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

著者	INADA Koichi
journal or publication title	鹿児島大学理学部紀要. 数学・物理学・化学
volume	29
page range	7-14
別言語のタイトル	ポアソン分布の母数の2段階最小判別情報縮約推定量について
URL	http://hdl.handle.net/10232/00007036

Rep. Fac. Sci. Kagoshima Univ. (Math., Phys. & Chem.),
No.29, 7-14, 1996.

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

Koichi INADA*

(Received September 10, 1996)

Abstract

A double stage estimation scheme is considered for lowering the mean square error for the estimator of a parameter λ of a 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, a double stage MDI estimator and a double stage MDI shrinkage estimator are suggested, and their mean square errors are derived and compared.

Key words: Double sample scheme, prior interval information, shrinkage type estimator, minimum discrimination information approach, mean square error efficiency.

1 Introduction

Given two succcessively drawn random samples from the poisson distribution with a parameter λ, let us suppose $X_{1j}, j = 1, 2, \ldots, n_1$ be the first sample, and $X_{2j}, j = 1, 2, \ldots, n_2$ be the second sample. Then the usual maximum likelihood estimator of λ is the pooled sample mean $\bar{X}_p = \sum_{i=1}^{2} \sum_{j=1}^{n_i} X_{ij}/N$, where $N = n_1 + n_2$. If we further assume that the prior knowledge about λ is available in the form of an initial estimate. Then it has long been known that the pooled sample mean is inadmissible under the quadratic loss (cf. Bickle [2]). Moreover, in such experiments where samples can be drawn in successon, the estimation by a double stage sampling scheme can reduce the number of observations (cf. Katti [5]). For reasons, we apply the MDI procedure for the

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

double stage sampling scheme to coming at a class of estimates which not only yeilds efficient estimators (comparing to $\bar{X}_p$) but also reduces the number of observations. The object of this paper is to propose and study yet another class of estimators which shrinks towards the preliminary conjectured interval.

2 Double stage MDI shrinkage estimator

Let $X_{1j}(j=1,2,\ldots,n_1)$ and $X_{2j}(j=1,2,\ldots,n_2)$ denote the first and second random samples independently and consecutively drawn from the poisson distribution with a parameter λ and suppose we have a prior interval information that λ lies in a known interval, say $\lambda\in[a,b]$. Our double sample estimation scheme for the parameter λ of the poisson distribution with the interval prior information is treated as a constrained optimization of the Kullback - Leibler discrimination information function.
Estimate λ that

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

where $I(f:g)$ is the Kullback-liebler [7] discrimination information

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

where $f(x|\lambda)$ and a referenced distribution $g(x|\hat{\lambda})$ respectivly denote a poisson p.d.f's with parameter 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 [8]. They have axiomatically shown that a 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 $\bar{X}_p$, the above constrained minimization bears analogy to the external constraints problem(ECP) in the "minimum discrimination information(MDI)" procedure (cf. Gokhale and Kullback [3]), so that we may call an estimator of λ obtained from this procedure as MDI estimator.

Assume that, without loss of generality, the initial estimate for λ is in the form of a preliminary conjectured interval, $[\lambda_0-\delta,\lambda_0+\delta]$, where $\lambda_0>0$ and $\delta>0$. Observing that if we use first sample X_{1j}'s, the Kullback-Liebler discrimination information (2.2) with a class of estimator $\hat{\lambda}_1=k_1\bar{X}_1+(1-k_1)\lambda_0$, $\bar{X}_1=\sum_{j=1}^{n_1}X_{1j}/N_1$ is

$$\text{(2.3)}\qquad I_1(f:g)=\sum_{j=1}^{n_1} f_{1j}(x_{1j}|\lambda)\ln\frac{f_{1j}(x_{1j}|\lambda)}{g_{1j}(x_{1j}|\hat{\lambda}_1)} = \lambda\ln\lambda/\hat{\lambda}_1-(\lambda-\hat{\lambda}_1),$$

