IMPROVED ESTIMATION IN LOGNORMAL MODELS*

by

Andrew L. Rukhin
Purdue University

Technical Report #84-38

Department of Statistics
Purdue University

September 1985
(Revised January 1986)

*Research was supported by the National Science Foundation, Grant MCS-8401996.

Andrew L. Rukhin is Professor, Department of Statistics, Purdue University, West Lafayette, IN 47907. The research for this article was supported by NSF Grant MCS-8401996. The original version of this paper was written at Department of Statistics, University of Rome where facilities were generously provided during author's sabbatical leave. The author also thanks the associate editor for many helpful suggestions.

IMPROVED ESTIMATION IN LOGNORMAL MODELS

BY

Andrew L. Rukhin
Statistics Department
Purdue University
West Lafayette, IN 47907

ABSTRACT

The estimation of the function $\exp(a\xi+b\sigma^2)$ of normal parameters ξ and σ^2 on the basis of a random sample $X_1,\ldots,X_n$ is considered. This function corresponds to the mean, the median and all moments of lognormal distribution. We show that the minimum variance unbiased estimator suggested by Finney in 1941 can be substantially improved in terms of mean square error. Similar result is established for the maximum likelihood estimator. We suggest for practical use the following generalized Bayes estimator

$$\delta(X,Y) = \exp(aX+(\gamma-\beta)Y) \sum_{k=0}^{\left[\frac{n+1}{2}\right]} \frac{(2m-k)!}{k!(m-k)!} (2\beta Y)^k \Big/ \sum_{k=0}^{\left[\frac{n+1}{2}\right]} \frac{(2m-k)!}{k!(m-k)!} (2\gamma Y)^k .$$

Here $X = \sum_1^n X_j/n$, $Y^2 = \sum_1^n (X_j-X)^2$, and constants β and γ are determined by formulas $\beta^2 = \gamma^2 - 2b + 2a^2/n$, $\gamma = 1.5(b-3a^2/(2n))$. This estimator is shown to be locally optimal for both small and large values of σ^2. The results of numerical study of the quadratic risk show the superiority of this estimator over the mentioned traditional procedures.

Key words: Lognormal distribution, generalized Bayes estimators, minimum variance unbiased estimators, maximum likelihood estimators, indamissibility, quadratic risk.

1. INTRODUCTION

Let random variable Z be normally distributed with mean ξ and variance σ^2, so that the distribution of exp(Z) is lognormal. The mean of this distribution, the median, the mode and all moments have the form $\theta = \exp(a\xi + b\sigma^2)$ for some constants a and b. For instance in the case of mean a = 1, b = 1/2; for the median a = 1, b = 0, etc.

In this paper, we consider the estimation problem for quadratic loss of the parametric function θ, as above, on the basis of normal random sample $X_1,\ldots,X_n$, $n \geq 2$. In other terms we are interested in estimating the mean (or any other moment) of lognormal distribution. This problem presents practical interest since lognormal distribution is a commonly accepted model in many applications, e.g. in economics, psychological studies, reliability etc. Also logarithmic transformation of variables, which is supposed to have normalizing effect, is widely used, so that it is of interest to find the mean of the original sample.

The estimation problem of θ has a long history. Finney (1941) has derived the best (minimum variance) unbiased estimator of θ whose properties were studied later by Bradu and Mundlak (1970) and Evans and Shaban (1974, 1976) (see also Aitchison and Brown; 1966; Ebbeler 1973; Dhrymes 1962; Laurent 1963; Thöni 1969; Shaban 1981). Confidence intervals for θ have been obtained in Land (1971, 1973).

In Section 2 we demonstrate that substantial improvements upon Finney's estimator and maximum likelihood estimator are possible. In Section 3 a class of generalized Bayes esimators is obtained. Within this class we determine the procedure which is locally optimal at zero and at infinity. Numerical study of the risk of these procedures is reported in Section 4.

