The Korean Communications
in Statistics Vol. 6, No. 1, 1999
pp. 237-249

Distance between the Distributions of the P-value and the Lower Bound of the Posterior Probability[1]

Hyun Sook Oh[2]

Abstract

It has been issued that the irreconcilability of the classical test for a point null and standard Bayesian formulation for testing such a point null. The infimum of the posterior probability of the null hypothesis is used as measure of evidence against the null hypothesis in Bayesian approach; here the infimum is over the family of priors on the alternative hypotheses which includes all density that are *a priori* reasonable. For iid observations from a multivariate normal distribution in p dimensions with an unknown mean and a covariance matrix propotional to the Identity, we consider the difference and the Wolfowitz distance of the distributions of the P-value and the lower bound of the posterior probability over the family of all normal priors. The Wolfowitz distance is interpreted as the average difference of the quantiles of the two distrbutions.

1. Introduction

It is very well known that in parametric testing problems where the null hypothesis is sharp, the standard Bayesian method and the classical method of testing the null hypothesis are sometimes hard to reconcile. There are many results to this effect in the literature; Berger and Sellke(1987) made an illuminating contribution wherein they show that for testing a sharp null hypothesis about the mean of a univariate normal distributions with a known variance, even the minimum posterior probability of the null hypothesis over really large classes of priors on the two-sided alternative can be significantly larger than the P-value of the common classical test, regardless of the sample size, provided the sharp null is assigned a probability of 0.5. Casella and Berger(1987) showed such a conflict appears to be germane to the sharp null. Berger and Sellke's phenomena has been extended and generalized for high dimension in Berger and Delampady(1987), Delampady(1989a, 1989b, 1990). Also, Oh and DasGupta(1998) considered in greater depth the role of the assumption that the apriori the sharp null has a probability of 0.5 in the Berger-Sellke phenomenon.

This article considers the problem of testing a sharp null hypothesis for iid observations

1) This work was financially supported by the Grant of Kyungwon University in 1998.
2) Assistant Professor, Dept. of Applied Statistics, Kyungwon University, Sungnam, Kyunggido, Korea

from multivariate normal distribution; $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ are iid random vectors in p dimension with the $N(\boldsymbol{\theta}, \sigma^2 I)$ distribution. A point null hypothesis $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ is tested against a two-sided alternative $H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$. π_0 denotes the Bayesian's apriori probability on H_0. The family of priors considered on the alternative is Γ_{MVN} = {all multivariate normal $N(\mathbf{0}, \tau^2 I)$ priors}. Without loss of generality let us assume $\boldsymbol{\theta}_0 = \mathbf{0}$. Since a sufficient statistic of $\boldsymbol{\theta}$ is $\overline{\boldsymbol{X}}$ which is distributed as $N(\boldsymbol{\theta}, \frac{\sigma^2}{n} I)$, for any $g(\boldsymbol{\theta}) \in \Gamma_{MVN}$, the posterior probability of H_0 is given by

$$P(H_0 | \boldsymbol{x}, g(\boldsymbol{\theta})) = \left[1 + \frac{(1-\pi_0)}{\pi_0} \frac{m_g(\overline{\boldsymbol{x}})}{f(\overline{\boldsymbol{x}} | \boldsymbol{\theta} = \mathbf{0})} \right]^{-1},$$

where $m_g(\overline{\boldsymbol{x}})$ and $f(\overline{\boldsymbol{x}} | \boldsymbol{\theta} = \mathbf{0})$ are the density functions of $N(\mathbf{0}, (\tau^2 + \frac{\sigma^2}{n}) I)$ and $N(\mathbf{0}, \sigma^2 I)$, respectively. The infimum of the posterior probability of H_0 is attained at $\hat{\tau}^2 = \max\left\{0, \frac{\overline{\boldsymbol{x}}^t \overline{\boldsymbol{x}}}{p} - \frac{\sigma^2}{n}\right\}$ at which $m_g(\overline{\boldsymbol{x}})$ is maximized over $g \in \Gamma_{MVN}$. Here $\hat{\tau}^2$ is called ML-II prior. Let $\boldsymbol{z} = \sqrt{n}\, \overline{\boldsymbol{x}}$. We then have the infimum of the posterior probability of H_0, $\underline{P}(H_0 | \boldsymbol{x}, \Gamma)$, as follows.

$$\underline{P}(H_0 | \boldsymbol{x}, \Gamma) = \begin{cases} \pi_0 & \text{if } \|z\|^2 \le p \\ \left(1 + \frac{1-\pi_0}{\pi_0} \frac{\exp(\|z\|^2/2)}{(e\|z\|^2/p)^{p/2}}\right)^{-1} & \text{if } \|z\|^2 > p. \end{cases}$$

