Statistics and Econometrics Series 16
March 2001

Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

COHERENCE OF THE POSTERIOR PREDICTIVE P-VALUE BASED ON THE POSTERIOR ODDS.

J. de la Horra Navarro and M.T. Rodríguez Bernal*

Abstract

It is well-known that classical p-values sometimes behave incoherently for testing hypotheses in the sense that, when $\Theta_0 \subset \Theta_0^{'}$, the support given to Θ_0 is greater than or equal to the support given to $\Theta_0^{'}$. This problem is also found for posterior predictive p-values (a Bayesian-motivated alternative to classical p-values). In this paper, it is proved that, under some conditions, the posterior predictive p-value based on theposterior odds is coherent, showing that the choice of a suitable discrepancy variable is crucial.

Keywords: Coherence, posterior odds, p-value, posterior predictive p-value.

*de la Horra, Departamento de Matemáticas, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Tfno: 91-3974945, e-mail: julian.delahorra@uam.es ; Rodríguez, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid, Calle Madrid 126, 28903 Getafe (Madrid), Spain, Tfno: 91-6245853, e-mail: mrodrigu@est-econ.uc3m.es.This research was partially supported by DGESIC (Spain) under grant PB97-0021.

1 Introduction

In the last years, several papers have been published analyzing possible alternatives to the classical p-value. The posterior predictive p-value is a Bayesian-motivated alternative. The concept was first introduced by Guttman (1967) and Rubin (1984), who used the posterior predictive distribution of a test statistic to calculate the tail-area probability corresponding to the observed value of the statistic. Such a tail-area probability is called posterior predictive p-value by Meng (1994) (who extended the concept by using discrepancy variables), whereas the tail-area probability used by Box (1980) can be called prior predictive p-value. More recently, Bayarri and Berger (1999) have introduced the conditional predictive p-value and the partial posterior predictive p-value.

The asymptotic behaviour of the posterior predictive p-value was studied by De la Horra and Rodríguez-Bernal (1997). Different aspects of the application of the posterior predictive p-value to the problem of goodness of fit were analyzed by Gelman et al. (1996) and De la Horra and Rodríguez-Bernal (1999).

The concept of posterior predictive p-value is briefly introduced in Section 2. The concept of coherence of a p-value is introduced in Section 3. Schervish (1996) showed that the classical p-value could behave incoherently as a measure of support for hypotheses. A similar problem was pointed out by Lavine and Schervish (1999) for Bayes factors and by De la Horra and Rodríguez-Bernal (2001) for the posterior predictive p-value.

The posterior predictive p-value based on the posterior odds is considered in Section 4. De la Horra and Rodríguez-Bernal (2000) proved the asymptotic optimality of this posterior predictive p-value. In this paper, it is proved that, under some conditions, this posterior predictive p-value behaves coherently as a measure of support.

Finally, some examples are given in Section 5.

2 Posterior predictive p-value

Let x be an observation from the random variable X taking values in $\mathcal{X}$ and having density function $f(x|\theta)$, where $\theta \in \Theta$. We want to test $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1 = \Theta - \Theta_0$.

The posterior predictive p-value is a Bayesian-motivated alternative to the classical p-value introduced by Guttman (1967) and Rubin (1984) and extended by Meng (1994). Let $\pi(\theta)$ be the prior density summarizing the prior information about θ, and let $D(x,\theta)$ be a discrepancy variable, where a discrepancy variable is a function $D : \mathcal{X} \times \Theta_0 \to \mathbb{R}^+$ measuring (in some reasonable way) the "discrepancy" between the observation x and the parameter θ. The concept of discrepancy variable $D(x,\theta)$ was introduced by Tsui and Weerahandi (1989) and is nothing but a generalization of a test statistic $D(x)$. In fact, the posterior odds we will use in Section 4 is simply a test statistic.

