Bayes and minimax estimation of parameters of Markov transition matrix

Babulal Seal, Sk Jakir Hossain

Department of Statistics, Burdwan University, Burdwan, W.B., India

Abstract. Based on sample observations, Bayes estimator and minimax estimator for the transition probability matrix are worked out in the cases where we have belief regarding the parameters. For example, where the states seem to be equal or not, are worked out using Dirichlet prior. In both cases priors are in accordance with our beliefs. Using the Bayes method, minimax estimator is also obtained.

1. Introduction

Internal migration model (Seal and Hossain, 2013b) was used for estimation of the parameters involved in analysing internal migration data and also testing (Seal and Hossain, 2013a) for the same. Even a measure of importance of the states and their estimation was developed by the authors (Seal and Hossain, 2013c). Now we look at another way of estimating the parameters of the transition matrix.

Internal migrational behavior of population is attempted with the Markov model. Estimation of parameters of transition probability matrix in classical way is available in literature. Now the knowledge of behaviors of migration pattern should be employed in finding Bayes procedure. For example, if we have belief that all states are almost equally important then we should try to add such beliefs into estimation procedures. In this paper we shall work with such notion e.g., Bayes and minimax. The importance of such work nowadays is important if we follow "In his report (MacPherson, 2004) on the 'Strengthening of United Nations: an agenda for further change' UN Secretary-General Kofi Annan identified migration as a priority issue".

For several years, a progressive new trend has been growing in applied statistics. It is becoming even more popular to build application-specific models that are designed to account for the hierarchical and latent structures inherent in any particular data generation mechanism. But the development of methodological and computational tools for statistical analysis began to bring such model fitting into routine practice.

In Section 2 Bayes estimation of the parameters under Dirichlet prior is given. Section 3 includes Bayes estimation where the states or places seem to be equal (i.e., they seem to be equally important). There is a way to find out minimax estimator through Bayes estimation, e.g. Equaliser Bayes is Minimax. In Section 4, Minimax estimator through Bayes procedure is obtained. Also it is compared with MLE.

2. Bayes estimation under conjugate prior

In this section we would like to evaluate the Bayes' estimator for the multinomial parameter $\mathbf{p}_i = (p_{i1}, p_{i2}, \cdots, p_{ik})$, for all $i = 1, 2, \cdots, k$ under Dirichlet prior, which is conjugate to the p.m.f. of $\mathbf{n}_i$, for all

$i = 1, 2, \cdots, k$. Bayes estimator is easier to find out under this prior but if there are sufficiently large sample then such assumption on prior is immaterial.

Result 2.1. *The Bayes estimator of p_{ij}'s under Dirichlet prior with constant vector $\alpha_i = (\alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{ik})$ for all $i = 1, 2, \cdots, k$ is $\delta(n_{ij}) = \frac{n_{ij} + \alpha_{ij}}{n_{i\cdot} + \alpha_{i\cdot}}$ for all $i, j = 1, 2, \cdots, k$.*

Proof. A natural conjugate prior for the parameter $\mathbf{p}_i$, for all $i = 1, 2, \cdots, k$ is given by the Dirichlet density with parameter $\alpha_{\mathbf{i}} = (\alpha_{i1}, \cdots, \alpha_{ik})$, for all $i = 1, 2, \cdots, k$. Thus the joint prior density is given by

$$\pi(\mathbf{p}) = \prod_{i=1}^{k} \pi(\mathbf{p}_i) = \prod_{i=1}^{k} \left(\frac{\Gamma\left(\sum_{j=1}^{k} \alpha_{ij}\right)}{\prod_{j=1}^{k} \Gamma(\alpha_{ij})} \prod_{j=1}^{k} p_{ij}^{\alpha_{ij}-1} \right).$$

The joint distribution of $\mathbf{n} = (\mathbf{n}_1^T, \mathbf{n}_2^T, \cdots \mathbf{n}_k^T)^T$ and $\mathbf{p} = (\mathbf{p}_1^T, \mathbf{p}_2^T, \cdots \mathbf{p}_k^T)^T$ is given by

$$f(\mathbf{n}|\mathbf{p})\pi(\mathbf{p}) = \prod_{i=1}^{k} f(\mathbf{n}_i|\mathbf{p}_i)\pi(\mathbf{p}_i) = \prod_{i=1}^{k} \left(\frac{n_{i\cdot}!}{\prod_{j=1}^{k} n_{ij}!} \frac{\Gamma\left(\sum_{j=1}^{k} \alpha_{ij}\right)}{\prod_{j=1}^{k} \Gamma(\alpha_{ij})} \prod_{j=1}^{k} p_{ij}^{n_{ij}+\alpha_{ij}-1} \right).$$

The marginal p.m.f. of $\mathbf{n}$ is

$$g(\mathbf{n}) = \int_{S_1} \cdots \int_{S_k} f(\mathbf{n}|\mathbf{p})\pi(\mathbf{p}) \prod_{i=1}^{k} \prod_{j=1}^{k} dp_{ij},$$

where the integration is carried out over the region

$$S_i = \left\{ p_{ij} : p_{ij} \geq 0 \text{ and } \sum_{j=1}^{k} p_{ij} = 1 \right\}, \text{ for all } i = 1, 2, \cdots$$

