Int. J. Contemp. Math. Sci., Vol. 1, 2006, no. 9, 419-438

Bayesian Approach to Estimation of Accuracy in Two-sided Testing

Hsiuying Wang

Institute of Statistical Science
Academia Sinica
Taipei 115, Taiwan

Abstract

For a two-sided hypothesis testing, it is known that there does not exist UMP tests. Uniformly most powerful unbiased (UMPU) tests and likelihood ratio test are the common approaches to two-sided testing. The p-values of these tests are usually used as evidence against null hypothesis. However, there are criticisms of p-values as a measure of evidence against null hypothesis for the two-sided testing problem in literature. Thus, in this paper, evidence measures derived from Bayesian approach are proposed to be the replacements of p-values and are investigated from both decision and testing perspectives. From decision theoretic framework, the proposed evidence measures can be demonstrated as admissible estimators, however, p-value of UMPU test are not admissible estimators; from testing aspect, the tests derived from p-values and the proposed evidence measures are shown to be UMPU tests. The Bayes estimator is better than p-value from theoretical aspect and it has the same merit as the p-value in testing point of view. Therefore, evidence measures derived from Bayesian approach are recommended in this paper.

Keywords: UMPU tests; Bayes estimators; p-value; Evidence measures

1 Introduction

Let X be a normal random variable with distribution $N(\theta, \sigma^2)$, where θ is unknown and σ^2 is known. Let Θ_0 be a subset of parameter space, which is an interval $[\theta_0, \theta_1]$ or contains only one point θ_0. Consider the two sided hypothesis testing

$$H_0 : \theta \in \Theta_0 \quad versus \quad H_1 : \theta \notin \Theta_0. \tag{1}$$

It is well known that there does not exist UMP tests for two-sided hypotheses testing. In this case, it is impossible to find a good one among all tests, then we focus on the class of unbiased tests instead of all tests. A test with power function $\beta(\theta)$ is unbiased if $\beta(\theta') \geq \beta(\theta'')$ for every $\theta' \in \Theta_0^c$ and $\theta'' \in \Theta_0$. An uniformly most powerful test (UMPU) is a UMP test within the class of unibased test. It is shown in (Lehmann (1997)) that UMPU test does exist for two-sided testing in exponential families. For $\Theta_0 = [\theta_0, \theta_1]$, a level α UMPU test of (1) is of the form

$$\phi(x) = \begin{cases} 1 & \text{when} \quad x \leq c_1 \quad \text{or} \quad x \geq c_2 \\ 0 & \text{when} \quad c_1 < x < c_2, \end{cases}$$

where the c's are determined by

$$E_{\theta_0}\phi(X) = E_{\theta_1}\phi(X) = \alpha. \tag{2}$$

When Θ_0 contains only one point θ_0, a level α UMPU test of (1) is of the form

$$\phi(x) = \begin{cases} 1 & \text{when} \quad x \leq c_1 \quad \text{or} \quad x \geq c_2 \\ 0 & \text{when} \quad c_1 < x < c_2, \end{cases}$$

where the c's are determined by

$$E_{\theta_0}\left[\phi(X)\right] = \alpha \quad \text{and} \quad E_{\theta_0}\left[X\phi(X)\right] = E_{\theta_0}(X)\alpha. \tag{3}$$

The set $\{x : \phi(x) = 1\}$ is the rejection region of the test. If the observations belong to the rejection region, the null hypothesis is rejected. In this approach, the significance level α should be specified first. Another way of reporting the results of a test is to report the p-value. The p-value for a sample point x is the smallest value of α for which the sample point will lead to rejection of H_0. Although the p-value is defined in terms of α levels, it is not a significance level. We usually use the p-values as a measure of data-based evidence against H_0. However, using only p-value as evidence against null hypothesis might lead to a wrong decision (see Berger and Wolpert (1988)). Berger and Sellke (1987) and

Berger and Delampady (1987) also have other criticisms for p-values. Thus, it is essential to set up a criterion to evaluate p-values. Schaafsma, Tolboom and Van der Menlen (1989) and Hwang, Casella, Robert, Wells and Farrel (1992) use the risk

$$E\left(r(X) - I\left(\theta \in \Theta_0\right)\right)^2 \tag{4}$$

to evaluate p-value, where

$$I\left(\theta \in \Theta_0\right) = \begin{cases} 1 & \text{if} \quad \theta \in \Theta_0 \\ 0 & otherwise \end{cases}$$

and $r(x)$ is an estimator of $I\left(\theta \in \Theta_0\right)$. In a testing problem, we have to decide to reject or accept the null hypothesis. Actually, this problem can be interpreted as the problem of estimating $I\left(\theta \in \Theta_0\right)$. Hence a squared error risk of the form (4) is a sensible criterion to evaluate evidence measures against null hypothesis.

