Econometrics Working Paper EWP0004
ISSN 1485-6441

Preliminary-Test and Bayes Estimation of a Location Parameter Under 'Reflected Normal' Loss

David E. A. Giles *

Department of Economics
University of Victoria
Canada

April, 2000

Abstract:

In this paper, we consider a simple preliminary-test estimation problem where the analyst's loss structure is represented by a 'reflected Normal' penalty function. In particular we consider the estimation of the location parameter in a Normal sampling problem, where a preliminary test is conducted for the validity of a simple restriction on this parameter. The exact finite-sample risk of this pre-test estimator is derived under 'reflected Normal' loss, and this risk is compared with those of the unrestricted and restricted Maximum Likelihood estimators of location under this loss structure. The paper draws comparisons between these results and those obtained under conventional quadratic loss. Some simple Bayesian analysis is also considered. The results extend naturally to the case of estimating the coefficients in a Normal linear multiple regression model.

Author Contact: Professor David Giles, Department of Economics, University of Victoria, P.O. Box 1700 STN CSC, Victoria, B.C., Canada, V8W 2Y2
E-mail: dgiles@uvic.ca; Voice: (250) 721 8540; FAX: (250) 721 6214

1. Introduction

In statistics and econometrics, the expression "preliminary-test estimation", refers to a situation where the choice of estimator for some parameter (vector) is essentially randomized through the prior application of an hypothesis test. The test need not relate to the parameters of interest - for example, it could relate to a set of nuisance parameters. Neither is it necessary that the same set of data be used for the prior test as for the primary estimation problem. This randomization of the choice of estimator complicates its sampling properties significantly, as was first recognized by Bancroft (1944). Extensive surveys of the subsequent literature on preliminary-test estimation are given by Bancroft and Han (1977) and Giles and Giles (1993).

Preliminary-test estimators generally have quite adequate (if not optimal) large-sample properties. For example, if the "component" estimators (between which a choice is made), and the prior test, are each consistent, then the pre-test estimator will also be consistent. On the other hand, as pre-test estimators are discontinuous functions of the data, it is well known (*e.g.*, Cohen (1965)) that they are inadmissible under conventional loss functions. This is because they are discontinuous functions of the data. Despite their inadmissibility, and the fact that generally they are not mini-max either, pre-test estimators are of considerable practical interest for at least two reasons. First, they represent estimation strategies of precisely the type that are frequently encountered in practice. Typically, in many areas of applied statistics, a model will be estimated and then the model will be subjected to various specification tests. Subsequently, the model in question may be simplified (or otherwise amended) and then re-estimated. Second, in all of the pre-test problems that have been considered in the literature, pre-test estimators dominate each of their "component estimators" over different parts of the parameter space, in terms of risk. Indeed, in some cases (*e.g.*, Giles, 1991) they may even dominate their components simultaneously over the *same* part of the parameter space.

Although the sampling properties of various preliminary-test estimators have been studied by a range of authors, little is known about their complete sampling distributions. The only exceptions appear to be the results of Giles (1992), Giles and Srivastava (1993), Ohtani and Giles (1996b), and Wan (1997). Generally the finite-sample properties of pre-test estimators have been evaluated in terms of risk, and usually the latter have been based on the assumption that the loss function is quadratic. Some exceptions (using absolute-error loss, the asymmetric "LINEX" loss function, or "balanced" loss) include the contributions of Ohtani *et al*. (1997), Giles *et al*. (1996), Ohtani and Giles (1996), Giles and Giles (1996), and Geng and Wan (2000), among others.

Despite its tractability and historical interest, two obvious practical shortcomings of the quadratic loss function are its unboundedness and its symmetry. Recently, Spiring (1993) has addressed the first of these weaknesses (and to a lesser degree the second), by analyzing the “reflected Normal” loss function. He motivates this loss with reference to problems in quality assurance (*e.g.*, Taguchi 1986). The reflected Normal loss function has the particular merit that it is bounded. It can readily be made asymmetric if this is desired for practical reasons. An alternative loss structure is the “bounded LINEX” (or “BLINEX”) loss discussed by Levy and Wen (1997a,b) and Wen and Levy (1999a,b). The BLINEX loss is both bounded and asymmetric.

In this paper, we consider a simple preliminary-test estimation problem where the analyst’s loss structure is “reflected Normal”. Specifically, we consider the estimation of the location parameter in a Normal sampling problem, where a preliminary test is conducted for the validity of a simple restriction on this parameter. The exact finite-sample risk of this pre-test estimator is derived under reflected Normal loss, and this risk is compared with those of both the unrestricted, and restricted, Maximum Likelihood estimators. This appears to be the first study of a pre-test estimator when the loss structure is bounded, and comparisons are drawn between these results and those obtained under conventional (unbounded) quadratic loss. Our results extend naturally to the case of estimating the coefficients in a Normal linear multiple regression model. Although we consider only a symmetric loss function in this paper, the extension to the asymmetric case is also straightforward.

In the next section we formulate the problem and the notation. Exact expressions for the risk functions are derived in section 3, and these are evaluated, illustrated and discussed in section 4. Some related Bayesian analysis is provided in section 5; and section 6 offers some concluding remarks, and suggests some directions for further research.