Also, since $\|Z\|^2$ is distributed as the Chisquare distribution with p degrees of freedom under H_0, the classical P-value is given by

$$p(\boldsymbol{z}) = P(X_p^2 \ge \|z\|^2).$$

Section 2 deals with the distribution of the difference between $\underline{P}(H_0 | \boldsymbol{x}, \Gamma_{NOR})$ and the P-value in one dimension. In section 3, the Wolfowitz distance of the unconditional and conditional(conditioned on P-value < p_0(fixed)) distributions of $\underline{P}(H_0 | \boldsymbol{x}, \Gamma_{MVN})$ and the P-value in each dimension is derived. The distance is interpreted as the average difference of the quantiles of the two distributions. Figure 1 is given for visual illustration of the difference between the P-value and $\underline{P}(H_0 | \boldsymbol{x}, \Gamma_{MVN})$ for $\pi_0 = 0.5$ in one dimension.

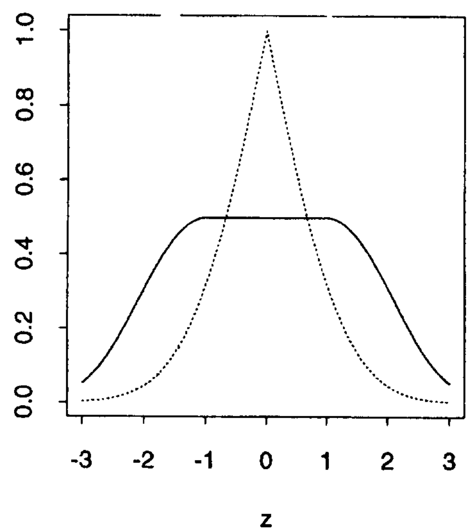

Figure 1(a)

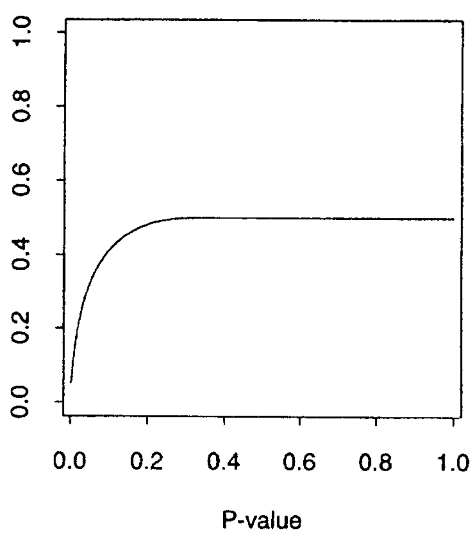

Figure 1(b)

Figure 1: (a) Plots of the P-value($\cdots$) and $\underline{P}(H_0|\mathbf{x}, \Gamma_{NOR})$(—) (b) The Plot of the $\underline{P}(H_0|\mathbf{x}, \Gamma_{NOR})$ vs. the P-value

2. Distribution of the difference between the P-value and the infimum posterior probability

Let us consider the distribution of the difference between $\underline{P}(H_0|\boldsymbol{x},\Gamma_{NOR})$ and the P-value in one dimension. Let $d(z)$ denote the difference of $\underline{P}(H_0|\boldsymbol{x},\Gamma_{NOR})$ and the P-value, $p(z)$. Then

$$d(z) = \underline{P}(H_0|\boldsymbol{x},\Gamma_{NOR}) - p(z)$$
$$= \begin{cases} \pi_0 - 2(1-\Phi(|z|)) & \text{if } |z|\leq 1 \\ \left(1+\frac{1-\pi_0}{\pi_0}\frac{\exp(z^2/2)}{\sqrt{e}|z|}\right)^{-1} & \text{if } |z|>1\,. \end{cases}$$

Clearly, $d(z)$ is increasing if $0\leq z\leq 1$. For $1<z<\infty$, $d(z)$ has exactly one maximum with $\lim_{z\to\infty} d(z)=0$. The proof of this with any dimension p will be given in Lemma 1 in section3.

