韓國 軍事運營分析 學會誌
第12卷 第1號, 1986. 6.

Estimators Which Shrink Towards a Point Other Than the Origin

Choi, An Jae*
Kim, Jai Young*

Abstract

In this study, estimators which shrink towards other than the origin are developed. We discuss the harmonic mean of the observations as a reasonable center and develop estimators which shrink the MLE(δ°) towards the harmonic mean under the loss function L_1. The percentage reduction of average loss(PRAL) of derived estimators compared to the usual estimator(MLE) under the loss function L_1 is estimated for p=4 and 8. Computer simulation shows that the risk performance of the derived estiamtors is remarkably good when the means are greater than three.

I. Introduction

Estimators which shrink towards the origin are expected to be good when the underlying parameters are small. The simulation study in Tsui and Press(1982) showed that the percentage of savings in risk over the MLE is about 25%–30% when the parameter λ_i's are in the range of (0, 4) for $p \geqslant 3$.

If some of the parameters are large very little improvement in risk over the MLE can be expected by shrinking towards the origin. Shrinkage estimators tend to have good risk improvement only near points to which they shrink. Thus if λ is thought to be away from the origin, an estimator shrinking towards the origin will not give good risk reduction. The choice of an appropriate prefixed point or a point determined by data itself should remedy this problem.

* Korea Military Academy

We consider the normalized loss function,

$$L_k = \sum_{i=1}^{p} (\delta_i - \lambda_i)^2/(\lambda_i)^k, \quad k=0, 1, 2, 3, \ldots\ldots$$

A question arises whether there are estimators for the simultaneous estimation of Poisson means.

The Poisson means estimation has been much less extensively studied than the normal means estimation. Only very recently, estimators dominating the MLE have been found which allow utilization of a prefixed point. The ability to assess estimators which shrink towards a pre-determined point is important. Tsui(1978), Tsui(1981), Tsui and Press(1982) and Ghosh, Hwang and Tsui(1983) considered this problem. Tsui(1978) and Tsui and Hudson(1981) constructed estimators which shift the MLE towards a prechosen point as well as points determined by data under the loss function L_0. Under the loss function $L_{k(>1)}$, Ghosh, Hwang and Tsui (1983) proposed better estimators which shrink towards a prefixed point $q=(q_1, \ldots, q_p)$, where the q_i's are nonnegative integrers. Their estimators are defined componentwise as

$$\delta_i^{G_1}(X) = X_i - \frac{(N(X-ke_i)-1) + (X_i^{(k)} - (q_i+k)^{(k)}) +}{(S+k)^i + (X_i^{(k)} - (q_i+k)^{(k)}) +} \tag{1}$$

where (1) $i=1, \ldots, p$, and e_i is vector whose i^{th} coordinate is one,

(2) $N(X)=\#[i: X_i > q_i + k]$, and $(a)_+ = \max[a, 0]$,

(3) $$(S+k)^i = \sum_{j\neq i} ((X_j + k)^{(k)} - (q_j+k)^{(k)}) \tag{2}$$

Better estimators for the loss functions $L_{k(>1)}$, which shrink towards the minimum of the observation are defined componentwise as

$$\delta_i^{G_2}(X)=X_i - \frac{(N(X-ke_i)-1)_+ (X_i^{(k)} - (X_{(1)}+k)^{(k)}) +}{\sum_{j\neq i} ((X_j+k)^{(k)} - (X_{(1)}+k)^{(k)})_+ + (X_i^{(k)} - (X_{(1)}+k)^{(k)})} \tag{3}$$

where (1) $i = 1, \ldots, p$,

(2) $N(X)=\#[i: X_i > X_{(1)} + k]$, and

(3) $X_{(1)} = \min[X_i]$.

The estimators proposed by Ghosh, Hwang and Tsui shift the MLE downward only, since a heavy penalty is imposed for overestimating small parameters λ_i under the loss functions $L_{k(>1)}$.

In a simulation study of Gosh, Hwang and Tsui (1983), q_i is chosen to be the integrer part of λ_i for the estimator δ^{G_1} under L_0. But the risk performances of δ^{G_1} are somewhat disappointing when the λ_i's are large but close to one another. The risk performances of δ^{G_2} are comparatively good among dominating estimators when the λ_i's are close to one another (see Table 2, 3. Ghosh, Hwang and Tsui (1983)).

A log transform is used for the Poisson data and the transformed data treated as approximately normally distributed. Under the squared error loss function L_0, an estimator δ^M which shifts the MLE to the geometric mean of the observations is considered by Ghosh, Hwang and Tsui(1983). We can see that δ^M for Poisson means estimation is similar to δ^L for normal means estimation under the loss function L_0.

We focus on the estimation of Poisson parameters under the normalized squared error loss function L_1. It is natural to find estimators:

(1) which shrink the usual estimator to a reasonable center, and

(2) which have good risk reduction when we have a correct prechosen point.

