BAYESIAN LOOK AHEAD ONE STAGE SAMPLING ALLOCATIONS FOR SELECTING THE LARGEST NORMAL MEAN*

by

Shanti S. Gupta and Klaus J. Miescke
Purdue University University of Illinois at Chicago

Technical Report #91-60C

Department of Statistics
Purdue University

October 1991
Revised March 1992
Revised January 1993

*This research was supported in part by the NSF Grant DMS-8923071 at Purdue University.

◇

BAYESIAN LOOK AHEAD ONE STAGE SAMPLING ALLOCATIONS FOR SELECTING THE LARGEST NORMAL MEAN*

by

Shanti S. Gupta and Klaus J. Miescke
Purdue University
West Lafayette, IN

University of Illinois at Chicago
Chicago, IL

ABSTRACT

¿From two independent normal populations with unknown means and a common known variance, samples of unequal sizes are observed at stage 1. The goal is to find that population with the larger mean. Using the Bayes approach, optimum allocations of m additional observations, at stage 2, are derived under the linear and the 0–1 loss.

AMS Subject Classification: Primary 62F07; secondary 62F15.

Key Words: Bayesian look ahead procedure; selection and ranking; normal populations.

*This research was supported in part by the NSF Grant DMS-8923071 at Purdue University

1. Introduction

Let $\mathcal{P}_1, \ldots, \mathcal{P}_k$ be $k \geq 2$ given normal populations with unknown means $\theta_1, \ldots, \theta_k \in \mathbb{R}$ and a common known variance $\sigma^2 > 0$. Suppose we want to find the population with the largest mean using a Bayes selection rule which is based on a known prior density $\pi(\boldsymbol{\theta})$ and a given loss function $L(\boldsymbol{\theta}, i)$, $\boldsymbol{\theta} \in \mathbb{R}^k$, $i \in \{1, 2, \ldots, k\}$. Assume that k independent samples of sizes $n_1, \ldots, n_k$, respectively, have been observed already at a first stage, and let m more observations be allowed to be taken at a future second stage. The problem considered is how to allocate these m observations in an optimum way among the k populations, given the information gathered so far. It should be pointed out that the special case of $n_1 = \ldots = n_k = 0$ represents the analogous problem of how to allocate m observations at a first stage.

Looking ahead one stage using the expected posterior Bayes risk, given all observations collected so far, does not only lead the way to an optimum allocation of observations in the future. It also provides a relative measure of how much better the decision can be expected to be after further sampling has been performed following this optimum allocation. In many empirical studies in marketing reseach (e.g. direct marketing), medical research (e.g. clinical trials) and social research (e.g. survey sampling), there are interim analyses performed at certain stages to decide if sampling should be continued, and if so, how to allocate observations.

Under the assumption of k independent normal priors and either a linear loss or a 0-1-loss, a solution to the problem has been obtained for the case of $k = 2$, which turns out to be already rather involved. It allocates observations in such a way that the posterior gets as close as possible to being decreasing in transposition (DT). Moreover, somewhat surprising, it does not depend at all on the observations gathered at the first stage. This fact implies that one can allocate in an optimum way one new observation at a time, until all m have been drawn, thereby arriving at the same allocation as in the former approach.

Selecting the population with the largest (overall) sample mean is usually called the natural selection rule, since it is the uniformly best permutation invariant selection procedure in the frequentist sense for a general class of loss functions, if the sample sizes are equal. However, for unequal sample sizes, the natural selection rule loses much of its quality, although it still remains intuitively appealing. Therefore, optimum sample size allocations for this rule have been considered in the frequentist approach by Bechhofer (1969), Dudewicz and Dalal (1975), and Bechhofer, Hayter, and Tamhane (1991). On the other hand, Bayes rules with normal priors turn out to have complicated forms that cannot be represented explicitly, except for those situations where the posterior is (DT). This has been shown in Gupta and Miescke (1988), where it was recommended to plan an experiment's sampling allocation in such a way as to make the posterior (DT). The present result adds now substantial support to this earlier recommendation, although thus far, however, only for the case of two populations.

