ESTIMATING A LINEAR FUNCTION OF THE NORMAL MEAN AND VARIANCE

By

Andrew L. Rukhin[1]

Technical Report #84-24

Department of Statistics
Purdue University

July, 1984
(Revised October 1985)

[1]Research supported by NSF Grant MCS-8101670

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

Key words and phrases: Estimation, normal parameters, quadratic loss, admissibility

ESTIMATING A LINEAR FUNCTION OF THE NORMAL MEAN AND VARIANCE

by

Andrew L. Rukhin[1]

Purdue University

Estimation of Linear Function

[1]Research supported by NSF Grant MCS-8101670

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

Key words and phrases: Estimation, normal parameters, quadratic loss, admissibility

SUMMARY

The estimation problem of a linear function, $\xi + b\,\sigma^2$, of normal parameters ξ and σ^2, is considered. The inadmissibility for quadratic loss of natural estimator of this function is proved for any real constant b, $b \neq 0$.

1. Introduction and Summary.

Let $x_1,\ldots,x_n$, $n \geqslant 2$ be a random sample from a normal distribution with mean ξ and variance σ^2. We consider the statistical estimation problem of a linear function $\theta = \xi + b\,\sigma^2$ where b is a given constant. If $b = 1/2$, $\exp(\theta)$ is the expected value of a lognormal random variable. The estimation problem for the means of transformed normal variables has been studied by many authors (see Neyman and Scott (1960), Hoyle (1968), Shimizu and Iwase (1981), Shimizu (1983)). Uniformly most accurate unbiased confidence intervals for θ, as above, have been obtained by Land (1971) (see also sequel papers of Johnson, Joshi and Land (1973) and Joshi (1974)).

In this paper we are concerned with point estimation of θ under quadratic loss. If $X = \sum_1^n x_i/n$ and $Y^2 = \sum_1^n (x_i - X)^2$, then (X,Y) is a version of complete sufficient statistic, and the unbiased estimator $\delta_0(X,Y)$ of θ has the form

$$\delta_0\,(X,Y) = X + b\,Y^2/(n-1).$$

However for $b \neq 0$ this estimator can be improved upon very easily. Indeed the risk function of a procedure $X + c\,b\,Y$, for any c, because of independence of X and Y has the form

$$E(X + c\,b\,Y^2-\theta)^2 = E(X-\xi)^2 + b^2E(cY^2 - \sigma^2)^2\ ,$$

so that the optimal choice of c is

$$c = c_0 = E_{01}\,Y^2/(E_{01}Y^4) = (n+1)^{-1}\ .$$

Thus we shall investigate the admissibility of the estimator

$$\delta_1(X,Y) = X + b\,Y^2/(n+1).$$

This estimator can be shown to be minimax if one rescales the loss function by the risk of δ_1, i.e. if $L_1(\theta,\delta) = (\theta - \delta)^2/E(\delta_1 - \theta)^2$. The admissibility question for procedure similar to δ_1 (the best equivariant estimator) as an estimator of linear function of mean and standard deviation, $\xi + b\sigma$, has been investigated by Zidek (1971), who established its inadmissibility for $b \neq 0$ (see also Rukhin (1983)). The inadmissibility of the best equivariant estimator of σ is also known (see Stein (1964), Brown (1968), Brewster and Zidek (1974), Strawderman (1974)). In view of these results the inadmissibility of δ_1 as an estimator of θ does not seem to be surprising. However the direct proof of this fact is not easy. For instance, the estimator $X+b\delta(X,Y)$ where δ is improved estimator of σ^2, typically does not improve upon δ_0. (This fact is due to the correlation between X and $\delta(X,Y)$). Thus we prove the inadmissibility of δ_0 by using essentially Stein's necessary condition for admissibility (see Stein (1955), Farrell (1968)). In the proof we discover that δ_0 is generalized Bayes procedure not only with respect to the "uniform" prior but also against many other prior distributions densities of which are related to the heat equation. However, basically the same priors have the inadmissible estimator of σ^2, $Y^2/(n+1)$ as the Bayes one, and δ_0 cannot be admissible as an estimator of θ.