Under the criterion (4), Hwang et al (1992) show that for one-sided testing problem, p-values of UMP tests are admissible estimators of $I(\theta \in H_0)$ for some distributions and, however, for two-sided testing problem, p-values of UMPU tests are inadmissible for general distributions. Hence, for a two-sided testing problem, there exist other better estimators for $I(\theta \in H_0)$. Bayes estimators with respect to some proper priors are good candidates because they are admissible estimators. It reveals that Bayes estimators are good measures of evidence against H_0 from decision theoretic perspective. However, it is perhaps unfair to discuss this problem only from the decision theoretic framework because the merit of using p-value of UMPU test as a measure of evidence is, from testing viewpoint, that the test is a UMPU test. Thus, in this paper, we will examine the Bayes estimators of $I(\theta \in H_0)$ from testing pointview and demonstrate that Bayes estimator has the same advantage as p-value from testing aspect. First, a α level test derived from an estimator

$r(x)$ of $I(\theta \in H_0)$ is of the form

$$\phi(x) = \begin{cases} 1 & \text{if} \quad r(x) \leq k \\ 0 & \text{if} \quad r(x) > k, \end{cases} \tag{5}$$

where k is a positive constant between 0 and 1 such that $\max_{\theta \in \Theta_0} E_\theta \phi(x) = \alpha$. An estimator $r(x)$ of $I(\theta \in H_0)$ represents a confidence of θ belonging Θ_0. Thus, a reasonable level α test derived from estimator $r(x)$ should be of the form (5). Consequently, if $r(x)$ in (5) is replaced by p-value of UMPU test, then $\phi(x)$ is a UMPU test, which is the advantage that people use the p-values as a measure of evidence against null hypothesis.

In this paper, the test derived from some Bayes estimators of $I(\theta \in \Theta_0)$ will be shown to be UMPU test. Thus, these Bayes estimators have the same merit as p-values from testing point of view, but it is superior to p-values from decision theoretical point of view. Combining these results, the Bayes estimators are strong competitors of p-values as measures of evidence against null hypothesis.

In this paper, the parameter space is assumed to be the natural parameter space $(-\infty, \infty)$. When the parameter space is restricted to some subset of natural parameter space, Woodroofe and Wang (2001) show that p-value of UMP test is inadmissible from decision theoretical aspect for one-sided testing problem of Poisson distribution and purposed a Bayes estimator to be a modified p-value. Moreover, for simple hypothesis versus simple hypothesis case, Wang (2004) demonstrated that some Bayes estimator is better than the p-value of most powerful test, which reveals that Bayes estimator is a good substitute of p-value when the parameter space is restricted. In this paper, Bayes estimators are demonstrated to be a good measure of evidence for two-sided testing when the parameter space is natural parameter space. Hence, Bayes estimators are recommended to be a substitute of p-value as a measure of evidence against null hypothesis.

2 The case of $\Theta_0 = [\theta_0, \theta_1]$

In this section, we consider the case of $\Theta_0 = [\theta_0, \theta_1]$. The other case that Θ_0 contains only one point is discussed in next section. For $\Theta_0 = [\theta_0, \theta_1]$, the test derived from the Bayes estimator $\eta(x)$ of $I(\theta \in \Theta_0)$ with respect to uniform prior $U(2\theta_0 - \theta_1, 2\theta_1 - \theta_0)$ will be shown to be a UMPU test. A main mathematics technique used in this section is solving equations of cubic polynomials.

Lemma 1. *(Jacobson (1985), p.260, exercise 2) Let $f(x) = x^3 - px^2 + qx - r$ be a cubic polynomial, where p, q and r are three real numbers (p, q and r $\in \Re$). Then $f(x)$ has three real distinct roots if*

$$\triangle(f) = p^2q^2 - 4p^3r + 18pqr - 4q^3 - 27r^2 > 0, \tag{6}$$

where $\triangle(f)$ is called the discriminant of $f(x)$.

Lemma 2. *Let*

$$\eta(x) = \int_{\theta_0}^{\theta_1} g(\theta - x)d\theta \Big/ \int_{2\theta_0-\theta_1}^{2\theta_1-\theta_0} g(\theta - x)d\theta,$$

where $g(x - \theta) = e^{-(x-\theta)^2/2}/\sqrt{2\pi}$. For every x, there exist a ω such that $g(x - \theta_0) = g(\omega - \theta_1)$ and $\eta(x) = \eta(\omega)$.

Proof. For any x, the only two solutions of ω such that

$$g(x - \theta_0) = g(\omega - \theta_1)$$

are $\omega = x - \theta_0 + \theta_1$ and $\omega = -x + \theta_0 + \theta_1$. Note that $\eta(x)$ and $\eta(\omega)$ can be rewritten as

$$\frac{\int_{\theta_0-x}^{\theta_1-x} e^{-\frac{t^2}{2}}dt}{\int_{2\theta_0-\theta_1-x}^{2\theta_1-\theta_0-x} e^{-\frac{t^2}{2}}dt}$$

and

$$\frac{\int_{\theta_0-\omega}^{\theta_1-\omega} e^{-\frac{t^2}{2}}dt}{\int_{2\theta_0-\theta_1-\omega}^{2\theta_1-\theta_0-\omega} e^{-\frac{t^2}{2}}dt}. \tag{7}$$

If w is replaced with $-x+\theta_0+\theta_1$ in (7), then by the fact that $e^{-t^2/2}$ is symmetric to zero, we have $\eta(\omega)=\eta(x)$. Thus, ω can be chosen as $-x+\theta_0+\theta_1$.

$\square$

Lemma 3. *Let*

$$d(x)=(1-k)\int_{\theta_0}^{\theta_1} e^{-(\theta-x)^2/2}d\theta - k\int_{2\theta_0-\theta_1}^{\theta_0} e^{-(\theta-x)^2/2}d\theta - k\int_{\theta_1}^{2\theta_1-\theta_0} e^{-(\theta-x)^2/2}d\theta. \tag{8}$$

(i) When $k<1/3$, $d(x)$ has a local maximum at $(\theta_0+\theta_1)/2$ for $\theta_1>\theta_0$.

(ii) When $k\geq 1/3$, $d(x)$ has a local maximum at $(\theta_0+\theta_1)/2$ for $\theta_1-\theta_0\geq \sqrt{\ln 3k}$.