2. General Framework and Notation

After a standard reduction of the data by sufficiency, the model assumptions can be summarized as follows: At $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k) \in \mathbb{R}^k$, $X_i \sim N(\theta_i, p_i^{-1})$ with $p_i^{-1} = \sigma^2/n_i$, and $Y_i \sim N(\theta_i, q_i^{-1})$ with $q_i^{-1} = \sigma^2/m_i$, are the sample means from the samples of population $\mathcal{P}_i$ at stage 1 and stage 2, respectively, $i = 1, \ldots, k$, which are altogether independent. A priori, the parameters $\boldsymbol{\Theta} = (\Theta_1, \ldots, \Theta_k)$ are random and follow a given prior distribution which will be later assumed to have $\Theta_i \sim N(\mu_i, \nu_i^{-1})$, $i = 1, \ldots, k$, independent. Let the loss function for selections at stage 1 and stage 2 be denoted by $L(\boldsymbol{\theta}, s)$, $\boldsymbol{\theta} \in \mathbb{R}^k$, $s = 1, \ldots, k$, which will be later assumed to be either linear or of the 0-1-type.

After $\boldsymbol{X} = \boldsymbol{x} \in \mathbb{R}^k$ has been observed at stage 1, every Bayes selection rule $d_1^*(\boldsymbol{x})$ can be found through

$$E\{L(\boldsymbol{\Theta}, d_1^*(\boldsymbol{x}))|\boldsymbol{X} = \boldsymbol{x}\} = \min_{i=1,\ldots,k} E\{L(\boldsymbol{\Theta}, i)|\boldsymbol{X} = \boldsymbol{x}\}. \tag{1}$$

Likewise, after $\boldsymbol{Y} = \boldsymbol{y} \in \mathbb{R}^k$ has been observed additionally at stage 2, every Bayes rule satisfies

$$E\{L(\boldsymbol{\Theta}, d_2^*(\boldsymbol{x}, \boldsymbol{y}))|\boldsymbol{X} = \boldsymbol{x}, \boldsymbol{Y} = \boldsymbol{y}\} = \min_{i=1,\ldots,k} E\{L(\boldsymbol{\Theta}, i)|\boldsymbol{X} = \boldsymbol{x}, \boldsymbol{Y} = \boldsymbol{y}\}. \tag{2}$$

There will be no need to consider randomized Bayes selection rules in the following since minimaxity and invariance concepts will not be used.

Many results for Bayes selection rules can be found in the literature. An overview is provided by Gupta and Panchapakesan (1979, 1991). Only recently, however, attention has been given also to non-symmetric models. The binomial case has been treated by Berger and Deely (1988) and by Abughalous and Miescke (1989), whereas the normal case is studied and discussed in Gupta and Miescke (1988). Rather than studying the properties of the Bayes selection rules d_1^* and d_2^* in details, let us assume that they have been derived already, are ready to be used, and all that is needed is to allocate sample sizes in an optimum manner.

Before entering each of the two stages, similar allocation problems arise which are closely related. Before entering stage 1, by looking ahead one stage, we would like to minimize the expected posterior risk subject to $n_1 + \ldots + n_k = n$, where n is the total number of observations allowed to be taken at stage 1. This leads to the criterion for $n_1, \ldots, n_k$

$$\min_{\substack{n_1,\ldots,n_k \\ n_1+\ldots+n_k=n}} E(\min_{i=1,\ldots,k} E\{L(\Theta, i)|\boldsymbol{X}\}). \tag{3}$$

Likewise, at the end of stage 1, the criterion for $m_1, \ldots, m_k$ with a total number of m observations allowed at stage 2 is the following

$$\min_{\substack{m_1,\ldots,m_k \\ m_1+\ldots+m_k=m}} E\{\min_{i=1,\ldots,k} E\{L(\Theta, i)|\boldsymbol{X} = \boldsymbol{x}, \boldsymbol{Y}\}|\boldsymbol{X} = \boldsymbol{x}\}. \tag{4}$$

In the next two sections, solutions to criterion (4) will be obtained under linear and 0-1-loss for $k = 2$ populations assuming independent normal priors for $\Theta_1, \ldots, \Theta_k$. Formally, or using a standard sequential updating argument as described in Berger (1985) p. 445, one can get solutions of (3) from those of (4) by setting $n_1 = \ldots = n_k = 0$ and then relabel m_i by n_i, $i = 1, \ldots, k$, and m by n. Thus, we need to consider only criterion (4) in the sequel.

