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

AMS 1980 subject classifications: Primary 62F10, Secondary 62C15, 62C20, 62F11

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 ξ/σ^{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 $x_1,\ldots,x_n$, $n \geq 2$, be a random sample from a normal distribution with mean ξ and variance σ^2. We consider the statistical estimation problem of a parametric function $\theta = \xi/\sigma^{2p}$ where p is a given real number. The most interesting case is when $p = 1/2$, so that θ 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 $p = 1$, θ is a natural parameter in the corresponding exponential family. The tests of the hypothesis about θ 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 $X = \sum_1^n x_j/n$ and $S = \sum_1^n (x_j - X)^2$, then (X,S) is a version of the complete sufficient statistic, and the unbiased estimator $\delta_U(X,S)$ has the form

$$\delta_U(X,S) = c_1 X/S^p ,$$

where $c_1 = 2^p\,\Gamma((n-1)/2)/\,\Gamma((n-1)/2 - p)$.

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 $\xi/\sigma = \eta$, and because of the independence of X and S it has the form

$$E_{\eta 1}(cX/S^p - \eta)^2 = \eta^2[c^2\, E_{01}\, S^{-2p} - 2cE_{01}\, S^{-p} + 1] + c^2 n^{-1} E_{01} S^{-2p} .$$

Assuming that $n > 4p + 1$ (otherwise the risk is infinite) we put

$$c_0 = \frac{E_{01}S^{-p}}{E_{01}S^{-2p}} = \frac{2^p\Gamma((n-1)/s-p)}{\Gamma((n-1)/2-2p-p)}. \tag{1.1}$$

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 δ_0 is better than the unbiased estimator δ_U. Moreover δ_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 δ_0. We prove that δ_0 is admissible if $p \geq 0$ 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 σ^{-p} (see Stein (1965), Brown (1968), Brewster and Zidek (1974), Strawderman (1974)). In Section 2 we show that δ_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 θ.

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 ξ and σ, which is traditionally used. The Bayes estimator $\delta_B(x,s)$ has the form

$$\delta_B(x,s) = \frac{2^p \int_{-\infty}^{\infty} \int_0^{\infty} \xi\tau^{n/2-p} \exp\{-n\tau(x-\xi)^2-\tau s\}\lambda(\xi,\tau)d\xi d\tau}{\int_{-\infty}^{\infty} \int_0^{\infty} \tau^{n/2-2p} \exp\{-n\tau(x-\xi)^2-\tau s\}\lambda(\xi,\tau)d\xi d\tau} .$$

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

$$\delta_B(x,s) = c_0 x/s^p$$

$$+ 2^p \int_0^{\infty} [xR + R_x/(2ny) - 2^{-p} c_0 xs^{-p}y^{-p}R]y^{n/2-p}e^{-ys} dy/\int_0^{\infty} Ry^{n/2-2p}e^{-ys} dy. \tag{2.2}$$

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

According to (2.2) δ_0 is a generalized Bayes procedure with respect to prior density λ if and only if

$$I = \int_0^{\infty} [xR + R_x/(2ny) - 2^{-p} c_0 xs^{-p}y^{-p}R]y^{n/2-p}e^{-ys} dy = 0 \tag{2.3}$$

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 $k = 1,2,\ldots$

$$r_k = \prod_{j=1}^{k} [B(j+n/2-2p,p)/B((n-1)/2-2p,p)-1]. \tag{2.5}$$

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 $p \leq 1$ and for $p > 1$ it is continuous for $0 \leq t < 1$ and has a positive jump $\overline{g} = \overline{g}_p$ at $t = 1$.

Now define the generalized prior distribution Λ_0 by the formula

$$d\Lambda_0 = (1 - \overline{g})\lambda(\xi,\tau)d\xi d\tau/\tau + \overline{g}\ d\varepsilon \tag{2.6}$$

where

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

In this case

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

$$= \int_0^1 \exp\{-nyx^2 t\}dG(t)$$

$$= \sum_{k=0}^{\infty} r_k(nx^2y)^k/k! \tag{2.7}$$

Therefore R solves (2.3) and δ_0 is the Bayes estimator with respect to the prior distribution Λ_0.

