Estimating the Mean of a Normal Distribution with Known Coefficient of Variation†

by

Leon Jay Gleser and John D. Healy
Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #407

February 1975

---

†Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant AFOSR-72-2350B at Purdue University. The United States Government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright hereon.

Estimating the Mean of a Normal Distribution with Known Coefficient of Variation

by

Leon Jay Gleser and John D. Healy
Purdue University

ABSTRACT

In the present paper, it is shown that for estimating the mean θ of a normal distribution $N(\theta, a\theta^2)$ with known coefficient of variation $\sqrt{a}$, Khan's [1] minimum variance estimator $\hat{\theta}_{LU}$ among all unbiased estimators linear in the sample mean $\bar{x}$ and sample standard deviation s is inadmissible under squared error loss. The estimator θ_{LMMS} which has uniformly minimum risk among all estimators linear in $\bar{x}$ and s is obtained. An admissible Bayes estimator $\hat{\theta}_B$ of θ under a natural prior is also obtained, along with simple upper and lower bounds, $\hat{\theta}_{B+}$ and $\hat{\theta}_{B-}$, for $\hat{\theta}_B$ which closely approximate $\hat{\theta}_B$ and which resemble in form the maximum likelihood estimator of θ. All of the estimators obtained are shown to be B.A.N. The domination of $\hat{\theta}_{LU}$ by $\hat{\theta}_{LMMS}$ in terms of risk under squared error is shown to be a consequence of a general inadmissibility of unbiased, scale invariant estimators of scale parameters under squared error loss.

Estimating the Mean of a Normal Distribution with Known Coefficient of Variation

by

Leon Jay Gleser and John D. Healy*
Purdue University

1. INTRODUCTION AND SUMMARY

If we sample from a normal distribution $N(\theta,\sigma^2)$ and attempt to estimate θ under a squared-error loss function, it is known that the sample mean $\bar{x}$ is an admissible estimator for θ (and also is UMV unbiased). However, suppose that we sample from a normal distribution in which it is known that σ^2 is a fixed given multiple, $\sigma^2 = a\theta^2$, of θ^2, and the unknown mean θ is positive ($\theta>0$). Put another way, suppose that the coefficient of variation $\sqrt{a}$ of the normal population is known. In this situation, Khan [1] found an unbiased estimator $\hat{\theta}_{LU}$ of θ which is a linear combination of the sample mean $\bar{x}$ and the sample standard deviation s, and which has minimum variance among all unbiased estimators linear in $\bar{x}$ and s. Since Khan's estimator does not equal $\bar{x}$, his result shows that $\bar{x}$ is inadmissible as an estimator of the mean θ of a normal distribution when the coefficient of variation is known.

We note that in the model $N(\theta,a\theta^2)$, the parameter θ acts as a scale parameter. Typically, in point estimation problems involving a scale parameter (and squared error loss), admissibility and unbiasedness are incompatible, in the sense that unbiased estimators are usually not admissible.

---

*Leon Jay Gleser is associate professor and John D. Healy is a graduate student in the Department of Statistics, Purdue University, Mathematical Sciences Building, West Lafayette, Indiana 47907. The research upon which this paper is based was supported by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant AFOSR-72-2350B at Purdue University.

This fact suggests that Khan's estimator is itself inadmissible. In Section 2, we find the estimator, $\hat{\theta}_{LMMS}$, which has minimum mean square error among all estimators that are a linear combination of $\bar{x}$ and s. Since $\hat{\theta}_{LMMS} \neq \hat{\theta}_{LU}$, our results demonstrate the inadmissibility of $\hat{\theta}_{LU}$. The estimator $\hat{\theta}_{LMMS}$ is no harder to compute than $\hat{\theta}_{LU}$, and has a simple risk function. It also trivially follows from the characterizations of $\hat{\theta}_{LU}$ and $\hat{\theta}_{LMMS}$ that $\hat{\theta}_{LMMS}$ has a uniformly smaller variance than $\hat{\theta}_{LU}$. We note that Khan [1] has shown that $\hat{\theta}_{LU}$ is best asymptotic normal (B.A.N.); in Section 2, we show that $\hat{\theta}_{LMMS}$ is also B.A.N.

