# IMPROVED ESTIMATION IN LOGNORMAL REGRESSION MODELS

by

Andrew L. Rukhin
Purdue University

Technical Report #85-27

Department of Statistics
Purdue University

October 1985

*This research supported by NSF grant DMS-8901996
AMS 1980 subject classifications: Primary 62F10, Secondary 62H12, 62J02, 62P20.

Improved Estimation in Lognormal Regression Models

by

Andrew L. Rukhin

Purdue University

Lognormal regression model with unknown error variance is considered. We give a class of estimators of the regression coefficients vector improving upon traditional estimator when the number of independent variables is at least three. The relationship between these estimators on one hand and James-Stein type estimators of the normal mean and improved estimators of the normal variance on another hand is discussed.

Key words and phrases: Lognormal regression models, inadmissibility, quadratic loss, James-Stein estimators, normal variance.

1. Introduction

In this paper we consider a lognormal regressiom model which is often used to describe a production process in economics (see Dhrymes (1962), Goldberger (1968), Zellner (1971)). Namely, we assume that

$$z_i = \log y_i = \theta_1 x_{i1} + \ldots + \theta_k x_{ik} + \varepsilon_i, \quad i = 1,\ldots,n \qquad (1.1)$$

where x's are given values of independent (explanatory) variables, $\theta$'s are unknown regression coefficients, and $\varepsilon$'s are independent normal errors each with zero mean and unknown variance $\sigma^2$. The observations y's in (1.1) typically represent economic output at n successive time moments.

Under this model the minimum variance linear unbiased estimate of $\theta$ is the least squares regression estimator $\tilde{\theta} = (X^T X)^{-1} X^T Z$. We assume that the design matrix X has full rank k. It is well known that $\tilde{\theta}$ has multivariate normal distribution with mean $\theta$ and covariance matrix $\sigma^2 \sum$ where $\sum = (X^T X)^{-1}$. Also the scaled residual sum of squares

$$(Z - X\tilde{\theta})^T (Z - X\tilde{\theta})/\sigma^2 = S^2/\sigma^2$$

has $\chi^2$ distribution with n-k degress of freedom.

In the problem of predicting next value of y or that of evaluating regression coefficients it is often of interest to estimate $E \exp(\theta_i) = \exp\{\theta_i + b_i \sigma^2\}$ rather than $E\tilde{\theta}_i = \theta_i$. For instance in the Health Insurance study mentioned by Duan (1983) one needs to make conclusions about mean dollar expenditures, not log-dollar-expenditures. Thus we consider here the estimation problem of the parametric vector $\xi = \theta + \sigma^2 \Sigma b$, where b is a given vector, on the basis of normal observation U with mean $\theta$ and covariance matrix $\sigma^2\Sigma$. It is also assumed that statistic $S^2$ such that $S^2/\sigma^2$ has $\chi^2$ distribution is available.

The loss function L is considered in this paper has the form

$$L(\theta,\sigma;\delta) = ||\delta - \xi||^2 = ||\delta - \theta - \sigma^2\Sigma b||^2. \quad (1.2)$$

All our results are also ture for more general quadratic loss $\Sigma w_i(\delta_i - \theta_i - \sigma^2 b_i)^2$ where $w_i$ are positive weights.

Traditional estimators $\delta$ of $\xi$ have the form

$$\delta(U,S) = U + S^2 d$$

for some vector d. It is easy to see that under (1.2) the best choice of d is $d = d_0 = \Sigma b/(\nu + 1)$ where $\nu$-1 is the number of degrees of freedom of $\chi^2$-distribution. Thus we shall study the admissibility of estimator

$$\delta_0(U,S) = U + S^2\Sigma b/(\nu+1).$$

Notice that we can and will assume that $\Sigma = I$. Indeed the inadmissibility of $\delta_0$ in the case $\Sigma = I$ is equivalent to inadmissibility of $\delta_0$ for arbitrary $\Sigma$. (If an estimator $\delta_1(U,S)$ is better than $\delta_0$ for $\Sigma = I$, then $\Sigma^{-\frac{1}{2}} \delta_1(\Sigma^{\frac{1}{2}}U,S)$ improves upon $\delta_0$ for arbitrary positive definite $\Sigma$). The inadmissibility of $\delta_0$ will be proven for $k \geq 3$.

