The Korean Communications
in Statistics Vol. 4, No. 1, 1997
pp. 185-191

Bayesian Estimation of Multinomial and Poisson Parameters Under Starshaped Restriction

Myongsik Oh[1]

Abstract

Bayesian estimation of multinomial and Poisson parameters under starshaped restriction is considered. Most Bayesian estimations in order restricted statistical inference require the high-dimensional integration which is very difficult to evaluate. Monte Carlo integration and Gibbs sampling are among alternative methods. The Bayesian estimation considered in this paper requires only evaluation of incomplete beta functions which are extensively tabulated.

1. Introduction

Incorporating prior information regarding a collection of parameters into a statistical inference is very similar to the basic tenet of Bayesian statistics. Thus the idea of using a Bayesian approach to order restricted problem is very appealing. Kraft and Eeden(1964) proposed a Bayesian approach to the problem of estimating a nondecreasing set of binomial parameters. Barlow, Bartholomew, Bremner and Brunk (1972) studied Bayesian estimation for independent samples from members of an exponential-type family and gave a theorem which yields the mode of the posterior distribution as an isotonic regression. Broffit (1984) considered a Bayesian approach to some order restricted problems which are motivated by graduation techniques in actuarial science. Sedransk, Monahan, and Chiu (1985) considered a problem of estimating an ordered set of multinomial probabilities.

Bayesian methods in order restricted statistical inference have been slow in developing. This is mainly due to the difficulty of handling proper prior distribution. In other words, as with many Bayesian methods, calculation of the moments of the posterior distribution requires the evaluation of high-dimensional integrals and is very difficult. Numerical methods are often neccessary to analyse models with constrained parameters. One method suggested by Sedransk et al. (1985) is to use Monte Carlo integration. Gelfand, Smith and Lee (1992) studied the technique using Gibbs sampling.

A vector $\mu=(u_1, \mu_2, \cdots, \mu_k)$ is said to be lower starshaped provided

1) Full-time Lecturer, Department of Statistics, Pusan University of Foreign Studies, Pusan, 603-738, Korea

$$\mu_1 \geq \frac{\mu_1+\mu_2}{2} \geq \cdots \geq \frac{\mu_1+\mu_2+\cdots+\mu_k}{k} \geq 0$$

with an analoguous restriction defining an upper starshaped vector. A vector which is lower starshaped except possibly for the nonnegativity requirement is termed decreasing on the average. Starshaped restriction is frequently encountered in reliability theory. Shaked (1979) derived the maximum likelihood estimate of vector of Poisson and normal means subject to starshaped restriction. Dykstra and Robertson (1982) studied the maximum likelihood estimation and likelihood ratio tests under starshaped restriction using independent samples from multinomial and Poisson population.

In this paper we will derive the Bayesian estimate of multinomial and Poisson parameters under starshaped restriction. Both square and zero-one loss functions are considered. Unlike the most Bayesian estimations in order restricted case, only evaluation of incomplete beta functions is required to find Bayes estimates under square loss function.

2. Multinomial problem

Suppose the outcome of an experiment must be one of k mutually exclusive events having probabilities $p_1, p_2, \cdots, p_k$ ($\sum_{i=1}^k p_i = 1$). We have n independent trials of the experiment. The prior distribution on $\mathbf{p} = (p_1, p_2, \cdots, p_k)$ is usually chosen to be Dirichlet distribution with known parameter $(\alpha_1, \alpha_2, \cdots, \alpha_k)$. Now we are interested in estimating p subject to

$$p_1 \geq \frac{p_1+p_2}{2} \geq \cdots \geq \frac{p_1+p_2+\cdots+p_{k-1}}{k-1} \geq \frac{1}{k} \ .$$

The prior distribution is given by

$$C(\alpha_1, \alpha_2, \cdots, \alpha_k) \prod_{i=1}^k p_i^{\alpha_i - 1}$$

on its support A_k, where

$$A_k = \{ \ \mathbf{p} \ : \ 0 \leq p_i \leq 1, i=1,2,\cdots,k, \sum_{i=1}^k p_i = 1, p_1 \geq \frac{p_1+p_2}{2} \geq \cdots \geq \frac{p_1+p_2+\cdots+p_{k-1}}{k-1} \geq \frac{1}{k} \ \} \ ,$$

and $C(\alpha_1, \alpha_2, \cdots, \alpha_k)^{-1} = \int_{A_k} \prod_{i=1}^k p_i^{\alpha_i - 1} \, d\mathbf{p}$. Then the posterior distribution of $\mathbf{p}$ given the data $(n_1, n_2, \cdots, n_k)$ is

$$C(\alpha_1+n_1, \alpha_2+n_2, \cdots, \alpha_k+n_k) \prod_{i=1}^k p_i^{\alpha_i + n_i - 1}$$

on its support A_k.

