Ann. Inst. Statist. Math.
37 (1985), Part A, 481-486

A RANDOMIZED SOLUTION FOR MULTI-BAYES ESTIMATES OF THE MULTINORMAL MEAN*

D. J. DE WAAL, P. C. N. GROENEWALD, J. M. VAN ZYL AND J. V. ZIDEK

(Received Dec. 14, 1983; revised Feb. 19, 1985)

Summary

This paper is concerned with the randomized solution for the estimation of the multinormal mean by two or more Bayesians. The optimum rule is found by maximizing the Kalai-Nash product in the case when randomization is necessary. It can be considered as an extension on the work by Weerahandi and Zidek [2], [3].

1. Introduction

Weerahandi and Zidek [3] consider the solution to the problem of estimating the mean θ of a multivariate normal population by two Bayesians if the posterior density of θ for the ith Bayesian is multivariate normal with mean θ_i and covariance matrix Σ_i. Assuming a utility function

$$U_i(\hat{\theta}, \theta)\propto\exp\left[-\frac{1}{2}(\theta-\hat{\theta})'W_i^{-1}(\theta-\hat{\theta})\right]$$

for the ith Bayesian where $\hat{\theta}$ is the estimate, W_i and Σ_i are given p.s. matrices. The expected utility for Bayesian i is

$$(1.1)\qquad U_i(\hat{\theta})=\mathrm{E}\,[U_i(\hat{\theta}, \hat{\theta}_i)]=\exp\left[-\frac{1}{2}(\theta_i-\hat{\theta})'\Lambda_i^{-1}(\theta_i-\hat{\theta})\right]$$

where $\Lambda_i=W_i+\Sigma_i$. The authors, referred to above, gave the randomized solution for $\Lambda_1=\Lambda_2$ and $\alpha_1=\alpha_2$ where α_1 and α_2 are defined in the Kalai-Nash product

$$(1.2)\qquad P(\delta)=\left[\int U_1(\hat{\theta})\delta d(\hat{\theta})\right]^{\alpha_1}\left[\int U_2(\hat{\theta})\delta d(\hat{\theta})\right]^{\alpha_2}, \qquad \alpha_1+\alpha_2=1,\ \alpha_i\geqq 0$$

Key words: Multi-Bayes estimates, Kalai-Nash product, multinormal, randomized rule.
* Financially supported by the CSIR and UOFS.

and δ is the randomized rule. They also gave the solution if $\alpha_1 \neq \alpha_2$ and $\Lambda_1 = \Lambda_2$. If $\Lambda_1 \neq \Lambda_2$, no explicit formulae are given for obtaining the randomized solution but results which will be of some potential value are given.

We will consider in this paper the randomized solution in general for the two Bayesian case and will also consider the problem of more than two Bayesians. Let $S = \{(U_1(\hat{\theta}), \cdots U_n(\hat{\theta}));\ \hat{\theta} \in R^p\}$ and $\bar{S}$ denote the convex hull of these points. $\mathscr{B}$ is the set of all Pareto-optimal points of S, and is a subset of the boundary of S denoted by ∂S. If $S \neq \bar{S}$ the optimal solution will be randomized. (Weerahandi and Zidek [2], [3]). If $U_n(\hat{\theta})$ is maximized with respect to $\hat{\theta}$, keeping $U_j(\hat{\theta})$ $(j \neq n)$ fixed, the estimator

$$\hat{\theta}_\lambda = \Lambda^{-1}\left(\Lambda_n^{-1}\theta_n + \sum_{i=1}^{n-1} \lambda_i \Lambda_i^{-1}\theta_i\right) \tag{1.3}$$

with $\Lambda = \left(\Lambda_n^{-1} + \sum_{i=1}^{n-1} \lambda_i \lambda_i^{-1}\right)$ and $\lambda_i \geqq 0$, $i = 1, \cdots, n-1$ is obtained, where $\lambda_1, \cdots, \lambda_{n-1}$ are the Lagrange multipliers. This estimator yields the Pareto-optimal set of utilities,

$$\mathscr{B} = \left\{U_i(\hat{\theta}_\lambda),\ i = 1, \cdots, n_1;\ \hat{\theta}_\lambda = \Lambda^{-1}\left(\Lambda_n^{-1}\theta_n + \sum_{i=1}^{n-1} \lambda_i \Lambda_i^{-1}\theta_i\right),\ \lambda_i \geqq 0\right\}. \tag{1.4}$$

The Kalai-Nash product (Nash [1], Weerahandi and Zidek [2]) is defined as

$$P(\hat{\theta}) = \prod_{i=1}^{n} [U_i(\hat{\theta})]^{\alpha_i}, \qquad \sum_i \alpha_i = 1,\ \alpha_i \geqq 0,\ i = 1, \cdots, n. \tag{1.5}$$

