (2) is approximately the same as the p–value in (3). Berger and Delampady (1987) seek conditions under which both p–values are approximately equal. From Bayesian perspective, this can be done when the posterior probabilities of the null hypotheses are close or, equivalently, when the Bayes factor in (2) is similar to the Bayes factor in (3). A relation between (2) and (3) with regard to the Bayes factor is given by Gómez–Villegas and Gómez Sánchez–Manzano (1992). There it is shown that the Bayes factor in (3) converges to the Bayes factor in (2) when b goes to zero. A difference between the use of Bayes factor and posterior odds in this framework can be seen in Levine and Casella (1996).

Let us suppose that our prior distribution is $\pi(\theta) \in \Gamma$, with Γ defined by (1). In the point null testing problem, we need a mixed prior distribution

$$\pi^*(\theta) = pI_{\{\theta_0\}}(\theta) + (1-p)\pi(\theta)I_{\{\theta\neq\theta_0\}}(\theta) \tag{4}$$

where $I_A(\theta) = 1$ if $\theta \in A$ and $I_A(\theta) = 0$ if $\theta \in A^c$. Whereas in (3) it is sufficient to choose $\pi(\theta) \in \Gamma$. Then, what we propose is to choose the value of p, in the mixed distribution (4), as

$$p = \int_{|\theta-\theta_0|\leq b} \pi(\theta)d\theta. \tag{5}$$

From now on we will note (4) by $\pi^*(\theta, b)$ making the dependence of b explicit.

This construction is based on the assumption that $\pi(\theta)$ represents our prior beliefs about θ but, as it is not possible to test (2) with $\pi(\theta)$, we approach (2) by (3) choosing a convenient value of b.

In the same way of Berger and Sellke (1987), we seek to minimize $\Pr(H_0^*|x)$ over the class Γ in (1). A reason to take the infimum is that for a small infimum the null hypothesis must be rejected according to the interpretation of the p–value. More reasons can be seen in Berger and Sellke (1987). Besides, this development is similar to that of Casella and Berger (1987) who reconcile Bayesian and frequentist evidence in the one–sided testing problem and we are interested in making clear the reason for the discrepancy between both approaches in the point null testing problem.

There is a substantial amount of literature about the reconciliation between p–values and posterior probabilities, some important references, besides the ones mentioned above, are Edwards et al. (1963), Pratt (1965), Dickey and Lienz (1970), DeGroot (1974), Bernardo (1980), Ghosh and Mukerjee (1992), Berger, Boukai and Wang (1997), Gómez–Villegas and Sanz (1998), Mukhopadhyay and Das Gupta (1997), Marden (2000), Sellke, Bayarri and Berger (2001) and Gómez–Villegas, Maín and Sanz(2002).

1.3. Justification and notation

The choice of p, the mass assigned to the point null hypothesis, as in (5) is basic for posterior calculations. A way of justifying this construction is by using the Kullback–Leibler information measure, $\delta(\pi^*|\pi) = \int \pi(\theta)\ln(\pi(\theta)/\pi^*(\theta))\,d\theta$, as a measure of discrepancy between π and π^*. With our method, when b goes to zero $\delta(\pi^*|\pi)$ also goes to zero while if p is constant then $\delta(\pi^*|\pi)$ is constant too. The detail of this justification can be seen in Gómez–Villegas and Sanz (2000).

We denote the likelihood function by $f(x|\theta)$, which is considered as a function of θ for the observed value x. We assume that π_0, the base prior, is unimodal with mode θ_0 and density $\pi_0(\theta)$ and that q, the contamination, has density $q(\theta)$ both with respect to the Lebesgue measure. Thus, any $\pi \in \Gamma$ as in (1) has density

$$\pi(\theta) = (1-\varepsilon)\pi_0(\theta) + \varepsilon q(\theta). \tag{6}$$

The marginal distribution of X with respect to the prior $\pi \in \Gamma$ is denoted by $m(x|\pi)$. Assuming the existence of all quantities in the problem, we have

$$m(x|\pi) = (1-\varepsilon)m(x|\pi_0) + \varepsilon m(x|q), \tag{7}$$

therefore, if the posterior distributions $\pi_0(\theta|x)$ and $q(\theta|x)$ exist, the posterior distribution of θ given x with respect to π is given by $\pi(\theta|x) = \lambda(x)\pi_0(\theta|x) + (1-\lambda(x))q(\theta|x)$, where $\lambda(x) = (1-\varepsilon)m(x|\pi_0)/m(x|\pi)$.