Proof. The first and second derivates of $d(x)$ with respect to x are

$$\begin{aligned}&-(1-k)\left[e^{-(\theta_1-x)^2/2}-e^{-(\theta_0-x)^2/2}\right]+k\left[e^{-(\theta_0-x)^2/2}-e^{-(2\theta_0-\theta_1-x)^2/2}\right]\\&+k\left[e^{-(2\theta_1-\theta_0-x)^2/2}-e^{-(\theta_1-x)^2/2}\right]\end{aligned} \tag{9}$$

and

$$\begin{aligned}&e^{-(\theta_1-x)^2/2}(x-\theta_1)-e^{-(\theta_0-x)^2/2}(x-\theta_0)+ke^{-(2\theta_0-\theta_1-x)^2/2}(x-2\theta_0+\theta_1)\\&-ke^{-(2\theta_1-\theta_0-x)^2/2}(x-2\theta_1+\theta_0).\end{aligned} \tag{10}$$

If x is replaced with $(\theta_0+\theta_1)/2$ in (10), then (10) is equal to

$$(\theta_1-\theta_0)e^{-9(\theta_1-\theta_0)^2/8}\left[3k-e^{(\theta_1-\theta_0)^2}\right]. \tag{11}$$

Since $(\theta_0+\theta_1)/2$ is a root of $\frac{\partial}{\partial x}d(x)$ and (10) is the second derivate of $d(x)$, if we can show that (11) is less than zero, then $d(x)$ has a local maximum at $(\theta_0+\theta_1)/2$. When $k<1/3$, (11) is less than zero. When $k\geq 1/3$, (11) is less than zero if and only if $\theta_1-\theta_0\geq\sqrt{\ln 3k}$. Hence the proof is completed. $\square$

Lemma 4. *Let $\xi(x) = F(x/2) - F(-x/2) + kF(-3x/2) - kF(3x/2)$, $x > 0$, where $F(\cdot)$ is the cumulative function of standard normal distribution $N(0,1)$. Then when $k < 1/3$, $\xi(x)$ is positive and, when $k \geq 1/3$, $\xi(x)$ is positive if and only if $x > \nu$, where ν satisfies $\xi(x) = 0$.*

Proof.

$$\frac{\partial}{\partial x}\xi(x) = e^{-x^2/8}(1 - 3ke^{-x^2}). \tag{12}$$

(i) When $k < 1/3$, $\frac{\partial}{\partial x}\xi(x) > 0$ for $-\infty < x < \infty$. And $\xi(x) = 0$ when $x = 0$ and $\xi(x) = 1 - k$ when $x = \infty$. The figure of $\xi(x)$ to x is Figure 1. Note that $\xi(x)$ is positive when x is greater than zero.

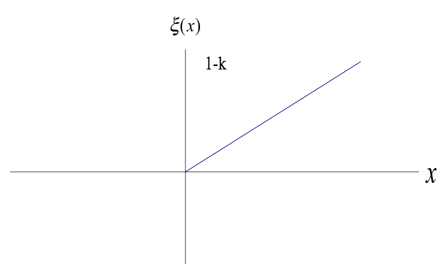

Figure 1: Plot of $\xi(x)$

(ii) When $k \geq 1/3$, $\frac{\partial}{\partial x}\xi(x) = 0$ if and only if $x = \infty$ and $\sqrt{\ln 3k}$. By the fact that $\xi(x) = 1 - k$ at $x = \infty$, $\xi(0) = 0$, $\frac{\partial}{\partial x}\xi(x) < 0$ when $0 < x < \sqrt{\ln 3k}$ and $\frac{\partial}{\partial x}\xi(x) > 0$ when $x > \sqrt{\ln 3k}$. It leads to $\xi(\sqrt{\ln 3k}) < 0$ and the figure of $\xi(x)$ to x is Figure 2. Note that ν is greater than $\sqrt{\ln 3k}$.

□

Combining Lemma 1-4, we have the main result in Theorem 1.

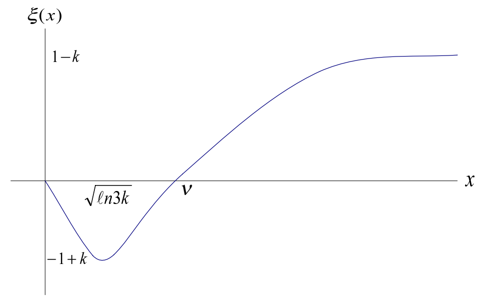

Figure 2: Plot of $\xi(x)$

Theorem 1. *Let X be a normal random variable with distribution $N(\theta, \sigma^2)$, where σ^2 is known. In hypotheses testing (1), consider the test*

$$\phi(x) = \begin{cases} 1 & \text{if} \quad \eta(x) \le k \\ 0 & \text{if} \quad \eta(x) > k, \end{cases}$$

where $\eta(x)$ is a Bayes estimator of $I(\theta \in H_0)$ with respect to uniform prior $U(2\theta_0 - \theta_1, 2\theta_1 - \theta_0)$ and k is a positive constant such that $E_{\theta_0}\phi(X) = \alpha$, then

(i) When $k \le 1/3$, $\phi(x)$ is α- level UMPU test for θ_0 and θ_1 satisfying $\theta_0 < \theta_1$.

(ii) When $k > 1/3$, $\phi(x)$ is a α-level UMPU test for θ_0 and θ_1 satisfying $\theta_1 - \theta_0 > \nu$, where ν is defined in Lemma 4.