In Section 2 of this paper we investigate the form of generalized Bayes estimators of θ and give a heuristic argument for the inadmissibility. This argument is made rigorous in Section 3 where the inadmissibility of δ_1 is established.

2. Generalized Bayes Estimators of θ.

In this Section we obtain a convenient expression for the generalized Bayes estimator of θ. For our purpose quadratic loss function of the form

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

turns out to be more suitable than ordinary quadratic loss.

Let $\lambda(\xi,\sigma)$ be a density of a (generalized) prior distribution over $\Theta = \{(\xi,\sigma), \sigma > 0\}$ with respect to uniform (right Haar) measure $d\xi\, d\sigma/\sigma$. (The latter choice is motivated mainly by tradition). The Bayes estimator $\delta_B(x,y)$ has the form

$$\delta_B(x,y) = \int_{-\infty}^{\infty}\int_{0}^{\infty} [\xi + b\sigma^2]\sigma^{-6} p((x-\xi)/\sigma, y/\sigma)\lambda(\xi,\sigma)d\xi\, d\sigma/\sigma$$

$$/ \int_{-\infty}^{\infty}\int_{0}^{\infty} \sigma^{-6} p((x-\xi)/\sigma,\ y/\sigma)\ \lambda(\xi,\sigma)d\xi\cdot d\sigma/\sigma\ .$$

Here for positive y and real x

$$p(x,y) = C\, y^{n-2} \exp\{- nx^2/2 - y^2/2\},$$

where C is the normalizing constant, is the density of the joint distribution of (X,Y) when $\xi = 0$, $\sigma = 1$.

Assuming that the following integrations by parts are legitimate one obtains

$$\int(\xi-x)\ \exp\{ - n(x-\xi)^2/2\sigma^2\}\lambda d\xi$$
$$= \sigma^2 n^{-1}\int \exp\{- n(x-\xi)^2/2\sigma^2\}\ \lambda_\xi d\xi,$$
$$\int(\xi-x)^2\ \exp\{- n(x-\xi)^2/2\sigma^2\}\lambda d\xi$$
$$= \sigma^2 n^{-1}\int\exp\{- n(x-\xi)^2/2\sigma^2\}[\lambda + \sigma^2 n^{-1}\lambda_{\xi\xi}]d\xi$$

and

$$(y^2 + n(x-\xi)^2)\ \int\exp\{- [n(x-\xi)^2 + y^2]/2\sigma^2\}\lambda\sigma^{-n-5}d\sigma$$
$$= \int[(n+2)\lambda - \sigma\lambda_\sigma]\ \exp\{- [n(x-\xi)^2 + y^2]/2\sigma^2\}\sigma^{-n-3}d\sigma.$$

Combining these formulas we derive the following useful representation for the Bayes estimator

$$\begin{aligned}\delta_B(x,y) &= x + by^2/(n+1) \\ &+ b(n+1)^{-1} \int_{-\infty}^{\infty}\int_{0}^{\infty} \exp\{-[n(x-\xi)^2 + y^2]/2\sigma^2\}\sigma^{-n-1} \\ &\times [\lambda_{\xi\xi} + n\lambda_\sigma/\sigma + (n+1)\lambda_\xi/(b\sigma^2)]\,d\xi d\sigma \\ &/ \int_{-\infty}^{\infty}\int_{0}^{\infty} \exp\{-[n(x-\xi)^2 + y^2]/2\sigma^2\}\sigma^{-n-5}\lambda\, d\xi d\sigma. \end{aligned} \tag{2.1}$$

It follows from (2.1) that $\delta_B(x,y) = \delta_1(x,y)$ if and only if the prior density λ satisfies the following parabolic differential equation

$$D\lambda = \lambda_{\xi\xi} + n\lambda_\sigma + (n+1)\lambda_\xi/(b\sigma^2) = 0. \tag{2.2}$$

For instance, $\lambda(\xi,\sigma) \equiv 1$, i.e. the "uniform" density is a solution of (2.2). However (2.2) admits many other solutions.