The prior mass assigned to the null hypothesis results, from (5) and (6), $p = (1-\varepsilon)p_0 + \varepsilon q_0$, where

$$p_0 = \int_{|\theta-\theta_0|\leq b} \pi_0(\theta)\, d\theta \quad \text{and} \quad q_0 = \int_{|\theta-\theta_0|\leq b} q(\theta)\, d\theta. \tag{8}$$

A classical measure of evidence against the null hypothesis, which depends on the observations, is the p–value. If there exists an appropriate statistic $T(X)$ for testing (3), for example a sufficient statistic, the p–value of the sample point, x, is $p(x) = \sup_{\theta\in H_0} Pr(|T(X)| > |T(x)| \quad |\theta)$. In particular, for testing (2), the p–value takes the form $p(x) = Pr(|T(X)| > |T(x)| \quad |\theta_0)$.

2. Unimodal contaminations

In this section we consider the ε–contaminated class as in (1), with Q, the class of contaminations as

$$Q = \{\text{All unimodal distributions with the same mode as } \pi_0(\theta)\} \tag{9}$$

class which is particularly reasonable if the base prior, $\pi_0(\theta)$, is also unimodal.

In order to find the infimum of the posterior probability of H_0^* over the class Γ, with the class of contaminations as (9), it is sufficient to find it over the much smaller class

$$\Gamma_U = \{\pi = (1-\varepsilon)\pi_0 + \varepsilon q,\ q \,\text{is}\, U(\theta_0, \theta_0 + k) \,\text{or}\, U(\theta_0 - k, \theta_0),\ \text{for some}\, k > 0\}, \tag{10}$$

as it is shown in Sivaganesan and Berger (1989). Where $U(a, b)$ denotes the uniform distribution on the interval (a, b).

In the following theorem we obtain the infimum of the posterior probability of the point null hypothesis for the class (10) given $\pi^*(\theta, b)$ by (4) and p computed as in (5).

Theorem 2.1. *For the hypotheses (2), if we take an arbitrary prior distribution* $\pi(\theta) \in \Gamma_U$ *as in (10) and a mixed prior distribution as (4) with the mass assigned* *to the null hypothesis according to (5), then the infimum of the posterior probability of* H_0^* *is attained for a* $\pi(\theta)$ *with* k *given by*

$$k = \frac{\{(1-\varepsilon)p_0 - p^2\}a - (1-\varepsilon)b\, m(x|\pi_0)}{p(1-p)f(x|w)} \tag{11}$$

where

i) $a = \int_{\theta_0}^{\theta_0+k} f(x|\theta)\, d\theta$ *and* $w = \theta_0 + k$ *if* $q(\theta)$ *is* $U(\theta, \theta_0 + k)$

ii) $a = \int_{\theta_0-k}^{\theta_0} f(x|\theta)\, d\theta$ *and* $w = \theta_0 - k$ *if* $q(\theta)$ *is* $U(\theta_0 - k, \theta_0)$

Proof: Computing the infimum of the posterior probability of H_0^*

$$Pr(H_0^*|x) = \frac{f(x|\theta_0)}{f(x|\theta_0) + \dfrac{1-p}{p} m(x|\pi)} \tag{12}$$

is just like computing the supremum of $G(k) = (1-p)/p\, m(x|\pi)$ over the class Γ_U.

Assuming $b \leq k$, by (5), we have

$$p = \int_{|\theta-\theta_0|\leq b} \pi(\theta)\, d\theta = (1-\varepsilon)p_0 + \varepsilon \frac{b}{k} \tag{13}$$

with p_0 given by (8), while $m(x|q) = a/k$.