Proof. Without loss of generality, we assume that $\sigma^2 = 1$. The rejection region of null hypothesis is

$$\left\{ x : \frac{\int_{\theta_0}^{\theta_1} f(\theta - x) d\theta}{\int_{2\theta_0 - \theta_1}^{2\theta_1 - \theta_0} f(\theta - x) d\theta} \le k \right\}, \tag{13}$$

which is equal to $\{x : d(x) \le 0\}$, where $d(x)$ is defined in (8). The first derivative of $d(x)$ is (9).

Let $m = \theta_1 - \theta_0$ and $y = e^{m(x-\theta_0)}$. Then, (9) can be rewritten as

$$ke^{\left(-\frac{(x-\theta_0)^2}{2} - 2m^2\right)} \left[y^3 - e^{3m^2/2} y^2/k + e^{2m^2} y/k - e^{3m^2/2} \right].$$

Let

$$f(y) = y^3 - \left(e^{3m^2/2}/k\right) y^2 + \left(e^{2m^2}/k\right) y - e^{3m^2/2}. \tag{14}$$

By Lemma 1, $\triangle(f(y)) = e^{3m^2}\left[e^{4m^2} - 8ke^{3m^2} - 27k^4\right]/k^4$, which is equal to

$$e^{3m^2}\left(e^{m^2} - 3k\right)^3\left(e^{m^2} + k\right)/k^4. \tag{15}$$

Note that k is greater than zero, thus, (15) is greater than zero if $e^{m^2} > 3k$. Then we will proceed to the proof by consider two cases : (i) $k < 1/3$, (ii) $k \geq 1/3$. First, we consider the case (i). Since $k < 1/3$, which leads to $e^{m^2} > 3k$, (15) is greater than zero for all m. Thus the three degree polynomial in (14) has three distinct real roots. Assume that the three roots are y_1, y_2 and $y_3 \quad (y_1 < y_2 < y_3)$. Hence $\frac{\partial}{\partial x}d(x)$ is less than zero when $y < y_1$ or $y_2 < y < y_3$ and $\frac{\partial}{\partial x}d(x)$ is greater than zero when $y_1 < y < y_2$ or $y > y_3$. It means that $d(x)$ is decreasing when $y < y_1$ and $y_2 < y < y_3$ and $d(x)$ is increasing when $y_1 < y < y_2$ and $y > y_3$. Moreover, $d(x)$ can be rewritten as

$$(1-k)\int_{\theta_0-x}^{\theta_1-x} e^{-t^2/2}dt - k\int_{2\theta_0-\theta_1-x}^{\theta_0-x} e^{-t^2/2}dt - k\int_{\theta_1-x}^{2\theta_1-\theta_0-x} e^{-t^2/2}dt. \tag{16}$$

When x goes to infinity and minus infinity, (16) goes to zero. Thus, combining the above arguments, $d(x)$ has one local maximum and two local minimum. Note that $(\theta_0+\theta_1)/2$ is one root of (9). By Lemma 3, $d(x)$ has a local maximum at $x = (\theta_0+\theta_1)/2$. $d((\theta_0+\theta_1)/2)$ is equal to $F((\theta_1-\theta_0)/2) - F((\theta_0-\theta_1)/2) + kF(3(\theta_0-\theta_1)/2) - kF(3(\theta_1-\theta_0)/2)$, where $F(\cdot)$ is the cumulative function of standard normal distribution. By Lemma 4, when $k < 1/3$, $d((\theta_0+\theta_1)/2)$ is always greater than zero for all θ_0 and θ_1 satisfying $\theta_1 - \theta_0 > 0$. Thus, the figure of $d(x)$ to y is Figure 3.

By Figure 3, it leads to that the region $\{x : d(x) \leq 0\}$ is

$$\{x : y < \mu_1^* \quad \text{or} \quad y > \mu_2^*\}, \tag{17}$$

where μ_1^* and μ_2^* are two values such that $d(\mu_1^*) = d(\mu_2^*) = 0$. (see Figure 3). Let $c_1 = \ln \mu_1^*/m + \theta_0$ and $c_2 = \ln \mu_2^*/m + \theta_0$. Then by the definition of y, (17)

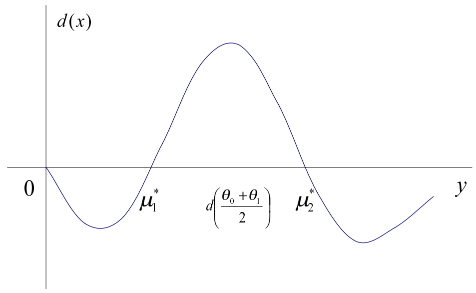

Figure 3: Plot of $d(x)$

is equivalent to

$$\{x : x < c_1 \quad \text{or} \quad x > c_2\}. \tag{18}$$

Note that k is chosen to be satisfied $E_{\theta_0}\phi(X) = \alpha$. Hence by the form of the rejection region (18) which is the same as the form of UMPU test and $E_{\theta_0}\phi(X) = \alpha$, the proof of case (i) will be completed if we prove $E_{\theta_1}\phi(X) = \alpha$. ¿From Lemma 2, for every x, there exists a ω such that $e^{-(x-\theta_0)^2/2} = e^{-(\omega-\theta_1)^2/2}$ and $\eta(x) = \eta(\omega)$. Then it will lead to $E_{\theta_1}\phi(X) = \alpha$.

Now we will consider case (ii) ($k \geq 1/3$). When $k \geq 1/3$, (15) is greater than zero only when $m = \theta_1 - \theta_0 > \sqrt{\ln 3k}$. And by Lemma 4, $d\left((\theta_0 + \theta_1)/2\right)$ is greater than zero only when $\theta_1 - \theta_0 > \nu$ and in Lemma 4, we know that $\nu > \sqrt{\ln 3k}$. Thus, when $\theta_1 - \theta_0 > \nu$, by a similar argument as in case (i), the proof is completed. □

The value of ν in Theorem 1 is defined in Lemma 4 and Table 1 is the value of ν corresponding to $k \geq 1/3$ by numerical calculation.