Let $z(u,t) = \lambda((u + a\log t)n^{-\frac{1}{2}}, tn^{-\frac{1}{2}})$, $a = (n+1)/(bn)$ for $b \neq 0$. Then

$$tz_{uu} + z_t = 0. \tag{2.3}$$

Notice that equation (2.3) is closely related to the adjoint heat equation, general solutions to which are known.

For instance, $z(u,t) = \exp(u^2/4t)t^{-1/2}$ and $z(u,t) = \exp(\alpha u-\alpha t^2)$ are solutions of (2.3). However the corresponding prior densities do not admit better approximations by proper densities in terms of posterior risk than the uniform density.

The same equations (2.2) and (2.3) are related to the estimation problem of σ^2. Indeed using the formulae of this Section one obtains the following representation for the Bayes estimator $\tilde{\delta}_B(x,y)$ of σ^2 under quadratic loss $(\tilde{\delta}/\sigma^2-1)^2$,

$$\tilde{\delta}_B(x,y) - y^2/(n+1)$$

$$= \int_{-\infty}^{\infty}\int_0^{\infty} \exp\{-[n(x-\xi)^2+y^2]/2\sigma^2\}\sigma^{-n-1}[\lambda_{\xi\xi}+n\lambda\sigma/\sigma]d\xi d\sigma$$

$$/ \int_{-\infty}^{\infty}\int_0^{\infty} \exp\{-[n(x-\xi)^2+y^2]2\sigma^2\}\sigma^{-n-5}\ \lambda d\xi d\sigma.$$

It follows from (2.4) that $\tilde{\delta}_B(x,y) = y^2/(n+1)$ if and only if λ satisfies (2.2) with a=0. It is also clear that $\delta_B(x,y) = \delta_1(x,y)$ for a prior density $\lambda(\xi,\sigma)$ if and only if $\tilde{\delta}_B(x,y) = y^2/(n+1)$ for the density $\tilde{\lambda}(\xi,\sigma) = \lambda(\xi + a\log\sigma,\sigma)$.

Before concluding this section we sum up our results.

Proposition. The generalized Bayes estimator $\delta_B(x,y)$ of $\theta = \xi + b\sigma^2$ under quadratic loss $(\theta-\delta)^2\sigma^{-4}$ and prior density λ has the form (2.1). If λ satisfies (2.2) then $\delta_B = \delta_1$.

3. Inadmissibility proof

Theorem. The procedure $\delta_1(x,y) = x + b\, y^2/(n+1)$ is admissible for estimating a linear function of normal parameters, $\theta = \xi + b\,\sigma^2$, under quadratic loss, if and only if $b\neq 0$.

Proof. It suffices to consider the case $b \neq 0$.

Because of the continuity of the risk functions the inadmissibility of δ_1 will be established if we show that for any sequence of finite measures Λ_m, m=2,3,... over Θ with positive densities λ_m, lim inf $\rho_m>0$,

$$\rho_m = \iint \sigma^{-4}[E(\delta_1-\theta)^2 - E(\delta_m-\theta)^2]d\Lambda_m(\xi,\sigma), \tag{3.1}$$

where δ_m is the Bayes estimator for the prior distribution Λ_m, and $\Lambda_m(B)>1$ for an open set B.

