Testing a Point Null Hypothesis against One-Sided for Non Regular and Exponential Families: The Reconcilability Condition to P-values and Posterior Probability

Parisa Zolfaghari [1]**, Rahim Chinipardaz** [1,2]**, and Jafar Esmaily** [3]**.**

[1] Department of Statistics, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran.

[2] Department of Statistics, Faculty of Mathematics and Computer Science, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

[3] Department of Mathematics, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran.

Abstract. In this paper, the reconcilability between the P-value and the posterior probability in testing a point null hypothesis against the one-sided hypothesis is considered. Two essential families, non regular and exponential family of distributions, are studied. It was shown in a non regular family of distributions; in some cases, it is possible to find a prior distribution function under which P-value and posterior probability are achieved. However, in the exponential family of distributions, this agreement is based on the complete monotonicity of a function of hazard rate.

Keywords. Completely Monotone, Exponential Family of Distributions, Non Regular Family of Distributions, P-values, Posterior Probability.

MSC: 62F15, 62F03.

_____

Parisa Zolfaghari (parisa-zolfaghari@iauahvaz.ac.ir)

Corresponding Author: Rahim Chinipardaz (chinipardaz_r@scu.ac.ir)

Jafar Esmaily (esmaily@iauahvaz.ac.ir)

1   Introduction

Different methods are used to evaluate a statistical hypothesis. The two main methods are significance testing and the Bayesian test. In a significance testing, the smallness of the P-value means that observations do not support the null hypothesis, and in the Bayesian test, the large posterior probability indicates the support of H_0. Bayesian statisticians have claimed that the smallness of the P-value cannot be considered as a reason for the rejection of H_0 at least when H_0 is a point null and tested against a two-sided alternative hypothesis. They have shown a hypothesis, for a given observation, rejected by P-value while having a high posterior probability. A necessary consequence of many Bayesians is that the P-value is a misleading measure of evidence and has to be banned or abandoned (see Lindley (1957), Nickerson (2000), and Mayo (2006)). For many other Bayesians, the p-value needs to be re-calibrated or requirs some transformation (see e. g. Sellke et al. (2001) and Efron and Gous (2001)). In a landmark paper, for testing the mean of a normal distribution, Edwards et al. (1963) demonstrated that the lower bound of the posterior probability of H_0 for a class of prior distributions is always higher than the corresponding P-value. This issue was expanded by Berger and Sellke (1987), Berger and Delampady (1987), and Chinipardaz (2003) who used different classes of priors and showed that the P-value overstates the evidence against H_0. Although Berger and Delampady (1987), Delampady (1989) and Verdinell and Wessermann (1996) verified that the posterior probability of an imprecise null hypothesis ($H_0 : |\theta - \theta_0| \leq \epsilon$ against $H_1 : |\theta - \theta_0| > \epsilon$ with small constant ϵ) leads to the same results even in the presence of nuisance parameter, many statisticians showed that for one-sided hypotheses testing, $H_0 : \theta \leq \theta_0 (\theta \geq \theta_0)$ against $H_1 : \theta > \theta_0 (\theta < \theta_0)$ two approaches are reconcilable (see e.g. Berger (1985) and Casella and Berger (1987)).

We consider a test $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0 (\theta < \theta_0)$. It is a point null hypothesis against the one-sided hypothesis and has attracted considerable attention in various applied settings. For example, in a dose-response experiment, the response is often expected to be non-decreasing with dose. In such a case, one may be interested in testing whether the mean response is equal or non-decreasing with dose, which is equivalent to testing the simple hypothesis against the one-sided hypothesis on the mean. In econometric modeling, econometric theory often suggests a linear inequality constraint on a normal linear regression model's coefficient. It is the test for linear equality against linear inequality. In this article, we followed this idea to give a prior distribution, if any, for which the P-value and the postorior probability reconciled. The materials organized as follows; in Section 2, we set up the problem for point null against a one-sided hypothesis. In Section 3, we look at the non regular family of

distributions and generalize the results in Chinipardaz and Noorbaloochi (2003) for a point null against a one-sided hypothesis. In Section 4, we follow the problem for an exponential family of distributions. We will show that deriving a prior distribution of an agreement is based on the complete monotonicity of the hazard rate. Finally, in Section 5, we present our discussion as well as the conclusion.

2 Setting up the Problem and Preliminary Results

Consider an observation X that is thought to have $f_\theta(x)$ distribution, where x and the parameter θ are real numbers. We wish to test $H_0 : \theta = \theta_0$ test against $H_1 : \theta > \theta_0$. In a Bayesian framework, one considers a discrete-continuous prior distribution

$$\pi(\theta) = \pi_0 I_{\{\theta_0\}}(\theta) + (1 - \pi_0)\pi_1(\theta)I_{(\theta_0,\infty)}(\theta),$$

where $\pi_0 = P(H_0)$ is the probability of H_0 to be true in prior and $I_A(w)$ is 0-1 indicator function, 1 if $w \in A$ and 0 otherwise. Here, $\pi_1(\theta)$ is the density, which describes how the prior mass is spread out over the alternative hypothesis. Then the posterior probability of H_0 is