In Section II, estimators similar to δ^M which shrink towards a reasonable center are derived under the loss function L_1. It is shown that the dervied estimator δ^1 dominates the MLE under the loss function $L_{k(>1)}$ for $p=2$, $\lambda_1+\lambda_2>3$. For $p>3$, asymptotic dominance over the MLE is proved. We derive an estimator $\delta^2(X)$ using the component X_i-h(X) instead of X_i, where h(X) is the harmonic mean of observations, and i=1, ... , p. The results of a computer simulation are described in Section III. PRAL (percentage reduction of average loss) over the MLE is calculated and compared.

II. Estimators which Shrink towards the Harmonic Mean (Constructing the Estimators)

For simultaneous estimation of normal means, Lindley's estimator δ^L shrinks the MLE towards the arithmetic mean of the observations. A similar estimator δ^M for Poisson means estimation which shrinks the MLE towards the geometric mean under the loss function L_0 is dervied by considering a log transformation. A natural question is thus: Can similar estimators be obtained for the loss function L_1? Note that δ^Z shrinks the MLE towards the origin and δ^{G_2} shrinks the MLE towards (X(1), ... $X_{(1)}$). To avoid the difficulty of the arbitrary origin, we derive an estimator which shrinks towards the harmonic mean under the loss function L_1.

Choosing the harmonic mean as a reasonable center is justified below:

(1) Consider a dominating estimator of the form (1-w(X))X, under the loss function L_1 which shrinks the MLE towards the origin. This estimator shrinks the MLE towards the point (0, . . ., 0). So we want to adjust this estimator by adding some function of X. Consider the estimator of the form.

$$\delta(X) = (1\text{-}w(X))\, X + c(X). \tag{4}$$

(2) Then

$$R(\delta, \lambda) - R(\delta^Z, \lambda) = E\left[\sum_{i=1}^{p} \frac{c^2 + 2c(X_i - wX_i - \lambda_i)}{\lambda_i}\right]$$

$$= E\left[c^2 \sum_{i=1}^{p} 1/\lambda_i + 2c \sum_{i=1}^{p} X_i/\lambda_i - 2w \sum_{i=1}^{p} X_i/\lambda_i - 2cp\right]$$

$$= E[\Delta], \text{ (say)}. \tag{5}$$

Minimizing the expression Δ with respect to c yields the optimal c as

$$c = \frac{p - \sum_{i=1}^{p} X_i/\lambda_i + w \sum_{i=1}^{p} X_i/\lambda_i}{\sum_{i=1}^{p} 1/\lambda_i} \tag{6}$$

Although c is a function of the unknown parameter λ, it can be estimated using the observations $X_1, \ldots, X_p$. In particular, if $\lambda_1, \ldots, \lambda_p$ are estimated by their MLE's, $X_1, \ldots, X_p$, we obtain $w(p/\sum_{i=1}^{p} X_i^{-1})$ as an estimator of c.

A better understanding of the structure of estimators under the loss function L_0 and L_1 can be obtained by investigating the structure of the Bayes estimator under L_0 and L_1. Let $\lambda_1, \ldots, \lambda_p$ have the common conjugate prior density which is gamma(α, β) with mean $\alpha\beta$, variance $\alpha\beta^2$, where $\alpha > 0$, $\beta > 0$. Under the loss function L_0, the Bayes estimator is given by

$$\delta_i^B(X) = (1 - 1/(\beta+1))\, X_i + (1/\beta+1))\, \alpha\beta, \quad i=1, \ldots, p. \tag{7}$$

Under the loss function L_1, the Bayes Bayes estimator is given by

$$\delta_i^B(X) = (1\text{-}1/(\beta + 1))X_i + (1/(\beta + 1))\, (\alpha\beta - \beta))\, (\alpha\beta - \beta),\ 1 = 1, \ldots, p. \tag{8}$$

(7) and (8) indicate that the Bayes estimator shrinks the MLE towards $\alpha\beta$ and $(\alpha\beta\text{-}\beta)$ under L_0 and L_1, respectively. If prior information about α, β is unknown we might consider estimating them from data following an empirical Bayes approach. The structure of the Bayes estimator under L_0 and L_1 indicates that the point towards which the MLE is shifted under the loss function L_1 is less than the point under the loss function L_0. This is another reason for choosing the harmonic mean as a reasonable center under the loss function L_1.

Using the harmonic mean as a reasonable center, we derive the estimator,

$$\delta_i^1(X) = X_i - wX_i + w(p/\sum_{i=1}^{p} 1/X_i)), \tag{9}$$

where $w=(p-1)/(Z+p-1)$, $\quad z = \sum_{i=1}^{p} X_i$.

