Journal of Mathematical Research and Exposition
Vol.9, No.4, Nov., 1989

Γ-Minimax Estimation of the Parameters of the Multinomial Distribution*

Chen Lanxiang

(Tong Ji University, Shanghai)

Abstract

Γ-minimax estimators of the parameters of the multinomial distribution are derived by restricting only the means of prior distirbution.

In this paper the problem of estimating the unknown parameter θ of a multinomial distribution is considered. If, there is precise information about the parameter which can be described by a prior π then usually the Bayes principle is applicable, i.e., a Bayes estimator δ with respect to the prior π is considered to be optimal. If, on the other hand, no prior information on parameter θ is available then the minimax principle can be used. In this paper an intermediate approach between the Bayes and the minimax principle is chosen. The use of the Γ-minimax principle is appropriate if vague prior information is available which can be described by a subset of all priors.

In [1] Γ-minimax principle is given. In [5] and [7] minimax estimators of the scale parameter θ lying in a bounded interval for normal distribution and Γ-distribution under squared error loss are derived In [4] and [6] Γ-minimax estimators of the parameters under the restriction of the parameter's moments are derived.

In this paper, Γ-minimax estimators of the parameters of the multinomial distribution are determined under the condition that the means of priors lie within some given bounds. Reader may compare these results with that in [2] and [3].

Let $X=(X_1, X_2, \cdots, X_k)$ have a multinomial distribution with parameter $n, \theta, \theta=(\theta_1, \theta_2, \cdots, \theta_k)$, i.e., $X \sim M(n, \theta_1, \cdots, \theta_k)$. X has a density $f(x,\theta)=f(x_1,\cdots, x_k|\theta_1,\cdots,\theta_k)=\frac{n!}{\prod_{i=1}^{k} x_i!}\theta_i^{x_i}$. Here sample space is $\mathscr{X}=\{(x_1,\cdots,x_k)\in N_0^k, \sum_{i=1}^{k} x_i=n\}$,

* Received Apr.31, 1987.

and Parameter space is $\Theta=\{(\theta_1,\cdots,\theta_k)\in[0,1]^k,\ \sum_{i=1}^{k}\theta_i=1\}$.

Any Borel probabliity measure π on the parameter space Θ is called a prior. Let Π be the set of all priors. The set of all (non-randomized) estimators, i.e., the set of all Borel measurable functions $\delta(X):\mathscr{X}\to\Theta$, is denoted by Δ.

Assume that the loss function of the estimator $\delta(x)=(\delta_1(x),\cdots,\delta_k(x))$ of $\theta=(\theta_1,\cdots,\theta_k)$ is

$$L(\theta,\delta(x))=\sum_{i=1}^{k}\{a_i[\theta_i-\delta_i(x)]^2+t_ix_i\} \tag{1}$$

with $a_i>0$, $t_i>0$, $i=1,2,\cdots,k$.

The risk fuction of δ is given by

$$R(\theta,\delta(x))=\int_{\mathscr{X}}L(\theta,\delta(x))f(x|\theta)\mathrm{d}x. \tag{2}$$

The Bayes risk of an estimator $\delta\in\Delta$ with respect to a prior π under loss (1) is defined by

$$r(\pi,\delta)=\int_{\Theta}R(\theta,\delta(x)\pi(\mathrm{d}\theta). \tag{3}$$

In this paper subset Γ of priors of the form

$$\Gamma=\{\pi\in\Pi.\ 0<\gamma_i<E_\pi\theta_i<\mu_i<1,\ \text{and}\ \sum_{i=1}^{k}\gamma_i<1<\sum_{i=1}^{k}\mu_i,\ i=1,2,\cdots,k\}$$

are considered.

An estimator $\delta^*\in\Delta$ with

$$\sup_{\pi\in\Gamma}r(\pi,\delta^*)=\inf_{\delta\in\Delta}\ \sup_{\pi\in\Gamma}r(\pi,\delta) \tag{4}$$

is called a Γ-minimax estimator. i.e., a Γ-minimax estimator minimizes the maximum Bayes risk with respect to the element of Γ.

A prior $\pi^*\in\Gamma$ with

$$\inf_{\delta\in\Delta}\ r(\pi^*,\delta)=\sup_{\pi\in\Gamma}\ \inf_{\delta\in\Delta}\ r(\pi,\delta) \tag{5}$$

is called least favourable in Γ.

