J. Stat. Appl. Pro. Lett. **2**, No. 2, 123-129 (2015)

Journal of Statistics Applications & Probability Letters
An International Journal

http://dx.doi.org/10.12785/jsapl/020204

A Most Powerful Test for Prior Distribution based on a Primary Sample

*Hamzeh Torabi**

Department of Statistics, Yazd University, 89175-741, Yazd, Iran

Received: 18 Dec. 2014, Revised: 4 Mar. 2015, Accepted: 8 Mar. 2015
Published online: 1 May 2015

Abstract: Testing of hypotheses is one of the main purpose of statistical inference. But up to the present, a goodness-of-fit test for two competing prior distributions has not introduced. In this paper, some preliminary concepts regarding to hypotheses testing for prior distribution based on a primary sample are introduced and then by an appropriate approach a version of Neyman-Pearson lemma to find a most powerful goodness-of-fit test for prior distribution is given. Finally, some examples are presented to clarify the method.

Keywords: Prior density function, Hypothesis testing, Critical region, Probability of Type I and II errors, Most powerful test, Goodness-of-fit test.

1 Introduction

One of the primary purpose of statistical inference is to test parametric hypotheses based on a random sample $\boldsymbol{X} = (X_1, \cdots, X_n)$ from a parametric population with a probability density function (henceforth PDF) $f(x|\theta)$, where θ is a constant value of a set Θ, i.e., parameter space; and according to the random sample, one must choose one of two hypotheses $H_0 : \theta \in \Theta_0$ or $H_0 : \theta \in \Theta_1$, where Θ_j's are two disjoint subsets of Θ; See e.g. [3,4,6,8-11].

In the Bayesian approach is assumed that the random variable θ has a prior distribution $\pi(\theta)$ with the support Θ. There are a few well established approaches that deal with prior uncertainty; For more details, see e.g. [1,2,5,7,12]. However, listed below are some of them:

– **Empirical Bayes**. In this approach, one trusts the model but wants to estimate the unknown prior parameters that named as hyper-parameters and denoted by ν. More precisely, suppose $\boldsymbol{X}|\theta$ has density $f(\boldsymbol{x}|\theta)$ and $\theta|\nu$ has prior density $\pi(\theta|\nu)$ and distribution $\Pi(\theta|\nu)$. Then the predictive or marginal density of $\boldsymbol{X}|\nu$ is given by

$$m(\boldsymbol{x}|\nu) = \int_{\Theta} f(\boldsymbol{x}|\theta)\Pi(d\theta|\nu).$$

This can be taken as a likelihood for ν and so ML-II prior density for θ is estimated by $\pi(\theta|\hat{\nu})$ where $\hat{\nu}$ maximizes $m(\boldsymbol{x}|\nu)$.

– **Hierarchical Bayes**. Instead of estimating hyper-parameters, in the two stages hierarchical Bayes approach, we put a prior on hyper-parameters. Let $\pi(\theta|\nu)$ be a first-stage prior with a hyper-parameter ν with range Ξ and let $\lambda(\nu)$ and $\Lambda(\nu)$ be prior density and distribution of ν, respectively. Then the marginal prior density function for θ is obtained by

$$\pi(\theta) = \int_{\Xi} \pi(\theta|\nu)\Lambda(d\nu).$$

– **Robust Bayes**. Converse of the above, a whole class of plausible priors $\pi \in \Gamma$ is considered instead of a single prior. This leads to a class of inferences instead of a single inference. If the inferences differ drastically, then attempts to revise Γ into a smaller class are tried.

– **Γ-Minimax**. Instead of a whole class of inferences arising from consideration of the class Γ of priors, a suitable minimax procedure by confining attention to the priors in Γ is considered.