The well-known classical p-value for testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$, when the discrepancy variable $D(x,\theta_0)$ is used, will play an important role in this paper:

$$p(x,\theta_0) = Pr\{y \in \mathcal{X} : D(y,\theta_0) \geq D(x,\theta_0)|\theta_0\} = \int_{A_{\theta_0}} f(y|\theta)dy,$$

where $A_{\theta_0} = \{y \in \mathcal{X} : D(y,\theta_0) \geq D(x,\theta_0)\}$.

We can now give the definition of posterior predictive p-value, such as it was introduced by Meng (1994):

Definition 1. *The posterior predictive p-value for testing $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1 = \Theta - \Theta_0$, when the discrepancy variable $D(x,\theta)$ is used, is defined as*

$$\begin{aligned} p(x,\Theta_0) &= Pr\{(y,\theta) \in \mathcal{X} \times \Theta_0 : D(y,\theta) \geq D(x,\theta)|x,\Theta_0\} \\ &= \int_A f(y,\theta|x,\Theta_0)dyd\theta, \end{aligned}$$

where $A = \{(y,\theta) \in \mathcal{X} \times \Theta_0 : D(y,\theta) \geq D(x,\theta)\}$.

We will need to express $p(x,\Theta_0)$ in an alternative way:

$$
\begin{aligned}
p(x,\Theta_0) &= \int_A f(y,\theta|x,\Theta_0)dyd\theta \\
&= \int_A f(y|\theta)\pi(\theta|x,\Theta_0)dyd\theta \\
&= \int_{\Theta_0}\left[\int_{A_\theta} f(y|\theta)dy\right]\pi(\theta|x,\Theta_0)d\theta \\
&= \int_{\Theta_0} p(x,\theta)\pi(\theta|x,\Theta_0)d\theta \\
&= \int_{\Theta_0} p(x,\theta)\frac{\pi(\theta|x)}{Pr(\Theta_0|x)}d\theta \\
&= \frac{\int_{\Theta_0} p(x,\theta)\pi(\theta|x)d\theta}{\int_{\Theta_0}\pi(\theta|x)d\theta}.
\end{aligned}
$$

In this way, $p(x,\Theta_0)$ is a ratio between the area under the posterior density $\pi(\theta|x)$, weighted by $p(x,\theta)$, over Θ_0, and the area under $\pi(\theta|x)$ over Θ_0. Therefore, the classical p-value $p(x,\theta)$ will play the role of a weight function.

3 Coherence

P-values are usually interpreted as a measure of support in favour of the null hypothesis $H_0 : \theta \in \Theta_0$. Schervish (1996) used the following definition of coherence for p-values:

Definition 2. *A measure of support for hypotheses is coherent if, when* $\Theta_0 \subset \Theta_0'$, *the support given to* Θ_0' *is greater than or equal to the support given to* Θ_0.

Schervish (1996) showed, through some examples, that the classical p-value could behave incoherently. A similar problem was pointed out by Lavine and Schervish (1999) for Bayes factors and by De la Horra and Rodríguez-Bernal (2001) for posterior predictive p-values.

In all the examples analyzed by De la Horra and Rodríguez-Bernal (2001) (where incoherences with posterior predictive p-values were detected), the discrepancy variable considered was $D(x,\theta) = |x-\theta|$. In principle, that seemed a natural choice, because $D(x,\theta) = |x-\theta|$ is the usual discrepancy variable used for testing the value of the mean of Normal observations (as in those examples), but incoherences may perhaps disappear if a Bayesian-motivated discrepancy variable is used.

