Statistica Sinica **9**(1999), 1131-1140

MULTIPLE HYPOTHESES TESTING WITH PARTIAL PRIOR INFORMATION

Jian Zhang

Academia Sinica, Beijing and Universite Catholique de Louvain

Abstract: By modifying the optimal criterion of Kudo (1967), an alternative theory is developed for incorporating prior probabilities into multiple hypotheses testing. Unlike Kudo's theory, an explicit optimal test can be obtained. It is shown that the optimal test is a natural extension of some well-known tests such as the maximum likelihood test.

Key words and phrases: Bayesian robust analysis, multiple hypotheses testing, partial prior information.

1. Introduction

Let $X = (x_1, \ldots, x_n)$ be a random sample of size n with a joint density $f(X, \theta)$ with respect to a σ finite measure μ, where θ is an unknown parameter lying in a subset Θ of a metric space, and X takes values in an Euclidean space Ω. Let $\{\Theta_i, 1 \leq i \leq m\}$ be a finite partition of Θ. We consider the problem of testing multiple hypotheses

$$H_j : \theta \in \Theta_j, 1 \leq j \leq m, \tag{1.1}$$

in which it is desired to decide to which one of the m subsets θ belongs on the basis of the sample X. The decision can be described by a test function $\phi(X) = (\phi_1(X), \ldots, \phi_m(X))$ with $\phi_j(X) \geq 0, 1 \leq j \leq m$, and $\sum_{j=1}^m \phi_j(X) = 1$. We accept H_j when $\phi_j(X) = 1$. For each $1 \leq j \leq m$, a non-negative function $L_j(\theta)$ is given to designate the loss corresponding to the acceptance of H_j, when θ is the value of the parameter. The commonly used loss function is

$$L_j(\theta) = \begin{cases} 0, & \theta \in \Theta_j, \\ L_{ji}, & \theta \in \Theta_i, \text{ and } i \neq j, \end{cases} \tag{1.2}$$

for $1 \leq j \leq m$. There are many testing problems with more than two hypotheses. For instance, when one wishes to test which of two trials is more effective, one frequently is faced with the third hypothesis representing indifference. Even in classical testing there may exist three hypotheses: the null hypothesis H_o; the alternative H_1; there is no significant evidence for accepting either H_o or H_1 (see

Berger (1985)). As pointed out by Lehmann (1986, p.380), in the problem of multiple comparisons, testing equality of m means as a simple choice between acceptance and rejection usually leaves many questions unanswered. In particular, when the hypothesis is rejected one would like to obtain more detailed information about the relative position of the means. Thus, a test of more than m hypotheses is required. Another type of common multiple hypotheses testing problem is that of classifying an observation into one of several populations (see Rao (1973, pp.491-497), or Anderson (1984, Chapter 6)).

In the literature, there are at least two approaches to multiple hypotheses testing— the Bayesian and minimax approaches (see, for example, Zacks (1971, Chapter 9), or Berger (1985)). The former uses prior information that can be completely and accurately quantified in terms of a single prior distribution, whereas the latter completely ignores the prior information. In some practical situations, there really exists prior information on the parameter. However, it is often found difficult to assign an appropriate prior distribution. The reason is that the prior information may be incomplete and the population underlying the prior information is usually not consistent with that of the sample (see Berger (1985, p.240), Zhang (1992, 1995)).

To take partial prior information into account, Hodges and Lehmann (1952) suggested a compromise between the Bayesian and minimax approaches. Blum and Rosenblatt (1967) studied the minimax estimate when the prior distribution is restricted to certain families of prior distributions. Kudo (1967) introduced the mean-max criterion to choose the estimate when the prior is defined on a subsigma field of sets. Recent treatments from the viewpoint of Bayesian robustness can be found in Berger (1985, 1990, 1994), and in the references therein.

