Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsta20

Maximum Average-Power (MAP) Tests

Lin-An Chen [a] , Hui-Nien Hung [a] & Chih-Rung Chen [a]
[a] Institute of Statistics, National Chiao Tung University , Hsinchu, Taiwan
Published online: 28 Aug 2007.

Communications in Statistics—Theory and Methods, 36: 2237–2249, 2007

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610920701215480

Taylor & Francis
Taylor & Francis Group

Inference

Maximum Average-Power (MAP) Tests

LIN-AN CHEN, HUI-NIEN HUNG,
AND CHIH-RUNG CHEN

Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan

The objective of this article is to propose and study frequentist tests that have maximum average power, averaging with respect to some specified weight function. First, some relationships between these tests, called maximum average-power (MAP) tests, and most powerful or uniformly most powerful tests are presented. Second, the existence of a maximum average-power test for any hypothesis testing problem is shown. Third, an MAP test for any hypothesis testing problem with a simple null hypothesis is constructed, including some interesting classical examples. Fourth, an MAP test for a hypothesis testing problem with a composite null hypothesis is discussed. From any one-parameter exponential family, a commonly used UMPU test is shown to be also an MAP test with respect to a rich class of weight functions. Finally, some remarks are given to conclude the article.

Keywords Maximum average-power test; Most Powerful test; Uniformly most powerful test; Uniformly most powerful unbiased test.

Mathematics Subject Classification Primary 62F03; Secondary 62F04.

1. Introduction

In statistical problems, there exist many different level α tests, i.e., tests whose Type I errors are bounded by α. For any hypothesis testing problem with a simple alternative hypothesis, a level α most powerful (MP) test is recommended since by definition it maximizes the power among all the level α tests. Similarly, for any hypothesis testing problem with a composite alternative hypothesis, a level α uniformly most powerful (UMP) test is also recommended when it exists. However, there often does not exist any level α UMP test except in some very special cases. See, e.g., Lehmann (1986, Ch. 3). For such a situation, one searches for most powerful test among a more restricted class such as unbiased tests. Even with the unbiasedness restriction, there often does not exist any level α uniformly most powerful unbiased (UMPU) test. See, e.g., Neyman and Pearson (1936, 1938) and Lehmann (1986, Ch. 4).

Received September 29, 2006; Accepted November 24, 2006
Address correspondence to Lin-An Chen, Institute of Statistics, National Chiao Tung University, Hsinchu 30050, Taiwan; E-mail: lachen@stat.nctu.edu.tw

For a more detailed description, we classify the hypothesis testing problems whose level α MP, UMP, or UMPU tests are proposed in the literature into the following four categories:

A. $(A1)$ $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$ and $(A2)$ $H_0 : \theta \geq \theta_0$ vs. $H_1 : \theta < \theta_0$, where θ_0 is a known real-valued constant;
B. $H_0 : \theta \leq \theta_1$ or $\theta \geq \theta_2$ vs. $H_1 : \theta_1 < \theta < \theta_2$, where both θ_1 and θ_2 are known real-valued constants with $\theta_1 < \theta_2$;
C. $H_0 : \theta_1 \leq \theta \leq \theta_2$ vs. $H_1 : \theta < \theta_1$ or $\theta > \theta_2$, where both θ_1 and θ_2 are known real-valued constants with $\theta_1 \leq \theta_2$;
D. $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \{\theta_1\}$, where both Θ_0 and $\{\theta_1\}$ are known disjoint sets with $\Theta_0 \neq \emptyset$.

For any hypothesis testing problem, any level α MP or UMP test is also a level α UMPU test, but a level α UMPU test is not necessarily a level α MP or UMP test. Some well-known results of level α MP, UMP, or UMPU tests proposed in the literature are listed as follows:

(i) From any family of distributions with the monotone likelihood ratio property, there exists a level α UMP test for any hypothesis testing problem belonging to Category *A*. See, e.g., Lehmann (1986, Theorem 2, Ch. 3).
(ii) From any one-parameter exponential family, there exists a level α UMP test for any hypothesis testing problem belonging to Category *B*. See, e.g., Lehmann (1986, Theorem 6, Ch. 3).
(iii) From any one-parameter exponential family, there exists a level α UMPU test for any hypothesis testing problem belonging to Category *C*. See, e.g., Lehmann (1986, Sec. 4.2).
(iv) From any multiparameter exponential family, there exists a level α UMPU test for any hypothesis testing problem belonging to any of Categories *A*, *B*, or *C*, where θ is the one-dimensional parameter of interest. See, e.g., Lehmann (1986, Theorem 3, Ch. 4).
(v) There exists a level α MP test for any hypothesis testing problem belonging to Category *D*. See, e.g., Lehmann (1986, Theorem 3, Appendix).

However, all the hypothesis testing problems belonging to Categories *A*, *B*, *C*, or *D* are still limited for statistical applications. Some relevant remarks are given in the following.

1. From one-parameter exponential families, Categories *A*, *B*, and *C* only cover hypothesis testing problems with an interval rather than non-interval null parameter space. For example, the hypothesis testing problem with the null hypothesis $H_0 : \theta \in [\theta_1, \theta_2] \cup [\theta_3, \theta_4]$ vs. the alternative $H_1 : \theta \in (-\infty, \theta_1) \cup (\theta_2, \theta_3) \cup (\theta_4, \infty)$ does not fall into any of Categories *A*, *B*, or *C*, where all θ_1, θ_2, θ_3, and θ_4 are known real-valued constants with $\theta_1 \leq \theta_2 < \theta_3 \leq \theta_4$.