The following lemma gives a necessary and sufficient condition for coherence and will be used in Section 4:

Lemma 1. *Let us consider* Θ_0 *and* Θ_0', *where* $\Theta_0 \subset \Theta_0'$. *Then,* $p(x, \Theta_0) \leq p(x, \Theta_0')$ *if and only if*

$$p(x, \Theta_0) \leq \frac{\int_{\Theta_0' - \Theta_0} p(x, \theta)\pi(\theta|x)d\theta}{\int_{\Theta_0' - \Theta_0} \pi(\theta|x)d\theta}$$

Proof. For any x, $p(x, \Theta_0) \leq p(x, \Theta_0')$ if and only if

$$\frac{\int_{\Theta_0} p(x, \theta)\pi(\theta|x)d\theta}{\int_{\Theta_0} \pi(\theta|x)d\theta} \leq \frac{\int_{\Theta_0} p(x, \theta)\pi(\theta|x)d\theta + \int_{\Theta_0' - \Theta_0} p(x, \theta)\pi(\theta|x)d\theta}{\int_{\Theta_0} \pi(\theta|x)d\theta + \int_{\Theta_0' - \Theta_0} \pi(\theta|x)d\theta}.$$

A little algebra leads to

$$\frac{\int_{\Theta_0} p(x, \theta)\pi(\theta|x)d\theta}{\int_{\Theta_0} \pi(\theta|x)d\theta} \leq \frac{\int_{\Theta_0' - \Theta_0} p(x, \theta)\pi(\theta|x)d\theta}{\int_{\Theta_0' - \Theta_0} \pi(\theta|x)d\theta},$$

that is,

$$p(x, \Theta_0) \leq \frac{\int_{\Theta_0' - \Theta_0} p(x, \theta)\pi(\theta|x)d\theta}{\int_{\Theta_0' - \Theta_0} \pi(\theta|x)d\theta}.$$

□

4 Main results

When we want to test $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1 = \Theta - \Theta_0$, a very natural discrepancy variable (from the Bayesian viewpoint) is the posterior odds:

$$D^*(x) = \frac{Pr(\Theta_1|x)}{Pr(\Theta_0|x).}$$

This discrepancy variable is simply a test statistic (because it does not depend on θ) and has a very clear Bayesian meaning.

De la Horra and Rodríguez-Bernal (2000) extended work by Thompson (1997) and proved that the posterior predictive p-value based on the posterior odds has good asymptotic properties.

In this section, the coherence of the posterior predictive p-value $p^*(x, \Theta_0)$ based on the posterior odds,

$$p^*(x, \Theta_0) = \frac{\int_{\Theta_0} p^*(x, \theta)\pi(\theta|x)d\theta}{\int_{\Theta_0} \pi(\theta|x)d\theta},$$

where $p^*(x, \theta) = Pr\{y \in \mathcal{X} : D^*(y) \geq D^*(x)|\theta\}$, will be studied.

First of all, we have to express the weight function $p^*(x, \theta)$ in an alternative way:

$$\begin{aligned} D^*(y) \geq D^*(x) \quad &\Longleftrightarrow \quad \frac{Pr(\Theta_1|y)}{Pr(\Theta_0|y)} \geq \frac{Pr(\Theta_1|x)}{Pr(\Theta_0|x)} \\ &\Longleftrightarrow \quad Pr(\Theta_0|y) \leq Pr(\Theta_0|x). \end{aligned} \tag{1}$$

Therefore:

$$p^*(x, \theta) = Pr\{y \in \mathcal{X} : Pr(\Theta_0|y) \leq Pr(\Theta_0|x)|\theta\}.$$

We can now prove that, under some conditions, the posterior predictive p-value based on the posterior odds behaves coherently.

Theorem 1. *Let x be an observation from the random variable X with density function $f(x|\theta)$, where $\theta \in \Theta \subset \mathbb{R}$. Let us assume that*

(i) $f(x|\theta)$ is a monotonically decreasing function in $|x - \theta|$,

(ii) The posterior density $\pi(\theta|x)$ is a monotonically decreasing function in $|\theta - g(x)|$, for some monotonically increasing function $g(x)$.

Then,

(a) For $\Theta_0 = (-\infty, a)$ and $\Theta'_0 = (-\infty, b)$ (with $a < b$), we have $p^*(x, \Theta_0) \leq p^*(x, \Theta'_0)$.

