A Festschrift for Herman Rubin
Institute of Mathematical Statistics
Lecture Notes – Monograph Series
Vol. 45 (2004) 228–236

An asymptotic minimax determination of the initial sample size in a two-stage sequential procedure

Michael Woodroofe[1]

University of Michigan

Abstract: When estimating the mean of a normal distribution with squared error loss and a cost for each observation, the optimal (fixed) sample size depends on the variance σ^2. A two-stage sequential procedure is to first conduct a pilot study from which σ^2 can be estimated, and then estimate the desired sample size. Here an asymptotic formula for the initial sample size in a two-stage sequential estimation procedure is derived–asymptotic as the cost of a single observation becomes small compared to the loss from estimation error. The experimenter must specify only the sample size, n_0 say, that would be used in a fixed sample size experiment; the initial sample size of the two-stage procedure is then the least integer greater than or equal to $\sqrt{n_0/2}$. The resulting procedure is shown to minimize the maximum Bayes regret, where the maximum is taken over prior distributions that are consistent with the initial guess n_0; and the minimax solution is shown to provide an asymptotic lower bound for optimal Bayesian choices for a wide class of prior distributions.

1. Introduction

It is indeed a pleasure to offer this tribute to Herman Rubin and to ponder his influence on my own work. I still remember the interest with which I read the papers on Bayes' risk efficiency [7] and [8] early in my career. From reading these papers, I gained an appreciation for the power of statistical decision theory and its interplay with asymptotic calculations that go beyond limiting distributions. These involved moderate deviations in the case of [7]. A central idea in [8] is the study of a risk functional, the integrated risk of a procedure with respect to a prior distribution that can vary over a large class. I have used this idea in a modified form in work on sequential point estimation and very weak expansions for sequential confidence intervals—[12, 13, 14], and the references given there. This idea is also present in Theorem 2 below. The connection between [12] and Bayes risk efficiency is notable here. The following is proved in [12], though not isolated: Suppose that it is required to estimate the mean of an exponential family with squared error loss and a cost for each observation and that the population mean is to be estimated by the sample mean. Then there is a stopping time which is Bayes risk non-deficient in the sense of [4]; that is, it minimizes a Bayesian regret asymptotically, simultaneously for all sufficiently smooth prior distributions.

The present effort combines tools from decision theory and asymptotic analysis to obtain a rule for prescribing the initial sample size in a two-stage sequential procedure for estimating the mean of a normal distribution. Unlike the fully sequential, or even three-stage, versions of the problem, Bayes risk non-deficiency is

[1]Statistics Department, University of Michigan, Ann Arbor, Michigan 48109, USA. e-mail: michaelw@stat.lsa.umich.edu

Keywords and phrases: Bayesian solutions, integrated risk and regret, inverted gamma priors, sequential point estimation, squared error loss.

AMS 2000 subject classifications: 62L12.

not possible with two-stage procedures, and the rule is obtained from minimaxity. The problem is stated in Section 2, and the minimax solution is defined. The rule requires the statistician to specify only the fixed sample size, n_0 say, that would have been used in a fixed sample size design, or to elicit such from a client. The minimax initial sample size is then the least integer that is greater than or equal to $\sqrt{n_0/2}$. The proof of asymptotic minimaxity is provided in Section 3. As explained in Section 4, the minimax solution is very conservative but, at least, provides an asymptotic lower bound for optimal Bayesian solutions for a wide class of prior distributions.

2. The problem

Let $X_1, X_2, \ldots \overset{ind}{\sim} \text{Normal}[\mu, \sigma^2]$, where $-\infty < \mu < \infty$ and $\sigma > 0$ are unknown, and consider the problem of estimating μ with loss of the form

$$L_a(n) = a^2\big(\bar{X}_n - \mu\big)^2 + n, \tag{1}$$

where $\bar{X}_n = (X_1 + \cdots + X_n)/n$. In (1), $a^2(\bar{X}_n - \mu)^2$ represents the loss due to estimation error, and n the cost of sampling. The units are so chosen that each observation costs one unit, and a is determined by the importance of estimation error relative to the cost of sampling. Also, the estimator has been specified as $\bar{X}_n$, leaving only the sample size n to be determined. If σ were known, then the expected loss plus sampling cost, $E_{\mu,\sigma^2}[L_a(n)] = a^2\sigma^2/n + n$, would be minimized when n is an integer adjacent to