To simplify notation, let $\alpha_i = p_i + \nu_i$ and $\mu_i(\boldsymbol{x}) = (\nu_i\mu_i + p_i x_i)/(\nu_i + p_i)$, $i = 1, \ldots, k$. Then, under the assumption of independent normal priors, i.e. under $\Theta_i \sim N(\mu_i, \nu_i^{-1})$, $i = 1, \ldots, k$, independent, the following conditional distributions will be relevant for solving criterion (4).

$$\begin{gathered}\text{Given } \boldsymbol{X} = \boldsymbol{x} \text{ and } \boldsymbol{Y} = \boldsymbol{y}, \ \Theta_1, \ldots, \Theta_k \text{ are independent with} \\ \Theta_i \sim N\left(\frac{\alpha_i\mu_i(\boldsymbol{x}) + q_i y_i}{\alpha_i + q_i}, \frac{1}{\alpha_i + q_i}\right), \ i = 1, \ldots, k.\end{gathered} \tag{5}$$

By setting $q_1 = \ldots = q_k = 0$ in (5), one gets also the conditional distribution of $\boldsymbol{\Theta}$, given $\boldsymbol{X} = \boldsymbol{x}$. Moreover,

$$\begin{aligned}&\text{Given } \boldsymbol{X} = \boldsymbol{x},\ Y_1, \ldots, Y_k \text{ are independent with} \\ &Y_i \sim N\left(\mu_i(\boldsymbol{x}), \frac{\alpha_i + q_i}{\alpha_i q_i}\right),\ i = 1, \ldots, k.\end{aligned} \tag{6}$$

The posterior distribution of $\Theta_1, \ldots, \Theta_k$ at stage 2, as given by (5), is (DT) if and only if

$$\alpha_1 + q_1 = \alpha_2 + q_2 = \ldots = \alpha_k + q_k. \tag{7}$$

It was recommended by Gupta and Miescke (1988) to plan an experiment in such a way that (7) is satisfied, because of two reasons. First, the Bayes rule is then of a very simple form, and second, the posterior information about the k unknown parameters is then equally and fairly balanced. The solutions of criterion (4) for $k = 2$ under linear and 0-1-loss, which are derived in the next two sections, turn out to be the same: Choose m_1 and m_2 subject to $m_1 + m_2 = m$ in such a way that one gets as close as possible to the (DT) configuration (7). Since this common solution does not depend on the observations $\boldsymbol{X} = \boldsymbol{x}$ at stage 1, the solution has been found also to the analogous sequential problem where one observation at a time has to be allocated. Using the latter m times leads to the same result as performing the former in one single step. Finally, one can see that this solution applies to the open sequential setting where m is not finite and stopping rules are employed. Whether these results can be extended to the case of $k \geq 3$ populations is an interesting question which has to be investigated in the future.

3. Linear Loss

In this section we assume that the loss in linear, i.e. $L(\boldsymbol{\theta}, i) = \theta_{[k]} - \theta_i$, $i = 1, \ldots, k$, $\boldsymbol{\theta} \in \mathbb{R}^k$, where $\theta_{[k]} = \max\{\theta_1, \ldots, \theta_k\}$. Criterion (4) reduces then to minimize as a function of $m_1, \ldots, m_k$, subject to $m_1 + \ldots + m_k = m$, the look ahead expected posterior risk

$$E\{\Theta_{[k]} | \boldsymbol{X} = \boldsymbol{x}\} - E\{\max_{i=1,\ldots,k} E\{\Theta_i | \boldsymbol{X} = \boldsymbol{x}, \boldsymbol{Y}\} | \boldsymbol{X} = \boldsymbol{x}\}. \tag{8}$$

For $k \geq 2$ and the independent normal priors introduced in section 2, this reduces further by using first (5) and then (6) to maximize as a function of $m_1, \ldots, m_k$, subject to $m_1 +$

$\ldots + m_k = m$, the following

$$E\{\max_{i=1,\ldots,k} \frac{\alpha_i \mu_i(\boldsymbol{x}) + q_i Y_i}{\alpha_i + q_i} | \boldsymbol{X} = \boldsymbol{x}\} = E\left(\max_{i=1,\ldots,k} \left[\mu_i(\boldsymbol{x}) + \left(\frac{q_i}{\alpha_i(\alpha_i + q_i)}\right)^{\frac{1}{2}} N_i\right]\right), \quad (9)$$

where $N_1, \ldots, N_k$ are independent standard normal generic random variables.