(b) For $\Theta_0 = (a, \infty)$ and $\Theta'_0 = (b, \infty)$ (with $b < a$), we have $p^*(x, \Theta_0) \leq p^*(x, \Theta'_0)$.

Proof. (a)

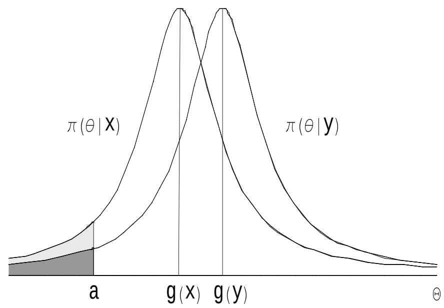

Figure 1

For $\Theta_0 = (-\infty, a)$, we have:

$$\begin{aligned} p^*(x,\theta) &= Pr\{y \in \mathcal{X} : Pr(\Theta_0|y) \leq Pr(\Theta_0|x)|\theta\} \\ &= Pr\{y \in \mathcal{X} : g(y) \geq g(x)|\theta\} \qquad (2) \\ &= Pr\{y \in \mathcal{X} : y \geq x|\theta\} \qquad (3) \end{aligned}$$

where:

(2) is easily seen in Figure 1, by $\pi(\theta|x)$ being a decreasing function in $|\theta - g((x)|$ (assumption (ii)),

(3) is true since $g(x)$ is an increasing function (assumption (ii)).

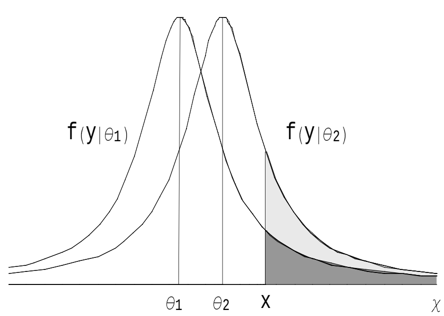

Figure 2

Moreover, $p^*(x,\theta) = Pr\{y \in \mathcal{X} : y \geq x|\theta\}$ is a monotonically increasing function in θ, by assumption (i) (see Figure 2).

This reasoning is also valid for $\Theta_0' = (-\infty, b)$ and, therefore, the weight function $p^*(x,\theta)$ will be the same for computing $p^*(x,\Theta_0)$ and $p^*(x,\Theta_0')$. We have:

$$\begin{aligned}
p^*(x,\Theta_0) &= \frac{\int_{\Theta_0} p^*(x,\theta)\pi(\theta|x)d\theta}{\int_{\Theta_0} \pi(\theta|x)d\theta} \\
&\leq \frac{\int_{\Theta_0} p^*(x,a)\pi(\theta|x)d\theta}{\int_{\Theta_0} \pi(\theta|x)d\theta} \\
&= p^*(x,a) = \frac{\int_{\Theta_0'-\Theta_0} p^*(x,a)\pi(\theta|x)d\theta}{\int_{\Theta_0'-\Theta_0} \pi(\theta|x)d\theta} \\
&\leq \frac{\int_{\Theta_0'-\Theta_0} p^*(x,\theta)\pi(\theta|x)d\theta}{\int_{\Theta_0'-\Theta_0} \pi(\theta|x)d\theta},
\end{aligned}$$

where the inequalities are true because $p^*(x,\theta)$ is an increasing function in θ. The result is now obtained by applying Lemma 1.

(b) The reasoning is analogous. □

Theorem 2. *Let x be an observation from the random variable X with density function $f(x|\theta)$, where $\theta \in \Theta \subset \mathbb{R}$. Let us assume that*

(i) $f(x|\theta)$ is a monotonically decreasing function in $|x-\theta|$,

(ii) The posterior density $\pi(\theta|x)$ is a monotonically decreasing function in $|\theta - cx|$, for some $c > 0$.

Then, for $\Theta_0 = (-a,a)$ and $\Theta_0' = (-b,b)$ (with $a < b$), we have $p^*(x,\Theta_0) \leq p^*(x,\Theta_0')$.

