# $\varepsilon$–contaminated priors in testing point null hypothesis: A procedure to determine the prior probability

**Miguel A. Gómez–Villegas**[1], **Luis Sanz**

*Departamento de Estadística e Investigación Operativa,*
*Universidad Complutense de Madrid. 28040 Madrid, Spain.*

**Abstract**

In this paper the problem of testing a point null hypothesis from the Bayesian perspective and the relation between this and classical approach is studied. A procedure to determine the mixed prior distribution is introduced and a justification for this construction based on a measure of discrepancy is given. Then, we compare a lower bound for the posterior probability, when the prior is in the class of $\varepsilon$–contaminated distributions, of the point null hypothesis with the p–value.

*AMS classification:* 62F15; 62A15.

*Keywords*: $\varepsilon$–contaminated class; Point null hypothesis; Posterior probability; $p$–values.

## 1. Introduction

### 1.1. $\varepsilon$–contaminated class

To carry out a Bayesian analysis concerning an unknown parameter, $\theta$, it is necessary to indicate the prior beliefs about $\theta$ through a prior distribution of probability. It is not usually the case that the prior information could be expressed in terms of a concrete probability distribution since this prior information is, frequently, vague. This lack of precision is the reason why, often, the prior information is expressed in terms of a class of distributions, $\Gamma$, in which we include all priors that look reasonable concerning our prior beliefs. Moreover, to compare the posterior probability of the null hypothesis with the p–value it looks reasonable to elicit a class of priors instead of a concrete prior distribution since the p–value doesn't use prior information.

An interesting way to describe prior beliefs is to consider the $\varepsilon$–contaminated class

[1]Corresponding author: Departamento de Estadística e Investigación Operativa, Facultad de Ciencias Matemáticas, Universidad Complutense de Madrid, 28040 Madrid, Spain.

given by

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

where $\pi_0$ is a particular prior distribution, the prior that one would use in a Bayesian analysis with only one prior distribution. $Q$ is the class of probability distributions that represents the possible –and reasonable– deviations of $\pi_0$. A fixed $\varepsilon$, with $0 \leq \epsilon \leq 1$, represents the degree of contamination that we want to introduce in $\pi_0$.

As to the class $Q$, there are several possibilities we can take into account. We are going to use the class of all probability distributions. Huber (1973) and Sivaganesan (1988) use this class in other context. Berger and Berliner (1986), Berger (1985), Sivaganesan and Berger (1989) and Berger (1994) give relevant information about other classes of contamination.

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–values with the infimum of the posterior probability, also relevant examples are contained in Section 2. Finally, in Section 3 a discussion and some other possible ways to apply the same idea are contained.

### 1.2. The problem

We consider the 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 $f(x|\theta)$ continuous in $\theta_0$. 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 $p = \frac{1}{2}$–,(see Robert, 1994, Ch. 5).

In 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 $\theta_0$. Examples can be seen in Berger (1985), Berger and Delampady (1987) and Lee (1989) among others.

An interesting discussion about the difference between (2) and (3), without using this mechanism, can be found in Lindley (1988) and discussion contained there.

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 $\Gamma$ defined in (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

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

This construction is based on the assumption that $\pi(\theta)$ represents our prior beliefs about $\theta$ 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 $\Gamma$ in (1). From (5) we have $p = (1-\varepsilon)p_0 + \varepsilon q_0$, where

$$p_0 = \int_{|\theta-\theta_0|\leq b} \pi_0(\theta)d\theta, \quad \text{and} \quad q_0 = \int_{|\theta-\theta_0|\leq b} q(\theta)d\theta. \tag{6}$$

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), Cox and Hinckley (1974), DeGroot (1974), Bernardo (1980), Rubin (1984), Ghost and Mukerjee (1992), Berger, Boukai and Wang (1997) and Mukhopadhyay and Das Gupta (1997).

### 1.3. Justification and notation

The choice of $p$, the mass assigned to the point null hypotheses, 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 $\pi$ and $\pi^*$.