We have proven the following

Theorem 1. The generalized Bayes estimator δ_B of $\theta = \xi/\sigma^{2p} = \xi(2\tau)^p$ under quadratic loss and prior density λ has form (2.2) where the function R is defined by (2.1). The estimator δ_0 is generalized Bayes for the prior distribution (2.6).

Notice that the support of the measure Λ_0 is the whole parameter space $\{(\xi,\tau),\ \tau > 0\}$. Also although Λ_0 is not a finite measure, for all τ

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

so that Λ_0 is "less improper" than the uniform distribution.

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

$$g(\eta) = \int_0^1 \exp\{-n\eta^2 t/(1-t)\}\ (1-t)^{-1/2}\ dG(t).$$

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

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

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 $\delta_0(X,S) = c_0X/S^p$ 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 $p \geqslant 0$.

Proof. If $p = 0$ the estimator δ_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 δ_0 will be proven if one finds a sequence h_m, $m = 1,2,\ldots$, of positive measurable functions $h_m \to 1$ as $m \to \infty$, $\iint h_m(\tau)d\Lambda_0(\xi,\tau) < \infty$, and

$$\rho_m = \int_{-\infty}^{\infty}\int_0^{\infty} \tau^{1-2p}[E(\delta_0-\theta)^2 - E(\delta_m-\theta)^2]h_m(\tau)d\Lambda_0(\xi,\tau) \to 0.$$

Here δ_m is the Bayes estimator for the proper prior distribution Λ_m, $d\Lambda_m = h_m d\Lambda$.

A straightforward calculation shows that

$$\rho_m = C \int_{-\infty}^{\infty} \int_0^{\infty} [\delta_m(x,s) - \delta_0(x,s)]^2 \, dx \, ds$$

$$\int_{-\infty}^{\infty}\int_0^{\infty} y^{n/2-2p+1} \exp\{-ny(x-\xi)^2-ys\} s^{(n-3)/2} d\Lambda_m(\xi,y).$$

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

$$R(x,y) = R_m(x,y) = \int_0^1 \exp\{-nyx^2t\}dG(t)h_m(y)$$

one has

$$\rho_m = C \int_{-\infty}^{\infty}\int_0^{\infty} \frac{\{\int_0^{\infty}[R+R_x/(2nxy) - 2^{-p}c_0\, s^{-p}y^{-p}R]y^{p-n/2}e^{-sy}dy\}^2}{\int_0^{\infty} y^{n/2-2p}\, R\, e^{-sy}\, dy}$$

$$x^2 \, dx \, s^{(n-3)/2} \, ds.$$

Assume that

$$\int_0^{\infty} h_m(y) \, dy/y < \infty. \tag{3.1}$$

Then

$$\int_{-\infty}^{\infty}\int_0^{\infty} h_m(\tau)d\Lambda_0(\xi,\tau) = \int_0^1 t^{-1/2}dG(t)\int_0^{\infty} h_m(y)dy/y < \infty.$$

A straightforward calculation shows that

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

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

$$\left./B((n-1)/2 - 2p,p)\right]^2$$

$$\leq \int_0^\infty \int_0^1 \int_0^1 v^{n/2-2p}(1-v)^{p-1} e^{-nyx^2tv} h_m(y)$$

$$[1-h_m(yv)/h_m(y)]^2(1-vt)dvdG(t)e^{-sy}y^{n/2}dy$$

$$\times \int_0^\infty \int_0^1 \int_0^1 v^{n/2-2p}(1-v)^{p-1} e^{-nyx^2tv} h_m(y)(1-vt)^{-1}dvdG(t)e^{-sy}y^{n/2-2p}dy.$$

Because of Lemma 2

$$\rho_m \leq C\int_{-\infty}^{\infty}\int_0^\infty \int_0^\infty \int_0^1 \int_0^1 v^{n/2-2p}(1-v)^{p-1}e^{-nyx^2tv}$$

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

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

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

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

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

$$= (\log v)^2\ m^{-2} \int_0^\infty h_m(y)\ dy/y = C(\log v)^2 m^{-1}$$

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 δ_0 is inadmissible for estimating θ under quadratic loss if $p < 0$. In fact the estimator δ of the form

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

where

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

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

improves upon δ_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:

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

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 $k = 0,1,2\ldots$

$$\int_0^1 t^k dG(t) = \prod_{j=1}^{k} (1-b_j) = (-1)^k r_k. \tag{5.1}$$

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 $\mathrm{Re}\, z > 2p - 1 - n/2$.

