The Korean Communications
in Statistics Vol. 7, No. 3, 2000
pp. 767-772

Improved Estimation of Poisson Means under Balanced Loss Function

Younshik Chung[1)]

Abstract

Zellner(1994) introduced the notion of a balanced loss function in the context of a general linear model to reflect both goodness of fit and precision of estimation. We study the perspective of unifying a variety of results both frequentist and Bayesian from Poisson distributions. We show that frequentist and Bayesian results for balanced loss follow from and also imply related results for quadratic loss functions reflecting only precision of estimation. Several examples are given for Poisson distribution.

Keywords : Balanced loss; Generalized Bayes estimator; Poisson mean; Posterior expected function; Quadratic loss.

1. Introduction

Let $y^t=(y_1,\dots,y_p)$ be an observation vector when the observations $Y_i, 1\le i\le p$, are independently obtained from Poisson distribution with mean θ_i as follows;

$$f(y_i|\theta_i)=\frac{e^{-\theta_i}\theta_i^{y_i}}{y_i\,!},\ y_i=0,1,\dots \tag{1.1}$$

Standard loss structure for estimation the parameter vector $\Theta=(\theta_1,\dots,\theta_p)\in\Omega\subset \boldsymbol{R}^p$ is

$$L(\Theta-\hat{\Theta})=\|\hat{\Theta}-\Theta\| \tag{1.2}$$

where $\hat{\Theta}$ denotes the estimator of Θ.

It was shown that in Roy and Mitra(1957), $\delta_i^0(Y)=Y_i$ is the unique minimum variance unbiased estimator(UMVUE) of θ_i . Hudson(1978), Hwang(1982) and Ghosh, Hwang and Tsui(1983) considered the problem of improving upon $\delta^0(Y)$ with ith component $\delta_i^0(Y)=Y_i$

1) Associate Professor, Department of Statistics, Pusan National University, Pusan, 609-735, Korea.
Email : yschung@hyowon.pusan.ac.kr

under the loss (1.2). For example, the improved estimator $\delta(Y)$ is given componentwisely by $\delta_i(Y)=\delta_i^0(Y)+g_i(Y)$ where

$$g_i(Y)=-\frac{\sum_{k=1}^{Y_i}\frac{1}{k}}{S},\quad S=\sum_{i=1}^{p}\left(\sum_{k=1}^{Y_i}\frac{1}{k}\right)^2 \tag{1.3}$$

This means that the risk function of $\delta(Y)$ is less than that of $\delta^0(Y)$.

Zellner(1994) proposed the balanced loss function as a means of incorporating both goodness of fit and precision of estimation in the evaluation of an estimator. A general form of the balanced loss for the above setup is

$$L_w(\Theta,\hat{\Theta})=w(y-\hat{\Theta})^t(y-\hat{\Theta})+(1-w)(\Theta-\hat{\Theta})^t(\Theta-\hat{\Theta}) \tag{1.4}$$

when $0\le w\le 1$ is the relative weight given to the goodness of fit portion of the loss and $1-w$ is the relative weight given to the precision of estimation portion. Note also that the loss function (1.2) is a special case of $L_w(\Theta,\hat{\Theta})$ when w is chosen to equal to zero. We will therefore refer to the loss function as $L_0(\Theta,\hat{\Theta})$. Under the balanced loss (1.4), Rodrigue and Zellner(1995) considered the estimation of the exponential mean time and Chung, Kim and Song(1998) investigated an admissible linear estimation of Poisson mean and Chung and Kim(1997) obtained the James-Stein type estimator of multivariate normal mean. Recently, Chung and Kim(1998) and Chung, Kim and Dey(1999) obtained the new class of minimax estimation of multivariate normal mean under the balanced loss(1.4).

We are interested in estimating θ using an estimator $\delta(Y)$ and the loss in estimating θ by $\delta(Y)$ will be denoted $L(\theta,\delta(Y))$. If we have two estimators δ_1 and δ_2 that are both functions of Y, then δ_1 is said to be better than δ_2 in terms of risk if $R(\delta_1,\theta)\le R(\delta_2,\theta)$ for all $\theta\in\Theta$, with strict inequality for some $\theta\in\Theta$ where $R(\delta,\theta)=E_\theta^Y[L(\delta(Y),\theta)]$. If δ_1 is better than δ_2, we say δ_1 dominates δ_2 in terms of risk.