This is a convex function of λ with global minimum at $\hat{\lambda}_1$. Introducing the external constraint $\lambda\in[\lambda_0-\delta,\lambda_0+\delta]$ to (2.3), we see that, in the interval $\bar{X}_1\in[\lambda_0-\delta/k_1,\lambda_0+\delta/k_1]$,

$\hat{\lambda}_1$ achives the gloval minimum of $I_1(f : g)$. Thus we may take $\bar{X}_1 \in [\lambda_0 - \delta/k_1, \lambda_0 + \delta/k_1]$ as the preliminary test region for the interval prior knowledge about λ. If we use this region in conjunction with the second samples X_{2j}'s, our double stage estimation scheme consists of following steps. By means of the MDI procedure (cf Inada and Tanaka [4]) based on the first sample, construct the test region $\bar{X}_1 \in [\lambda_0 - \delta/k_1, \lambda_0 + \delta/k_1]$ in the estimation space of λ when $\hat{\lambda}_1$ belongs to the region. Otherwise, we take the second sample and use the MDI procedure based on the both samples, i.e.,

$$\text{minimize } I_p(f : g) = \text{minimize} \sum_{i=1}^{2} \sum_{j=1}^{n_i} f_{ij}(x_{ij}|\lambda) \ln \frac{f_{ij}(x_{ij}|\lambda)}{g_{ij}(x_{ij}|\hat{\lambda}_p)}$$

subject to $\lambda \in [\lambda_0 - \delta, \lambda_0 + \delta]$, where $\hat{\lambda}_p = k_2\bar{X}_p + (1 - k_2)\lambda_0$ *and* $0 < k_2 \leq 1$.

The above double stage estimation scheme leads to a class of double stage MDI shrinkage estimator of λ:

$$(2.4) \quad T_{DMDI}(k_1, k_2) = \begin{cases} k_1\bar{X}_1 + (1 - k_1)\lambda_0 & \text{if } \bar{X}_1 \in R_1 \\ k_2\bar{X}_p + (1 - k_2)\lambda_0 & \text{if } \bar{X}_1 \in \bar{R}_1,\ \bar{X}_p \in R_2 \\ \lambda_0 + \delta & \text{if } \bar{X}_1 \in \bar{R}_1,\ \bar{X}_p > \lambda_0 + \delta/k_2 \\ \lambda_0 - \delta & \text{if } \bar{X}_1 \in \bar{R}_1,\ 0 < \bar{X}_p < \lambda_0 - \delta/k_2 \end{cases}$$

where $\bar{X}_p$ is the unconstrained pooled maximum likelihood estimator of λ, and $R_i = [\lambda_0 - \delta/k_i, \lambda + \delta/k_i],\ 0 < k_i \leq 1\ (i = 1, 2)$.

$T_{DMDI}(k_1, k_2)$ can be viewed as a form of shrinkage estimators of λ which shrinks towards the constrained interval, $\lambda \in [\lambda_0 - \delta, \lambda_0 + \delta]$. It is shown that if $k_1 = k_2 = 1$, the double stage MDI shrinkage estimator in (2.4) reduces to the double stage MDI estimator. And it is also shown that if $n_2 = 0$, the double stage MDI shrinkage estimator reduces to $T_n(k)$ in [4]. The mean square error for $T_{DMDI}(k_1, k_2)$ which is a function of the true value λ and the region R_1 and R_2 is

