MINIMAX ESTIMATION OF THE LOCATION PARAMETER OF THE SHIFTED NORMAL DISTRIBUTION WITH UNKNOWN SCALE

by A. L. Rukhin
Purdue University

Department of Statistics
Division of Mathematical Sciences

Mimeograph Series # 515

November 1977

Summary

Estimation of location with unknown scale

The problem of estimating the location parameter ξ with unknown scale σ in the family $\{\sigma^{-1}\, p((\cdot-\xi)/\sigma),\ \xi \in R^1,\ \sigma > 0\}$ is considered. It is shown that if $p(u) = (2\pi)^{-\frac{1}{2}} \exp\{-(u-b)^2/2\}$, then the best equivariant estimator of ξ for quadratic loss is admissible if and only if $b = 0$.

MINIMAX ESTIMATION OF THE LOCATION PARAMETER OF THE SHIFTED NORMAL DISTRIBUTION WITH UNKNOWN SCALE

By A. L. Rukhin
Purdue University

I. INTRODUCTION. Suppose that the statistician wishes to measure an unknown quantity ξ with measurements subject to additive random errors, the variance of which is unknown. In other words the observations X_j have the following form

(1.1) $$X_j = \xi + \sigma\varepsilon_j \qquad j = 1,\ldots,n,$$

where ξ and σ are unknown and $\varepsilon_1,\ldots,\varepsilon_n$ are i.i.d. random variables. The classical theory of measurement assumes that the distribution of these variables is normal with zero mean and variance one. The present paper deals with the more general case where ε_j is normal with known mean b and known variance τ, j=1,...n. If $b > 0$ this is an attempt to model a commonly observed situation where inability to measure precisely (without bias) is related to inability to measure **reliably** (with small error variation).

In the above situation, $X_1,\ldots,X_n$, $n \geq 2$, are i.i.d. random variables each with the density $\sigma^{-1} p((\cdot-\xi)/\sigma)$, where ξ and σ are unknown location and scale parameters and p is the normal density

(1.2) $$p(u) = (2\pi)^{-\frac{1}{2}} \tau^{-1} \exp(-(u-b)^2/2\tau^2).$$

AMS 1970 subject classifications. 62F10, 62C99.

Key words and phrases. Equivariant estimators, estimation of the location parameter with unknown scale parameter, inadmissibility.

The problem of estimating ξ is invariant under the affine group, and when $b = 0$, the best equivariant estimator (coinciding, of course, with the sample mean) is known to be admissible if the loss is quadratic. Moreover admissibility of the sample mean in this situation is a characteristical property of the normal law (cf. Kagan and Zinger (1973)). The situation is quite different if $b \neq 0$. We will give in this paper for each $b \neq 0$ a class of minimax estimators of ξ for quadratic loss which improve upon the best equivariant estimator.

The problem of admissibility of the best equivariant estimator of the scale parameter σ with unknown ξ and $b = 0$ has been studied by several authors. Stein (1964) showed that the best equivariant estimator for σ^2 is inadmissible. Brown (1968) proved in a more general setting that the best equivariant estimator of the αth power of σ in the presence of an unknown location parameter is inadmissible for a large class of loss functions. Similar results were obtained by Brewster and Zidek (1974) and Strawderman (1974). We will indicate the connection of the problem of estimating σ with ξ unknown to that of estimating ξ with unknown σ and $b \neq 0$.

2. A CLASS OF MINIMAX ESTIMATORS OF ξ.

In this section we produce a class of minimax estimators of ξ when σ is unknown and $b \neq 0$. Let $X = n^{-1} \sum_1^n X_j$ and $Y^2 = n^{-1} \sum_1^n (X_j - X)^2$ be sufficient statistics for ξ and σ.

If $\delta(X,Y)$ is an equivariant estimator of ξ under the affine group based on X and Y, then $\delta(cX + d, cY) = c\delta(X,Y) + d$ for all d and $c > 0$. This implies that δ has the form $\delta(X,Y) = X - \lambda Y$ for some constant λ. If the loss is measured by $(\delta-\xi)^2/\sigma^2$, an examination of the risk of δ reveals that the best choice for λ is

$$\lambda_0 = \frac{\Gamma(\frac{n+1}{2})n^{\frac{1}{2}}}{\Gamma(\frac{n+2}{2})2^{\frac{1}{2}}\tau} b = ab.$$

This estimator $\delta_0(X,Y) = X - \lambda_0 Y$ has a constant risk and is known to be minimax.

We consider estimators of the form

$$\delta(X,Y) = X-bY\ \phi(X/Y) \tag{2.1}$$

for some measurable function ϕ.