We will consider estimator of the form $\delta(y)=\delta^0(y)+g(y)$ where $g(y)\in\Omega$. Assume that $g(y)$ is chosen so that $E[g(y)^t(\delta^0(y)-y)]=0$. We will show that if $\delta(y)=\delta^0(y)+g(y)$ dominates the UMVUE $\delta^0(y)$ in terms of risk for the loss function L_0, then $\delta_w(y)=\delta^0(y)+(1-w)g(y)$ dominates the UMVUE $\delta^0(y)$ in terms of risk for the loss function L_w and vice versa. Also we will show that if $\pi(\Theta)$ is generalized prior distribution which gives rise to a generalized Bayes procedure for each w, it gives rise to a dominating generalized Bayes procedure for each w.

The paper is organized as follows. Section 2 gives a simple identity connecting $R_w(\Theta, \delta)$ and $R_0(\Theta, \delta)$. The general results are obtained in Section 3. Section 4 is devoted to examples for Poisson distributions.

2. Simple Identity

We need the following lemma for proving the main results. For notational conveniences, define, for $0 \leq w \leq 1$, the difference of loss between $\delta^0(\mathbf{y})$ and $\delta^0(\mathbf{y})+(1-w)g(\mathbf{y})$ by

$$\Delta_w(\Theta, \delta^0(\mathbf{y})+(1-w)g(\mathbf{y})) = L_w(\Theta, \delta^0(\mathbf{y})+(1-w)g(\mathbf{y})) - L_w(\Theta, \delta^0(\mathbf{y}). \quad (2.1)$$

In particular, $\Delta_0(\Theta, \delta^0(\mathbf{y})+g(\mathbf{y})) = L_0(\Theta, \delta^0(\mathbf{y})+g(\mathbf{y})) - L_0(\Theta, \delta^0(\mathbf{y}))$ denotes the difference of loss between $\delta^0(\mathbf{y})$ and $\delta^0(\mathbf{y})+g(\mathbf{y})$.

Lemma 2.1. Assume that $g(y) \in \Omega$. Then

$$\begin{aligned}\Delta_w(\Theta, \delta^0(\mathbf{y})+(1-w)g(\mathbf{y})) &= (1-w)^2\Delta_0(\Theta, \delta^0(\mathbf{y})+g(\mathbf{y})) \\ &\quad + 2w(1-w)g(\mathbf{y})^t(\delta^0(\mathbf{y})-\mathbf{y})\end{aligned} \quad (2.2)$$

Proof.

$$\begin{aligned}\Delta_w(\Theta, \delta^0(\mathbf{y})+(1-w)g(\mathbf{y})) &= [w\|\delta^0(\mathbf{y})+(1-w)g(\mathbf{y})-\mathbf{y}\|^2 \\ &\quad + (1-w)\|\delta^0(\mathbf{y})+(1-w)g(\mathbf{y})-\Theta\|^2 \\ &\quad - (w\|\delta^0(\mathbf{y})-\mathbf{y}\|^2+(1-w)\|\delta^0(\mathbf{y})-\Theta\|^2) \\ &= [w(1-w)^2\|g(\mathbf{y})\|^2+(1-w)(1-w)^2\|g(\mathbf{y})\|^2 \\ &\quad + 2w(1-w)g(\mathbf{y})^t(\delta^0(\mathbf{y})-\mathbf{y})+2(1-w)^2g(\mathbf{y})^t(\delta^0(\mathbf{y})-\Theta)] \\ &= [(1-w)^2\|g(\mathbf{y})\|^2+2(1-w)^2g(\mathbf{y})^t(\delta^0(\mathbf{y})-\Theta)] \\ &\quad + 2w(1-w)g(\mathbf{y})^t(\delta^0(\mathbf{y})-\mathbf{y}) \\ &= (1-w)^2\Delta_0(\Theta, \delta^0(\mathbf{y})+g(\mathbf{y}))+2w(1-w)g(\mathbf{y})^t(\delta^0(\mathbf{y})-\mathbf{y})\end{aligned} \quad (2.3)$$

This completes the proof of the lemma.

Remark 2.1. For Poisson distribution, $g(\mathbf{y})^t(\delta^0(\mathbf{y})-\mathbf{y})=0$ since $\delta^0(Y)=Y$. Therefore, under the balanced loss in (1.4), it follows from (2.3) that the risk difference function between $\delta^0(\mathbf{y})$ and $\delta^0(\mathbf{y})+(1-w)g(\mathbf{y})$ is given by

$$\begin{aligned}\Delta_w(\Theta) &= R_w(\Theta, \delta^0(\mathbf{y})+(1-w)g(\mathbf{y}))-R_w(\Theta, \delta^0(\mathbf{y})) \\ &= (1-w)^2E\Delta_0(\Theta, \delta^0(\mathbf{y})+g(\mathbf{y}))\end{aligned} \quad (2.4)$$

where $R_w(\Theta, \hat{\Theta}) = EL_w(\Theta, \hat{\Theta})$ denotes the risk function under the loss in (1.3).