Although the simplicity of $\hat{\theta}_{LMMS}$ recommends it as an estimator of θ, it is very likely that $\hat{\theta}_{LMMS}$ is itself inadmissible within the class of all estimators of θ. This conjecture arises by noting that $\hat{\theta}_{LMMS}$ is a linear combination of the minimal sufficient statistics in a non-linear parametric problem. On the other hand, Bayes estimators against priors which put positive probability on every open set are known to be admissible. Thus, in Section 3 we are led to consider the Bayes estimator $\hat{\theta}_B$ against a certain prior, the inverted-gamma prior, which is customarily used to represent prior opinion about the variance of a normal distribution. (The fact that $\theta > 0$ precludes using a normal prior for θ.) The estimator $\hat{\theta}_B$ cannot be expressed in closed form, but is easily computed using continued fractions. Further, we find two closed-form estimators $\hat{\theta}_{B+}$ and $\hat{\theta}_{B-}$ which closely bound $\hat{\theta}_B$ above and below, respectively. The form of each of these estimators resembles that of the maximum likelihood estimator $\hat{\theta}_{MLE}$. It is shown that all three of these estimators, $\hat{\theta}_B$, $\hat{\theta}_{B+}$, and $\hat{\theta}_{B-}$ are B.A.N.

2. LINEAR MINIMUM MEAN SQUARE ESTIMATION

We assume that a random sample $x_1, x_2, \ldots, x_n$ of fixed size $n \geq 2$ is taken from the normal distribution $N(\theta, a\theta^2)$, where $\theta > 0$ is unknown and $a > 0$ is known. Let

$$\bar{x} = n^{-1} \sum_{i=1}^{n} x_i, \qquad s^2 = n^{-1} \sum_{i=1}^{n} (x_i - \bar{x})^2, \tag{2.1}$$

and let

$$c_n = \sqrt{n}\ (\Gamma(\tfrac{n-1}{2})/(2a)^{1/2}\ \Gamma(\tfrac{n}{2})). \tag{2.2}$$

It is known (Khan [1]) that $T_1 = \bar{x}$ and $T_2 = c_n s$ are both unbiased estimators of θ, i.e.,

$$E_\theta T_1 = E_\theta T_2 = \theta; \tag{2.3}$$

that

$$\mathrm{Var}_\theta\ T_1 = \frac{a\theta^2}{n}, \quad \mathrm{Var}_\theta\ T_2 = \theta^2 \Big[\frac{(n-1)}{2} \frac{\Gamma^2(\frac{n-1}{2})}{\Gamma^2(\frac{n}{2})} - 1\Big]; \tag{2.4}$$

and that T_1 and T_2 are uncorrelated (actually, independent). We remark in passing that the minimal sufficient statistic for θ is $(\bar{x}, s)$, or equivalently (T_1, T_2), but that the family of distributions of the minimal sufficient statistic is not complete. [Rather than use Khan's [1] more complicated proof, we note that lack of completeness follows directly from (2.3).]

From (2.3) it is readily apparent that all estimators of the form

$$\hat{\theta}(a) = \alpha T_1 + (1-\alpha) T_2, \qquad -\infty < \alpha < \infty, \tag{2.5}$$

are unbiased estimators of θ. Khan [1] showed that among the subclass of estimators of the form (2.5) for which $0 \leq \alpha \leq 1$, the estimator

$$\hat{\theta}_{LU}(\alpha) = \alpha^* T_1 + (1-\alpha^*)T_2,$$

where

$$\alpha^* = \frac{\mathrm{Var}_\theta T_2}{\mathrm{Var}_\theta T_1 + \mathrm{Var}_\theta T_2} = \frac{[\frac{n-1}{2} \frac{\Gamma^2(\frac{n-1}{2})}{\Gamma^2(\frac{n}{2})} - 1]}{\frac{a}{n} + (\frac{n-1}{2} \frac{\Gamma^2(\frac{n-1}{2})}{\Gamma^2(\frac{n}{2})} - 1)}, \tag{2.7}$$

has smallest variance (and thus smallest risk under squared-error loss). Actually, as we shall see, the estimator (2.6) has minimum variance (and minimum risk) within the entire class of estimators of the form (2.5).

Now, consider the class of all estimators of θ which are linear in $\bar{x}$ and s. A typical such estimator can be written in the form

$$\hat{\theta}(\alpha_1,\alpha_2) = \alpha_1 T_1 + \alpha_2 T_2, \tag{2.8}$$

since $T_1 = \bar{x}$ and $T_2 = c_n s$. We now find the estimators which have, respectively, smallest risk among all estimators of the form (2.5), and smallest risk among all estimators of the form (2.8). We do this by means of a lemma which has generality beyond that of the present problem.