In this paper we are mainly concerned with the partial prior model (Kudo (1967)), in which a statistician is informed of nothing but the prior probabilities of m hypotheses, say $p_i, 1 \leq i \leq m$, with $\sum_{i=1}^m p_i = 1$, and that the prior distribution π belongs to the following class of prior distributions

$$PM = \{\pi : \int_{\Theta_i} \pi(d\theta) = p_i, \quad 1 \leq i \leq m\}.$$

PM is of substantial interest in that the prior quantiles might be easy to specify for practitioners. Berliner and Goel (1990) determined ranges of posterior probabilities of sets with respect to this class of priors. Related studies were made by DeRobertis and Hartigan (1981), Manski (1981), Cano, Hernandez and Moreno (1985), Lambert and Duncan (1986), among others. There have been relatively few works on the problem of Bayesian robust hypothesis testing. A recent paper in this direction was provided by Berger and Mortora (1994) who considered the situation, in the presence of nuisance parameters, from the posterior decision point of view. In this paper, we investigate the following question: which procedure for testing (1.1) is optimal under the above partial prior model, and in what sense is it optimal?

To motivate our procedure, we consider the simplest case: $H_1 : \theta \in \Theta_1$ versus $H_2 : \theta \in \Theta_2$, where Θ_1 and Θ_2 are bounded. Suppose there are prior probabilities p_1 and p_2 on Θ_1 and Θ_2, respectively. For $i = 1, 2$, let π_i be a uniform distribution on Θ_i, $\pi = \pi_1 p_1 + \pi_2 p_2$ the prior distribution, and let $m = \int_{\Theta_1} f(X,\theta)\pi_1(d\theta)p_1 + \int_{\Theta_2} f(X,\theta)\pi_2(d\theta)p_2$. Then the posterior density of θ is

$$\pi(\theta|X) = \begin{cases} \frac{f(X,\theta)\pi(\theta)p_1}{m}, & \theta \in \Theta_1, \\ \frac{f(X,\theta)\pi(\theta)p_2}{m}, & \theta \in \Theta_2. \end{cases}$$

The ratio of the maximum posterior densities of Θ_1 and Θ_2 is

$$R = C\,\frac{\sup_{\Theta_1} f(X,\theta)p_1}{\sup_{\Theta_2} f(X,\theta)p_2},$$

where c is a constant. As a function of the sample X, R can be used to test H_1 versus H_2. When $p_1 = p_2 = 0.5$, this test reduces to the well-known Maximum Likelihood Ratio Test (see, for example, Cox and Hinkley (1974)). The procedure suggested here is an extension of the above ratio test.

In the next sections, we first establish a frequentist Bayesian criterion to measure the risk of a testing procedure $\phi(X)$ with respect to PM. An optimal test is developed in terms of this criterion. We show that, suitably choosing prior information, we can recover the Maximum Likelihood Ratio Test and the Bayes procedures in Zacks (1971, p.425) and Rao (1973, p.493). Then we investigate the related questions for the quantile class of priors and the ϵ-corrupted class of priors.

2. The Main Theory

2.1. An alternative criterion and test

Kudo (1967) suggested the use of the maximum Bayes risk with respect to PM to measure the risk of a decision procedure. The optimal solution with respect to this criterion seems not easy to solve. In the following, an alternative criterion is proposed. The main advantage of the new criterion is that the corresponding optimal test has an explicit formula which is a natural extension of several classical tests, including the Maximum Likelihood Ratio Test.