2. INADMISSIBILITY RESULTS

Let $X = \sum_1^n X_j/n$, $Y^2 = \sum_1^n (X_j-X)^2$ be a version of the complete sufficient statistic for (ξ,σ). Notice that

$$E \exp(aX) = \exp(a\xi + a^2\sigma^2/(2n))$$

and

$$EY^{2k} = \sigma^{2k}\Gamma(k+(n-1)/2)\ 2^k/\Gamma((n-1)/2).$$

Therefore

$$\delta_u(X,Y) = e^{aX}\Gamma((n-1)/2) \sum_{k=0}^{\infty} (ry)^{2k}/[k!\Gamma(k+(n-1)/2)2^k],$$

$r^2 = 2(b-a^2/2n)$, is the unbiased estimator. More conveniently if $b > a^2/(2n)$,

$$\delta_u(X,Y) = e^{aX}\Gamma((n-1)/2)\ I_{(n-3)/2}(rY)(rY/2)^{-(n-3)/2} = e^{aX}h(rY), \qquad (2.1)$$

where $I_{(n-3)/2}$ is the modified Bessel function of the first kind.

Bradu and Mundlak (1970) and Evans and Shaban (1974) have studied the variance (or the quadratic risk) of δ_u. In somewhat different notation they showed that

$$E\, h^2(rY) = e^{r^2\sigma^2} h(r^2\sigma^2). \tag{2.2}$$

Thus

$$E(\delta_u(X,Y) - \theta)^2 = \theta^2(h(r^2\sigma^2)\, \exp(a^2\sigma^2/n) - 1).$$

An undesirable feature of estimator (2.1) in the case when $b < a^2/(2n)$ is that it takes negative values with positive probability (see Teekens and Koerts (1972)). Thus δ_u is not admissible in this case; for instance, the procedure $\max(0,\delta_u)$ is better. We show that in the case $b \geq a^2/(2n)$ estimator (2.1) and a related estimator (2.5) are also inadmissible and substantial improvements over them for small samples are possible.

To this end we consider estimators of the form

$$\delta(X,Y) = e^{aX} g(Y) \tag{2.3}$$

where g is a positive function with finite second moment.

The mean square error of procedures (2.3) has the form

$$\begin{aligned} E(\delta-\theta)^2 &= Ee^{2aX}g^2(Y) - 2\theta Ee^{aX}g(Y) + \theta^2 \\ &= \theta^2[\exp\{-2(b-a^2/n)\,\sigma^2\}\, E(g(Y)-\exp(c\sigma^2))^2 \\ &\quad + 1 - \exp(-a^2\sigma^2/n)], \end{aligned} \tag{2.4}$$

where $c = b-3a^2/(2n)$. Thus for estimators (2.3) the estimation problem of θ reduces to that of $\exp(c\sigma^2)$.

From (2.4) it is clear that δ_u can have excessively large risk. Indeed (2.4) suggests that the choice of $g(Y)$ being unbiased estimator of $\exp(c\sigma^2)$ (and not $\exp(r^2\sigma^2)$) is more appropriate.

The resulting estimator

$$\delta_1(X,Y) = e^{aX}h((2c)^{1/2}Y) \tag{2.5}$$

where h is defined by (2.1) is uniformly better than δ_u. Indeed

$$E(h(qY) - e^{c\sigma^2})^2$$

$$= \exp\{q^2\sigma^2\}\ [h(q^2\sigma^2\ )-1]$$

$$+\ (\exp(c\sigma^2) - \exp\{q^2\sigma^2/2\})^2.$$

For $c > 0$, let $q = q_0 = (2c)^{1/2}$; then

$$E(h(q_0Y)-e^{c\sigma^2})^2 < E(h(qY) - e^{c\sigma^2})^2$$

for any $q > q_0$. Because of (2.4) and the monotonicity of the function $e^t[h(t) - 1]$ this shows the inadmissibility of δ_u.

