韓國數學敎育學會誌<數學敎育>
1984. 12, Vol. XXIII, No. 1

Journal of the Korea Society of Mathematical Education
Dec. 1984, Vol. XXIII, No. 1

Simultaneous Estimation of Poisson Means

by Seung-Ho Lee*
Ajou University, Suweon, Korea

1. Introduction

Let $X_1,\dots,X_p$ be p independent random variables, and assume that the probability density of X_i with respect to some measure μ_i is $f_i(x_i|\theta_i)$, $i=1,\dots,p$, where $\theta=(\theta_1,\dots,\theta_p)$ is some unknown parameter. We use the notation

$$X_i \overset{\text{indep.}}{\frown\!\smile} f_i(x_i|\theta_i),\ \ i=1,\dots,p \tag{1.1}$$

to indicate this. The measure μ_i is assumed to be Lebesgue measure when X_i has an absolutely continuous distribution, and is taken to be the counting measure on nonnegative integers when X_i has a discrete distribtion.

It is desired to estimate $\theta(\theta_1,\dots,\theta_p)$ on the basis of $X=(X_1,\dots,X_p)$. The parameter space is clearly $\Theta=\Omega_1\times\dots\times\Omega_p\subset R^p$ (p-dimensional Euclidean space). Let $a=(a_1,\dots,a_p)$ be an available action (i.e., an estimate of θ) and assume that the action space is $\mathscr{A}\subset R^p$. When action is taken and θ is the true parameter value, it is assumed that a loss $L(\theta,a)$ is incurred, where $L(\theta,a)$ is a real valued function defined on $\Theta\times\mathscr{A}$.

A (nonrandomized) estimator $d(X)=(d_1(X),\dots,d_p(X))$ is a function from the sample space to $\mathscr{A}$, which estimates θ by $d(X)$ when X is observed. The risk function $R(\theta,d)$ of an estimator d is defined to be

$$R(\theta,d)=E_\theta L(\theta,d(X)). \tag{1.2}$$

An estimator d^* is defined to be ***as good as*** d if

$$R(\theta,d^*)\le R(\theta,d) \tag{1.3}$$

for all $\theta\in\Theta$. The estimator d^* is said to be ***better than*** d (or ***dominates*** d) if, in addition to (1.3),

$$R(\theta,d^*)<R(\theta,d) \tag{1.4}$$

for some $\theta\in\Theta$. The estimator d is ***admissible*** if there exists no better estimator, and is ***inadmissible*** otherwise.

When $X=(X_1,\dots,X_p)$ has a p-variate normal distribution with mean $\theta=(\theta_1,\dots,\theta_p)$ and the identity covariance matrix, i.e.,

$$X_i \overset{\text{indep.}}{\frown\!\smile} N(\theta_i,1),\ \ i=1,\dots,p. \tag{1.5}$$

Stein (1955) considered the problem of estimating $\theta=(\theta_1,\dots,\theta_p)$ based on X under the loss function,

$$L(\theta,a)=\sum_{i=1}^{p}(\theta_i-a_i)^2, \tag{1.6}$$

* This work was supported in part by the Research Fund of the Ministry of of Education, Korean Government, in 1983.

and proved the surprising result that the usual estimator $d^0(X)=X$, which is the unbiased minimum variance estimate (UMVE) of θ is inadmissible when $p\geq 3$. A better estimator d^* was found in James and Stein (1960), which has the form,

$$d^*(X)=\left\{1-(p-2)/\sum_{i=1}^{p} X_i^2\right\} X. \tag{1.7}$$

Since then, a considerable amount of work by a number of researchers has gone into finding the estimator better than the usual estimator in more general settings. For the normal distributions, the results in the most general setting obtained so far can be found in Berger, et. al. (1977) and Gleser (1979). And Efron and Morris (1973) used Bayesian ideas to evaluate Stein's estimator by introducing the "relative saving loss" as the foundations of their analysis of the normal means problem.

Stein's phenomenon has also been observed for many other distributions. In the simultaneous estimation of several independent Poisson distributions, Clevension and Zidek (1975), Maritz (1969), and Lee (1982), among others, developed estimators improving upon the usual estimators.

In this paper, we consider the problem of simultaneous estimation of means of several independent Poisson distributions. In Section 2, some of the existing estimators are briefly reviewed. And in Section 3, a new estimator is proposed by means of empirical Bayes ideas. In Section 4, this estimator is compared with the existing estimators under normalized squared error loss by Monte Carlo methods.