To begin with, let $\theta_o \in \Theta$ denote the unknown true parameter. In light of the likelihood principle (see, for instance, Berger (1985, p.27)), given sample X, a θ for which $f(X,\theta)$ is large is more “likely” to be the true θ than a θ for which $f(X,\theta)$ is small. Note that the loss associated with the acceptance of H_j is

$L_j(\theta)$ when θ_o is θ. Hence, given the sample X, $L_j(\theta)f(X,\theta)$ can be used as the weighted loss of accepting H_j and taking θ as the true value of the parameter. Averaging the weighted loss with respect to the prior distribution π yields the average weighted loss of the acceptance of H_j given X and π. Then, given X, the maximum weighted loss of the acceptance of H_j under the partial model $\pi \in PM$ is

$$\sup_{\pi \in PM} \int_\Theta f(X,\theta)L_j(\theta)\pi(d\theta).$$

This gives the expected relative maximum weighted loss of the acceptance of H_j,

$$\begin{aligned}&\int \phi_j(X)\frac{\sup_{\pi\in PM}\int_\Theta L_j(\theta)f(X,\theta)\pi(d\theta)}{\sum_{k=1}^m \sup_{\pi\in PM}\int_\Theta L_k(\theta)f(X,\theta)\pi(d\theta)} f(X,\theta_o)\mu(dX)\\ &= \int \phi_j(X)\frac{\sup_{\pi\in PM} m(X|\pi)\int_\Theta L_j(\theta)m(d\theta|X)}{\sum_{k=1}^m \sup_{\pi\in PM} m(X|\pi)\int_\Theta L_k(\theta)m(d\theta|X)} f(X,\theta_o)\mu(dX),\end{aligned} \tag{2.1}$$

where $m(X|\pi)$ and $m(d\theta|X)$ are the marginal density of X and the posterior distribution of θ given X, respectively. The risk of a decision ϕ can be measured by

$$R(\phi,\theta_o) = \sum_{j=1}^m \int \phi_j(X)\frac{\sup_{\pi\in PM}\int_\Theta L_j(\theta)f(X,\theta)\pi(d\theta)}{\sum_{k=1}^m \sup_{\pi\in PM}\int_\Theta L_k(\theta)f(X,\theta)\pi(d\theta)} f(X,\theta_o)\mu(dX). \tag{2.2}$$

The decision ϕ^o with $R(\phi^o,\theta_o) = \min R(\phi,\theta_o)$ for all $\theta_o \in \Theta$ is called the PM optimal test of the multiple hypotheses (1.1). An explicit form of such a test is obtained in the following theorem.

Theorem 2.1. *Suppose that $f(X,\theta)$ is measurable with respect to (X,θ), and that for $1 \le i \ne j \le m$ and each X, $L_j(\theta)f(X,\theta)$ is continuous on Θ_i with respect to θ. For $1 \le j \le m$, assume*

$$M_j(X) = \sum_{k=1}^m p_k \sup_{\Theta_k} L_j(\theta)f(X,\theta) < \infty.$$

Then the PM optimal test $\phi^o = (\phi_1^o,\ldots,\phi_m^o)$ has

$$\phi_j^o = \begin{cases} 1, & M_j(X) = \min_{1\le k\le m} M_k(X),\\ 0, & \text{otherwise},\end{cases}$$

$1 \le j \le m$. We choose $\phi_i^o(X)$, $\phi_j^o(X)$ such that $\phi_i^o(X)\phi_j^o(X) = 0$ if $M_j(X) = M_i(X)$ for some $i < j$.

Proof of Theorem 2.1. Observe that

$$\sup_{\pi\in PM}\int_\Theta L_j(\theta)f(X,\theta)\pi(d\theta) = \sum_{k=1}^m p_k \sup_{\theta\in\Theta_k} L_j(\theta)f(X,\theta).$$

For $\theta_o \in \Theta$, we have

$$R(\phi, \theta_o) = \sum_{j=1}^{m} \int \phi_j(X) \frac{M_j(X)}{\sum_{k=1}^{m} M_k(X)} f(X, \theta_o) \mu(dX)$$

which attains the minimum at $\phi = \phi^o$.

Remark. If for each j, Θ_j is a finite set, then Theorem 2.1 holds. If for each j, Θ_j is a singleton, then the PM optimal test is just the Bayes solution discussed in Rao (1973, p.493).