$$N = a\sigma,$$

and in many ways the problem is one of estimating N. This will be done using the sample variances

$$S_n^2 = \left(\frac{1}{n-1}\right)\sum_{i=1}^{n}\big(X_i - \bar{X}_n\big)^2$$

for $n \geq 2$. Interest in two-stage sequential procedures for estimation originated with Stein's famous paper [9]. The problem has a long history, much of which is included in Chapter 6 of [5], but there seems to be no general agreement on the choice of the initial sample size m in two-stage procedures. Some additional references are provided in the last section.

A *two-stage procedure* consists of a pair $\delta = (m, \tilde{N})$ where $m \geq 2$ is an integer and $\tilde{N} = \tilde{N}(S_m^2)$ is an integer valued random variable for which $\tilde{N} \geq m$. The estimator of μ is then $\bar{X}_{\tilde{N}}$. For example, letting $\lceil x \rceil$ denote the least integer that is at least x,

$$\hat{N}_a = \max\{m, \lceil aS_m \rceil\} \tag{2}$$

satisfies the conditions for any $m \geq 2$. The choice of m has to be subjective at some level, because there is no data available when it is chosen. Here it is required only that the experimenter specify a prior guess, u say for σ, or even just the guess $n_0 = au$ for N. This seems a very modest requirement, since a fixed sample size experiment would have to include a prior guess for N. Given the prior guess, it is shown that

$$m_a = \max\left\{2, \left\lceil\sqrt{\frac{1}{2}n_0}\right\rceil\right\} \tag{3}$$

leads to a procedure that minimizes the maximum Bayes' regret in the class of prior distributions for which σ has prior mean u.

3. The theorem

The risk of a two stage procedure $\delta = (m, \tilde{N})$ is $R_a(\delta; \sigma^2) = E_{\mu,\sigma^2}[L_a(\tilde{N})]$. Using the Helmut transformation (for example, [11, p. 106]), it is easily seen that

$$R_a\big(\delta; \sigma^2\big) = E_{\sigma^2}\left[\frac{a^2\sigma^2}{\tilde{N}} + \tilde{N}\right], \tag{4}$$

which depends on σ^2, but not on μ. The difference

$$r_a\big(\delta, \sigma^2\big) = E_{\sigma^2}\left[\frac{a^2\sigma^2}{\tilde{N}} + \tilde{N}\right] - 2N,$$

is called the regret.

If ξ is a prior distribution over $[0, \infty)$, write P_ξ and E_ξ for probability and expectation in the Bayesian model, where $\sigma^2 \sim \xi$ and $S_2^2, S_3^2, \ldots$ are jointly distributed random variables; and write P_ξ^m and E_ξ^m for conditional probability and expectation given S_m^2. Then the integrated risk of a two-stage procedure δ with respect to a ξ is

$$\bar{R}_a(\delta, \xi) = \int_0^\infty R_a\big(\delta; \sigma^2\big)\xi\big\{d\sigma^2\big\} = E_\xi\left[\frac{a^2\sigma^2}{\tilde{N}} + \tilde{N}\right]$$

possibly infinite; and if $\int_0^\infty \sigma\xi\{d\sigma^2\} < \infty$, then the integrated regret of δ with respect to ξ is

$$\bar{r}(\delta, \xi) = \int_0^\infty r_a\big(\delta; \sigma^2\big)\xi\big\{d\sigma^2\big\} = E_\xi\left[\frac{a^2\sigma^2}{\tilde{N}} + \tilde{N} - 2N\right]$$

again possibly infinite. As noted above, the experimenter must specify $E_\xi(N)$, or equivalently, $E_\xi(\sigma)$. In fact, it is sufficient for the experimenter to specify an upper bound. For a fixed $u \in (0, \infty)$, let $\Xi = \Xi_u$ be the class of prior distributions for which

$$\int_0^\infty \sigma\xi\big\{d\sigma^2\big\} \le u; \tag{5}$$

and let $\Xi^o = \Xi_u^o$ be the class of ξ for which there is equality in (5). Also, let δ^a be the procedure $(m_a, \hat{N}_a)$ defined by (2) and (3) with $n_0 = au$.

