

**A RESTRICTED MINIMAX DETERMINATION OF THE
INITIAL SAMPLE SIZE IN STEIN'S AND RELATED
TWO-STAGE PROCEDURE**

JOON SANG LEE AND MICHAEL WOODROOFE

Statistics Department, The University of Michigan, USA

When the variance is known, a level $1 - \alpha$ confidence interval of specified width $2h > 0$ for the mean of a normal distribution requires a sample of size at least $\eta = c^2\sigma^2/h^2$, where c is the upper $(1 - \frac{1}{2}\alpha)^{th}$ quantile of the standard normal distribution. If the variance is unknown, then such an interval may be constructed using Stein's double sampling procedure in which an initial sample of size $m \geq 2$ is drawn and used to estimate η . Here it is shown that if the experimenter can specify a prior guess, η_0 say, for η , then $\sqrt{\frac{1}{2}(1 + c^2)\eta_0}$ is an approximately minimax choice for the initial sample size. The formulation is, in fact, more general and includes point estimation with equivariant loss as well as interval estimation.

Key words and Phrases: Decision problem; minimax solution; sequential point and interval estimation;

AMS 2000 Classification: 62L12

1. Introduction

Stein's (1945) two-stage procedure for setting a confidence interval of prescribed width $2h > 0$ and confidence $1 - \alpha$ for the mean of a normal distribution may be described as follows: Let X_1, X_2, \dots be independent normal variables with unknown mean μ and variance σ^2 ; and let \bar{X}_n and S_n^2 denote the sample mean and variance of X_1, \dots, X_n , so that $n\bar{X}_n = X_1 + \dots + X_n$ and $(n-1)S_n^2 = (X_1 - \bar{X}_n)^2 + \dots + (X_n - \bar{X}_n)^2$. First take an initial sample of size $m \geq 2$ and compute S_m^2 ; next let

$$T_m = \max\left\{m, \left\lceil \frac{c_{m-1}^2 S_m^2}{h^2} \right\rceil\right\},$$

where c_m is the upper $1-\alpha/2$ quantile of the t-distribution with m degrees of freedom and $\lceil x \rceil$ denotes the least integer that is greater than or equal to x ; then take an additional $T_m - m$ observations and declare $[\bar{X}_{T_m} - h, \bar{X}_{T_m} + h]$ to be the interval. Stein showed that $P_{\mu, \sigma^2}[|\bar{X}_{T_m} - \mu| \leq h] \geq 1 - \alpha$ for all μ and σ^2 for any $m \geq 2$. Of course, if σ^2 were known, then a confidence interval of width $2h$ and confidence $1 - \alpha$ could be constructed by taking a sample of size at least

$$\eta = \frac{c^2 \sigma^2}{h^2}, \quad (1)$$

where c is the upper $1 - \alpha/2$ quantile of the standard normal distribution. Stein's procedure effectively estimates η by T_m .

The present paper centers on the choice of m in Stein's procedure and related ones. The choice of m has to be subjective at some level, because there is no data when it is chosen. It is required here that the experimenter specify a prior expectation, σ_0^2 say, for σ^2 . That is, it is assumed that the experimenter has a prior distribution ξ for which

$$\int_0^\infty \sigma^2 \xi\{d\sigma\} = \sigma_0^2. \quad (2)$$

Equivalently, the experimenter must specify the prior expectation $\eta_0 = c^2 \sigma_0^2 / h^2$ for η . No other details of the prior are required.

Let Ξ_0 be the class of prior distributions for which (2) holds. Then a corollary to the main result asserts that

$$m_h = \lceil \sqrt{\left(\frac{1+c^2}{2}\right)\eta_0} \rceil \quad (3)$$

is an asymptotically minimax choice for m within the class Ξ_0 in the following sense:

$$\begin{aligned} \inf_{m \geq 2} \sup_{\xi \in \Xi_0} \int_0^\infty E_{\sigma^2}(T_m) \xi\{d\sigma\} &= \eta_0 + 2\sqrt{\left(\frac{1+c^2}{2}\right)\eta_0} + o(\sqrt{\eta_0}) \\ &= \sup_{\xi \in \Xi_0} \int_0^\infty E_{\sigma^2}(T_{m_h}) \xi\{d\sigma\} \end{aligned} \quad (4)$$

as $h \downarrow 0$. For example, if $\alpha = .05$, $h = .1$, and $\sigma_0 = 1$, then $\eta_0 = 384.16$ and the approximately minimax choice of m is $m_h = 31$. While only a

prior guess for σ^2 is required, the minimax choice (3) is sensitive to that choice, since m_h depends linearly on σ_0 . This is unfortunate, but in some sense unavoidable, since only prior knowledge/opinion is available when m is chosen.