Table 1

k	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9
ν	0.39	0.78	1.04	1.26	1.46	1.66	1.86	2.07	2.30	2.56	2.88	3.29

3 The case of $\Theta_0 = \theta_0$

If the null hypothesis sample space contains only one point, the likelihood ratio test is a common tool for testing two-sided hypotheses. For the normal distribution, the likelihood ratio test and UMPU test are the same. Theorem 2 and Theorem 3 are the main results of this section. Without loss of generality, σ^2 is assumed to be 1 in Theorem 2 and Theorem 3.

Theorem 2. *Let $X_1, \cdots, X_n$ be iid according to $N(\theta, 1)$. For testing (1), assume that the null parameter space Θ_0 contains only one point θ_0, then the test*

$$\phi(x) = \begin{cases} 1 & \text{if} \quad \eta_1(x) \leq k \\ \\ 0 & \text{if} \quad \eta_1(x) > k \end{cases}$$

is a level α UMPU test, which is also a likelihood ratio test, where $\eta_1(x) = me^{-\sum_{i=1}^{n}(x_i-\theta_0)^2/2} / \left(e^{-\sum_{i=1}^{n}(x_i-t_1)^2/2} + me^{-\sum_{i=1}^{n}(x_i-\theta_0)^2/2} + e^{-\sum_{i=1}^{n}(x_i-t_2)^2/2}\right)$, $t_1 < \theta_0 < t_2$, $\theta_0 - t_1 = r$, $t_2 - t_1 = 2r$, $k = me^{cnr+nr^2/2} / \left(1 + e^{2cnr} + me^{cnr+nr^2/2}\right)$, *$m$ $(0 < m < \infty)$ and r $(0 < r < \infty)$ are some constants and c is the upper $\alpha/2$ quantile of $N(0, 1)$.*

Proof. The rejection region of $\phi(x)$ is equivalent to

$$\left\{x : m(1-k)e^{-\sum_{i=1}^{n}(x_i-\theta_0)^2/2} - ke^{-\sum_{i=1}^{n}(x_i-t_1)^2/2} - ke^{-\sum_{i=1}^{n}(x_i-t_2)^2/2} \leq 0\right\}. \tag{19}$$

By dividing $e^{-\sum_{i=1}^{n}(x_i-\theta_0)^2/2}$ in both side of (19) and let $y = e^{n(\bar{x}-\theta_0)r}$, (19) is equivalent to

$$\{y : h(y) \geq 0\}, \tag{20}$$

where

$$h(y) = ke^{-nr^2/2}\left[y^2 - m(1-k)e^{nr^2/2}y/k + 1\right].$$

Thus, $\frac{\partial}{\partial y}h(y) = ke^{-nr^2/2}\left[2y - m(1-k)e^{nr^2/2}/k\right]$ and let y_0 be the root of $\frac{\partial}{\partial y}h(y) = 0$. Note that $\frac{\partial^2 h(y)}{\partial^2 y} = 2ke^{-nr^2/2}$, which is always greater than zero.

Hence $h(y_0)$ has a local minimum at y_0 and $h(y_0) = 1 - e^{nr^2}(1-k)^2m^2/(4k^2)$. By the definition of y, we need only consider $h(y)$ with respect to positive part of y. By the facts that $h(0) > 0$, $h(y) > 0$ when y goes to infinity, $\frac{\partial}{\partial y}h(y) < 0$ when $0 < y < y_0$ and $\frac{\partial}{\partial y}h(y) > 0$ when $y > y_0$. Moreover, since $k = me^{cnr+nr^2/2}/\left(1 + e^{2cnr} + me^{cnr+nr^2/2}\right)$ satisfying $e^{nr^2}m^2 > 4k^2/(1-k)^2$, which leads to $h(y_0) < 0$. Therefore, the figure of $h(y)$ to y is Figure 4.

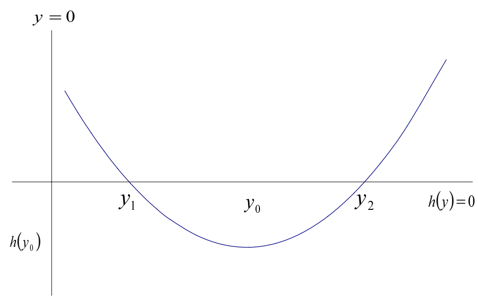

Figure 4: Plot of $h(y)$

The two roots of $h(y) = 0$ are

$$y_1 = \left(e^{nr^2/2}(1-k)m - \sqrt{-4k^2 + e^{nr^2}(1-k)^2m^2}\right)/2k.$$

and

$$y_2 = \left(e^{nr^2/2}(1-k)m + \sqrt{-4k^2 + e^{nr^2}(1-k)^2m^2}\right)/2k.$$

By (20), the rejection region of $\phi(x)$ is $\{y : y < y_1 \quad \text{and} \quad y > y_2\}$. By the definition of y, $(\bar{x} - \theta) = (\ln y)/nr$, which leads that the rejection region is $\{\bar{x} : (\bar{x} - \theta_0) < (\ln y_1)/nr \quad \text{and} \quad (\bar{x} - \theta_0) > (\ln y_2)/nr\}$. By straightforward calculation, $y_1y_2 = 1$, which means that $\phi(x)$ is a likelihood ratio test because of $\ln y_1/nr = -\ln y_2/nr$. By Casella and Berger (1990), the rejection region of a UMPU test for a normal distribution is also the same as this rejection region. Hence $\phi(x)$ is also a UMPU test. Moreover, according to α-level likelihood ratio test, $\ln y_2/n^2r = -\ln y_1/n^2r = c$, where c

is a upper quantile of $N(0,1)$. Then, by solving the above two equations, $k = e^{cnr+nr^2/2}m/\left(1+e^{2cnr}+e^{cnr+nr^2/2}m\right)$. □