Lemma 1 If there exists a $\pi^*\in\Gamma$ and $\delta^*(X)$ is Bayes estimator of θ with respect to π^* which satisfy

$$r(\pi^*,\delta^*)=\sup_{\pi\in\Gamma}\ r(\pi,\delta^*). \tag{6}$$

Then $\delta^*(X)$ is a Γ-minimax estimator of θ and π^* is least favourable in Γ.

Proof By the definition of Bayes estimator and the given conditions,

$$\sup_{\pi\in\Gamma}\ \inf_{\delta\in\Delta}\ r(\pi,\delta)>\inf_{\delta\in\Delta}\ r(\pi^*,\delta)=r(\pi^*,\delta^*)=\sup_{\pi\in\Gamma}r(\pi,\delta^*)>\inf_{\delta\in\Delta}\sup_{\pi\in\Gamma}r(\pi,\delta),$$

On the other hand,

$$\inf_{\delta\in\Delta}\ \sup_{\pi\in\Gamma}r(\pi,\delta)>\sup_{\pi\in\Gamma}\ \inf_{\delta\in\Delta}\ r(\pi,\delta) \tag{7}$$

from (6) and (7), the assertion follows.

Lemma 2. Let fuctions $h_i(T)(i=1,2,\cdots,k)$ be defined as

$$h_i(T)=\frac{1}{2\sqrt{n}}[n-(T-nt_i)(n+\sqrt{n})^2\cdot a_i^{-1}],$$

and

$$d_i=(n+\sqrt{n})^{-2}\cdot a_i(n-2n\gamma_i)+nt_i,\quad e_i=(n+\sqrt{n})^{-2}\cdot a_i(n-2n\mu_i)+nt_i$$

and

$$c_i(T)=\begin{cases}\sqrt{n}\,\gamma_i, & T>d_i.\\ h_i(T), & e_i<T<d_i.\\ \sqrt{n}\,\mu_i, & T<e_i.\end{cases}$$

Then there exists exactly one T_0, such that $\sum_{i=1}^{k}c_i(T_0)=\sqrt{n}$.

Proof. It is obvious that $T>d_i$ is equivalent to $h_i(T)<\sqrt{n}\,\gamma_i$ and $T>e_i$ is equivalent to $h_i(T)<\sqrt{n}\,\mu_i$. Since $h_i(T)$ is continuously decreasing, $c_i(T)$ and $\sum_{i=1}^{k}c_i(T)$ are also continuously decreasing.

If $T>\max_i d_i$, then $\sum_{i=1}^{k}c_i(T)=\sqrt{n}\sum_{i=1}^{k}\gamma_i<\sqrt{n}$, and if $T<\min_i e_i$, then $\sum_{i=1}^{k}c_i(T)=\sqrt{n}\sum_{i=1}^{k}\mu_i>\sqrt{n}$. Hence there exists exactly one $T_0\in[\min_i e_i,\ \max_i d_i]$ such that $\sum_{i=1}^{k}c_i(T_0)=\sqrt{n}$.

Now we choose prior distribution π of θ with the density

$$f(\theta)=f(\theta_1,\cdots,\theta_k)=\frac{\Gamma(\sum_{i=1}^{k}c_i)}{\prod_{i=1}^{k}\Gamma(c_i)}\prod_{i=1}^{k}\theta_i^{c_i-1},\ (c_i>0,i=1,\cdots,k). \tag{8}$$

We can easily find its density of the posterior distribution of θ with respect to π

$$f(\theta|x)=f(\theta_1,\cdots,\theta_k|x_1,\cdots,x_k)=\frac{\Gamma(\sum_{i=1}^{k}c_i+n)}{\prod_{i=1}^{k}\Gamma(c_i+x_i)}\prod_{i=1}^{k}\theta_i^{x_i+c_i-1} \tag{9}$$

and Bayes estimator $\delta^{\pi}(X)=(\delta_1^{\pi}(X),\cdots,\delta_k^{\pi}(X))$ of θ with respect to π under loss fuction (1) with

$$\delta_i^{\pi}(X)=\frac{X_i+c_i}{n+\sum_{i=1}^{k}c_i},\qquad (i=1,\cdots,k). \tag{10}$$