There has been continuing interest in two-stage procedures for estimating means since [7], but few explicit recommendations for the initial sample size. On one hand there has been interest in asymptotic properties of two-stage procedures. Much of this work is described in [6] and does include some order of magnitude restrictions on the initial sample size. If a lower bound for σ is known, say $\sigma \geq \sigma_* > 0$, then m must be at least $m_* = c^2\sigma_*^2/h^2$ and $\lceil m_* \rceil$ is suggested as an initial sample size. This remark can be useful when $\sigma_* \gg h$. See [5] for an example. There has also been some recent interest in Bayesian solutions, [3], [4] and [8], when the prior distribution can be specified completely. A special case of our main result appears in [12], the case of point estimation with squared error loss, as described in the next section. For that case, it is shown that a statistician who can specify the prior distribution completely may be led to take a larger initial sample size than the minimax choice.

The proof of (4) requires finding an appropriate decision problem, one for which Stein's procedure is the (asymptotic minimax) solution. The decision problem is identified in Section 2. The main result is stated there, and (4) proved. Section 3 contains some preliminary lemmas. The first two of these, at least, are intuitive. The main result is then established in Section 4. The proof is substantially more involved than the special case in [12], where simple explicit expressions were used extensively.

2. The Main Result

The proof of Stein's result uses the observation: If $t = t(S_m^2) \geq m$ is an (measurable) integer valued function of S_m^2 , then the conditional distribution of \bar{X}_t given S_m^2 is normal with mean μ and variance σ^2/t . It then

4

follows that

$$P_{\mu, \sigma^2} [|\bar{X}_t - \mu| \leq h] = E_{\sigma^2} [2\Phi\left(\frac{h\sqrt{t}}{\sigma}\right) - 1], \quad (5)$$

where Φ denotes the standard normal distribution function. Using extensions of (5) to fully sequential procedures, Woodroffe [11] showed a close connection between the problem of setting a fixed width confidence interval and an optimal stopping problem in which the statistician must find a stopping time t with respect to S_2^2, S_3^2, \dots to minimize $\eta E_{\sigma^2} [K(t/\eta)]$, where

$$K(x) = \frac{2[1 - \Phi(c\sqrt{x})]}{c\varphi(c)} + x,$$

η is as in (1) with $\Phi(c) = 1 - \alpha/2$, and φ the standard normal density. This connection is elaborated in the proof of Corollary 1 at the end of this section.

A similar form arises in point estimation. To see how, suppose that it is required to estimate μ with a loss of the form $A|\bar{X}_n - \mu|^{2p}$ for estimation error, where $p > 0$, and unit cost for each observation. Here A determines the importance of estimation error relative to the cost of sampling and is assumed to be large. If $m \geq 2$, $t = t(S_m^2) \geq m$, and a sample of size t is taken then, as in (5), the expected loss plus sampling cost is

$$E[A|\bar{X}_t - \mu|^{2p} + t] = E_{\sigma^2} \left[\frac{A\gamma_p\sigma^{2p}}{t^p} + t \right],$$

where $\gamma_p = 2^p\Gamma(\frac{1}{2} + p)/\sqrt{\pi}$, and this is of the form $\eta E_{\sigma^2} [K(t/\eta)]$ with $\eta = (Ap\gamma_p)^{\frac{1}{p+1}}\sigma^{\frac{2p}{p+1}}$ and $K(x) = 1/(px^p) + x$.

Observe that for both the interval and point estimation problems, K is of the form

$$K(x) = K_0(x) + x, \quad (6)$$

where K_0 is a non-negative decreasing strictly convex function for which $K'_0(1) = -1$ (so that $K(x)$ is minimized when $x = 1$), and

$$\eta = a\sigma^{2q}, \quad (7)$$

where $0 < a, q < \infty$. In the point estimation problem, $K_0(x) = 1/(px^p)$, $q = p/(p+1)$, and $a = (Ap\gamma_p)^{1/(p+1)}$. For interval estimation, $K_0(x) = 2[1 - \Phi(c\sqrt{x})]/[c\varphi(c)]$, $q = 1$, and $a = c^2/h^2$.