$$P(H_0|x) = P(\theta = \theta_0|x) = \frac{\pi_0 f_{\theta_0}(x)}{\pi_0 f_{\theta_0}(x) + (1 - \pi_0)m_{\pi_1}(x)},$$

where $m_{\pi_1}(x) = \int_{\theta_0}^\infty f_\theta(x)\pi_1(\theta)d\theta$ is the predictive function on (θ_0, ∞). To have equal posterior probability and P-value, $p(x)$, we require

$$\int_{\theta_0}^\infty f_\theta(x)\pi_1(\theta)d\theta = \frac{\pi_0}{1 - \pi_0}\frac{1 - p(x)}{p(x)}f_{\theta_0}(x).$$

To be an impartial statistician, the prior odds ratio, $\frac{1-\pi_0}{\pi_0}$, is taken to be one. Therefore, to reconcile between Bayesians and frequentists, we have to seek a prior, $\pi_1(\theta)$, on (θ_0, ∞) satisfying

$$\int_{\theta_0}^\infty f_\theta(x)\pi_1(\theta)d\theta = \frac{1 - p(x)}{p(x)}f_{\theta_0}(x). \tag{2.1}$$

Similar results can be obtained in the same manner for $H_0 : \theta = \theta_0$ against $H_1 : \theta < \theta_0$ with the prior distribution

$$\pi(\theta) = \pi_0 I_{\{\theta_0\}}(\theta) + (1 - \pi_0)\pi_1(\theta)I_{[0,\theta_0)}(\theta).$$

In this case, we seek for $\pi_1(\theta)$ on $[0, \theta_0)$ such that

$$\int_0^{\theta_0} f_\theta(x)\pi_1(\theta)d\theta = \frac{1 - p(x)}{p(x)} f_{\theta_0}(x). \qquad (2.2)$$

One might argue that this is a pertinacious requirement and should be simplified such that this condition is satisfied only with a set of observations, say C, under which the null hypothesis rejected. There is

$$C = \{x; p(x) < \alpha\},$$

where α is a small pre-assigned value.

Examples 2.1. Consider $X_1, \ldots, X_n \sim U(0, \theta)$. To test $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$, the considered test statistic is $T(X) = X_{(n)} = \max_{1 \le i \le n}\{X_i\}$ and P-value is given by

$$p(t) = P_{\theta_0}(X_{(n)} \ge t) = \int_t^{\theta_0} \frac{nu^{n-1}}{\theta_0^n}du = 1 - \left(\frac{t}{\theta_0}\right)^n,$$

where t is the observed value of T. Now, we seek $\pi_1(\theta)$ satisfying

$$\int_{\theta_0^+}^{\infty} f_\theta(t)\pi_1(\theta)d\theta = \frac{1 - p(t)}{p(t)} f_{\theta_0}(t) = \frac{(\frac{t}{\theta_0})^n}{1 - (\frac{t}{\theta_0})^n} \frac{nt^{n-1}}{\theta_0^n}.$$

The left hand side of this relation is

$$\int_{\theta_0^+}^{\infty} \frac{nt^{n-1}}{\theta^n}\pi_1(\theta)d\theta = nt^{n-1}\int_{\theta_0}^{\infty} \frac{1}{\theta^n}\pi_1(\theta)d\theta.$$

There is no prior distribution of $\pi_1(\theta)$ satisfying this condition.

In this example for $H_0 : \theta = \theta_0$ against $H_1 : \theta < \theta_0$ the P-value is $P_{\theta_0}(X_{(n)} \le \mathbf{t}) = \left(\frac{t}{\theta_0}\right)^n$. Besides, the right hand side of equation (2.2) is $nt^{n-1}\left[\left(\frac{1}{t}\right)^n - \left(\frac{1}{\theta_0}\right)^n\right]$ and we need $\pi_1(\theta)$ to satisfy

$$\int_t^{\theta_0} \frac{nt^{n-1}}{\theta^n}\pi_1(\theta)d\theta = nt^{n-1}\int_t^{\theta_0} \frac{1}{\theta^n}\pi_1(\theta)d\theta = nt^{n-1}\left[\left(\frac{1}{t}\right)^n - \left(\frac{1}{\theta_0}\right)^n\right].$$

It can be demonstrated that using $\pi_1(\theta) = \frac{n}{\theta}$, the posterior probability of H_0 and P-value are reconciled. To show that this is not the only case satisfying this condition, in the following, we look at a non regular family of distributions.

3 Non Regular Family of Distributions

Suppose we have the random variable X from a non regularfamily of distributions which satisfies one of the folllowing forms:

$(i) f_\theta(x) = Q(\theta)h(x)$, $x \leq \theta$, $Q(\theta)$ is a nonincreasing function of θ and $h(x)$ is a nondecreasing function of x.

$(ii) f_\theta(x) = Q(\theta)h(x)$, $x \geq \theta$, $Q(\theta)$ is a nondecreasing function of θ and $h(x)$ is a nonincreasing function of x.