Lemma 5. *The 4th degree polynomial equation* $y^4+e^{3nr^2/2}y^3-m(1-k)/ke^{2nr^2}y^2+e^{3nr^2/2}y+1=0$ *has 4 real roots at* r_1, r_2, r_3 *and* r_4 *if* nr^2 *is large, where*
$r_1 = -\frac{1}{4}e^{3nr^2/2} - A - \frac{1}{2}(B-C)^{\frac{1}{2}}, r_2 = -\frac{1}{4}e^{3nr^2/2} - A + \frac{1}{2}(B-C)^{\frac{1}{2}}$
$r_3 = -\frac{1}{4}e^{3nr^2/2} + A - \frac{1}{2}(B+C)^{\frac{1}{2}}, r_4 = -\frac{1}{4}e^{3nr^2/2} + A + \frac{1}{2}(B+C)^{\frac{1}{2}}$,

$$A = \frac{1}{2}\left(2+\frac{1}{4}e^{3nr^2} - e^{2nr^2}(-1+k)m/k\right)^{\frac{1}{2}},$$

$$B = -2+\frac{1}{2}e^{3nr^2} - e^{2nr^2}(-1+k)m/k$$

and

$$C = \frac{-8e^{3nr^2/2} - e^{9nr^2/2} + 4e^{7nr^2/2}(-1+k)m/k}{4\sqrt{2+\frac{1}{4}e^{3nr^2} - e^{2nr^2}(-1+k)m/k}}.$$

Proof. By straightforward calculation, $(y-r_1)(y-r_2)(y-r_3)(y-r_4)$ is equal to the lefthand side of the equation. The method of solving $4th$ degree of polynomial referres to Zwillinger, Krantz and Rosen (1996). □

Theorem 3. *Let* $X_1, \cdots, X_n$ *be iid according to* $N(\theta,1)$. *For testing (1), assumed that the null parameter space contains only one point* θ_0, *then the test*

$$\phi(X) = \begin{cases} 1 & if \quad \eta_2(X) \le k \\ 0 & if \quad \eta_2(X) > k \end{cases}$$

is a level α *UMPU test, which is also a likelihood ratio test, where* $\eta_2(x) = me^{-(x-\theta_0)^2/2}/\left[\left(\sum_{j=1}^{4} e^{-(x-t_j)^2/2}\right) + me^{-(x-\theta_0)^2/2}\right]$, $t_1 < t_2 < \theta_0 < t_3 < t_4$, $t_2 - t_1 = \theta_0 - t_2 = t_3 - \theta_0 = t_4 - t_3 = r$, $k = e^{2nr(c+4r)}m/(e^{6nr^2} + e^{2nr(2c+3r)} + e^{3nr(2c+5r)/2} + e^{cnr+15nr^2/2} + e^{2nr(c+4r)}m)$, *$m$ and r are some constants and c is the upper $\alpha/2$ quantile of* $N(0,1)$.

Proof. The rejection region of $\phi(X)$ is

$$\left\{X : me^{-\sum_{i=1}^{n}(x_i-\theta_0)^2/2} \Big/ \left(me^{-\sum_{i=1}^{n}(x_i-\theta_0)^2/2} + \sum_{j=1}^{4} e^{-\sum_{i=1}^{n}(x_i-t_j)^2/2}\right) \leq k\right\},$$

which is equivalent to

$$\{X : m(1-k)e^{-\sum_{i=1}^{n}(x_i-\theta_0)^2/2} - k\sum_{j=1}^{4} e^{-\sum_{i=1}^{n}(x_i-\theta_j)^2/2} \leq 0\}. \tag{21}$$

By dividing $e^{-\sum_{i=1}^{n}(x_i-\theta_0)^2/2}$ in both side, (21) is equal to

$$\{X : m(1-k) - k\left(\sum_{j=1}^{4} e^{n(\bar{x}-\theta_0)(t_j-\theta_0)-n(\theta_0-t_j)^2/2}\right) \leq 0\}. \tag{22}$$

Let $y = e^{n(\bar{x}-\theta_0)r}$. By the fact that $t_2 - t_1 = \theta_0 - t_2 = t_3 - \theta_0 = t_4 - t_3 = r$, (22) can be rewritten as

$$\left\{-m(1-k) + k\left(y^{-2}e^{-2nr^2} + y^{-1}e^{-nr^2/2} + ye^{-nr^2/2} + y^2e^{-2nr^2}\right) \geq 0\right\},$$

which can also be written as

$$\left\{X : ky^{-2}e^{-2nr^2}\left(y^4 + e^{3nr^2/2}y^3 - m(1-k)/ke^{2nr^2}y^2 + e^{3nr^2/2}y + 1\right) \geq 0\right\}. \tag{23}$$

Let

$$h(y) = y^4 + e^{3nr^2/2}y^3 - m(1-k)/ke^{2nr^2}y^2 + e^{3nr^2/2}y + 1. \tag{24}$$