Theorem 1. For the null distribution of $d(Z)$, the median is $\pi_0-0.5$ for $\pi_0\leq 0.5$; for $\pi_0>0.5$, the median m_d is the unique root of the equation,

$$m_d = d(\Phi^{-1}(5/4+(m_d-\pi_0)/2)).$$

Proof : Since $d(1)=\pi_0-2(1-\Phi(1))$, $\pi_0-0.5<d(1)$. So

$$\underline{P}(d(Z)\leq\pi_0-0.5) = \underline{P}(\pi_0-2(1-\Phi(Z))\leq\pi_0-0.5)$$
$$= \underline{P}(2(1-\Phi(Z))>0.5)=0.5.$$

Now for $\pi_0>0.5$, let m_d denote the median. We claim $m_d<d(1)$; for otherwise,

$$\underline{P}(d(Z)\leq m_d) \geq \underline{P}(\pi_0-2(1-\Phi(Z))\leq\pi_0-2(1-\Phi(1)))$$
$$>0.5,$$

which is a contradiction. Thus

$$\underline{P}(d(Z)\leq m_d) = \underline{P}(\pi_0-2(1-\Phi(Z))\leq m_d)+\underline{P}(d(Z)\leq m_d). \quad (1)$$

But, since $d(z)$ is monotone decreasing for $z<d(1)$ and since $m_d<d(1)$,

$$\underline{P}(d(Z)\leq m_d) = P(|Z|\geq d^{-1}(m_d)).$$

Then (1) implies that

$$m_d = d(\Phi^{-1}(5/4+(m_d-\pi_0)/2))\,. \quad \square$$

The assertion of this result is intriguing. In particular, with $\pi_0 = 0.5$, $\underline{P}(H_0|\boldsymbol{x}, \Gamma_{NOR})$ will be larger than the P-value in exactly 50% of a long sequence of experiments. Table 1 shows the median, mean and standard deviation of the distribution of d(Z) for various values of π_0.

Table 1 : Summary of the distribution of $d(Z)$

π_0	median	mean	std. dev.
0.1	-0.4	-0.407149	0.278297
0.2	-0.3	-0.313462	0.269735
0.3	-0.2	-0.218825	0.263053
0.4	-0.1	-0.12309	0.258303
0.5	0.	-0.0260617	0.255529
0.6	0.072535	0.1162675	0.254794
0.7	0.173107	0.1315536	0.256259
0.8	0.276329	0.1581646	0.260371
0.9	0.38355	0.1967748	0.268511

3. Wolfowitz distance of distributions of the P-value and the infimum posterior probability

Let F and G be two different c.d.f.s. Then $W(F, G) = \int_{-\infty}^{\infty} |F(x) - G(x)|\, dx$ is called the Wolfowitz distance of F and G. Since $W(F, G) = \int_0^1 |F^{-1} - G^{-1}|(x)\, dx$ (see Dudley(1989)), $W(F, G)$ is the average of the absolute difference of quantiles of F and G. In this section the Wolfowitz distance of the P-value and the infimum of the posterior probability of H_0 are derived for the family of multivariate normal priors on H_1, Γ_{MVN}, in each dimension.

3.1 Unconditional distributions

Let F be the c.d.f. of $\underline{P}(H_0|\boldsymbol{x}, \Gamma_{MVN})$ and let G be the c.d.f of the P-value in p-dimension. Since the P-value is uniformly distributed on (0, 1), it follows that G(x) = x for $0 \le x \le 1$. If we define

$$\lambda(t) = \left(1 + \frac{1-\pi_0}{\pi_0} \frac{e^{t/2}}{(et/p)^{p/2}}\right)^{-1}, \tag{2}$$

then

$$P(H_0|X,\Gamma_{MVN})=\begin{cases}\pi_0 & \text{if } \|z\|^2 \le p \\ \lambda(\|z\|^2) & \text{if } \|z\|^2 > p\end{cases}.$$

Since $\lambda(t)$ is decreasing in t for t > p, it follows that F(x) = 1 for $x \ge \pi_0$. For $x < \pi_0$, $F(x) = P(\lambda(\|z\|^2) \le x) = P(\chi_p^2 \ge \lambda^{-1}(x))$. The following Lemma is needed for the proof of Theorem 2.

Lemma 1. $\rho(y)=\lambda(y)-P(\chi_p^2\ge p)=0$ has the only root and $\rho(y)$ changes sign once from the negative to the positive at all on $[0,\infty)$ for any p.

proof: Let us consider $\frac{\partial}{\partial y}\rho(y)$. Note that

$$\frac{\partial}{\partial y}\lambda(y)=\frac{\frac{C_1}{2}y^{p/2-1}e^{y/2}(p-y)}{(y^{p/2}+C_1e^{y/2})^2} \quad \text{and} \quad \frac{\partial}{\partial y}P(\chi_p^2\ge y)=-C_2y^{p/2-1}e^{-y/2},$$

where $C_1=\frac{1-\pi_0}{\pi_0}\left(\frac{p}{e}\right)^{p/2}$ and $C_2=(\Gamma(p/2)2^{p/2})^{-1}$.