It is straightforward to obtain the following lemma based on the definition of P-value.

Lemma 3.1. *To test* $H_0 : \theta = \theta_0$ *against* $H_1 : \theta > \theta_0 (\theta < \theta_0)$ *in non regular family of distributions, P-value is given by:*

$$p - value = \begin{cases} Q(\theta_0)G_1(x) & H_1 : \theta > \theta_0 \ (Q(\theta) \text{ nonincreasing}) \\ Q(\theta_0)G_2(x) & H_1 : \theta < \theta_0 \ (Q(\theta) \text{ nonincreasing}) \\ Q(\theta_0)G_3(x) & H_1 : \theta > \theta_0 \ (Q(\theta) \text{ nondecreasing}) \\ Q(\theta_0)G_4(x) & H_1 : \theta < \theta_0 \ (Q(\theta) \text{ nondecreasing}), \end{cases}$$

where $G_1(x) = \int_x^{\theta_0} h(u)du$, $G_2(x) = \int_0^x h(u)du$, $G_3(x) = \int_x^\infty h(u)du$, $G_4(x) = \int_{\theta_0}^x h(u)du$.

Proof. For the case of $H_1 : \theta > \theta_0$, $p(x) = P_{\theta_0}(X \geq x) = Q(\theta_0)G_i(x), i = 1, 3$, and for $H_1 : \theta < \theta_0$, $p(x) = P_{\theta_0}(X \leq x) = Q(\theta_0)G_i(x), i = 2, 4$, are P-values. It should be noted that the case of $p(x) = 0$ is dropped because the posterior probability of H_0 also vanishes. □

Using this lemma, to set equal P-values with corresponding posterior probabilities

$$\frac{Q(\theta_0)h(x)}{Q(\theta_0)h(x) + \int_{\Theta_1} Q(\theta)h(x)\pi_1(\theta)d\theta} = Q(\theta_0)G_i(x), i = 1, 2, 3, 4.$$

and after some manipulations, we get Table1. For all the cases, we need to have $\pi_1(\theta)$ satisfying the equation in the last column of Table 1.

3.1 Distribution of θ

There are four cases that can be considered to obtain $\pi_1(\theta)$. However, we show that only two of them lead to the prior distribution $\pi_1(\theta)$. Consider the second case where

Table 1: Relations (2.1) or (2.2) when P-value and posterior probability are taken to be equal.

$Q(\theta)$	H_1	Constraint	(2.1) or (2.2)
nonincreasing	$\theta > \theta_0$	$x \leq \theta$	$G_1(x) \int_{\theta_0}^{\infty} Q(\theta)\pi_1(\theta)d\theta + Q(\theta_0)G_1(x) = 1$
nonincreasing	$\theta < \theta_0$	$x \leq \theta$	$G_2(x) \int_{x}^{\theta_0} Q(\theta)\pi_1(\theta)d\theta + Q(\theta_0)G_2(x) = 1^*$
nondecreasing	$\theta > \theta_0$	$x \geq \theta$	$G_3(x) \int_{\theta_0}^{x} Q(\theta)\pi_1(\theta)d\theta + Q(\theta_0)G_3(x) = 1$
nondecreasing	$\theta < \theta_0$	$x \geq \theta$	$G_4(x) \int_{0}^{\theta_0} Q(\theta)\pi_1(\theta)d\theta + Q(\theta_0)G_4(x) = 1$

$H_0 : \theta = \theta_0$ is tested against $H_1 : \theta < \theta_0$ when $Q(\theta)$ is a nonincreasing function of θ. Define $\bar{\pi}_1(\theta) = Q(\theta)\pi_1(\theta)$ and set $F(x) = Q(\theta_0) + \int_{x}^{\theta_0} \bar{\pi}_1(\theta)d\theta$. In the second row of Table 1, marked by $*$, we have

$$F(x)G_2(x) = 1. \tag{3.1}$$

Therefore, $Q(\theta_0) + \int_{x}^{\theta_0} \bar{\pi}_1(\theta)d\theta = \frac{1}{G_2(x)}$ and differentiate with respect to x results in

$$-\bar{\pi}_1(x) = \frac{-h(x)}{G_2^2(x)} \implies \bar{\pi}_1(x) = \frac{h(x)}{G_2^2(x)},$$

and therefore, $\pi_1(\theta) = \frac{h(\theta)}{G_2^2(\theta)Q(\theta)}$.

We can use this prior distribution to obtain the posterior probability. One can check the result by putting this prior distribution to have equal P-value and posterior probability. In the second row of Table 1:

$$
\begin{aligned}
G_2(x) \int_{x}^{\theta_0} Q(\theta)\frac{h(\theta)}{G_2^2(\theta)Q(\theta)}d\theta + Q(\theta_0)G_2(x) &= G_2(x)\frac{-1}{G_2(\theta)}\Big|_{x}^{\theta_0} + Q(\theta_0)G_2(x) \\
&= \left[\frac{-G_2(x)}{G_2(\theta_0)} + 1\right] + Q(\theta_0)G_2(x) \\
&= 1.
\end{aligned}
$$

Since $\int_{0}^{\theta_0} h(x)Q(\theta_0)dx = 1$, we have $G_2(\theta_0) = \frac{1}{Q(\theta_0)}$. The reader can follow this method to obtain $\pi_1(\theta)$ for $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$ when $Q(\theta)$ is a nondecreasing

function of θ. Interestingly, for the first and forth rows in Table 1 ($H_1 : \theta > \theta_0$ with $Q(\theta)$ a nonincreasing function and $H_1 : \theta < \theta_0$ with $Q(\theta)$ a nondecreasing function) the method fails to obtain $\pi_1(\theta)$.