Proof.

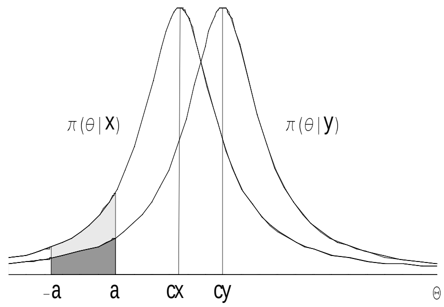

Figure 3

For $\Theta_0 = (-a, a)$, we have:

$$\begin{aligned} p^*(x,\theta) &= Pr\{y \in \mathcal{X} : Pr(\Theta_0|y) \le Pr(\Theta_0|x)|\theta\} \\ &= Pr\{y \in \mathcal{X} : |cy| \ge |cx|\,|\theta\} \qquad (4) \\ &= Pr\{y \in \mathcal{X} : |y| \ge |x|\,|\theta\} \end{aligned}$$

where (4) is readily seen in Figure 3, by assumption (ii).

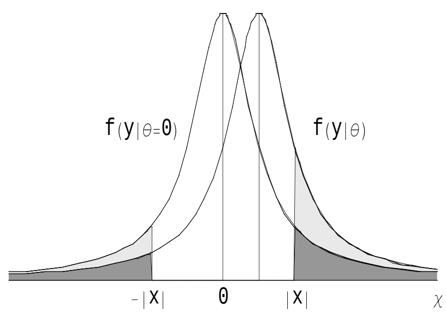

Figure 4

Moreover, $p^*(x,\theta) = Pr\{y \in \mathcal{X} : |y| > |x|\,|\theta\}$ is a monotonically increasing function in $|\theta|$, by assumption (i) (see Figure 4). Therefore, $p^*(x,\theta)$ is a function of the form shown in Figure 5.

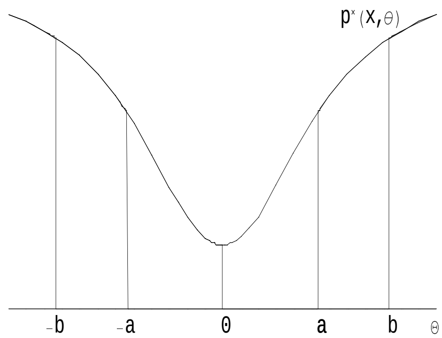

Figure 5

We remark that the reasoning is also valid for $\Theta_0' = (-b, b)$ and, therefore, the weight function $p^*(x,\theta)$ will be the same for computing $p^*(x,\Theta_0)$ and $p^*(x,\Theta_0')$.

We have:

$$
\begin{aligned}
p^*(x,\Theta_0) &= \frac{\int_{\Theta_0} p^*(x,\theta)\pi(\theta|x)d\theta}{\int_{\Theta_0} \pi(\theta|x)d\theta} \\
&\leq \frac{\int_{\Theta_0} p^*(x,a)\pi(\theta|x)d\theta}{\int_{\Theta_0} \pi(\theta|x)d\theta} = p^*(x,a) \\
&= \frac{\int_{\Theta_0'-\Theta_0} p^*(x,a)\pi(\theta|x)d\theta}{\int_{\Theta_0'-\Theta_0} \pi(\theta|x)d\theta} \\
&\leq \frac{\int_{\Theta_0'-\Theta_0} p^*(x,\theta)\pi(\theta|x)d\theta}{\int_{\Theta_0'-\Theta_0} \pi(\theta|x)d\theta},
\end{aligned}
$$

where the inequalities are true because $p^*(x,\theta)$ is an increasing function in $|\theta|$ (see Figure 5). The result is now obtained by applying Lemma 1. $\square$

Comments

a) It is important to remark why we have obtained coherence in these cases. The key in the proofs is the form of the weight function $p^*(x,\theta)$. For instance, the form of $p^*(x,\theta)$ in Theorem 2 is shown in Figure 5. If we would use the discrepancy variable $D(x,\theta) = |x-\theta|$ (as in De la Horra and

Rodríguez-Bernal (2001)) instead of $D^*(x) = \frac{Pr(\Theta_1|x)}{Pr(\Theta_0|x)}$ (as in this paper), it can be seen that $p(x,\theta)$ would be of the form shown in Figure 6, and the proof would not be possible.