THEOREM. Assume that $b \neq 0$ and that it is desired to estimate ξ using the loss $(\delta-\xi)^2/\sigma^2$. If $\phi(z)$ satisfies the inequality

$$a \geq \phi(z) \geq \frac{(\mu bz)^2-2(n+1)a|bz| + 4(n+1)}{2(n+1)|bz|}$$

($\mu = n^{\frac{1}{2}}/\tau$), when $|bz + 2a(n+1)\mu^{-2}| \leq 2(n+1)\mu^{-2}[a^2-\mu^2(n+1)^{-1}]^{\frac{1}{2}}$, and $\phi(z) = a$ otherwise, then the estimator (2.1) is minimax.

PROOF. The risk of the estimator (2.1) clearly depends only on the parameter $\eta = \xi/\sigma + b$, so that without loss of generality we can take $\sigma=1$. Note that

$$\begin{aligned} R(\eta,\delta) &= E_\eta(X-by\phi(X/Y)-\eta + b)^2 = \\ &E_\eta(X-\eta)^2 + b^2E_\eta(Y\phi(X/Y)-1)^2- \\ &-2bE_\eta(X-\eta)(Y\phi(X/Y)-1). \end{aligned} \tag{2.2}$$

Since $E_\eta X = \eta$ and X and Y are independent we see that the minimaxity of the estimator (2.1) is equivalent to the following inequality

$$\begin{aligned} \Delta &= R(\eta,\delta_0)-R(\eta,\delta) = \\ &= b^2[E_\eta(aY-1)^2-E_\eta(Y\phi(X/Y)-1)^2] + \\ &+ 2b\ E_\eta(X-\eta)Y\phi(X/Y) \geq 0. \end{aligned} \tag{2.3}$$

From (2.3) it follows that

(2.4) $$\Delta = E_\eta[a-\phi(X/Y)][b^2Y^2(a + \phi(X/Y))-2b^2Y-2bY(X-\eta)] =$$

$$= \frac{\exp\{-\mu^2\eta^2/2\}\mu^{n+1}}{(2\pi)^{\frac{1}{2}}\Gamma(^n/2)2^{^n/2-1}} \sum_{k=0}^{\infty} \frac{\eta^k\mu^{2k}}{k!} \int_{-\infty}^{\infty} [a-\phi(z)]z^k dzx$$

$$\int_0^{\infty} [b^2y^2(a + \phi(z))-2b^2y-2by(zy-\eta)]y^{n+k}\exp\{-\mu^2(1 + z^2)y^2/2\}dy =$$

$$= \frac{\exp\{-\mu^2\eta^2/2\}}{(2\pi)^{\frac{1}{2}}\Gamma(^n/2)2^{^n/2-1}\mu^2} \sum_{k=0}^{\infty} \frac{\eta^k\mu^k}{k!} \int_{-\infty}^{\infty} [a-\phi(z)]z^k(1 + z^2)^{-(n+k+3)/2}dzx$$

$$\int_0^{\infty} [b^2y^2(a + \phi(z)-2zb^{-1})-2b^2\mu y(1+z^2)^{\frac{1}{2}} + 2bk(z + z^{-1})]y^{n+k}e^{-y^2/2}dy.$$

Since for all integer k

$$k\int_0^{\infty} y^{n+k}\, e^{-y^2/2}\, dy = \int_0^{\infty} y^{n+k+2}\, e^{-y^2/2}\, dy-(n+1)\int_0^{\infty} y^{n+k}\, e^{-y^2/2}dy,$$

the relation (2.4) can be rewritten in the following way

(2.5) $$\Delta = \frac{\exp\{-\mu^2\eta^2/2\}b^2}{(2\pi)^{\frac{1}{2}}\Gamma(^n/2)2^{^n/2-1}\mu^2} \sum_{k=0}^{\infty} \frac{\eta^k\mu^k}{k!} \int_{-\infty}^{\infty} [a-\phi(z)]z^k x$$

$$x(1+z^2)^{-(n+k+3)/2}dz \int_0^{\infty} [(a+\phi(z) + 2/bz)y^2-2\mu(1+z^2)^{\frac{1}{2}}y-2(n+1)(z+z^{-1})/b]x$$

$$y^{n+k}e^{-y^2/2}dy =$$

$$= \frac{\exp\{-\mu^2\eta^2/2\}b^2}{(2\pi)^{\frac{1}{2}}\Gamma(^n/2)2^{^n/2-1}\mu^2} \int_{-\infty}^{\infty}\int_0^{\infty} [a-\phi(z)]\exp\{zy\eta\mu/(1+z^2)^{\frac{1}{2}}\}x$$

$$(1+z^2)^{-(n+3)/2}y^n[(a+\phi(z)+2/bz)y^2-2\mu(1+z^2)^{\frac{1}{2}}y-$$

$$2(n+1)(z+z^{-1})/b]e^{-y^2/2}dy\, dz.$$

From (2.5) it is evident that Δ is nonnegative if

$$[a-\phi(z)][(a+\phi(z)+2/bz)y^2-2\mu(1+z^2)^{\frac{1}{2}}y-2(n+1)(z+z^{-1})/b] =$$
$$= [a-\phi(z)][(a+\phi(z)+2/bz)(y-\mu(1+z^2)^{\frac{1}{2}}(a+\phi(z)+2/bz)^{-1})^2-$$
$$- \mu^2(1+z^2)(a+\phi(z)+2/bz)^{-1}-2(n+1)(1+z^2)/bz] \geq 0.$$

Now it is clear that if ϕ is chosen in such a way that $\phi(z) \leq a$, and on the set where $\phi(z) < a$

$$(2.6) \qquad -2(n+1)(bz)^{-1} \geq \mu^2(a+\phi(z)+2(bz)^{-1})^{-1} \geq 0,$$

then the estimator (2.1) is minimax.