Examples 3.1. Consider $X \sim U(0, \theta)$ and the hypothesis testing procedure $H_0 : \theta = \theta_0$ against $H_1 : \theta < \theta_0$. From the second row of Table 1, we have

$$\pi_1(\theta) = \frac{h(\theta)}{G_2^2(\theta)Q(\theta)} = \frac{1}{\frac{1}{\theta}(\int_0^\theta d\theta)^2} = \frac{1}{\theta}.$$

Using this prior distribution, we get

$$P(\theta = \theta_0|x) = \frac{\frac{1}{\theta_0}}{\frac{1}{\theta_0} + \int_x^{\theta_0} \frac{1}{\theta}\frac{1}{\theta}d\theta} = \frac{x}{\theta_0} = P_{\theta_0}(X \le x) = p(x).$$

3.2 Extension to a Random Sample

Let $X_1, \ldots, X_n$ be members of the non regular family of distributions, taken the form (*i*) or (*ii*). The sufficient statistic, $X_{(n)}(X_{(1)})$, functions as the test statistic in both significance testing and the Bayesian test. As the non regular family of distributions is close under the sufficient statistic, the distribution of $X_{(n)}$ and $X_{(1)}$ are, respectively,

$$f_{X_{(n)}}(t|\theta) = Q^n(\theta)m_1(t), \quad t \le \theta$$
$$f_{X_{(1)}}(t|\theta) = Q^n(\theta)m_2(t), \quad t \ge \theta,$$

where $m_1(t) = nh(t)N_1^{n-1}(t)$, $m_2(t) = nh(t)N_2^{n-1}(t)$, $N_1(t) = \int_0^t h(u)du$ and $N_2(t) = \int_t^\infty h(u)du$. To test $H_0 : \theta = \theta_0$ agaist $H_1 : \theta > \theta_0(\theta < \theta_0)$, we have the following lemma.

Lemma 3.2. *To test* $H_0 : \theta = \theta_0$ *against* $H_1 : \theta > \theta_0(\theta < \theta_0)$ *in non regular family of distributions, P-value is given by:*

$$p - value = \begin{cases} Q^n(\theta_0)M_1(t) & H_1 : \theta > \theta_0 \ (Q(\theta) \text{ nonincreasing}) \\ Q^n(\theta_0)M_2(t) & H_1 : \theta < \theta_0 \ (Q(\theta) \text{ nonincreasing}) \\ Q^n(\theta_0)M_3(t) & H_1 : \theta > \theta_0 \ (Q(\theta) \text{ nondecreasing}) \\ Q^n(\theta_0)M_4(t) & H_1 : \theta < \theta_0 \ (Q(\theta) \text{ nondecreasing}), \end{cases}$$

where $M_1(t) = \int_t^{\theta_0} m_1(u)du$, $M_2(t) = \int_0^t m_1(u)du$, $M_3(t) = \int_t^\infty m_2(u)du$, $M_4(t) = \int_{\theta_0}^t m_2(u)du$.

Now, to equate P-value and posterior probability, Table 1 is modified to Table 2.

Table 2: Relations (2.1) or (2.2) when P-value and posterior probability are taken to be equal.

$Q(\theta)$	H_1	Constraint	(2.1) or (2.2)
nonincreasing	$\theta > \theta_0$	$t \leq \theta$	$M_1(t) \int_{\theta_0}^{\infty} Q^n(\theta)\pi_1(\theta)d\theta + Q^n(\theta_0)M_1(t)=1$
nonincreasing	$\theta < \theta_0$	$t \leq \theta$	$M_2(t) \int_{t}^{\theta_0} Q^n(\theta)\pi_1(\theta)d\theta + Q^n(\theta_0)M_2(t) = 1^*$
nondecreasing	$\theta > \theta_0$	$t \geq \theta$	$M_3(t) \int_{\theta_0}^{t} Q^n(\theta)\pi_1(\theta)d\theta + Q^n(\theta_0)M_3(t) = 1$
nondecreasing	$\theta < \theta_0$	$t \geq \theta$	$M_4(t) \int_{0}^{\theta_0} Q^n(\theta)\pi_1(\theta)d\theta + Q^n(\theta_0)M_4(t) = 1$

Here, the same method can be applied to obtain the prior distribution. In this case for $H_0 : \theta = \theta_0$ against $H_1 : \theta < \theta_0$ in $f_\theta(x) = Q(\theta)h(x), x \leq \theta$,we have

$$\pi_1(\theta) = \frac{m_1(\theta)}{M_2^2(\theta)Q^n(\theta)},$$

and for $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$, in $f_\theta(x) = Q(\theta)h(x), x \geq \theta$

$$\pi_1(\theta) = \frac{m_2(\theta)}{M_3^2(\theta)Q^n(\theta)},$$

where $m_1(\theta), m_2(\theta), M_2^2(\theta)$ and $M_3^2(\theta)$ are given as before.