It must be noted that p depends on q through $q_0 = b/k$, so the infimum of $Pr(H_0^*|x)$ in q can be computed, by (7), as the supremum in $q \in Q$ of

$$G(k) = \frac{1-p}{p} \left\{ (1-\varepsilon)m(x|\pi_0) + \varepsilon \frac{a}{k} \right\} \tag{14}$$

with p given by (13).

In order to find the value of k for which $G(k)$ in (14) is maximized, we obtain $G'(k)$ and setting this equal to zero it is straightforward to verify that

$$(1-\varepsilon)m(x|\pi_0)b + kp(1-p)f(x|w) + \{p^2 - (1-\varepsilon)p_0\}a = 0. \tag{15}$$

Equation (15) will have solution $k \geq 0$ only if:

1. $p^2 - (1-\varepsilon)p < 0$, and
2. $|\{p^2 - (1-\varepsilon)p_0\}a| > (1-\varepsilon)b\, m(x|\pi_0)$.

Condition 1 is reasonable since we can think that p is close to p_0 and, then, 1 is equivalent to require $p < 1-\varepsilon$, and this looks reasonable if we are thinking that the amount of error, ε, is not large. Nevertheless, if $\varepsilon = 1$ or $p^2 - (1-\varepsilon)p_0 \geq 0$ then $G'(k)$ will be positive in any case and, therefore, the supremum will be achieved when k goes to ∞. Condition 2 is due to mathematical reasons.

From (15) we obtain (11) the value of k which maximizes $G(k)$ and, so, we will find $\inf_{\pi\in\Gamma_U} Pr(H_0^*|x)$. $\Box$

In order to see how theorem 2.1 works, the following example for the normal model is considered.

Example 2.1. Let us suppose that $X|\theta$ is $N(\theta,\sigma^2)$ distributed, with σ^2 known, and $\pi_0(\theta)$ is $N(\mu,\tau^2)$, with both parameters known. If $X_1,\ldots,X_n$ is a random sample of X, then $\overline{X}$ is $N(\theta,\sigma^2/n)$ distributed and $m(\overline{x}|\pi_0)$ is $N(\mu,\tau^2+\sigma^2/n)$.

Besides $p_0 = \Phi\{(\theta_0+b-\mu)/\tau\} - \Phi\{(\theta_0-b-\mu)/\tau\}$, Φ denoting the standard cumulative distribution function.

Table 1 shows, choosing $\sigma^2=1$, $\tau^2=2$, $\theta_0=\mu=0$ and $n=10$, for some specific values of $t=\sqrt{n}(\overline{x}-\theta_0)/\sigma$ and b, the values of k, from (11), in which the infimum of the posterior probability of the point null hypothesis is achieved.

The last column in Table 1 is computed from

$$\inf_{\pi\in\Gamma_U} Pr(H_0^*|\overline{x}) = \left\{1 + \frac{\sup_{\pi\in\Gamma_U} G(k)}{f(\overline{x}|\theta_0)}\right\}^{-1}. \tag{16}$$

Besides, we can get now the values of b and k, say b^* and k^*, so that the p–values and $\inf_{\pi\in\Gamma_U} Pr(H_0^*|\overline{x}) = \underline{Pr}(H_0^*|\overline{x})$ match. This could be done from the expression

$$p(\overline{x}) = \left\{1 + \frac{\sup_{\pi\in\Gamma_U} G(k)}{f(\overline{x}|\theta_0)}\right\}^{-1}, \tag{17}$$

and then the prior probability depends on the data but we can avoid it, replacing $p(\overline{x})$ by the significance level of the test, α. Moreover, the infimum of the posterior probability of H_0^* and the p–value are closed if the value chosen for b is close to b^*,