Our next theorem shows that δ^1 dominates the MLE under the loss function L_1, for p=2, $\lambda_1 + \lambda_2 > 3$. To prove this we need the following lemma.

Lemma 1

Let $Z \sim$ Poisson(Λ), then

$$E[\frac{(Z-1)(Z-2)}{(Z+1)} - \frac{(Z-1)\Lambda}{(Z+1)}] < 0 \qquad \text{for } \Lambda > 3. \tag{10}$$

Proof

$$E[\frac{Z^2 - 3Z+2}{Z+1}] = \sum_{z=0}^{\infty} \frac{(z^2 - 3z + 2)e^{-\Lambda}\Lambda^z}{(z+1)!}$$

$$= (1/\Lambda) [\sum_{y=1}^{\infty} (y^2 - 5y+6)e^{-\Lambda}\Lambda^y/y!]$$

$$= (1/\Lambda) [\Lambda^2+\Lambda-5\Lambda+6-6e^{-\Lambda}]$$

$$= \Lambda - 4 + 6\Lambda^{-1} - 6e^{-\Lambda}\Lambda^{-1}. \tag{11}$$

$$\Lambda E[\frac{Z-1}{Z+1}] = \Lambda \sum_{i=1} (z-1)e^{-\Lambda}\Lambda^z/(z+1)!.$$

Let y=z+1, then

$$\Lambda E[\frac{Z-1}{Z+1}] = \sum_{y=1}^{\infty} (y-2)\, e^{-\Lambda})$$

$$= \Lambda-2-(-2e^{-\Lambda})$$

$$= \Lambda-2+2e^{-\Lambda}. \tag{12}$$

Let (11) be $A(\Lambda)$ and (12) be $B(\Lambda)$, respectively. Then, $A(\Lambda)-B(\Lambda) = -2+6\Lambda^{-1}-2e^{-\Lambda}-6e^{-\Lambda}-6e^{-\Lambda}\Lambda^{-1}$ < 0 for $\Lambda > 3$, (actually for $\Lambda > 2.576$).

Theorem 2.

δ^1 (X) given in (9) dominates the MLE under the loss function L_1 for $p = 2, \lambda_1 + \lambda_2 > 3$.

Proof

Let $D = R(\delta^1, \lambda) - R(\delta_0, \lambda)$ under the loss function L_1.

Then

$$D = E[\sum_{i=1}^{2} [w^2(X_i-h)^2 - 2w(X_i-h)(X_i-\lambda_i)]/\lambda_i]$$

where $h(X) = 2/(\sum_{i=1}^{2} 1/X_i)$, $w=1/(Z+1)$.

We have

$$D =, E[\frac{w^2(X_1{}^2-4X_1^2X_2/Z+4X_1^2X_2^2/Z^2)-2w(X_1^2-2X_1^2X_2/Z)}{\lambda_1}$$

$$+\frac{w^2(X_2^2-4X_1X_2^2/Z+4X_1^2X_2^2/Z^2)-2w(X_2^2-2w(X_2^2-2x_1X_2^2/Z)}{\lambda_2}$$

$$+ 2w(X_1-2X_1X_2/Z)+2w(X_2-2X_1X_2/Z)].$$

Note that

$$E[X_1^2X_2 \mid Z] = E[X_1-1)X_2+X_1X_2 \mid Z], \tag{13}$$

and

$$E[X_1^2X_2^2 \mid Z] = E[X_1(X_1-1)X_2(X_2-1) + X_1^2X_2 + X_1X_2^2 - X_1X_2 \mid Z]. \tag{14}$$

Let $\Lambda=\lambda_1+\lambda_2$ and $\theta_i = \lambda_i/\Lambda$, $i=1, 2$. Reparameterizing λ as (Λ, θ), we have

$$D = E[(\Lambda\theta_1)^{-1} [w_2(Z^2\theta_1^2 + Z\theta_1^2 - \frac{4}{Z}(A(\theta, Z)\theta_1 + B(\theta, Z)\theta_1 + B(\theta, Z))$$

$$+w^2(\frac{4}{Z^2}(A(\theta, Z) + B(\theta, Z) + C(\theta, Z)))$$

$$-2w(Z^2\theta_1^2 + Z\theta_2 - Z\theta_1^2 - (2/Z)A(\theta, Z)\theta_1)$$

$$+2w\Lambda(Z\theta_1^2 - 2(Z-1)\theta_1^2\theta_2)]$$

$$+(\Lambda\theta_2)^{-1} [w^2(z^2\theta_2^2 + Z\theta_2 - Z\theta_2^2 \frac{4}{Z}(A(\theta, Z)\theta_2 + B(\theta, Z))$$

$$+w^2(\frac{4}{Z^2}(A(\theta, Z) + B(\Theta, Z) + C(\theta, Z)))$$

$$-2w(Z^2\theta_2^2+Z\theta_2-Z\theta_2^2-(2/Z)A(\theta, Z)\theta_2)$$

$$+2w\Lambda(Z\theta_2^2-2(Z-1)\theta_1\theta_2^2)]] \tag{15}$$

where (1) $A(\theta, Z) = Z(Z-1)(Z-2)\theta_1\theta_2$,
(2) $B(\theta, Z) = Z(Z-1)\theta_1\theta_2$, and
(3) $C(\theta, Z) = Z(Z-1)(Z-2)(Z-3)\theta_1^2\theta_2^2$.