Examples 3.2. To extend example 3.1 for n random variables, consider $X_1, \ldots, X_n \sim U(0, \theta)$. For testing the hypothesis $H_0 : \theta = \theta_0$ against $H_1 : \theta < \theta_0$, we used the test statistic $T(X) = X_{(n)}$ then we have $Q(\theta) = \theta^{-1}$. Therefore,

$$\pi_1(\theta) = \frac{n\theta^{n-1}}{\frac{1}{\theta^n}(\int_0^\theta nu^{n-1}du)^2} = \frac{n}{\theta}.$$

Using this prior distribution, we have

$$P(\theta = \theta_0|t) = \frac{\frac{nt^{n-1}}{\theta_0^n}}{\frac{nt^{n-1}}{\theta_0^n} + \int_t^{\theta_0} \frac{nt^{n-1}}{\theta^n} \frac{n}{\theta} d\theta} = (\frac{t}{\theta_0})^n = P_{\theta_0}(X_{(n)} \leq t) = p(t).$$

Examples 3.3. Let $X_1, \dots, X_n$ are independent with pdf $f(x) = \lambda e^{-\lambda(x-\theta)}, x \geq \theta$. To test $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$, we used the test statistics $T(X) = X_{(1)}$ then we have $Q(\theta) = e^{n\lambda\theta}$. Therefore,

$$\pi_1(\theta) = \frac{n\lambda e^{-n\lambda\theta}}{e^{-2n\lambda\theta}e^{n\lambda\theta}} = n\lambda.$$

Using this prior distribution, we have

$$P(\theta = \theta_0|t) = \frac{n\lambda e^{-n\lambda(t-\theta_0)}}{n\lambda e^{-n\lambda(t-\theta_0)} + \int_{\theta_0}^{t} n\lambda e^{-n\lambda(t-\theta)}d\theta} = e^{-n\lambda(t-\theta_0)} = P_{\theta_0}(X_{(1)} \geq t) = p(t).$$

3.3 Other Alternatives

The results in the previous sections can be considered for two special cases in the alternative hypothesis.

3.3.1 Point Null against Point Alternative Hypothesis

Consider point null against point alternative hypothesis tests, $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$. Without loss of generality, we assume $\theta_1 > \theta_0$ and $Q(\theta)$ is a nonincreasing function of θ. Then, $p(x) = Q(\theta_0)G(x)$, where $G(x) = \int_x^{\theta_0} h(u)du$. The H_0 posterior probability is

$$P(\theta = \theta_0|x) = \frac{\pi_0 Q(\theta_0)h(x)}{\pi_0 Q(\theta_0)h(x) + (1-\pi_0)Q(\theta_1)h(x)} = \frac{Q(\theta_0)}{Q(\theta_0) + \frac{1-\pi_0}{\pi_0}Q(\theta_1)}.$$

For the equality of $p(x)$ and $p(\theta = \theta_0|x)$ we need π_0 to satisfy

$$Q(\theta_0) + \frac{1-\pi_0}{\pi_0}Q(\theta_1) = \frac{1}{G(x)},$$

for all x, i.e. the prior distribution depends on x which is meaningless.

This statement holda true no only for this class of distributions, but also for all distributions.

3.3.2 Null Point against Closed Interval Alternative Hypothesis

Another case is $H_0 : \theta = \theta_0$ against $H_1 : \theta \in (\theta_0, b]$ where b is some real finite value. This case includes a number of real examples.

Because the bounds of integral in third rows of Table 1 and Table 2 are changed from x to b, the method fails to obtain a $\pi_1(\theta)$ prior distribution.

4   Exponential Family of Distributions

Consider one-parameter exponential family of distributions as

$$f_\theta(x) = e^{\eta(\theta)T(x)}P(\theta)h(x),$$

where η and P are real-valued functions of the parameter and $T(x)$ is a real-valued statistic. Often, it is more convenient to use canonical form given by

$$f_\theta(x) = e^{\theta T(x)}Q(\theta)h(x),$$

or even simpler natural form as

$$f_\theta(x) = e^{-\theta x}Q(\theta)h(x). \tag{4.1}$$

As the exponential family of distributions is closed under-sampling, we use only one observation, x. The result is not different because $T(X)$ will belong to the exponential family of distributions.