2. Formulation of the Problem

The problem that we consider here is cast in simple terms to facilitate the exposition. However, the reader will recognize that it generalizes trivially to more interesting situations, such as the estimation of the coefficient vector in a standard linear multiple regression model when potentially there are exact linear restrictions on the coefficients. In that sense, our analysis here extends that of Judge and Bock (1978), and others[1], through the consideration of a different loss structure. We will be concerned with the estimation of the location parameter, μ, in a Normal population with unknown scale parameter. We have a simple random sample of values:

$$x_i \sim \text{ i.i.d. } N[\mu, \sigma^2] \quad ; \quad i = 1, 2, 3, \ldots\ldots\ldots., n$$

and we will be concerned with a prior "t-test" of

$$H_0 : \mu = \mu_0 \quad vs. \quad H_A : \mu \neq \mu_0 .$$

The choice of estimator for μ will depend on the outcome of this preliminary test. If H_0 were false we would use the unrestricted maximum likelihood estimator (UMLE), $\mu_1 = (\sum_{i=1}^{n} x_i / n) = \bar{x}$. On the other hand, if H_0 were true we would use the restricted maximum likelihood estimator (RMLE), which here is just μ_0 itself. So, the preliminary-test estimator (PTE) of μ in this situation is

$$\mu_p = [I_R(t) \times \mu_1] + [I_A(t) \times \mu_0],$$

where "t" is the usual t-statistic for testing H_0, defined as

$$t = [\mu_1 - \mu_0] / [s^2 / n]^{½}$$

and

$$s^2 = [\sum_{i=1}^{n} (x_i - \bar{x})^2] / (n-1) .$$

It will be more convenient to use $F = t^2$ as the test statistic, recalling that $F \sim F(1, n\text{-}1; \lambda)$ where the non-centrality parameter is $\lambda = n\delta^2 / (2\sigma^2)$, and $\delta = (\mu - \mu_0)$. So, the PTE of μ may be written as

$$\mu_p = [I_R(F) \times \mu_1] + [I_A(F) \times \mu_0] , \qquad \textbf{(1)}$$

where $I_\Omega(.)$ is an indicator function taking the value unity if its argument is in the subscripted interval, and zero otherwise. In our case, the rejection region is the set of values $R = \{ F : F > c_\alpha \}$ and the "acceptance (non-rejection) region for the "t-test" is $A = \{ F : F \leq c_\alpha \}$, where c_α is critical value for a chosen significance level, α. It should be noted that

$$I_R (F) = [1 - I_A (F)] , \quad \textbf{(2)}$$

$$[I_R (F) \times I_A (F)] = 0 , \quad \textbf{(3)}$$

and

$$[I_\Omega (F)]^p = I_\Omega (F) ; \Omega = \{A, R\}; \; p \geq 1. \quad \textbf{(4)}$$

If we let θ be a scalar parameter to be estimated, and let τ be a statistic used as an estimator of θ, then the "reflected Normal" loss function is defined in the following way:

$$\mathcal{L}(\tau, \theta) = K \{1 - \exp [- (\tau - \theta)^2 \; / (2\gamma^2)]\} , \quad \textbf{(5)}$$

where K is the maximum loss; and γ is a pre-assigned shape parameter that controls the rate at which the loss approaches its upper bound. For example, if we set $\gamma = (\Delta/4)$, for some Δ, then $\mathcal{L} \geq (0.9997K)$ for all values $\tau > \theta \pm \Delta$. The "reflected Normal" loss structure arise in the context of "M-estimation" (*e.g.*, Huber, 1977), and in the context of robust estimation its influence function is known to have rather good properties. Figure 1 compares the "reflected Normal" and conventional quadratic loss functions.

3. Derivation of the Risk Functions

The risk function of τ as an estimator of θ is $\Re(\tau) = E [\mathcal{L}(\tau, \theta)]$, where expectation is taken over the sample space. Let us consider the risks of the RMLE, UMLE and PTE estimators of μ in turn.

RMLE:

In our case the risk of the RMLE is trivial, as μ_0 is a constant, and is simply

$$\Re(\mu_0) = K \{1 - \exp [- \delta^2 \; / (2\gamma^2)]\}. \quad \textbf{(6)}$$

UMLE:

$$\Re(\mu_1) = K\int_{-\infty}^{\infty} \{1 - \exp[-(\mu - \mu_1)^2 / (2\gamma^2)]\} p(\mu_1)\, d\mu_1 , \qquad \textbf{(7)}$$

where

$$p(\mu_1) = (n / 2\pi\sigma^2)^{½} \exp[-(\mu_1 - \mu)^2 / (2\sigma^2 / n)]. \qquad \textbf{(8)}$$

Substituting (8) in (7), completing the square, and using the result that a Normal density integrates to unity, we obtain:

$$\Re(\mu_1) = K \{1 - \gamma / [(\sigma^2 / n) + \gamma^2]\}. \qquad \textbf{(9)}$$

Before proceeding to the derivation of the risk function for the PTE of μ, some comments on the risks of these "component estimators" are in order. First, as is the case with a quadratic loss function, $\Re(\mu_1)$ does not depend on the (squared) "estimation error", δ^2, and hence is also constant with respect to λ. Second, as is also the case with a quadratic loss function for this problem, $\Re(\mu_0)$ is an increasing function of δ^2 (or λ). Under quadratic loss this risk increases linearly with λ, and so it is unbounded. Here, however, it increases from zero at a decreasing rate, and *limit* $[\Re(\mu_0)] = K$ when $\delta^2 \rightarrow \infty$. That is, the risk of the RMLE is bounded. Finally, equating $\Re(\mu_0)$ and $\Re(\mu_1)$, we see that the risk functions intersect at

$$\delta^2 = -2\gamma^2 \ln\{\gamma / [(\sigma^2 / n) + \gamma^2]\}.$$

These results are reflected in the figures in the next section.