Notice that the admissibility of this estimator even for $k = 1$ is not evident. In fact the natural estimator of $\sigma^2$, $S^2/(\nu+1)$ is not admissible when $\theta$ is unknown (see Stein (1964), Brown (1968), Brewster and Zidek (1974), Strawderman (1974)). Rukhin (1986) proved that $\delta_0$ is inadmissible when $k = 1$ by showing that Stein's necessary condition for admissibility is not met. However explicit improvements over $\delta_0$ are not known when $k = 1$. Analogously to the estimation problem of a normal mean, the traditional estimator is inadmissible when $k \geq 3$, and we demonstrate this fact by constructing a class of procedures with uniformly smaller risk. The main difficulty is that our problem is not invariant under usual groups of transformations and that the risk is not a function of maximal invariant as in many other multivariate estimation problems.

## 2. Inadmissibility Result

In this Section U denotes normal random k-dimensional vector with unknown mean $\theta$ and covariance matrix $\sigma^2 I$, and $S^2/\sigma^2$ is a chi-square random variable with $\nu-1$ degrees of freedom. To estimate parametric vector $\theta + \sigma^2 b$, where b is given, we consider a class of estimators of the form

$$\delta(U,S) = (1-rS^2/\| U \|^2)\, U + S^2(1 - \phi(\| U \| /S))b/(\nu+1)). \quad (2.1)$$

where r is a constant and $\phi$ is a positive differentiable function.

Theorem. Estimators (2.1) improve upon $\delta_0$ for loss function (1.2) if $0 < r < 2(k-2)/(\nu+1)$ and $\phi(Z) = \psi(Z)(1 + Z^2)^{-c}$, $c > 0$, where $\psi$ is a nonincreasing function

$$0 \leq \psi(Z) \leq D/A,$$

with

$$D = D(c,\nu) = \min\{2/(2+c),$$

$$[B((\nu+7)/2+c,\tfrac{1}{2})\, B((\nu+3)/2,2) - B((\nu+3)/2+c,\tfrac{1}{2})\, B((\nu+4)2,2)]$$

$$/B((\nu+3)/2,2)\, B((\nu+7)/2+2c,\tfrac{1}{2})\}, \quad (2.2)$$

and $A = A(r,\nu)$ is defined by (2.7).

Proof. One has for an estimator (2.1)

$$R(\theta,\sigma;\delta)$$

$$= \sum_1^k E_{\theta\sigma}(U_i - rS^2U_i/\| U \|^2 - \theta_i)^2$$

$$+ \| b \|^2 E_{\theta\sigma}\, (S^2(1-\phi)/(\nu+1) - \sigma^2)^2$$

$$+ 2\sum_{i}^{k} b_i E_{\theta\sigma}(U_i - rS^2U_i/\| U \|^2 - \theta_i)(S^2(1-\phi)/(\nu+1)-\sigma^2)$$

$$= R_1 + \| b \|^2 R_2 + 2\sum b_i Q_i. \tag{2.3}$$

Integrating by parts (see Efron and Morris (1976), Stein (1981)) one obtains

$$R_1 = \sigma^2 k - 2rE_{\theta\sigma}(U_i-\theta_i)U_i/\| U \|^2 E_{\theta\sigma}S^2 + r^2 E_{\theta\sigma}\| U \|^{-2} E_{\theta\sigma}S^4$$

$$= \sigma^2[k+(r^2(\nu^2-1) - 2r(\nu-1)(k-2))E_{\eta 1} \| U \|^{-2}]$$

with $\eta = \theta/\sigma$.