$$\begin{aligned} & MSE(T_{DMDI}(k_1, k_2)\ |\lambda, R_1, R_2) \\ &= E[(k_1\bar{X}_1 + (1 - k_1)\lambda_0 - \lambda)^2 | \bar{X}_1 \in R_1] Pr(\bar{X}_1 \in R_1) \\ &+ E[(k_2\bar{X}_p + (1 - k_2)\lambda_0 - \lambda)^2 | \bar{X}_p \in \bar{R}_1,\ \bar{X}_p \in R_2] Pr(\bar{X}_1 \in \bar{R}_1,\ \bar{X}_p \in R_2) \\ &+ (\lambda_0 + \delta - \lambda)^2 Pr(\bar{X}_1 \in \bar{R}_1,\ \bar{X}_p > \lambda_0 + \delta/k_2) \\ &+ (\lambda_0 - \delta - \lambda)^2 Pr(\bar{X}_1 \in \bar{R}_1,\ 0 < \bar{X}_p < \lambda_0 - \delta/k_2) \\ (2.5) \quad &= \sum_{x=[n_1(\lambda_0 - \delta/k_1)]+1}^{[n_1(\lambda_0+\delta/k_1)]} \left(\frac{k_1}{n_1} x + (1 - k_1)\lambda_0 - \lambda \right)^2 Pr(X = x) \\ &+ \sum_{x=0}^{[n_1(\lambda_1 - \delta/k_1)]} \sum_{y=[N(\lambda_0 - \delta/k_2)]+1}^{[N(\lambda_0+\delta/k_2)]} \left(\frac{k_2}{N} y + (1 - k_2)\lambda_0 - \lambda \right)^2 Pr(X = x,\ Y = y) \end{aligned}$$

$$+ \sum_{x=[n_1(\lambda+\delta/k_1)]+1}^{\infty} \sum_{y=[N(\lambda_0-\delta/k_2)]+1}^{N(\lambda_0+\delta/k_2)]} \left(\frac{k_2}{N}y + (1-k_2)\lambda_0 - \lambda\right)^2 Pr(X=x,\ Y=y)$$

$$+ (\lambda_0+\delta-\lambda)^2 \sum_{x=0}^{[n_1(\lambda_0-\delta/k_1)]} \sum_{y=[N(\lambda_0+\delta/k_2)]+1}^{\infty} Pr(X=x,\ Y=y)$$

$$+ (\lambda_0+\delta-\lambda)^2 \sum_{x=[n_1(\lambda_0+\delta/k_1)]+1}^{\infty} \sum_{y=[N(\lambda_0+\delta/k_2)]+1}^{\infty} Pr(X=x,\ Y=y)$$

$$+ (\lambda_0-\delta-\lambda)^2 \sum_{x=0}^{[n_1(\lambda_0-\delta/k_1)]} \sum_{y=0}^{[N(\lambda_0-\delta/k_2)]} Pr(X=x,\ Y=y)$$

$$+ (\lambda_0-\delta-\lambda)^2 \sum_{x=[n_1(\lambda_0+\delta/k_1)]+1}^{\infty} \sum_{y=0}^{N[(\lambda_0-\delta/k_2)]} Pr(X=x,\ Y=y)$$

where the distributions of $X = \sum_{j=1}^{n_1} X_{1j}$ and $Y = \sum_{i=1}^{2}\sum_{j=1}^{n_i} X_{ij}$ are poissons with parameter $n_1\lambda$, $N\lambda$ respectively, and $[x]$ denotes the greatest integer $\leq x$. When $n_1(\lambda_0-\delta/k_1)$ and $N(\lambda_0-\delta/k_2)$ are integers, $[n_1(\lambda_0/\delta/k_1)]$ and $[N(\lambda_0-\delta/k_2)]$ in (2.5) should be changed to $n_1(\lambda-\delta/k_1)-1$ and $N(\lambda_0-\delta/k_2)-1$, respectively.