2. From any multiparameter exponential family, there is no level α UMPU test proposed in the literature for any hypothesis testing problem not belonging to any of Categories *A*, *B*, or *C*. For example, the hypothesis testing problem with null hypothesis $H_0 : (\theta_1, \theta_2)^T \in [\theta_{11}, \theta_{12}] \times [\theta_{21}, \theta_{22}]$ vs. the alternative $H_1 : (\theta_1, \theta_2)^T \notin [\theta_{11}, \theta_{12}] \times [\theta_{21}, \theta_{22}]$ does not belong to any of Categories *A*, *B*, or *C*, where $(\theta_1, \theta_2)^T$ is the two-dimensional parameter of interest and all θ_{11}, θ_{12}, θ_{21}, and θ_{22} are known real-valued constants with $\theta_{11} \leq \theta_{12}$ and $\theta_{21} \leq \theta_{22}$.

3. Although there exists a level α MP test for any hypothesis testing problem belonging to Category D, there often does not exist any UMP or UMPU test for a hypothesis testing problem with a composite alternative hypothesis.

In this article, we would like to answer the following question: Is it possible to define a "best" level α test among all level α tests such that it always exists and has higher powers for parameters in the alternative parameter space in some appropriate sense? In fact, we would like to maximize the average of powers for all parameters in the alternative parameter space with respect to (w.r.t.) a prespecified weight function. We motivate the idea by the following example.

Consider the situation where we would like to rank all the students in a class. In practice, the most commonly used method of ranking all the students in a class is as follows: First calculate the weighted average of scores of taken courses for each student in the class and then rank all the students by their weighted averages, where the weights are usually proportional to the credits of taken courses. At the moment, a student in the class is called best if he or she has the highest weighted average of scores of taken courses in the class. Following the spirit of the UMP criterion, a student in the class is called uniformly best if he or she has the highest score for each course offered in the class. A best rather than uniformly best student in the class is usually ranked the first for at least two important reasons. The first one is that when there exists a uniformly best student in the class, he or she is also a best student in the class. The second one is that there always exists at least one best student in the class, but there does not necessarily exist any uniformly best student in the class. These motivate us to see if there always exists a "best" level α test among all level α tests for any hypothesis testing problem in some appropriate sense. In this article, for any hypothesis testing problem with a composite alternative hypothesis, instead of finding a level α UMP or UMPU test, we would like to find a level α maximum average-power (MAP) test which maximizes the average of powers among all level α tests w.r.t. some weight function.

There are a few frequentist approach research articles which consider the idea of maximizing the average power. Martin and Silva (1994) and Martin et al. (1998) investigate the criterion for 2×2 contingency tables and use it to choose the best among several non-randomized tests. Moreover, Martin and Tapia (1999) solved the same problem under a multinomial model. These tests may be best among the tests they studied. However, they are not level α MAP when compared with all the level α tests including randomized tests (or only compared with all the level α non randomized tests). They focus on non randomized tests which are scientifically reasonable, but it creates a tremendously analytical difficulty for the reason stated below. Since the Type I error is not necessarily the prespecified nominal level α, it can only be numerically determined whether one can add more points to the rejection region (and make the test more powerful) without causing its maximum Type I error larger than the nominal level.

Another related article, Brown et al. (1995), which will be commented in the conclusion, focuses on confidence interval and confidence set construction. None of these articles provide the fundamental existence result of any level α MAP test for a continuous or discrete case (when compared with all the level α tests including randomized tests), which is the aim of the article.

The study of average power (and hence MAP) would become important in many modern statistical applications, such as those related to microarray experiments and Genomics, in which the number of unknown parameters is huge and could be 10,000 or more. To compare the power function of tests (possibly in multiple testing framework) which depends on huge number of parameters, the only practical way is to compare the average power. See, for example, Cui et al. (2005).

Now we discuss some advantages of the approach of searching for a level α MAP test.

(i) There exists a level α MAP test w.r.t. any weight function for any hypothesis testing problem. See Theorem 3.1. However, there often does not exist any level α UMP or UMPU test for a hypothesis testing problem. See, e.g., Lehmann (1986, Chs. 3, 4).
(ii) When there exists a level α MP or UMP test for some hypothesis testing problem, it is also a level α MAP test w.r.t. any weight function for the same hypothesis testing problem. See Theorem 2.2.
(iii) A level α MAP test w.r.t. any weight function for any hypothesis testing problem with a simple null hypothesis could be easily obtained by utilizing the Neyman–Pearson fundamental lemma. See Neyman and Pearson (1933) and Theorem 4.1.
(iv) Sometimes, a level α MAP test w.r.t. a weight function for a hypothesis testing problem with a finite number of elements in the null parameter space could be obtained by utilizing a generalization of the Neyman–Pearson fundamental lemma. See Lehmann (1986, Ch. 3) and Theorem 5.1.
(v) From any one-parameter exponential family, a commonly used level α UMPU test for any hypothesis testing problem belonging to Category C is also a level α MAP test w.r.t. a rich class of weight functions for the same hypothesis testing problem. See Example 5.2.

The article is organized as follows. In Sec. 2, MAP tests are proposed for testing general statistical hypotheses. Some relationships between MAP tests and MP or UMP tests are presented. In Sec. 3, the existence of an MAP test for any hypothesis testing problem is shown. In Sec. 4, an MAP test for any hypothesis testing problem with a simple null hypothesis is constructed, including some interesting classical examples. In Sec. 5, a MAP test for a hypothesis testing problem with a composite null hypothesis is discussed. From any one-parameter exponential family, a commonly used UMPU test is shown to be also an MAP test with respect to a rich class of weight functions. Finally, some concluding remarks are given in Sec. 6.