* Corresponding author e-mail: htorabi@yazd.ac.ir

This paper is not related to the above approaches. Specifically, it is applicable to model selection that are the following two approaches:

–**Bayesian**. Suppose there are two competing models:

$$\begin{cases} \text{Model 0: } \boldsymbol{X}|\theta \text{ has density } f_0(\boldsymbol{x}|\theta) \text{ and } \theta \text{ has prior density } \pi_0 \\ \text{Model 1: } \boldsymbol{X}|\theta \text{ has density } f_1(\boldsymbol{x}|\theta) \text{ and } \theta \text{ has prior density } \pi_1. \end{cases}$$

Of course, if f_0 and f_1 are the same, then π_0 and π_1 must differ. Then, the Bayesian approach is straight forward: The Bayes factor of Model 0 relative to Model 1 is given by

$$BF_{01}(\boldsymbol{x}) = \frac{m_0(\boldsymbol{x})}{m_1(\boldsymbol{x})} = \frac{\int_\Theta f_0(\boldsymbol{x}|\theta)\Pi_0(d\theta)}{\int_\Theta f_1(\boldsymbol{x}|\theta)\Pi_1(d\theta)}.$$

–**Frequentist**. Suppose that $\boldsymbol{X}|\theta$ has density $f(\boldsymbol{x}|\theta)$. Consider now two competing priors:

$$\begin{cases} H_0 : \theta \sim \pi_0 \\ H_1 : \theta \sim \pi_1, \end{cases} \tag{1}$$

where π_0 and π_2 are two different priors not necessarily in one family of distributions; We call the last testing problem as *prior hypotheses testing* (henceforth PHT).

Based on a primary sample, if one can introduce a testing method for PHT, i.e., the main attempt of this paper, then in the next step, *Bayesian statisticians may use the correct prior distribution to make the ordinary Bayesian statistical inference based on the main random sample.*

The paper is organized as follows:

In Section 2, we provide some definitions and preliminaries regarding to prior hypotheses testing. A Neyman-Pearson lemma for goodness-of-test for prior distribution is given in Section 3, and finally, some examples are presented in Section 4.

2 prior hypotheses testing

In this section, we give some concepts for prior hypotheses testing.

Definition 2.1. Assume that $\boldsymbol{X} = (X_1, \cdots, X_n)$ is a random sample from a parametric population with PDF $f(.|\theta)$ in which θ is a random variable and has a prior density function π_j under H_j, $j = 1, 2$. We define the weighted probability density function (henceforth WPDF) of $\boldsymbol{X}$ under H_j by

$$f_j(\boldsymbol{x}) = \left[\int_0^{\sup_{\theta\in\Theta}\pi_j(\theta)} \int_{\{\theta\in\Theta|\pi_j(\theta)>r\}} f(\boldsymbol{x}|\theta)d\theta dr\right] / \left[\int_0^{\sup_{\theta\in\Theta}\pi_j(\theta)} \int_{\{\theta\in\Theta|\pi_j(\theta)>r\}} d\theta dr\right],$$

if all integrals are finite; Substitute $\int_{\{\theta\in\Theta|\pi_j(\theta)>r\}}$ by $\sum_{\{\theta\in\Theta|\pi_j(\theta)>r\}}$ in the case of discrete prior distribution.

Remark 2.1. Not that $f_j(\boldsymbol{x})$ can be assumed as a joint PDF but not the marginal PDF of $\boldsymbol{X}$, since $f_j(\boldsymbol{x})$ is nonnegative and hence using the Fubini theorem, we have

$$\begin{aligned} \int_{\mathbb{R}^n} f_j(\boldsymbol{x})d\boldsymbol{x} &= \left[\int_{\mathbb{R}^n}\int_0^{\sup_{\theta\in\Theta}\pi_j(\theta)} \int_{\{\theta\in\Theta|\pi_j(\theta)>r\}} f(\boldsymbol{x}|\theta)d\theta dr d\boldsymbol{x}\right] / \left[\int_0^{\sup_{\theta\in\Theta}\pi_j(\theta)} \int_{\{\theta\in\Theta|\pi_j(\theta)>r\}} d\theta dr\right] \\ &= \left[\int_0^{\sup_{\theta\in\Theta}\pi_j(\theta)} \int_{\{\theta\in\Theta|\pi_j(\theta)>r\}} \left(\int_{\mathbb{R}^n} f(\boldsymbol{x}|\theta)d\boldsymbol{x}\right) d\theta dr\right] / \left[\int_0^{\sup_{\theta\in\Theta}\pi_j(\theta)} \int_{\{\theta\in\Theta|\pi_j(\theta)>r\}} d\theta dr\right] \\ &= \left[\int_0^{\sup_{\theta\in\Theta}\pi_j(\theta)} \int_{\{\theta\in\Theta|\pi_j(\theta)>r\}} 1 d\theta dr\right] / \left[\int_0^{\sup_{\theta\in\Theta}\pi_j(\theta)} \int_{\{\theta\in\Theta|\pi_j(\theta)>r\}} d\theta dr\right] \\ &= 1. \end{aligned}$$