The main difference between the classical Bayes test and the PM optimal test lies in the use of the prior distribution. The Bayes approach incorporates the whole prior distribution into the test while the PM optimal test does not. In practice, there are at least two situations to consider. In one, several prior quantiles can be specified directly from the partial prior information. In another, the prior distribution can be selected, but with some uncertainties, for example, about its tails. In the former situation, the PM optimal test can be directly used. In the latter situation, we can first select part of the prior distribution, for example, the corresponding prior probabilities of hypotheses. Then we use the PM optimal test. Thus, we can expect that the robust behavior of the PM optimal test with respect to the prior distribution is better than that of the Bayes test. Although the Bayesian robust analysis in Berger (1985) can be used in these situations, some additional efforts may be necessary to come to a unique decision. The following example highlights this point.

Example 2.1. (see Berger (1985), p.166). Suppose one wishes to classify the intelligence (IQ) level θ of a child as $\Theta_1 : \theta < 90$ (below average), $\Theta_2 : 90 \leq \theta \leq 110$ (average), or $\Theta_3 : \theta > 110$ (above average) by an IQ test. The loss function is defined by

$$L_1(\theta) = \begin{cases} 0, & \text{if } \theta < 90, \\ \theta - 90, & \text{if } 90 \leq \theta \leq 110, \\ 2(\theta - 90), & \text{if } \theta > 110, \end{cases}$$

$$L_2(\theta) = \begin{cases} 90 - \theta, & \text{if } \theta < 90, \\ 0, & \text{if } 90 \leq \theta \leq 110, \\ \theta - 110, & \text{if } \theta > 110, \end{cases}$$

$$L_3(\theta) = \begin{cases} 2(110 - \theta), & \text{if } \theta < 90, \\ 110 - \theta, & \text{if } 90 \leq \theta \leq 110, \\ 0, & \text{if } \theta > 110. \end{cases}$$

Assume that the test result x follows $N(\theta, 100)$, a normal distribution with mean θ and variance 100. Assume also that, in the population as a whole, θ is distributed according to F. The question is which class the IQ level of a child belongs to if he scores 115 (i.e., $x = 115$) on the test.

Bayes approach: Suppose that there are two possible choices of F, $F = N(100, 15^2)$ and $F = N(100, 10^2)$, due to incompleteness of the prior information. Then, by the theory in Berger (1985, p.166), we obain two different Bayes decisions, $\theta \in \Theta_3$ and $\theta \in \Theta_2$ when we adopt $F = N(100, 15^2)$ and $F = N(100, 10^2)$, respectively. Which decision should we take? We might use the PM optimal test to answer the question.

PM optimal test: Let $f(115, \theta) = 0.1 \exp(-(115-\theta)^2/200)/\sqrt{2\pi}$, and p_i denote the prior probabilities on Θ_i, $i = 1, 2, 3$. Then

$$M_1 = p_2 \sup_{90\le\theta\le110}(\theta - 90)f(115, \theta) + p_3 \sup_{\theta>110} 2(\theta - 90)f(115, \theta),$$
$$M_2 = p_1 \sup_{\theta<90}(90 - \theta)f(115, \theta) + p_3 \sup_{\theta>110}(\theta - 110)f(115, \theta),$$
$$M_3 = p_1 \sup_{\theta<90} 2(110 - \theta)f(115, \theta) + p_2 \sup_{90\le\theta\le110}(110 - \theta)f(115, \theta).$$

When $F = N(100, 15^2)$, we have $p_1 = p_3 = 0.2514$, $p_2 = 0.4972$, $M_1 = 0.888$, $M_2 = 0.269$ and $M_3 = 0.086$. The decision of the PM optimal test is $\theta \in \Theta_3$. When $F = N(100, 10^2)$, we have $p_1 = p_3 = 0.1587$, $p_2 = 0.6826$, $M_1 = 0.820$, $M_2 = 0.170$ and $M_3 = 0.105$. The decision of the PM optimal test is still $\theta \in \Theta_3$. Thus, given uncertainty about the selected prior distributions, the PM optimal test seems preferable to the Bayes approach because of its robustness.