There is a problem here because $\pi(\theta)$ is a density but $\pi^*(\theta)$ is not a density. We can sort out the problem considering two measures on $(\mathcal{R}, \mathcal{B}_{\mathcal{R}})$. For all $A$ in the Borel $\sigma$–field we define the following measures

$$\mu(A) = \int_A \pi(\theta)\, d\lambda(\theta) \quad \text{and} \quad \mu^*(A) = \int_A \pi^*(\theta)\, d\lambda(\theta) = \begin{cases} p + (1-p)\mu(A) & \text{if} \quad \theta_0 \in A \\ (1-p)\mu(A) & \text{if} \quad \theta_0 \in A^c \end{cases}.$$

The first measure, $\mu(A)$, is originated by the density $\pi(\theta)$ and the second, $\mu^*(A)$, by $\pi^*(\theta)$, where $p$ is given by (5) and $\lambda$ is the Lebesgue measure.

It is easy to prove that $\mu$ is absolutely continuous with respect to $\mu^*$ ( $\mu << \mu^*$ ), so it exists $d\mu/d\mu^*$, the Radon-Nikodym derivative of $\mu$ with respect to $\mu^*$. Besides, it is straightforward to see that

$$\frac{d\mu}{d\mu^*}(\theta) = \begin{cases} 0 & \text{if} \quad \theta = \theta_0 \\ \frac{1}{1-p} & \text{if} \quad \theta \neq \theta_0 \end{cases} \tag{7}$$

Now, using that $\mu << \mu^*$, we can define the discrepance between $\mu$ and $\mu^*$ as $\delta(\mu^*|\mu) = \int_\Theta (\ln(d\mu/d\mu^*))\, d\mu$. Then, by (7), we have $\delta(\mu^*|\mu) = -\ln(1-p)$.

Several comments are in order. First, when $b$ goes to zero then $p$, according to (5), goes to zero too and the discrepancy between $\mu^*$ and $\mu$ also goes to zero. This is a justification to construct (4) in this way and it makes reasonable the replacement of (2) by (3). Secondly, if $p = 1/2$ is employed, instead of using the value of p given in (5), then the discrepancy between $\mu^*$ and $\mu$ is $\delta(\mu^*|\mu) = 0.693$; which is perhaps a high discrepancy. Finally, the suitable choice of $b$, which depends on the problem we are dealing with, is perhaps more intuitive than just selecting an arbitrary value of $p$.

We denote the likelihood function by $f(x|\theta)$, which is considered as a function of $\theta$ for the observed value $x$. 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 $m(x|\pi) = (1-\varepsilon)m(x|\pi_0) + \varepsilon m(x|q)$, hence, if the posterior distributions $\pi_0(\theta|x)$ and $q(\theta|x)$ exist, the posterior distribution of $\theta$ given $x$ with respect to $\pi$ is given by $\pi(\theta \mid x) = \lambda(x)\pi_0(\theta|x) + (1-\lambda(x))q(\theta|x)$, where $\lambda(x) = (1-\varepsilon)(m(x|\pi_0))/m(x|\pi)$.

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. Arbitrary contaminations

In this section we obtain, in Theorem 1, a lower bound for the posterior probability of the point null hypothesis, given $\pi^*$ by (4) and with $p$ computed according to (5). In order to achieve the infimum of the posterior probability sufficient conditions, when $\pi \in \Gamma$, are established in Theorem 2. Finally, this Section contains several examples where we compare the lower bounds of the posterior probability with the p–values for different values of $b$. Gómez–Villegas and Sanz (1998) give a survey about these ideas using the class of all unimodal and symmetric distributions.