Estimator (2.5) has been suggested by Evans and Shaban (1976) who tabulated its quadratic risk. However this estimator is also inadmissible.

To explain it notice that if for small Y

$$g(Y) \sim 1 + gY^2$$

then for small σ

$$E(g(Y) - \exp(c\sigma^2))^2 \sim \sigma^4[g^2(n^2-1)-2gc(n-1)+c^2].$$

Therefore for small σ the optimal choice of the coefficient g is

$$g = g_0 = c/(n+1). \tag{2.6}$$

This fact explains the local optimality of the estimator

$$\delta_3(X,Y) = \exp\{aX+cY^2/(n+1)\} \tag{2.7}$$

the risk of which has been evaluated by Evans and Shaban (1976). It also shows that both estimators δ_u and δ_1 can be improved for small σ. For instance, the estimator δ_2 with

$$g_2(Y) = h([2c(n-1)/(n+1)]^{1/2}Y) \tag{2.8}$$

is locally better than these estimators.

The following estimator δ_0 with

$$g_0(Y) = \sum_{k=0}^{\infty} (cY^2)^k \, \Gamma(k+(n-1)/2)/[k!\Gamma(2k+(n-1)/2)2^k] \qquad (2.9)$$

is also locally optimal. The motivation for this estimator is the following:

$$Y^{2k}\Gamma(k+(n-1)/2)/[\Gamma(2k+(n-1)/2)2^k]$$

is the best estimator under quadratic loss of σ^{2k} among all estimators proportional to Y^{2k}. Thus each term of the sum in (2.9) is the best estimator of the corresponding term in the Taylor expansion of $\exp(c\sigma^2)$.

To complement the asymptotical study of the risk we consider also large values of σ. It can be shown that if as $Y \to \infty$

$$g(Y) \sim Ce^{dY} Y^v$$

with some constants C, v $(v > -n+2)$, and d, then for estimator (2.3)

$$(2\sigma^2)^{-1} \log[E(\delta-\theta)^2/E(\delta_0-\theta)^2] \to \max(d^2,c)-2(b-a^2/(2n)). \qquad (2.10)$$

Since

$$h(Y) \sim Ce^Y Y^{-(n-2)/2}$$

it follows that δ_2 and δ_0 are better than δ_1 and δ_u for large values of σ.

In section 3 we obtain a generalized Bayes estimator δ_B which improves upon δ_1 (and therefore upon δ_u).

This fact will be derived from the following.

Theorem 1. Let $g(Y) = g_1(Y) - 2f(Y)$ where $g_1(Y) = h((2c)^{1/2}Y)$, and f is a nonnegative function such that

(a) $f(Y)Y^{-2}$ is nondecreasing

(b) $h((2c)^{1/2}Y) - f(Y)$ is increasing

(c) $EY^2 f(Y) \leq 2c\sigma^4 \exp(c\sigma^2)$.

Then for all σ

$$E(g(Y) - \exp(c\sigma^2))^2 < E(g_1(Y)^{1/2}Y) - \exp(c\sigma^2))^2.$$

Proof.

$$E(g_1(Y) - \exp(c\sigma^2))^2 - E(g(Y) - \exp(c\sigma^2))^2$$

$$= 4Ef(Y)(g_1(Y) - f(Y) - \exp(c\sigma^2)).$$

According to condition (b) for any fixed σ the function $g_1(Y)-f(Y)-\exp(c\sigma^2)$ changes sign from negative to positive at most once. Because of (a) our Theorem will be proven if we show that

$$EY^2(g_1(Y) - f(Y) - \exp(c\sigma^2)) \geq 0. \tag{2.11}$$

By differentiating in σ the identity

$$Eg_1(Y) = \exp(c\sigma^2)$$

one obtains

$$EY^2 g_1(Y) = (n-1+2c\sigma^2)\sigma^2 \exp(c\sigma^2).$$

Thus (2.11) means that

$$EY^2 f(Y) \leq 2c\sigma^4 \exp(c\sigma^2)$$

which is exactly assumption (c) of Theorem 1.