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

$$\int_0^1 x^z dF_j(x) = (1-b_j)[1-B(j+z+n/2+1-2p,p)\, b^{-1}]^{-1} \tag{5.2}$$

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

$$\phi_m(u) = \underset{A_u}{\int\ldots\int} \phi_1(u/(u_1\ldots u_{m-1})) \prod_{j=1}^{m-1} \phi_1(u_j)du_j/u_j$$

where integration is over the set

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

and

$$\phi_1(u) = u^{n/2-2p}(1-u)^{p-1}, \qquad 0 < u < 1.$$

Thus

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

$$= [B(n/2 - 2p + z,\ p)]^m.$$

Notice that since $(1 - u)^{p/2} < 1$ for $0 < u < 1$

$$\phi_m(u) \leq u^{n/2-2p} \int \underset{A_u}{\cdots} \int \prod_{1}^{m-1} du_j/u_j \tag{5.2}$$
$$= u^{n/2-2p}(-\log u)^{m-1}/(m-1)!\ .$$

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 ε_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

$$\sum_j \int_0^1 \left(\log u - \int_0^1 \log u \; dF_j\right)^2 dF_j < \infty,$$

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 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 $\Pi_{j=0}^{\infty} f_k$, where f_j is the largest jump of F_j, vanishes. Clearly

$$f_j = 1 - b_j$$

and

$$\log f_j \sim - \Gamma((n - 1)/2 - p)/[j^p \Gamma((n - 1)/2 - 2p)].$$

Thus $\Pi_{j=0}^{\infty} f_j$ vanishes if and only if $p \leqslant 1$. If $p > 1$, the distribution function G has a jump $\bar{g} = \Pi_{j=0}^{\infty} f_j$ at $t = 1$, and is continuous for $t > 1$.

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

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

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

References

Blyth, C. R. (1951. On minimax statistical decision procedures and their admissibility, Ann. Math. Statist. 22, 22-42.

Brewster, J. F. and Zidek, J. V. (1974). Improving on equivariant estimators, Ann. Statist. 2, 21-38.

Brown, L. (1968). Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. Ann. Math. Statist. 39, 29-48.

Brown, L. (1979). A heuristic method for determining admissibility of estimators -- with applications. Ann. Statist. 7, 960-994.

Hennequin, P. L. and Tortrat, A. (1965). Theorie des Probabilites et Quelques Applications. Masson et Cie Edit., Paris.

Linnik, Yu. V. (1968). Statistical Problems with Nuisance Parameters, Amer. Math. Soc., Providence.

Rukhin, A. L. (1983). Admissibility and minimaxity results in the estimation problem of exponential quantiles, Technical Report #83-54, Department of Statistics, Purdue University, submitted to Annals of Statistics.

Rukhin, A. L. (1984a). Estimation of linear function of the normal mean and variance, Technical Report #84-24, Department of Statistics, Purdue University.

Rukhin, A. L. (1984b). Estimating a ratio of normal parameters, Technical Report #84-25, Department of Statistics, Purdue University.

Rukhin , A. L. (1985). Quadratic estimators of quadratic functions of normal parameters, Technical Report #85-2, Department of Statistics, Purdue University.

Stein, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean, Ann. Inst. Statist. Math. 16, 155-160.

Stein, C. (1965). Approximation of improper prior measures by prior probability measures, in Bernoulli, Bayes, Laplace Anniversary Volume, Springer-Verlag, New York, 217-241.

Strawderman, W. E. (1974). Minimax estimation of powers of the variance of a normal population under squared error loss, Ann. Statist. 2, 190-198.