In the statement of the main result, the statistician must specify a pair $\delta = (m, t)$, where $m \geq 2$ is an integer, the initial sample size, and $t = t(S_m^2)$ is a measurable integer valued function for which $t \geq m$. The loss incurred when δ is used is taken to be $\eta K(t/\eta)$, where K and η are as in (6) and (7), and the risk is then

$$R_a(\delta; \sigma^2) = \eta E_{\sigma^2} \left[K\left(\frac{t}{\eta}\right) \right]. \quad (8)$$

The function K_0 is required to be a twice continuously differentiable convex function for which

$$K_0(x) = O(x^{-\ell})$$

as $x \downarrow 0$ for some $\ell > 0$.

For a given σ_0 , let Ξ_0 be the class of prior distributions ξ for which

$$\int_0^\infty \sigma^{2q} \xi \{d\sigma\} = \sigma_0^{2q}.$$

Further, let

$$\begin{aligned} r_a(\delta; \sigma^2) &= R_a(\delta; \sigma^2) - K(1)\eta, \\ \bar{R}_a(\delta; \xi) &= \int_0^\infty R_a(\delta; \sigma^2) \xi \{d\sigma\}, \end{aligned}$$

and

$$\bar{r}_a(\delta; \xi) = \int_0^\infty r_a(\delta; \sigma^2) \xi \{d\sigma\}$$

for $\xi \in \Xi_0$. Finally, let $\eta_0 = a\sigma_0^{2q}$, the prior expectation of η , and let δ_a be the procedure defined by

$$m_a = \lceil q\sqrt{K''(1)\eta_0} \rceil \quad (9)$$

and

$$t_a = \max\{m_a, \lceil aS_m^{2q} \rceil\}. \quad (10)$$

Theorem 2.1. *With the notation and assumptions of the previous two paragraphs,*

$$\inf_{\delta} \sup_{\xi \in \Xi_0} \bar{r}_a(\delta; \xi) = 2q\sqrt{K''(1)\eta_0} + o(\sqrt{\eta_0}) = \sup_{\xi \in \Xi_0} \bar{r}_a(\delta_a; \xi) \quad (11)$$

as $a \rightarrow \infty$.

The proof of Theorem 2.1) is presented in Section 4. The following corollary contains (4) as a special case. In its statement $0 < \alpha < 1$ and $0 < \sigma_0^2 < \infty$ are fixed, $a = c^2/h^2$, $\eta_0 = c^2\sigma_0^2/h^2$, and $h \downarrow 0$.

Corollary 2.1. *If $\delta_h = (m_h, t_h)$ are any procedures for which*

$$P_{\mu, \sigma^2} [|\bar{X}_{t_h} - \mu| \leq h] \geq 1 - \alpha \quad (12)$$

for all μ and σ^2 , then

$$\sup_{\xi \in \Xi_0} \int_0^\infty E_{\sigma^2}(t_h) \xi \{d\sigma\} \geq \eta_0 + 2\sqrt{\left(\frac{1+c^2}{2}\right)\eta_0} + o(\sqrt{\eta_0}) \quad (13)$$

as $h \downarrow 0$. Moreover, there is equality in (13) if m_h is as in (3) and $t_h = T_{m_h}$.

Proof. If δ_h satisfies (12), then $E_{\sigma^2}[K_0(t/\eta)] \leq \alpha/[c\varphi(c)]$,

$$\bar{R}_a(\delta_h; \xi) \leq \frac{\alpha\eta_0}{c\varphi(c)} + E_\xi(t_h) = K(1)\eta_0 + E_\xi(t_h - \eta_0),$$

and, therefore,

$$\int_0^\infty E_{\sigma^2}(t_h) \xi \{d\sigma\} - \eta_0 \geq \bar{r}_a(\delta_h; \xi)$$

for all $\xi \in \Xi_0$. The inequality asserted in (13) now follows directly from the theorem, since $q = 1$ and $K''(1) = (1+c^2)/2$ for interval estimation.