Table 1: *Values of k where the infimum of $Pr(H_0^*|\overline{x})$ is achieved for some values of t and b*

t	b	k	$\sup_{\pi\in\Gamma_U} G(k)$	$\underline{Pr}(H_0^*\|\overline{x})$
1.645	0.1	1.081	5.52499	0.05573
	0.2	1.112	2.58070	0.11217
	0.3	1.150	1.60304	0.16902
1.960	0.1	1.175	5.43849	0.03286
	0.2	1.201	2.54499	0.06770
	0.3	1.233	1.58374	0.10450
2.596	0.1	1.375	5.14862	0.00836
	0.2	1.394	2.41652	0.01764
	0.3	1.420	1.50861	0.02664
3.291	0.1	1.598	4.74991	0.00118
	0.2	1.614	2.23486	0.00250
	0.3	1.632	1.39895	0.00399

since the infimum is a continuous function of b. Now, the infimum is attained when the contamination distribution, $q(\theta)$, is uniform in $U(\theta_0, \theta_0 + k^*)$ or in $U(\theta_0 - k^*, \theta_0)$.

Table 2 shows those values of b^* and k^* for some specific $t = \sqrt{n}(\overline{x} - \theta_0)/\sigma$. Besides, we see that $\underline{Pr}(H_0^*|\overline{x})$ is close to $\inf_{\pi\in\Gamma_U} Pr(H_0|\overline{x}) = \underline{Pr}(H_0|\overline{x})$.

Table 2: *Values of b and k that make equal the p–value with $\underline{Pr}(H_0|\overline{x})$, $\underline{Pr}(H_0^*|\overline{x})$, infimum over Γ_U*

t	b^*	k^*	$\underline{Pr}(H_0^*\|\overline{x})$ = p–value	$\underline{Pr}(H_0\|\overline{x})$	$\underline{Pr}(H_0^*\|\overline{x}, p = 0.5)$
1.645	0.1785	1.122	0.1	0.10298	0.45062
1.960	0.1498	1.180	0.05	0.05310	0.32957
2.596	0.1184	1.381	0.01	0.01081	0.11477
3.291	0.0855	1.589	0.001	0.00107	0.01868

Table 2 shows, too, that if the prior mass assigned to the null hypothesis is $p = 0.5$, the infimum of the posterior probability of H_0^* is much more larger than the p–value, but if the values of b in (5) are close to b^*, then Bayesian and classical approaches are numerically close. Moreover, table 2 shows that if $p = 0.5$, the posterior probability of the point null hypothesis disagree with the posterior probability of the interval hypothesis.

3. Unimodal and symmetric contaminations

If the base prior, π_0, is unimodal and symmetric it may be reasonable to require that the contaminations be also unimodal and symmetric. Thus, in this section, we deal with the class (1) and Q, the class of contaminations, will be

$$Q = \{\text{All symmetric unimodal distributions with the same mode as}\, \pi_0\}$$

This ε–contaminated class will be denoted Γ_{US}. Sivaganesan and Berger (1989) prove that in order to find the infimum of the posterior probability of H_0^* over the class Γ_{US} it is sufficcient to do it over the class

$$\Gamma_{US} = \{\pi = (1-\varepsilon)\pi_0 + \varepsilon q,\, q \in U(\theta_0 - k, \theta_0 + k)\, \text{for some}\, k > 0\}, \qquad (18)$$

based on representing a symmetric unimodal density as a mixture of symmetric uniforms.

Theorem 3.1 shows the value of k for the uniform distribution where the infimum is attained.

Theorem 3.1. *For the hypotheses (2), if $\pi(\theta) \in \Gamma_{US}$ as in (18) and a mixed prior distribution as (4), with the mass assigned to the null by (5), is used, the infimum of the posterior probability of H_0^* is attained for a $\pi(\theta)$ with k given by*

$$k = \frac{\{(1-\varepsilon)p_0 - p^2\}a - 2(1-\varepsilon)bm(x|\pi_0)}{p(1-p)\{f(x|\theta_0+k) + f(x|\theta_0-k)\}} \qquad (19)$$

where $a = \int_{\theta_0-k}^{\theta_0+k} f(x|\theta)\, d\theta$

Proof: The infimum of $Pr(H_0^*|x)$ over Γ_{US} will be obtained, as in theorem 2.1, computing the supremum of $G(k) = (1-p)/pm(x|\pi)$.