Theorem 1. *For any* $u \in (0, \infty)$.

$$\inf_\delta \sup_{\xi\in\Xi} \bar{r}(\delta; \xi) \sim \sqrt{2n_0} \sim \sup_{\xi\in\Xi} \bar{r}\big(\delta^a; \xi\big)$$

as $a \to \infty$.

Proof. The proof will consist of showing first that

$$\limsup_{a\to\infty} \sup_{\xi\in\Xi} \frac{1}{\sqrt{a}}\bar{r}\big(\delta^a; \xi\big) \le \sqrt{2u} \tag{6}$$

and then that

$$\liminf_{a\to\infty} \sup_{\xi\in\Xi^o} \inf_\delta \frac{1}{\sqrt{a}}\bar{r}(\delta; \xi) \ge \sqrt{2u} \tag{7}$$

This is sufficient, since $\inf_\delta \sup_{\xi\in\Xi} \ge \sup_{\xi\in\Xi^o} \inf_\delta$. In the proofs of (6) and (7), there is no loss of generality in supposing that $u = 1$.

The Upper Bound. From (4) and (2),

$$R_a\big(\delta^a;\sigma^2\big) \le a\sigma^2 E_{\sigma^2}\left[\frac{1}{S_{m_a}}\right] + aE_{\sigma^2}(S_{m_a}) + m_a + 1. \tag{8}$$

Here

$$E_{\sigma^2}(S_m) = C(m)\sigma, \tag{9}$$

where

$$C(m) = \frac{\Gamma(\frac{m}{2})}{\sqrt{\frac{m-1}{2}}\Gamma(\frac{m-1}{2})}$$

and Γ is the Gamma-function; and, similarly,

$$E\bigg(\frac{1}{S_m}\bigg) = \sqrt{\frac{m-1}{m-2}}\frac{1}{C(m-1)\sigma}. \tag{10}$$

A version of Stirling's Formula asserts that

$$\log\Gamma(z) = \bigg(z - \frac{1}{2}\bigg)\log(z) - z + \frac{1}{2}\log(2\pi) + \frac{1}{12z} + O\bigg(\frac{1}{z^3}\bigg)$$

as $z \to \infty$. See, for example, [1, p. 257]. It then follows from simple algebra that

$$C(m) = 1 - \frac{1}{4m} + O\bigg(\frac{1}{m^2}\bigg). \tag{11}$$

Let a be so large that $m_a \ge 3$. Then, combining (8) and (11),

$$\begin{aligned} R_a\big(\delta^a;\sigma^2\big) &\le a\sigma\left[\sqrt{\frac{m_a-1}{m_a-2}}\frac{1}{C(m_a-1)} + C(m_a)\right] + m_a + 1 \\ &= 2a\sigma + \frac{a\sigma}{2m_a} + m_a + 1 + a\sigma \times O\bigg(\frac{1}{m_a^2}\bigg), \end{aligned}$$

where $O(1/m)$ is a function only of m. So, for every $\xi \in \Xi = \Xi_1$,

$$\bar{r}_a(\delta^a;\xi) \le \frac{a}{2m_a} + m_a + 1 + a \times O\bigg(\frac{1}{m_a^2}\bigg) \le \sqrt{2a} + O(1),$$

establishing (6), since $n_0 = 2a$ when $u = 1$.

The Lower Bound. The lower bound (7) will be established by finding the Bayes procedure and a lower bound for the Bayes regret

$$\bar{r}_a(\xi) = \inf_{\delta} \bar{r}_a(\delta;\xi)$$

for a general prior distribution ξ and then finding priors $\xi_a \in \Xi^o$ for which $\liminf_{a\to\infty} \bar{r}_a(\xi_a)/\sqrt{a} \ge \sqrt{2}$.