The parameter space considered in this paper is $\theta = [0, \infty)$ or $(-\infty, 0]$. Without loss of generality, suppose that $H_0 : \theta = \theta_0$ is tested against $H_1 : \theta > \theta_0$ (or $H_1 : \theta < \theta_0$). It can be demonstrated that

$$
\begin{aligned}
m_{\pi_1}(x) &= \int_{\Theta_1} f_\theta(x)\pi_1(\theta)d\theta \\
&= \int_{\Theta_1} e^{-\theta x}h(x)Q(\theta)\pi_1(\theta)d\theta \\
&= h(x)\int_{\Theta_1} e^{-\theta x}Q(\theta)\pi_1(\theta)d\theta.
\end{aligned}
$$

For the equality of P-value and posterior probability under H_0, it is required that

$$\int_{\Theta_1} e^{-\theta x}Q(\theta)\pi_1(\theta)d\theta = \frac{1 - p(x)}{p(x)}\frac{f_{\theta_0}(x)}{h(x)} = \phi(x),$$

where $p(x) = P_{\theta_0}(X \leq x)$ or $P_{\theta_0}(X \geq x)$ depend on $H_1 : \theta > \theta_0$ or $H_1 : \theta < \theta_0$, respectively. To continue, define $\bar{\pi}(\theta) = Q(\theta)\pi_1(\theta)I_{\Theta_1}(\theta)$. Then,

$$\int_0^\infty e^{-\theta x}\bar{\pi}(\theta)d\theta = \frac{1 - p(x)}{p(x)}\frac{f_{\theta_0}(x)}{h(x)} = \phi(x),$$

and $\phi(x)$ is the Laplace transform of $\bar{\pi}(\theta)$. Therefore, there is a proper prior distribution for θ if and only if $\phi(x)$ is completely monotone and $\phi(0) = 1$, i.e.

$$(-1)^n\phi^{[n]}(x) \geq 0, \quad \phi(0) = 1, \tag{4.2}$$

and there is prior distribution, not necessary proper, when $\phi(x)$ is completely monotone and $\phi(0) \neq 1$. Here $\phi^{[n]}(x)$ is the nth derivative of $\phi(x)$ with respect to x. If $\phi(x)$ is discrete, the result is true with $\Delta\phi(x) = \phi(x + 1) - \phi(x)$ and $\Delta^{[j]}(x) = \Delta\Delta^{[j-1]}(x)$ (see Appendix). To check the complete monotonicity of $\phi(x)$, one can remove the positive and constant value $Q(\theta_0)$ for simplicity from (4.1).

Examples 4.1. Let X has the exponential distribution with pdf $f_\theta(x) = \theta e^{-\theta x}, x > 0, \theta > 0$. To test $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$, $p(x) = P_{\theta_0}(X \leq x) = 1 - e^{-\theta_0 x}$ and

$$\phi(x) = \frac{1 - F_{\theta_0}(x)}{F_{\theta_0}(x)}f_{\theta_0}(x) = \frac{\theta_0 e^{-2\theta_0 x}}{1 - e^{-\theta_0 x}}.$$

We have

$$\begin{aligned}
\phi^{[n]}(x) &= \theta_0\left[(-2\theta_0)^n e^{-2\theta_0 x} + (-3\theta_0)^n e^{-3\theta_0 x} + \cdots\right] \\
&= \theta_0\sum_{j=2}^\infty(-j\theta_0)^n e^{-j\theta_0 x} = \theta_0^{n+1}\sum_{j=2}^\infty(-j)^n e^{-j\theta_0 x}.
\end{aligned}$$

Hence $(-1)^n\phi^{[n]}(x) \geq 0, \phi(0) = \infty \neq 1$. Therefore, there is an improper prior distribution to reconcile these two approaches. Now, suppose we want to test $H_0 : \theta = \theta_0$ against $H_1 : \theta < \theta_0$. In this case

$$\phi(x) = \frac{F_{\theta_0}(x)}{1 - F_{\theta_0}(x)}f_{\theta_0}(x) = \theta_0(1 - e^{-\theta_0 x}) \Rightarrow (-1)^n\phi^{[n]}(x) = (-1)\theta_0^{(n+1)}e^{-\theta_0 x},$$

$(-1)^n\phi^{[n]}(x) \leq 0$ which is not completely monotone and therefore, there is no $\pi(\theta)$ for the agreement of P-value and the posterior probability.

Examples 4.2. Let X follows the geometric distribution with pdf $f_p(x) = pq^x$, $x = 0, 1, \ldots, q = 1 - p$. Let $\theta = -\ln(1 - p)$, then $p = 1 - e^{-\theta}$ and $f(x|\theta) = e^{-\theta x}(1 - e^{-\theta})$ belongs to an exponential family of distributions.