Substitute $\int_{\mathbb{R}^n}$ by $\sum_{\mathbb{R}^n}$ in the discrete cases. Hence $f_j(\boldsymbol{x})$ is a joint PDF.

Remark 2.2. Let $T(\boldsymbol{X})$ be a sufficient statistic for θ. Using the factorization criterion, we have $f(\boldsymbol{x}|\theta) = g(t|\theta)h(\boldsymbol{x})$, where $t = T(\boldsymbol{x})$ and $g(t|\theta)$ can be considered as the PDF of $T(\boldsymbol{X})$. Therefore

$$f_j(\boldsymbol{x}) = h(\boldsymbol{x}) \left[\int_0^{\sup_{\theta\in\Theta} \pi_j(\theta)} \int_{\{\theta\in\Theta|\pi_j(\theta)>r\}} g(t|\theta) d\theta dr \right] / \left[\int_0^{\sup_{\theta\in\Theta} \pi_j(\theta)} \int_{\{\theta\in\Theta|\pi_j(\theta)>r\}} d\theta dr \right],$$

Let $g_j(t) = f_j(\boldsymbol{x})/h(\boldsymbol{x})$. $g_j(t)$ may be considered as the WPDF of $T(\boldsymbol{X})$ under H_j, $j = 0, 1$.

Remark 2.3. If H_j is a simple hypothesis, i.e., $H_j : \theta = \theta_j$, then the priors must be taken $\pi_j(\theta) = 1$ if $\theta = \theta_j$ and zero otherwise, i.e., a degenerate distribution. In this case $\sup_{\theta\in\Theta} \pi_j(\theta) = 1$ and $\{\theta \in \Theta|\pi_j(\theta) > r\} = \{\theta_j\}$ for any $0 < r \leq 1$ and thus $\sum_{\{\theta\in\Theta|\pi_j(\theta)>r\}} f(\boldsymbol{x}|\theta) = \sum_{\theta\in\{\theta_j\}} f(\boldsymbol{x}|\theta) = f(\boldsymbol{x}|\theta_j)$, then $f_j(\boldsymbol{x}) = f(\boldsymbol{x}|\theta_j)$, $j = 0, 1$, i.e, we confront an ordinary joint pdf of $\boldsymbol{X}$.

In PHT such as the classical hypotheses testing, we must give a test function $\Phi(\boldsymbol{X})$, based on the sample $\boldsymbol{X}$. In the following, we define the test function.

Definition 2.2. Let $\boldsymbol{X}$ be a random sample with the PDF $f(\boldsymbol{x}|\theta)$. $\Phi(\boldsymbol{X})$ is called a test function if it is the probability of rejecting H_0 providing to $\boldsymbol{X} = \boldsymbol{x}$ is observed.

Definition 2.3. Let $\Phi(\boldsymbol{X})$ be a test function. The probability of Type I and II errors related to $\Phi(\boldsymbol{X})$ for the prior testing problem (1) is defined by $\alpha_\Phi = E_0[\Phi(\boldsymbol{X})]$, and $\beta_\Phi = 1 - E_1[\Phi(\boldsymbol{X})]$, respectively, in which $E_i[\Phi(\boldsymbol{X})]$ is the expected value of $\Phi(\boldsymbol{X})$ over the WPDF $f_j(\boldsymbol{x})$, $j = 0, 1$.