PTE:

The derivation of the risk of the preliminary-test estimator of μ is somewhat more complex, and we will use the following result from Clarke (1986, Appendix 1):

Lemma

Let w be a non-central chi-square variate with g degrees of freedom and non-centrality parameter θ, let $\phi(\cdot)$ be any real-valued function, and let n be any real value such that $n > (-g/2)$. Then:

$$E[w^n \phi(w)] = 2^n \sum_{m=0}^{\infty} (e^{-\theta}\theta^m / m!) [\Gamma(½(g + 2n + 2m)) / \Gamma(½(g + 2m))] E[\phi(\chi^2(g + 1 + 2m))].$$

■

Now, to evaluate $\Re(\mu_p)$, we note that is can be written as

$$\Re(\mu_p) = K\{1 - E(\exp[-(\mu - \mu_p)^2 / (2\gamma^2)])\}$$

$$= K \sum_{r=1}^{\infty} \{[E(\mu_p - \mu)^{2r}] / [(-1)^r (2\gamma^2)^r (r!)]\}.$$

From (1), we have:

$$(\mu_p - \mu) = [I_R(F) \times \mu_1] + [I_A(F) \times \mu_0] - \mu$$

$$= I_R(F)(\mu_1 - \mu) + \delta I_A(F),$$

and so, using (2) to (4):

$$(\mu_p - \mu)^{2r} = (\mu_1 - \mu)^{2r} + [\delta^{2r} - (\mu_1 - \mu)^{2r}] I_A(F); \quad r = 1, 2, 3, \ldots\ldots$$

Then,

$$\Re(\mu_p) = K \sum_{r=1}^{\infty} \{[E(\mu_1 - \mu)^{2r} + E\{I_A(F)[\delta^{2r} - (\mu_1 - \mu)^{2r}]\}] / [(-1)^r (2\gamma^2)^r (r!)]\}. \quad \textbf{(10)}$$

Now, from the moments of the Normal distribution (e.g., Zellner, 1971, pp. 364-365),

$$E(\mu_1 - \mu)^{2r} = [2r(\sigma^2 / n)^r / (2\pi)] \Gamma(r + ½); \quad r = 1, 2, 3, \ldots\ldots. \quad \textbf{(11)}$$

Also,

$$E\{I_A(F)[\delta^{2r} - (\mu_1 - \mu)^{2r}]\} = E\{I_A(F)[\delta^{2r} - \sum_{j=0}^{2r} \delta^j (\mu_1 - \mu_0)^{2r-j} ({}^{2r}C_j)]\}$$

$$= E\{ I_A(F)[\delta^{2r} - \sum_{j=0}^{2r} \delta^j (n\chi^2_{(1;\lambda)}/\sigma^2)^{r-j/2}(^{2r}C_j)]\}, \qquad \textbf{(12)}$$

where $\chi^2_{(q;\ \lambda)}$ denotes a non-central chi-square variate with "q" degrees of freedom, and non-centrality parameter, λ ($= n\delta^2/(2\sigma^2)$).

Recalling that $F = [(n-1)(\chi^2_{(1;\lambda)}/\chi^2_{(n-1;0)})]$, where the two chi-square variates are independent, we can re-express (12) as:

$$E\{ I_A(F)[\delta^{2r} - (\mu_1 - \mu)^{2r}]\} = \delta^{2r}\,\Pr.[(\chi^2_{(1;\lambda)}/\chi^2_{(n-1;0)}) < c^*_\alpha]$$

$$- \sum_{j=0}^{2r} \{ \delta^j (^{2r}C_j)(n/\sigma^2)^{r-j/2} E[I_A((n-1)\chi^2_{(1;\lambda)}/\chi^2_{(n-1;0)})(\chi^2_{(1;\lambda)})^{r-j/2}]\}, \qquad \textbf{(13)}$$

where $c^*_\alpha = [c_\alpha/(n-1)]$.

The expectation in (13) can be evaluated by repeatedly using the result of Clarke (1986), stated in the above Lemma, and the independence of the associated chi-square variates[2]:

$$E[I_A((n-1)\chi^2_{(1;\lambda)}/\chi^2_{(n-1;0)})(\chi^2_{(1;\lambda)})^{r-j/2}]$$

$$= 2^{r-j/2} \sum_{i=0}^{\infty} (e^{-\lambda}\lambda^i/i!)[\Gamma(½(1+2r-j+2i))/\Gamma(½(1+2i))]$$

$$\times \Pr.[(\chi^2_{(1+2r-j+2i;\lambda)}/\chi^2_{(n-1)}) < c^*_\alpha]. \qquad \textbf{(14)}$$

Finally, using the results in (11) - (14) we can write the risk of the pre-test estimator, (10), as:

$$\Re(\mu_p) = K \sum_{r=1}^{\infty} [1 / \{(-1)^r (2\gamma^2)^r (r!)\}] \Big\{ [2r(\sigma^2/n)^r/(2\pi)]\, \Gamma(r + ½)$$

$$+ \delta^{2r} \Pr.[(\chi^2_{(1;\lambda)} / \chi^2_{(n-1;0)}) < c^*_\alpha] - \sum_{j=0}^{2r} \{\delta^j ({}^{2r}C_j)(n/\sigma^2)^{r-j/2}\, 2^{r-j/2}$$

$$\times \sum_{i=0}^{\infty} (e^{-\lambda}\lambda^i / i!)\, [\Gamma(½(1 + 2r - j + 2i)) / \Gamma(½(1 + 2i))]\, \Pr.[(\chi^2_{(1+2r-j+2i;\lambda)} / \chi^2_{(n-1)}) < c^*_\alpha]\} \Big\}$$