Note that

$$A(\theta, Z) + B(\theta, Z) = Z(Z-1)^2\theta_1\theta_2. \tag{16}$$

Then

$$D=E[\Lambda^{-1}[w^2(Z^2\theta_1+Z-Z\theta_1-4(D(\theta, Z)-4(Z-1)\theta_2)$$

$$+w^2\frac{4}{Z}(D(\theta, Z)(Z-3)\theta_2+(Z-1)^2\theta_2))$$

$$-2w(Z^2\theta_1+Z-Z\theta_1-2D(\theta, Z)-4(Z-1)\theta_1)$$

$$+w^2\frac{4}{2}(D(\theta, Z)(Z-3)\theta_1+(Z-1)^2\theta_1))$$

$$-2w(Z^2\theta_2+Z-Z\theta_2-2D(\theta, Z)) \quad +2w\Lambda(Z\theta_2-2(Z-1)\theta_1\theta_2)]]$$

where $D(\theta, Z) = (Z-1)(Z-2)\theta_1\theta_2$, so that

$$D = E[\Lambda^{-1}[(w^2-2w)(Z^2+Z)+2w\Lambda Z+4w^2(Z-1)((Z-1)/Z)-1)$$

$$+w^2((4/Z)(Z-3)-8)D(\theta, Z)+8w(Z-1)((Z-2)-\Lambda\theta_1\theta_2]]. \tag{17}$$

Note that

(1) $(4/Z)(Z-3)-8<0$,
(2) $E[2(Z-1)(Z-2)-w(Z-1)\Lambda] < 0$ (by Lemma 1), and
(3) $0<\theta_1\theta_2<1/4$.

Now,

$$D\leqslant E[\Lambda^{-1}[w^2(Z^2+Z+(4/Z)(Z-1)^2-4(Z-1))-2w(Z^2+Z-Z\Lambda)]]$$

$$\text{(since } E[-wZ^2+wZ\Lambda]<0)$$

$$\leqslant E[\Lambda^{-1}[w^2(Z^2+Z+\frac{4}{Z}(Z-1)^2-4(Z-1))-2wZ]]$$

$$= E[\Lambda^{-1}(-Z^2-Z-4+4/Z)/(Z+1)^2]$$

$$\leqslant 0 \text{ for } Z\geqslant 1. \tag{18}$$

The next theorem shows that the estimator δ^1 also dominates the MLE under the loss functions $L_{k(>2)}$.

Theorem 3.

δ^1 (X) dominates δ^0 under the loss functions $L_{k(>1)}$ for $p=2$, $\Lambda>3$.

Proof

Let $D_k = R(\delta^1, \lambda) - R(\delta^0, \lambda)$ under the loss L_k. Using a method similar to that of the proof in Theorem 2,

$$\begin{aligned} D_k = E[\Lambda^{-k}[&(w^2-2w)(Z^2-Z) + 2w\Lambda Z)(f_1(\theta)) + (w^2-2w)Z(f_2(\theta)) \\ &+4w^2((Z-1)^2/Z)-(Z-1))(f_3(\theta)) \\ &+w^2(-4(Z-1)(Z-2)(f_4(\theta)) \\ &+w^2(4/Z)(Z-1)(Z-3)(f_5(\theta)) \\ &+4w(Z-1)(Z-2-\Lambda)(f_4(\theta))]] \end{aligned} \tag{19}$$

where (1) $f_1(\theta) = \theta_1^{2-k} + \theta_2^{2-k}$

(2) $f_2(\theta) = \theta_1^{1-k} + \theta_2^{1-k}$

(3) $f_3(\theta) = \theta_1^{1-k}\theta_2 + \theta_1\theta_2^{1-k}$

(4) $f_4(\theta) = \theta_1^{2-k}\theta_2 + \theta_1\theta_2^{2-k}$, and

(5) $f_5(\theta) = \theta_1^{2-k}\theta_2^2 + \theta_1^2\theta_2^{2-k}$

Note that $f_4(\theta) > f_5(\theta)$.