Assuming $b \leq k$, the prior mass, p, assigned to the null hypothesis is, given by (5), now

$$p = (1-\varepsilon)p_0 + \varepsilon\frac{b}{k} \tag{20}$$

The marginal distribution of X given π is, by (7),

$$m(x|\pi) = (1-\varepsilon)m(x|\pi_0) + \varepsilon\frac{a}{2k} \tag{21}$$

Then, by (21) and (20)

$$G(k) = \frac{k(1-\varepsilon)m(x|\pi_0) + \varepsilon a/2}{k(1-\varepsilon)p_0 + \varepsilon b} - (1-\varepsilon)m(x|\pi_0) - \varepsilon\frac{a}{2k}, \tag{22}$$

Expression (22) depends only on k for b fixed and x observed. Then, by differentiating (22), results

$$(1-\varepsilon)m(x|\pi_0)b + \frac{1}{2}kp(1-p)\{f(x|\theta_0+k) + f(x|\theta_0-k)\} + \frac{1}{2}\{p^2 - (1-\varepsilon)p_0\}a = 0$$

and (19) is obtained. □

Theorem 3.1 gives us the value of k which maximizes $G(k)$ and then the value that minimizes the posterior probability of H_0^*.

Example 3.1. Consider a random variable X with $N(\theta, \sigma^2)$ distribution, σ^2 known, and the base prior $\pi_0(\theta)$ is $N(\mu, \tau^2)$ for μ and τ^2 given. If $X_1, \ldots, X_n$ is a random sample of size n, the sample mean, $\overline{X}$, is $N(\theta, \sigma^2/n)$ distributed and $m(x|\pi_0)$ is $N(\mu, \tau^2 + \sigma^2/n)$. Moreover, the prior mass assigned to H_0^*, p, is given by (20) with $p_0 = \Phi\{(\theta_0 + b - \mu)/\tau\} - \Phi\{(\theta_0 - b - \mu)/\tau\}$.

Then, equation (19) gives solutions in k for fixed values of $\overline{x}$ and b. Table 3 shows the values of k for which the infimum of the posterior probability of the point null hypothesis is attained for some especific values of $t = \sqrt{n}|\overline{x} - \theta_0|/\sigma$ and different values of b. Calculus are done for $\varepsilon = 0.2$, $\sigma^2 = 1$, $\tau^2 = 2$, $\theta_0 = \mu = 0$ and $n = 10$.

The last column in Table 3 is obtained from

$$\inf_{\pi \in \Gamma_{US}} Pr(H_0^*|\overline{x}) = \left(1 + \frac{\sup_{\pi \in \Gamma_{US}} G(k)}{f(\overline{x}|\theta_0)}\right)^{-1}.$$

Table 3: *Values of k where the infimum of $Pr(H_0^*|\overline{x})$ is achieved for some values of t and b*

p–value	t	b	k	$\sup_{\pi\in\Gamma_{US}} G(k)$	$\underline{Pr}(H_0^*\|\overline{x})$
0.1	1.645	0.1	1.570	4.40095	0.06898
		0.2	$+\infty$	2.08649	0.13503
		0.3	$+\infty$	1.32996	0.19689
0.05	1.960	0.1	1.546	4.31036	0.04111
		0.2	5.073	2.03466	0.08327
		0.3	$+\infty$	1.29449	0.12493
0.01	2.596	0.1	1.634	4.08511	0.01051
		0.2	1.782	1.92363	0.02206
		0.3	4.477	1.21042	0.03462
0.001	3.291	0.1	1.785	3.78266	0.00148
		0.2	1.846	1.78289	0.00314
		0.3	1.963	1.11884	0.00499

A couple of observations can be made about the results shown in Table 3. Firstly, the distribution $q(\theta) \in Q$, for which the infimum is attained, depends on the value fixed for b and the observation $\overline{x}$, and in some cases the infimum is attained for the uniform improper contamination. Secondly, if b takes moderate small values, the infimum of the posterior probability of H_0^* is close to the p–value independently of the observed value $\overline{x}$, so if $t = 1.96$ and $b \in (0.1, 0.2)$ the infimum of the posterior probability is in $(0.04111, 0.08327)$, close to the p–value, 0.05.