<u>Lemma 2.1</u>. Let T_1 and T_2 be any two uncorrelated and unbiased estimators of a parameter θ. Assume that the ratios

$$v_i \equiv \theta^{-2}\, \mathrm{Var}_\theta T_i \tag{2.9}$$

are independent of θ, i = 1,2. Then the estimator,

$$T^*_{LU} = \frac{v_2 T_1 + v_1 T_2}{v_1 + v_2}, \tag{2.10}$$

has uniformly (over θ) minimum risk,

$$R(\theta,T^*_{LU}) = \theta^2 \frac{v_1v_2}{v_1+v_2} = \theta^2 \frac{1}{(\frac{1}{v_1}+\frac{1}{v_2})}, \tag{2.11}$$

among all unbiased estimators linear in T_1 and T_2, while the estimator,

$$T^*_{LMMS} = \frac{v_2T_1 + v_1T_2}{v_1+v_2+v_1v_2}, \tag{2.12}$$

has uniformly minimum risk,

$$R(\theta,T^*_{LMMS}) = \theta^2 \frac{v_1v_2}{v_1+v_2+v_1v_2} = \theta^2 \frac{1}{(\frac{1}{v_1}+\frac{1}{v_2}+1)}, \tag{2.13}$$

among all estimators linear in T_1 and T_2.

Proof. For any estimator T,

$$R(\theta,T) = E_\theta(T-\theta)^2 = Var_\theta T + (E_\theta T-\theta)^2.$$

By the given, $E_\theta T_1 = E_\theta T_2 = \theta$, $Var_\theta T_i = \theta^2 v_i$, and T_1 and T_2 are uncorrelated. Thus

$$\begin{aligned} R(\theta,\alpha_1T_1+\alpha_2T_2) &= \alpha_1^2 \, Var_\theta T_1 + \alpha_2^2 \, Var_\theta T_2 + \theta^2(\alpha_1+\alpha_2-1)^2 \\ &= \theta^2[\alpha_1^2v_1+\alpha_2^2v_2 + (\alpha_1+\alpha_2-1)^2], \end{aligned} \tag{2.14}$$

and since (2.14) is quadratic in α_1 and α_2, we can minimize (2.14) with respect to α_1 and α_2 by standard techniques of partial differentiation. The resulting coefficients α_1^* and α_2^* that minimize (2.14) are $\alpha_i^* = v_i/(v_1+v_2+v_1v_2)$, $i = 1,2$. The results (2.12) and (2.13) now follow by substitution and simplification. For the linear combination $\alpha_1T_1+\alpha_2T_2$ to be unbiased, we must have $\alpha_1 + \alpha_2 = 1$. Thus, from (2.14), we have that for any unbiased linear combination $\alpha T_1+(1-\alpha)T_2$, $R(\theta,\alpha T_1+(1-\alpha)T_2) = \theta^2(\alpha^2v_1+(1-\alpha)^2v_2)$. Again, this expression is a quadratic in α, and standard differentiation techniques show that the minimum of the expression is achieved for $\alpha = v_2(v_1+v_2)^{-1}$. The results (2.10) and (2.11) now easily follow. Since both T^*_{LU} and T^*_{LMMS} are independent of θ, the optimality properties (minimization of risk) for these estimators hold uniformly in θ. Q.E.D.

The condition that the v_i are independent of θ (see (2.9)) holds in any problem (including the present one) in which θ is a scale parameter for the distributions of T_1 and T_2. In all such problems, we see from the Lemma that T^*_{LMMS} does not equal T^*_{LU} (indeed, T^*_{LMMS} can be described as a shrinking of T^*_{LU} toward 0), and that T^*_{LMMS} has strictly smaller risk than T^*_{LU}. The difference

$$R(\theta,T^*_{LU}) - R(\theta,T^*_{LMMS}) = \frac{\theta^2}{(\frac{1}{v_1} + \frac{1}{v_2})(1 + \frac{1}{v_1} + \frac{1}{v_2})}$$

between the risks of T^*_{LU} and T^*_{LMMS} is most positive (i.e. most in favor of T^*_{LMMS}) when either v_1 or v_2 are large, while this difference is negligible when both v_1 and v_2 are small. The ratio

$$R(\theta,T^*_{LU})/R(\theta,T^*_{LMMS}) = \frac{(\frac{1}{v_1} + \frac{1}{v_2} + 1)}{(\frac{1}{v_1} + \frac{1}{v_2})}$$

shows a similar trend, being largest (favoring T^*_{LMMS}) when either v_1 or v_2 is large, and being near 1 when both v_1 and v_2 are small.