It can be shown by completing the square in the exponent of $P(\hat{\theta})$, that among the nonrandomized rules in the n Bayesian case, the value of $\hat{\theta}$ that maximizes $P(\hat{\theta})$ is: Choose with certainty the action

$$\hat{\theta} = \left(\sum_{i=1}^{n} \alpha_i \Lambda_i^{-1}\right)^{-1}\left(\sum_{i=1}^{n} \alpha_i \Lambda_i^{-1}\theta_i\right) \tag{1.6}$$

which is a weighted average of θ_i, $i = 1, \cdots, n$. It is clear from (1.6) and (1.4) that the point which maximizes (1.5) is an element of $\mathscr{B}$. If S is not convex, the optimal solution lies on the boundary $\bar{S}$ and randomization is necessary so that (1.6) is no longer optimal (Weerahandi and Zidek [3]). In the next section we will examine the randomized solution for two Bayesians and a result that will be needed is, in general, the derivative

$$\frac{dU_j(\hat{\theta}_\lambda)}{dU_i(\hat{\theta}_\lambda)} = -\lambda_i \frac{U_j(\hat{\theta}_\lambda)}{U_i(\hat{\theta}_\lambda)}. \tag{1.7}$$

2. A general randomized solution for bi-Bayes estimates of the multinormal mean

We will denote $U_i(\hat{\theta}_\lambda)$ by $U_i(\lambda)$ in this section. We assume that the set S of all feasible 2-tuples is not convex and that $(U_1(\lambda_1), U_2(\lambda_1)) \in \mathscr{B}$ and $(U_1(\lambda_2), U_2(\lambda_2)) \in \mathscr{B}$ are any two points that belong to $\mathscr{B}$. Defining Bayesian i's expected utility as $\delta U_i(\lambda_1)+(1-\delta)U_i(\lambda_2)$, $i=1, 2$, if the decision rule is δ, it means that randomization is permitted and any point on the line segment joining $(U_1(\lambda_1), U_2(\lambda_1))$ and $(U_1(\lambda_2), U_2(\lambda_2))$ is feasible. Maximizing the Kalai-Nash product

$$P(\delta)=\prod_{i=1}^{2} [\delta U_i(\lambda_1)+(1-\delta)U_i(\lambda_2)]^{\alpha_i}$$

with respect to δ then gives

$$\delta=\frac{\alpha_1 U_2(\lambda_2)[U_1(\lambda_1)-U_1(\lambda_2)]+\alpha_2 U_1(\lambda_2)[U_2(\lambda_1)-U_2(\lambda_2)]}{[U_2(\lambda_2)-U_2(\lambda_1)][U_1(\lambda_1)-U_1(\lambda_2)]} . \tag{2.1}$$

The problem is to find the two points

$$(U_1(\lambda_1), U_2(\lambda_1)) \quad \text{and} \quad (U_1(\lambda_2), U_2(\lambda_2))$$

and therefore the optimal randomized rule δ^* such that

$$P(\delta^*)=\sup\{P(\delta),\ \delta \text{ randomized}\} .$$

It is clear that these two points must be the tangent points of the tangent line

$$U_2(\lambda)=g(\lambda)U_1(\lambda)+c(\lambda)$$

to $\mathscr{B}$; where

$$g(\lambda)=\frac{dU_2(\lambda)}{dU_1(\lambda)} \tag{2.2}$$

$$=-\lambda\frac{U_2(\lambda)}{U_1(\lambda)} \qquad \text{(from 1.6)}$$

and

$$c(\lambda)=U_2(\lambda)-g(\lambda)U_1(\lambda) . \tag{2.3}$$

Since there must be two λ's, say λ_1 and λ_2 for which $g(\lambda_1)=g(\lambda_2)$ and $c(\lambda_1)=c(\lambda_2)$, a way to find these λ's will be to plot $g(\lambda)$ and $c(\lambda)$ for $\lambda \geqq 0$ and the point of intersection of the graph will give the two λ's. If S is convex there cannot be two points in $\mathscr{B}$ which yield the same tangent line (see Figure 1), and the graph of $g(\lambda)$ against $c(\lambda)$ for $\lambda \geqq 0$

does not intersect, which means that randomization is not required. The procedure will be illustrated by an example.