2. Simultaneous Estimators of Poisson Means

Suppose $X_1,\dots,X_p$ are independent Poisson random variables with means $\lambda_1,\dots,\lambda_p$ respectively, i.e.,

$$X_i \overset{\text{indep.}}{\sim} f(x_i|\lambda_i)=e^{-\lambda_i}\lambda_i^{x_i}/\Gamma(x_i+1),\ \ \lambda_i>0,\ \ i=1,\dots,p. \tag{2.1}$$

We wish to estimate the vector $\lambda=(\lambda_1,\dots,\lambda_p)$ with loss measured by the normalized squared error loss,

$$L(\lambda, d)=\sum_{i=1}^{p} \lambda_i^{-1}(\lambda_i-d_i)^2, \tag{2.2}$$

where d_i is the estimate of λ_i. We allow our estimate of λ_i to depend on the entire vector of observations $X=(X_1,\dots,X_p)$. The usual estimator $d_i^0(X)=X_i$, which is the UMVE of λ_i, is known to be a minimax with risk function $R(\lambda, d^0)=p$ for every value of λ.

Intuitively, one guesses that shrinking the usual estimator d^0 toward some point might yield a better estimator as in the normal means problem of Stein (1955).

Clevenson and Zidek (1975) obtained an estimator which dominates the usual estimator $d^0(X)=X$, when $p>2$. The new estimator d^{CZ} has the form,

$$d_i^{CZ}(X)=\left\{1-\frac{p}{\sum X_i+p}\right\} X_i,\ \ i=1,\dots,p \tag{2.3}$$

which shrinks d^0 toward zero.

On the otherhand, assuming the common independent prior for λ_i, Maritz (1969) obtained a Bayes estimator d^M, which has the form,

$$d_i^M(X)=X_i-\left(\frac{T^2-\bar{X}}{T^2}\right)(X_i-\bar{X}),\ \ i=1,\dots,p \tag{2.4}$$

where $\bar{X}=\sum_{i=1}^{p} X_i/p$ and $T^2=\sum_{i=1}^{p} (X_i-\bar{X})^2/p$.

Note that the estimator $d^M(X)$ of Maritz (1969) shrinks the usual estimator $d^0(X)$ toward $\bar{X}$ in each coordinate.

In the same vein, Lee (1982) proposed another estimator $d^L(X)$ which shrinks $d^0(X)$ toward $(\bar{X},\ldots,\bar{X})$,

$$d_i^L(X)=X_i-\left(\frac{\bar{X}}{\bar{X}+S^2}\right)(X_i-\bar{X}),\ i=1,\ldots,p \tag{2.5}$$

where $S^2=\sum_{i=1}^{p} (X_i-\bar{X})^2/(p-1)$.

The improvements upon the usual estimator d^0 achieved by using these estimators are not uniform on the parameter space Θ. The behaviors of these estimators are to be investigated in details in Section 4.

3. An Empirical Bayes Approach

Let λ_i, $i=1,\ldots,p$, are themselves independently exponentially distributed with pdf's

$$\pi(\lambda_i|\beta)=\frac{1}{\beta}e^{-\lambda_i/\beta}. \tag{3.1}$$

Then, under the distributional assumptions (2.1) and (3.1) it is easy to show that the Bayes estimator of λ_i is given by

$$d_i^*(X)=(1-B)X_i,\ i=1,\ldots,p \tag{3.2}$$

where $B=(1+\beta)^{-1}$.

And the Bayes risk $r(B,d^*)$ of the estimator d^* is

$$r(B,d^*)\equiv E_B R(\lambda,d^*(X))=(1-B)p \tag{3.3}$$

where 'E_B' indicates the expectation under the distribution (3.1) with $\beta=1/B-1$. This should be compared with the Bayes risk of d^0, $r(B,d^0)=p$. The 'savings' obtained by using d^* instead of d^0 are $r(B,d^0)-r(B,d^*)=Bp$. If β is large then B is small, but as β approaches zero, B approaches one and the savings become considerable.

If B (or β) is not known, one can attempt to estimate B from the data. Under (2.1) and (3.1) $Y=\sum_{i=1}^{p} X_i$ is a sufficient statistic for B, which has the Poisson distribution with mean $\sum_{i=1}^{p} \lambda_i$. Thus it is natural to consider an estimator of the form,

$$d=(1-\hat{B}(Y))X \tag{3.4}$$

wherer $\hat{B}(Y)$ is a reasonable estimate of B.

It is obvious that the premium paid for estimating B rather than knowing its exact value can be expressed in terms of the "relative savings loss" (RSL) due to Efron and Morris (1973) as follows:

$$RSL(B,d)\equiv\frac{r(B,d)-r(B,d^*)}{r(B,d^0)-r(B,d^*)}=\frac{r(B,d)-(1-B)p}{Bp}. \tag{3.5}$$

Now a lemma is introduced which is fundamental to pursue the empirical Bayes approach.