3. Main Results

Theorem 3.1. If $\delta^0(\mathbf{y})+g(\mathbf{y})$ dominates $\delta^0(\mathbf{y})$ in terms of risk under the loss (1.2), then $\delta^0(\mathbf{y})+(1-w)g(\mathbf{y})$ dominates $\delta^0(\mathbf{y})$ in terms of risk under the balanced loss (1.4) for $0\le w\le 1$.

Proof. Recall that the risk difference function between $\delta^0(\mathbf{y})$ and $\delta^0(\mathbf{y})+(1-w)g(\mathbf{y})$ is given by

$$\begin{aligned}\Delta_w(\Theta) &= R_w(\Theta, \delta^0(\mathbf{y})+(1-w)g(\mathbf{y}))-R_w(\Theta, \delta^0(\mathbf{y}))\\ &= (1-w)^2 E\Delta_0(\Theta, \delta^0(\mathbf{y})+g(\mathbf{y}))\end{aligned} \tag{3.1}$$

Since $\delta^0(\mathbf{y})+g(\mathbf{y})$ dominates $\delta^0(\mathbf{y})$ in terms of risk under the loss (1.2), the risk difference between $\delta^0(\mathbf{y})$ and $\delta^0(\mathbf{y})+(1-w)g(\mathbf{y})$ under the loss (1.2) is $E\Delta_0(\Theta, \delta^0(\mathbf{y})+g(\mathbf{y}))$ which is less than zero and so $\Delta_w(\Theta)\le 0$.

Hence, the class of estimators which dominates the UMVUE estimator $\delta^0(\mathbf{y})$ under the usual estimation loss leads to a class which dominates under the balanced loss L_w (and vice versa). An immediate corollary which follows from the convexity of L_0 is

Corollary 3.2. If $\delta^0(\mathbf{y})+g(\mathbf{y})$ dominates $\delta^0(\mathbf{y})$ in terms of risk under L_{w_0}, it dominates $\delta^0(\mathbf{y})$ under L_w in terms of risk for all $0\le w_0\le w$.

Note that it also follows that if $\delta^0(\mathbf{y})+(1-w)g(\mathbf{y})$ dominates $\delta^0(\mathbf{y})$ in terms of risk under L_w, then $\delta^0(\mathbf{y})+(1-w')g(\mathbf{y})$ dominates $\delta^0(\mathbf{y})$ in terms of risk under $L_{w'}$ for all $0\le w\le w'$.

Equally general results for generalized Bayes estimators are possible as the following result shows.

Theorem 3.3. Let $\pi(\Theta)$ be a generalized prior for Θ. If $\delta^0(\mathbf{y})+g(\mathbf{y})$ minimizes the posterior loss under L_0 (and is therefore generalized Bayes), then $\delta^0(\mathbf{y})+(1-w)g(\mathbf{y})$ minimizes the posterior loss under L_w.

Proof. Let $E^{\Theta|\mathbf{y}}$ denote the posterior expected value given $\mathbf{y}$. Suppose $\delta^0(\mathbf{y})+g(\mathbf{y})$ is generalized Bayes under L_0. Then, $E^{\Theta|\mathbf{y}}L_0(\Theta, \delta^0(\mathbf{y})+g(\mathbf{y}))\le E^{\Theta|\mathbf{y}}L_0(\Theta, \delta^0(\mathbf{y})+h(\mathbf{y}))$ for all $h(\mathbf{y})\in\Omega$. It follows from Lemma (2.1) that

$$E^{\Theta|\mathbf{y}}L_0(\Theta, \delta^0(\mathbf{y})+(1-w)g(\mathbf{y}))\le E^{\Theta|\mathbf{y}}L_0(\Theta, \delta^0(\mathbf{y})+(1-w)h(\mathbf{y})) \tag{3.2}$$

and hence since $(1-w)h(\mathbf{y})$ is a general function belongs to Ω for $0\le w\le 1$, that

$\delta^0(y)+(1-w)g(y)$ is generalized Bayes for loss L_w. The Lemma follows trivially for $w=1$ since $\delta^0(y)$ is generalized Bayes.

The following result follows directly from Theorem 3.1 and Theorem 3.3.

Corollary 3.4. Suppose $\delta^0(y)+g_\pi(y)$ is generalized Bayes estimator for the loss L_0 which dominates $\delta^0(y)$ for this loss. Then, $\delta^0(y)+(1-w)g_\pi(y)$ is the generalized Bayes estimator for this same prior for the loss L_w and dominates $\delta^0(y)$ for this loss.