The equality asserted in the corollary follows from a direct calculation. Clearly, $T_m \leq c_{m-1}^2 S_m^2/h^2 + m + 1$ for any m and, therefore,

$$E_\xi(T_m - \eta_0) \leq \left(\frac{c_{m-1}^2 - c^2}{h^2}\right)\sigma_0^2 + m + 1$$

for any $\xi \in \Xi_0$. From [1], p. 949, or first principles, $c_m = c + (c + c^3)/4m + O(1/m^2)$ as $m \rightarrow \infty$ and, therefore,

$$\left(\frac{c_{m-1}^2 - c^2}{h^2}\right)\sigma_0^2 = \left(\frac{c^2 + c^4}{2mh^2}\right)\sigma_0^2 + O\left(\frac{1}{m^2h^2}\right) = \left(\frac{1+c^2}{2m}\right)\eta_0 + O\left(\frac{1}{m^2h^2}\right).$$

Letting $m = m_h$ and combining the last two expressions now shows that there is equality in (13) when m_h is as in (3). \diamond

3. Preliminary Lemmas

In the statements of lemmas below q is fixed,

$$g_{\alpha,\beta}(y) = \frac{\beta^{\frac{1}{2}\alpha} y^{\frac{1}{2}\alpha-1} e^{-\frac{1}{2}\beta y}}{2^{\frac{1}{2}\alpha} \Gamma(\frac{1}{2}\alpha)}$$

$$\nu_\alpha = \int_0^\infty y^q g_{\alpha,1}(y) dy = \frac{2^q \Gamma(\frac{1}{2}\alpha + q)}{\Gamma(\frac{1}{2}\alpha)}.$$

and

$$\bar{K}_\alpha(x) = \int_0^\infty K\left(\frac{xy^q}{\nu_\alpha}\right) g_{\alpha,1}(y) dy, \quad (14)$$

where Γ denotes the gamma function. Here $\bar{K}_\alpha(x)$ is defined for all $0 < \alpha, x < \infty$, though possibly infinite. Properties of \bar{K}_α are central to the proof.

In the proofs of the following Lemmas, use is made of Stirling's Formula,

$$\log \Gamma(z) = \left(z - \frac{1}{2}\right) \log(z) - z + \frac{1}{2} \log(2\pi) + \frac{1}{12z} + O\left(\frac{1}{z^3}\right) \quad (15)$$

as $z \rightarrow \infty$. See, for example, [1], p. 267. In particular, it follows from Stirling's Formula that

$$\nu_\alpha = \alpha^q + O(1)$$

as $\alpha \rightarrow \infty$.

Lemma 3.1. *i) The integral in (14) converges for all $\alpha > 2lq$.*

ii) If $\bar{K}_\alpha(x) < \infty$ for some $0 < \alpha, x < \infty$, then $\bar{K}_{\alpha'}(x') < \infty$ for all $\alpha' \geq \alpha$ all x' .

iii) In this case, \bar{K}_α is a twice continuously differentiable, strictly convex function, and $\bar{K}_\alpha(x) > K(x)$ for all x .

iv) Finally,

$$\bar{K}_\alpha(x) = K(x) + \frac{q^2 x^2 K''(x)}{\alpha} + o\left(\frac{1}{\alpha}\right) \quad (16)$$

uniformly in x on compact subintervals of $(0, \infty)$ as $\alpha \rightarrow \infty$.

Proof. Clearly, $\bar{K}_\alpha(x) < \infty$ iff $\int_0^1 K_0(xy^q/\nu_\alpha)g_{\alpha,1}(y)dy < \infty$. The first assertion follows directly. For ii) and iii), write

$$\bar{K}_\alpha = \int_0^\infty K\left(\frac{y^q}{\nu_\alpha}\right)g_{\alpha,x^{-\frac{1}{q}}}(y)dy. \quad (17)$$

The second assertion follows since $g_{\alpha',x'^{-\frac{1}{q}}}(y)/g_{\alpha,x^{-\frac{1}{q}}}(y)$ remains bounded as $y \rightarrow 0$ when $\alpha' \geq \alpha$ and $0 < x, x' < \infty$. The continuous differentiability asserted in iii) also follows from (17), since $g_{\alpha,\beta}$ is an exponential family in β ; \bar{K}_α is, in fact, analytic in $x > 0$. See Brown [2], pp. 34-36. That \bar{K}_α is strictly convex and $\bar{K}_\alpha(x) > K(x)$, then follow from the convexity of K , the positivity of $g_{\alpha,1}$, and Jensen's Inequality. Relation (16) may be formally obtained by applying the delta method and rigorously established along the lines of Chapter 3 in [9]. \diamond