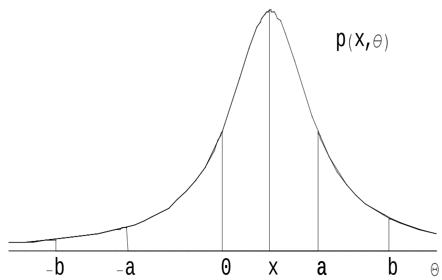

Figure 6

b) Assumption (i) in Theorem 1 includes a large class of location families that are common in statistical analysis. Assumption (ii) seems plausible with $g(x)$ identified as a *natural* point estimate.

c) The result in Theorem 2 can be translated to the case in which $\Theta_0 = (\theta_0 - a, \theta_0 + a)$ and $\Theta_0' = (\theta_0 - b, \theta_0 + b)$ (with $a < b$). We have just to rewrite the problem by considering the observation $X' = X - \theta_0$ (instead of X), and by considering the null hypotheses $(-a, a)$ (instead of $(\theta_0 - a, \theta_0 + a)$) and $(-b, b)$ (instead of $(\theta_0 - b, \theta_0 + b)$).

5 Examples

The following examples show some important cases in which Theorems 1 and 2 apply.

Example 1.

$$\left.\begin{array}{c} X \sim N(\theta, \sigma^2) \quad (\sigma^2 \ \mathit{known}) \\ \pi(\theta) \propto 1 \end{array}\right\} \Longrightarrow \pi(\theta|x) \sim N(x, \sigma^2).$$

Assumptions in Theorems 1 and 2 are fulfilled by taking $g(x) = x$ *and* $c = 1$.

We take $\sigma^2 = 1$ for obtaining the following tables (by simulation):

	$\Theta_0 = (-1, \infty)$		$\Theta_0 = (0, \infty)$		$\Theta_0 = (1, \infty)$	
x	$D^*(x)$	$p^*(x)$	$D^*(x)$	$p^*(x)$	$D^*(x)$	$p^*(x)$
-2	5.3030	0.0810	42.9558	0.0109	739.7970	0.0007
-1	1.0000	0.2404	5.3030	0.0788	42.9558	0.0114
0	0.1886	0.4091	1.0000	0.2619	5.3030	0.0786
1	0.0233	0.4860	0.1886	0.4193	1.0000	0.2375
2	0.0014	0.5071	0.0233	0.4833	0.1886	0.4159

	$\Theta_0 = (-1.5, 1.5)$		$\Theta_0 = (-1, 1)$		$\Theta_0 = (-0.5, 0.5)$	
x	$D^*(x)$	$p^*(x)$	$D^*(x)$	$p^*(x)$	$D^*(x)$	$p^*(x)$
0.1	0.1568	0.9372	0.4700	0.9301	1.6235	0.9230
0.2	0.1646	0.8752	0.4857	0.8606	1.6599	0.8473
0.3	0.1779	0.8120	0.5123	0.7911	1.7218	0.7724
0.4	0.1967	0.7511	0.5504	0.7296	1.8108	0.6999
0.5	0.2216	0.6987	0.6009	0.6639	1.9296	0.6310
1	0.4593	0.4447	1.0953	0.3902	3.1368	0.3381
1.5	1.0054	0.2916	2.3077	0.1968	6.3581	0.1491
2	2.2435	0.1608	5.3571	0.0928	15.5023	0.0561
2.5	5.3042	0.0789	14.0207	0.0359	45.7285	0.0172
3	13.9692	0.0348	43.0171	0.0120	166.307	0.0042

