Estimating a Ratio of Normal Parameters
By
Andrew Rukhin

Technical Report #84-25

Department of Statistics
Purdue University

July, 1984
Revised May, 1985

Research supported by NSF Grant MCS-8101670


Key words and phrases: point estimation, normal parameters, quadratic loss, admissibility, generalized Bayes estimators, coefficient of variation
Summary

Estimating a Ratio of Normal Parameters

The estimation problem of a function of normal parameters $\xi/\sigma^2p$ is considered. We prove that a "natural" estimator of this ratio is admissible for quadratic loss if and only if $p$ is nonnegative.
1. Introduction and Summary.

Let $x_1, \ldots, x_n$, $n > 2$, be a random sample from a normal distribution with mean $\xi$ and variance $\sigma^2$. We consider the statistical estimation problem of a parametric function $\theta = \xi/\sigma^2 p$ where $p$ is a given real number. The most interesting case is when $p = 1/2$, so that $\theta$ is a dimensionless characteristic, which is a reciprocal of a commonly used coefficient of variation. The case of general $p$ presents certain interest from the point of view of statistical decision theory. Since the admissibility proof for $p = 1/2$ is not any easier we consider the general situation. Notice also that if $p = 1$, $\theta$ is a natural parameter in the corresponding exponential family. The tests of the hypothesis about $\theta$ with invariant power function have been studied by Linnik (1968), Theorem 3.3.1, to whose memory the author would like to dedicate this work.

In this paper we use quadratic loss function of the form

$$L(\xi, \sigma; \delta) = (\theta - \delta)^2 \sigma^{4p-2}$$

which is invariant under scale transformations. If $X = \frac{1}{n} \sum_{i=1}^{n} x_i$ and $S = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{X})^2$, then $(X, S)$ is a version of the complete sufficient statistic, and the unbiased estimator $\hat{\theta}(X, S)$ has the form

$$\hat{\theta}(X, S) = c_1 X / S^p,$$

where $c_1 = 2p \frac{\Gamma((n-1)/2)}{\Gamma((n-1)/2 - p)}$.

However for $p \neq 0$ and sufficiently large $n$ this estimator can be improved upon very easily. Indeed the risk function of any procedure $cX/S^p$ depends only on $\xi/\sigma = n$, and because of the independence of $X$ and $S$ it has the form

$$E_n(cX/S^p - n)^2 = n^2(c^2 E_0 S^{-2p} - 2c E_0 S^{-p} + 1) + c^2 n^{-1} E_0 S^{-2p}.$$

Assuming that $n > 4p + 1$ (otherwise the risk is infinite) we put
\[ c_0 = \frac{E_{01} S^{-p}}{E_{01} S^{-2p}} = \frac{2^{p} \Gamma((n-1)/2-p)}{\Gamma((n-1)/2-2p-p)}. \]  

(1.1)

If \( c > c_0 \) then the estimator \( cX/SP \) has larger risk than \( \delta_0(X,S) = c_0X/SP \). Since \( c_1 > c_0 \) the estimator \( \delta_0 \) is better than the unbiased estimator \( \delta_U \). Moreover \( \delta_0 \) is admissible within the class of procedures \( cX/SP \), and it can be shown to be minimax for properly rescaled quadratic loss function. Therefore we shall investigate the admissibility of \( \delta_0 \). We prove that \( \delta_0 \) is admissible if \( p > 0 \) and inadmissible if \( p < 0 \). The admissibility part seems to be surprising in view of the inadmissibility of \( c_0/SP \) as an estimator of \( \sigma^{-p} \) (see Stein (1965), Brown (1968), Brewster and Zidek (1974), Strawderman (1974)). In Section 2 we show that \( \delta_0 \) is generalized Bayes procedure with respect to a prior which admits a good approximation (in terms of posterior risk) by probability priors. Exactly this fact is known to be responsible for admissibility (see Stein (1965)), and our admissibility proof in Section 3 is just a modification of the standard one (Blyth (1951)). Similar admissibility phenomenon also happens in the estimation problem of exponential parameters (see Rukhin (1983)) and other functions of normal parameters (Rukhin (1984a), (1985)).