Then

$$\begin{aligned} D_k \leqslant E[\Lambda^{-k} f_1(\theta)[&((w^2-2w)(z^2-2w)(z^2-Z)+2w\Lambda Z)\ (w^2-2w)Z(f_2(\theta)/f_1(\theta)) \\ &+4w^2(Z-1)(-1/Z)(f_3(\theta)/f_1(\theta)) \\ &+4w^2(Z-1)(Z-2)(-3/Z)(f_4(\theta)/f_1(\theta)) \\ &+4w(Z-1)(Z-2-\Lambda)(f_4(\theta)/f_1(\theta))]]. \end{aligned} \tag{20}$$

Note also that $f_2(\theta)/f_1(\theta) > 2$,

$f_3(\theta)/f_1(\theta) > 1$, and

$f_4(\theta)/f_1(\theta) > 2\theta_1\theta_2$.

Now we have the inequality,

$$D_k \leq E[\Lambda^{-k} f_1(\theta)[(w^2-2w)(Z^2-Z)+2w\Lambda Z+2(w^2-2w)Z+4w^2(Z-1)(-1/Z)$$
$$+4w^2(Z-1)(Z-2)(-3/Z)\theta_1\theta_2+8w(Z-2)(Z-2-\Lambda)\theta_1\theta_2]] \qquad (21)$$

≤ 0 (by (17) in Theorem 2).

The next theorem gives the asymptotic result for the estimator $\delta^1(X)$ for $p > 2$

Theorem 4.

Let $D_1(\lambda) = R(\delta^1, \lambda) - R(\delta^Z, \lambda)$ under loss L^1 and let $\lambda_i = nk_i$, $k_i > 0$, $i=1, \ldots, p$. Then the asymptotic improvement of δ^1 over δ^0 is

$$\lim_{n\to\infty} D_1(\lambda) \cong - \frac{(p-1)^2 p^2}{(\sum_{i=1}^{p} k_i)^2 \sum_{i=1}^{p} (1/k_i)} \qquad (22)$$

where k_i, $i=1, \ldots, p$, are fixed.

Proof

$$D_1(\lambda) = E[\sum_{i=1}^{p} [w^2h^2 + 2wh(X_i - wX_i - \lambda_i)]/\lambda_i]$$
$$= E[w^2h^2 \sum_{i=1}^{p} 1/\lambda_i + 2wh \sum_{i=1}^{p} X_i/\lambda_i - 2w^2h \sum_{i=1}^{p} X_i/\lambda_i - 2whp],$$

where $w=(p-1)/(\sum_{i=1}^{p} X_i + p-1)$, and h is the harmonic mean of the observations.

Let $Y_i = X_i/n$, $i=1, \ldots, p$, then

$E[Y_i]=k_i$ and $\mathrm{Var}[Y_i] = k_i/n$.

So, $\underset{\sim}{Y} \to \underset{\sim}{k}$ in probability or distribution.

We know

$$w = (p-1)/(Z+p-1)$$
$$= \frac{(p-1)/p}{\overline{X}+(p-1)/p}$$

$$< \frac{(p-1)/p}{\overline{X}}, \text{ where } \overline{x} = \sum_{i=1}^{p} X_i/p,$$

and $h \leq \overline{X}$.

Note that $wh \leqslant (p-1)/p$.

Therefore

$$nw^2h^2 \sum_{i=1}^{p} 1/\lambda_i \leqslant ((p-1)/p)^2 \sum_{i=1}^{p} 1/k_i.$$

and

$$2nw^2h \sum_{i=1}^{p} X_i/\lambda_i \leqslant 2(p-1)/p)\, nw(\sum_{i=1}^{p} X_i) \sum_{j=1}^{p} (1/\lambda_j)$$

$$\leqslant 2((p-1)/p)\,(p-1) \sum_{j=1}^{p} (1/k_j).$$

By bounded convergence theorem,

$$E[n[w^2h^2 \sum_{i=1}^{p} 1/\lambda_i - 2w^2h \sum_{i=1}^{p} (X_i/\lambda_i)]]$$

$$\frac{(p-1)^2\, p^2}{(\Sigma\, k_i)^2\, \Sigma\, 1/k_i} - 2\,\frac{(p-1)^2\, p^2}{(\Sigma\, k_i)^2\, \Sigma\, 1/k_i}$$

$$\to \frac{(p-1)^2\, p^2}{(\sum_{i=1}^{p} k_i)^2 \sum_{i=1}^{p} 1/k_i} \tag{23}$$

Need to show

$$2E[nwh \sum_{i=1}^{p} (X_i/\lambda_i - 1)] \to 0$$

or $$2E[n(wh-w^*h^*) \sum_{i=1}^{p} (X_i/\lambda_i-1)] \to 0,$$

where w^*, h^* are w, h with λ's replacing X's.

Let

$$F_n(k) = \left(\frac{p-1}{\sum_{i=1}^{p} k_i + (p-1)/n}\right)\left(\frac{p}{\sum_{i=1}^{p} 1/k_i}\right) \tag{24}$$

then $n(wh-w^*h^*) = F_n(Y)-F_n(k)$.