Thus

$$\frac{\partial}{\partial y}\rho(y)=0$$
$$\Leftrightarrow (y-p)\frac{C_1}{2}=C_2(y^pe^{-y}+2C_1y^{p/2}e^{-y/2}+C_1^2). \tag{3}$$

Since the expression on the right in (3) is decreasing of y for y > p, $\frac{\partial}{\partial y}\rho(y)=0$ has the only one root. Hence there exists exactly one extremum on $[p,\infty)$ and it can be shown to be a maximum. Now, since $\lim_{y\to 0}\rho(y)=-1$ and $\lim_{y\to\infty}\rho(y)=0$, there must be one root of $\rho(y)=0$. Moreover, $\rho(y)$ changes sign once from the negative to the positive obviously. □

Theorem 2. If $\pi_0 > P(\chi_p^2\ge p)$,

$$W(F,G)=0.5+\pi_0^2-\pi_0[1+P(\chi_p^2\ge p)]+\int_p^\infty f_{p(x)}\lambda(x)dx\ ;$$

if $\pi_0 \le P(\chi_p^2 \le p)$,

$$W(F,G)=0.5-\pi_0[1-P(\chi_p^2\ge p)]-[P(\chi_p^2\ge a(\pi_0))]^2 - \int_p^{a(\pi_0)} f_{p(x)}\lambda(x)dx+\int_{a(\pi_0)}^\infty f_{p(x)}\lambda(x)dx,$$

where $f_p(x)$ denotes the density function of Chisquare distribution with p degrees of freedom and $a(\pi_0)$ is a unique root of the equation $\lambda(y)=P(\chi_p^2\ge y)$ for a given p.

proof : By the definition of F and G,

$$W(F,G)=\int_0^{\pi_0}|P(\chi_p^2\ge\lambda^{-1}(x))-x|dx+\int_{\pi_0}^1(1-x)\,dx$$
$$=-\int_p^\infty|P(\chi_p^2\ge y)-\lambda(y)|\lambda'(y)\,dy+\frac{1}{2}(1-\pi_0)^2\quad(\textit{by change of variable}).$$

Let $a(\pi_0)$ be a unique root of $\lambda(y)-P(\chi_p^2\ge y)=0$. By Lemma 1,

if $a(\pi_0)\le p$, $\lambda(y)\ge P(\chi_p^2\ge y)$ for $y>p$;

if $a(\pi_0)>p$, $\lambda(y)\le P(\chi_p^2\ge y)$ for $p<y\le a(\pi_0)$ and $\lambda(y)>P(\chi_p^2\ge y)$ for $y\ge a(\pi_0)$.

Also, note that

$$a(\pi_0)\le p\Leftrightarrow\lambda(p)-P(\chi_p^2\ge p)>0\Leftrightarrow\pi_0>P(\chi_p^2\ge p).$$

(i) Assume that $\pi_0>P(\chi_p^2\ge p)$.

Then

$$\int_p^\infty(P(\chi_p^2\ge y)-\lambda(y))\lambda'(y)\,dy$$
$$=-P(\chi_p^2\ge p)\lambda(p)+0.5\,\lambda(p)^2+\int_p^\infty f_{p(y)}\lambda(y)dy\quad(\textit{by integration by parts})$$
$$=-\pi_0P(\chi_p^2\ge p)+0.5\pi_0^2+\int_p^\infty f_{p(y)}\lambda(y)dy.$$

Thus

$$W(F,G)=0.5+\pi_0^2-\pi_0[1+P(\chi_p^2\ge p)]+\int_p^\infty f_{p(x)}\lambda(x)dx.$$

(ii) Assume $\pi_0\le P(\chi_p^2\ge y)$.

Then

$$\int_p^\infty(P(\chi_p^2\ge y)-\lambda(y))\lambda'(y)\,dy$$
$$=-\int_p^{a(\pi_0)}(P(\chi_p^2\ge y)-\lambda(y))\lambda'(y)\,dy-\int_{a(\pi_0)}^\infty(\lambda(y)-P(\chi_p^2\ge y))\lambda'(y)\,dy$$
$$=-P(\chi_p^2\ge a(\pi_0))\lambda(a(\pi_0))+P(\chi_p^2\ge p)\lambda(p)+\lambda(a(\pi_0))^2-0.5\lambda(p)^2$$
$$=-P(\chi_p^2\ge a(\pi_0))\lambda(a(\pi_0))-\int_p^{a(\pi_0)}f_p(y)\lambda(y)dy+\int_{a(\pi_0)}^\infty f_p(y)\lambda(y)dy\quad(\textit{by integration by parts})$$
$$=\pi_0P(\chi_p^2\ge p)-0.5\pi_0^2-[P(\chi_p^2\ge a(\pi_0))]^2-\int_p^{a(\pi_0)}f_p(y)\lambda(y)dy+\int_{a(\pi_0)}^\infty f_{p(y)}\lambda(y)dy.$$