Finding the Bayes procedure is not difficult. If the initial sample size is $m \ge 2$, then $\tilde{N}$ should be chosen to minimize the posterior expected loss $E_\xi^m[a^2\sigma^2/n + n]$ with respect to n. Clearly,

$$E_\xi^m\left[\frac{a^2\sigma^2}{n} + n\right] = \frac{a^2 V_m}{n} + n = 2a\sqrt{V_m} + \frac{1}{n}\big(n - a\sqrt{V_m}\big)^2 \tag{12}$$

where

$$V_m = E_\xi^m(\sigma^2)$$

So, (12) is minimized when n is the larger of m and an integer adjacent to $a\sqrt{V_m}$, leaving

$$\bar{r}_a(\xi) = \inf_{m\ge 2} E_\xi\left\{2a\sqrt{V_m} + \frac{1}{m}\big(m - a\sqrt{V_m}\,\big)_+^2 + \eta(a,m)\right\} - 2a,$$

where $(x)_+^2$ denotes the square of the positive part of x and $0 \le \eta(a,m) \le 1/m$. An alternative expression is

$$\bar{r}_a(\xi) = \inf_{m\ge 2} E_\xi\left\{2a\big[\sqrt{V_m} - U_m\big] + \frac{1}{m}\big(m - a\sqrt{V_m}\,\big)_+^2 + \eta(a,m)\right\}, \tag{13}$$

where

$$U_m = E_\xi^m(\sigma)$$

and $E_\xi(U_m) = E_\xi(\sigma) = 1$.

Suppose now that ξ is an inverted Gamma prior with density

$$\frac{1}{\Gamma(\frac{1}{2}\alpha)}\left(\frac{\beta}{2\sigma^2}\right)^{\frac{1}{2}\alpha}\exp\left[-\frac{\beta}{2\sigma^2}\right]\frac{1}{\sigma^2}, \tag{14}$$

where $\alpha > 1$ and $\beta > 0$. Equivalently $1/\sigma^2$ has a Gamma distribution with parameters $\alpha/2$ and $\beta/2$. Then

$$E(\sigma) = \frac{\Gamma(\frac{\alpha-1}{2})}{\Gamma(\frac{1}{2}\alpha)}\sqrt{\frac{\beta}{2}}. \tag{15}$$

Letting

$$W_m = (m-1)S_m^2$$

and applying (15) to the posterior distributions then leads to

$$U_m = \frac{\Gamma(\frac{\alpha+m-2}{2})}{\Gamma(\frac{\alpha+m-1}{2})}\sqrt{\frac{\beta+W_m}{2}}$$

and

$$V_m = E_\xi^m(\sigma^2) = \frac{\beta+W_m}{\alpha+m-3} = B(\alpha+m-1)^2 \times U_m^2, \tag{16}$$

where

$$B(m) = \sqrt{\frac{m-1}{m-2}}\,C(m) = 1 + \frac{1}{4m} + O\left(\frac{1}{m^2}\right). \tag{17}$$

In order for the ξ of (14) to be in $\Xi^o = \Xi_1^o$, α and β must be so constrained that the right side of (15) equals one. Then $E_\xi(U_m) = 1,\ E_\xi(\sqrt{V_m}) = B(\alpha+m-1)$, and

$$\begin{aligned}
\bar{r}_a(\xi) &= \inf_{m\ge 2} E_\xi\left\{2a\big[B(\alpha+m-1)-1\big]U_m + \frac{1}{m}\big(m - a\sqrt{V_m}\,\big)_+^2 + \eta(a,m)\right\} \\
&\ge \inf_{m\ge 2}\left\{2a\big[B(\alpha+m-1)-1\big] + (1-\epsilon)^2 m P_\xi\big[a\sqrt{V_m} \le \epsilon m\big]\right\}
\end{aligned}$$

for any $\epsilon > 0$.