To test $H_0 : p = p_0$ against $H_1 : p < p_0$, the P-value is $p(x) = P_{p_0}(X \geq x) = q_0^x$, then

$$\phi(x) = \frac{F_{p_0}(x-1)}{1 - F_{p_0}(x-1)} p_0 e^{-x \ln(\frac{1}{1-p_0})} = p_0(1 - q_0^x)$$

$$\Delta\phi(x) = \phi(x+1) - \phi(x) = p_0^2 q_0^x.$$

Therefore, $(-1)\Delta\phi(x) < 0$ and there is no prior distribution to reconcile posterior probability and P-value. Now, we consider the case where $H_1 : p > p_0$. We have

$$\phi(x) = \frac{1 - F_{p_0}(x)}{F_{p_0}(x)} p_0 e^{-x \ln(\frac{1}{1-p_0})} = \frac{p_0 q_0^{2x+1}}{1 - q_0^{x+1}} \Rightarrow \Delta\phi(x) = p_0 \left[\sum_{j=2}^{\infty} q_0^{jx+(j-1)}(q_0^j - 1) \right] < 0.$$

So, $(-1)\Delta\phi(x) > 0$. Similarly, $\Delta^{[n]}\phi(x) = p_0 \sum_{j=2}^{\infty} q_0^{jx+(j-1)}(q_0^j - 1)^n$, which results in

$(-1)^n \Delta^{[n]}\phi(x) > 0$ showing that there is a prior disribution under which posterior probabilty and P-value are equal.

Table 3 presents our investigation for some common-used distributions of the exponential family of distributions.

4.1 The Prior Distribution

The method presented in Section 4 can only be used to check whether a prior distribution exists for equality of p-value and posterior probability. It does not determine, however, the structure of the prior distribution. Obtaining the prior distribution requires other techniques such as those given in Section 3. To follow this, we consider Example 4.2.

Examples 4.3. It was shown that in example 4.2, for $H_0 : p = p_0$ against $H_1 : p > p_0$, there exists a prior distribution, say $\pi_1(\theta)$ satisfy:

$$\frac{p_0(1-p_0)^x}{p_0(1-p_0)^x + \int_{p_0}^1 p(1-p)^x \pi_1(p) dp} = 1 - (1 - p_0)^{x+1}, \tag{4.3}$$

Table 3: The existence or non-existence of prior distribution for some exponential family of distributions. ("*" means that we could check only some derivatives)

Distribution	H_1	There exists $\pi(\theta)$
Exponential	$\theta > \theta_0$	Yes
Exponential	$\theta < \theta_0$	No
Poisson	$\theta > \theta_0$	No
Poisson	$\theta < \theta_0$	Yes *
Pareto	$\theta > \theta_0$	Yes
Pareto	$\theta < \theta_0$	No
Geometric	$\theta > \theta_0$	Yes
Geometric	$\theta < \theta_0$	No
Weibull	$\theta > \theta_0$	Yes
Weibull	$\theta < \theta_0$	No
Binomial	$\theta > \theta_0$	No
Binomial	$\theta < \theta_0$	Yes *

or

$$\int_{p_0}^{1} p(1-p)^x \pi_1(p)dp = \frac{p_0(1-p_0)^x}{1-(1-p_0)^{x+1}} - p_0(1-p_0)^x = \frac{p_0(1-p_0)^{2x+1}}{1-(1-p_0)^{x+1}}. \tag{4.4}$$

Differentiating (4.4) with respect to p_0 in both sides, we have

$$-p_0(1-p_0)^x \pi_1(p_0) = \frac{(1-p_0)^{2x}[1-2p_0-2xp_0]}{1-(1-p_0)^{x+1}} - \frac{(x+1)p_0(1-p_0)^{3x+1}}{[1-(1-p_0)^{x+1}]^2}.$$

Therefore,

$$\pi_1(p) = -\frac{(1-p)^x[1-2p-2xp]}{p[1-(1-p)^{x+1}]} + \frac{(x+1)(1-p)^{2x+1}}{[1-(1-p)^{x+1}]^2}.$$

Although check (4.2) complex manipulations are needed. More importantly, $\pi_1()$ changes with x, which indicates that the prior distribution is obtainable. Table 4 presents some prior distributions under which the p-value and posterior probability are equal.

It should be noted that since the prior distribution must be independent of x, there is no prior distribution to match the P-value and the posterior probability for each x.
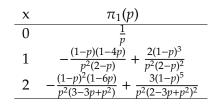
Table 4: The prior distribution to match p-value and posterior probability

x	$\pi_1(p)$
0	$\frac{1}{p}$
1	$-\frac{(1-p)(1-4p)}{p^2(2-p)} + \frac{2(1-p)^3}{p^2(2-p)^2}$
2	$-\frac{(1-p)^2(1-6p)}{p^2(3-3p+p^2)} + \frac{3(1-p)^5}{p^2(2-3p+p^2)^2}$

5   Conclusion and Comments

In this paper, we have a fresh look at the paradoxical result in testing the significance testing and the Bayesian tests. We considered the point null hypothesis against the one-sided hypothesis testing. We aim to search for a prior distribution in Bayesian testing so that the P-value and posterior probability are equal. Two different families of distributions have been examined. In the non regular family of distributions, it was shown that there are no such prior distributions, in general. However, in some cases, there are such priors, proper or improper. Obtaining such a prior for the exponential family of distributions leads to complete monotonicity of a function of the hazard rate. In most cases in this family, such suitable priors do not exist, or they are very complicated and rare to find.

References

Berger, J. O. (1985), Statistical Decision Theory and Bayesian Analysis, 2nd ed. Berlin, Springer Verlag.