It follows that

$$R_1 \leq \sum_i E_{\theta\sigma}(U_i - \theta_i)^2 = k\sigma^2$$

if

$$0 < r < 2(k-2)/(\nu+1)$$

and the optimal choice of r is

$$r = r_0 = (k-2)/(\nu+1)$$

Now we consider the last term in (2.3). Integrating by parts one obtains

$$Q_i = rE_{\theta\sigma}(U_i - rS^2U_i/\| U \|^2 - \theta_i)(S^2(1-\phi)/(\nu+1)-\sigma^2)$$

$$= rE_{\theta\sigma}U_i\phi S^4/[\| U \|^2 (\nu+1)]$$

$$- E_{\theta\sigma}(U_i - \theta_i)S^2\phi/(\nu+1)$$

$$= \sigma^3(\nu+1)^{-1}E_{\eta 1}U_iS\| U \|^{-2}(rS^3\phi - \| U \| \phi') . \qquad (2.4)$$

Notice that with a normalizing constant $f(\| U \|)$

$$\| U \| \ E_{\eta 1} \{S\phi'/U\} = f(\| U \|) \ \| U \| \int_0^\infty s^{\nu-1} e^{-s^2/2}\phi'(\| U \| /s)ds$$

$$= f(\| U \| ) \int_0^\infty [(\nu+1)s^\nu - s^{\nu+2}] e^{-s^2/2} \phi(\| U \| /s)ds$$

$$= E_{\eta 1} \{[(\nu+1)S^2-S^4]\phi/U\},$$

so that

$$Q_i = \sigma^3(\nu+1)^{-1} E_{\eta 1}U_iS^2\| U \|^{-2} [(r+1)S^2-\nu-1].$$

Thus $R(\theta,\sigma;\delta)/\sigma^2$ is a quadratic polynomial in $\sigma$, and

$$R(\theta,\sigma:\delta) < R(\theta,\sigma;\delta_0)$$

if

$$[(\nu+1)^{-1} \sum_1^k b_i \, E_{\eta 1} U_i S^2\phi \| U \|^{-2}((r+1)S^2-\nu-1)]^2$$

$$\leq \| b \|^2 \, r(\nu-1)(2(k-2)-r(\nu+1)) \, E_{\eta 1} \| U \|^{-2}$$

$$\times \, E_{\eta 1}[(S^2/(\nu+1)-1)^2 - (S^2(1-\phi)/(\nu+1) -1)^2] \, . \qquad (2.5)$$

Because of Cauchy-Schwarz inequality

$$[\sum_1^k b_i E_{\eta 1} U_i S^2\phi \| U \|^{-2} ((r+1)S^2-\nu-1)]^2$$

$$\leq \| b \|^2 (\nu^2-1)E_{\eta 1} \| U \|^{-2} ((r+1)S^2-\nu-1)^2 E_{\eta 1} S^4\phi$$

$$= \| b \|^2 (\nu^2-1)(r^2(\nu^2+1)+2\nu) \, E_{\eta 1} S^4\phi E_{\eta 1} \| U \|^{-2} ,$$

so that (2.5) would follow from the following inequality

$$E_{\eta 1}S^4\phi \le 2r(2(k-2)-r(\nu+1))(\nu+1)^{-2}(r2(\nu^2+1)+2\nu)^{-1}$$

$$\times E_{\eta 1}S^2\phi(S^2-S^2\phi/2-\nu-1)$$

or for all $\eta$

$$E_{\eta 1}S^2\phi(S^2-AS^2\phi-\nu-1) \tag{2.6}$$

where

$$A = \tfrac{1}{2} + (r^2(\nu^2+1)+2\nu)(\nu+1)^2/[2r(2(k-2)-r(\nu+1))]\,. \tag{2.7}$$

To prove (2.6) we use the approach of Strawderman's paper (1974) where a class of functions $\phi$ satisfying to (2.6) for $A = \frac{1}{2}$ has been found. Analysis of Strawderman's proof of Theorem 1 shows that functions $\phi$ specified in the condition of our Theorem meet condition (2.6).