2. Generalized Bayes Estimators of \( \theta \).

We assume in this section that \( p \) is a positive number. We also use a convenient reparameterization: \( \tau = 1/(2\sigma^2) \). Let \( \lambda(\xi,\tau) \) be a density of a (generalized) prior distribution over \( \{ (\xi,\tau), \tau > 0 \} \) with respect to invariant measure \( d\xi \, d\tau/\tau \). Notice that the latter corresponds to the right Haar measure for \( \xi \) and \( \sigma \), which is traditionally used. The Bayes estimator \( \delta_B(x,s) \) has the form
\[
\delta_B(x,s) = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \xi^n \exp\{-n(x-\xi)^2 - \tau s\} \lambda(\xi,\tau) d\xi d\tau}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \xi^n \exp\{-n(x-\xi)^2 - \tau s\} \lambda(\xi,\tau) d\xi d\tau}.
\]

Denote

\[
R(x,y) = \int_{-\infty}^{\infty} \exp\{-ny(x-\xi)^2\} \lambda(\xi,y) d\xi.
\]  \tag{2.1}

Then (assuming the convergence of needed integrals)

\[
\int_{-\infty}^{\infty} \xi \exp\{-ny(x-\xi)^2\} \lambda(\xi,y) d\xi = xR(x,y) + R_x(x,y)/(2ny),
\]

where \(R_x(x,y) = \frac{\partial}{\partial x} R(x,y)\). Thus

\[
\delta_B(x,s) = c_0 x/s \quad + \quad 2^p \int_{0}^{\infty} [xR_x/(2ny) - 2^{-p} c_0 x s^{-p_y - p_R} y^{n/2 - p_e - ys} \quad dy/\int_{0}^{\infty} R_y^{n/2 - 2p_e - ys} \quad dy. \tag{2.2}
\]

Notice that the representation (2.2) holds also for some prior distributions which are not absolutely continuous in which case \(\lambda\) should be interpreted as a generalized function.

According to (2.2) \(\delta_0\) is a generalized Bayes procedure with respect to prior density \(\lambda\) if and only if

\[
I = \int_{0}^{\infty} [xR_x/(2ny) - 2^{-p} c_0 x s^{-p_y - p_R} y^{n/2 - p_e - ys} \quad dy = 0 \tag{2.3}
\]

The equation (2.3) has a "trivial" solution \(R(x,y) = y^{-3/2}\) which corresponds to the conjugate (not uniform) generalized prior density \(\lambda(\xi,\tau) = \tau^{-1}\). However (2.3) admits many other solutions. Let

\[
R(x,y) = \sum_{k=0}^{\infty} r_k (nx^2y)^k/k!,
\]
where \( r_0 = 1 \) and for \( k = 1, 2, \ldots \)

\[
    r_k = \prod_{j=1}^{k} \left[ B(j+n/2-2p,p)/B((n-1)/2-2p,p)-1 \right].
\]  

(2.5)

Term by term, differentiation and integration yields \( I = 0 \). In the appendix we show that \( \{-1\}^k r_k, k = 0, 1, \ldots \) is a moment sequence of a distribution function \( G = G_p \) over the interval \([0, 1]\), i.e. for \( k = 0, 1, \ldots \)

\[
    \int_0^1 t^k dG(t) = (-1)^k r_k.
\]

The distribution function \( G_p \) is continuous for \( p < 1 \) and for \( p > 1 \) it is continuous for \( 0 < t < 1 \) and has a positive jump \( \bar{g} = \bar{g}_p \) at \( t = 1 \).

Now define the generalized prior distribution \( \Lambda_0 \) by the formula

\[
    d\Lambda_0 = (1 - \bar{g})\lambda(\xi, t)d\xi d\tau/\tau + \bar{g} d\varepsilon
\]  

(2.6)

where

\[
    \lambda(\xi, \tau) = (n_\xi/\pi)^{1/2} \int_0^1 \exp\{-n_\tau\xi^2 t/(1-t)\}(1-t)^{-1/2} dG(t).
\]

In this case

\[
    R(x, y) = (ny/\pi)^{1/2} \int_{-\infty}^{\infty} \exp\{-ny\xi^2 t/(1-t) - ny(x-\xi)^2\}d\xi \int_0^1 (1-t)^{-1/2}dG(t)
\]

\[
    = \int_0^1 \exp\{-ny^2 t\}dG(t)
\]

\[
    = \sum_{k=0}^{\infty} r_k(nx^2 y)^k/k!
\]

(2.7)

Therefore \( R \) solves (2.3) and \( \delta_0 \) is the Bayes estimator with respect to the prior distribution \( \Lambda_0 \).
We have proven the following

Theorem 1. The generalized Bayes estimator $\delta_B$ of $\theta = \xi/\sigma^{2p} = \xi(2\tau)^p$ under quadratic loss and prior density $\lambda$ has form (2.2) where the function $R$ is defined by (2.1). The estimator $\delta_0$ is generalized Bayes for the prior distribution (2.6).