Let $c_0=\sum_{i=1}^{k}c_i$. Then the risk fuction of $\delta^{\pi}(X)$ is

$$R(\theta,\delta^{\pi}(X))=\sum_{i=1}^{k}a_i\left[\frac{c_0^2-n}{(n+c_0)^2}\right]\theta_i^2+\sum_{i=1}^{k}\left[a_i\frac{n-2c_0\cdot c_i}{(n+c_0)^2}+nt_i\right]\theta_i+\sum_{i=1}^{k}\frac{a_ic_i^2}{(n+c_0)^2}.$$

Theorem 1 Assume $X\sim M(n,\theta_1,\cdots,\theta_k)$, $\pi^*\subset\Pi$, whose density is given in (8) with c_i replaced by $c_i(T_0)$. Let $\delta^*(X)=(\delta_1^*(X),\cdots,\delta_k^*(X))$, where $\delta_i^*(X)=(n+\sqrt{n})^{-1}(X_i+c_i(T_0))$, $(i=1,\cdots,k)$. Then π^* is the least favourable prior distribution in Γ and $\delta^*(X)$ is Γ-minimax estimator of θ, where $c_i(T_0)$ is given in Lemma 2.

Proof. For the distribution with density (8)

$$E_\pi\theta_i=\left(\sum_{i=1}^{k}c_i\right)^{-1}c_i .$$

By Lemma 2, we have $\sum_{i=1}^{k}c_i(T_0)=\sqrt{n}$, and for π^*, $\gamma_i<E_{\pi^*}\theta_i=\frac{c_i(T_0)}{\sqrt{n}}<\mu_i$ $(i=1,\cdots,k)$, which implies $\pi^*\subset\Gamma$. And also it is easy to see that $\delta^*(X)$ is Bayes estimator of θ with respect to π^*.

A short calculation yields

$$r(\pi,\delta^*)=\sum_{i=1}^{k}\left[a_i\frac{n-2\sqrt{n}\,c_i(T_0)}{(n+\sqrt{n})^2}+nt_i\right]E_\pi\theta_i+\sum_{i=1}^{k}\frac{a_ic_i^2(T_0)}{(n+\sqrt{n})^2}$$

$$=\sum_{i=1}^{k}B_iE_\pi\theta_i+T_0+\sum_{i=1}^{k}\frac{a_ic_i^2(T_0)}{(n+\sqrt{n})^2},$$

where $$B_i=a_i\cdot\frac{n-2\sqrt{n}c_i(T_0)}{(n+\sqrt{n})^2}-T_0+nt_i .$$

The following three cases are to be considered.

1) If $c_i(T_0)=\sqrt{n}\mu_i$ then $h_i(T_0)>\sqrt{n}\mu_i$, i.e., $B_i>0$,

2) if $c_i(T_0)=\sqrt{n}\gamma_i$ then $h_i(T_0)<\sqrt{n}\gamma_i$, i.e., $B_i<0$,

3) if $c_i(T_0)=h_i(T_0)$, then $B_i=0$.

In order to find $\sup_{\pi\in\Gamma}r(\pi,\delta^*)$, it is obvious that for $B_i<0$ we must choose $E_\pi\theta_i=\gamma_i$ and for $B_i>0$ we must choose $E_\pi\theta_i=\mu_i$, hence.

$$\sup_{\pi\in\Gamma}r(\pi,\delta^*(x))=\sum_{i=1}^{k}B_i\frac{C_i(T_0)}{\sqrt{n}}+T_0+\sum_{i=1}^{k}\frac{a_ic_i^2(T_0)}{(n+\sqrt{n})^2}=r(\pi^*,\delta^*(x)).$$

By Lemma 1 π^* is least favourable in Γ and $\delta^*(X)$ is Γ-minimax estimator of θ.