2.2. Extensions and special cases

Theorem 2.1 can be extended to the following situation, which occurs for example when the prior density is specified by the histogram approach or by several quantiles (see Berger (1985), p.77).

Suppose that for each Θ_i, $1 \le i \le m$, there exist a partition $\cup_{k=1}^{t_i}\Theta_{ik}$ and prior probabilities p_{ik} on Θ_{ik}, $1 \le k \le t_i$. This partial prior information model is denoted by PMR. Then, under PMR and the loss function $L_i(\theta), 1 \le i \le m$, the optimal test of the multiple hypotheses (1.1) can be obtained in the following way. Consider the problem of testing $\sum_{i=1}^m t_i$ hypotheses $H_{ik} : \theta \in \Theta_{ik}$, $1 \le k \le t_i$, $1 \le i \le m$ with the following loss function: if $\theta \in \Theta_{ik}$, then the loss of accepting H_{ij} is 0 and the loss of accepting H_{vj} is $L_i(\theta)$. Let ϕ_{ik}^o, $1 \le k \le t_i, 1 \le i \le m$ be the optimal test. The desired decision is of the form $\phi_i^o = \sum_{k=1}^{t_i} \phi_{ik}^o$ for $1 \le i \le m$. In particular, as $\min_{1\le i\le m} t_i$ tends to ∞, the limit of the PM optimal test is the Bayes procedure in Zacks (1971, p.425).

In some practical situations, there may be misspecification of the prior probabilities, and some of the prior probabilities may be missing. From the robust viewpoint, we can use the following two classes of priors to express such uncertainty and modify Theorem 2.1 in these cases.

(i) A quantile class of priors

$$QPM = \{\pi : l_i \le \int_{\Theta_i} \pi(d\theta) \le u_i, i = 1, \ldots, m\},$$

where $0 \le l_i \le u_i \le 1$, $i = 1, \ldots, m$ are predetermined constants (see Berger (1990)).

For each $1 \le j \le m$, we sort $\sup_{\Theta_i} L_j(\theta) f(X,\theta)$, $i = 1, \ldots, m$, from the largest to the smallest. For simplicity, we assume

$$\sup_{\Theta_1} L_j(\theta) f(X,\theta) \ge \cdots \ge \sup_{\Theta_m} L_j(\theta) f(X,\theta).$$

Write

$$s = \max\{k \ge 1 : \sum_{i=1}^{k} (u_i - l_i) \le 1 - \sum_{i=1}^{m} l_i\}.$$

Set $p^o_{ij} = u_i$, $i = 1, \ldots, m$, if $s = m$; $p^o_{ij} = u_i$, $i = 1, \ldots, s$, $p^o_{s+1,j} = l_{s+1} + 1 - \sum_{i=1}^{s} u_i + \sum_{i=s+1}^{m} l_i$ and $p^o_{ij} = l_i$, $i = s+2, \ldots, m$. Then, under QPM, the conclusion of Theorem 2.1 still holds when for each $1 \le j \le m$, $p_i, 1 \le i \le m$, in M_j is replaced by $p^o_{ij}, 1 \le i \le m$.

(ii) An ϵ-corrupted class, EPM, in which there exists a subset N_o of $\{1, \ldots, m\}$ and a positive number ϵ such that $\sum_{i \in N_o} p_i = 1 - \epsilon$ and p_i is determined only when $i \in N_o$. Then, under EPM, the conclusion of Theorem 2.1 remains true when M_j is replaced by

$$\sum_{i \in N_o} p_i \sup_{\Theta_i} L_j(\theta) f(X,\theta) + \epsilon \sup_{\cup_{i \notin N_o, i \ne j} \Theta_i} L_j(\theta) f(X,\theta),$$

for $1 \le j \le m$.