Choose ε so that the ball of radius ε about k does not contain 0. Let B denote the ball.

On B, ∇F is bounded so

$$n(wh-w^*h^*) < K \sum_{i=1}^{p} (Y_i-k_i)^2, \text{ where K is a constant.}$$

By Schwarz's inequality

$$E[\sum_{i=1}^{p} (Y_i-k_i)^2 \sum_{i=1}^{p} (Y_i/k_i)-1)] \to 0.$$

On B^C, $n(wh-w^*h^*) < 2n(p-1)/p$.
But $n(Pr(Y \quad B^C)) \to 0$. (by Markov inequality)

Therefore

$$\lim_{n\to\infty} D_1(\lambda) \approx \frac{(p-1)^2 p^2}{(\sum_{i=1}^{p} k_i)^2 \sum_{i=1}^{p} 1/k_i} \tag{25}$$

This result indicates that δ^1 has better risk performance than that of δ^Z (and necessarily than that of δ^0) when $\lambda_1 = \ldots = \lambda p$, $p\to\infty$ and $n \to \infty$.

In the estimator $\delta^1(X)$, we use $w(X)=(p-1)/(\sum_{i=1}^{p} X_i+p-1)$ which is defined for the estimator δ^Z. Using the deviations X_i-h(X), i=1, . . . , p, rather than x_i in the estimator δ^Z, we derive the next estimator,

$$\delta_i^2(X) = X_i - \frac{P-1}{\Sigma X_i - h(x) + p - 1}(X_i-h(X)), \tag{26}$$

where i=1, ... , p. If the minimum observation X(1) is 0, then the estimator δ^1 1 and δ^2 are the same as the estimator δ^Z.

The estimators δ^1 and δ^2 shrink the MLE upward and downward. Even though the estimators δ^1 and δ^2 do not improve upon the MLE, their risk performance is substantially better under several considered cases.

III. Simulation Results

The simulation result of Gosh, Hwang and Tsui(1983) shows that the estimator δ^{G_2} which shrinks the MLE towards the minimum of the Poisson observations has better risk reduction than that of δ^Z in almost all situations when the parameters are close to one another. But the estimator δ^Z which shrinks the MLE towards the origin has the merit that δ^Z dominates the MLE under the variety of loss functions L_k, k=2, 3, . . .

To provide justification for the proposed estimators δ^1 and δ^2, we performed the following computer simulation. We proved that the estimator δ_1 dominates δ_0 under the various loss functions for p=2, $\Lambda > 3$. We want to compare the risk performances of δ^{G_2} and δ^2 and the risk performances of δ^Z and δ^1. Generated values of λ are described in tables 1 abd 2. The simulation results are described in Table 3 and Tabel 4 for p=4 and p=8, respectively. We calculate the percentage reduction of average loss(PRAL) over the MLE under the loss function L_1.

From Tables 3 and 4, we see that the average PRAL is an increasing function of p, the number of independent Poisson populations. Also, we see that the PRAL decreases as the range of the λ_i's increase. The proposed estimators δ^1 and δ^2 shrink the MLE towards the harmonic mean upward and downward. So, if the λ_i's are small then δ^1 and δ^2 will not give an improved PRAL, since the structure of the loss function indicates that a heavy penalty is imposed for overestimating small parameters λ_i.

Although, the estimator δ^2 does not dominate the MLE the simulation result indicates that δ^2 performs remarkably well when the λ_i's are close to one another as well as when the λ_i's are relatively far apart, while the PRAL of δ^{G_2} decreases as the size of the λ_i's increases. The number of improved cases in loss over the estimator δ^{G_2} of the estimator δ^2 is impressive when the λ_i's are bigger than 3. Comparing estimators δ^Z and δ^1 which dominate the MLE under a variety of loss functions $L_{k>1}$, we see that the risk performance of δ^1 is better than that of δ^Z in almost all cases when the λ_i's are bigger than 3.

Table 1.

Generated Values of λ for p=4

range of the parameters	generated values of λ
(0, 5)	(1.2, 2.6, 3.2, 4.4) (1.8, 1.9, 3.0, 3.7) (0.5, 1.4, 2.4, 3.8)
(3, 8)	(3.9, 4.3, 6.5, 7.3) (4.4, 6.1, 6.3, 6.9) (3.4, 5.2, 7.0, 7.1)
(6, 11)	(6.6, 7.5, 7.7, 10.9) (6.5, 8.6, 9.5, 10.1) (6.2, 6.9, 0, 10.5)
(6, 11)	(9.3, 10.1, 10.6, 6, 12.2) (6.5, 8.6, 9.5, 10.1) (6.2, 6.9, 9.0, 10.5)
(0, 8)	(0.8, 1.4, 2.2, 7.3) (1.3, 2.8, 3.7, 7.0) (1.4, 4.4, 5.7, 7.7)
(3, 11)	(6.9, 7.3, 7.7, 9.9) (5.0, 5.3, 5.8, 8.5) (3.2, 8.2, 9.7, 10.6)
(6, 14)	(6.9, 9.8, 11.2, 11.3) (6.6, 8.7, 10.7, 12.6) (6.2, 6.9, 11.1, 13.8)