2. Maximum Average-Power Tests

In this article, let X be the observation which is distributed according to a probability density function (pdf) $f(\cdot;\theta)$ w.r.t. some σ-finite measure μ defined on $(\mathscr{X},\mathscr{F})$ for $\theta\in\Theta$, where $\mathscr{X}$ is the sample space, $\mathscr{F}$ is a σ-field on $\mathscr{X}$, θ is the unknown parameter, and Θ is the parameter space. Suppose that we are interested in testing the null hypothesis $H_0:\theta\in\Theta_0$ vs. the alternative $H_1:\theta\in\Theta_1$, where $\{\Theta_0,\Theta_1\}$ is a non trivial partition of Θ, i.e., Θ_0 and Θ_1 are disjoint non empty sets with $\Theta_0\cup\Theta_1=\Theta$. ϕ is called a test if ϕ is a measurable function w.r.t. $\mathscr{F}$ and $0\le\phi(x)\le 1$ for $x\in\mathscr{X}$. For a test ϕ, β_ϕ is called the power function of ϕ if

$$\beta_\phi(\theta)=\int_{\mathscr{X}}\phi(x)f(x;\theta)\mu(dx)$$

for $\theta \in \Theta$. Let α be a known constant with $0 < \alpha < 1$, e.g., 0.05, and let Φ_α denote the class of all level α tests.

Throughout this article, we assume that Condition 2.1 is met.

Condition 2.1. Assume that $f(x;\theta)$, $(x,\theta) \in \mathscr{X} \times \Theta_1$, is measurable w.r.t. $\mathscr{F} \times \mathscr{G}$ and that Λ is a probability measure defined on $(\Theta_1, \mathscr{G})$, where $\mathscr{X} \times \Theta_1$ denotes the product space of $\mathscr{X}$ and Θ_1, $\mathscr{G}$ a σ-field on Θ_1, and $\mathscr{F} \times \mathscr{G}$ the product σ-field of $\mathscr{F}$ and $\mathscr{G}$.

We make the following definitions.

Definition 2.1. Any probability measure Λ satisfying Condition 2.1 is called a weight function on Θ_1.

Definition 2.2. For a test ϕ, $\int_{\Theta_1} \beta_\phi(\theta)\Lambda(d\theta)$ $(\equiv \bar{\beta}_{\phi,\Lambda})$ is called the average power of ϕ w.r.t. the weight function Λ.

Definition 2.3. A test ϕ_α^* is called a level α *MAP* test w.r.t. the weight function Λ if $\phi_\alpha^* \in \Phi_\alpha$ and $\bar{\beta}_{\phi,\Lambda} \le \bar{\beta}_{\phi_\alpha^*,\Lambda}$ for all $\phi \in \Phi_\alpha$.

For the simple alternative hypothesis $\Theta_1 = \{\theta_1\}$, the only probability measure defined on Θ_1 is the trivial one with $\Lambda(\{\theta_1\}) = 1$. In such a situation, a level α MAP test w.r.t. the unique weight function Λ is equivalent to a level α MP test for the same hypothesis testing problem. As a result, our interest is mainly focused on any hypothesis testing problem with a composite alternative hypothesis.

A relationship between MAP tests and MP or UMP tests for any hypothesis testing problem is given in the following theorem.

Theorem 2.2. *Consider any hypothesis testing problem with the null hypothesis* $H_0 : \theta \in \Theta_0$ *vs. the alternative* $H_1 : \theta \in \Theta_1$. *The following are equivalent*:

(i) ϕ_α^* *is a level* α *MP or UMP test*;
(ii) ϕ_α^* *is a level* α *MAP test w.r.t. any weight function* Λ *on* Θ_1.

Proof. It is obvious that (i) implies (ii). Now, suppose that (ii) is true but that (i) fails. Then there exist a level α test ϕ and a parameter θ_1 in the alternative parameter space such that $\beta_\phi(\theta_1) > \beta_{\phi_\alpha^*}(\theta_1)$. Let $\mathscr{G}$ and Λ satisfy Condition 2.1 such that $\Lambda(\{\theta_1\}) = 1$, e.g., $\mathscr{G}$ is the finest σ-field on the alternative parameter space. Since

$$\bar{\beta}_{\phi,\Lambda} = \beta_\phi(\theta_1) > \beta_{\phi_\alpha^*}(\theta_1) = \bar{\beta}_{\phi_\alpha^*,\Lambda},$$

ϕ_α^* is not a level α MAP test w.r.t. Λ, contradicting (ii) and completing the proof. □

From Theorem 2.2, any level α UMP test proposed in the literature for any hypothesis testing problem is also a level α MAP test w.r.t. any weight function for the same hypothesis testing problem. For example, from any family of distributions with the monotone likelihood ratio property, there exists a level α UMP test for any hypothesis testing problem belonging to Category *A*, which implies that it is also a level α MAP test w.r.t. any weight function for the same hypothesis testing problem.

A relationship between MAP tests and MP tests for any hypothesis testing problem is given in the following theorem.

Theorem 2.3. *Consider any hypothesis testing problem with the null hypothesis* $H_0 : \theta \in \Theta_0$ *vs. the alternative* $H_1 : \theta \in \Theta_1$. *For any weight function* Λ *on* Θ_1, *the following are equivalent*:

(i) ϕ^*_α *is a level* α *MAP test w.r.t.* Λ;
(ii) ϕ^*_α *is a level* α *MP test for the hypothesis testing problem with the same null hypothesis* $H_0 : \theta \in \Theta_0$ *vs. the different simple alternative* $H_1' : X$ *has pdf*

$$f_\Lambda(x) = \int_{\Theta_1} f(x;\theta)\Lambda(d\theta) \tag{1}$$

w.r.t. μ.