Observe that $B(\alpha + m - 1)$ is positive and bounded away from 0 for $0 < \alpha \leq 1$ for each fixed $m \geq 2$. It follows that the term in braces on the right side of (13) is of order a for each fixed $m \geq 2$ when ξ is an inverted gamma prior with $0 < \alpha \leq 2$ and, therefore, that the minimizing $m = m_a$ approaches ∞ as $a \to \infty$. So, $\inf_{m \geq 2}$ in (13) can be replaced by $\inf_{m \geq m_0}$ for any m_0 for all sufficiently large a.

The marginal distribution of W_m is of the form

$$P_\xi[W_m \leq w] = \int_0^w \frac{1}{\beta} g\left(\frac{z}{\beta}\right) dz,$$

where

$$g(z) = \frac{\Gamma(\frac{\alpha+m-1}{2})}{\Gamma(\frac{\alpha}{2})\Gamma(\frac{m-1}{2})} \frac{z^{\frac{m-3}{2}}}{(1+z)^{\frac{\alpha+m-1}{2}}}.$$

Clearly,

$$\begin{aligned} \int_c^\infty g(z)dz &\leq \frac{\Gamma(\frac{\alpha+m-1}{2})}{\Gamma(\frac{\alpha}{2})\Gamma(\frac{m-1}{2})} \int_c^\infty \left(\frac{1}{1+z}\right)^{1+\frac{1}{2}\alpha} dz \\ &= \frac{2\Gamma(\frac{\alpha+m-1}{2})}{\alpha\Gamma(\frac{\alpha}{2})\Gamma(\frac{m-1}{2})} \left(\frac{1}{1+c}\right)^{\frac{1}{2}\alpha} \end{aligned}$$

for all $c > 0$. So, there is a constant K for which

$$\int_c^\infty g(z)dz \leq \frac{Km}{\sqrt{1+c}}$$

for all $1 < \alpha \leq 2$, $m \geq 2$ and $c > 0$. Let ξ_a be an inverted gamma prior with $\beta_a = o(a^{-2})$ and α_a determined by the condition that $E_{\xi_a}(\sigma) = 1$. Then $\alpha_a \to 1$ as $a \to \infty$. If $\epsilon > 0$ is given, then

$$\begin{aligned} P_{\xi_a}\left[a\sqrt{V}_m \geq \epsilon m\right] &= P_{\xi_a}\left[W_m \geq \epsilon^2 \frac{m^2(m-2)}{a^2} - \beta_a\right] \\ &\leq \frac{2Ka}{\epsilon}\sqrt{\beta_a} \leq \epsilon \end{aligned}$$

for all $m \geq 3$ and sufficiently large a. It follows that for any $m_0 \geq 3$,

$$\bar{r}(\xi_a) \geq \inf_{m \geq 2} \left\{2a\left[B(\alpha_a + m - 1) - 1\right] + (1-\epsilon)^3 m\right\}. \tag{18}$$

for all sufficiently large a. From (11) and (17) there is an m_0 for which $B(m) \geq l + (1-\epsilon)/4m$ for all $m \geq m_0 - 1$. Then

$$\bar{r}(\xi_a) \geq (1-\epsilon) \inf_{m \geq m_0} \left[\frac{a}{2m} + (1-\epsilon)^2 m\right] \geq (1-\epsilon)^2 \sqrt{2a}$$

for all sufficiently large a. Relation (7) follows since $\epsilon > 0$ was arbitrary. □

4. The minimax solution as a lower bound

The minimax solution is very conservative in that it specifies a very small initial sample size. For example, if the prior guess for the best fixed sample size is 100, then the asymptotic minimax solution calls for an initial sample size of only 8; and if the prior guess is increased to 1000, then the initial sample size increases only to 23. The asymptotic minimax solution approximates the Bayes procedure when the σ is small with high probability, but still has a fixed mean, as is clear from the nature of the inverted gamma prior that was used to obtain the lower bound. A statistician who can specify more about the prior distribution will take a larger initial sample size for large a and incur a smaller regret. For example, if $\sigma \ge \sigma_0 > 0$ with prior probability one, then the optimal initial sample size is at least $a\sigma_0$, and the Bayes regret is of order one as $a \to \infty$, assuming that σ has a finite prior mean. A more detailed study of the asymptotic properties of Bayes procedures suggests that optimal m is closely related to the behavior of the prior density near $\sigma^2 = 0$, a relationship that might be difficult to specify or elicit from a client. The inverted gamma priors (14) are an extreme case since the prior density approaches zero very rapidly as $\sigma^2 \to 0$ in this case. An advantage of the asymptotic minimax solution, of course, is that it does not require the statistician to elicit detailed prior information from a client.