(15)

4. Some Illustrative Evaluations

The risk functions for our various estimators under a conventional quadratic loss function are well known[3], and for the comparative purposes they are illustrated in Figure 2. The risk functions for the restricted and unrestricted maximum likelihood estimators under "reflected Normal" loss, as in (6) and (9), are easily evaluated for particular choices of the parameters and sample size, and these are illustrated in Figure 3. In particular, we see there that $\Re(\mu_0)$ is bounded above by K. The evaluation of the risk of the preliminary-test estimator is rather more tedious, but it can readily be verified by simulation. Some examples of this appear in Figures 4 to 6. There, the range of values for δ^2 is such that the boundedness of $\Re(\mu_0)$ is not visually apparent.

In particular, Figure 4 compares $\Re(\mu_p)$ with $\Re(\mu_0)$ and $\Re(\mu_1)$ for a small sample size (n = 10), and illustrative parameter values. The general similarity between these results, and their counterparts under quadratic loss (as in Figure 2), is striking. In particular, there is a region of the parameter space where μ_p is least preferred among the three estimators under consideration. Similarly, there are regions where each of μ_0 and μ_1 are least preferred. There are regions where either μ_0 or μ_1 are most preferred among the three estimators, but there is no region of the parameter space where the pre-test estimator is preferred over both μ_0 and μ_1 simultaneously. The effect of increasing the sample size from n = 10 to n = 100 can be seen by comparing Figures 4 and 5. In each of these figures the convergence of $\Re(\mu_p)$ to $\Re(\mu_0)$, as $\delta^2 \rightarrow \infty$, is as expected. The preliminary test has a power function that approaches unity in this case, so in the limit the PTE and the UMLE of μ coincide.

Figure 6 depicts the effects of increasing the significance level for the preliminary test from 5% to 10%. A larger significance level implies that greater weight is given to the UMLE when δ^2 (when H_0 is true). An increase in the size of the test also increases the power, so μ_p also gives greater weight to μ_1 when $\delta^2 > 0$. It is clear that in principle it should be possible bound the region in which $\Re(\mu_p)$ and $\Re(\mu_0)$ intersect, as is done in the quadratic loss case by Judge and Bock (1978, p.73). However, the extremely complex nature of the expression for the former risk function (in (15)) makes this rather impractical from an analytical standpoint.

5. Bayesian Estimation

In this section we briefly consider the Bayes estimator of μ under reflected normal loss, as this estimator has the desirable attribute of being admissible, if the prior p.d.f. is "proper". We will take the Bayes estimator, τ_B, of a parameter θ to be the "minimum (posterior) expected loss" (MELO) estimator. That is, τ_B minimizes

$$EL = \int_{-\infty}^{\infty} \mathfrak{L}(\tau_B, \theta)\, p(\theta \mid x)\, d\theta\,, \qquad \textbf{(16)}$$

where $p(\theta \mid x)$ is the posterior p.d.f. for θ, given $x = (x_1, \ldots\ldots, x_n)$. If $p(\theta)$ is the prior p.d.f. for θ, it is well known[4] that the τ_B that minimizes (15) will also minimize the Bayes risk,

$$Er(\theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(\theta)\, \mathfrak{L}(\tau_B, \theta)\, p(x \mid \theta)\, dx\, d\theta\,, \qquad \textbf{(17)}$$

as long as the Bayes risk is finite. As is also well known, if $\mathfrak{L}(\tau_B, \theta)$ is quadratic, then τ_B is the mean of $p(\theta \mid x)$.

However, when $\mathfrak{L}(\tau, \theta)$ is reflected normal, there is no simple closed-form expression for the τ_B that minimizes (15), for an arbitrary posterior p.d.f.. Of course, more progress can be made for specific posterior cases. Let us consider some particular choices of prior (and hence posterior) p.d.f. for μ in our problem., assuming that σ^2 is known. The case of unknown σ^2 does not lend itself to a simple analysis, and is not considered further below.

Case (i): Known Variance, Conjugate Prior

Recalling the normality of our data, the natural-conjugate prior for μ is N[μ', σ'] and the posterior for μ is N[μ'', σ'']. It is well known[5] that

$$\mu'' = [\mu_1 \sigma'^2 + \mu'(\sigma^2 / n)] / [\sigma'^2 + (\sigma^2 / n)] \quad \textbf{(18)}$$

and

$$\sigma''^2 = [\sigma'^2 \sigma^2 / n] / [\sigma'^2 + (\sigma^2 / n)] \, . \quad \textbf{(19)}$$

So, under quadratic loss, the Bayes estimator of μ is $\mu_B = \mu''$. The Bayes estimator under the reflected normal loss is the μ_B that minimizes

$$EL = \int_{-\infty}^{\infty} \mathfrak{L}(\mu_B, \mu)\, p(\mu \mid x)\, d\mu \ , \quad \textbf{(20)}$$

where

$$\mathfrak{L}(\mu_B, \mu) = K \{1 - \exp[-(\mu_B - \mu)^2 / (2\gamma^2)]\} \, , \quad \textbf{(21)}$$

and

$$p(\mu \mid x) = \left(\sigma'' \sqrt{2\pi}\right)^{-1} \exp[-(\mu - \mu'')^2 / (2\sigma''^2)] \, ; \qquad -\infty < \mu < \infty \quad \textbf{(22)}$$

Substituting (21) and (22) in (20); setting the derivative of EL with respect to μ_B equal to zero; completing the square on μ; and solving for μ_B, it emerges after a little manipulation that $\mu_B = \mu''$. So, for this case the Bayes estimator of μ is the same under either quadratic or reflected normal loss functions.