Lemma 3.2. *For all $\alpha > 0$,*

$$\inf_{0 < x < \infty} \bar{K}_\alpha(x) > K(1). \quad (18)$$

If \bar{K}_α is finite, then $\bar{K}_\alpha(x)$ attains its minimum at a unique value x_α . As $\alpha \rightarrow \infty$, $x_\alpha \rightarrow 1$ and

$$\inf_{x > 0} \bar{K}_\alpha(x) = \bar{K}_\alpha(1) + o\left(\frac{1}{\alpha}\right). \quad (19)$$

Proof. Since (18) is clear when $\bar{K}_\alpha \equiv \infty$, it suffices to consider α for which \bar{K}_α is finite. For such α , $\lim_{x \rightarrow \infty} \bar{K}_\alpha(x) = \infty$ and

$$\lim_{x \downarrow 0} \bar{K}'_\alpha(x) = \lim_{x \downarrow 0} K'(x) < 0,$$

so that the infimum cannot be attained as $x \downarrow 0$ or $x \rightarrow \infty$. That the infimum is attained at a unique point then follows from the strict convexity, and $\bar{K}_\alpha(x_\alpha) > K(x_\alpha) \geq K(1)$, establishing (18). Next $\inf_{|x-1| \geq \epsilon} \bar{K}_\alpha(x) \geq \inf_{|x-1| \geq \epsilon} K(x) > \bar{K}_\alpha(1)$ for all sufficiently large α for any $\epsilon > 0$ and, therefore, that $x_\alpha \rightarrow 1$ as $\alpha \rightarrow \infty$. Relation (19) then follows from (16),

since

$$\begin{aligned}\bar{K}_\alpha(x_\alpha) &= K(x_\alpha) + \frac{q^2 x_\alpha^2 K''(x_\alpha)}{\alpha} + o\left(\frac{1}{\alpha}\right) \\ &\geq K(1) + \frac{q^2 K''(1)}{\alpha} + o\left(\frac{1}{\alpha}\right) \\ &= \bar{K}_\alpha(1) + o\left(\frac{1}{\alpha}\right),\end{aligned}$$

where the inequality uses $K(x) \geq K(1)$ and $x_\alpha \rightarrow 1$. \diamond

Let $\xi_{\alpha,\beta}$ be the prior distribution for σ under which $\theta = 1/\sigma^2$ has density $g_{\alpha,\beta}$. Also, write $E_{\alpha,\beta}$ for unconditional expectation in the Bayesian model and $E_{\alpha,\beta}^m$ for conditional expectation given S_m^2 , when σ has prior $\xi_{\alpha,\beta}$. Then

$$U_0 := \int_0^\infty \sigma^{2q} \xi_{\alpha,\beta} \{d\sigma\} = \int_0^\infty \theta^{-q} g_{\alpha,\beta}(\theta) d\theta = \frac{\beta^q}{\nu_{\alpha-2q}}$$

for $\alpha > 2q$, and $\xi \in \Xi_0$ iff $\beta = \nu_{\alpha-2q}^{1/q} \sigma_0^2$, in which case $\eta_0 = aU_0$. Similarly,

$$\int_0^\infty \eta K\left(\frac{x}{\eta}\right) \xi_{\alpha,\beta} \{d\sigma\} = a \int_0^\infty \theta^{-q} K\left(\frac{x\theta^q}{a}\right) g_{\alpha,\beta}(\theta) d\theta = \eta_0 \bar{K}_{\alpha-2q}\left(\frac{x}{\eta_0}\right).$$

Inverted gamma priors $\xi_{\alpha,\beta}$ are conjugate to scaled chi-squared distributions, and

$$U_m := E_{\alpha,\beta}^m(\sigma^{2q}) = \frac{\beta_m^q}{\nu_{\alpha_m-2q}},$$

where

$$\begin{aligned}\alpha_m &= \alpha + m - 1, \\ \beta_m &= \beta + (m-1)S_m^2.\end{aligned}$$

Lemma 3.3. *If $\alpha > 2(\ell+1)q$ and $x > 0$, then*

$$E_{\alpha,\beta}[\eta K\left(\frac{x}{\eta}\right)] = \bar{\eta} \bar{K}_{\alpha-2q}\left(\frac{x}{\bar{\eta}}\right),$$

where $\bar{\eta} = aE_{\alpha,\beta}(\sigma^{2q})$.