For the case of $k = 2$ populations, using the identity $\max\{v, w\} = (v+w)/2 + |v-w|/2$, (9) has the following simple representation

$$\frac{1}{2}(\mu_1(\boldsymbol{x}) + \mu_2(\boldsymbol{x})) + \frac{1}{2}E(|\mu_1(\boldsymbol{x}) - \mu_2(\boldsymbol{x}) + \gamma N|), \quad (10)$$
$$\text{where } \gamma^2 = \frac{q_1}{\alpha_1(\alpha_1 + q_1)} + \frac{q_2}{\alpha_2(\alpha_2 + q_2)}, \text{ and } N \sim N(0, 1).$$

Finally, one can see that the function $E(|\mu + \sigma N|) = 2\mu[\Phi(\mu/\sigma) + (\mu/\sigma)^{-1}\varphi(\mu/\sigma) - 0.5]$ is increasing $\sigma > 0$ for every fixed $\mu \in \mathbb{R}$. As usual, Φ and φ denote the c.d.f. and density of $N(0, 1)$, respectively. Therefore, to maximize (10), subject to the given side condition, one has to maximize γ^2 as a function of m_1 and m_2, subject to $m_1 + m_2 = m$. Substituting for q_2 the condition $q_2 = (m/\sigma^2) - q_1$ makes γ^2 a function of q_1 alone, which can be seen easily to be increasing (decreasing) whenever $\alpha_1 + q_1 < (>) \alpha_2 + q_2$.

To summarize, the following result has been proved.

Theorem 1. *Under linear loss, the optimum look ahead m observations Bayes allocation rule for selecting the larger mean of two normal populations with a common known variance $\sigma^2 > 0$ is to minimize the absolute difference of $\alpha_1 + q_1$ and $\alpha_2 + q_2$, i.e. to get the posterior as close as possible to the (DT) configuration.*

To conclude this section, let us look at the more general situation where the two populations have, rather than a common known variance $\sigma^2 > 0$, known positive variances σ_1^2 and σ_2^2, say. In this situation, the optimum allocation of m observations is to minimize the absolute difference of $\sigma_1(\alpha_1 + q_1)$ and $\sigma_2(\alpha_2 + q_2)$. However, under the side condition of $q_1 + q_2 = q$, where q is fixed, the optimum allocation turns out to be the same as in the theorem. Thus, Theorem 1 has to be considered as a result on optimum allocation of sampling information rather than that of sample sizes.

4. 0-1-Loss

In this section we assume that the loss is of the 0-1 type, i.e. $L(\boldsymbol{\theta}, i)$ is equal to zero if $\theta_i = \theta_{[k]}$, and equal to one otherwise. Criterion (4) reduces then to minimize, as a function of $m_1, \ldots, m_k$, subject to $m_1 + \ldots + m_k = m$, the look ahead expected posterior risk

$$1 - E\{\max_{i=1,\ldots,k} P\{\Theta_i = \Theta_{[k]} | \boldsymbol{X} = \boldsymbol{x}, \boldsymbol{Y}\} | \boldsymbol{X} = \boldsymbol{x}\}. \tag{11}$$

For $k \geq 2$ and the independent normal priors introduced in section 2, the expectation expression in (11) can be represented in the following form, which can be derived by using (5) and (6), and which has to be maximized as a function of $m_1, \ldots, m_k$, subject to $m_1 + \ldots + m_k = m$.

$$E\left(\max_{i=1,\ldots,k} \int_{\mathbb{R}} \prod_{j \neq i} \Phi\left(\sqrt{\alpha_j + q_j}\left[\frac{z}{\sqrt{\alpha_i + q_i}} + \mu_i(\boldsymbol{x}) - \mu_j(\boldsymbol{x}) + \gamma_i N_i - \gamma_j N_j\right]\right) \varphi(z) dz\right), \tag{12}$$

where $\gamma_t^2 = q_t/(\alpha_t(\alpha_t + q_t))$, $t = 1, \ldots, k$, and where $N_1, \ldots, N_k$ are independent standard normal generic random variables.