Thus

$$W(F,G)=0.5-\pi_0[1-P(\chi_p^2\ge p)]-[P(\chi_p^2\ge a(\pi_0))]^2-\int_p^{a(\pi_0)}f_{p(x)}\lambda(x)dx+\int_{a(\pi_0)}^\infty f_{p(x)}\lambda(x)dx.$$

□

Table 2 shows the values of W(F, G) for various values of π_0 and p. The distance, W(F, G), is large for too small or too large π_0 and as π_0 tends to 0.5 from the extremes the distance becomes short. Also note that W(F, G) is minimized at π_0=0.5 which is considered as 'objective' choice of prior probability on H_0. On the other hand, distances become short as p tends to infinity.

3.2 Conditional distributions

In this section, let us consider the case that the P-value is moderately small for which we have more interest in comparing $\underline{P}(H_0|\boldsymbol{x}, \Gamma_{MVN})$ and the P-value. We have the following illustrative example.

Example 1 Let F^c denote the conditional c.d.f. of $\underline{P}(H_0|\boldsymbol{x}, \Gamma_{NOR})$ and let G^c denote the conditional c.d.f of P-value, given P-value < 0.2. Let $h(y) = (1 + \frac{1-\pi_0}{\pi_0}\frac{e^{y^2/2}}{\sqrt{ey}})^{-1}$, y > 0.

Then

$$F^c(x) = 10(1-\Phi(h^{-1}(x))), \quad 0 < x \le h(z_0),$$

where z_0 is a root of $2(1-\Phi(z)) = 0.2$. Also,

$$G^c(x) = 5x, \quad 0 < x < 0.2.$$

Since h(y) is increasing in π_0 for a fixed y > 0, $h(z_0) \ge 0.2$ if $\pi_0 \ge \pi_0^*$, where π_0^* satisfies $h(z_0) = 0.2$. The approximate value of z_0 and π_0^* are 1.28155 and 0.211955 respectively. Also, since $h(y) - 2(1-\Phi(y))$ has only one root and it changes sign once from the negative to the positive at all for y > 0 (see Lemma 1 for the proof of this in general multi dimension),

$$h(y) > 2(1-\Phi(y)) \quad \text{for } y > z_0, \text{ if } . \pi_0 \ge \pi_0^*$$

Let $h_0 = h(z_0)$. Thus, if $\pi_0 \ge \pi_0^*$, the Wolfowitz distance of F^c and G^c is

$$\begin{aligned}
W(F^c, G^c) &= 5\Big[\int_0^{0.2} |2(1-\Phi(h^{-1}(x))) - x|\,dx + \int_{0.2}^{h_0} (0.2 - 2(1-\Phi(h^{-1}(x))))\,dx\Big] \\
&= 5\Big[-\int_{h^{-1}(0.2)}^{\infty} |2(1-\Phi(y)) - h(y)| h'(y)\,dy \\
&\quad + \int_{z_0}^{h^{-1}(y)} (2(1-\Phi(y)) - 0.2) h'(y)\,dy\Big] \quad (\textit{by change of variable}) \\
&= 5\Big[-\int_{h^{-1}(0.2)}^{\infty} (h(y) - 2(1-\Phi(y))) h'(y)\,dy \\
&\quad + \int_{z_0}^{h^{-1}(y)} (2(1-\Phi(y)) - 0.2) h'(y)\,dy\Big] \\
&= 5\Big[-0.02 + 2\int_{z_0}^{\infty} \phi(y) h(y)\,dy\Big] \quad (\textit{by integration by parts}).
\end{aligned}$$

Now, assume that $\pi_0 < \pi_0^*$ and let $a(\pi_0)$ be a unique root of the equation

$h(y)=2(1-\Phi(y))$ for $y>0$. Again, since $h(y)-2(1-\Phi(y))$ has only one root and it changes sign once from the negative to the positive at all for y > 0 and $h(z_0)<2(1-\Phi(z_0))$, $a(\pi_0)>z_0$. Thus, if $\pi_0<\pi_0^*$,

$$\begin{aligned} W(F^c,G^c) = & \ 5\Big[-\int_{z_0}^{a(\pi_0)}(2(1-\Phi(y))-h(y))h'(y)\,dy \\ & -\int_{a(\pi_0)}^{\infty}(h(y)-2(1-\Phi(y)))h'(y)\,dy+\int_{h_0}^{0.2}(0.2-x)\,dx\Big] \\ = & \ 5\Big[0.02-[2(1-\Phi(a(\pi_0)))]^2-2\int_{z_0}^{a(\pi_0)}\phi(y)h(y)\,dy+2\int_{a(\pi_0)}^{\infty}\phi(y)h(y)\,dy\Big]. \end{aligned}$$