The following result shows that the asymptotic minimax solution (3) provides an asymptotic lower bound for optimal Bayesian solutions for a very large class of prior distributions.

Theorem 2. *Suppose that $\xi\{0\} = 0$, that ξ has a continuously differentiable density on $(0, \infty)$, and that $\int_0^\infty \sigma^2 \xi\{d\sigma^2\} < \infty$. Let $m_a^* = m_a^*(\xi)$ be an optimal initial sample size for ξ. Then*

$$\lim_{a\to\infty} \frac{m_a^*}{\sqrt{a}} = \infty. \tag{19}$$

Proof. As above, there is no loss of generality in supposing that $\int_0^\infty \sigma\xi\{d\sigma^2\} = 1$. By (13),

$$\bar{r}_a(\xi) = \inf_{m\ge 2} \left[2ab_m + c_m(a) + \eta(a, m)\right],$$

where $b_m = E_\xi[\sqrt{V_m} - U_m]$, $c_m(a) = E_\xi[(m - a\sqrt{V_m})_+^2]/m$, and $0 \le \eta(a, m) \le 1/m$. Then

$$2a[b_{m_a^*} - b_{2m_a^*}] \le c_{2m_a^*}(a) + \frac{1}{2m_a^*},$$

since $2ab_m + c_m(a) + \eta(a, m)$ is minimized when $m = m_a^*$ and $0 < \eta(a, m) \le 1/m$. By Lemmas 1 and 2 below,

$$c_m(a) \le mP_\xi\left[\sigma \le \frac{m}{a}\right] \tag{20}$$

and

$$b_m - b_{2m} \ge \frac{\epsilon}{m} \tag{21}$$

for some $\epsilon > 0$ that does not depend on m. Combining the last three equations,

$$\frac{2a\epsilon}{m_a^*} \le 2m_a^* P_\xi\left[\sigma \le \frac{2m_a^*}{a}\right] + \frac{1}{2m_a^*}$$

and, therefore,

$$\frac{m_a^*}{\sqrt{a}} \ge \sqrt{\frac{\epsilon}{2P_\xi[\sigma \le 2m_a^*/a]}},$$

for all sufficiently large a. Relation (19) follows directly, completing the proof, except for the proofs of the lemmas. □

Lemma 1. *Relation* (20) *holds.*

Proof. Using Jensen's Inequality twice, $(m-a\sqrt{V_m})_+^2 \leq [E_\xi^m(m-a\sigma)]_+^2 \leq E_\xi^m[(m-a\sigma)_+^2]$. So,

$$c_m(a) \leq \frac{1}{m}E_\xi\left[(m-a\sigma)_+^2\right] \leq mP_\xi\left[\sigma \leq \frac{m}{a}\right],$$

as asserted. □

Lemma 2. *There is an $\epsilon > 0$ for which relation (21) holds.*

Proof. Since $E_\xi(U_m) = E_\xi(\sigma)$ for all m, $b_m - b_{2m} = E_\xi[\sqrt{V_m} - \sqrt{V_{2m}}]$. Next, since $V_m - V_{2m} = 2\sqrt{V_m}(\sqrt{V_m} - \sqrt{V_{2m}}) - (\sqrt{V_{2m}} - \sqrt{V_m})^2$ and $E_\xi^m(V_{2m} - V_m) = 0$,

$$b_m - b_{2m} = E_\xi\left[\frac{(\sqrt{V_{2m}} - \sqrt{V_m})^2}{2\sqrt{V_m}}\right].$$