As it happened in the case of unimodal contaminations, it is possible to determine a value of b, say b^*, such that the p–value and the infimum of the posterior probability of H_0^* match. Table 4 shows the values of b^* and the respectives of k, say k^*.

For example, if $t = 1.96$ and $b^* = 0.12$ is chosen in (5), the infimum of the posterior probability is attained for the uniform distribution in the interval $(\theta_0 - 1.6, \theta_0 + 1.6)$ and this infimum and the p–value match, whereas if the posterior probability of the point null is computed with $p = 0.5$, as it is usually done in the literature, gives 0.38817 which is too far from the p–value.

Table 4: *Values of b and k that make equal the p–value with $\underline{Pr}(H_0|\overline{x})$, $\underline{Pr}(H_0^*|\overline{x})$, infimum over Γ_{US}*

t	b^*	k^*	$\underline{Pr}(H_0^*\|\overline{x})$ = p–value	$\underline{Pr}(H_0\|\overline{x})$	$\underline{Pr}(H_0\|\overline{x}, p = 0.5)$
1.645	0.1461	3.85	0.1	0.10811	0.51364
1.960	0.1213	1.60	0.05	0.05115	0.38817
2.596	0.0954	1.63	0.01	0.01014	0.14231
3.291	0.0688	1.77	0.001	0.00104	0.02357

Table 4 shows too the values of the infimum of the posterior probability of the interval hypothesis, H_0, $\underline{Pr}(H_0|\overline{x}) = \int_{|\theta-\theta_0|\leq b} \pi(\theta|\overline{x})\, d\theta$, where $\pi(\theta|\overline{x})$ is given by (12), then

$$\underline{Pr}(H_0|\overline{x}) = \frac{1-\varepsilon}{m(\overline{x}|\pi)} \int_{|\theta-\theta_0|\leq b} f(\overline{x}|\theta)\pi_0(\theta)\, d\theta + \frac{\varepsilon}{m(\overline{x}|\pi)} \int_{|\theta-\theta_0|\leq b} \frac{f(\overline{x}|\theta)}{2k}\, d\theta,$$

with $f(\overline{x}|\theta)$ the density of the $N(\overline{x}, \sigma^2/n)$ distribution and $m(\overline{x}|\pi)$ is given by (7). □

4. Comments

As it is shown, the ε–contaminated class allows an acceptable Bayesian approach, both analytical and intuitive, to the problem of testing point null hypothesis.

The procedure to determine the prior mass assigned to the point null hypothesis using a small interval of length $2b$, centered in θ_0, and to compute the probability assigned by $\pi(\theta)$ to this interval, allows us to obtain values of the posterior probability of the point null hypothesis that are closer to the p–value.

Moreover, the case in which the prior mass assigned to the point null hypothesis is 0.5 is a particular case for some value of b. In other words, there is a value of b for which the prior mass assigned to the point null is 0.5.

Then, the difference between the p–value and the posterior probability for the problem of testing a point null hypothesis is not due to using a mixed prior distribution but rather to the choice of the prior mass, p, in the mixed distribution, usually $p = 0.5$. Small values than this, depending on the sample model, allow us a better approximation between the p–value and the posterior probability.

Acknowledgements

This paper was supported by a grant from Ministerio de Ciencia y Tecnología (Spain), number BFM2003–04279.

References

1. Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis.* Springer Verlag, New York.

2. Berger, J.O. (1994). An overview over Robust Bayesian analysis (with discussion). *Test,* 3:5–124.

3. Berger, J.O. and Berliner, L.M. (1986). Robust Bayes and empirical Bayes analysis with ε–contaminated priors. *Ann. Statist.*,14, 2:461–486.