Example 2.1. Let

$$\theta_1=\begin{bmatrix}0\\0\end{bmatrix},\quad \theta_2=\begin{bmatrix}1\\1.5\end{bmatrix},\qquad \Lambda_1^{-1}=2I_2\,,\quad \Lambda_2^{-1}=I_2\,.$$

Then $\mathscr{B}$ looks as follows:

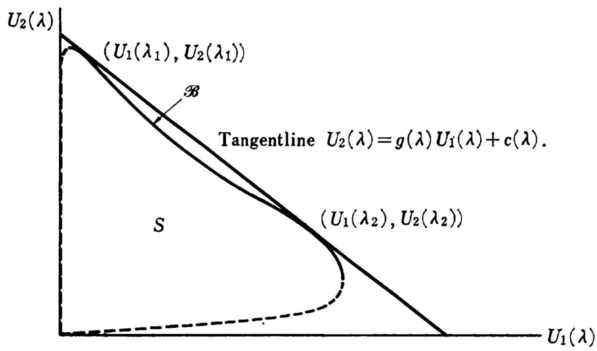

Fig. 1. $\mathscr{B}=\{(U_1(\lambda), U_2(\lambda)):\ \hat{\theta}_\lambda=\Lambda^{-1}(\Lambda_2^{-1}\theta_2+\lambda\Lambda_1^{-1}\theta_1),\ \lambda\geqq 0\}$
$S=\{(U_1(\hat{\theta}), U_2(\hat{\theta})),\ \hat{\theta}\in R^2\}$.

In Figure 2 $g(\lambda)$ and $c(\lambda)$ are plotted for various values of λ.

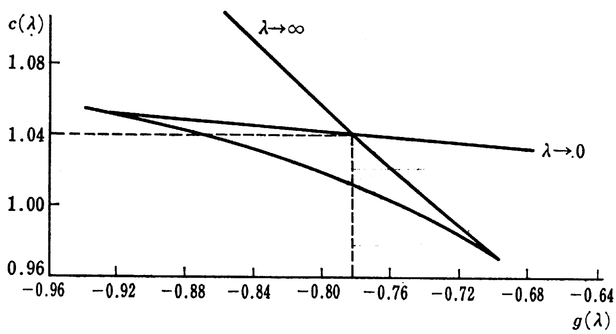

Fig. 2. Plot of $g(\lambda)$ and $c(\lambda)$ for $\lambda\geqq 0$.

From Figure 2 we obtain the approximate values of c and g at the point of intersection and from that determine the original values of λ, say λ_1 and λ_2, that yielded those c and g values. By additional calculations and linear interpolation more accurate values are obtained. These are given in the following table:

Table 1.

λ	$g(\lambda)$	$c(\lambda)$	$U_1(\lambda)$	$U_2(\lambda)$
$\lambda_1=1.8416$	-0.782	1.04	0.8623	0.3660
$\lambda_2=0.0593$	-0.782	1.04	0.0745	0.9819

From Table 1 it follows that

$$\hat{\theta}_{\lambda_1}=\begin{bmatrix}0.2135\\0.3203\end{bmatrix}, \qquad \hat{\theta}_{\lambda_2}=\begin{bmatrix}0.8939\\1.3409\end{bmatrix}.$$

The two Bayesians have to randomize between these two estimates by solving for δ^* from (2.1). It they agree on $\alpha_1=\alpha_2$, it follows that

$$\delta^*=0.7498.$$

Hence the optimal randomized rule is to estimate θ by $\hat{\theta}_{\lambda_1}$ with probability $\delta^*=0.7498$ or by $\hat{\theta}_{\lambda_2}$ with probability $1-\delta^*=0.2502$. The value of the Kalai-Nash product at δ^* is

$$P(\delta^*)=0.5882.$$

If we compare this with the Kalai-Nash product at the optimum nonrandomized estimate

$$\hat{\theta}^*=\begin{pmatrix}1/3\\1/2\end{pmatrix}, \qquad \text{say, calculated from (1.3) we get } P(\hat{\theta}^*)=0.5818,$$

which is smaller than $P(\delta^*)$. The expected utilities for the two Bayessians corresponding to the optimum nonrandomized estimate $\hat{\theta}^*$, are $U_1(\hat{\theta}^*)=0.6969$ and $U_2(\hat{\theta}^*)=0.4857$ where for δ^* they are from Table 1:

$$U_1(\delta^*)=0.8623(0.7498)+0.0745(0.2502)=0.6652$$

$$U_2(\delta^*)=0.3660(0.7498)+0.9819(0.2502)=0.5201.$$

Bayesian 1 is on the average doing better by not randomizing but to choose $\hat{\theta}^*$ where Bayesian 2 will gain by randomizing. The Kalai-Nash of course, is a maximum if the randomized rule is chosen.

UNIVERSITY OF THE ORANGE FREE STATE
UNIVERSITY OF THE ORANGE FREE STATE
UNIVERSITY OF THE ORANGE FREE STATE
BRITISH COLUMBIA

REFERENCES

[1] Nash, J. P. (1950). The bargaining problem, *Econometrica*, 18, 155-162.

[2] Weerahandi, S. and Zidek, J. V. (1981). Multi-Bayesian statistical decision theory, *Roy. Statist. Soc.*, A, **144**, 85-93.

[3] Weerahandi, S. and Zidek, J. V. (1983). Elements of Multi-Bayesian decision theory, *Ann. Statist.*, **11**, 1032-1046.