Remark 2.4. Using Remark 2.2, we conclude that in the case of simple hypothesis against simple alternative, i.e.,

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta = \theta_1, \end{cases}$$

as in the Neyman-Pearson lemma, the above definition of α_Φ and β_Φ gives the classical probability of errors.

Remark 2.5. Regarding to definitions of error sizes, it is concluded that prior hypotheses testing (1) is really equivalent to the following ordinary hypotheses testing

$$\begin{cases} H_0' : \boldsymbol{X} \sim f_0 \\ H_1' : \boldsymbol{X} \sim f_1 \end{cases} \tag{2}$$

Definition 2.4. A prior hypotheses testing problem with a test function Φ is said to be a test of (significance) level α if $\alpha_\Phi \leq \alpha$, where $\alpha \in [0, 1]$. We call α_Φ as the size of Φ.

Definition 2.5. A prior test Φ of level α is said to be the most powerful test of level α if $\beta_\Phi \leq \beta_\Phi^*$, for all test Φ^* of level α.

3 Neyman-Pearson lemma for PHT

In this section, a version of Neyman-Pearson lemma for PHT is stated and proved.

Theorem 3.1. Let $\boldsymbol{X} = (X_1, ..., X_n)$ be a random sample with observed value $\boldsymbol{x} = (x_1, ..., x_n)$ and the PDF $f(\boldsymbol{x}|\theta)$. For testing

$$\begin{cases} H_0 : \ \theta \sim \pi_0 \\ H_1 : \ \theta \sim \pi_1, \end{cases} \tag{3}$$

a) any test with test function

$$\Phi(\boldsymbol{x}) = \begin{cases} 1, & \text{if } f_0(\boldsymbol{x})/f_1(\boldsymbol{x}) < k \\ \delta(\boldsymbol{x}), & \text{if } f_0(\boldsymbol{x})/f_1(\boldsymbol{x}) = k \\ 0, & \text{if } f_0(\boldsymbol{x})/f_1(\boldsymbol{x}) > k, \end{cases} \tag{4}$$

for some $k \geq 0$ and $0 \leq \delta(\boldsymbol{x}) \leq 1$, is the MP test of level α, where $\alpha = \alpha_\Phi$.
If $k = 0$, then the test

$$\Phi(\boldsymbol{x}) = \begin{cases} 1, \text{ if } f_0(\boldsymbol{x}) = 0 \\ 0, \text{ if } f_0(\boldsymbol{x}) > 0, \end{cases} \tag{5}$$

is the MP test of size zero.

b) for $0 \leq \alpha \leq 1$, there is a test of form (4) or (5) with $\delta(\boldsymbol{x}) = \delta$ (a constant), for which $\alpha_{\Phi} = \alpha$.

Proof. Regarding to the definitions of f_j, α and β, where were stated in Section 2, and also the equivalency of testing problem (3) and (2), all parts are proved from the classical Neyman-Pearson lemma; See e.g. Lehmann and Romano [9], pp. 60-61 or Shao [11] 394-395. □

Using Remark 2.2, the following corollary is resulted.

Corollary 3.1 Under the conditions of Theorem 3.1, the MP test function for testing (3) is

$$\Phi(t) = \begin{cases} 1, & \text{if } g_0(t)/g_1(t) < k \\ \delta(t), & \text{if } g_0(t)/g_1(t) = k \\ 0, & \text{if } g_0(t)/g_1(t) > k, \end{cases} \qquad (6)$$

for some $k \geq 0$, where $t = T(\boldsymbol{x})$ is the observation of a sufficient statistic $T(\boldsymbol{X})$ for θ.

4 Some Examples

In this section, we present two examples to clarify the theoretical discussions so far.