4. Berger, J.O., Boukai, B. and Wang, Y. (1997). Unified frequentist and Bayesian testing of a precise hypothesis, (with discussion). *Statistical Science*, 12, 3:133–160.

5. Berger, J.O. and Delampady, M. (1987). Testing precise hyphoteses. *Statistical Science*, 2, 3:217–352.

6. Berger, J.O. and Sellke, T. (1987). Testing a point null hyphoteses: The irreconciliability of p values and evidence, (with discussion). *J. Amer. Statist. Assoc.* 82:112–122.

7. Bernardo, J.M. (1980). A Bayesian analysis of classical hypothesis testing (with discussion). In J. M. Bernardo, M. H. DeGroot, D. V. Lindley, A. F. M. Smith, eds., *Bayesian Statistics, 1*, pp. 605–647. University Press, Valencia.

8. Casella, G. and Berger, R.L. (1987). Reconciling Bayesian and frequentist evidence in the One-Sided Testing Problem, (with discussion). *J. Amer. Statist. Assoc.* 82:106–111.

9. DeGroot, M.H. (1974). Reaching a consensus. *J. Amer. Statist. Assoc.* 68:966–969.

10. Dickey, J.M. and Lienz, B.P. (1970). The weighted likelihood ratio, sharp hypotheses about chances, the order of a Markov chain. *Ann. Math. Statist.* 41:214–226.

11. Edwards, W., Lindman, H. and Savage, L.J. (1963). Bayesian statistical inference for psychological research. Psychol. Rev. 70, 193–242. Reprinted in J. B. Kadane, eds., 1984 *Robustness of Bayesian Analysis*, pp. 1–62 North–Holland, Amsterdam.

12. Ghosh, J.K. and Mukerjee, R. (1992). Non–informative priors (with discussion). In J.M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds., *Bayesian Statistics 4,*, pp. 195–210. University Press, Oxford.

13. Gómez–Villegas, M.A. and Gómez Sánchez-Manzano, E.(1992). Bayes Factor in Testing Precise Hyphoteses. *Commun. Statist. - Theory Meth.* 21(6): 1707-1715.

14. Gómez–Villegas, M.A. and Sanz, L. (1998). Reconciling Bayesian and frequentist evidence in the point null testing problem. *Test 7,* 1:207–216.

15. Gómez–Villegas, M.A. and Sanz, L. (2000). ε–contaminated priors in testing point null hypothesis: a procedure to determine the prior probability. *Statist. and Probability Letters* 47, 1:53–60.

16. Gómez–Villegas, M.A., Maín, P. and Sanz, L. (2002). A suitable Bayesian approach in testing point null hypothesis: some examples revisited. *Commun. Statist.– Theory Meth.*, 31(2):201–217.

17. Jeffreys, H. (1967). *Theory of Probability.* 3rd ed. Oxford Univ. Press, London.

18. Lee, P.M. (1994). *Bayesian Statistics: An Introduction.* Charles Griffin, London.

19. Levine, R.A. and Casella, G. (1996). Convergence of posterior odds. *Journal of Statistical Planning and Inference*, 55:231–344.

20. Marden, J.I. (2000). Hypothesis testing: from p–values to Bayes factors. *J. Amer. Statist. Assoc.*, 82:112–122.

21. Mukhopadhyay, S.and Dasgupta, A. (1997). Uniform approximation of Bayes solutions and posteriors: frequentistly valid Bayes inference. *Statistics and Decisions* 15:51–7.

22. Pratt, J.W. (1965). Bayesian interpretation of standard inference statements. *J. Roy. Statist. Soc. B* 27:169–203.

23. Sellke, T., Bayarri, M.J. and Berger, J.O. (2001) Calibration of p Values for Testing Precise Null Hypotheses. *The American Statistician,* 55(1):62–71.

24. Sivaganesan, S. and Berger, J.O. (1989). Ranges of posterior measures for priors with unimodal contaminations. *Ann. Statist 17,* 2:868–889.