The expected sample size is given by

$$E(n|\lambda, R_1, R_2) = (n_1+n_2) - n_2 Pr(\bar{X}_1 \in R_1) < N, \tag{2.6}$$

where $Pr(\bar{X}_1 \in R_1) = \sum_{s^*=[R_1^*]} Pr(S^*=s^*)$, $[R_1^*]$ denotes a set of integer points in $[n_1(\lambda-\delta/k_1), n_1(\lambda+\delta/k_1)]$ and the distribution of $S^* = \sum_{j=1}^{n_1} X_{1j}$ is a poisson with a parameter $n_1\lambda$. The strict inequality (2.6) shows that we can achieve reduction in sample size by using the double stage MDI estimator. Table 1, Table 2 and Table 3 show the probability of avoiding the second sample, $Pr(\bar{X}_1) \in R_1$, and the expected percentage of the overall sample saved, $Pr(\bar{X}_1 \in R_1)n_2/(n_1+n_2) \times 100$ for a preliminary conjectured interval [0.0, 0.4] with $k_1 = 1$, $k_1 = 0.5$ and $k_1 = 0.1$ respectively. In accordance with the definition of R_1, it may be easily seen that the smaller value of k_1 in $T_{DMDI}(k_1, k_2)$ gives the larger probability of avoiding the second sample, thereby reducing the overall sample size. This fact is shown in Table I, Table 2 and Table 3. However, since our primary concern of using $T_{DMDI}(k_1, k_2)$ is lowering the MSE over the region of preliminary conjectured interval, the choice of k_1 for reduction in sample size should be made in conjunction with the efficiency of the double stage estimator.

The mean square error of $\bar{X}_p$, based on a fixed sample of size $E(n|\lambda, R_1, R_2)$, is $\lambda/E(n|\lambda, R_1, R_2)$. We therefore define the efficiency of $T_{DMDI}(k_1, k_2)$ relative to $\bar{X}_p$ by

$$Eff(T_{DMDI}(k_1, k_2)) = \frac{MSE(T_{DMDI}(k_1, k_2\ |\lambda, R_1, R_2))E(n|\lambda, R_1, R_2)}{\lambda}. \tag{2.7}$$

The behavior of the efficiency $Eff(T_{DMDI}(k_1, k_2))$ are plotted against values of λ under a prior interval information $\lambda \in [0.0, 0.4]$ for $n_1 = 5$, $n_2 = 15$ in Fig.1, for

$n_1 = 10$, $n_2 = 10$ in Fig.2 and for $n_1 = 15$, $n_2 = 5$ in Fig.3. Since $Eff(T_{DMDI}(k_1, k_2))$ for other combinations of the parameters ($k_1, k_2, \lambda, n_1, n_2$) revealed the same pattern as Figures 1 - 3, they are not shown in the figure.

Table 1. Probability of avoiding a second sample and the percentage of the overall sample saved for a preliminary conjectured interval [0.0, 0.4] with $k_1 = 1$.

$k_1 = 1$	$n_1 = 15$, $n_2 = 5$		$n_1 = 10$, $n_2 = 10$		$n_1 = 5$, $n_2 = 15$	
λ	Prob.	Percentage	Prob.	Percentage	Prob.	Percentage
0.1	.999	24.97	.996	49.81	.985	73.92
0.2	.966	24.16	.947	47.36	.919	68.97
0.3	.831	20.77	.815	40.76	.808	60.66
0.4	.606	15.15	.628	31.44	.676	50.75
0.5	.378	9.45	.440	22.02	.543	40.78
0.6	.206	5.16	.285	14.25	.423	31.73
0.7	.101	2.54	.172	8.64	.320	24.06
0.8	.045	1.14	.099	4.98	.238	17.85
0.9	.019	.48	.054	2.74	.173	13.01
1.0	.007	.19	.029	1.46	.124	9.34

Table 2. Probability of avoiding a second sample and the percentage of the overall sample saved for a preliminary conjectured interval [0.0, 0.4] with $k_1 = 0.5$.