For each π_0, the value of $W(F^c,G^c)$ is given on the first column (p=1) in Table 2. The distance increases as π_0 goes away from near 0.2. So if the P-value < 0.2 is given, i.e., there is a significant evidence against H_0 at level 20%, then the choice of $\pi_0=0.2$ seems to give the shortest distance of F^c and G^c. Also, note that the distances of F^c and G^c are small compared. □

Let us consider the distance between the conditional distributions of the P-value and $\underline{P}(H_0|\boldsymbol{x},\Gamma_{MVN})$ in general multi dimension. Let F^c be the conditional c.d.f of $\underline{P}(H_0|\boldsymbol{x},\Gamma_{MVN})$ and let G^c be the conditional c.d.f of P-value given that the P-value is less than or equal to p_0(fixed). Let t_0 be a value such that $p_0=P(\chi_p^2\ge t_0)$. Assume that $p_0\le P(\chi_p^2\ge p)$ and recall $\lambda(t)$ in (2). Then the conditional distribution of $\underline{P}(H_0|\boldsymbol{x},\Gamma_{MVN})$ is given by

$$F^c(x)=\begin{cases}\dfrac{P(\chi_p^2\ge\lambda^{-1}(x))}{p_0} & \text{if } 0<x\le\lambda(t_0)\\ 1 & \text{if } x\ge\lambda(t_0)\end{cases}$$

and

$$G^c(x)=\begin{cases}\dfrac{x}{p_0} & \text{if } 0<x<p_0\\ 1 & \text{if } x\ge p_0\end{cases}.$$

Theorem 3. Given P-value $<p_0$ with $p_0\le P(\chi_p^2\ge p)$,

let $\pi_0^*=\dfrac{p_0}{p_0+(1-p_0)e^{-t_0/2}(et_0/p)^{p/2}}$, where t_0 is the $(1-p_0)th$ percentile of Chisquare distribution of p degrees of freedom.

Then,

$$W(F^c, G^c) = \begin{cases} \frac{1}{p_0}[-0.5p_0^2 + \int_{t_0}^{\infty} f_{p(x)}\lambda(x)dx] & \text{if } \pi_0 \geq \pi_0^* \\ \frac{1}{p_0}\{0.5p_0^2 - [P(\chi_p^2 \geq a(\pi_0))]^2 \\ \quad - \int_{t_0}^{a(\pi_0)} f_p(x)\lambda(x)dx + \int_{a(\pi_0)}^{\infty} f_p(x)\lambda(x)dx\} & \text{if } \pi_0 < \pi_0^*, \end{cases}$$

where $f_p(x)$ denotes the density function of Chisquare distribution with p degrees of freedom and $\lambda(y) = P(\chi_p^2 \geq y)$ for a given p.

proof : Let $\lambda_0 = \lambda(t_0)$. Note that since λ_0 is a increasing function of π_0, $\lambda_0 \geq p_0 \Leftrightarrow \pi_0 \geq \pi_0^*$, where π_0^* satisfies the equation $\lambda_0 = p_0$. Then

$$\pi_0^* = \frac{p_0}{p_0 + (1-p_0)e^{-t_0/2}(et_0/p)^{p/2}}$$

(i) Assume that $\pi_0 \geq \pi_0^*$.

Then, by the definition of F^c and G^c,

$$\begin{aligned} W(F^c, G^c) &= \frac{1}{p_0}[\int_0^{p_0} |P(\chi_p^2 \geq \lambda^{-1}(x) - x| dx \\ &\quad + \int_{p_0}^{\lambda_0} (p_0 - P(\chi_p^2 \geq \lambda^{-1}(x) dx] \\ &= \frac{1}{p_0}[-\int_{\lambda^{-1}(p_0)}^{\infty} (P(\chi_p^2 \geq y) - \lambda(y))\lambda'(y) dy] \\ &\quad \frac{1}{p_0}[-\int_{t_0}^{\lambda^{-1}(p_0)} (p_0 - P(\chi_p^2 \geq y))\lambda'(y) dy] \\ &= \frac{1}{p_0}[-0.5p_0^2 + \int_{t_0}^{\infty} f_p(y)\lambda(y) dy] \end{aligned}$$

(ii) Assume that $\pi_0 < \pi_0^*$.