Proof. The theorem can be proved easily by recognizing that the average power $\bar{\beta}_{\phi,\Lambda}$ of a test ϕ is the power w.r.t. $f_\Lambda(x)$, i.e.,

$$\begin{aligned}\bar{\beta}_{\phi,\Lambda} &= \int_{\Theta_1}\left[\int_{\mathscr{X}} \phi(x)f(x;\theta)\mu(dx)\right]\Lambda(d\theta)\\ &= \int_{\mathscr{X}}, \phi(x)\left[\int_{\Theta_1} f(x;\theta)\Lambda(d\theta)\right]\mu(dx).\\ &= \int_{\mathscr{X}} \phi(x)f_\Lambda(x)dx,\end{aligned} \tag{2}$$

by Fubini's theorem. □

From Theorem 2.3, any hypothesis testing problem with a composite alternative hypothesis could change to the hypothesis testing problem with the same null hypothesis vs. a different simple alternative hypothesis. Thus, the method proposed in the literature in particular Neyman–Pearson fundamental lemma to obtain a level α MP test might be utilized to obtain a level α MAP test.

3. Existence of MAP Tests

In this section, a main result of the article concerning the existence of a level α MAP test w.r.t. any weight function for any hypothesis testing problem is given as follows.

Theorem 3.1. *Consider any hypothesis testing problem with the null hypothesis* $H_0 : \theta \in \Theta_0$ *vs. the alternative* $H_1 : \theta \in \Theta_1$. *There exists a level* α *MAP test w.r.t. any weight function* Λ *on* Θ_1.

Proof. From Theorem 2.3, a level α MAP test w.r.t. Λ is equivalent to a level α MP test for a different hypothesis testing problem with the same null hypothesis $H_0 : \theta \in \Theta_0$ vs. another simple alternative $H_1' : X$ has pdf $f_\Lambda(x)$ defined in (1), $x \in \mathscr{X}$, w.r.t. μ.

Let $\{\phi_n\}_{n=1}^{\infty}$ be a sequence of level α tests such that $\lim_{n\to\infty} \bar{\beta}_{\phi_n,\Lambda} = \sup_{\phi\in\Phi_\alpha} \bar{\beta}_{\phi,\Lambda}$. Then, by Eq. (2),

$$\lim_{n\to\infty} \int_{\mathscr{X}} \phi_n(x) f_\Lambda(x)\mu(dx) = \sup_{\phi\in\Phi_\alpha} \int_{\mathscr{X}} \phi(x) f_\Lambda(x)\mu(dx).$$

By the weak compactness theorem, e.g., see Lehmann (1986, Theorem 3, Appendix), there exist a subsequence $\{n_i\}_{i=1}^{\infty}$ of $\{1, 2, 3, \ldots\}$ and a test ϕ_α^* such that ϕ_{n_i} converges weakly to ϕ_α^* as $i \to \infty$. Since ϕ_{n_i} converges weakly to ϕ_α^* as $i \to \infty$ and $f_\Lambda(x)$, $x \in \mathscr{X}$, is an integrable function w.r.t. μ, it follows that

$$\lim_{i\to\infty} \int_{\mathscr{X}} \phi_{n_i}(x) f_\Lambda(x)\mu(dx) = \int_{\mathscr{X}} \phi_\alpha^*(x) f_\Lambda(x)\mu(dx),$$

which implies that $\lim_{i\to\infty} \bar{\beta}_{\phi_{n_i},\Lambda} = \bar{\beta}_{\phi_\alpha^*,\Lambda}$ by Eq. (2). Since the level α test ϕ_{n_i} converges weakly to ϕ_α^* as $i \to \infty$ and $f(\cdot;\theta)$ is an integrable function w.r.t. μ for $\theta \in \Theta_0$,

$$\beta_{\phi_\alpha^*}(\theta) = \int_{\mathscr{X}} \phi_\alpha^*(x) f(x;\theta)\mu(dx) = \lim_{i\to\infty} \int_{\mathscr{X}} \phi_{n_i}(x) f(x;\theta)\mu(dx) = \lim_{i\to\infty} \beta_{\phi_{n_i}}(\theta) \le \alpha$$

for $\theta \in \Theta_0$, which implies that ϕ_α^* is also a level α test. As

$$\bar{\beta}_{\phi_\alpha^*,\Lambda} = \lim_{i\to\infty} \bar{\beta}_{\phi_{n_i},\Lambda} = \sup_{\phi\in\Phi_\alpha} \bar{\beta}_{\phi,\Lambda},$$

ϕ_α^* is a level α MAP test w.r.t. Λ, completing the proof. □

In this article, we assume that Λ is a probability. When Λ is an infinite measure, then typically the average power is infinite, since even the smallest power function is bounded below by a positive number such as α. Hence, any test is MAP and it is pointless to study such a criterion.

4. MAP Tests for a Simple Null Hypothesis

In this section, consider any hypothesis testing problem with a simple null hypothesis. The Neyman–Pearson fundamental lemma (Neyman and Pearson, 1933) provides a level α MP test for any hypothesis testing problem with a simple null hypothesis vs. a simple alternative hypothesis. See, e.g., Lehmann (1986, Theorem 1, Ch. 3). Utilizing the Neyman–Pearson fundamental lemma, a level α MAP test w.r.t. a weight function for any hypothesis testing problem with a simple null hypothesis is constructed as follows.

Theorem 4.1. *Consider any hypothesis testing problem with the simple null hypothesis $H_0 : \theta = \theta_0$ vs. the alternative $H_1 : \theta \in \Theta_1$. For any weight function Λ on Θ_1, there exists a level α MAP test ϕ_α^* w.r.t. Λ such that $\beta_{\phi_\alpha^*}(\theta_0) = \alpha$ and*

$$\phi_\alpha^*(x) = \begin{cases} 1 & \text{for } f_\Lambda(x) > C_\alpha f(x;\theta_0), \\ 0 & \text{for } f_\Lambda(x) < C_\alpha f(x;\theta_0), \end{cases}$$

where C_α is a non negative constant and $f_\Lambda(x)$ is as defined in (1).