This Theorem can be used to show that δ_2 is better than δ_1. Also an analogue of Theorem 1 can be proven for the maximum likelihood estimator $\hat{\delta}(X,Y) = \exp\{aX+bY^2/n\}$, the quadratic risk of which is infinite of $\sigma^2 > n/(4b)$.

We formulate these results as

Theorem 2. The minimum variance unbiased eatimator δ_u, the estimator δ_1 defined by (2.5) and the maximum likelihood estimator $\hat{\delta}$ of $\theta = \exp\{a\xi + b\sigma^2\}$ are inadmissible for quadratic loss. Estimator (2.8) improves upon δ_u and δ_1 for positive $c = b-3a^2/(2n)$.

3. BAYES ESTIMATORS OF θ.

Let $\lambda(\xi,\sigma)$ be (generalized) prior density. Then Bayes estimator $\delta_B(X,Y)$ for the loss function $(\delta/\theta-1)^2$ has the form

$$\delta_B(X,Y) = \frac{\iint \sigma^{-n}\exp\{-a\xi-b\sigma^2-[n(X-\xi)^2+Y^2]/(2\sigma^2)\}\lambda(\xi,\sigma)d\xi d\sigma}{\iint \sigma^{-n}\exp\{-2a\xi-2b\sigma^2-[n(X-\sigma)^2+Y^2]/(2\sigma^2)\}\lambda(\xi,\sigma)d\xi d\sigma}.$$

Notice that the generalized Bayes estimator with respect to traditional non-informative prior, $\lambda(\xi,\sigma) = \sigma^{-1}$, does not exist. If $\lambda(\xi,\sigma) = \lambda(\sigma)$, i.e., the prior is "uniform" in ξ, then

$$\delta_B(X,Y) = e^{aX} \frac{\int \sigma^{-n+1}\exp\{-\sigma^2(b-a^2/2n)-Y^2/(2\sigma^2)\}\lambda(\sigma)\, d\sigma}{\int \sigma^{-n+1}\exp\{-2(b-a^2/n)\sigma^2-Y^2/(2\sigma^2)\}\lambda(\sigma)\, d\sigma} \tag{3.1}$$

$$= e^{aX} g(Y).$$

Thus the generalized Bayes procedures in this case have the form (2.3) The Bayesian estimation problem has been considered by Zellner (1971) who derived the form of posterior density and performed a comparison between Bayesian and non-Bayesian results.

We consider prior densities of the form

$$\lambda(\sigma) = \sigma^{-2\nu+n-2}\exp\{-\sigma^2[\gamma^2/2-2(b-a^2/n)]\},\ \gamma^2 > 4(b-a^2/n). \tag{3.2}$$

To evaluate δ_B we need the following formula (see Erdelyi (1954) p. 313, (17))

$$\int_0^\infty u^{-2\nu-1}\exp\{-Au^2/2-B/(2u^2)\}du = (A/B)^{\nu/2}K_\nu((AB)^{1/2}) \tag{3.3}$$

where K_ν is modified Bessel function of the third kind.

Because of (3.3) one obtains with $\beta^2 = \gamma^2-2c$

$$g(Y) = (\beta/\gamma)^\nu\ K_\nu(\beta Y)/K_\nu(\gamma Y). \tag{3.4}$$

If $\nu = m + 1/2$ where m is a positive integer then

$$g(Y) = \exp((\gamma-\beta)Y) \sum_{k=0}^{m} \frac{(2m-k)!}{k!(m-k)!} (2\beta Y)^k \Big/ \sum_{k=0}^{m} \frac{(2m-k)!}{k!(m-k)!} (2\gamma Y)^k. \tag{3.5}$$