Then

$$\begin{aligned} W(F^c, G^c) &= \frac{1}{p_0}[\int_0^{\lambda_0} |P(\chi_p^2 \geq \lambda^{-1}(x) - x| dx + \int_{\lambda_0}^{p_0} (P_0 - x)dx] \\ &= \frac{1}{p_0}[-\int_{t_0}^{\infty} |P(\chi_p^2 \geq y) - \lambda(y)| \lambda'(y) dy + p_0(p_0 - \lambda_0) - 0.5(p_0^2 - \lambda_0^2)]. \end{aligned}$$

Let $a(\pi_0)$ be a unique root of the equation $\lambda(y) = P(\chi_p^2 \geq y)$. Since $\lambda_0 < p_0$ for $\pi_0 < \pi_0^*$, $a(\pi_0) > t_0$ by Lemma 1. Thus

$$\begin{aligned} &-\int_{t_0}^{\infty} |P(\chi_p^2 \geq y) - \lambda(y)| \lambda'(y) dy \\ &= -\int_{t_0}^{a(\pi_0)} (P(\chi_p^2 \geq y) - \lambda(y))\lambda'(y) dy - \int_{a(\pi_0)}^{\infty} (\lambda(y) - P(\chi_p^2 \geq y))\lambda'(y) dy \\ &= -0.5\lambda_0^2 - [P(\chi_p^2 \geq a(\pi_0))]^{2+} p_0\lambda_0 - \int_{t_0}^{a(\pi_0)} f_p(y)\lambda(y)dy + \int_{a(\pi_0)}^{\infty} f_p(y)\lambda(y) dy. \end{aligned}$$

This proves the result.

□

Table 3 shows the values of $W(F^c, G^c)$ for various values of π_0 and p for $p_0 = 0.2$.

As in Example 1, it can be said that the distance increases as π_0 goes away from near $p_0 = 0.2$. So if the P-value < p_0 is given, i.e., there is a significant evidence against H_0 at level $100\,p_0\%$, then the choice of $\pi_0 = p_0$ seems to give the shortest distance of F^c and G^c in p-dimension. Also, distances become short as p tends to infinity. That is, Bayes-classical conflict distance decreases as p increases, which is known as Lindley's paradox.

References

[1] Berger, J. O. and Sellke, T. (1987). Testing a Point Null Hypothesis: The irreconcilability of Significance levels and Evidence, *Journal of American Statistical Association.* 82 112-122.

[2] Berger,J. O. and Delampady, M. (1987). Testing precise hypotheses (with discussion), *Statistical. Science.* 2 317-348.

[3] Casella, G. and Berger, R. L. (1987). Reconciling Bayesian and Frequentist Evidence in the One-sided Testing Problem (with disscussion), *Journal of American Statistical Association.* 82 106-111.

[4] Delampady, M. (1989a). Lower bounds on Bayes factors for invariant testing situations, *Journal of Multivariate Analysis.* 28 227-246

[5] Delampady, M. (1989b). Lower bounds on Bayes factors for interval null hypotheses, *Journal of American Statistical Association.* 84 120-124.

[6] Delampady, M. (1990). Bayesian hypothesis testing with symmetric and unimodal priors, *Technical Report #90-47, Purdue Univ., West Laf.*

[7] Dudley, R. M. (1989). *Real Analysis and Probability,* Wardsworth and Brooks/Cole, Advanced Books and Software, Pacific Grove, California.

[8] Oh, H. S. and DasGupta, A. (1999). Comparison of the P-value and posterior probability of a sharp null hypothesis, *Journal of Statistical Planning and Inferences* 76 93-107

Table 2 : Value of W(F, G)

π_0	p=1	p=2	p=3	p=4	p=5	p=6	p=7
0.1	0.40768	0.408904	0.409625	0.410103	0.410448	0.410730	0.410925
0.15	0.364839	0.365186	0.365647	0.366022	0.366323	0.366570	0.366776
0.2	0.327574	0.326292	0.325988	0.325908	0.325903	0.325970	0.325962
0.25	0.296446	0.293563	0.292395	0.291758	0.291356	0.291078	0.290875
0.3	0.271205	0.267117	0.265220	0.264090	0.263325	0.262767	0.262337
0.35	0.251363	0.246617	0.244215	0.245722	0.241683	0.240908	0.240301
0.4	0.236874	0.231655	0.228969	0.227259	0.226047	0.225129	0.224403
0.45	0.227707	0.222086	0.219187	0.217341	0.216028	0.215031	0.214239
0.5	0.223894	0.217944	0.214870	0.212910	0.211516	0.210456	0.209614
0.55	0.22547	0.219276	0.216069	0.214021	0.212563	0.211455	0.210574
0.6	0.232481	0.226137	0.222843	0.220783	0.219238	0.218096	0.217189
0.7	0.263045	0.256747	0.253448	0.251336	0.249829	0.248680	0.247765
0.8	0.316258	0.310603	0.307620	0.305699	0.304324	0.303274	0.302436
0.9	0.39347	0.389431	0.387260	0.385854	0.384841	0.384066	0.383445