Proof. The proof depends on the simple relation

$$\theta^{-q} g_{\alpha,\beta}(\theta) = \frac{\beta^q}{\nu_{\alpha-2q}} g_{\alpha-2q,\beta}(\theta).$$

10

It follows that $\bar{\eta} = a\beta^q/\nu_{\alpha-2q}$ and

$$\begin{aligned} E_{\alpha,\beta}[\eta K(\frac{x}{\eta})] &= \int_0^\infty a\theta^{-q} K(\frac{x\theta^q}{a}) g_{\alpha,\beta}(\theta) d\theta \\ &= \frac{a\beta^q}{\nu_{\alpha-2q}} \int_0^\infty K(\frac{x\theta^q}{a}) g_{\alpha-2q,\beta}(\theta) d\theta \\ &= \bar{\eta} \int_0^\infty K(\frac{x\theta^q}{\bar{\eta}\nu_{\alpha-2q}}) g_{\alpha-2q,1}(\theta) d\theta \\ &= \bar{\eta} \bar{K}_{\alpha-2q}(\frac{x}{\bar{\eta}}), \end{aligned}$$

as asserted. ◇

Lemma 3.4. *There is a constant C for which*

$$P_{\alpha,\beta}[U_m > u] \leq C(\beta u^{-\frac{1}{q}})^{\frac{1}{2}\alpha}$$

for all $u > 0$, provided that $m \geq 2$ and $\alpha + m - 1 \geq 2q + 1$.

Proof. Let $Z_m = \beta_m/\beta$. Then the marginal density of Z_m is

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

for $1 \leq z < \infty$. This is at most $C_0 m^{\frac{1}{2}\alpha}/z^{1+\frac{1}{2}\alpha}$ for all $\alpha > 2q$ and all $z \geq 1$ for some constant C_0 . So,

$$P_{\alpha,\beta}[Z_m > z] \leq \frac{C_0 m^{\frac{1}{2}\alpha}}{qz^{\frac{1}{2}\alpha}}$$

for all $\alpha > 2q$ and $z \geq 1$. The Lemma is an easy consequence of this and (15), since

$$P_{\alpha,\beta}[U_m > u] = P_{\alpha,\beta}[Z_m > \frac{\nu_{\alpha_m-q}^{\frac{1}{q}} u^{\frac{1}{q}}}{\beta}].$$

4. Proof of the Main Result

It suffices to show that

$$\sup_{\xi \in \Xi_0} \inf_{\delta} \bar{r}_a(\delta; \xi) \geq 2q\sqrt{K''(1)\eta_0} + o(\sqrt{\eta_0}) \quad (20)$$

and

$$\sup_{\xi \in \Xi_0} \bar{r}_a(\delta_a; \xi) \leq 2q\sqrt{K''(1)\eta_0} + o(\sqrt{\eta_0}) \quad (21)$$

The Upper Bound. Relation (21) is established first. With m_a and t_a as in (9) and (10), let a be so large that $m_a \geq 2\ell q$. Then it is clear from the monotonicity of K_0 that

$$K\left(\frac{t_a}{\eta}\right) \leq K\left[\left(\frac{S_{m_a}^2}{\sigma^2}\right)^q\right] + \frac{m_a + 1}{\eta}.$$

So,

$$R_a(\delta_a; \sigma^2) \leq \eta \bar{K}_{m_a-1}(y_a) + m_a + 1,$$

where $y_a = \nu_{m_a-1}/(m_a - 1)^q$ and, therefore,

$$\sup_{\xi \in \Xi_0} \bar{r}_a(\delta_a; \xi) \leq \eta_0[\bar{K}_{m_a-1}(y_a) - K(1)] + m_a + 1,$$