Lemma. *Let* $\lambda_i \overset{indep.}{\sim}$ *Exponential* $(1/\beta)$, *and* $X_i|\lambda_i \overset{indep.}{\sim}$ *Poisson* (λ_i) *for* $i=1,\cdots,p$ *and*

$d_i=(1-\hat{B}(Y))X_i,\ \ Y=\sum_{i=1}^{p} X_i$. *Then*

$$RSL(B,d)=E_B\left[\frac{\hat{B}(Z+1)-B}{B}\right]^2 \tag{3.6}$$

where $Z=Y|B$ *has the negative binomial distribution with pdf*

$$f(z|B)=\binom{z+p}{z}B^{p+1}(1-B)^z. \tag{3.7}$$

Proof. See Lee (1983).

By virtue of the lemma, the empirical Bayes problem to estimate B from the data reduces to more familiar forms. On the basis of $Z\sim NB(p+1,\ 1-B)$, one wishes to estimate B with loss function,

$$L(\hat{B},B)=\{(\hat{B}-B)/B\}^2. \tag{3.8}$$

Let $\pi_a(B)$ be the prior distribution on B having density $(1-a)B^{-a}$ on $(0,1]$, $a<1$. Then the Bayes rule is

$$\hat{B}_a(Y)=\frac{p-a}{Y+p+1-a}. \tag{3.9}$$

This leads to an estimator of the form,

$$d^a(X)=\left\{1-\frac{p-a}{\sum_{i=1}^{p} X_i+p+1-a}\right\}X. \tag{3.10}$$

To assess the Bayes risk of this type of estimators explicitly, it is needed to calculate the negative moments of the negative binomial random variables. These are very tedious and we have not attempted to carry out the details.

We now propose an estimator d^{LS} as the one with $a=0$ in (3.10) on the basis of the empirical results,

$$d^{LS}(X)=\left\{1-\frac{p}{\sum_{i=1}^{p}X_i+p+1}\right\}X. \tag{3.11}$$

4. Monte Carlo Comparisons

We were unable to find the risk functions in closed forms, and hence a Monte Carlo study has been carried out to compare the estimators d^{CZ}, d^M, d^L, d^{LS} and d^0 for various parameter configurations.

For a specified parameter configuration, we plot the risk of each estimator estimated by Monte Carlo methods based on 500 independent Poisson random variates generated by VAX-11/750 at Seoul National University. These results are given in Figure A and B.

From the Figures,we observe the followings.

When the parameters are not quite different, d^L and d^M are seemed to perform well with risks less than the half of that of d^0. But, when the differences among the parameters are significant, the risks of d^L and d^M are exceeding the risk of d^0, the minimax value of the estimators. These comes from the fact that d^L and d^M depend on the sample moments which are very sensitive to the extreme values.

On the otherhand, the behaviors of d^{CZ} and d^{LS} are stable in the sense that the improvements are prominent near zero and the risks increase strictly to that of d^0 as the parameters increase independent of their configurations. Thus d^{CZ} and d^{LS} are dominating the usual estimator d^0.

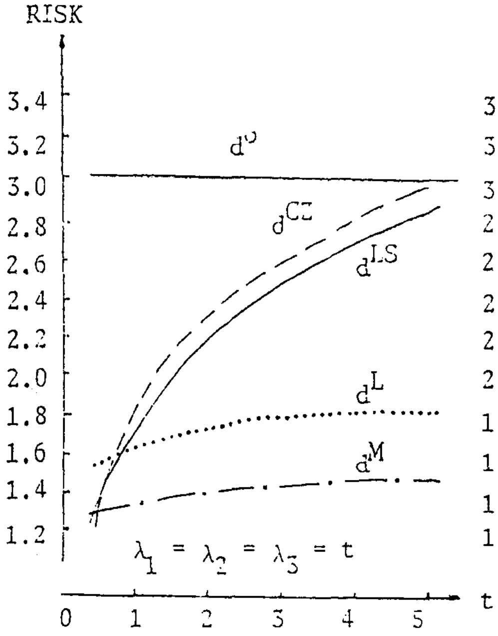

Fig. A-1.