Example 4.1. Let $X_1, X_2, \cdots, X_n$ be a random sample from a normal population with mean μ and known variance σ^2, i.e., $N(\mu, \sigma^2)$. We have

$$f(x|\mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}, \quad \mu \in \mathbb{R}, \quad \sigma > 0.$$

In two cases $\mu_1 > \mu_0$ and $\mu_1 < \mu_0$, we want to find a MP test for testing problem

$$\begin{cases} H_0 : \mu \sim \pi_0 \\ H_1 : \mu \sim \pi_1, \end{cases}$$

based on the random sample $\boldsymbol{X}$, where

$$\pi_j(\mu) = \frac{1}{\tau\sqrt{2\pi}} \, \mathrm{e}^{-(\mu-\mu_j)^2/(2\tau^2)}, \quad j = 0, 1, \quad \mu \in \mathbb{R}, \ \tau > 0,$$

and μ_j, $j = 1, 2$ and τ are all known. Note that under H_j, $\theta \sim N(\mu_j, \tau^2)$, $j = 0, 1$.

Ii is remarkable that if $\tau \to 0$, then the above testing problem tends to the testing problem

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu = \mu_1. \end{cases}$$

We know that $T(\boldsymbol{X}) = \bar{X}$ is a sufficient statistic for μ and also $T \sim N(\mu, \sigma^2/n)$; i.e.,

$$g(t|\mu) = \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left\{ -\frac{(t-\mu)^2}{2\sigma^2/n} \right\}.$$

But it is easy to show that $\{\mu|\pi_j(\mu) > r\} = (\mu_j - \mu^*(r), \mu_j + \mu^*(r))$, where $\mu^*(r) = \tau\sqrt{-2\log\left(r\tau\sqrt{2\pi}\right)}$, and $\sup_{\mu\in\mathbb{R}} \pi_j(\mu) = 1/(\tau\sqrt{2\pi})$. Hence

$$\begin{aligned} g_j(t) &= \int_0^{1/(\tau\sqrt{2\pi})} \int_{\mu_j-\mu^*(r)}^{\mu_j+\mu^*(r)} g(t|\mu) d\mu dr \Big/ \int_0^{1/(\tau\sqrt{2\pi})} \int_{\mu_j-\mu^*(r)}^{\mu_j+\mu^*(r)} d\mu dr \\ &= \int_0^{1/(\tau\sqrt{2\pi})} \int_{\mu_j-\mu^*(r)}^{\mu_j+\mu^*(r)} g(t|\mu) d\mu dr \Big/ \int_0^{1/(\tau\sqrt{2\pi})} 2\mu^*(r) dr. \end{aligned}$$

Thus we must consider the following test

$$\begin{cases} H_0 : T \sim g_0 \\ H_1 : T \sim g_1. \end{cases}$$

But using corollary 3.1, the MP test function is like as (6). It is easy to show that $g_0(t)/g_1(t)$ is decreasing (increasing) in t if $\mu_1 > \mu_0$ ($\mu_1 < \mu_0$); i.e., the MP critical region in the cases $\mu_1 > \mu_0$ and $\mu_1 < \mu_0$ are the set of $\boldsymbol{X}$'s where $T(\boldsymbol{X}) > c$ and $T(\boldsymbol{X}) < c$, respectively, in which c is determined by size of test and the PDF $g_0(.)$. Hence, in the case $\mu_1 > \mu_0$ ($\mu_1 < \mu_0$), we have $c = G_0^{-1}(1-\alpha)$ ($c = G_0^{-1}(\alpha)$), where G_0^{-1} is the inverse function of the corresponding CDF of $g_0(.)$.

Let $\mu_0 = 0$, $\mu_1 = 1$ and $\sigma^2 = 16$. Figures 1 and 2 show the plots of g_0 and g_1 for the special case $\tau = 4$ and $n = 25$.

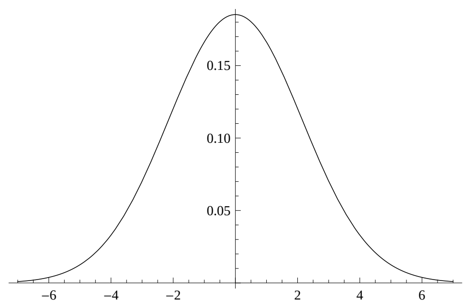

Fig. 1: The plot of g_0 for $n = 25$, $\mu_0 = 0$, $\mu_1 = 1$ and $\sigma^2 = 16$.