In the present $N(\theta,a\theta^2)$ problem, we see from (2.4) that

$$v_1 = \frac{a}{n}, \quad v_2 = [\frac{(n-1)}{2}\frac{\Gamma^2(\frac{n-1}{2})}{\Gamma^2(\frac{n}{2})} - 1]. \tag{2.15}$$

Since both v_1 and v_2 tend monotonically to 0 at rate n^{-1} as $n \to \infty$, we see that there is little choice between $\hat{\theta}_{LU} \equiv T^*_{LU}$ and $\hat{\theta}_{LMMS} = T^*_{LMMS}$ in large samples, while use of $\hat{\theta}_{LMMS}$ is clearly indicated when n is of small or moderate size.

Since

$$\hat{\theta}_{LMMS} = \left(\frac{v_1+v_2}{v_1+v_2+v_1v_2}\right)\hat{\theta}_{LU} \tag{2.16}$$

and

$$\lim_{n\to\infty} n^{1/2}\left(\frac{v_1+v_2}{v_1+v_2+v_1v_2} - 1\right) = 0,$$

it follows that $n^{1/2}(\hat{\theta}_{LMMS} - \hat{\theta}_{LU}) \xrightarrow{P} 0$. Since Khan (1968) has shown that $\hat{\theta}_{LU}$ is B.A.N., we conclude that $\hat{\theta}_{LMMS}$ is also B.A.N.

Remark I: We could have also shown that $\hat{\theta}_{LMMS}$ has smaller risk than $\hat{\theta}_{LU}$ by noting from (2.11), (2.15) and (2.16) that

$$\hat{\theta}_{LMMS} = \left(\frac{1}{1 + \frac{v_1v_2}{v_1+v_2}}\right)\hat{\theta}_{LU}$$

$$= \frac{1}{\left(1 + \frac{\mathrm{Var}_\theta\hat{\theta}_{LU}}{\theta^2}\right)}\hat{\theta}_{LU},$$

and using the following general lemma.

Lemma 2.2. If T is any unbiased estimator of a parameter θ whose variance $\mathrm{Var}_\theta T$ has the property that $v = \theta^{-2}\mathrm{Var}_\theta T$ is independent of θ, then T is inadmissible under squared error loss, and is uniformly improved upon by the biased estimator $(1+v)^{-1}T$.

Proof. Since

$$R(\theta,\frac{1}{1+v}T) = \left(\frac{1}{1+v}\right)^2 \mathrm{Var}_\theta T + \left(\frac{1}{1+v}\theta - \theta\right)^2$$

$$= \frac{v}{(1+v)^2}\theta^2 + \left(\frac{1}{1+v} - 1\right)^2\theta^2$$

$$= \theta^2 \frac{v}{1+v} < \theta^2 v = R(\theta,T),$$

the inadmissibility of θ is established. Q.E.D.

Lemma 2.2 is constantly rediscovered in the literature in various contexts. For example, using Lemma 2.2, we can immediately conclude that the unbiased estimator s^2 of the variance σ^2 of the normal distribution is inadmissible, and [since $\mathrm{Var}_{\sigma^2}\, s^2 = 2\sigma^2/(n-1)$] that

$$\left(\frac{1}{1+\frac{2}{n-1}}\right)s^2 = \frac{n-1}{n+1}\, s^2 = \frac{\sum_{i=1}^{n} (x_i-\bar{x})^2}{n+1}$$

is superior to s^2 in terms of risk for all σ^2. In general, Lemma 2.2 shows that any unbiased estimator T of a parameter θ whose distribution involves θ as a scale parameter must be inadmissible.