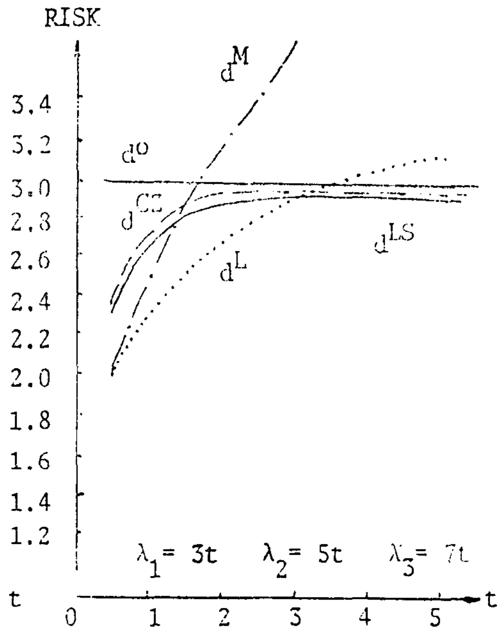

Fig. A-2.

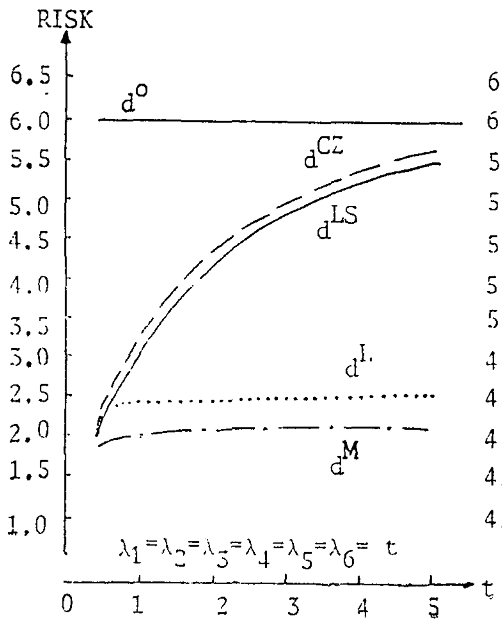

Fig. B-1.

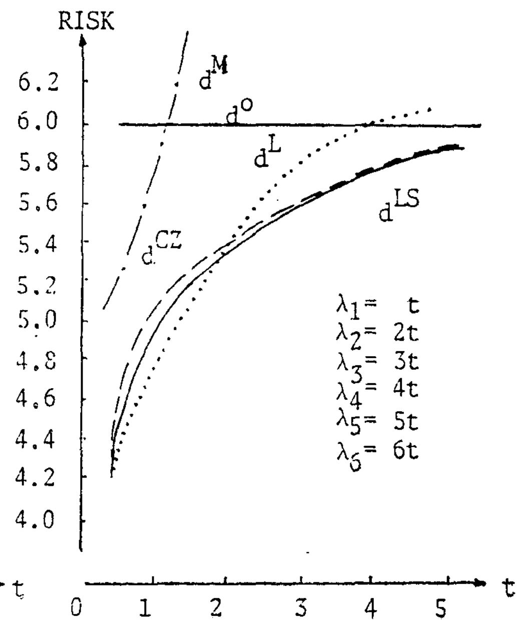

Fig. B-2.

Abstract

A problem of estimating the means of Poisson populations using independent samples is considered. The total loss is the sum of component, normalized squared error losses. An empirical Bayes estimator is derived and compared, by Monte Carlo methods, with existing estimators which are proposed as improving estimators upon the usual one. Monte Carlo results show that the performance of the derived estimator is satisfactory over the whole parameter space.

References

1. Berger, J., Bock, M.E., Brown, L.D., Casella, G. and Glesser, L., Minimax estimation of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix, ***Ann. Statist.***, 5 (1977), 763-771.
2. Clevension, M.L. and Zidek, J.V., Simultaneous estimation of the means of independent Poisson laws, *Jour. Amer. Statist. Assoc.*, **70** (1975), 698-705.
3. Efron, B. and Morris, C., Stein's estimation rule and its competetors - An empirical Bayes approach, *Jour. Amer. Statist. Assoc.*, **68** (1973), 117-130.
4. Glesser, L.J., Minimax estimation of a normal mean vector when the covariance matrix is unknown, *Ann. Statist.*, **4** (1979), 838-846.
5. James, W. and Stein, C., Estimation with quadratic loss, *Pro. Fourth Berkely Symposium Math. Statist. Prob.*, University of California Press, **1** (1960), 361-379.
6. Lee,K.W., On simultaneous estimation of Poisson means, Ms Thesis, Dept. of Comp. Sc.and Statist., Seoul National University, 1982.
7. Lee, S.H., An empirical Bayes approach to simultaneous estimation for Poisson variates, *Ajou University Review*, 5 (1983), 69-71.
8. Maritz, J.S., Empirical Bayes estimators for the Poisson distributions, *Biometrika*, **56** (1969), 349-359.
9. Stein, C., Inadmissibility of the usual estimators for the mean of a multivariate normal distribution, *Proc. Third Berkeley Symposium Math. Statist. Prob.*, University of California Press, **1** (1955), 197-206.