$k_1 = 0.5$	$n_1 = 15$, $n_2 = 5$		$n_1 = 10$, $n_2 = 10$		$n_1 = 5$, $n_2 = 15$	
λ	Prob.	Percentage	Prob.	Percentage	Prob.	Percentage
0.1	.999	24.99	.999	49.99	.998	74.86
0.2	.998	24.97	.995	49.77	.981	73.57
0.3	.982	24.57	.966	48.32	.934	70.07
0.4	.916	22.90	.889	44.46	.857	64.28
0.5	.776	19.41	.762	38.10	.757	56.81
0.6	.587	14.68	.606	30.31	.647	48.54
0.7	.397	9.92	.449	22.48	.536	40.24
0.8	.242	6.05	.313	15.66	.433	32.51
0.9	.135	3.38	.206	10.33	.342	25.67
1.0	.069	1.74	.130	6.50	.265	19.87

Table 3. Probability of avoiding a second sample and the percentage of the overall sample saved for a preliminary conjectured interval [0.0, 0.4] with $k_1 = 0.1$.

$k_1 = 0.1$	$n_1 = 15,\ n_2 = 5$		$n_1 = 10,\ n_2 = 10$		$n_1 = 5,\ n_2 = 15$	
λ	Prob.	Percentage	Prob.	Percentage	Prob.	Percentage
0.1	1.000	25.00	1.000	50.00	1.000	75.00
0.2	1.000	25.00	1.000	50.00	1.000	75.00
0.3	1.000	25.00	1.000	50.00	1.000	75.00
0.4	.999	25.00	1.000	50.00	.999	74.99
0.5	.999	24.99	1.000	50.00	.999	74.99
0.6	.999	24.98	.999	49.99	.999	74.99
0.7	.997	24.93	.999	49.99	.999	74.97
0.8	.988	24.71	.999	49.99	.999	74.93
0.9	.964	24.12	.999	49.97	.997	74.81
1.0	.917	22.92	.998	49.92	.994	74.59

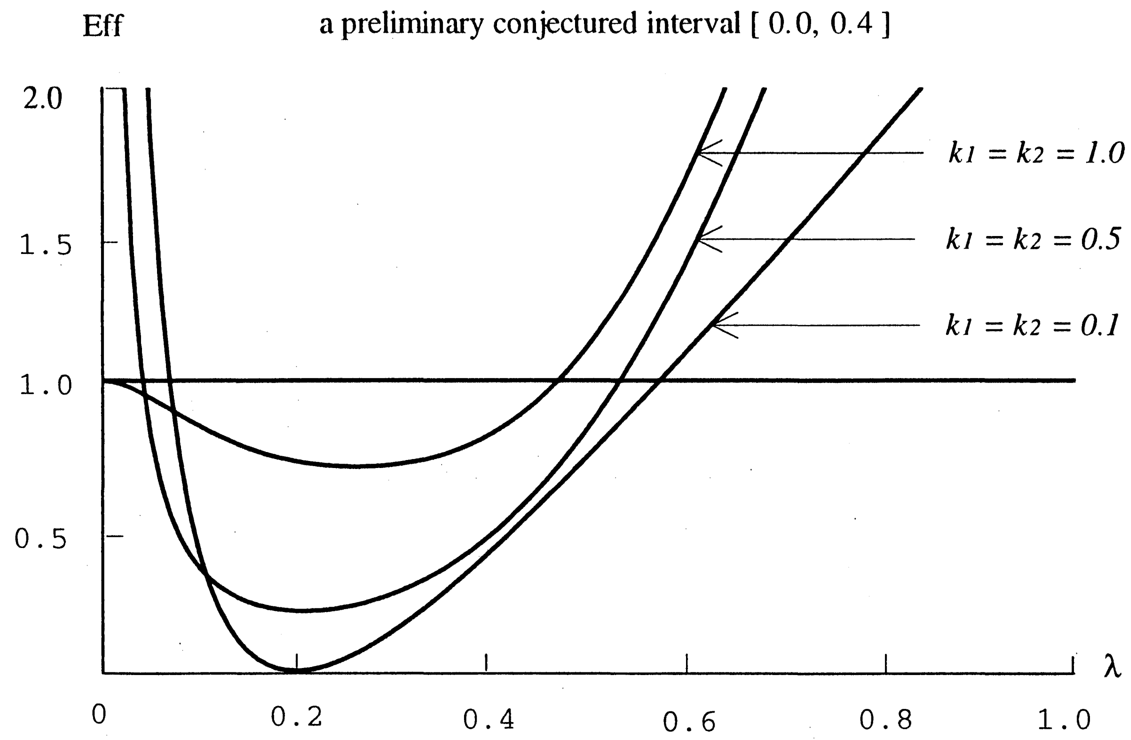

Fig. 1. Efficiencies of the DMDI estimator for $n_1 = 5$ and $n_2 = 15$.

