A Minimum Discrimination Information Shrinkage Estimator of a Parameter of a Poisson Distribution

Rep. Fac. Sci. Kagoshima Univ. (Math., Phys. & Chem.),
No. 28, 15-20, 1995.

A Minimum Discrimination Information Shrinkage Estimator of a Parameter of a Poisson Distribution

Koichi INADA* and Koichi TANAKA *

(Received September 29, 1995)

Abstract

An estimation scheme is considered for lowering the mean square error for the estimator of a parameter λ of the Poisson distribution in a region around a prior interval information $\lambda \in [a, b]$. The minimum discrimination information (MDI) approach is employed to conduct the scheme. This approach provides a simple way of specifying the prior interval information about λ, and also allows to consider a shrinkage type estimator. By means of the approach, the MDI estimator and a class of shrinkage to interval estimators are suggested. The estimators compare favorably with the previously proposed estimator in terms of mean square error efficiency.

Key words: Minimum discrimination information approach, Kullback-Leibler discrimination information measure, prior interval information, shrinkage type estimator, mean square error efficiency.

1 Introduction

If $X_1, X_2, \ldots, X_n$ are independent random variables and X_i has a poisson distribusion with parameter λ $(i = 1, 2, \ldots, n)$, then the maximum likelihood estimator of λ is the overall frequency $\bar{X}_n = \sum_{i=1}^{n} X_i/n$. The estimator $\bar{X}_n$ is unbiased with variance λ/n. In fact, it is the minimum variance unbiased estimator of λ. However, for the problem of estimating λ, if we have a preliminary conjecture that $\lambda \in [a, b]$, and if that information is strong compared to other available evidence, it would be better to use an estimator $\hat{\lambda}_n$ which combines the preliminary conjectured prior information in the estimation space around which accuracy seems most crucial. So that $\hat{\lambda}_n$ has smaller mean square error (MSE) than $\bar{X}_n$ for all λ in the preliminary conjectured interval [a, b], even though its

* Department of Mathematics, Faculty of Science, Kagoshima University, Kagoshima 890, Japan.

MSE is greater than that of $\bar{X}_n$ if λ is not in interval. Thus, if the preliminary conjectured interval [a, b] contains λ , we gain in MSE efficiency by using $\hat{\lambda}_n$ instead of $\bar{X}_n$.

One method to utilize this kind of interval information is through the shrinkage type estimator based on MSE criterion. The problem of estimating the normal parameters has been considered by Gokhale et al.(1993), Inada and Kim (1992), and Thompson(1968). When the prior knowledge about the poisson λ is $\lambda \in [a, b]$, Thompson(1968) has given the following shrinkage estimator:

$$\hat{\lambda}_T = \bar{X}_n + \frac{\hat{\sigma}^2}{2(b-a)} \ln \frac{(\bar{X}_n - b)^2 + \hat{\sigma}^2}{(\bar{X}_n - a)^2 + \hat{\sigma}^2}, \quad \text{where } \hat{\sigma}^2 = \frac{\bar{X}_n}{n}. \tag{1.1}$$

Disadvantages of Thompson's shrinkage estimator may be specified on two grounds. First, its MSE sometimes takes larger value than that of $\bar{X}_n$, even though λ takes some value in $[a, b]$. Secondly, it is not robust about the prior information of λ in a sense that it fails to secure the property of uniformly smaller MSE(compared to $\bar{X}_n$) in or around the prior interval so that in the case of a wrong conjecture about the prior interval, even if it is not crucial, the estimator may give us an inefficient estimation (compared to $\bar{X}_n$) of λ.

The object of this paper is to propose and study yet another class of estimators which shrinks toward the preliminary conjectured interval, and eliminates those problems attached to Thompson's estimator. In section 2, we obtain the MDI estimator and a class of shrinkage to interval estimators for this problem. In section 3, we compare the suggested estimators with Thompson's estimator in terms of mean square error efficiency. Section 4 contains some conclusions and further research topics of interest related with this study.