For the case of $k = 2$ populations, using the identity $\max\{\Phi(v - w), \Phi(w - v)\} = \Phi(|v - w|)$, (12) has the following simple representation

$$\begin{aligned} &E(\Phi(|a + bN|)), \quad \text{where} \\ &a = \left[\frac{(\alpha_1 + q_1)(\alpha_2 + q_2)}{\alpha_1 + q_1 + \alpha_2 + q_2}\right]^{\frac{1}{2}} |\mu_1(\boldsymbol{x}) - \mu_2(\boldsymbol{x})|, \quad \text{and} \\ &b = \left[\frac{q_1\alpha_2(\alpha_2 + q_2) + q_2\alpha_1(\alpha_1 + q_1)}{\alpha_1\alpha_2(\alpha_1 + q_1 + \alpha_2 + q_2)}\right]^{\frac{1}{2}}. \end{aligned} \tag{13}$$

It has been shown below of (10) that the function $E(|a + bN|)$, where $N \sim N(0,1)$, is increasing in $|b|$ for every fixed $a \in \mathbb{R}$. An analogous result does not hold for the function $E(\Phi(|a + bN|))$. An investigation in this direction shows, on the contrary, that $E(\Phi(|a + bN|))$ is increasing (decreasing) in $|b|$ for small (large) values of $|a|$. Therefore, another more involved approach has to be taken to derive the optimum allocation.

At this point, it is convenient to express a and b in (13) as functions of, say, $\omega = \frac{1}{2}(q + \alpha_2 - \alpha_1) - q_1$, where $q = m/\sigma^2$, by incorporating the side condition $q_1 + q_2 = q$. This leads to

$$a(\omega) = \frac{1}{2}h(\omega)\Delta, \; b(\omega) = \left[\frac{1}{\rho}h^2(\omega) - 1\right]^{\frac{1}{2}}, \tag{14}$$

$$\text{where } \Delta = |\mu_1(\boldsymbol{x}) - \mu_2(\boldsymbol{x})|, \rho = 4\alpha_1\alpha_2/(\alpha_1+\alpha_2), \text{ and}$$

$$h(\omega) = \left[((q+\alpha_1+\alpha_2)^2 - 4\omega^2)/(q+\alpha_1+\alpha_2)\right]^{\frac{1}{2}},$$

$$(-q+\alpha_2-\alpha_2)/2 \le \omega \le (q+\alpha_2-\alpha_1)/2.$$

The crucial tool for finding the optimum allocation turns out to be the following derivative

$$\begin{aligned} &\frac{d}{d\omega}\int_{\mathbf{R}} \Phi(|a(\omega)+b(\omega)z|)\varphi(z)dz \\ &= \int_{\mathbf{R}} \text{sign } (a(\omega)+b(\omega)z)(a'(\omega)+b'(\omega)z) \\ &\quad \times \varphi(a(\omega)+b(\omega)z)\varphi(z)dz. \end{aligned} \tag{15}$$

After some further steps of standard calculations one can see that (15) has the following representation, where for convenience the argument ω is suppressed in the functions a, b and their respective derivatives a', b'.

$$\frac{1}{\sqrt{1+b^2}}\varphi\left(\frac{a}{\sqrt{1+b^2}}\right)\int_{\mathbf{R}} \text{sign } \left(y + \frac{a}{b\sqrt{1+b^2}}\right)\left(\frac{b'}{\sqrt{1+b^2}}y + a' - \frac{abb'}{1+b^2}\right)\varphi(y)dy. \tag{16}$$

This holds for any differentiable $a(\omega)$ and $b(\omega)$. For the particular forms of a and b as functions of h, as given in (14), further simplification is attained by the fact that $a'(1+b^2) \equiv abb'$, and by expressing a, b, a' and b' by h and h' using (14). This leads to the following form of (16)

$$\frac{\sqrt{\rho}h'}{h\sqrt{h^2-\rho}}\varphi\left(\sqrt{\rho}\frac{\Delta}{2}\right)\int_{\mathbf{R}} \text{sign } \left(y + \frac{\rho}{\sqrt{h^2-\rho}}\frac{\Delta}{2}\right)y\varphi(y)dy. \tag{17}$$

Evaluation of the integral is straightforward, and the final result can be summarized in the following

Lemma. $\frac{d}{d\omega}\int_{\mathbf{R}} \Phi(|a(\omega)+b(\omega)z|)\varphi(z)dz$

$$= \frac{2\sqrt{\rho}h'(\omega)}{h(\omega)\sqrt{h^2(\omega)-\rho}}\varphi\left(\frac{\sqrt{\rho}h(\omega)}{\sqrt{h^2(\omega)-\rho}}\frac{\Delta}{2}\right).$$