Known asymptotical formulas for Bessel functions show that for $\nu > 1$ and small Y,

$$(\beta/\gamma)^{\nu} K_{\nu}(\beta Y)/K_{\nu}(\gamma Y)$$

$$\sim 1 + cY^2/[2(\nu-1)].$$

Therefore for positive c the best choice (for small σ) $\nu = \nu_0$ is such that

$$\nu_0 - 1 = (n+1)/2$$

Thus if n is an even number, g has form (3.5). Since for large Y

$$g(Y) \sim \exp\{(\gamma-\beta)Y\} \; (\beta/\gamma)^{\nu-1/2},$$

formula (2.10) suggests the optimal choice of γ for large σ:

$$\gamma - (\gamma^2 - 2c)^{1/2} = c^{1/2}.$$

We denote by δ_B the corresponding Bayes estimator, and study its mean squared error in Section 4. One can check that all conditions of Theorem 1 are met, so that δ_B improves upon δ_1 and δ_u.

We formulate our results here as

Theorem 3. The generalized Bayes estimator δ_B of θ for prior density $\lambda(\xi,\sigma) = \lambda(\sigma)$ has form (3.1). If $\lambda(\sigma)$ is given by (3.2) then g has form (3.4). The choice of the parameters γ and ν in (3.2) for positive c which minimize the risk function for small and large values of σ are

$$\nu_0 = (n+3)/2 \tag{3.6}$$

and

$$\gamma_0 = 3c^{1/2}/2 \ . \qquad (3.7)$$

4. NUMERICAL RESULTS

The relative mean square errors $E(\delta-\theta)^2/\theta^2$ were evaluated for estimators δ_u, δ_1, δ_2, δ_3, δ_0 and δ_B which are defined by formulae (2.1), (2.5), (2.8), (2.7), (2.9) and (3.5), (3.6), (3.7). In these calculations we put n=4, 6, 8 and 10, and a=1, b=1/2, which corresponds to the lognormal mean estimation.

For all sample sizes considered δ_B clearly exhibits the best behavior. Notice that for larger values of σ the competing estimators δ_2 and δ_3 will have considerably larger risk than δ_B. Also notice that in all cases considered δ_2 is preferable to δ_u, δ_1 and δ_3. The UMVU estimator δ_u exhibits a poor performance and cannot be recommended in practice. Its relative quadratic risk is even worse for smaller values of a.

Mehran (1973) had shown that the variance of δ_u is numerically close to the variance of unbiased (inadmissible) estimator $n^{-1} \sum_1^n \exp(X_j)$. The latter procedure is a particular case of a consistent nonparametric estimator (so-called smearing estimate, see Duan (1983)), and this fact can be interpreted as another argument against UMVU estimator in this problem. Notice however that the bias of all other estimators in this study is negative, so that our conclusions are valid only if underestimation of θ is not more heavily penalized than is suggested by the quadratic loss choice.

Comparing risks of δ_u and smearing estimate with the Cramer-Rao bound gives a useful inequality for the function h in (2.1):

$$\sigma^2(1+\sigma^2/2)/n < h((n-1)\sigma^2/n)e^{\sigma^2/n}-1 \quad\quad (4.1)$$
$$< (e^{\sigma^2}-1)/n.$$

Inequality (4.1) gives useful bounds for quadratic risks of estimators δ_1 and δ_2. To evaluate these risks for small values of $\sigma(\sigma<1)$ we used the power series, for larger values of σ the recurrent formulae for modified Bessel functions provide accurate numerical results.