First consider the Bayes estimation under square loss function. The Bayes estimate of p_i, denoted by $\hat{p}_i$, is the posterior mean of p_i, i.e.,

$$\begin{aligned}\hat{p}_i &= E(\, p_i \mid (n_1, n_2, \cdots, n_k)\,) \\ &= \int_{A_k} p_i \cdot C(\alpha_1+n_1, \alpha_2+n_2, \cdots, \alpha_k+n_k) \prod_{i=1}^{k} p_i^{\alpha_i+n_i-1}\, d\mathbf{p} \\ &= \frac{C(\alpha_1+n_1, \cdots, \alpha_k+n_k)}{C(\alpha_1+n_1, \cdots, \alpha_{i-1}+n_{i-1}, \alpha_i+n_i+1, \alpha_{i+1}+n_{i+1}, \cdots, \alpha_k+n_k)}\,.\end{aligned}$$

To calculate $\hat{p}_i$ we need to evaluate $C(\cdot)$. The evaluation of $C(\cdot)$ requires high-dimensional integration and seems to be intractable. This can be, however, obtained easily by using the following lemma.

Lemma 1. Let $\mathrm{IB}_u(a, b) = \int_0^u \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}\, dx$. Then

$$C(\alpha_1, \alpha_2, \cdots, \alpha_k)^{-1} = \int_{A_k} \prod_{i=1}^{k} p_i^{\alpha_i-1}\, d\mathbf{p} = \prod_{i=1}^{k-1} \frac{\Gamma(\sum_{j=1}^{i}\alpha_j)\Gamma(\alpha_{i+1})}{\Gamma(\sum_{j=1}^{i+1}\alpha_j)}[1- \mathrm{IB}_{i/(i+1)}(\sum_{j=1}^{i}\alpha_j, \alpha_{i+1})]\,.$$

Proof. Consider a one-to-one transformation $\theta_i = \sum_{j=1}^{i} p_j \,/\, \sum_{j=1}^{i+1} p_j$ for $i=1, \cdots, k-1$. Then we have $p_1 = \prod_{j=1}^{k-1}\theta_j$, $p_i = (1-\theta_{i-1})\prod_{j=i}^{k-1}\theta_j$ for $i=2, \cdots k-1$, and $p_k = 1-\theta_{k-1}$. Then the restriction on θ_i's including basic restriction becomes $i/(i+1) \le \theta_i \le 1$ for $i=1, \cdots, k-1$. It is not difficult to show that $\left|\frac{\partial(p_1, \cdots, p_{k-1})}{\partial(\theta_1, \cdots, \theta_{k-1})}\right| = \prod_{i=2}^{k-1}\theta_i^{\,i-1}$. Note that θ_i's are turned out to be statistically independent. Now we have

$$\begin{aligned}\int_{A_k} \prod_{i=1}^{k} p_i^{\alpha_i-1}\, d\mathbf{p} &= \prod_{i=1}^{k-1} \int_{\frac{i}{i+1}}^{1} \theta_i^{\sum_{j=1}^{i}\alpha_j-1}(1-\theta_i)^{\alpha_{i+1}-1}\, d\theta_i \\ &= \prod_{i=1}^{k-1} \frac{\Gamma(\sum_{j=1}^{i}\alpha_j)\Gamma(\alpha_{i+1})}{\Gamma(\sum_{j=1}^{i+1}\alpha_j)} \int_{\frac{i}{i+1}}^{1} \frac{\Gamma(\sum_{j=1}^{i+1}\alpha_j)}{\Gamma(\sum_{j=1}^{i}\alpha_j)\Gamma(\alpha_{i+1})}\, \theta_i^{\sum_{j=1}^{i}\alpha_j-1}(1-\theta_i)^{\alpha_{i+1}-1}\, d\theta_i \\ &= \prod_{i=1}^{k-1} \frac{\Gamma(\sum_{j=1}^{i}\alpha_j)\Gamma(\alpha_{i+1})}{\Gamma(\sum_{j=1}^{i+1}\alpha_j)}[1- \mathrm{IB}_{i/(i+1)}(\sum_{j=1}^{i}\alpha_j, \alpha_{i+1})].\end{aligned}$$

This competes the proof.