**Theorem 1.** *Consider the hyphoteses introduced in* (2), *an arbitrary prior distribution* $\pi(\theta) \in \Gamma$ *as in (1) and a mixed prior distribution as (4) with the mass assigned to the null hypothesis according to (5). Then*

$$Pr(H_0^*|x) \geq \left(1 + \frac{1-(1-\varepsilon)p_0}{(1-\varepsilon)p_0} r(x)\right)^{-1}, \tag{8}$$

*where*

$$r(x) = (1-\varepsilon)\frac{m(x|\pi_0)}{f(x|\theta_0)} + \varepsilon \frac{\sup_{\theta \neq \theta_0} f(x|\theta)}{f(x|\theta_0)}.$$

*Proof:* Computing a lower bound of the posterior probability of $H_0^*$

$$Pr(H_0^*|x) = \frac{f(x|\theta_0)}{f(x|\theta_0) + \dfrac{1-p}{p} m(x|\pi)} \tag{9}$$

is just like computing an upper bound of $(1-p)/p\, m(x|\pi)$ when $\pi \in \Gamma$. We remark that, by the construction of $\pi^*(\theta)$, $p$ depends on $q$ through $q_0$. So the infimum of $Pr(H_0^*|x)$ in $q$ can be computed as the supremum in $q \in Q$ of

$$\frac{1-p}{p} m(x|\pi) = \left[\frac{1}{(1-\varepsilon)p_0 + \varepsilon q_0} - 1\right]\left[(1-\varepsilon)m(x|\pi_0) + \varepsilon m(x|q)\right]. \tag{10}$$

With $p_0$ and $q_0$ given by (6) and the supremum of (10) is always less or equal than the product of

$$\sup_{q\in Q}\left[\frac{1}{(1-\varepsilon)p_0+\varepsilon q_0}-1\right] \tag{11}$$

and $\sup_{q\in Q}[(1-\varepsilon)m(x|\pi_0)+\varepsilon m(x|q)]$. But (11) is equal to $\dfrac{1}{(1-\varepsilon)p_0}-1$ and

$$m(x|q)=\int_\Theta f(x|\theta)q(\theta)d\theta \le \sup_{\theta\neq\theta_0} f(x|\theta)\int_\Theta q(\theta)d\theta = \sup_{\theta\neq\theta_0} f(x|\theta).$$

Then, immediately we obtain (8). □

Theorem 1 gives a lower bound for the posterior probability of the null hypothesis and a first question is when the infimum is achieved by a distribution of the class $\Gamma$ in (1). The answer is given by the following theorem.

**Theorem 2.** *Let $\hat\theta_n$ be the maximum likelihood estimator of $\theta$ when $\theta$ is in $H_1^*$. If $\hat\theta_n\in I_b^c$ and, for fixed $\delta$, $\int_{\hat\theta_n-\delta}^{\hat\theta_n+\delta} f(x|\theta)\,d\theta$ is approximated by $2\delta f(x|\hat\theta_n)$, then the distribution given by $\overline{\pi}(\theta)=(1-\varepsilon)\pi_0(\theta)+\varepsilon\overline{q}(\theta)$, where $\overline{q}(\theta)$ is uniform in $\left(\hat\theta_n-\delta,\hat\theta_n+\delta\right)$, satisfies*

$$\inf_{\pi\in\Gamma} Pr(H_0^*|x)=Pr(H_0^*|x,\overline{\pi})=\left(1+\frac{1-(1-\varepsilon)p_0}{(1-\varepsilon)p_0}r(x)\right)^{-1} \tag{12}$$

*Proof:* By (9), we need to compute $p$ and $m(x|\overline{\pi})$. But

$$\begin{aligned} p &= \int_{|\theta-\theta_0|\le b}\overline{\pi}(\theta)\,d\theta \\ &= (1-\varepsilon)\int_{|\theta-\theta_0|\le b}\pi_0(\theta)\,d\theta+\varepsilon\int_{|\theta-\theta_0|\le b}\overline{q}(\theta)\,d\theta=(1-\varepsilon)p_0. \end{aligned}$$

and $m(x|\overline{\pi})=(1-\varepsilon)m(x|\pi_0)+\varepsilon m(x|\overline{q})$ , where