Case (ii): Known Variance, Diffuse Prior

In this case the (improper) prior p.d.f. for μ is

$$p(\mu) \propto \text{constant}; \qquad -\infty < \mu < \infty$$

and the corresponding posterior p.d.f. is well known[6] to be

$$p(\mu \mid x) = \sqrt{n}\left(\sigma\sqrt{2\pi}\right)^{-1} \exp[-n(\mu - \mu_1)^2 / (2\sigma^2)] . \qquad \textbf{(23)}$$

That is, the posterior is $N[\mu_1 , \sigma^2 / n]$, and so the Bayes estimator of μ under quadratic loss is $\mu_B = \mu_1$. The corresponding estimator under reflected normal loss is obtained by substituting (23) and (22) in (20); setting the derivative of EL with respect to μ_B equal to zero; completing the square on μ; and solving for μ_B. It transpires after some manipulation that $\mu_B = \mu_1$, so for this case as well the Bayes estimator of μ is the same under either quadratic or reflected normal loss functions.

6. Concluding Remarks

Preliminary-test estimation is commonplace, but often little attention is paid to the implications that such prior testing has for the sampling properties of estimators. When these implications have been studied, generally the analysis has been in terms of the risk function of the pre-test estimator and its "component" estimators. The majority of this risk analysis has been based on very restrictive loss functions, such as quadratic loss. One aspect of such loss structures is that they are symmetric with respect to the "direction" of the estimation error, and this may be unrealistic in practice. This condition has been relaxed by several authors, as is discussed in Section 1. Another feature of conventional loss functions (and the asymmetric ones that have been considered in a pre-test context) is that they are unbounded as the estimation error grows. This may also be unrealistic in practice. The (bounded) "reflected normal" loss function is considered in this paper, in the context of estimating a Normal mean after a pre-test of a simple restriction. With this loss structure the risk of the restricted maximum likelihood "estimator" is also bounded, in contrast to the situation under quadratic loss. In other respects, however, the *qualitative* risk properties of the pre-test estimator under reflected normal loss are the same as under quadratic loss. Interestingly, the Bayes estimator of the mean is the same under both loss structures, with either a conjugate or diffuse prior, at least in the case where the precision of the process is known.

Figure 1: Reflected Normal & Quadratic Loss Functions
(n = 10; sigma = 1; K = 5; gamma = 1)

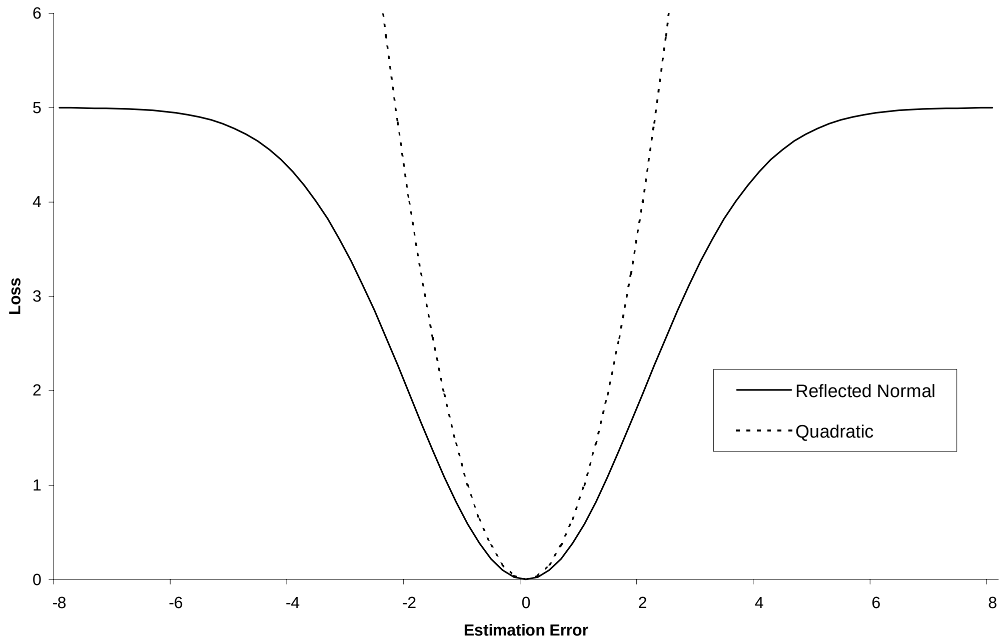


Figure 2: Risks Under Quadratic Loss
(n=10; sigma = 1; K = 1; gamma = 1)