Notice that the support of the measure $\Lambda_0$ is the whole parameter space ${(\xi, \tau), \tau > 0}$. Also although $\Lambda_0$ is not a finite measure, for all $\tau$

$$\int_{-\infty}^{\infty} \lambda(\xi, \tau) d\xi < \infty,$$

so that $\Lambda_0$ is "less improper" than the uniform distribution.

In terms of parameters $n = \xi/\sigma$ and $\sigma$, $d\Lambda_0(\xi, \tau) = g(n) \, dn \, d\sigma/\sigma$, where

$$g(n) = \frac{1}{\Gamma(1-t)^{-1/2}} \int_{0}^{\infty} \exp(-n^2 t/(1-t)) (1-t)^{-1/2} \, dG(t).$$

It can be shown that as $n \to \infty$

$$g(n) \sim C|n|^{2p-n/2-1}.$$

Notice that $2p-n/2-1<-3/2$, so that $g$ possesses the properties of generalized prior densities for admissible scale equivariant estimators of normal variance discussed in Brown (1979) p. 991. Similar (but different from $g$) prior densities have been also used by Brewster and Zidek (1974) and Strawderman (1974).

3. Admissibility Result.

Theorem 2. The procedure $\delta_0(X,S) = c_0X/SP$ based on a random sample of size $n$, $n > 4p+1$, is admissible for estimating the ratio of normal parameters, $\theta = \xi/\sigma^{2p}$, under quadratic loss if $p > 0$. 
Proof. If \( p = 0 \) the estimator \( \delta_0 \) coincides with \( X \) which is known to be an admissible estimator of \( \theta = \xi \). Therefore we assume \( p > 0 \).

Because of the continuity of the risk functions the admissibility of \( \delta_0 \) will be proven if one finds a sequence \( h_m, m = 1, 2, \ldots, \) of positive measurable functions \( h_m \rightarrow 1 \) as \( m \rightarrow \infty \), \( \int h_m(\tau)\,d\Lambda_0(\xi, \tau) < \infty \), and

\[
\rho_m = \int_{-\infty}^{\infty} \int_{0}^{\infty} \tau^{1-2p}\left[ E(\delta_0 - \theta)^2 - E(\delta_m - \theta)^2 \right] h_m(\tau)\,d\Lambda_0(\xi, \tau) + 0.
\]

Here \( \delta_m \) is the Bayes estimator for the proper prior distribution \( \Lambda_m, d\Lambda_m = h_m\,d\Lambda \).

A straightforward calculation shows that

\[
\rho_m = C \int_{-\infty}^{\infty} \int_{0}^{\infty} \left[ \delta_m(x, s) - \delta_0(x, s) \right]^2 \,dx\,ds
\]

\[
\int_{-\infty}^{\infty} \int_{0}^{\infty} y^{n/2-2p+1} \exp\left[ -ny(x-\xi)^2 - ys(n-3)/2 \right] d\Lambda_m(\xi, y).
\]

Here \( C \) denotes a generic positive constant which depends only on \( n \) and \( p \).

Calculation similar to this done in (2.7) shows that with

\[
R(x, y) = R_m(x, y) = \int_{0}^{1} \exp\left[ -nyx^2t \right] dG(t)h_m(y)
\]

one has

\[
\rho_m = C \int_{-\infty}^{\infty} \int_{0}^{\infty} \frac{\left[ \int_{0}^{\infty} [R+R_X/(2nxy) - 2^{-p-1} s^{-p-1} y^{p-n/2} e^{-sy} dy] \right]^2}{\int_{0}^{\infty} y^{n/2-2p} R e^{-sy} dy} \,x^2 \,dx \,s(n-3)/2 \,ds.
\]

Assume that

\[
\int_{0}^{\infty} h_m(y) \,dy/y < \infty.
\]