Berger, J. O. and Delampady, M. (1987), Testing Precise Hypotheses (with discussion). *Statistical Science*, **2**, 317-352.

Berger, J. O. and Sellke, T. (1987), Testing a Point Null hypotheses: The Irreconcilability of P-values and Evidence (with discussion). *Journal of the American Statistical Association*, **82**, 112-139.

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

Chinipardaz, R. (2003), The Discrepancy of P-values and Posterior Probability in Poisson Distribution. *Pakistan Journal of Statistics*, **19**, 301-313.

Chinipardaz, R. and Noorbaloochi, S. (2003), The reconcilability of P-values and posterior probability in nonregular distributions. *Far east Journal of Theoretical Statistics*, **10**, 112-122.

DeGroot, M. H. (1973), Dowing what comes naturally: Interpreating a tail area as a posterior probability or likelihood ratio. *Journal of the American Statistical Association*, **68**, 966-969.

Delampady, M. (1989), Lower bound on Bayes Factors for Interval Null Hypotheses. *Journal of the American Statistical Association*, **84**(405), 120-124.

Dickey, J. M. (1977), Is the tail area useful as an approximate Bayes factor ?. *Journal of the American Statistical Association*, **72**, 138-142.

Edwards, W., Lindman, H. and Savage, L. J. (1963), Bayesian Statistical Inference for Psychological Research. *Psychological Review*, **70**(3), 193-242.

Efron, B. and Gous, A. (2001), Scale of evidence for model selection: Fisher versus Jeffrey Model Selection, In: Lecture Notes - Monograph Series. *Institute of Mathematical Statistics*, **38**, 208-256.

Feller, W. (1971), An Introduction to Probability Theory and its Applications. Vol. 2, New York, John Wiley.

Lindley, D. V., (1957), A Statistical Paradox. *Biometrika*, **64**, 207-213.

Mayo, D. (2006), Philosophy of statistics. In: S. Sarkar, S. and Pfeifer, J. Eds. The Philosophy of Science: an encyclopedia. London: Routledge, 802–815.

Nickerson, R. S. (2000), Null hypothesis significance testing: A review of on old and continuing controversy. *Psychological Methods*, **5**(2), 241-301.

Sellk, T., Bayarri, M. J. and Berger, J. O. (2001), Calibration of P-values for Testing Precise Null Hypotheses. *The American Statistician*, **55**, 62-71.

Verdinell, I. and Wessermann, L. (1998), Bayesian goodness of fit testing using infinite dimensional exponential families. *The Annals of Statistics*, **26**(4), 1215-1241.

Appendix

A Compleletly Monotone

Definition A.1. The function of ϕ on $(0, \infty)$ is said to be a completely monotone function if all of it's derivatives exist and

$$(-1)^n \phi^{[n]}(\lambda) \geq 0 \quad \lambda > 0,$$

is hold.

Lemma A.1. *The function ϕ on $(0, \infty)$ has Laplas transform of a distribution function if and only if ϕ is completely monotone and $\phi(0) = 1$.*

Proof. Feller (1971). □

Lemma A.2. *The function ϕ on $(0, \infty)$ is completely monotone if and only if*

$$\phi(\lambda) = \int_0^\infty e^{-\lambda x} dF(x), \quad \lambda > 0.$$

In this case, F is a nonnegative and increasing function but its measure is not necessarily finite.

Proof. Feller (1971). □

Definition A.2. The numerical sequence (finite or infinite) such as $c_0, c_1, \ldots, c_n$ is said to be a completely monotone if we have

$$(-1)^r \Delta^{[r]}(c_r) \geq 0 \quad r = 0, 1, \ldots,$$

where $\Delta c_i = c_{i+1} - c_i$. Note that $\Delta^r = \Delta \Delta^{r-1}$.

Lemma A.3. *The moments of $\{c_k\}$ of a distribution is a sequence of completely monotone functions with $c_0 = 1$. Conversely, a sequence of completely monotone functions such as $\{c_k\}$ with $c_0 = 1$ is a unique moments of distribution function. Function ϕ on $(0, \infty)$ is completely monotone if and only if*

$$\phi(\lambda) = \int_0^\infty e^{-\lambda x} dF(x), \quad \lambda > 0,$$

and $\phi(0) = 1$.

Proof. Feller (1971). □

Examples A.1.

$$g(x) = \frac{1}{x} \implies g^{[n]}(x) = \begin{cases} Q(nx^{-n} & n = 2k \\ nx^{-n}(-1)^n & n = 2k+1, \end{cases}$$

$g(x)$ is completely monotone, $g(0) = 1 \implies g(x) = \frac{1}{x} = \int_0^\infty e^{-xy} dF(y) = \int_0^\infty e^{-xy} dy$ F is nonnegative and nondecreasing function but its measure on $(0, \infty)$ is not finite.

Examples A.2.

$$a_n = \frac{1}{n} \implies \Delta a_n = \frac{-1}{n(n+1)}, \dots$$

$$\Delta^{[k]} a_n = \frac{(-1)^k k!}{n(n+1)(n+2)\cdots(n+k)}.$$