$$m(x|\overline{q})=\int_\theta f(x|\theta)\overline{q}(\theta)\,d\theta=\frac{1}{2\delta}\int_{\hat\theta_n-\delta}^{\hat\theta_n+\delta} f(x|\theta)\,d\theta\simeq f(x|\hat\theta_n),$$

then we obtain (12). □

It is interesting to note that the real restriction in this theorem is $\widehat{\theta}_n\in I_b^c$, since the approximation of the integral is always possible by choosing a sufficiently small value of $\delta$. When $\widehat{\theta}_n$ is in $I_b$, (8) becomes an strict inequality.

**Example 2.1.** Let us suppose that $X|\theta$ is distributed $N(\theta,\sigma^2)$, with $\sigma^2$ known, and $\pi_0(\theta)$ has a $N(\mu,\tau^2)$ distribution, 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, $m(\overline{x}|\pi_0)$ is $N(\mu,\tau^2+\sigma^2/n)$ and $\sup_{\theta\neq\theta_0} f(x|\theta)=\sqrt{n}/(\sigma\sqrt{2\pi})$, either $x=\theta_0$ or $x\neq\theta_0$. Besides $p_0=\Phi((\theta_0+b-\mu)/\tau)-\Phi((\theta_0-b-\mu)/\tau)$, $\Phi$ denoting the standard cumulative distribution function.

Table 1 shows, with $\sigma^2 = 1$, $\tau^2 = 2$, $\theta_0 = \mu = 0$ and $n = 10$, the values of the lower bound given by (8) for some specific values of $t = \sqrt{n}(\overline{x} - \theta)/\sigma$ and some $b$.

*Table 1 goes here*

We can observe, too, that if we take an adequate value of $b$, the values of the lower bound given by (8) are close to the respective p–values. For example, if we take $b$ between 0.2 and 0.3 it can be seen that these lower bounds are approximately equal to the p–values. □

A way to choose $b$ is to make the lower bound in (8), with $p$ as in (5), agree with the p–value. This could be done if we obtain $b$ from the expression

$$p(x) = \left(1 + \frac{1-(1-\varepsilon)p_0}{(1-\varepsilon)p_0} r(x)\right)^{-1}$$

but this implies that

$$p_0 = \frac{1}{1-\varepsilon}\left(1 + \frac{1-p(x)}{p(x)r(x)}\right)^{-1}, \tag{13}$$

and the prior probability depends on the data. A possibility to avoid that the prior distribution becomes data dependent is to replace $p(x)$ by the significance level of the test, $\alpha$. Moreover, if the value choosen for $b$ is close to that one obtained by (13), then we get that the infimum of the posterior probability and the p–value are closed, since the infimum is a continuous function of $b$.

Jeffreys (1967, pg. 274) has dealt with this problem, normal likelihood with known variance, using the Cauchy distribution as prior. We deal with this situation in the following example.

**Example 2.2.** Let $X$ have a $N(\theta, 1)$ distribution with $\theta$ unknown. Let $\pi_0(\theta)$ be the base prior distribution of $\theta$, i.e. a $Cauchy(0, 2)$ distribution. To test $H_0^* : \theta = 0$ versus $H_1^* : \theta \neq 0$, with a random sample of size 10 and $\varepsilon = 0.2$. Table 2 shows the values of $b$ providing lower bounds for the posterior probability of $H_0^*$ close to the p–values. It can be observed that, in this case, the values of $b$, that make the p–value be near the posterior probability, are now slightly larger than in Example 2.1. □

*Table 2 goes here*

## 3. Comments

The results that we have obtained are consequence of a methodology based on the relation between point null and interval null hypothesis. The discrepancy used in this paper between $\pi(\theta)$ and $\pi^*(\theta)$ justifies the choice of $p$ as in (5) with a suitable value of $b$. According to this procedure, the mixed prior distribution $\pi^*(\theta)$, used in the point null problem, is close to the continuous prior $\pi(\theta)$, used in the interval problem, as we stated in 1.3.