Then

\[
\int_{-\infty}^{\infty} \int_{0}^{\infty} h_m(\tau)\,d\Lambda_0(\xi, \tau) = \int_{0}^{1} t^{1-1/2} dG(t) \int_{0}^{\infty} h_m(y) \,dy/y < \infty.
\]
A straightforward calculation shows that

\[
\left[ \int_0^\infty \left[ R + R_x/(2nx) - 2^p c_0(sy)^{-pR} \right] y^{n/2-p} e^{-sy} dy \right]^2
\]

\[
= \left[ \int_0^\infty \int_0^1 v^{n/2-2p(1-v)p^{-1}} e^{-nxy^2tv} [h_m(y) - h_m(yv)] dv \right] \left. dG(t) y^{n/2-p} e^{-sy} dy \right] / B((n-1)/2 - 2p, p)
\]

\[
< \int_0^\infty \int_0^1 v^{n/2-2p(1-v)p^{-1}} e^{-nxy^2tv} h_m(y)
\]

\[
[1-h_m(yv)/h_m(y)]^2 (1-vt) dG(t) e^{-syt} y^{n/2} dy
\]

\[
\times \int_0^\infty \int_0^1 v^{n/2-2p(1-v)p^{-1}} e^{-nxy^2tv} h_m(y) (1-vt)^{-1} dG(t) e^{-syt} y^{n/2-2p} dy.
\]

Because of Lemma 2

\[
\rho_m < C \int_0^\infty \int_0^1 v^{n/2-2p(1-v)p^{-1}} e^{-nxy^2tv}
\]

\[
[1-h_m(yv)/h_m(y)]^2 dG(t) h_m(y) e^{-syt} y^{n/2} dy x^2 dx s^{(n-3)/2} ds
\]

\[
= C \int_0^\infty \int_0^1 v^{n/2-2p(1-v)p^{-1}} h_m(y) [1-h_m(yv)/h_m(y)]^2 dv dy / y.
\]

Here we have used the fact that \( \int_0^1 t^{-3/2} dG(t) < \infty \).

Now we define the sequence:

\[
h_m(y) = (1 + (\log y)^2 m^{-2})^{-1}
\]

so that (3.1) is satisfied. Also

\[
\int_0^\infty h_m(y) [1-h_m(yv)/h_m(y)]^2 dy / dy
\]

\[
= \int_0 [h_m^2(y)/h_m(yv^{-1})] dy / y - \int_0 h_m(y) dy / y
\]

\[
= (\log v)^2 m^{-2} \int_0 h_m(y) dy / y = C (\log v)^2 m^{-1}
\]
and

\[ \rho_m \leq C m^{-1} \to 0 \]
as \( m \to \infty \).

For completeness sake we formulate the following result proof of which can be found in Rukhin (1984b).

Theorem 3. The estimator \( \delta_0 \) is inadmissible for estimating \( \theta \) under quadratic loss if \( p < 0 \). In fact the estimator \( \delta \) of the form

\[ \delta(x,s) = \delta_0(x,s) - 2\delta_0(x,s)h(|x(n/s)|^{1/2}) \]

where

\[ h(z) = \max\{0,1-(1+z^2)^{1-p}z^{-2d_n}\} \]

\[ d_n = \max_{k \geq 0} k B(n/2+k-p,-p)/B((n-1)/2-p,-p) \]

improves upon \( \delta_0 \).

Acknowledgement. The author is grateful to Bill Strawderman for interesting discussion. Thanks are also due to the referee for many helpful remarks. In particular he pointed out a gap in the original proof of Theorem 2.
Appendix

We use here the following notation:

\[ b = B(((n-1)/2-2p), p), \quad b_j = b^{-1} B(j+n/2+1-2p, p). \]

Lemma 1. For any \( p, 0 < p < (n-1)/4 \), and \( n \geq 2 \), there exists a distribution function \( G(t) \) over the interval \([0,1]\) such that for all \( k = 0, 1, 2 \ldots \)

\[ \int_0^1 t^k dG(t) = \sum_{j=1}^k (1-b_j) = (-1)^k \gamma_k. \quad (5.1) \]

Also

\[ \int_0^1 t^{-3/2} dG(t) < \infty. \]

For \( p < 1 \), the function \( G \) is continuous. For \( p > 1 \), this function is continuous for \( 0 < t < 1 \) and has a positive jump \( \bar{g} \) at \( t = 1 \),

\[ \bar{g} = \prod_{j=0}^\infty (1-b_j). \]

Proof. Let for complex \( z \)

\[ \phi(z) = \prod_{j=0}^\infty (1-b_j)(1-B(j+z+n/2+1-2p, p)b^{-1}). \]

It is easy to see that \( \phi(z) \) is an analytic function of \( z \) in the region \( \Re z > 2p - 1 - n/2 \).

We show that \( \phi(z) \) is a Mellin transform of a distribution \( G \), i.e.

\[ \int_0^1 t^z dG(z) = \phi(z). \]
Since \( \phi(k) = (-1)^k r_k, \ k = 0, 1, \ldots, \) the first part of Lemma 1 will then be proven.