## 3. Discussion

Theorem of Section 2 shows that an improvement upon traditional estimator $U + S^2b/(\nu+1)$ of $\theta + \sigma^2 b$ is obtained by combining Stein-James improvement $(1-rS^2/\|U\|^2)U$ over $U$ as an estimator of $\theta$ and Strawderman's improvement $S^2(1-\phi)$ over $S^2$ as an estimator of $\sigma^2$. It seems to be natural to expect that such a combination will be an improvement itself. However this is far from being true in general, namely, a linear combination of improved estimators of $\theta$ and $\sigma^2$ will not always provide a better estimator of the combination of $\theta$ and $\sigma^2$. (For instance, $U+S^2(1-\phi)b/(\nu+1)$ is not better than $U+S^2b/(\nu+1)$ for any $\phi$ when $k = 1$, see Rukhin (1986)). The key fact which guarantees the improved character of estimator (2.1) is uncorrelatedness of Stein-James estimator $(1-rS^2/\|U\|^2)U$ and $S^2/(\mu+1)-\sigma^2$ which is essentially formula (2.4).

Also notice that the technique of solving differential inequalities which arise from the unbiased estimates of the risk function does not work in our case. Indeed let $\delta = \delta_0-2h$, where $h = h(U,S)$ is a smooth vector-function. Then

$$\sigma^{-2}E_{\theta\sigma}[(\delta_0-\xi)^2-(\delta_0-2h-\xi)^2] = 4E_{\theta\sigma}Dh$$

where

$$Dh = -(\nu-3)\ \|h\|^2\ S^{-2}-2\ \|h\|\ S^{-1}\frac{\partial}{\partial s}\ \|h\|$$

$$+\sum\frac{\partial}{\partial u_i}h_i + (\nu+1)^{-1}\sum_1^k b_i[S\frac{\partial}{\partial s}\ h_i - 2h_i].$$

However it can be shown that the differential inequality $Dh \geq 0$ does not have any nontrivial solutions.

Of course for practical use of estimators (2.1) statistician has to specify the origin and the scale unit so that $\delta$ will differ from $\delta_0$ and the improvement of $\delta$ over $\delta_0$ will be substantial (see Berger (1982) for the case $b \equiv 0$). For this purpose a wider class of better estimators may be needed.

Some other loss functions, like,

$$L_T(\theta,\sigma;\delta) = \sum (\delta_i - \exp(\theta_i+b_i\sigma^2))^2$$

are also of interest. Notice however that the best unbiased estimator is badly inadmissible even when $k = 1$ (see Rukhin (1985)).

# REFERENCES

Berger, J. (1982). Selecting a minimax estimator of a multivariate normal mean. Ann. Statist. 10, 81-92.

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.

Dhrymes, P. J. (1962). On devising unbiased estimators for the parameters of the Cobb-Douglass production function. Econometrica, 30, 297-304.

Duan, N. (1983). Smearing estimate: a nonparametric retransformation method. J. Amer. Statist. Assoc., 78, 605-610.

Efron, B., and Morris, C. (1976). Families of minimax estimators of the mean of a multivariate normal distribution. Ann. Statist. 4, 11-21.

Goldberger, A. S. (1963). The interpretation and estimation of Cobb-Douglas functions. Econometrica 36, 464-472.

Rukhin, A. (1985). Improved estimation in lognormal models, submitted to J. Amer. Statist. Assoc.

Rukhin, A. (1986). Estimating a linear function of the normal mean and variance. Sankhya. Ser A, 48.

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. (1981). Estimation of the mean of a multivariate normal distribution. Ann. Statist. 9, 1135-1151.

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

Zellner, A. (1971). Bayesian and non-Bayesian analysis of the lognormal distribution and lognormal regression. J. Amer. Statist. Assoc. 66, 327-330.