$\frac{\partial}{\partial y}h(y) = 4y^3 + 3e^{3nr^2/2}y^2 - 2m(1-k)/ke^{2nr^2}y + e^{3nr^2/2}$. By Lemma 1, let $p = -3e^{3nr^2/2}/4$, $q = -2m(1-k)e^{2nr^2}/(4k)$ and $r = -e^{3nr^2/2}/4$, then we have the discriminant

$$\begin{aligned}\Delta\left(\frac{\partial}{\partial y}h(y)\right) &= e^{3nr^2}[-108 + 108e^{2nr^2}(-1+k)m/k + 9e^{4nr^2}(-1+k)^2m^2/k^2 + \\ &\quad e^{3nr^2}(-27 - (-1+k)^3m^3/k^3)]/64.\end{aligned}$$

When nr^2 is large enough, the leading term of $\Delta\left(\frac{\partial}{\partial y}h(y)\right)$ is $9e^{7nr^2}(1-k)^2m^2$, which is positive. By Lemma 1, it leads that $\frac{\partial}{\partial y}h(y)$ has three distinct roots, say y_0, y_1, y_2. Note that $h(y) > 0$ if y is zero and $h(y) > 0$ if y goes to infinity.

Moreover, by Lemma 5, the equation $h(y) = 0$ has four roots r_1, r_2, r_3 and r_4. By straightforward calculation, $r_1 \times r_2 = 1$ and $r_3 \times r_4 = 1$. Note that it is obvious that r_1 is negative and r_4 is positive if n is large enough because the terms involving $e^{3nr^2/2}$ are leading terms in these roots. Since $r_1 \times r_2$ is positive and $r_3 \times r_4$ is also positive, thus, r_2 is negative and r_3 is positive. Combining the above arguments, the figure of $h(y)$ with respect to y is figure 5.

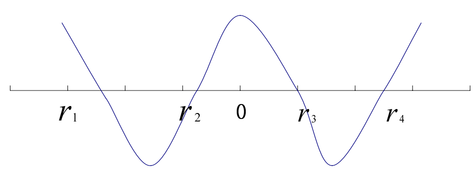

Figure 5: Plot of $h(y)$

By the defintion of y, y is always positive. Thus, the figure of $h(y)$ with respect to $y > 0$ is

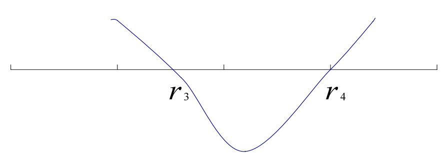

Figure 6: Plot of $h(y)$

By (23), the rejection region is $\{y : 0 < y < r_3 \quad \text{or} \quad y > r_4\}$, which is corresponding to

$$\{x : \bar{x} - \theta_0 < \ln r_3/nr \quad \text{or} \quad \bar{x} - \theta_0 > \ln r_4/nr\}\,. \tag{25}$$

Note that $r_3 \times r_4 = 1$. Thus (25) is equal to $\{x : \bar{x} - \theta_0 < -\ln r_4/nr \quad \text{or} \quad \bar{x} - \theta_0 >$ $\ln r_4/nr\}$, which is a UMPU test and also a likelihood ratio test. Moreover, $-\ln r_3/nr = \ln r_4/nr = c$, where c is a upper $\alpha/2$ quantile of $N(0,1)$. By solving the above two equations, we have $k = e^{2nr(c+4r)}m/(e^{6nr^2} + e^{2nr(2c+3r)} + e^{3nr(2c+5r)/2} + e^{cnr+15nr^2/2} + e^{2nr(c+4r)}m)$. □

The priors of the Bayes estimators in Theorem 2 and Theorem 3 are $I_{(\theta=t_1,t_2)}$ and $I_{(\theta=t_1,t_2,t_3,t_4)}$ respectively, where

$$I_{(\theta=t_1,\cdots,t_g)} = \begin{cases} 1 & \text{if} \quad \theta = t_1, t_2, \cdots t_g \\ 0 & \text{otherwise.} \end{cases}$$

These points $\{t_i, i = 1, \cdots, g\}$ are suggested to be some possible values of θ if we have information of θ. In Theorem 2 and Theorem 3, g is chosen to be 2 or 4. If g is chosen to be greater than 4, then it is difficult to have general results because by Abel's theorem, one of the principal results of Galois Theory, there are no general solutions for solving gth degree polynomial equations for $g \geq 5$ (Fraleigh (1982), Herstein (1975)). But, although it is impossible to establish general results for $g \geq 5$, we will investigate it from numerical study.

Proposition 1. *If $\eta_1(x)$ in Theorem 2 is replaced with $me^{-\sum_{i=1}^n (x_i-\theta_0)^2/2}/$ $(\sum_{j=1}^g e^{-\sum_{i=1}^n (x_i-t_j)^2/2} + me^{-\sum_{x=1}^n (x_i-\theta_0)^2/2})$, where g is an even positive integer, $t_i - \theta_0 = (i - (g/2+1))\,r$ for $i \leq g/2$ and $t_i - \theta_0 = (i - g/2)\,r$ for $i > g/2$. Then the rejection region $\{x : \eta_1(x) \leq k\}$ is equal to $\{y : -y^{g/2}m(1-k)/k + \sum_{j=1}^{g/2} \left(y^{g/2-j}e^{-nr^2j^2/2} + y^{g/2+j}e^{-nr^2j^2/2}\right) > 0\}$, where $y = e^{n(\bar{x}-\theta_0)r}$.*