Example 2.

$$\left.\begin{array}{ll} X \sim N(\theta, \sigma^2) & (\sigma^2 \ \textit{known}) \\ \pi(\theta) \sim N(\mu, \tau^2) & \end{array}\right\} \Longrightarrow \pi(\theta|x) \sim N(\mu(x), \tau^2(x)),$$

$$\textit{with} \ \left\{ \begin{array}{r} \mu(x) = \frac{\sigma^2}{\sigma^2+\tau^2}\mu + \frac{\tau^2}{\sigma^2+\tau^2}x \\ \tau^2(x) = \frac{\sigma^2\tau^2}{\sigma^2+\tau^2} \end{array} \right.$$

If we take $g(x) = \frac{\sigma^2}{\sigma^2+\tau^2}\mu + \frac{\tau^2}{\sigma^2+\tau^2}x$ and $c = \frac{\tau^2}{\sigma^2+\tau^2}$, assumptions in Theorem 1 are always fulfilled, and assumptions in Theorem 2 are fulfilled when $\mu = 0$.

We take $\mu = 0$ and $\tau^2 = \sigma^2 = 1$ for obtaining the following tables (by simulation):

	$\Theta_0 = (-1, \infty)$		$\Theta_0 = (0, \infty)$		$\Theta_0 = (1, \infty)$	
x	$D^*(x)$	$p^*(x)$	$D^*(x)$	$p^*(x)$	$D^*(x)$	$p^*(x)$
-2	1.0000	0.0707	11.7146	0.0124	426.5570	0.0008
-1	0.3153	0.2466	3.1710	0.0905	58.0060	0.0135
0	0.0854	0.4796	1.0000	0.3103	11.7146	0.1001
1	0.0172	0.6602	0.3154	0.5688	3.1710	0.3532
2	0.0023	0.7867	0.0854	0.7707	1.0000	0.6640

	$\Theta_0 = (-1.5, 1.5)$		$\Theta_0 = (-1, 1)$		$\Theta_0 = (-0.5, 0.5)$	
x	$D^*(x)$	$p^*(x)$	$D^*(x)$	$p^*(x)$	$D^*(x)$	$p^*(x)$
0.1	0.0356	0.9344	0.1881	0.9278	0.9253	0.9229
0.2	0.0370	0.8686	0.1925	0.8580	0.9375	0.8475
0.3	0.0394	0.8034	0.1999	0.7904	0.9581	0.7727
0.4	0.0429	0.7408	0.2103	0.7229	0.9872	0.6994
0.5	0.0474	0.6793	0.2239	0.6592	1.0253	0.6298
1	0.0881	0.4258	0.3453	0.3782	1.3733	0.3361
1.5	0.1698	0.2450	0.5835	0.1917	2.0932	0.1505
2	0.3157	0.1326	1.0094	0.0866	3.4883	0.0549
2.5	0.5671	0.0683	1.7693	0.0340	6.2591	0.0169
3	1.0000	0.0263	3.1745	0.0109	12.1043	0.0043

Example 3.

$$\left.\begin{array}{l} X \sim f(x|\theta) = \begin{cases} 1 + x - \theta & \text{if } x \in (\theta - 1, \theta) \\ 1 - x + \theta & \text{if } x \in [\theta, \theta + 1) \\ 0 & \text{otherwise} \end{cases} \\ \pi(\theta) \propto 1 \end{array}\right\}$$

$$\Longrightarrow \pi(\theta|x) = \begin{cases} 1 + \theta - x & \text{if } \theta \in (x - 1, x) \\ 1 - \theta + x & \text{if } \theta \in [x, x + 1) \\ 0 & \text{otherwise} \end{cases}$$

Assumptions in Theorems 1 and 2 are fulfilled by taking $g(x) = x$ and $c = 1$. We get the following tables (by simulation):