4. Examples

A very general class of estimators improving on $\delta^0(Y)=(Y_1,\ldots,Y_p)$ follows from Hwang(1982), see also Ghosh, Hwang and Tsui(1983).

Theorem 4.1. (Hwang, 1982). Let $y^t=(y_1,\ldots,y_p)$ be an observation vector when the observations Y_i, $1\le i\le p$, are independently obtained from Poisson distribution with mean θ_i as follows

$$f(y_i|\theta_i)=\frac{e^{-\theta_i}\theta_i^{y_i}}{y_i!},\quad y_i=0,1,\ldots \tag{4.1}$$

Standard loss structure for estimationg the parameter vector $\Theta=(\theta_1,\ldots,\theta_p)\in\Omega\subset \boldsymbol{R}^p$ is

$$L(\Theta-\hat{\Theta})=\|\hat{\Theta}-\Theta\| \tag{4.2}$$

where $\hat{\Theta}$ denotes the estimator of Θ . Then the estimator $\delta^0(Y)+g(Y)$ improves on $\delta^0(Y)$ in terms of risk provided

$$\sum_{i=1}^{p}\delta_i^0(Y)[g_i(Y)-g_i(Y-e_i)]+\frac{g_i^2(Y)}{2}\le 0 \tag{4.3}$$

where $g(Y)=(g_1(Y),\ldots,g_p(Y))$ and e_i denotes ith coordinate vector whose ith component is one and the remaining entries are all zero.

Remark 4.1. Famous solutions in Theorem 4.1 is obtained as

$$g_i(Y)=-c(Y)h_i(y_i)/D,\quad i=1,\ldots,p \tag{4.4}$$

where

$$h_i(y_i)=\sum_{i=1}^{y_i}\frac{1}{y_i},\qquad D=\sum_{j=1}^{p}h_j(y_j)h_j(y_j+1),\quad 0\le c(y)\le 2. \tag{4.5}$$

Alternatively, we may choose $\sum_{j=1}^{p}h_j^2(y_j)$ as D which is same as the form in (1.3) with

$c(y)=1$.

Then we may give a fairly general theorem for balanced loss in $L_w(\theta, \hat{\theta})$. This is setting by combining Theorem 4.1. and Remark 4.1 with Theorem 3.1.

Theorem 4.2. Let $g(y)$ satisfy the conditions in (4.4). Then $\delta^0(Y)+(1-w)g(Y)$ dominates $\delta^0(Y)$ in terms of risk for $L_w(\theta, \hat{\theta})$ in (1.4) for $0 \le w \le 1$.

REFERENCES

[1] Chung, Y. and Kim, C.(1997). Simultaneous estimation of the multivariate normal mean under balanced loss function, *Communications in Statistics-Theory and Methods* Vol. 26, 1599-1611.

[2] Chung, Y. and Kim, C.(1998). Minimax estimation of multivariate normal mean under balanced loss function, *Far East Journal of Theoretical Statistics*, Vol. 2, 143-156.

[3] Chung, Y., Kim, C. and Dey, D.K.(1999). A new class of minimax estimator of multivariate normal mean vector under balanced loss function, *Statistics and Decisions*, Vol. 17, 255-266.

[4] Chung, Y., Kim, C. and Song, S.(1998). Linear estimators of Poisson mean under balanced loss functions, *Statistics and Decisions*, Vol. 16, 245-258.

[5] Ghosh, M., Hwang, J.T. and Tsui, K.W.(1983). Construction of improved estimators in multiparameter estimation for discrete exponential families, *Annals of Statistics*, Vol. 11, 351-367.

[6] Hudson, H.M.(1978). A natural identity for exponential families with applications in multi-parameter estimation, *Annals of Statistics*, Vol. 6, 473-484.

[7] Hwang, J.T.(1982). Improving upon standard estimators in discrete exponential families with application to Poisson and negative binomial laws, *Annals of Statistics*, Vol. 10, 857-867.

[8] Rodrigue, J. and Zellner, A.(1995). Weighted balanced loss function for the exponential time to failure, *Communications in Statistics-Theory and Methods*, Vol. 23, 3609-3616.

[9] Roy, J. and Mitra, S.K.(1957). Unbiased minimum variance estimation in a class of discrete distribution, *Sanhkaya A.*, 371-378.

[10] Zellner, A.(1994). *Bayesian and non-Bayesian estimation using balanced loss function. Statistical decision Theory and Related Topics V,* (J.O. Berger and S.S. Gupta Eds.) New York: Springer-verlag, 377-390.