Theorem 4.1 can be applied to that a simple null hypothesis based on observations distributed according to a one-parameter exponential family. We provide three examples of two-sided tests. The first is to test the normal mean, the second the normal variance, and the last the binomial probability. In all three cases, by Theorem 4.1, the rejection region of a MAP can be written as

$$RT \equiv f_\Lambda(x)/f(x, \theta_0) > C_\alpha.$$

Hence below the effort is devoted to the calculation of the ratio RT.

Example 4.1 (Testing the Normal Mean). Let $X \equiv (X_1, \ldots, X_n)^T$, where X_i's are the independent observations having $N(\theta, \sigma_0^2)$ distribution with σ_0 is known. Suppose that our interest is to test $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$. Then there does not exist any level α UMP test, but there does exist a level α UMPU test. See, e.g., Lehmann (1986, p. 136).

Consider the a $N(m, v^2)$ weight function Λ. Let $T \equiv \sqrt{n}(\overline{X} - \theta_0)/\sigma_0$ and let $f_T(\cdot; \theta)$ denote the pdf of T. Under H_0, T has the standard normal distribution. The function $f_\Lambda(t) \equiv \int f_T(t; \theta)\Lambda(d\theta)$, equals the pdf of $N(\sqrt{n}(m - \theta_0)/\sigma_0, 1 + nv^2/\sigma_0^2)$. Here we use the simple fact that if $\overline{X}$ given θ has the $N(\theta, \sigma_0^2/n)$ distribution and θ has the $N(m, v^2)$ distribution, then $\overline{X}$ has the unconditional $N(m, v^2 + \sigma_0^2/n)$ distribution. Consequently, it is easy to see that the ratio $RT \equiv f_\Lambda(t)/f_T(t; \theta_0)$ is a strictly increasing function of $|t - \sigma_0(\theta_0 - m)/(\sqrt{n}v^2)|$. Let ϕ_α^* be the test which rejects H_0 if and only if $\sqrt{n}|\overline{X} - \theta_0 - \sigma_0^2(\theta_0 - m)/(nv^2)|/\sigma_0 \geq C_\alpha$, where

$$\Phi\left(\frac{\sigma_0(\theta_0 - m)}{\sqrt{n}v^2} + C_\alpha\right) - \Phi\left(\frac{\sigma_0(\theta_0 - m)}{\sqrt{n}v^2} - C_\alpha\right) = 1 - \alpha.$$

Consequently, it follows from Theorem 4.1 that ϕ_α^* is a level α MAP test w.r.t. Λ.

Example 4.2 (Testing the Binomial Probability). Let X be the observation having the binomial $(n; \theta)$ distribution. Suppose that our interest is to test $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$. Then there does not exist any level α UMP test, but there does exist a level α UMPU test. See, e.g., Lehmann (1986, p. 138).

Consider the beta(a, b) weight function Λ with pdf

$$\lambda(\theta) \propto \theta^{a-1}(1 - \theta)^{b-1}.$$

Then, for $x \in \{0, 1, \ldots, n\}$,

$$f_\Lambda(t) \equiv \int_0^1 f(x; \theta)\Lambda(d\theta) \propto \frac{\Gamma(a + x)\Gamma(b + n - x)}{x!(n - x)!}.$$

Thus, $RT \equiv f_\Lambda(t)/f(x; \theta_0)$ is a strictly increasing function of $\Gamma(a + x)\Gamma(b + n - x)(1 - \theta_0)^x/\theta_0^x$ which is defined as $g(x)$. Let $k \equiv -a + (a + b + n - 1)\theta_0$. Note that, for $x < n$, $g(x + 1)/g(x)$ is greater than one if and only if $x > k$. Let ϕ_α^* be a size α test such that

$$\phi_\alpha^*(x) = \begin{cases} 1 & \text{for } g(x) > g(C_\alpha), \\ 0 & \text{for } g(x) < g(C_\alpha). \end{cases}$$

The numbers C_α and $\phi^*_\alpha(x)$ when $g(x) = g(C_\alpha)$ are chosen so that $\phi^*_\alpha(x)$ has Type I error α. Consequently, it follows from Theorem 4.1 that ϕ^*_α is a level α MAP test w.r.t. Λ.

Theorem 4.1 can be applied to multiparameter cases as well. Here we focus on the simple null hypotheses. From the location-scale family of univariate normal distributions, a level α MAP test w.r.t. any normal-inverse-gamma weight function for any hypothesis testing problem with a simple null hypothesis is given as follows.

Example 4.3 (Testing the Normal Mean and Variance). Let $X \equiv (X_1, \ldots, X_n)^T$, where X's are the independent observations having $N(\mu, \sigma^2)$ distribution. Suppose that our interest is to test $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$, where θ_0 $(\equiv (\mu_0, \sigma_0^2)^T)$. Then there does not exist any level α UMP test and there is no level α UMPU test proposed in the literature.