2 Minimum discrimination information estimator

Let $X_1, X_2, \ldots, X_n$ be independent random variables and X_i has a poisson distribution with parameter λ $(i = 1, 2, \ldots, n)$ and suppose we have a preliminary conjectured prior information about λ in the form of an interval $\lambda \in [a, b]$. Our sample estimation scheme for the poisson λ with the interval prior information is treated as a constrained optimization of the Kullback-Leibler discrimination information function.
Estimate λ that

$$\text{minimizes } I(f:g) \quad \text{subject to } \lambda \in [a, b] \tag{2.1}$$

where $I(f:g)$ is the Kullback-Leibler(1951) discrimination information

$$I(f:g) = \sum_{x=0}^{\infty} f(x|\lambda) \ln \frac{f(x|\lambda)}{g(x|\hat{\lambda})}. \tag{2.2}$$

Here $f(x|\lambda)$ and referenced distribution $g(x|\hat{\lambda})$ respectively denote poisson p.d.f.'s with paramcters λ and $\hat{\lambda}$, where $\hat{\lambda}$ is an estimator of λ. Motivation of choosing the disparity measure (2.2) can be found in Shore and Johnson(1980). They have axiomatically shown that principle of minimum discrimination information is uniquely correct method for inductive inference when new information is given in the form of expected value. When we take $\hat{\lambda}$ as the unconstrained maximum likelihood estimator $\bar{X}_n$, the above constrained minimization bears analogy to the external constraints problem(ECP) in the "minimum discrimination information(MDI)" procedure(cf. Gokhale and Kullback (1978)), so that we may call an estimator of λ otained from this procedure as the MDI estimator.

Set $a = \lambda_0 - \delta$ and $b = \lambda_0 + \delta$, and use a class of estimators which shrinks towards λ_0 for $\hat{\lambda}$ then we have following result from the constrained optimization (2.1).

Theorem 1 *Let a preliminary conjectured interval for* λ *be* $\lambda \in [\lambda_0 - \delta, \lambda_0 + \delta]$, $\lambda_0 > 0$ *and* $\delta > 0$, *and let* $\hat{\lambda} = k\bar{X}_n + (1-k)\lambda_0$, $0 < k \leq 1$, *for the reference distribution* $g(x_i|\hat{\lambda})$. *Then the constrained optimization (2.1) yields a class of shrinkage to interval estimators of the poisson* λ *that includes the MDI estimator* (T_{MDI}):

$$T_n(k) = \begin{cases} \lambda_0 - \delta & if \;\; 0 \leq \bar{X}_n < \lambda_0 - \delta/k \\ k\bar{X}_n + (1-k)\lambda_0 & if \;\; \lambda_0 - \delta/k \leq \bar{X}_n \leq \lambda_0 + \delta/k \\ \lambda_0 + \delta & if \;\; \bar{X}_n > \lambda_0 + \delta/k \end{cases} \tag{2.3}$$

where $\bar{X}_n$ *is the maximum likelihood estimator of* λ.

Proof Summing the right hand side of (2.2) with respect to x's, we have

$$I(f : g) = \lambda \ln \frac{\lambda}{\hat{\lambda}} - \{\lambda - \hat{\lambda}\},$$

where $\hat{\lambda} = k\bar{X}_n + (1-k)\lambda_0$. This is a convex function of λ with global minimum at $\lambda = \hat{\lambda}$. Subject to the external constraint $\lambda \in [\lambda_0 - \delta, \lambda_0 + \delta]$, $I(f : g)$ is minimized by $\lambda = \lambda_0 - \delta$ if $0 \leq \bar{X}_n < \lambda_0 - \delta/k$, $\lambda = k\bar{X}_n + (1-k)\lambda_0$ if $\lambda_0 - \delta/k \leq \bar{X}_n \leq \lambda_0 + \delta/k$ and $\lambda = \lambda_0 + \delta$ if $\bar{X}_n > \lambda_0 + \delta/k$. Thus we get the above estimator. We can easily confirm that $T_n(k)$ reduces to the MDI estimator T_{MDI} when $k = 1$. □

Corollary 1 *If true* λ *is in the preliminary conjectured region,* $\lambda \in [\lambda_0 - \delta, \lambda_0 + \delta]$, *then* $MSE(T_{MDI}) < MSE(\bar{X}_n)$.