The right side is easily analyzed. From Stirling's Formula, $y_a = 1 + O(1/m_a)$. Then $\bar{K}_{m_a-1}(y_a) = K(y_a) + q^2 y_a^2 K''(y_a)/m_a + o(1/m_a)$, by Lemma 3.1, and $K(y_a) + q^2 y_a^2 K''(y_a)/m_a = K(1) + q^2 K''(1)/m_a + o(1/m_a)$ by the continuity of K'' . This leaves

$$\begin{aligned} \sup_{\xi \in \Xi_0} \bar{r}_a(\delta_a; \xi) &\leq \frac{q^2 K''(1)\eta_0}{m_a} + m_a + 1 + o\left(\frac{\eta_0}{m_a}\right) \\ &\leq 2q\sqrt{K''(1)\eta_0} + o(\sqrt{\eta_0}) \end{aligned}$$

where the last step uses the definition (9) of m_a . This establishes (21).

The Lower Bound. Let $\beta = \beta_a = a^{-2/q}$, and let $\alpha = \alpha_a$ be determined by $\beta^q/\nu_{\alpha-2q} = \sigma_0^{2q}$. Then the inverted gamma priors $\xi_{\alpha,\beta}$ are in Ξ_0 , $\beta \rightarrow 0$, $\alpha \downarrow 2q$ as $a \rightarrow \infty$, and

$$P_{\alpha,\beta}[U_m > \frac{\epsilon m}{a}] \leq C\left[\beta\left(\frac{a}{\epsilon m}\right)^{\frac{1}{q}}\right]^{\frac{1}{2}\alpha} \leq \frac{C}{\epsilon\sqrt{a}} \quad (22)$$

for all $m \geq 2$, $0 < \epsilon < 1$, and large a . It will be shown that

$$\inf_{\delta} \bar{r}(\delta; \xi_{\alpha,\beta}) \geq 2q\sqrt{K''(1)\eta_0} + o(\sqrt{\eta_0}). \quad (23)$$

To establish this, first observe that

$$\inf_{\delta} \bar{R}(\delta; \xi_{\alpha,\beta}) = \inf_{m \geq 2} E_{\alpha,\beta} \left\{ \inf_{t \geq m} E_{\alpha,\beta}^m \left[\eta K\left(\frac{t}{\eta}\right) \right] \right\} \quad (24)$$

12

and

$$E_{\alpha,\beta}^m[\eta K(\frac{t}{\eta})] = \bar{\eta}_m \bar{K}_{\alpha_m-2q}(\frac{t}{\bar{\eta}_m})$$

for fixed $m \geq 2$ and $t \geq m$, where $\bar{\eta}_m = aU_m$. So, for fixed m ,

$$\inf_{t \geq m} E_{\alpha,\beta}^m[\eta K(\frac{t}{\eta})] \geq I_m + \mathbb{I}_m,$$

where

$$I_m = \inf_{x>0} \bar{\eta}_m \bar{K}_{\alpha_m-2q}(x)$$

$$\mathbb{I}_m = \bar{\eta}_m [\bar{K}_{\alpha_m-2q}(\frac{m}{\bar{\eta}_m}) - \inf_{x>0} \bar{K}_{\alpha_m-2q}(x)] \mathbf{1}_{\{x_m^* < m/\bar{\eta}_m\}},$$

and x_m^* is the value at which $\bar{K}_{\alpha_m-2q}(x)$ is minimized. Clearly, $E_{\alpha,\beta}(U_m) = U_0$ and, therefore, $E_{\alpha,\beta}(\bar{\eta}_m) = \eta_0$. So,

$$E_{\alpha,\beta}(I_m) = \eta_0 \inf_{x>0} \bar{K}_{\alpha_m-2q}(x).$$

Now, $\inf_{\alpha \leq \alpha^o} \inf_{x>0} \bar{K}_\alpha(x) > K(1)$ for any α^o by Lemma 3.2. Combining this with (21), it follows that the infimum in (24) can be restricted of $m \geq m_0$ for any given m_0 for all sufficiently large a .