Consider the normal-inverse-gamma (a, d, m, v^2) weight function Λ, with pdf

$$\lambda(\theta) \propto \frac{1}{\sigma^{d+3}} \exp\left[-\frac{(\mu - m)^2 + av^2}{2v^2\sigma^2}\right].$$

Let $T \equiv (\overline{X}, \sum_{i=1}^n (X_i - \overline{X})^2)^T$ $(\equiv (T_1, T_2)^T)$ and let $f_T(\cdot; \theta)$ denote the pdf of T. Then, for $t = (t_1, t_2)^T$,

$$f_T(t; \theta) \propto \exp\left[-\frac{n(t_1 - \mu)^2}{2\sigma^2}\right] t_2^{(n-3)/2} \exp\left(-\frac{t_2}{2\sigma^2}\right)$$

and

$$f_\Lambda(t) \equiv \int f_T(t; \theta)\Lambda(d\theta) \propto \frac{t_2^{(n-3)/2}}{[h(t)]^{(d+n)/2}},$$

where $h(t) \equiv a + n(t_1 - m)^2/(1 + nv^2) + t_2$. Thus, $RT \equiv f_\Lambda(t)/f_T(t; \theta_0)$ is a strictly increasing function of $g(t_1, t_2) \equiv \exp\{[n(t_1 - \mu_0)^2 + t_2]/(2\sigma_0^2)\}/[h(t)]^{(n+d)/2}$ which is defined as $g(t)$. Let ϕ^*_α be the size α test which rejects H_0 if and only if $g(t_1, t_2) \geq C_\alpha$, where C_α is chosen so that the test have Type I error α. Consequently, it follows from Theorem 4.1 that ϕ^*_α is a level α MAP test w.r.t. Λ.

5. MAP Tests for a Composite Null Hypothesis

In this section, we focus on constructing the MAP test for a composite null hypothesis, which is a more difficult problem. However, in some situations, testing the composite null hypothesis can be reduced to a null hypothesis with finitely many points, called the finite null hypothesis in this article. We can then apply the generalized Neyman–Pearson fundamental Lemma. See, for example, Lehmann (1986, Theorem 5, Ch. 3).

For the finite null hypothesis, we have the following theorem.

Theorem 5.1. *Consider a hypothesis testing problem with the finite-element composite null hypothesis* $H_0 : \theta \in \{\theta_{01}, \theta_{02}, \ldots, \theta_{0k}\}$ *vs. the alternative* $H_1 : \theta \in \Theta_1$. *For a weight* *function Λ on Θ_1, if there exist non negative constants $C_{\alpha,1}, C_{\alpha,2}, \ldots, C_{\alpha,k}$ and a test ϕ_α^* such that $\beta_{\phi_\alpha^*}(\theta_i) = \alpha$ for $i = 1, 2, \ldots, k$ with*

$$\phi_\alpha^*(x) = \begin{cases} 1 \quad \textit{for} \ \int_{\Theta_1} f(x;\theta)\Lambda(d\theta) > \sum_{i=1}^{k} C_{\alpha,i} f(x;\theta_i), \\ 0 \quad \textit{for} \ \int_{\Theta_1} f(x;\theta)\Lambda(d\theta) < \sum_{i=1}^{k} C_{\alpha,i} f(x;\theta_i), \end{cases}$$

then ϕ_α^* is a level α MAP test w.r.t. Λ.

Utilizing Lehmann (1986, Theorem 7, Ch. 3), we have the following theorem.

Theorem 5.2. *Consider a hypothesis testing problem with the composite null hypothesis $H_0 : \theta \in \Theta_0$ vs. the alternative $H_1 : \theta \in \Theta_1$. For weight functions Λ_0 on Θ_0 and Λ_1 on Θ_1, if there exists a level α test ϕ_α^* such that it is also a level α MAP test w.r.t. Λ for testing the simple null hypothesis H_0' : X has pdf $\int_{\Theta_0} f(x;\theta)\Lambda_0(d\theta)$, $x \in \mathscr{X}$, w.r.t. μ vs. the alternative $H_1 : \theta \in \Theta_1$, then ϕ_α^* is a level α MAP test w.r.t. Λ.*

There is a simple application of Theorem 5.2 to test the normal mean. A commonly used level α UMPU test for a hypothesis testing problem with an interval composite null hypothesis versus a two-sided alternative hypothesis is shown to be also a level α MAP test w.r.t. a rich class of weight functions as follows.

Example 5.1 (Testing the Normal Mean in a Composite Null Hypothesis). Let $X \equiv (X_1, \ldots, X_n)^T$, where X_i's are the independent observations having an $N(\theta, \sigma_0^2)$ distribution with σ_0 is known. Suppose that our interest is to test $H_0 : \theta_1 \le \theta \le \theta_2$ vs. $H_1 : \theta < \theta_1$ or $\theta > \theta_2$. Then there does not exist any level α UMP test, but there does exist a level α UMPU test. See, e.g., Lehmann (1986, p. 135).

Consider a weight function Λ on $(-\infty, \theta_1) \cup (\theta_2, \infty)$ $(\equiv \Theta_1)$ with $\Lambda((-\infty, a]) = \Lambda([\theta_1 + \theta_2 - a, \infty))$ for all $a \in (-\infty, \theta_1)$, i.e., symmetric w.r.t. $(\theta_1 + \theta_2)/2$ $(\equiv \theta_0)$. Moreover, consider another weight function Λ_0 on $[\theta_1, \theta_2]$ $(\equiv \Theta_0)$ with $\Lambda_0(\{\theta_1\}) = \Lambda_0(\{\theta_2\}) = 1/2$. Let $T \equiv \sqrt{n}(\overline{X} - \theta_0)/\sigma_0$ and let $f_T(\cdot;\theta)$ denote the pdf of T. Let $\omega(\theta) \equiv \sqrt{n}(\theta - \theta_0)/\sigma_0$. Then,

$$f_T(t;\theta) \propto \exp\left\{-\frac{[t - \omega(\theta)]^2}{2}\right\}.$$

Consequently,

$$\begin{aligned} &\int_{\Theta_1} f_T(t;\theta)\Lambda(d\theta) \\ &= \int_{(-\infty,\theta_1)} [f_T(t;\theta) + f_T(t; 2\theta_0 - \theta)]\Lambda(d\theta) \\ &\propto \exp\left(-\frac{t^2}{2}\right) \int_{(-\infty,\theta_1)} \{\exp[t\omega(\theta)] + \exp[-t\omega(\theta)]\} \exp\left\{-\frac{[\omega(\theta)]^2}{2}\right\}\Lambda(d\theta). \end{aligned}$$