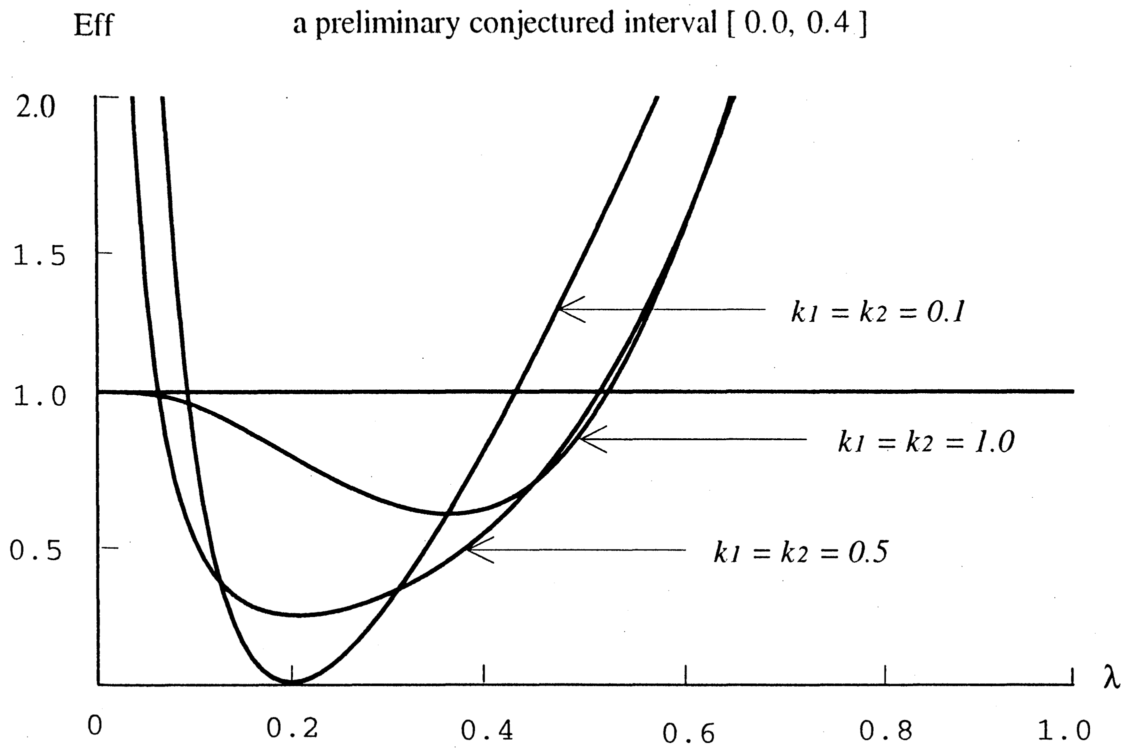

Fig. 2. Efficiencies of the DMDI estimator for $n_1 = 10$ and $n_2 = 10$.