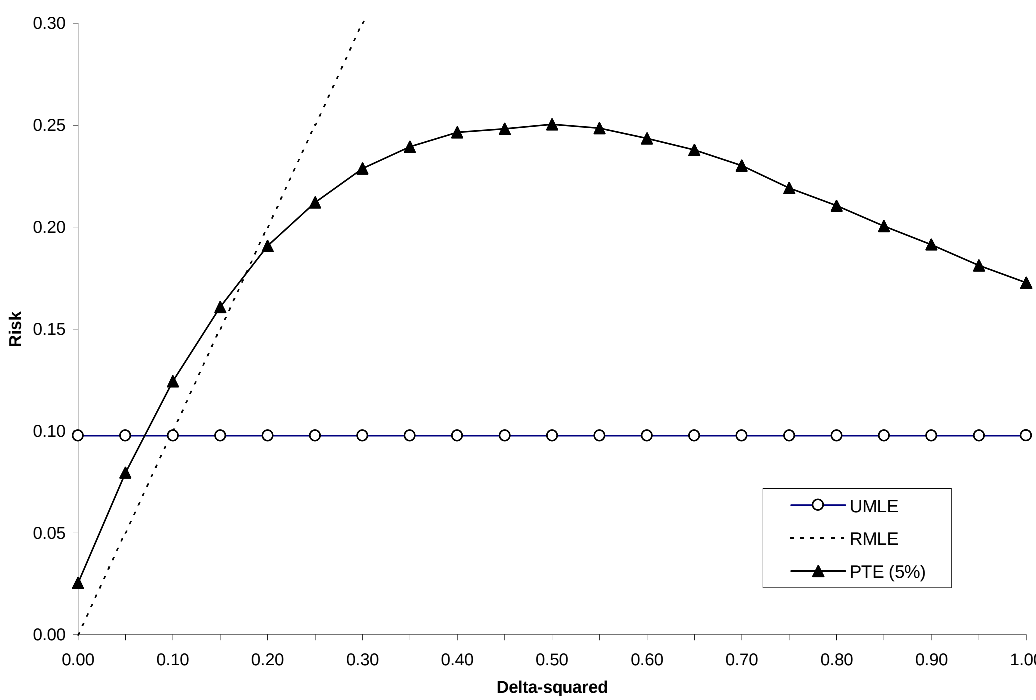


Figure 3: Risk Under Reflected Normal Loss
(n = 10; sigma = 1; K = 1; gamma = 1)

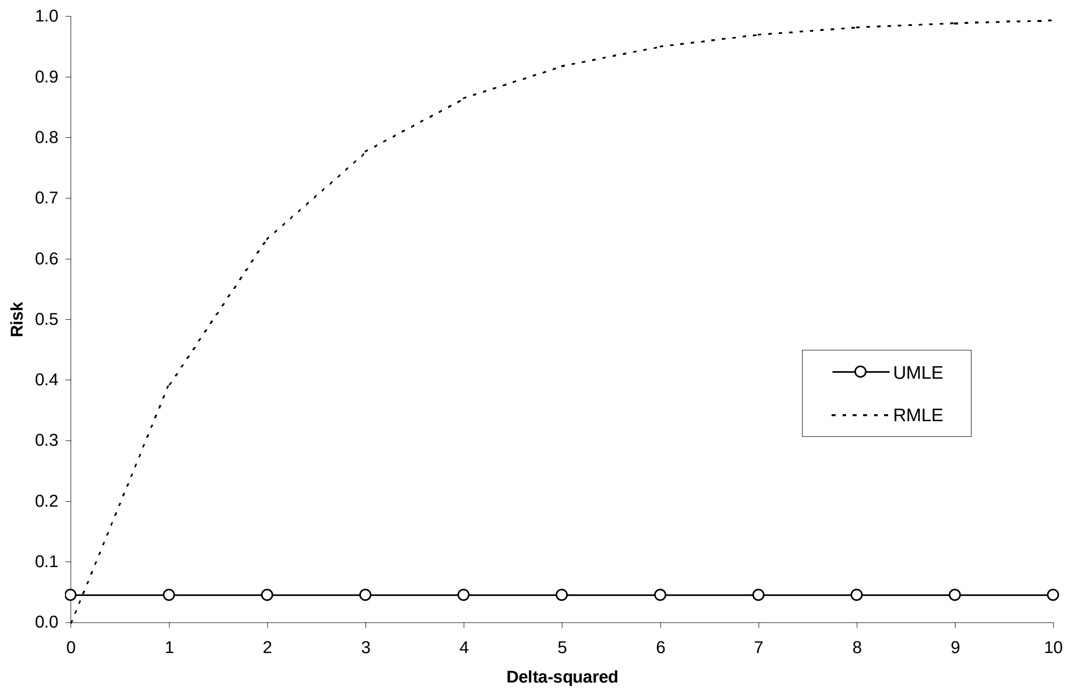


Figure 4: Risks Under Reflected Normal Loss
(n = 10; sigma = 1; K = 1; gamma = 1)