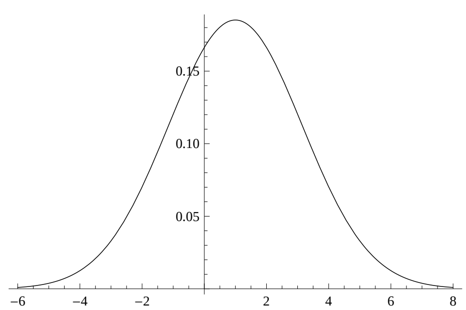

Fig. 2: The plot of g_0 for $n = 25$, $\mu_0 = 0$, $\mu_1 = 1$ and $\sigma^2 = 16$.

Note that $g_j(t)$ is a unimodal and symmetric PDF about μ_j, $j = 1,2$.
For the size $\alpha = 0.05$, Table 1 summarizes c and β (Type II error) for some various values of n and τ^2.

Table 1: The values of c and β for $n = 25, 50$ and 100.

$n\tau^2$		4	2	1	0.5	0.25	0.01	$\mu = 0$ versus $\mu = 1$
25	c	3.354	2.672	2.106	1.7564	1.552	1.327	1.316
	β	0.881	0.848	0.806	0.761	0.721	0.657	0.654
50	c	3.417	2.505	1.889	1.4898	1.242	0.945	0.931
	β	0.878	0.838	0.780	0.705	0.626	0.462	0.451
100	c	3.353	2.417	1.771	1.3361	1.053	0.678	0.658
	β	0.876	0.832	0.763	0.660	0.533	0.217	0.196

Table 1 illustrates that if $\tau \to 0$ then all results are completely coincided with the ordinary simple case, i.e., $H_0 : \mu = 0$ versus $H_1 : \mu = 1$, because in this case the MP test of size 0.05 rejects H_0 if $T > c$, where $c = 1.645(4/\sqrt{n})$

and $\beta = \Phi(\sqrt{n}(c-1)/4)$, where $\Phi(.)$ is the CDF of the standard normal distribution . From the table, it is also concluded that if $n \to \infty$, then $\beta \to 0$.

Example 4.2. Let $X_1, X_2, \cdots, X_n$ be a random sample from a Bernoulli distribution with parameter θ, i.e., $Ber(\theta)$, $0 < \theta < 1$.

It is interested to find a MP test for testing problem

$$\begin{cases} H_0: \ \theta \sim \pi_0 \\ H_1: \ \theta \sim \pi_1 \end{cases} \tag{7}$$

where

$$\pi_j(\theta) = \frac{1}{2\sigma_j} \sin\left((\theta - \theta_j)/\sigma_j\right), \quad 0 < \theta_j < \theta < \theta_j + \sigma_j \pi < 1, \ j = 0, 1.$$

It is obvious that the PDF π_j is unimodal, therefore $\pi_j(\theta) > r$ is equivalent to $\theta \in (L_j(r), U_j(r))$, where $L_j(r) = \theta_j + \sigma_j \arcsin(2r\sigma_j)$ and $U_j(r) = \theta_j + \sigma_j(\pi - \arcsin(2r\sigma_j))$. On the other hand, it can be shown that

$$\sup_{\theta \in (0,1)} \pi_j(\theta) = \frac{1}{2\sigma_j} =: m.$$

In addition, we know that $T(\boldsymbol{X}) = \sum_{i=1}^{n} X_i$ is a sufficient statistic for θ and also $T \sim B(n, \theta)$; i.e.,

$$g(t|\theta) = \binom{n}{t} \theta^t (1-\theta)^{n-t}, \quad t = 0, 1, ..., n.$$