Proof In the situation $\lambda \in [\lambda_1, \lambda_2]$, where $\lambda_1 = \lambda_0 - \delta$ and $\lambda_2 = \lambda_0 + \delta$, we see that $|\bar{X}_n - \lambda| > |\lambda_1 - \lambda|$ for $\bar{X}_n < \lambda_1$ and $|\bar{X}_n - \lambda| > |\lambda_2 - \lambda|$ for $\bar{X}_n > \lambda_2$. Therefore,

$$\begin{aligned} E[(\bar{X}_n - \lambda)^2] &> E[(\lambda_1 - \lambda)^2\, I_{[0,\lambda_1)}(\bar{X}_n)] + E[(\bar{X}_n - \lambda)^2\, I_{[\lambda_1,\lambda_2]}(\bar{X}_n)] \\ &+ E[(\lambda_2 - \lambda)^2\, I_{(\lambda_2,\infty)}(\bar{X}_n)]. \end{aligned}$$

Thus we have $MSE(T_{MDI}) < MSE(\bar{X}_n)$. □

The mean square error for $T_n(k)$ is

$$
\begin{aligned}
MSE[T_n(k)] &= \sum_{x=0}^{[\lambda_0-\delta/k]} (\lambda_0-\delta-\lambda)^2 Pr(\bar{X}=x) \\
&+ \sum_{x=[\lambda_0-\delta/k]+1}^{[\lambda_0+\delta/k]} (kx+(1-k)\lambda_0-\lambda)^2 Pr(\bar{X}=x) \\
&+ \sum_{x=[\lambda_0+\delta/k]+1}^{\infty} (\lambda_0+\delta-\lambda)^2 Pr(\bar{X}=x)
\end{aligned} \tag{2.4}
$$

$$
\begin{aligned}
&= \frac{1}{n^2}\left\{ \sum_{y=0}^{[\eta_0-\Delta/k]} (\eta_0-\Delta-\eta)^2 Pr(Y=y) \right. \\
&+ \sum_{y=[\eta_0-\Delta/k]+1}^{[\eta_0+\Delta/k]} (ky+(1-k)\eta_0-\eta)^2 Pr(Y=y) \\
&+ \left. \sum_{y=[\eta_0+\Delta/k]+1}^{\infty} (\eta_0+\Delta-\eta)^2 Pr(Y=y) \right\},
\end{aligned} \tag{2.5}
$$

where $[x]$ denotes a greatest integer $\leq x$ and putting $\eta = n\lambda$, $\eta_0 = n\lambda_0$ and $\Delta = n\delta$, the distribution of $Y = \sum_{i=1}^{n} X_i$ is a poisson with parameter η . When $\eta_0 - \Delta/k$ is integer, $[\eta_0 - \Delta/k]$ in (2.5) should be changed to $[\eta_0 - \Delta/k] - 1$ in computations. For Thompson's estimator $\hat{\lambda}_T$ given by (1.1), its mean square error is

$$
MSE(\hat{\lambda}_T) = \frac{1}{n^2}\sum_{y=0}^{\infty}\left\{ y + \frac{y}{4\Delta} ln \frac{(y-\eta_0-\Delta)^2+y}{(y-\eta_0+\Delta)^2+y} - \eta \right\}^2 Pr(Y=y) \tag{2.6}
$$

and the mean square error of $\bar{X}_n$ is

$$
MSE(\bar{X}) = \frac{\eta}{n^2}. \tag{2.7}
$$

3 Numerical comparisons

In order to make a comparison among the MDI estimator T_{MDI}, $T_n(k)$ and Thompson's estimator $\hat{\lambda}_T$ and $\bar{X}_n$, we define a reciprocal measure of the efficiency of the estimator T relative to $\bar{X}_n$ by $MSE(T)/MSE(\bar{X}_n)$. The values of $MSE(T)/MSE(\bar{X}_n)$ are plotted against values of η for $\eta_0 = 6$ and $\Delta = 6$ in Fig.1 and for $\eta_0 = 8$ and $\Delta = 6$ in Fig.2. Several behaviours of $MSE(T)/MSE(\bar{X}_n)$ may be pointed out from these figures.