Table 2.
Generated Values of λ for p=8

range of the parameters	generated values of λ
(0, 5)	(0.9, 1.4, 3.0, 3.0, 4.3, 4.7, 4.8, 4.9) (0.4, 0.6, 1.0, 1.3, 1.9, 2.9, 3.2, 4.7) (1.2, 1.6, 1.8, 2.8, 9.3, 5.4, 4.4, 4.6)
(3, 8)	(3.1, 4.1, 4.5, 5.1, 5.5, 6.1, 6.5) (3.2, 3.3, 4.2, 5.5, 6.4, 6.8, 7.2, 7.4) (3.2, 4.2, 5.6, 6.5, 6.6, 6.6, 6.9, 7.6)
(6. 11)	(6.6, 7.9, 8.4, 8.7, 8.8, 8.9, 3, 9.6, 10.8) (6.2, 6.5, 5.7, 7, 7.7, 9.6, 9.8, 10.4, 10.6) (6.1, 6.7, 7.3, 7.6, 8.5, 9.3, 9.5, 10.5)
(9, 14)	(9.4, 9.8, 10.1, 10.5, 11.6, 12.3, 12.8, 13.9) (10.2, 10.4, 10.7, 11.7, 11.7, 12.7, 12.9, 13.7) (9.1, 9.2, 10.7, 11.6, 12.0, 13.4, 13.8, 13.9)
(0, 8)	(1.8, 3.6, 3.7, 4.6, 6.3, 6.9, 7.6, 7.7) (1.8, 3.0, 4.3, 5.4, 5.9, 6.0, 7.1, 8.0) (0.6, 1.9, 3.4, 7.6, 6.6, 6.7, 6.2, 7.9)
(3, 11)	(4.3, 4.5, 4.8, 6.5, 6.7, 7.2, 8.6, 10.7) (3.1, 3.4, 4.5, 5.1, 7.0, 8.8, 10.1, 10.6) (3.2, 5.6, 5.9, 7.6, 8.5, 8.7, 9.9, 10.3)
(6, 14)	(7.5, 8.1, 8.5, 8.0, 9.5, 9.6, 9.6, 10.5, 12.8) (6.3, 7.8, 9.2, 10.3, 11.4, 12.0, 12.0, 13.7) (6.4, 7.2, 7.4, 8.3, 9.4, 12.8, 13.0, 13.8)

Table 3
PRAL over the MLE under L_1 for p=4

Range of the Parameters	δ^Z	δ^{G_2}	δ^1	δ^2
(0, 5)	21 19 23 (21)	19 22 20 (20)	17 19 17 (18)	21 27 17 (22)
(3, 8)	10 11 11 (11)	16 19 16 (17)	14 15 14 (14)	25 29 24 (26)
(6, 11)	5 7 7 (6)	14 15 14 (14)	11 11 11 (11)	25 25 24 (25)
(9, 14)	6 3 5 (5)	15 14 13 (14)	9 8 8 (8)	26 25 22 (24)
(0, 8)	18 14 11 (14)	13 14 11 (13)	14 11 8 (11)	13 13 7 (11)
(3, 11)	8 11 7 (9)	15 16 9 (14)	12 14 8 (11)	26 26 10 (21)
(6, 14)	3 7 6 (5)	12 12 10 (11)	9 9 9 (9)	22 20 16 (19)

† The first PRAL in each cell ia based on the first generated value in the Table 5, the second is based on the second and the third is based on the third, respectively.
The parenthetical values are the averages for the three sample PRALs in the cell.

Table 4

PRAL over the MLE under L_1 for p=8

Range of the Parameters	δ^Z	δ^{G_2}	δ^1	δ^2
(0, 5)	18 29 21 (23)	23 28 28 (26)	17 28 22 (22)	20 28 26 (25)
(3, 8)	12 12 11 (12)	23 21 21 (22)	19 19 18 (19)	32 30 31 (31)
(6, 11)	8 9 8 (8)	21 20 20 (20)	14 15 15 (15)	34 32 33 (33)
(9, 14)	7 6 7 (7)	19 19 17 (18)	12 11 11 (11)	33 33 30 (32)
(0, 8)	14 14 13 (14)	19 20 14 (18)	16 16 10 (14)	22 23 9 (18)
(3, 11)	12 12 9 (11)	20 17 16 (18)	17 15 15 (16)	29 33 24 (25)
(6, 14)	7 6 6 (6)	19 15 14 (16)	14 12 12 (13)	32 26 24 (27)

\+ The first PRAL in each cell ia based on the first generated value in the Table 6, the second is based on the second and the third is based on the third, respectively.
The parenthetical values are the averages for the three sample PRALs in the cell.