Hence

$$\begin{aligned} g_j(t) &= \int_0^m \int_{L_j(r)}^{U_j(r)} g(t|\theta) d\theta dr \Big/ \int_0^m \int_{L_j(r)}^{U_j(r)} d\theta dr \\ &= \int_0^m \int_{L_j(r)}^{U_j(r)} g(t|\theta) d\theta dr \Big/ \int_0^m (U_j(r) - L_j(r)) dr. \end{aligned}$$

Thus, the PHT (7) is equivalent to

$$\begin{cases} H_0: T \sim g_0 \\ H_1: T \sim g_1. \end{cases}$$

Let $n = 5$, $\theta_0 = 0$, $\theta_1 = 0.5$ and $\sigma_0 = \sigma_1 = 1/(2\pi)$. It is easy to verify that $g_0(t)/g_1(t)$ is decreasing in t; Also, see Table 2.

Table 2: The PDFs g_0 and $g_1(t)$ and their ratio for $n = 5$.

t	0	1	2	3	4	5
$g_0(t)$	0.289	0.341	0.236	0.104	0.027	0.003
$g_1(t)$	0.003	0.027	0.104	0.236	0.341	0.289
$g_0(t)/g_1(t)$	96.333	12.630	2.269	0.441	0.079	0.010

Hence the structure of the MP critical region is as $T \geq c$. Hence the MP test at size $\alpha = 0.03$ rejects H_0 if $T \geq 4$.

Note π_j is unimodal and symmetric about $\theta_j + \pi\sigma_j/2$. Thus if the Bayesian statistician believes that under H_j, θ is near to $\theta_j^* \in (0,1)$, then he may choose a appropriate μ_j and a small enough σ_j, such that $\theta_j^* = \theta_j + \pi\sigma_j/2$, since the support of π_j is $(\mu_j, \mu_j + \pi\sigma_j)$. In this case, the PHT tends to ordinary simple versus simple testing problem $H_0 : \theta = \theta_0^*$ versus $H_1 : \theta = \theta_1^*$ and the ordinary MP test is also obtained.

5 Conclusion

In this paper, we introduced some concepts and definitions regarding to prior hypotheses testing problem. Then a Neyman-Pearson lemma for finding a most powerful goodness-of-fit test for prior hypotheses testing was introduced and finally, two examples were presented.

Acknowledgement

The author is grateful to the anonymous referees for a careful checking of the details and for helpful comments that improved this paper.

References

[1] W.M. Bolstad, *Introduction to Bayesian Statistics*, Second Edition, John Wiley, 2007.
[2] B.P. Carlin, T.A. Louis, *Bayesian Methods for Data Analysis*, Third Edition. Boca Raton, FL: Chapman and Hall/CRC, 2008.
[3] G. Casella, R.L. Berger, *Statistical Inference*, 2nd Edition, Duxbury Press, 2002.
[4] M.H. DeGroot, M.J. Schervish, *Probability and Statistics (third ed.)*. Addison-Wesley, 2002.
[5] A. Gelman, J.B Carlin, H.S. Stern, D.B. Rubin, *Bayesian Data Analysis*, Second Edition. Boca Raton, FL: Chapman and Hall/CRC, 2003.
[6] R.V. Hogg, A.T., Craig, *Introduction to Mathematical Statistics*, 5th Edition, Eaglewood Cliffs, NJ: Prentice-Hall, 1995.
[7] P.M. Lee, *Bayesian Statistics: An Introduction*. Second Edition, 1997.
[8] E.L. Lehmann, G. Casella, *Theory of Point Estimation*, Second Edition, Springer-Verlag, 1998.
[9] E.L. Lehmann, J.P Romano, *Testing Statistical Hypotheses*, Third Edition, Chapman-Hall.
[10] Mood, A.M., Graybill, F.A. and Boes, D.C. (1974). *Introduction to the Theory of Statistics*, Third Edition, McGRAW-HILL, 2005.
[11] J. Shao, *Mathematical Statistics*, Second Edition, Springer-Verlag, 2003.
[12] R.L. Winkler, *Introduction to Bayesian Inference and Decision*, Second Edition, Probabilistic Publishing, 2003.