The procedure works well when the prior $\pi(\theta)$ is in the class of $\varepsilon$–contaminated distributions, in the sense that the lower bound for the posterior probability of the point null hypothesis can be close to the p–value as it is shown in the examples that we have dealt with. Similar results were obtained by Gómez–Villegas and Sanz (1998) in a different context.

The value of $b$ must be chosen, in general, as one of the intermediate values in tables 1 and 2, when the sample model is normal. If the sample model has much heavier tails than normal, the values of $b$ needed to obtain agreement will be greater than those obtained in examples 1 and 2 and if we use, in this case, suitable small values of $b$, as it seems more appropriate, the p–values will be greater than the lower bounds of the posterior probability.

Finally, in the light of our results, it seems that the discrepancy observed in testing point null hypothesis between Bayesian and classical approach becomes more acute by using $p = 0.5$ in the mixed distribution.

Other classes of distributions should be studied as prior distributions with this methodology and more research is necessary in order to establish conditions for the choice of b.

## Acknowledgements

Parts of this paper were written while the first author, M.A. Gómez–Villegas, was visiting the Department of Statistical Science, in the University College of London. This author would like to thank this Institution for its hospitality and particularly Professor A. P. Dawid for his useful comments.

The authors wish to thank the editor and the referees for helpful comments and valuable suggestions.

## References

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

Berger, J.O. (1994), An overview over Robust Bayesian analysis, (with discussion), *Test*

**3**, 1, 5–124.

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

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.

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

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.

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

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.

Cox, D.R. and Hinckley, D.V. (1974), *Theoretical Statistics* (Chapman and Hall, London).

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

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.

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

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

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

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

Huber, P.J. (1973), The use of Choquet capacities in statistics, *Bull. Internat. Statist. Inst.* **45**, 181–191.

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

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

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

Lindley, D.V. (1988), Statistical Inference Concerning Hardy-Weinberg equilibrium, *Bayesian Statistics 3, J.M. Bernardo et al. Oxford University Press.* 307–326.

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

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

Robert, Ch. P. (1994), *The Bayesian Choice: A Decision–Theoretic Motivation* (Springer Verlag. New York).

Rubin, D.B. (1984), Bayesianly justifiable and relevant frequency calculations for the applied statiscian, *Ann. Statist.* **12**, 1151–1172

Sivaganesan, S. (1988), Range of the posterior measures for priors with arbitrary contaminations, *Commun. Statist. - Theory Methods* **17**, 1591–1612.

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

Table 1: Lower bounds for posterior probabilities of $H_0^*$ when X $\sim$ Normal

| | t | | | |
|---|---|---|---|---|
| | 1.645 | 1.960 | 2.576 | 3.291 |
| $b$ =0.1 | 0.0325 | 0.0189 | 0.0046 | 0.0006 |
| $b$ =0.2 | 0.0657 | 0.0387 | 0.0097 | 0.0013 |
| $b$ =0.3 | 0.0994 | 0.0595 | 0.0151 | 0.0021 |
| $b$ =0.4 | 0.1335 | 0.0812 | 0.0209 | 0.0028 |
| $b$ =0.5 | 0.1678 | 0.1037 | 0.0272 | 0.0038 |
| p–value | 0.1000 | 0.0500 | 0.0100 | 0.0010 |

Table 2: Lower bounds for posterior probabilities of $H_0^*$ when X $\sim$ $N(\theta, 1)$ and $\pi_0(\theta)$ is $Cauchy(0, 2)$

| t | 1.645 | 1.960 | 2.596 | 3.291 |
|---|---|---|---|---|
| $b$ | 0.4 | 0.4 | 0.3 | 0.3 |
| $\underline{Pr}(H_0^*\|t)$ | 0.0898 | 0.0533 | 0.0103 | 0.0013 |
| p–value | 0.1 | 0.05 | 0.01 | 0.001 |