Table 1

Mean Square Errors of Estimators δ_u, δ_1, δ_2, δ_3, δ_0 and δ_B

		δ_u	δ_1	δ_2	δ_3	δ_0	δ_B
n=4	σ^2= .05	.00063	.00062	.00062	.00062	.00062	.00060
	.10	.00251	.00250	.00250	.00250	.00250	.00248
	.25	.01612	.01552	.01552	.01554	.01552	.01551
	.50	.07091	.06099	.06083	.06108	.06904	.06087
	1.0	.40783	.22513	.22479	.22725	.22513	.22553
	2.0	8.07714	.67023	.65937	.66647	.66009	.65764
n=6	.05	.00042	.00042	.00042	.00042	.00042	.00041
	.10	.00168	.00167	.00167	.00167	.00165	.00163
	.25	.01075	.01042	.01039	.01047	.01042	.01038
	.50	.04717	.04192	.04182	.04247	.04184	.04182
	1.0	.26546	.16831	.16693	.17394	.16788	.16667
	2.0	4.17172	.62874	.60456	.60578	.61156	.57481
n=8	.05	.00031	.00031	.00031	.00031	.00031	.00031
	.10	.00126	.00125	.00125	.00125	.00125	.00125
	.25	.00806	.00783	.00783	.00794	.00789	.00780
	.50	.03536	.03217	.03201	.03310	.03209	.03202
	1.0	.19645	.14032	.13449	.14638	.14021	.13408
	2.0	2.68503	.60994	.56314	.56896	.57750	.51950
n=10	.05	.00025	.00025	.00025	.00025	.00025	.00025
	.10	.00101	.00100	.00100	.00100	.00100	.00100
	.25	.00645	.00630	.00630	.00641	.00630	.00630
	.50	.02801	.02750	.02614	.02741	.02679	.02603
	1.0	.15582	.12095	.11021	.12904	.11107	.11278
	2.0	1.94102	.55367	.54097	.54286	.55262	.47715

REFERENCES

Aitchison, J. and Brown, J. A. C. (1966), The Lognormal Distribution, Cambridge, England: Cambridge University Press.

Bradu, D. and Mundlak, Y. (1970). "Estimation in Lognormal Linear Models," Journal of the American Statistical Association, 65, 198-211.

Dhrymes, P. J. (1962), "On Devising Unbiased Estimators for the Parameters of the Cobb-Douglas Production Function", Econometrica, 30, 297-304.

Duan, N. (1983), "Smearing Estimate: a Nonparametric Retransformation Method", Journal of the American Statistical Association, 78, 605-610.

Ebbeler, Donald H. (1973), "A Note on Large-Sample Approximations in Lognormal Linear Models," Journal of the American Statistical Association, 68, 231-4.

Erdelyi, A. (1954), Table of Integral Transforms, Vol. 1, New York: McGraw-Hill Book Company, Inc.

Evans, I. G. and Shaban, S. A. (1974), "A Note on Estimation in Lognormal Models," Journal of the American Statistical Association, 69, 779-781.

__________ (1976), "New Estimators (of smaller MSE) for Parameters of a Lognormal Distribution", Biometrische Zeitschriff, 18, 453-466.

Finney, D. J. (1941), "On the Distribution of a Variate Whose Logarithm Is Normally Distributed," Supplement to the Journal of the Royal Statistical Society, 155-61.

Land, C. E. (1971), "Confidence Intervals for Linear Functions of the Normal Mean and Variance," Annals of Mathematical Statistics, 42, 1187-1205.

__________ (1973), "Standard Confidence Limits for Linear Functions of the Normal Mean and Variance," Journal of the American Statistical Association, 68, 960-963.

Laurent, A. G. (1963), "The Lognormal Distribution and the Translation Method", Journal of the American Statistical Association, 58, 231-235.

Mehran, F. (1973), "Variance of the MVUE for the Lognormal Mean", Journal of the American Statistical Association, 68, 726-727.

Shaban, S. A. (1981), "On the Estimation of the Ratio of Means and Other Characteristics of Two Log Normal Variates", Biometrical Journal, 23, 357-369.

Teekens, R. and Koerts, J. (1972), "Some Statistical Implications of the Log Transformation of Multiplicative Models," Econometrica, 40, 793-819.

Zellner, A. (1971), "Bayesian and Non-Bayesian Analysis of the Lognormal Distribution and Lognormal Regression", Journal of the American Statistical Association, 66, 327-330.