π_0	p=8	p=9	p=10	p=15	p=20	p=25
0.1	0.411099	0.411245	0.411371	0.411808	0.412077	0.412265
0.15	0.366951	0.367102	0.367535	0.367717	0.368028	0.368250
0.2	0.326002	0.326043	0.326083	0.326261	0.326398	0.326506
0.25	0.290719	0.290569	0.290496	0.290190	0.290032	0.289937
0.3	0.261994	0.261713	0.261476	0.260688	0.260230	0.259924
0.35	0.239809	0.239400	0.239054	0.237872	0.237165	0.236682
0.4	0.223811	0.223315	0.222891	0.221532	0.220545	0.219933
0.45	0.213589	0.213044	0.212577	0.210955	0.209958	0.209266
0.5	0.208923	0.208343	0.207846	0.206119	0.205057	0.204319
0.55	0.209851	0.209243	0.208732	0.206912	0.205799	0.205024
0.6	0.216444	0.215817	0.215280	0.213413	0.212263	0.211463
0.7	0.247013	0.246381	0.245839	0.243949	0.242784	0.241972
0.8	0.301747	0.301167	0.300669	0.298929	0.297927	0.297102
0.9	0.382934	0.382502	0.382131	0.380830	0.380022	0.379456

Table 3 : Values of W(F^c, G^c) conditioning on P-value < 0.2

π_0	p=1	p=2	p=3	p=4	p=5	p=6	p=7
0.1	0.03717	0.041505	0.043897	0.045432	0.046517	0.047335	0.047981
0.15	0.02244	0.021644	0.022219	0.022874	0.023462	0.023970	0.024410
0.2	0.0356611	0.026060	0.022357	0.020415	0.019227	0.018428	0.017858
0.25	0.0711361	0.056729	0.050003	0.045916	0.043100	0.041007	0.039370
0.3	0.108668	0.091790	0.083875	0.079056	0.075758	0.073251	0.071314
0.35	0.147503	0.128335	0.119304	0.113791	0.109977	0.107136	0.104912
0.4	0.187754	0.166499	0.156436	0.150278	0.146010	0.142826	0.140331
0.45	0.22955	0.206442	0.195447	0.188699	0.184014	0.180515	0.177769
0.5	0.273046	0.248354	0.236543	0.229273	0.224215	0.220423	0.217461
0.55	0.318425	0.292461	0.279971	0.272259	0.266883	0.262856	0.259690
0.6	0.365909	0.339037	0.326033	0.317977	0.312350	0.308127	0.304803
0.7	0.46838	0.441072	0.427681	0.419322	0.413454	0.409035	0.405547
0.8	0.583823	0.558670	0.546133	0.538236	0.532657	0.528438	0.525095
0.9	0.718981	0.700360	0.690867	0.684809	0.680493	0.677207	0.674591

π_0	p=8	p=9	p=10	p=15	p=20	p=25
0.1	0.048506	0.048944	0.049319	0.050598	0.051369	0.051899
0.15	0.024794	0.025129	0.025431	0.026540	0.027271	0.027801
0.2	0.017432	0.017101	0.016844	0.016082	0.015722	0.015526
0.25	0.038045	0.036947	0.036008	0.032820	0.030906	0.029592
0.3	0.069744	0.068443	0.067329	0.063544	0.061269	0.059706
0.35	0.103108	0.101611	0.100330	0.095971	0.093348	0.091544
0.4	0.138306	0.136626	0.135185	0.130281	0.127326	0.125291
0.45	0.175540	0.173688	0.172100	0.166686	0.163420	0.161168
0.5	0.215046	0.213040	0.211317	0.205438	0.201886	0.199435
0.55	0.257114	0.254972	0.253131	0.246843	0.243037	0.240408
0.6	0.302095	0.299843	0.297904	0.291277	0.287259	0.284480
0.7	0.402698	0.400326	0.396279	0.391262	0.386990	0.384028
0.8	0.522358	0.520075	0.518096	0.511297	0.507138	0.504243
0.9	0.672439	0.670643	0.669072	0.663665	0.660336	0.658006