From Laplace's method, for example, [6],

$$V_m = S_m^2 + O\left(\frac{1}{m}\right)$$

*w.p.*1 (P_{σ^2}) for each $\sigma^2 > 0$ and, therefore, *w.p.*1 (P_ξ). That $\sqrt{m}(\sqrt{V_{2m}} - \sqrt{V_m})$ has a non-degenerated limiting distribution follows directly, and then

$$\liminf_{m\to\infty} mE_\xi\left[\frac{(\sqrt{V_{2m}} - \sqrt{V_m})^2}{2\sqrt{V_m}}\right] > 0$$

by Fatou's Lemma. Relation (21) follows. □

5. Remarks and acknowledgments

The smoothness condition on the prior in Theorem 2 can probably be relaxed. In the proof, it was used to derive the relation $V_m - S_m^2 = O(1/m)$ *w.p.*1, and this is a smaller order of magnitude that is needed.

If ξ is an inverted gamma prior with a fixed $\alpha > 1$ and $\beta > 0$, then

$$r_a(\xi) + \frac{a}{m_a^*(\xi)} = O\left[\sqrt{\log(a)}\,\right].$$

This may be established by combining techniques from the proofs of Theorems 1 and 2.

Bayesian solutions to two-stage sequential estimation problems have been considered by several authors—notably [2, 3], and [10].

The normality assumption has been used heavily, to suggest the estimators for μ and σ^2 and also for special properties of these estimators in (4), (9) and (10). It is expected that similar results may be obtained for multiparameter exponential families and other loss structures, and such extensions are currently under investigation in the doctoral work of Joon Lee. Extensions to a non-parametric context are more speculative.

It is a pleasure to acknowledge helpful conversations with Bob Keener, Joon Sang Lee, and Anand Vidyashankar and helpful comments from Anirban Dasgupta. This research was supported by the National Science Foundation and the U.S. Army Research Office.

References

[1] Abromowitz, M. and Stegun, I. (1964). *Handbook of Mathematical Functions.* National Bureau of Standards.

[2] Cohen, A. and Sackrowitz, H. (1985). Bayes double sampling procedures, *Ann. Statist.* **12**, 1035–1049. MR751290

[3] Cohen, A. and Sackrowitz, H. (1994). Double sampling estimation when cost depends on the parameter, *Statistical Decision Theory and Related Topics* **V**, 253–266. MR1286306

[4] Hodges, J. L. Jr. and Lehmann, E. L. (1970). Deficiency, *Ann. Math. Statist.* **41**, 783–801. MR272092

[5] Malay, G., Mukhopadhyay, N. and Sen, P. K. (1997). *Sequential Estimation.* Wiley. MR1434065

[6] Johnson, R. (1970). Asymptotic expansions associated with posterior distributions, *Ann. Math. Statist.* **41**, 851–864. MR263198

[7] Rubin, H. and Sethuraman, J. (1965). Probabilities of moderate deviations, *Sankhya, Ser. A* **27**, 325–346. MR203783

[8] Rubin, H. and Sethuraman, J. (1965). Bayes risk efficiency, *Sankhya, Ser. A* 347–356. MR207112

[9] Stein, C. (1945). A two-sample test for a linear hypothesis whose power is independent of the variance, *Ann. Math. Statist.* **16**, 243–258. MR13885

[10] Treder, R. P. and Sedransk, J. (1996). Bayesian double sampling plans with normal distributions, *J. Amer. Statist. Assn* **91**, 782–790. MR1395745

[11] Woodroofe, M. (1982). *Non-linear Renewal Theory in Sequential Analysis.* S.I.A.M. MR660065

[12] Woodroofe, M. (1985). Asymptotic local minimaxity in sequential point estimation, *Ann. Statist.* **13**, 676–688. Correction: **17**, p. 452. MR790565

[13] Woodroofe, M. (1986). Very weak expansions for sequential confidence intervals, *Ann. Math. Statist.* **14**, 1049–1067. MR856805

[14] Woodroofe, M. and Coad, D. S. (1997). Corrected confidence intervals for sequentially designed experiments, *Statistica Sinica* **7**, 53–74. MR1441144