If we compare the MDI estimator with Thompson's estimator in terms of MSE efficiency, it appears that Thompson's estimator is better when the value of η locates in the middle parts of the interval, but it get worse than the MDI estimator as the value of η lies around the edge of the interval. These figures also show that, unlike Thompson's estimator, the MDI estimator produces uniformly lower MSE than $\bar{X}_n$ over the preliminary conjectured interval and tends to sustain this property in the outside vicinity of this interval. It is seen that $T_n(k)$ with lower value of k achieves lower minimum value of $MSE(T_n(k))/MSE(\bar{X}_n)$ and $T_n(k)$ with higher value of k has wider effective interval for which $MSE(T_n(k))/MSE(\bar{X}_n) < 1$

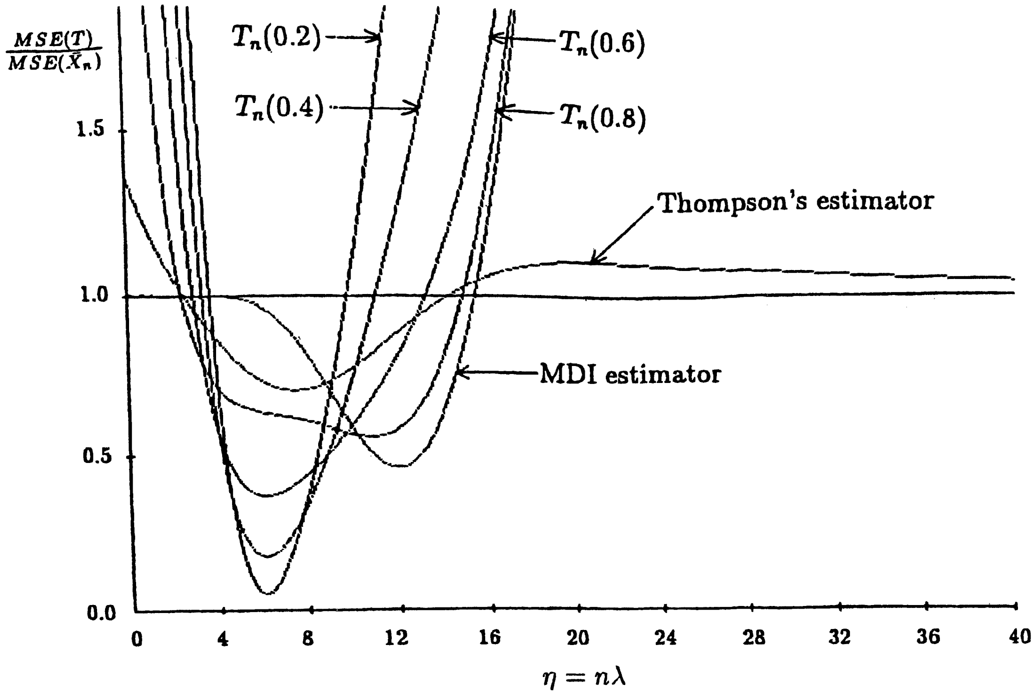

Fig. 1. Efficiencies of the MDI estimator, Thompson's estimator and the four MDI shrinkage estimators for η_0=6 and Δ=6.