The inequalities (2.5) can be achieved on the set $\{z: |bz+2a(n+1)\mu^{-2}| \leq 2(n+1)\mu^{-2}(a^2-{}^2(n+1)^{-1})^{\frac{1}{2}}\}$ with any function ϕ such that

$$(2.7) \qquad a \geq \phi(z) \geq \mu^2|bz|/2(n+1) + 2/|bz| - a.$$

Note that all manipulations in (2.4) and (2.5) are legitimate for such functions ϕ. The theorem is proven.||

Remark: If the inequalities in (2.6) are strict, then the corresponding minimax estimators do not have constant risk, and are hence better than the best equivariant estimator δ_0.

The estimator belonging to the class described in the theorem and maximizing the quantity Δ in (2.5) for $\eta = \eta_0$ has the form

$$\phi_0(z) = \frac{z}{b} + \frac{\int_0^\infty y^{n+1}\exp\{\eta_0\mu zy(1+z^2)^{-\frac{1}{2}}-y^2/2\}dy}{\int_0^\infty y^{n+2}\exp\{\eta_0\mu zy(1+z^2)^{\frac{1}{2}}-y^2/2\}dy}\,\mu(1+z^2)^{\frac{1}{2}}(1-\eta_0 b^{-1})$$

if the inequality (2.7) holds. ($\phi_0(z)$ must of course be truncated at a or $[\mu^2|bz|/2(n+1) + 2/|bz| - a]$ if (2.7) is violated.)

3. HEURISTIC RATIONALE. The formula (2.3) explains why the best equivariant estimator $\delta_0(X,Y) = X - abY$ is inadmissible, at least for large b. To see this, note that aY is the best scale equivariant estimator of σ for the loss function $(\delta/\sigma - 1)^2$. This estimator is known to be inadmissible if ξ is unknown and there exist functions ϕ such that

$$E_\eta(aY - 1)^2 \geq E_\eta(Y\phi(X/Y) - 1)^2$$

for all η. Thus the first term in (2.3) is positive if ϕ corresponds to the estimator improving upon aY. The second term cannot be large if ϕ varies slowly enough, and for sufficiently large b is negligible compared to the first term. It can indeed be proven that for sufficiently large b estimators of the form $X - abY\,\phi(X/Y)$ improve upon $\delta_0(X,Y)$ when ϕ corresponds to the Stein estimator [3] or some of Strawderman's estimators [4]. The fact that all known scale estimators improving upon aY have the form $Y\,\phi(X/Y)$ motivates the consideration of the estimators (2.1). Moreover for large b the relative improvement of the estimator $X-abY\,\phi(X/Y)$ upon $\delta(X,Y)$ is approximately equal to the relative improvement of $y\,\phi(X/Y)$ upon aY:

$$\frac{R(\eta,\delta_0) - R(\eta,\delta)}{R(\eta,\delta_0)} = \frac{E_\eta(aY-1)^2 - E_\eta(Y\phi(X/Y)-1)^2}{E_\eta(aY-1)^2} + 0\left(\frac{1}{|b|}\right).$$

The formula (2.2) can be used to prove admissibility of $\delta_0(X,Y)$ in the class of all location equivariant estimators of the form $X - \gamma(Y)$. Indeed

$$E_{\xi,\sigma}(X-\gamma(Y)-\xi)^2 =$$

$$E_{\xi,\sigma}(X-\gamma(Y)-\xi-\sigma b + \sigma b)^2 =$$

$$E_{\xi,\sigma}(X-\xi-\sigma b)^2 + E_{\xi,\sigma}(\gamma(Y)-\sigma b)^2 -$$

$$-2\,E_{\xi,\sigma}(X-\xi-\sigma b)(\gamma(Y)-\sigma b).$$

The last term in this relation is equal to zero because of the independence of X and Y. Thus inadmissibility of the estimator X - abY among estimators of the form X - $\gamma(Y)$ would imply inadmissibility of the estimator aY for the scale parameter σ of the normal law with zero mean within the class of all estimators depending only on Y. But this is not the case, so that δ_0 is admissible in the class of location equivariant estimators.

REFERENCES

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

[2] 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.

[3] Kagan, A. M. and Zinger, A. A. (1973). Sample mean as an estimator of the location parameter in presence of the nuisance scale parameter, Sankhyā, 35 447-454.

[4] 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.

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