Proof of (i) and (ii). The key step is to calculate

$$\sup_{QPM} \sum_{k=1}^{m} \int_{\Theta_k} L_j(\theta) f(X,\theta) \pi(d\theta)$$

and

$$\sup_{EPM} \sum_{k=1}^{m} \int_{\theta \in \Theta_k} L_j(\theta) f(X,\theta) \pi(d\theta),$$

which is straightforward. The proof is completed.

The following special cases are useful to illustrate Theorem 2.1 and its extensions.

Special cases. Assume that $L_j(\theta)$ is of the form (1.2). Without loss of generality, assume $L_{1m} \ne 0$. Then

$$M_j(X) = \sum_{i \ne j}^{m} p_i L_{ji} \sup_{\theta \in \Theta_i} f(X,\theta)$$

and the PM optimal test depends on $L_{ji}, 1 \le i,\ j \le m$ only through the ratios L_{ji}/L_{1m}, $1 \le i,\ j \le m$.

In particular, if $m = 2$, we have

$$M_1(X) = p_2 L_{12} \sup_{\theta\in\Theta_2} f(X,\theta), \quad M_2(X) = p_1 L_{21} \sup_{\theta\in\Theta_1} f(X,\theta),$$

$$\phi_1^o(X) = \begin{cases} 1, & p_1 \sup_{\Theta_1} f(X,\theta)/p_2 \sup_{\theta\in\Theta_2} f(X,\theta) \ge \frac{L_{12}}{L_{21}}, \\ 0, & \text{otherwise}, \end{cases} \tag{2.3}$$

and $\phi_2^o = 1 - \phi_1^o$. It means that we narrow or relax the rejection region of the classical maximum likelihood ratio test according to the ratio p_2/p_1.

Let $m = 3$, $L_{ij} = 1$, $L_{ii} = 0$, for $1 \le i \ne j \le 3$ (0-1 loss). Usually Θ_2 is an indifference region. Then, we have

$$\begin{aligned} M_1 &= \sup_{\theta\in\Theta_2} f(X,\theta)p_2 + \sup_{\theta\in\Theta_3} f(X,\theta)p_3, \\ M_2 &= \sup_{\theta\in\Theta_1} f(X,\theta)p_1 + \sup_{\theta\in\Theta_3} f(X,\theta)p_3, \\ M_3 &= \sup_{\theta\in\Theta_1} f(X,\theta)p_1 + \sup_{\theta\in\Theta_2} f(X,\theta)p_2. \end{aligned}$$

For simplicity, assume $M_i \ne M_j$ for $i \ne j$. Then, for $i = 1, 2, 3$,

$$\phi_i^o = \begin{cases} 1, \sup_{\theta\in\Theta_i} f(X,\theta)p_i > \max_{j\ne i} \sup_{\theta\in\Theta_j} f(X,\theta)p_j, \\ 0, \ \text{otherwise} . \end{cases} \tag{2.4}$$

Let $m = 3$ and assume that $L_{12} = 0 = L_{21}$, $L_{13} = L_{23}$, and $L_{31} = L_{32}$, that is, we actually wish to test the null hypothesis $\theta \in \Theta_3$ against the alternative $\theta \in \Theta_1 \cup \Theta_2$ with the prior probabilities on Θ_i, $1 \le i \le 3$. Then, we have

$$\begin{aligned} M_1(X) &= M_2(X) = L_{13} \sup_{\theta\in\Theta_3} f(X,\theta)p_3, \\ M_3(X) &= L_{31}(\sup_{\theta\in\Theta_1} f(X,\theta)p_1 + \sup_{\theta\in\Theta_2} f(X,\theta)p_2); \end{aligned}$$

$$\begin{aligned} &\phi_1^o(X) + \phi_2^o(X) = \begin{cases} 1, \frac{\sup_{\theta\in\Theta_1} f(X,\theta)p_1}{\sup_{\theta\in\Theta_3} f(X,\theta)p_3} + \frac{\sup_{\theta\in\Theta_2} f(X,\theta)p_2}{\sup_{\theta\in\Theta_3} f(X,\theta)p_3} \ge \frac{L_{13}}{L_{31}}, \\ 0, \ \text{otherwise} , \end{cases} \\ &\phi_1^o(X)\phi_2^o(X) = 0 \text{ and } \phi_3^o(X) = 1 - \phi_1^o(X) - \phi_2^o(X). \end{aligned} \tag{2.5}$$