Considering only large a and m ,

$$E_{\alpha,\beta}(I_m) = \eta_0 [K(1) + \frac{q^2 K''(1)}{\alpha + m - 1} + o(\frac{1}{m})]$$

Next,

$$E_{\alpha,\beta}(\mathbb{I}_m) = \mathbb{I}_{1,m} - \mathbb{I}_{2,m},$$

where

$$\mathbb{I}_{1,m} = \int_{x_m^* < m/\bar{\eta}_m} \bar{\eta}_m [\bar{K}_{\alpha_m-2q}(\frac{m}{\bar{\eta}_m}) - K(1)] dP_{\alpha,\beta}$$

and

$$\mathbb{I}_{2,m} = \int_{x_m^* < m/\bar{\eta}_m} \bar{\eta}_m [\inf_{x>0} \bar{K}_{\alpha_m-2q}(x) - K(1)] dP_{\alpha,\beta}.$$

Here

$$\begin{aligned} |\mathbb{I}_{2,m}| &\leq \int_{x_m^* < m/\bar{\eta}_m} \bar{\eta}_m \left[\frac{q^2 K''(1)}{m} + o\left(\frac{1}{m}\right) \right] dP_{\alpha,\beta} \\ &\leq \frac{m}{x_m^*} \left[\frac{q^2 K''(1)}{m} + o\left(\frac{1}{m}\right) \right] \\ &= O(1). \end{aligned}$$

For $\mathbb{I}_{1,m}$ observe that $x_m^* \leq m/\bar{\eta}_m$ iff $\bar{\eta}_m \leq m/x_m^*$ and recall that $x_m^* \rightarrow 1$. So, for any $0 < \epsilon < 1$,

$$\mathbb{I}_{1,m} \geq (1 - \epsilon)mP_{\alpha,\beta}[\bar{\eta}_m \leq \frac{\epsilon m}{K(1)}] \geq (1 - \epsilon)^2 m$$

for all $m \geq m_0$ for all sufficiently large a , by (22). This leaves

$$\begin{aligned} \inf_{\delta} \bar{r}(\delta; \xi_{\alpha,\beta}) &\geq \inf_{m \geq m_0} \left\{ \left[\frac{q^2 K''(1)}{m} + o\left(\frac{1}{m}\right) \right] \eta_0 + (1 - \epsilon)^2 m + O(1) \right\} \\ &= 2(1 - \epsilon)q\sqrt{K''(1)\eta_0} + o(\sqrt{\eta_0}), \end{aligned}$$

establishing (23) since ϵ was arbitrary.

5. Acknowledgements

Thanks to Anirban Das Gupta, Bob Keener, and Jaiyang Sun for helpful discussions. This research was supported by the National Science Foundation and U.S. Army Research Office.

References

1. Abromowitz, Milton and Irene Stegun (1964). *Handbook of Mathematical Functions*. National Bureau of Standards.
2. Brown, Larry (1986). *Fundamentals of Statistical Exponential Families with Applications in Statistical Decision Theory*. IMS Lecture Notes - Monograph Series, **9**.
3. Cohen, Arthur and Harold Sackrowitz (1985). Bayes double sampling procedures. *Ann. Statist.*, **12**, 1035-1049.
4. Cohen, Arthur and Harold Sackrowitz (1985). Double sampling estimation when cost depends on parameters. *Statistical Decision Theory and Related Topics*, **V**, 253-266.

5. Denne, Jonathan and Christopher Jennison (1999). Estimation the sample size for a t -test using an internal pilot. *Stat. Med.*, **18**, 1575-1585.
6. Ghosh, Malay, Nitus Mukhopadhyay, and Pranab Kumar Sen (1997). *Sequential Estimation*. Wiley.
7. Stein, Charles (1945). A two-stage test for a linear hypothesis whose power is independent of the variance. *Ann. Math. Statist.*, **16**, 243-258.
8. Treder, R.P. and J. Sedransk (1996). Bayesian Sequential Two-Phase Sampling. *J. Amer. Statist. Assn.*, **91**, 782-790.
9. Van der Vaart, Aad (1998). *Asymptotic Statistics*. Cambridge University Press
10. Woodroffe, Michael (1982). *Non-linear Renewal Theory in Sequential Analysis*. S.I.A.M.
11. Woodroffe, Michael (1986). Asymptotic optimality in sequential interval estimation. *Adv. Appl. Math.*, **7**, 70-79.
12. Woodroffe, Michael (2004). An asymptotic minimax determination of the initial sample size in a two-stage procedure. *IMS Lecture Notes - Monograph Series*, **45**, 228-236