Using the Lemma 1 we can show that, for $i=1, \cdots, k-1$,

$$\hat{p}_i = \frac{\prod_{l=1}^{k-1} \frac{\Gamma(\sum_{j=1}^{l}(a_j'+n_j))\Gamma(a_{l+1}'+n_{l+1})}{\Gamma(\sum_{j=1}^{l+1}(a_j'+n_j))}[1- \mathrm{IB}_{l/(l+1)}(\sum_{j=1}^{l}(a_j'+n_j), a_{l+1}'+n_{l+1})]}{\prod_{l=1}^{k-1} \frac{\Gamma(\sum_{j=1}^{l}(a_j+n_j))\Gamma(a_{l+1}+n_{l+1})}{\Gamma(\sum_{j=1}^{l+1}(a_j+n_j))}[1- \mathrm{IB}_{l/(l+1)}(\sum_{j=1}^{l}(a_j+n_j), a_{l+1}+n_{l+1})]}$$

$$= \frac{a_i+n_i}{\sum_{i=1}^{k}(a_i+n_i)} \prod_{l=\max(1,i-1)}^{k-1} \frac{[1- \mathrm{IB}_{l/(l+1)}(\sum_{j=1}^{l}(a_j'+n_j), a_{l+1}'+n_{l+1})]}{[1- \mathrm{IB}_{l/(l+1)}(\sum_{j=1}^{l}(a_j+n_j), a_{l+1}+n_{l+1})]},$$

where $a_l'=a_l$ for $l\neq i$, $a_i'=a_i+1$.

As the Bayes estimate under zero-one loss function of **p**, the mode of the posterior distribution can be obtained by finding the isotonic regression of a vector whose ith component is $(a_i+n_i)/[\sum_{i=1}^{k}(a_i+n_i)]$ provided $a_i+n_i-1>0$, $i=1,2,\dots,k$. Using the reparametrization scheme used in Lemma 1 (See also Dykstra and Robertson, 1982) we can easily find the Bayes estimate. Let $\hat{\theta}_i=\sum_{j=1}^{i}(a_j+n_j) \;/\; \sum_{j=1}^{i+1}(a_j+n_j)$ for $i=1,2,\dots,k-1$. Then the evaluation of $\mathbf{p}$ at $\boldsymbol{\theta} = \overline{\boldsymbol{\theta}} = (\overline{\theta}_1,\dots,\overline{\theta}_{k-1})$, where $\overline{\theta}_i=\max\{\hat{\theta}_i, i/(i+1)\}$, is the Bayes estimate. We note that for the case of flat prior, i.e., $a_i=1$ for all i, the Bayes estimate agrees with maximum likelihood estimate.

3. Poisson problem

Suppose we have a random sample of size n from each of k Poisson populations having means $\lambda_1,\lambda_2,\dots,\lambda_k$. The prior distribution on each λ_i is usually chosen to be a Gamma distribution with parameters a_i, and β. Specifically, $f(\lambda_i)=\frac{1}{\Gamma(a_i)\beta^{a_i}}\lambda_i^{a_i-1}e^{-\lambda_i/\beta}$, $\lambda_i>0$. We assume that a_i's and β are known. Now we are interested in estimating λ_i's subject to

$$\lambda_1 \geq \frac{\lambda_1+\lambda_2}{2} \geq \cdots \geq \frac{\lambda_1+\lambda_2+\cdots+\lambda_k}{k} \geq 0 .$$

The prior distribution of λ_i's incoporated with the above starshaped restriction is

$$C(a_1, a_2, \dots, a_k;\beta)\prod_{i=1}^{k}\frac{1}{\Gamma(a_i)\beta^{a_i-1}}\lambda_i^{a_i-1}e^{-\lambda_i/\beta}$$

on its support A_k, where

$$A_k=\{\boldsymbol{\lambda} : \lambda_i>0, i=1,2,\dots,k, \lambda_1\geq\frac{\lambda_1+\lambda_2}{2}\geq\cdots\geq\frac{\lambda_1+\lambda_2+\cdots+\lambda_k}{k}\geq 0\},$$

and

$$C(\alpha_1,\alpha_2,\dots,\alpha_k;\beta)^{-1}=\int_{A_k}\prod_{i=1}^{k}\frac{1}{\Gamma(\alpha_i)\beta^{\alpha_i-1}}\lambda_i^{\alpha_i-1}e^{-\lambda_i/\beta}\,d\boldsymbol{\lambda}.$$