Acknowledgements

The work was partly supported by the NNSF of China and the contract "Projet d'Actions de Recherche Concertées" no. 93/98-164. The author greatly appreciates the Institute of Statistics, Universite Catholique de Louvain for its hospitality and support. The author is especially grateful to Professor I. Gijbels, the Chair-Editor and referees for helpful comments which led to the improvement of presentation.

References

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis.* 2nd edition. John Wiley, New York.

Berger, J. (1985). *Statistical Decision Theory and Bayesian Analysis.* Springer-Verlag, New York.

Berger, J. (1990). Robust Bayesian analysis: sensitivity to the prior. *J. Statist. Plann. Inference* **25**, 303-328.

Berger, J. (1994). Bayesian robustness in bi-dimensional models: prior independence. *J. Statist. Plann. Inference* **40**, 161-176.

Berger, J. and Mortora, J. (1994). Robust Bayesian hypothesis testing in the presence of nuisance parameters. *J. Statist. Plann. Inference* **40**, 357-373.

Berliner, L. M. and Goel, P. (1990). Incorporating partial prior information: ranges of posterior probabilities. In *Bayesian and Likelihood Methods in Statistics and Econometrics: Essays in Honor of George A. Barnard* (Edited by S. Geisser, J. Hodges, F. J. Press and A. Zellner), 397-406. North-Holland, Amsterdam.

Blum, J. R. and Rosenblatt, J. (1967). On partial a priori information in statistical inference. *Ann. Math. Statist.* **38**, 1671-1678.

Cox, D. R. and Hinkley, D. V. (1974). *Theoretical Statistics.* Chapman and Hall, New York.

Cano, J. A., Hernandez, A. and Moreno, E. (1985). Posterior measures under partial prior information. *Statist.* **2**, 219-230.

DeRobertis, L. and Hartigan, J. A. (1981). Bayesian inference using intervals of measures. *Ann. Statist.* **1**, 235-244.

Hodges, J. L. Jr. and Lehmann, E. L. (1952). The use of previous experience in reaching statistical decisions. *Ann. Math. Statist.* **23**, 396-407.

Kudo, H. (1967). On partial prior information and the property of parametric sufficiency. *Proc. Fifth Berkely Symp. Statist. and Prob.* **1**, 251-265. University of California Press, Berkeley, CA.

Lambert, D. and Duncan, G. (1986). Single-parameter inference based on partial prior information. *Canad. J. Statist.* **14**, 297-305.

Lehmann, E. L. (1986). *Testing Statistical Hypotheses.* 2nd edition. Wiley, New York.

Manski, C. F. (1981). Learning and decision making when subjective probabilities have subjective domains. *Ann. Statist.* **9**, 59-65.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications.* 2nd edition. Wiley, New York.

Zacks, S. (1971). *The Theory of Statistical Inference.* Wiley, New York.

Zhang, J. (1992). On Bayes estimates of impact dispersion of reentrant vehicle, II. *Proceedings of the Second Conference of the Chinese Society of Industry and Applied Mathematics.* Hangzhou, China.

Zhang, J. (1995). On Bayes estimates of impact dispersion of reentrant vehicle, I. *Math. Statist. Appl. Prob.* **10**, 21-26.

Institute of Systems Science, Academia Sinica, Beijing 100080, China.

E-mail: jzhang@iss01.iss.ac.cn

(Received October 1997; accepted December 1998)