Remark II: Once we note that θ acts as a scale parameter for the $N(\theta, a\theta^2)$ distribution, we might think of transforming the data $x_1,\ldots,x_n$ by taking logarithms. (In analysis of variance problems, where the situation that the standard deviation is a fixed constant multiple of the mean is observed, this logarithmic transformation is frequently utilized to "normalize" the data.) The logarithmic transformation, if well defined, changes scale parameter problems to location parameter problems. However, since the x_i's can be negative, the logarithmic transformation is not well defined in the present problem. Taking logarithms of $|x_i|$ may reduce the data too much to retain efficiency in estimation of θ, while taking logarithms of x_i^+ (the positive part of x_i) gets us into complicated distributional problems involving truncated normal distributions when θ is small, and also may lose much of the information provided by the data. In any case, the location parameter of the distribution of $\log x_i$, when x_i is positive, is $\log \theta$, and the loss function customarily used for location parameter problems is squared error loss in the location parameter, $\log \theta$, which does not directly indicate loss in estimating θ. For these reasons, we have not investigated the transformation approach to estimating θ in the present problem.

3. A BAYES ESTIMATOR

If we were considering estimating the mean θ for a general normal population $N(\theta,\sigma^2)$, the customary approach would be to make use of a (conjugate) normal prior for θ. In the present case, however, θ is positive, and a prior normal distribution for θ does not make much sense. Noting again that θ acts as a scale parameter for the $N(\theta,a\theta^2)$ distribution, and thus that θ can also be regarded as a measure of variation, we are led to consider using a class of prior distributions, the inverted gamma distributions, which have previously been used in estimating variances of normal populations. Thus, consider a prior density $\psi(\theta)$ for θ of the form

$$\psi(\theta) = \begin{cases} \dfrac{\theta^{-r-1} e^{-w/\theta} w^r}{\Gamma(r)}, & \text{if } \theta > 0, \\ 0 , & \text{otherwise,} \end{cases} \tag{3.1}$$

where $w > 0$ is a given positive number, and $r > 0$ is a given integer. Since this density assigns positive probability to every open set in the parameter space $\Theta = (0,\infty)$, we know that the Bayes estimator,

$$\hat{\theta}_B = E(\theta|x_1,\ldots,x_n) = E(\theta|\bar{x},s^2), \tag{3.2}$$

will be admissible under squared-error loss.

The joint density function of $\bar{x}$, s^2, and θ is

$$g(\bar{x},s^2,\theta) = \psi(\theta)\ f(\bar{x},s^2|\theta), \tag{3.3}$$

where

$$f(\bar{x},s^2|\theta) = \frac{(s^2)^{(n-3)/2}\, n^{n/2} \exp - \frac{n}{2a\theta^2}[(\bar{x}-\theta)^2+s^2]}{\Gamma(\frac{n-1}{2})\, \sqrt{2\pi}\, (a\theta^2)^{n/2}}, \tag{3.4}$$

for $-\infty < \bar{x} < \infty$, $0 < s^2$. Since

$$\hat{\theta}_B = E(\theta|\bar{x},s^2) = \frac{\int_0^\infty \theta\, g(\bar{x},s^2,\theta)d\theta}{\int_0^\infty g(\bar{x},s^2,\theta)d\theta},$$

we obtain, after simplification of the integrals,

$$\hat{\theta}_B = \frac{\sqrt{u}\int_0^\infty \tau^{n+r-2} e^{-\frac{1}{2}(\tau-z)^2} d\tau}{\int_0^\infty \tau^{n+r-1} e^{-\frac{1}{2}(\tau-z)^2} d\tau} \tag{3.5}$$

$$\equiv \sqrt{u}\, h_{n+r-1}(z),$$

where $\tau = \theta^{-1}$,

$$u = a^{-1}\sum_{i=1}^n x_i^2, \qquad z = (-w + a^{-1}\sum_{i=1}^n x_i)(u)^{-1/2}. \tag{3.6}$$

Two useful recursive methods for calculating $h_{n+r-1}(z)$ can be obtained through use of integration by parts. Both recursive solutions always converge, but the first relation gives faster convergence when $z \geq 0$, and the second converges faster when $z \leq 0$.