The sign of the derivative in the lemma coincides with the sign of $h'(\omega) = -4\omega/(h(\omega)(q + \alpha_1 + \alpha_2))$, and thus will the sign of ω itself. This shows that $E(\Phi(|a(\omega) + b(\omega)N|))$, with $N \sim N(0,1)$, is maximized by letting ω be as close as possible to zero, i.e. by letting q_1 as close as possible to $(q + \alpha_2 - \alpha_1)/2$. The optimum allocation is thus seen to be the same as the one for linear loss given in Theorem 1. To summarize, the following result has been shown.

Theorem 2. *Under 0-1-loss, the optimum look ahead m observations Bayes allocation rule for selecting the larger mean of two normal populations with a common known variance $\sigma^2 > 0$ is to minimize the absolute difference of $\alpha_1 + q_1$ and $\alpha_2 + q_2$, i.e. to get the posterior as close as possible to the (DT) configuration.*

As it has been done at the end of section 3, let us look at the more general situation where the two populations have, rather than a common known variance $\sigma^2 > 0$, known positive variances σ_1^2 and σ_2^2, say. All results in this section up to and including the lemma can be extended to this more general situation by letting $\omega = m_1$, $q_1(\omega) = \omega/\sigma_1^2$, $q_2(\omega) = (m - \omega)/\sigma_2^2$, and

$$h(\omega) = \frac{4(\alpha_1 + q_1(\omega))(\alpha_2 + q_2(\omega))}{\alpha_1 + q_1(\omega) + \alpha_2 + q_2(\omega)}, \quad 0 \le \omega \le m. \tag{18}$$

The optimum allocation of m observations is, as before under linear loss, to minimize the absolute difference of $\sigma_1(\alpha_1 + q_1)$ and $\sigma_2(\alpha_2 + q_2)$. However, under the side condition of $q_1 + q_2 = q$, where q is fixed, the optimum allocation turns out to be the same as in the theorem. Thus, also Theorem 2 has to be considered as a result on optimum allocation of sampling information rather than that of sample sizes.

References

[1] Abughalous, M. M. and Miescke, K. J. (1989). On selecting the largest success probability under unequal sample sizes. *Journal of Statistical Planning and Inference*, **21**, 53–68.

[2] Bechhofer, R. E. (1969). Optimal allocation of observations when comparing several treatments with a control. *Multivariate Analysis II*, P. R. Krishnaiah ed., Academic Press, New York, 465-473.

[3] Bechhofer, R. E., Hayter, A. J. and Tamhane, A. C. (1991). Designing experiments for selecting the largest normal mean when the varianves are known and unequal: Optimal sample size allocation. *Journal of Statistical Planning and Inference*, **28**, 271–289.

[4] Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis.* Second Edition. Springer Verlag, New York.

[5] Berger, J. O. and Deely, J. (1988). A Bayesian approach to ranking and selection of related means with alternatives to AOV methodology. *Journal of the American Statistical Association*, **83**, 364–373.

[6] Dudewicz, E.J.and Dalal, S.R. (1975). Allocation of observations in ranking and selection with unequal variances. *Sankhya*, **B-37**, 28–78.

[7] Gupta, S. S. and Miescke, K. J. (1988). On the problem of finding the largest normal mean under heteroscedasticity. In: *Statistical Decision Theory and Related Topics IV*, S. S. Gupta and J. O. Berger eds., Springer Verlag, New York, Vol. 2, 37–49.

[8] Gupta, S. S. and Panchapakesan, S. (1991). Sequential ranking and selection procedures. In: *Handbook of Sequential Analysis*, B. K. Ghosh and P. K. Sen eds., M. Dekker, New York, 363–379.

[9] Gupta, S. S. and Panchapakesan, S. (1979). *Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations.* J. Wiley, New York.

Prof. Shanti S. Gupta
Department of Statistics
Purdue University
West Lafayette, IN 47907, U.S.A.

Prof. Klaus J. Miescke
Department of Mathematics, Statistics, and Computer Science
University of Illinois at Chicago
P.O. Box 4348
Chicago, IL 60680, U.S.A.