Then the posterior distribution of λ_i's given data, $(x_{ij}, j=1,\dots,n, i=1,2,\dots,k)$, is

$$C(\alpha_1+n\bar{x}_1,\dots,\alpha_k+n\bar{x}_k;\frac{\beta}{1+n\beta})\prod_{i=1}^{k}\frac{1}{\Gamma(\alpha_i+n\bar{x}_i)}(\frac{1+n\beta}{\beta})^{\alpha_i-1}\lambda_i^{\alpha_i+n\bar{x}_i-1}e^{-(\frac{1}{\beta}+n)\lambda_i},$$

where $\bar{x}_i=\sum_{j=1}^{n}x_{ij}/n$. Then the Bayes estimates under square loss function of λ_i, $\hat{\lambda}_i$, is given by

$$\frac{\beta(\alpha_i+n\bar{x}_i)}{1+n\beta}\;\frac{C(\alpha_1+n\bar{x}_1,\dots,\alpha_k+n\bar{x}_k;\frac{\beta}{1+n\beta})}{C(\alpha_1+n\bar{x}_1,\dots,\alpha_{i-1}+n\bar{x}_{i-1},\alpha_i+n\bar{x}_i+1,\alpha_{i+1}+n\bar{x}_{i+1},\dots\alpha_k+n\bar{x}_k;\frac{\beta}{1+n\beta})}$$

Now we evaluate $C(\cdot\,;\cdot)$. This result can be found in a straightforward fashion using the results in Section 2 together with the fact that the conditional distribution of independent Gamma variables, given their sum, is a Dirichlet distribution. Consider a one-to-one transformation. Let $\phi_i=\lambda_i/\sum_{i=1}^{k}\lambda_i, i=1,\dots,k-1$, and $\phi_k=\sum_{i=1}^{k}\lambda_i$. Then $\lambda_i=\phi_i\phi_k$, $i=1,\dots,k-1$, and $\lambda_k=\phi_k(1-\sum_{i=1}^{k-1}\phi_i)$. The restriction including basic restriction becomes

$$0\leq\phi_i\leq 1, i=1,\dots,k-1, \sum_{i=1}^{k-1}\phi_i\leq 1, \phi_k>0,\ \phi_1\geq\frac{\phi_1+\phi_2}{2}\geq\cdots\geq\frac{\phi_1+\phi_2+\cdots+\phi_{k-1}}{k-1}\geq\frac{1}{k}.$$

Note that $|\frac{\partial(\lambda_1,\dots,\lambda_{k-1})}{\partial(\phi_1,\dots,\phi_{k-1})}|=\phi_k^{k-1}$. Hence we have

$$C(\alpha_1,\alpha_2,\dots,\alpha_k;\beta)^{-1}=\int_{B_k}\frac{\Gamma(\sum_{i=1}^{k}\alpha_i)}{\prod_{i=1}^{k}\Gamma(\alpha_i)}\prod_{i=1}^{k-1}\phi_i^{\alpha_i-1}(1-\sum_{i=1}^{k-1}\phi_i)^{\alpha_k-1}\,d\boldsymbol{\phi},$$

where

$$B_k=\{\boldsymbol{\phi} : 0\leq\phi_i\leq 1, i=1,\dots,k, \sum_{i=1}^{k}\phi_i=1,\ \phi_1\geq\frac{\phi_1+\phi_2}{2}\geq\cdots\geq\frac{\phi_1+\phi_2+\cdots+\phi_{k-1}}{k-1}\geq\frac{1}{k}\}.$$

Note that $C(\alpha_1,\alpha_2,\dots,\alpha_k;\beta)$ does not depend upon β.

Using Lemma 1 we have

$$C(\alpha_1, \alpha_2, \cdots, \alpha_k; \beta)^{-1} = \frac{\Gamma(\sum_{j=1}^{k} \alpha_j)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \prod_{i=1}^{k-1} \int_{\frac{i}{i+1}}^{1} y_i^{\sum_{j=1}^{i} \alpha_j - 1} (1-y_i)^{\alpha_{i+1}-1} dy_i$$
$$= \prod_{i=1}^{k-1} [1 - \text{IB}_{i/(i+1)}(\sum_{j=1}^{i} \alpha_j, \alpha_{i+1})].$$

Hence we have, for $i = 1, 2, \cdots, k$,

$$\hat{\lambda}_i = \frac{\beta(\alpha_i + n\bar{x}_i)}{1+n\beta} \prod_{l=\max(1,i-1)}^{k-1} \frac{[1 - \text{IB}_{l/(l+1)}(\sum_{j=1}^{l}(\alpha_j' + n\bar{x}_j), \alpha_{l+1}' + n\bar{x}_{l+1})]}{[1 - \text{IB}_{l/(l+1)}(\sum_{j=1}^{l}(\alpha_j + n\bar{x}_j), \alpha_{l+1} + n\bar{x}_{l+1})]},$$

where $\alpha_l' = \alpha_l$ for $l \neq i$, $\alpha_i' = \alpha_i + 1$.