	$\Theta_0 = (-1, \infty)$		$\Theta_0 = (0, \infty)$		$\Theta_0 = (1, \infty)$	
x	$D^*(x)$	$p^*(x)$	$D^*(x)$	$p^*(x)$	$D^*(x)$	$p^*(x)$
-1.5	7.0000	0.0601	∞	0.0000	∞	0.0000
-1	1.0000	0.2573	∞	0.0000	∞	0.0000
-0.5	0.1429	0.4289	7.0000	0.0629	∞	0.0000
0	0.0000	1.0000	1.0000	0.2546	∞	0.0000
0.5	0.0000	1.0000	0.1429	0.4369	7.0000	0.0654
1	0.0000	1.0000	0.0000	1.0000	1.0000	0.2509
1.5	0.0000	1.0000	0.0000	1.0000	0.1429	0.4547

	$\Theta_0 = (-2, 2)$		$\Theta_0 = (-1, 1)$		$\Theta_0 = (-0.5, 0.5)$	
x	$D^*(x)$	$p^*(x)$	$D^*(x)$	$p^*(x)$	$D^*(x)$	$p^*(x)$
1	0.0000	1.0000	1.0000	0.2558	7.0000	0.0616
1.1	0.0050	0.5077	1.4691	0.2102	11.5000	0.0394
1.2	0.0204	0.4927	2.1250	0.1620	21.2222	0.0225
1.3	0.0471	0.4562	3.0816	0.1210	49.0000	0.0097
1.4	0.0869	0.4479	4.5556	0.0932	199.0000	0.0025

We remark that if we take $\pi(\theta) \sim U(-M, M)$ in this example (instead of $\pi(\theta) \propto 1$), the same posterior density is obtained, provided that $-M < x - 1$ and $x + 1 < M$. In other words, the same results are obtained by taking a uniform density as prior density, provided that its support is large enough.

References

Bayarri, M.J. and J.O. Berger (1999). Quantifying surprise in the data and model verification. In *Bayesian Statistics 6* (Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. eds.), Oxford University Press, Oxford, 53-82 (with discussion).

Box, G.E.P. (1980). Sampling and Bayes' inference in scientific modelling and robustness. *Journal of the Royal Statistical Society Series A* 143, 383-430.

De la Horra, J. and M.T. Rodríguez-Bernal (1997). Asymptotic behaviour of the posterior predictive p-value. *Communications in Statistics-Theory and Methods* 26, 2689-2699.

De la Horra, J. and M.T. Rodríguez-Bernal (1999). The posterior predictive p-value for the problem of goodness of fit. *Test* 8, 117-128.

De la Horra, J. and M.T. Rodríguez-Bernal (2000). Optimality of the posterior predictive p-value based on the posterior odds. *Communications in Statistics-Theory and Methods* 29, 181-192.

De la Horra, J. and M.T. Rodríguez-Bernal (2001). Posterior predictive p-values: what they are and what they are not. *Test* (to appear).

Gelman, A., X.L. Meng and H. Stern (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica* 6, 733-807 (with discussion).

Guttman; I. (1967). The use of the concept of a future observation in goodness-of-fit problems. *Journal of the Royal Statistical Society Series B* 29, 83-100.

Lavine, M. and M.J. Schervish (1999). Bayes factors: what they are and what they are not. *The American Statistician* 53, 119-122.

Meng, X.L. (1994). Posterior predictive p-values. *Annals of Statistics* 22, 1142-1160.

Rubin, D.B. (1984). Bayesianly justifiable and relevant frequency calculations for the applied statistician. *Annals of Statistics* 12, 1151-1172.

Schervish, M. (1996). P-values: what they are and what they are not. *The American Statistician* 50, 203-206.

Thompson, P. (1997). Bayes p-values. *Statistics and Probability Letters* 31, 267-274.

Tsui, K.W. and S. Weerahandi (1989). Generalized p-values in significance testing of hypothesis in the presence of nuisance parameters. *Journal of the American Statistical Association* 84, 602-607.