Suppose $\Gamma_0\subset\Gamma$ satisfies the following restriction

$$\Gamma_0=\{\pi\in\Pi,\ E_\pi\theta_i=\tau_i,\ 0<\tau_i<1 \text{ and } \sum_{i=1}^{k}\tau_i=1\}.$$

Corollary. Let $X\sim M(n,\theta_1,\cdots,\theta_k)$, and $\pi_0\subset\Pi$ which has the density (8) with $c_i=\sqrt{n}\tau_i$. Let $\delta^0(X)=(\delta_1^0(X),\cdots,\delta_k^0(X))$, where $\delta_i^0(X)=(n+\sqrt{n})^{-1}(X_i+\sqrt{n}\tau_i)$. Then π_0 is least favourable in Γ_0 and $\delta^0(X)$ is Γ_0-minimax estimator of θ

Proof. Now the Bayes risk function of $\delta^0(X)$ with respect to π_0 is

$$r(\pi,\delta^0)=\sum_{i=1}^{k}[a_i\frac{n-2n\tau_i}{(n+\sqrt{n})^2}+nt_i]\tau_i+\sum_{i=1}^{k}\frac{a_i\cdot n\cdot\tau_i^2}{(n+\sqrt{n})^2}$$

It is a constant. From Lemma 1 follows the assertion.

Remark 1 If the restriction on Γ are allowed to loose a little, so that

$$\sum_{i=1}^{k}\gamma_i=1 \quad \text{or} \quad \sum_{i=1}^{k}\mu_i=1,$$

then corresponding c_i should be chosen as $c_i=\sqrt{n}\gamma_i$ or $c_i=\sqrt{n}\mu_i$, the result is simltar to Γ_0.

Remark 2. By assuming $k=2$ in corollary, we obtain the Γ_0-minimax estimator of Binomial distribution, which is the result in (4).

Acknowledgement *The author would like to thank Dr. Eichenaur and Prof. Lehn for their help.*

References

[1] Berger, J.O., Statistical Decision Theory and Bayes Analysis (1985) 2nd. ed., Berlin-Heidelberg-New York Springer.

[2] Trybula, S., Math. statist. and prob. (1962) 3,225-228.

[3] Rutkowska, M., Zastos. mat. (1977) 16, 9—12.

[4] Jackson, D.A., O'donovan. T.M., Zimmer. W.J. and deely, j.j., Biometrika (1970) 57. 439—443.

[5] Casella, G. and Strawderman, W., Ann. statist. (1981) 9,868—876.

[6] Solomon. D.L., Journal of the American Statistical Association (1972) 647—649.

[7] Ghosh, T.K. and Singh .R., Ann. inst. statist. Math. (1970) 22,51—55.

多项分布参数的Γ-极大极小估计

陈 兰 祥

（同济大学应用数学系，上海）

摘 要

本文讨论了多项分布的参数在参数的先验一阶矩的若干限止条件下的Γ-极小极大估计，它有别于毫无先验信息下的通常极小极大估计，又有别于在确切的先验分布下的Bayes估计．本文主要证明了下述定理：

设$X=(X_1,\cdots,X_k)\sim$多项分布$M(n,\theta_1,\cdots,\theta_k)$，相应的先验分布族为$\Gamma=\{\pi\in\Pi,\ 0<\gamma_i<E_\pi\theta_i<\mu_i<1$，且$\sum_{i=1}^{k}\gamma_i<1<\sum_{i=1}^{k}\mu_i\}$．在损失函数$L(\theta,\delta(X))=\sum_{i=1}^{k}[a_i(\theta_i-\delta_i(x))^2+t_ix_i]$下，则存在唯一的一组$c_i(T_0)>0$，$(i=1,\cdots,k)$，设$\pi^*$的密度为

$$f(\theta_1,\cdots,\theta_k)=\frac{\Gamma(\sum_{i=1}^{k}c_i(T_0))}{\prod_{i=1}^{k}\Gamma(c_i(T_0))}\prod_{i=1}^{k}\theta_i^{c_i(T_0)-1}\quad（这里，\ \sum_{i=1}^{k}c_i(T_0)=\sqrt{n}）$$

以及$\delta^*(X)=(\delta_1^*(X),\cdots,\delta_k^*(X))$其中

$$\delta_i^*(X)=(n+\sqrt{n})^{-1}(X_i+c_i(T_0))\quad(i=1,\cdots,k)$$

则$\delta^*(X)$是$\theta=(\theta_1,\cdots,\theta_k)$的Γ-极小极大估计，而$\pi^*$是Γ-最不利分布．