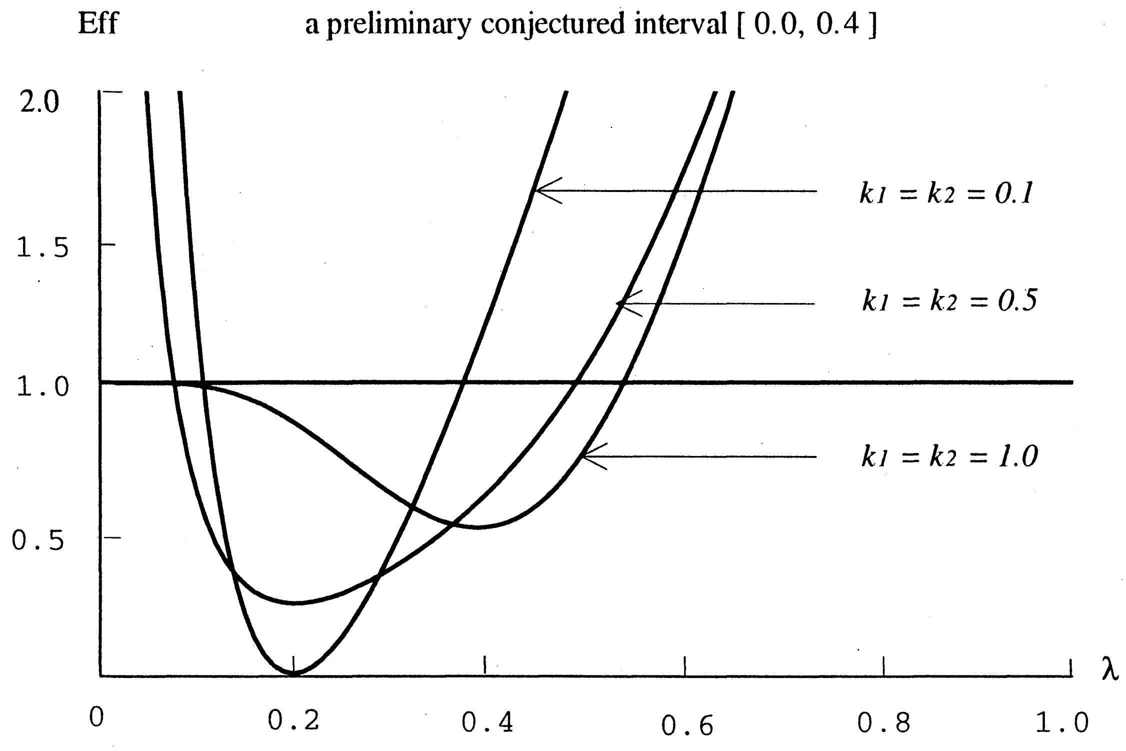

Fig. 3. Efficiencies of the DMDI estimator for $n_1 = 15$ and $n_2 = 5$.

In general, it could be seen from the study of the efficiency that (i) the length of the effective interval of λ increases as k_1 and k_2 increase for $n_1 \geq n_2$ and $k_1 = k_2 = 1$ achieves the longest effective interval of λ, (ii) the double stage MDI estimator achieves its minimum efficiency in the preliminary conjectured interval as k_1 and k_2 take small values.

References

[1] J. C. Arnold and H. A. Al-Bayyati, On double-stage estimation of the mean using prior knowledge, Biometrics, **26** (1970), 787-800.

[2] P. J. Bickle, Minimax estimation of the mean of a normal distribution when the parameter space is restricted, Ann. of Statist., **9** (1981), 1301-1309.

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

[4] K. Inada and K. Tanaka, A minimum discrimination information shrinkage estimator of a parameter of a Poisson distribution, Rep. Fac. Sci. Kagoshima Univ. (Math. Phys. Chem.), **28** (1995), 15-20.

[5] S. K. Katti, Using of some a prior knowledge in the estimation of means from double samples, Biometrics, **18** (1962), 139-147.

[6] H. J. Kim and K. Inada, On double stage minimum discrimination information estimators of the interval constrained normal mean, Comm. Statist. Theory and Methods, **10** (1994), 2733-2753.

[7] S. Kullback and R. A. Liebler, On information and sufficiency, Ann. Math. Statist., **22** (1951), 79-86.

[8] 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.

[9] V. B. Waiker and S. K. Katti, On a two-stage estimate of the mean, Journal of the American Statistical Association, **66** (1971), 75-81.