If distributions \( F_j, \ j = 0, 1, \ldots, \) are such that

\[
\int_0^1 x^z dF_j(x) = (1-b_j)[1-B(j+z+n/2+1-2p,p) b^{-1}]^{-1}
\]

(5.2)

and if infinite multiplicative convolution of \( F_j, \ j = 0, 1, \ldots, \) exists, then it satisfies (5.1).

First let us demonstrate the existence of such distributions \( F_j. \) Denote

\[
\phi_m(u) = \int \ldots \int \frac{\phi_1(u/(u_1 \ldots u_{m-1}))}{u_j} \prod_{j=1}^{m-1} \phi_1(u_j) du_j / u_j
\]

where integration is over the set

\[
A_u = \{(u_1, \ldots, u_{m-1}), \ 0 < u_i < 1, \ i = 1, \ldots, m-1, u_1 \ldots u_{m-1} > u\}
\]

and

\[
\phi_1(u) = u^{n/2-2p}(1-u)^{p-1}, \quad 0 < u < 1.
\]

Thus

\[
\int_0^1 u^{z-1} \phi_m(u) du = \left[ \int_0^1 u^{z-1} \phi_1(u) du \right]^m
\]

\[
= [B(n/2 - 2p + z, p)]^m.
\]
Notice that since \((1-u)^{p/2} < 1\) for \(0 < u < 1\)

\[
\phi_m(u) \leq u^{n/2-2p} \prod_{j=1}^{m-1} A_j \frac{du_j}{u_j}
\]

\[
= u^{n/2-2p}(-\log u)^{m-1}/(m-1)!
\]  

(5.2)

Now let

\[
\phi(u) = \sum_{m=1}^{\infty} \phi_m(u) b^{-m}.
\]

This series converges. Indeed because of (5.2)

\[
\phi(u) \leq u^{n/2-2p} \sum_{m=1}^{\infty} (-\log u)^{m-1}/[b^m(m-1)!]
\]

\[
= u^{n/2-2p} \exp(-\log u/b)/b.
\]

For positive \(z\)

\[
\int_0^1 u^{z-1} \phi(u)du
\]

\[
= \sum_{m=1}^{\infty} [B(z+n/2-2p, p)b^{-1}]^m
\]

\[
= B(z+n/2-2p, p)b^{-1} [1-B(z+n/2-2p, p)b^{-1}]^{-1}.
\]
Now if
\[ dF_j(u) = (1-b_j)(u^j \phi(u)du + d\varepsilon_1(u)), \]
where the distribution \( \varepsilon_1 \) puts unit mass at \( u = 1 \), then (5.2) is valid.

Notice that the infinite convolution of \( F_1, F_2, \ldots \) converges since, as is easy to check,
\[ \sum_j \int_0^1 \log u \, dF_j < \infty \]
and
\[ \sum_j \int_0^1 (\log u - \int_0^1 \log u \, dF_j)^2 \, dF_j < \infty, \]
which in our case is a sufficient condition for the convolution's convergence (see Hennequin and Tortrat (1965) p.202).

Thus the distribution function \( G \) such that \( \phi(z) \) is its Mellin transform exists. Since \( \phi(z) \) is analytic for \( \text{Re } z > -3/2 \), \( \int_0^1 t^{-3/2} dG(t) < \infty \).

It is known (cf. Hennequin and Tortrat (1965) p.205) that \( G \) is continuous if \( u^\infty_{j=0} f_k \), where \( f_j \) is the largest jump of \( F_j \), vanishes. Clearly
\[ f_j = 1 - b_j \]
and
\[ \log f_j \sim - \Gamma((n - 1)/2 - p)/[j^{1/2}(n - 1)/2 - 2p]]. \]
Thus $\prod_{j=0}^{\infty} f_j$ vanishes if and only if $p < 1$. If $p > 1$, the distribution function $G$ has a jump $\bar{g} = \prod_{j=0}^{\infty} f_j$ at $t = 1$, and is continuous for $t > 1$.

Lemma 2. If the distribution function $G$ is defined as in Lemma 1, then

$$\int_0^1 e^{zt}dG(t)$$

$$= \int_0^1 \int_0^1 \frac{1}{v^{n/2-2p(1-v)p-1}}e^{ztv(1-tv)^{-1}}dG(t)dv b^{-1}. \tag{5.4}$$

This formula is proved by expanding both sides of (5.4) in powers of $z$. 
References