The proof of Proposition 1 is by straightforward calculation. For $g = 6$, the rejection region is $\{y : y^6 + e^{5nr^2/2}y^5 + e^{4nr^2}y^4 - m(1-k)/ke^{9nr^2/2}y^3 + e^{4nr^2}y^2 + e^{5nr^2/2}y + 1 \geq 0\}$. From numerical calculation, the equation $y^6 + e^{5nr^2/2}y^5 + e^{4nr^2}y^4 - m(1-k)/ke^{9nr^2/2}y^3 + e^{4nr^2}y^2 + e^{5nr^2/2}y + 1 = 0$ only has two positive roots for most values of k, m, n, and r. For example, $k = 0.1$,

$m = 1$ and $nr^2 = 10$, the two positive roots of the equation are 0.000749 and 1335.17. Then by a similar argument as in Theorem 3, the tests derived from most Bayes estimators of the form $\eta_1(x)$ corresponding to $g = 6$ are UMPU tests. For $g = 8$, from numerical calculation and a similar argument as above, the tests derived from $\eta_1(x)$ corresponding to $g = 8$ for most values of k, m, n and r are UMPU tests. According to the numerical analysis of $g = 6$ and 8, it is very possible that the tests derived from $\eta_1(x)$ corresponding to $g > 8$ are UMPU tests.

4 Simulation results

In this section, the mean squared error of the proposed Bayes estimators and p-value are compared for the case of $\Theta_0 = \theta_0$. The mean squared error of an estimator $r(X)$ of $I(\theta \in \Theta_0)$ is $E(r(X) - I(\theta \in \Theta_0))^2$. Let MSE 1, MSE 2 and MSE 3 denote the mean squared errors of $\eta_1(x)$, $\eta_2(x)$ and p-value respectively.

Table 2: The table is mean squared errors of three estimators of $I(\theta = 0)$ when $m = 0.4$, $r = 1$ and $n = 5$. The simulation is based on 3000 replicates.

θ	MSE 1	MSE 2	MSE 3
-2	0.0002420	0.0002331	0.0001998
-1	0.0429601	0.0427160	0.0492836
0	0.3520266	0.3532986	0.3404814
1	0.0391925	0.0389499	0.0448714
2	0.0002121	0.0002037	0.00017611

Table 3: The table is mean squared errors of three estimators of $I(\theta = 0)$ when $m = 1.5$, $r = 1$ and $n = 5$. The simulation is based on 3000 replicates.

θ	MSE 1	MSE 2	MSE 3
-2	0.0009975	0.0009100	0.0000988
-1	0.1213554	0.120094	0.0464444
0	0.1445338	0.1462684	0.3320093
1	0.115851	0.114582	0.0443172
2	0.0009083	0.0008311	0.0001173

Table 4: The table is mean squared errors of three estimators of $I(\theta = 0)$ when $m = 3$, $r = 1$ and $n = 15$. The simulation is based on 3000 replicates.

θ	MSE 1	MSE 2	MSE 3
-2	3.30×10^{-10}	3.29×10^{-10}	1.10×10^{-12}
-1	0.04351728	0.04351725	0.00096786
0	0.01856288	0.01856291	0.329691
1	0.03704426	0.03704423	0.001205216
2	1.85×10^{-9}	1.85×10^{-9}	8.018×10^{-12}

Table 5: The table is mean squared errors of three estimators of $I(\theta = 0)$ when $m = 0.01$, $r = 1$ and $n = 15$. The simulation is based on 3000 replicates.

θ	MSE 1	MSE 2	MSE 3
-2	1.35×10^{-15}	1.34×10^{-15}	3.69×10^{-13}
-1	0.0008959	0.0008959	0.0010636
0	0.346827	0.346827	0.3302176
1	0.0007896	0.0007896	0.0010917
2	1.108×10^{-14}	1.106×10^{-14}	3.91×10^{-12}

From Table 2-5, we know that m and r can be selected to let $\eta_1(x)$ and $\eta_2(x)$ dominate p-value when $\theta = \theta_0$ or $\theta \neq \theta_0$. The selection of m and r can depend on the practical situations.

References

[1] Berger, J. O. and Wolpert, R. L. (1984). The Likelihood Principle. Institute of Mathematical Statistics, Hayward, CA.

[2] Casella, G. and Berger, R. L. (1990). Statistical Inference. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA.

[3] Fraleigh, J. B. (1967). A First Course in Abstract Algebra,Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont.

[4] Herstein, I. N. (1975). Topics in Algebra, Second edition. Xerox College Publishing, Lexington, Mass.-Toronto, Ont.

[5] Hwang, J. T. and Casella, G. and Robert, C. and Wells, M. T. and Farrell, R. H. (1992). Estimation of accuracy in testing. Ann. Statist. 20, 490-509.

[6] Jacobson, N. (1985). Basic algebra. I. Second edition. W. H. Freeman and Company, New York.

[7] Lehmann, E. L. (1997). Testing statistical hypotheses. Reprint of the 1986 second edition. Springer-Verlag, New York.

[8] Schaafsma, W., Tolboom, J. and Van der Menlen, B. (1989). Discussing truth or falsity by computing a q-value. In Statistical Data Analysis and Inference (Y. Dodge, ed.) 85-100. North-Holland, Amsterdam.

[9] Wang, H. (2004). Improved estimation of accuracy in simple hypothesis versus simple alternative testing. Journal of Multivariate Analysis, 90, 269-281.

[10] Woodroofe, M. and Wang, H. (2000). The problem of low counts in a signal plus noise model. Ann. Statist. 28, 1561-1569.

[11] Zwillinger, D., Krantz, S. G. and Rosen, K. H. (1996). CRC standard mathematical tables and formulae. 30th edition. CRC Press, Boca Raton, FL.

Received: March 15, 2006