A straightforward calculation shows that

$$\rho_m = \int_{-\infty}^{\infty}\int_0^{\infty} [\delta_m(x,y) - \delta_1(x,y)]^2 dxdy$$
$$\int_{-\infty}^{\infty}\int_0^{\infty} \sigma^{-6} p((x-\xi)/\sigma, y/\sigma)\ d\Lambda_m(\xi,\sigma).$$

Because of (2.1) one has with a generic constant C and D defined by (2.2)

$$\rho_m = C \int_{-\infty}^{\infty}\int_0^{\infty} [\int_{-\infty}^{\infty}\int_0^{\infty} \exp\{-[n(x-u)^2 + y^2]/2t^2\} D\lambda_m t^{-n-1} dt\ du]^2$$
$$/[\int_{-\infty}^{\infty}\int_0^{\infty} \exp\{-[n(x-u)^2 + y^2]/2t^2\}\lambda_m\ t^{-n-5} dt\ du] y^{n-2} dy\ dx$$
$$= C \int_{-\infty}^{\infty}\int_0^{\infty} [\int_{-\infty}^{\infty}\int_0^{\infty} \exp\{-[n(x-u)^2 + y^2]/2t^2\}[\tilde{\lambda}_{uu} + n\tilde{\lambda}_t] t^{-n-1} dt\ du]^2$$
$$/[\int_{-\infty}^{\infty}\int_0^{\infty} \exp\{-[n(x-u)^2 + y^2]/2t^2\}\tilde{\lambda}_m t^{-n-5} dt\ du]\ y^{n-2} dy\ dx.$$

where $\tilde{\lambda}(u,t) = \tilde{\lambda}_m(u,t) = \lambda_m(u + a \log t,\ t)$.

Thus

$$\rho_m = C \iint \sigma^{-4} [E(Y^2/(n+1) - \sigma^2)^2 - E(\tilde{\delta}_m - \sigma^2)^2] d\ \tilde{\Lambda}_m(\xi,\sigma)$$

where $\tilde{\Lambda}_m$ is the distribution defined by $\tilde{\lambda}_m$ and $\tilde{\delta}_m$ is the corresponding Bayes estimator of σ^2. Because of the inadmissibility of $Y^2/(n+1)$ as an estimator of σ^2

$$\liminf \rho_m > 0$$

and our Theorem is proved.

Acknowledgement

This paper was written while the author was on leave at Rutgers University. It is a pleasant duty to thank colleagues at the Department of Statistics at Rutgers University for interesting discussion. In particular the problem was suggested by Professor A. Cohen to whom the author is especially grateful. Thanks are also due to Jeesen Chen for several critical remarks and questions.

REFERENCES

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

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

Farrell, R. H. (1968): On a necessary and sufficient condition for admissibility of estimators when stricly convex loss is used. Ann. Math. Statist., 39, 23-28.

Hoyle, M.H. (1968): The estimation of variances after using a Gaussianating transformation. Ann. Math. Statist., 39, 1125-1143.

Johnson, B.R., Joshi, V.M. and Land, C.E. (1973): A note on two-sided confidence interval for linear functions of the normal mean and variance. Ann. Statist., 4, 413-418.

Joshi, V.M. (1976): Confidence intervals for linear functions of the normal mean. Ann. Statist., 4, 413-418.

Land, C.E. (1971): Confidence intervals for linear functions of the normal mean and variance. Ann. Math. Statist., 42, 1187-1205.

Neyman, J. and Scott, E.L. (1960): Correction for bias introduced by a transformation of variables. Ann. Math. Statist., 31, 643-655.

Rukhin, A.L. (1983): A class of minimax estimators of a normal quantile. Stat. and Prob. Letters, 1, 217-222.

Shimizu, K. (1983): Variances of UMVU estimators for means and variances after using a normalizing transformation. Commun. Statist. - Theor. Math., 12, 975-985.

Shimizu, K. and Iwase, K. (1981): Uniformly minimum variance unbiased estimation in lognormal and related distributions. Commun. Statist. Theor. Meth., 10, 1127-1147.

Stein, C. (1955): A necessary and sufficient condition for admissibility Ann. Math. Statist., 26, 518-522.

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.

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

Zidek, J. (1971): Inadmissibility of a class of estimators of a normal quantile. Ann. Math. Statist., 42, 1444-1447.