References

1. Albert, J. (1979). Robust Bayes Estimation. Mimeograph Series. No. 79-9, Dept of Stat, Purdue University.
2. Albert, J. (1981). Simultaneous Estimation of Poisson Means. Journal of Multivariate Analysis, 11,400-417.
3. Albert, J. (1981). Pseudo-Bayes Estimation of Multinominal Proportions. Commun. in Statist. Theory and Methods. A 10(16), 1587-1611.
4. Berger, J. (1977). Admissible minimax estimation of normal mean with arbitrary quadratic loss. Ann. of Statist. 4. 223-226.
5. Berger, J. (1978). Generalized Bayes estimation in multivariate problems. Ann. of Statist. 6. 783-801.
6. Berger, J. (1979). Multivariate Estimation with Nonsymmetric Loss Functions. In optimizing Methods in Statistics, J.S. Rustagi (Ed). Academic Press, New York.
7. Berger, J. (1980). Statistical Decision

Theory. Springer-Verlag, New York.

8. Bock, J. (1975). Minimax estimators of the Mean of Multivariate Normal Distribution. Ann. of Statist. 3. 209-218.
9. Brown, L. (1975). Estimation with Incompletely Specified Loss Functions (the case of several location parameters). J. Amer. Statist. Assoc. 70.417-427.
10. Clevenson, M. and Zidek, J. (1975). Simultaneous Estimation of the Means of Independent Poisson Laws. J. Amer. Statist. Assoc. 70,698-705.
11. Feinberg, S. and Hollan, P. (1973). Simultaneous Estimation of Multinomial Cell Probabilities. J. Amer. Statist. Assoc. 68. 683-691.
12. Ghosh, M. (1983). Estimation of Multiple Poisson Means: Bayes and Empirical Bayes. Statistics and Decisions. 1, 183-195.
13. Ghosh, M., Hwang, J. and Tsui, K. (1983). Construction of Improved Estimators in Multiparameter Estimation for Discrete Exponential Families. Ann. of Statist. 11,351-367.
14. Hodges, J. and Lehmann, E. (1951). Some application of the Cramer-Rao inequality. Proc. 2nd Berk. Symp. Math. Statist. Prabab, 13-20.
15. Hudson, h. (1974). Empirical Bayes Estimation. Technical Report, No. 58, Dept. of Stat., Stanford University.
16. Hudson, H. (1978). A Natural Identity for Exponential Families with Applications in Multiparameter Estimation. Ann. of Statist. 6.473-484.
17. Hudson, H. and Tsui, K. (1981). Simultaneous Poisson Estimators for a priori Hypothesis about Means. J. Amer. Statist. Assoc. 76,182-187.
18. Hwang, J. (1982). Improving upon Standard Estimators in Discrete Exponential Families with Applications to Poisson and Negative Binomial Cases. Ann. of Statist. 10,857-867.
19. James, W. and Stein, C. (1961). Estimation with quadratic Loss. Proc. 4th Berc. Symp. Math. Statist. Probab. 1. 361-379. Univ. of California Press.
20. Johnson, B. (1971). On the Admissible Estimators for Certain Fixed Sample Binomial Problems. Ann. of Statist. 42, 1579-1587.
21. Lehman, E. (1983). Theory of Point Estimation. Wiley, New York.
22. Lindley, D. (1962). Discussion of a paper by Stein, C. J. of Royal Statistica Society(B). 24,285-287.
23. Peng, J. (1975). Simultaneous Estimation of the Parameters of Independent Poisson Distributions. Technical Report, No. 78, Dept. of Stat., Stanford University.
24. Stein, C. (1956). Inadmissibility of the usual estimator of multivariate normal distribution. Proc. 3rd Berc. Symp. Math. Statist. Probab. 1. 197-206. Univ. of Calfornia Press.
25. Stein, C. (1981). Estimation of the Means of a Multivariate Ditribution. Ann. of Statist. 9.1135-1151.
26. Tsui, K. (1978). Simultaneous Estimation of Several Poisson Parameters under Squared Error Loss. Technical Report

No. 37, Dept. of Stat., Univ of California, Riverside.

27. Tsui, K. (1981). Difference Inequalities and Inadmissiblity of the Usual Estimator in Multiparameter Estimation of Discrete Exponential Families. Technical Report No. 653. Dept. of Stat., Univ. of Wisconsin.

28. Tsui, K. amd Press, S. (1982) . Simultaneous Estimation of Several Poisson Parameters under K-normalized Squared Error Loss. Ann. of Statist. 10.93-100.