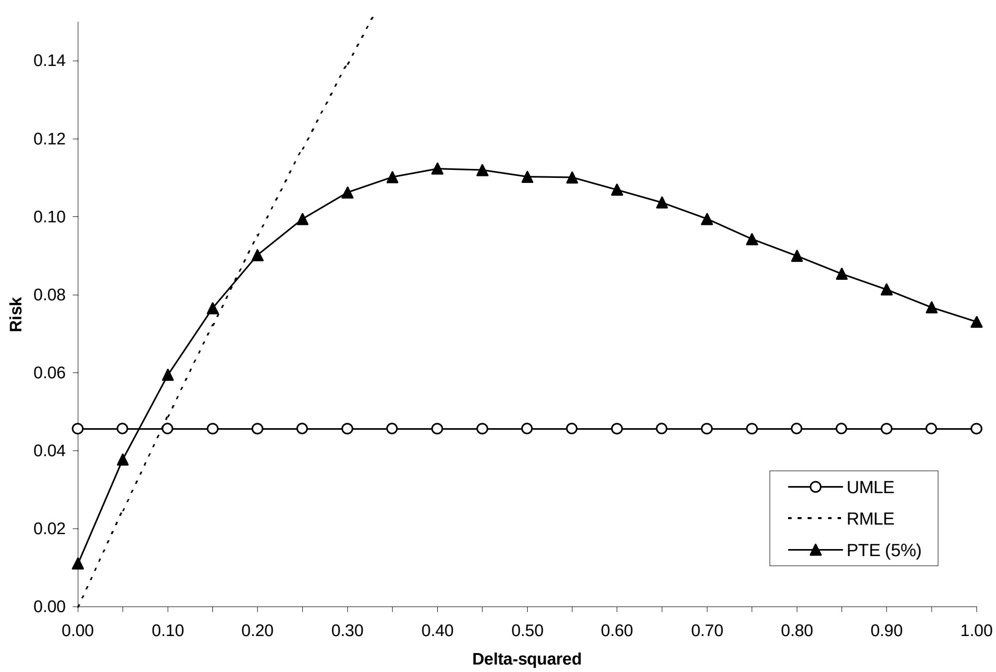


Figure 5: Risks Under Reflected Normal Loss (n = 100; sigma = 1; K = 1; gamma = 1)

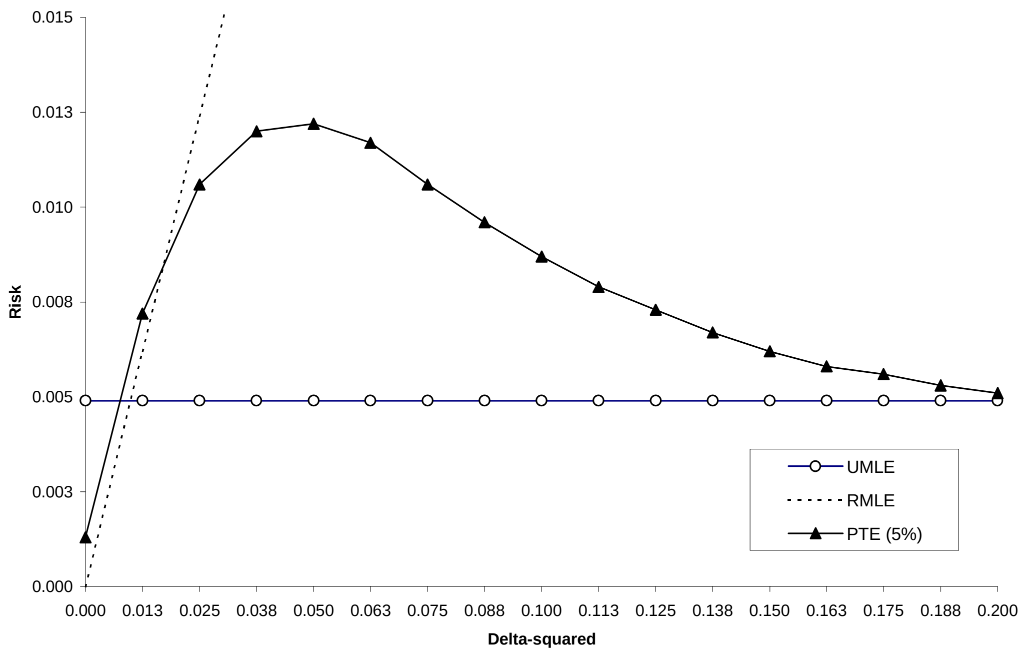
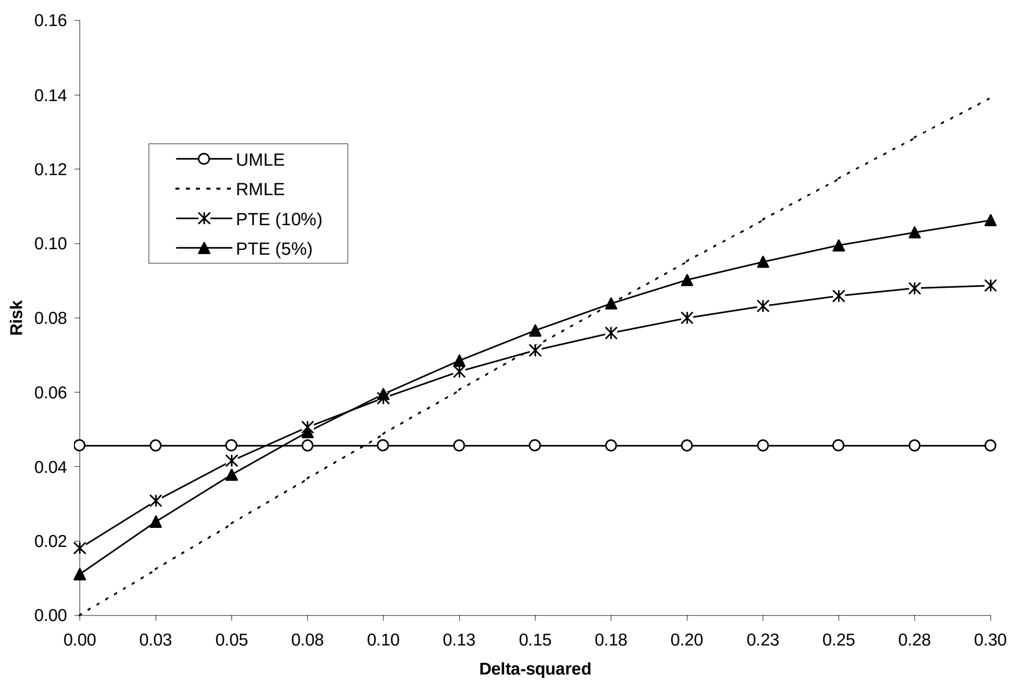


Figure 6: Risks Under Reflected Normal Loss (n = 10; sigma = 1; K = 1; gamma = 1)



References

Bancroft, T. A. , 1944, "On the Biases in Estimation Due to the Use of Preliminary Tests of Significance", *Annals of Mathematical Statistics*15, 190-204.