Thus, for $-\infty < t < \infty$,

$$\frac{\int_{\Theta_1} f_T(t;\theta)\Lambda(d\theta)}{\int_{\Theta_0} f_T(t;\theta)\Lambda_0(d\theta)} \propto \int_{(-\infty,\theta_1)} \frac{\exp[t\omega(\theta)]+\exp[-t\omega(\theta)]}{\exp[t\omega(\theta_1)]+\exp[-t\omega(\theta_1)]}\exp\left\{-\frac{[\omega(\theta)]^2}{2}\right\}\Lambda(d\theta).$$

It is easy to show that $\{\exp[t\omega(\theta)]+\exp[-t\omega(\theta)]\}/\{\exp[t\omega(\theta_1)]+\exp[-t\omega(\theta_1)]\}$ is a strictly increasing function of $|t|$ for $\theta < \theta_1$, which implies that $\int_{\Theta_1} f_T(t;\theta)\Lambda(d\theta)/\int_{\Theta_0} f_T(t;\theta)\Lambda_0(d\theta)$ is also a strictly increasing function of $|t|$. Since $\beta_{\phi_\alpha^*}(\theta_1) = \beta_{\phi_\alpha^*}(\theta_2) = \alpha$, ϕ_α^* is a level α MAP test w.r.t. Λ for testing the simple null hypothesis H_0' : T has pdf $\int_{\Theta_0} f_T(t;\theta)\Lambda_0(d\theta)$ w.r.t. μ vs. the alternative $H_1 : \theta \in \Theta_1$. Consequently, it follows from Theorem 5.2 that ϕ_α^* is a level α MAP test w.r.t. any prior Λ which is symmetric w.r.t. the mid point of θ_1 and θ_2. Note that this MAP result does not restrict to the unbiased tests. Hence this is a new result beyond UMPU tests.

A more interesting question is to see if, from any one-parameter exponential family, a commonly used level α UMPU test for a hypothesis testing problem with an interval null hypothesis vs. a two-sided alternative hypothesis is also a level α MAP test w.r.t. a rich class of weight functions, which is to be presented in the next example.

Example 5.2 (The One-Parameter Exponential Family)**.** Let X be the observation with pdf

$$f(x;\theta) = h(x)\exp[\theta T(x) - c(\theta)].$$

Suppose that our interest is to test $H_0 : \theta_1 \le \theta \le \theta_2$ vs. $H_1 : \theta < \theta_1$ or $\theta > \theta_2$. Then there does not exist any level α UMP test, but there does exist a level α UMPU test ϕ_α^* such that $\beta_{\phi_\alpha^*}(\theta_1) = \beta_{\phi_\alpha^*}(\theta_2) = \alpha$ with

$$\phi_\alpha^*(x) = \begin{cases} 1 & \text{for } T(x) < C_{\alpha,1} \text{ or } T(x) > C_{\alpha,2}, \\ \gamma_{\alpha,i} & \text{for } T(x) = C_{\alpha,i},\ i = 1, 2, \\ 0 & \text{for } C_{\alpha,1} < T(x) < C_{\alpha,2}. \end{cases}$$

See, e.g., Lehmann (1986, p. 135).

For simplicity of notation, set $C_\alpha \equiv (C_{\alpha,1} + C_{\alpha,2})/2$, $\theta_0 \equiv (\theta_1+\theta_2)/2$, $T_\alpha(x) \equiv T(x) - C_\alpha$, $h_1(x) \equiv h(x)\exp[\theta_0 T(x)]$, $\omega(\theta) \equiv \theta - \theta_0$, and $c_\alpha(\theta) \equiv c(\theta) - C_\alpha\omega(\theta)$. Then

$$f(x;\theta) = h_1(x)\exp[\omega(\theta)T_\alpha(x) - c_\alpha(\theta)].$$

Consider a weight function Λ on $(-\infty,\theta_1)\cup(\theta_2,\infty)$ $(\equiv\Theta_1)$ with

$$\frac{\Lambda(d(2\theta_0-\theta))}{\Lambda(d\theta)} = \exp[c_\alpha(2\theta_0-\theta) - c_\alpha(\theta)].$$

Moreover, consider another weight function Λ_0 on $[\theta_1,\theta_2]$ $(\equiv\Theta_0)$ with $\Lambda_0(\{\theta_1\}) = p_\alpha$ and $\Lambda_0(\{\theta_2\}) = 1 - p_\alpha$, where $p_\alpha \equiv \exp[c_\alpha(\theta_1) - c_\alpha(2\theta_0-\theta_1)]/\{1+\exp[c_\alpha(\theta_1) -$

$c_\alpha(2\theta_0 - \theta_1)]\}$. Then,

$$\begin{aligned}
&\int_{\Theta_1} f(x;\theta)\Lambda(d\theta)\\
&\quad = \int_{(-\infty,\theta_1)} \left[f(x;\theta) + f(x; 2\theta_0 - \theta)\frac{\Lambda(d(2\theta_0-\theta))}{\Lambda(d\theta)}\right]\Lambda(d\theta)\\
&\quad = h_1(x)\int_{(-\infty,\theta_1)} \{\exp[T_\alpha(x)\omega(\theta)] + \exp[-T_\alpha(x)\,\omega(\theta)]\}\exp[-c_\alpha(\theta)]\Lambda(d\theta).
\end{aligned}$$