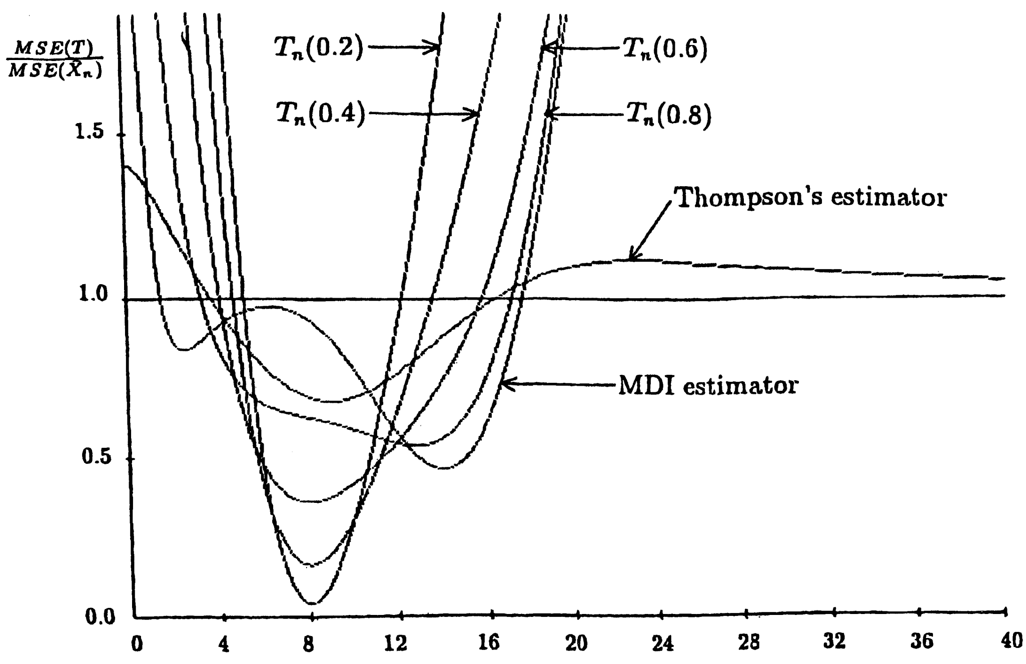

Fig. 2. Efficiencies of the MDI estimator, Thompson's estimator and the four MDI shrinkage estimators for η_0=8 and Δ=6.

4 Concluding remarks

We have proposed the MDI estimator and a class of shrinkage to interval estimators $T_n(k)$ for the poisson λ. This is done by virtue of the MDI approach which is widely applicable to a class of estimation problems where a prior interval information about the parameter of concerned is available. Numerical comparisons showed that the MDI estimator compared favorably with Thompson's estimator: (i) In every case of the preliminary conjectured interval ($\lambda \in [a, b]$), unlike Thomson's estimator $\hat{\lambda}_T$, the MDI estimator dominates $\bar{X}_n$ uniformly in MSE when the true value of λ lies in the interval. (ii) This dominance continues to exist in the outside vicinity of the interval. Thus the use of the MDI estimator assures that, to some extent, we will not be misled by a wrong prior information. It may be also pointed out that the suggested estimation scheme can easily extended to some other distributions so that they may yield estimators with MSE efficiency higher than one relative to usual maximum likelihood estimator. In particular, the MDI approach could possibly be extended to other distributions in exponential family. Such research interests are currently being investigated.

References

[1] D. V. Gokhole, and S. Kullback, The minimum discrimination information approach in analyzing categorical data, Comm. Statist. Theory and Methods, **7** (1978), 987–1005.

[2] D. V. Gokhale, K. Inada and H. J. Kim, A minimum discrimination information estimator of preliminary conjectured normal variance, Ann. Inst. Statist. Math, **45** (1993), 129–136.

[3] S. Kullback, and R. A. Leibler, On information and sufficiency, Ann. Math. Statist., **22** (1951), 79–86.

[4] K. Inada and H. J. Kim, A minimum discrimination information shrinkage to interval estimator of the normal mean, Bull. Inform. Cyber., **25** (1992), 61–69.

[5] J. E. Shore and R. W. Johnson, Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy, IEEE Trans. Inform. Theory, IT**26**(1) (1980), 26–37.

[6] J. R. Thompson, Accuracy borrowing in the estimation of the mean by shrinkage to interval, J. Amer. Statis. Assoc., **63** (1968b), 953–963.