<u>Relation I</u> $$h_{i+1}(z) = \frac{1}{i\, h_i(z) + z}, \tag{3.7}$$

<u>Relation II</u> $$h_i(z) = \left(\frac{1}{h_{i+1}(z)} - z\right)\frac{1}{i}. \tag{3.8}$$

We remark that these relations hold and yield recursive solutions even when the parameter $r > 0$ of the prior (3.1) is not an integer. When r is an integer, the initial value for (3.7) is

$$h_1(z) = (\sqrt{2\pi}\,\Phi(z))/(\sqrt{2\pi}\, z\Phi(z) + e^{-z^2/2}), \tag{3.9}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal N(0,1) distribution. Also, when r is an integer, and $z \leq 0$, we can use (3.8) to write $h_i(z)$ as an infinite regular continued fraction:

$$h_i(z) = [a_0^{(i)}, a_1^{(i)}, a_2^{(i)}, a_3^{(i)}, \ldots] \tag{3.10}$$

$$= a_0^{(i)} + \cfrac{1}{a_1^{(i)} + \cfrac{1}{a_2^{(i)} + \ldots}},$$

where for $k > 0$,

$$a_k^{(i)} = \begin{cases} |z| \prod_{\ell=1}^{\frac{1}{2}(k+1)} \left(\frac{i+2\ell-2}{i+2\ell-1}\right), & k \text{ odd}, \\ |z| \prod_{\ell=1}^{\frac{1}{2}(k)} \left(\frac{i+2\ell-1}{i+2\ell-2}\right), & k \text{ even}, \end{cases} \tag{3.11}$$

while $a_0^{(i)} = |z|/i$. A proof of this assertion appears in an Appendix.

We note from the definition (3.5) of $h_i(z)$ and from Liapounov's inequality (Loéve [3; p. 172]) that

$$h_i(z) \geq h_{i+1}(z), \quad \text{all } i \geq 2. \tag{3.12}$$

Applying (3.8) and (3.12), we find that

$$(i-1)h_i^2(z) + zh_i(z) - 1 \leq 0 \leq ih_i^2(z) + zh_i(z) - 1,$$

from which we conclude that

$$\frac{1}{2}\left(\frac{-z}{i} + \sqrt{\left(\frac{z}{i}\right)^2 + \frac{4}{i}}\right) \leq h_i(z) \leq \frac{1}{2}\left(\frac{-z}{i-1} + \sqrt{\left(\frac{z}{i-1}\right)^2 + \frac{4}{i-1}}\right). \tag{3.13}$$

Let

$$\hat{\theta}_{B-} = \frac{u^{1/2}}{2}\left(\frac{-z}{n+r-1} + \sqrt{\left(\frac{z}{n+r-1}\right)^2 + \frac{4}{n+r-1}}\right), \tag{3.14}$$

$$\hat{\theta}_{B+} = \frac{u^{1/2}}{2}\left(\frac{-z}{n+r-2}\right) + \sqrt{\left(\frac{z}{n+r-2}\right)^2 + \frac{4}{n+r-2}}.$$

It follows from (3.5) and (3.13) that

$$\hat{\theta}_{B-} \leq \hat{\theta}_B \leq \hat{\theta}_{B+}. \tag{3.15}$$

As can be seen by inspection of (3.14), when n+r is large, $\hat{\theta}_{B-}$ and $\hat{\theta}_{B+}$ are very close to one another in value, and hence very close in value to $\hat{\theta}_B$. We note that when w = 0 and r = 1, $\hat{\theta}_{B-} = \hat{\theta}_{MLE}$, while when w = 0 and r = 2, $\hat{\theta}_{B+} = \hat{\theta}_{MLE}$, where

$$\hat{\theta}_{MLE} = \frac{-\bar{x} + [4as^2 + (1+4a)(\bar{x})^2]^{1/2}}{2a} \tag{3.16}$$

is the maximum likelihood estimator of θ. Further, it is easily shown that for fixed w, r, and θ,

$$\underset{n\to\infty}{\text{plim}} \sqrt{n}\,(\hat{\theta}_{B-} - \hat{\theta}_{MLE}) = \underset{n\to\infty}{\text{plim}} \sqrt{n}(\hat{\theta}_{B+} - \hat{\theta}_{MLE}) = 0 \tag{3.17}$$

Since it is known that $\hat{\theta}_{MLE}$ is B.A.N., it follows from (3.17) that $\hat{\theta}_{B-}$ and $\hat{\theta}_{B+}$ are likewise B.A.N. Finally, application of (3.15) allows us to conclude that $\hat{\theta}_B$ is B.A.N.