Thus,

$$\frac{\int_{\Theta_1} f(x;\theta)\Lambda(d\theta)}{\int_{\Theta_0} f(x;\theta)\Lambda_0(d\theta)} \propto \int_{(-\infty,\theta_1)} \frac{\exp[T_\alpha(x)\omega(\theta)] + \exp[-T_\alpha(x)\omega(\theta)]}{\exp[T_\alpha(x)\omega(\theta_1)] + \exp[-T_\alpha(x)\omega(\theta_1)]}\exp[-c_\alpha(\theta)]\Lambda(d\theta).$$

It is easy to show that $\{\exp[t\omega(\theta)] + \exp[-t\omega(\theta)]\}/\{\exp[t\omega(\theta_1)] + \exp[-t\omega(\theta_1)]\}$ is a strictly increasing function of $|t|$ for $\theta < \theta_1$, which implies that $\int_{\Theta_1} f(x;\theta)\Lambda(d\theta)/\int_{\Theta_0} f(x;\theta)\Lambda_0(d\theta)$, is also a strictly increasing function of $|T_\alpha(x)|$. Since ϕ^*_α is a UMPU test, we have that $\beta_{\phi^*_\alpha}(\theta_1) = \beta_{\phi^*_\alpha}(\theta_2) = \alpha$ and $\beta_{\phi^*_\alpha}(\theta) \le \alpha$ for $\theta_1 < \theta < \theta_2$. Also, ϕ^*_α is a strictly increasing function of $|T_\alpha(x)|$, we have that ϕ^*_α is a level α MAP test w.r.t. Λ for testing the simple null hypothesis $H'_0 : X$ has pdf $\int_{\Theta_0} f(x;\theta)\Lambda_0(d\theta)$, $x \in \mathscr{X}$, w.r.t. μ vs. the alternative $H_1 : \theta \in \Theta_1$. Consequently, it follows from Theorem 5.2 that ϕ^*_α is a level α MAP test w.r.t. Λ.

6. Remarks

In this article, we establish some fundamental results concerning the existence of an MAP test, the test that achieves the maximum average power. This optimal test always exists, and in this regard the corresponding criterion is more useful than other criteria such as those leading to UMP or UMPU tests.

For a simple null hypothesis, it is easy to construct an MAP test; however, for a composite null hypothesis, special care needs to be taken in order to construct an MAP test. Although we have some results for such cases in this article, further investigation is needed for other situations.

In some discrete cases, we constructed randomized tests which are MAP. From the scientific point of view, it will be important to construct non randomized tests that have good average power such as those in Martin and Silva (1994), Martin et al. (1998), and Martin and Tapia (1999). Due to the discrete nature, the analytical problem is very difficult. However, further investigation in the direction may be fruitful, especially if developed with the help of high-power computing. Perhaps one may start out with the randomized MAP test and try to construct a non-randomized test to approximate it in some sense.

On the other hand, there is an interesting theory in Brown et al. (1995), which is related to our MAP approach. They focus on confidence interval and confidence set construction however. To simplify the discussion below, we write about the confidence interval construction although similar comments apply to the confidence set construction. By using the idea of finding an MAP test outlined in this article, they constructed a confidence interval that minimizes the average length among all the confidence intervals with coverage probability no less than a certain nominal level such as 0.95. The identity (Ghosh–Pratt identity) that relates the MAP test to the result of Brown et al. (1995) is the fact that maximizing the average power w.r.t. a weight function leads to minimizing the average length w.r.t. the same weight function. Further investigation in the confidence interval and confidence set construction approach for more complicated multiple parameter cases than those in Brown et al. (1995) may be interesting.

Drawing from the above discussion, the use of MAP approach in statistical inferences is a topic which has received only cursory attention. Further research and attention in application are required.

Acknowledgments

We are grateful to Professor J. T. Gene Hwang for his stimulating comments and suggestions, as well as for reviewing the manuscript.

We are also grateful to the editor and one referee for comments which improved the presentation of this article. This research work was partially supported by the National Science Council of the Taiwan Grant Numbers NSC 95-2118-M-009-007 and NSC 95-2118-M-009-004.

References

Brown, L. D., Casella, G., Hwang, J. T. G. (1995). Optimal confidence sets, bioequivalence, and the limacon of Pascal. *J. Amer. Statist. Assoc.* 90:880–889.

Cui, X., Hwang, J. T. G., Qiu, J., Blades, N. J., Churchill, G. A. (2005). Improved statistical tests for differential gene expression by shrinking variance components estimates. *Biostatistics* 6:59–75.

Lehmann, E. L. (1986). *Testing Statistical Hypotheses*. 2nd ed. New York: John Wiley & Sons.

Martin, A. A., Silva, M. A. (1994). Choosing the optimal unconditioned test for comparing two independent proportions. *Comput. Statist. Data Anal.* 17:555–574.

Martin, A. A., Tapia, G. J. M. (1999). Optimal unconditional test in 2×2 multinomial trials. *Comput. Statist. Data Anal.* 31:311–321.

Martin, A. A., Sanchez, Q. M. J., Silva, M. A. (1998). Fisher's mid p-value arrangement in 2×2 comparative trials. *Comput. Statist. Data Anal.* 29(1):107–115.

Neyman, J., Pearson, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses. *Phil. Trans. Roy. Soc. Ser. A* 231:289–337.

Neyman, J., Pearson, E. S. (1936). Contributions to the theory of testing statistical hypotheses. *Statist. Res. Mem.* 1:1–37.

Neyman, J., Pearson, E. S. (1938). Contributions to the theory of testing statistical hypotheses. *Statist. Res. Mem.* 2:25–57.