Bancroft, T. A. and C-P. Han, 1977, "Inference Based on Conditional Specification: A Note and Bibliography", *International Statistical Review*, 45, 117-127.

Clarke, J. A. (1986), *Some Implications of Estimating a Regression Scale Parameter After A Preliminary Test of Restrictions*, M.Ec. Minor Thesis, Department of Econometrics and Operations Research, Monash University.

Cohen, A., 1965, "Estimates of the Linear Combinations of Parameters in the Mean Vector of a Multivariate Distribution", *Annals of Mathematical Statistics*, 36, 299-304.

Geng, W. J. and A. T. K. Wan, 2000, "On the Sampling Performance of an Inequality Pre-Test Estimator of the Regression Error Variance Under LINEX Loss", *Statistical Papers*, forthcoming.

Giles, D. E. A., 1992, "The Exact Distribution of a Simple Pre-Test Estimator", in W. E. Griffiths *et al*. (eds.), *Readings in Econometric Theory and Practice: In Honor of George Judge*, North-Holland, Amsterdam, 57-74.

Giles, D. E. A., 1993, "Pre-Test Estimation in Regression Under Absolute Error Loss", *Economics Letters*, 41, 339-343.

Giles, D. E. A. and V. K. Srivastava, 1993, "The Exact Distribution of a Least Squares Regression Coefficient Estimator After a Preliminary t-Test", *Statistics and Probability Letters*, 16, 59-64.

Giles, J. A., 1991, "Pre-testing for Linear Restrictions in a Regression Model with Spherically Symmetric Disturbances, *Journal of Econometrics*, 50, 377-398.

Giles, J. A. and D. E. A. Giles, 1993, "Pre-Test Estimation in Econometrics: Recent Developments", *Journal of Economic Surveys*, 7, 145-197.

Giles, J. A. and D. E. A. Giles, 1996, "Estimation of the Regression Scale After a Pre-Test for Homoscedasticity Under LINEX Loss", *Journal of Statistical Planning and Inference*, 50, 21-35.

Giles, J. A., D. E. A. Giles and K. Ohtani, 1996, "The Exact Risks of Some Pre-Test and Stein-Type Estimators Under Balanced Loss", *Communications in Statistics*, *A*, 25, 2901-2924.

Huber, P. J. (1977), *Robust Statistical Procedures*, SIAM, Philadelphia.

Judge, G. G. and M. E. Bock (1978), *The Statistical Implications of Pre-Test and Stein-Rule Estimators in Econometrics*, Wiley, New York.

Levy, M. S. and D. Wen (1997a), "Bayesian Estimation Under the BLINEX Loss", unpublished mimeo.

Levy, M. S. and D. Wen (1997b), "BLINEX: A Bounded Asymmetric Loss Function", unpublished mimeo.

Ohtani, K. and D. E. A. Giles, 1996a, "On the Estimation of Regression 'Goodness of Fit' Under Absolute Error Loss", *Journal of Quantitative Economics*, 12, 17-26.

Ohtani, K. and J. A. Giles, 1996b, "The Density Function and MSE Dominance of the Pre-Test Estimator in a Heteroscedastic Linear Regression Model With Omitted Variables", *Statistische Hefte*, 37, 323-342.

Ohtani, K., D. E. A. Giles and J. A. Giles, 1997, "The Exact Risk Performance of a Pre-Test Estimator in a Heteroscedastic Linear Regression Model Under the Balanced Loss Function", *Econometric Reviews*, 16, 119-130.

Raiffa, H. and R. Schlaifer (1961), *Applied Statistical Decision Theory*, M.I.T. Press, Cambridge, MA.

Spiring, F. A. (1993), "The Reflected Normal Loss Function", *Canadian Journal of Statistics*, 21, 321-330.

Taguchi, G (1986), *Introduction to Quality Engineering: Designing Quality Into Products and Processes*, Kraus, White Plains, N.Y..

Wan, A. T. K., 1997, "The Exact Density and Distribution Functions of the Inequality Constrained and Pre-Test Estimators", *Statistical Papers*, 38, 329-341.

Wen, D. and M. S. Levy (1999a), "Admissibility of Bayes Estimates Under BLINEX Loss for the Normal Mean Problem", unpublished mimeo.

Wen, D. and M. S. Levy (1999b), "BLINEX: A Bounded Asymmetric Loss Function With Application to Bayesian Estimation", unpublished mimeo.

Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics*, Wiley, New York.

Footnotes

* An earlier version of this paper was presented at the University of Victoria Econometrics Colloquium, and benefitted accordingly. I am especially grateful to Judith Giles for sharing her wealth of knowledge regarding the properties of preliminary-test strategies.

1. See Giles and Giles (1993) for detailed references.
2. Further details of the proof of this result are available from the author on request.
3. For example, see Chapter 3 of Judge and Bock (1978), and Giles and Giles (1993).
4. For instance, see Zellner (1971, pp. 24-26).
5. See Raiffa and Schlaifer (1961, p. 55) and Zellner (1971, pp. 14-15).
6. For example, see Zellner (1971, p. 20).