The risk functions of $\hat{\theta}_{B-}$, $\hat{\theta}_{B+}$, $\hat{\theta}_{MLE}$, and $\hat{\theta}_B$ cannot be expressed in closed form. However, since $\hat{\theta}_B$ is known to be admissible (provided w > 0, r > 0), we know that there must be a region of values of θ for which $R(\theta,\hat{\theta}_B) < R(\theta,\hat{\theta}_{LMMS})$. The shape of the prior distribution (3.1) suggests that this region of θ-values lies near θ = 0. Since $\hat{\theta}_{B-}$ and $\hat{\theta}_{B+}$ are close to $\hat{\theta}_B$ when n+r is large, we would expect that both $\hat{\theta}_{B-}$ and $\hat{\theta}_{B+}$ would also improve upon $\hat{\theta}_{LMMS}$ in some region of θ-values, when moderate sample sizes are used.

APPENDIX

For positive constants a_0, a_1, $a_2, \ldots, a_m$, we define a finite continued fraction to be

$$c_m \equiv [a_0, a_1, \ldots, a_m] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ldots + \cfrac{1}{a_m}}},$$

while for the sequence a_0, a_1, $a_2, \ldots$ of positive constants, a regular continued fraction is defined to be

$$c \quad [a_0, a_1, a_2, \ldots] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ldots}} = \lim_{m \to \infty} c_m,$$

provided the limit of c_m exists. From Khinchin [2; pp. 6, 10], we have the following results.

<u>Theorem</u>. For c_m to converge to c, it is necessary and sufficient that the series $\sum_{k=0}^{\infty} a_k$ diverge. Further,

(a) $c_i \leq c_{i+2}$, if i is even,

(b) $c_j \geq c_{j+2}$, if j is odd,

(c) $c_i \leq c_j$, if i is even and j is odd.

To demonstrate that (3.10) holds when $z < 0$, we note from (3.8) that

$$h_i(z) = \frac{1}{ih_{i+1}} - \frac{z}{i} = \frac{1}{i\left(\frac{1}{(i+1)h_{i+2}} - \frac{z}{i+1}\right)} - \frac{z}{i} = \ldots .$$

Continuing in this fashion, we find that for all $m > 1$,

$$h_i = [a_0^{(i)}, a_1^{(i)}, \ldots, a_{m-1}^{(i)}, a_m^{(i)}(1-(zh_{i+m}(z))^{-1})]. \tag{A.1}$$

where $a_0^{(i)}, a_1^{(i)}, \ldots, a_m^{(i)}$ are defined by (3.11). It can be shown that the series defined by (3.11) diverges. Thus, our Theorem shows that

$$\lim_{m\to\infty} [a_0^{(i)}, a_1^{(i)}, \ldots, a_m^{(i)}] \equiv \lim_{m\to\infty} c_m^{(i)} = c^{(i)} \tag{A.2}$$

exists. Therefore, given $\varepsilon > 0$, there exists M_ε such that $m > M_\varepsilon$ implies that

$$|c_{m-2}^{(i)} - c_{m-1}^{(i)}| \leq \frac{\varepsilon}{2}, \; |c_{m-1}^{(i)} - c^{(i)}| \leq \frac{\varepsilon}{2}. \tag{A.3}$$

Choose $m > M_\varepsilon$ to be even. Then by our Theorem, part (a), the fact that $z < 0$, and (A.1), we have

$$h_i(z) \geq c_{m-2}^{(i)}, \tag{A.4}$$

and similarly by part (b) of our Theorem,

$$h_i(z) \leq c_{m-1}^{(i)}. \tag{A.5}$$

However, by part (c) of the Theorem,

$$c_{m-2}^{(i)} \leq c_{m-1}^{(i)}, \tag{A.6}$$

and thus by (A.3) - (A.6),

$$\begin{aligned} |h_i(z) - c^{(i)}| &\leq |h_i(z) - c_{m-1}^{(i)}| + |c_{m-1}^{(i)} - c^{(i)}| \\ &\leq |c_{m-1}^{(i)} - c_{m-2}^{(i)}| + |c_{m-1}^{(i)} - c^{(i)}| \leq \varepsilon. \end{aligned}$$

Since ε is arbitrary, we conclude that $h_i(z) = c^{(i)}$.

REFERENCES

[1] Khan, R. A., "A Note on Estimating the Mean of a Normal Distribution With Known Coefficient of Variation," Journal of the American Statistical Association 63 (September, 1968), 1039-1041.

[2] Khinchin, A. Ya., Continued Fractions, Chicago and London: University of Chicago Press, 1964.

[3] Loève, Michel, Probability Theory, Princeton: D. Van Nostrand Company, Inc., 1960.