The Bayes estimator under zero-one loss function of λ_i, the mode of posterior distribution can be obtained from Theorem 3.1 of Shaked (1979). See also Dykstra and Robertson (1982).

$$\hat{\lambda}_i = \frac{\sum_{j=1}^{k}(\alpha_j + n\bar{x}_j - 1)}{1/\beta + n} \cdot \max\{\frac{\alpha_i + n\bar{x}_i - 1)}{\sum_{j=1}^{k}(\alpha_j + n\bar{x}_j - 1)}, \frac{i}{i+1}\},\ i = 1, \ldots k-1,$$

$$\hat{\lambda}_k = \frac{\sum_{j=1}^{k}(\alpha_j + n\bar{x}_j - 1)}{1/\beta + n} \cdot [1 - \sum_{i=1}^{k-1} \max\{\frac{\alpha_i + n\bar{x}_i - 1)}{\sum_{j=1}^{k}(\alpha_j + n\bar{x}_j - 1)}, \frac{i}{i+1}\}].$$

It should be noticed that if the assumption of equal sample sizes is relaxed the aforementioned reparametrization scheme is no longer valid for Bayes estimation. We briefly discusse the estimation procedure for the case of unequal sample sizes. Let n_i be the sample size of random sample from ith population. Let $w_i = cn_i$ for known positive constant c. Recall that the lower starshaped restriction is decreasing on the average with nonnegativity requirement. Then the starshaped restriction mentioned above can be modified as

$$\lambda_1 \geq \frac{w_1\lambda_1 + w_2\lambda_2}{w_1 + w_2} \geq \cdots \geq \frac{w_1\lambda_1 + \cdots + w_k\lambda_k}{w_1 + \cdots + w_k} \geq 0 .$$

Suppose the prior distribution of λ_i is a Gamma distribution with shape parameter α_i and scale parameter β/w_i. Note that the scale parameters are proportional to sample sizes. Using the similar reparametrization scheme we can show that Bayes estimate under square loss function, $\hat{\lambda}_i$, $i = 1, \cdots, k,$, is given by

$$\hat{\lambda}_i = \frac{\beta(a_i + n_i \bar{x}_i)}{w_i + n_i \beta} \prod_{l=\max(1,i-1)}^{k-1} \frac{[1 - IB_{\sum_{j=1}^{l} w_j / \sum_{j=1}^{l+1} w_j}(\sum_{j=1}^{l}(a'_j + n_j \bar{x}_j), a'_{l+1} + n_l \bar{x}_{l+1})]}{[1 - IB_{\sum_{j=1}^{l} w_j / \sum_{j=1}^{l+1} w_j}(\sum_{j=1}^{l}(a_j + n_j \bar{x}_j), a_{l+1} + n_l \bar{x}_{l+1})]},$$

where $a'_l = a_l$ for $l \neq i$, $a'_i = a_i + 1$.

References

[1] Barlow, R.E., D.J. Bartholomew, J.M. Bremner, and H.D. Brunk (1972). *Statistical Inference under Order Restrictions,* John Wiley and Sons, London.

[2] Broffit, J.D. (1984). A Bayes estimator for ordered parameters and isotonic Bayesian graduation, *Scandinavian Acuarial Journal,* 1984, 231-247.

[3] Dykstra, R.L., and T. Robertson (1982). Order restricted statistical tests on multinomial and Poisson parameters: The starshaped restriction, *The Annals of Statistics*, Vol. 10, 1246-52.

[4] Gelfand, A.E., A.F.M. Smith, and T.-M. Lee (1992). Bayesian analysis of constrained parameter and truncated data problems using Gibbs sampling, *Journal of the American Statistical Association,* Vol. 87, 523-532.

[5] Kraft, C.H., and C. van Eeden (1964). Bayesian bioassay, *Annals of Mathematical Statistics,* Vol. 35. 886-890.

[6] Sedransk, J., J. Monahan, and H.Y. Chiu (1985). Bayesian estimation of finite population parameters in categorical data models incoporating order restrictions, *Journal of the Royal Statistical Society*, Series B, Vol. 47, 519-527.

[7] Shaked, M. (1979). Estimation of star-shaped sequences of Poisson and normal means, *The Annals of Statistics*, Vol. 7, 729-741.