After integration we have the marginal p.m.f. of $\mathbf{n}$ is

$$g(\mathbf{n}) = \prod_{i=1}^{k} \left[\frac{n_{i\cdot}!}{\prod_{j=1}^{k} n_{ij}!} \frac{\Gamma\left(\sum_{j=1}^{k} \alpha_{ij}\right)}{\prod_{j=1}^{k} \Gamma(\alpha_{ij})} \frac{\prod_{j=1}^{k} \Gamma(n_{ij} + \alpha_{ij})}{\Gamma\left(\sum_{j=1}^{k} (n_{ij} + \alpha_{ij})\right)} \right].$$

Then the posterior distribution of $\mathbf{p}$ is given by

$$\pi(\mathbf{p}|\mathbf{n}) = \frac{f(\mathbf{n}|\mathbf{p})\pi(\mathbf{p})}{g(\mathbf{n})} = \prod_{i=1}^{k} \frac{f(\mathbf{n}_i|\mathbf{p_i})\pi(\mathbf{p_i})}{g(\mathbf{n}_i)} = \prod_{i=1}^{k} \left[\frac{\Gamma\left(\sum_{j=1}^{k} (n_{ij} + \alpha_{ij})\right)}{\prod_{j=1}^{k} \Gamma(n_{ij} + \alpha_{ij})} \prod_{j=1}^{k} p_{ij}^{n_{ij}+\alpha_{ij}-1} \right], \tag{1}$$

which is the product of k independent Dirichlet distribution with parameter α_i', $i = 1, 2, \cdots, k$, where $\alpha_i' = (\alpha_{i1}', \cdots, \alpha_{ik}')^T = (n_{i1} + \alpha_{i1}, \cdots, n_{ik} + \alpha_{ik})^T$, for all $i = 1, 2, \cdots, k$.

Now we are interested to find out the posterior mean of p_{ij}'s. Let r be a specified value of i, then from (1) we have

$$\pi(\mathbf{p_r}|\mathbf{n_r}) = \frac{\Gamma\left(\sum_{j=1}^{k}(n_{rj} + \alpha_{rj})\right)}{\prod_{j=1}^{k}\Gamma(n_{rj} + \alpha_{rj})} \prod_{j=1}^{k} p_{rj}^{n_{rj}+\alpha_{rj}-1}.$$

Let j_o be the specified value of j, then

$$E(p_{rj_o}|\mathbf{n_r}) = \frac{\Gamma\left(\sum_{j=1}^{k}(n_{rj} + \alpha_{rj})\right)}{\prod_{j=1}^{k}\Gamma(n_{rj} + \alpha_{rj})} \int_{S_r} \prod_{\substack{j=1 \\ j \neq j_o}}^{k} p_{rj}^{n_{rj}+\alpha_{rj}-1} p_{rj_o}^{n_{rj_o}+\alpha_{rj_0}} \prod_{j=1}^{k} dp_{rj}$$

$$= \frac{\Gamma\left(\sum_{j=1}^{k}(n_{rj} + \alpha_{rj})\right)}{\prod_{j=1}^{k}\Gamma(n_{rj} + \alpha_{rj})} \frac{\Gamma(n_{rj_o} + \alpha_{rj_o} + 1)\prod_{\substack{j=1 \\ j \neq j_o}}^{k}\Gamma(n_{rj} + \alpha_{rj})}{\Gamma\left(\sum_{j=1}^{k}(n_{rj} + \alpha_{rj}) + 1\right)}$$

$$= \frac{\alpha_{rj_o} + n_{rj_o}}{\sum_{j=1}^{k}\alpha_{rj} + n_{r.}}.$$

Since r and j_o are specified value of i and j respectively, so we have,

$$E(p_{ij}) = \frac{\alpha_{ij} + n_{ij}}{\sum_{j=1}^{k}\alpha_{ij} + n_{i.}}, \quad \text{for all } i, j = 1, 2, \cdots, k.$$

Therefore the Bayes estimator of p_{ij} under sum of squared error loss is

$$\delta(n_{ij}) = \frac{\alpha_{ij} + n_{ij}}{\sum_{j=1}^{k}\alpha_{ij} + n_{i.}} = \frac{\alpha_{ij} + n_{ij}}{\alpha_{i.} + n_{i.}},$$

where $\alpha_{i.} = \sum_{j=1}^{k}\alpha_{ij}$, for all $i = 1, 2, \cdots, k$. $\quad\square$

3. Bayes estimation where the states seem to be equal

In this section we would like to evaluate the Bayes estimator under the constraint that all the states are equally important, i.e. $p_{11} = p_{22} = \cdots = p_{kk} = p$, where the value of p is unknown.

We have parameters $\mathbf{p} = (p_{11}, p_{12}, \cdots, p_{1k}; p_{21}, p_{22}, \cdots, p_{2k}; \cdots; p_{k1}, p_{k2}, \cdots, p_{kk})$ and our interest is in $(p_{11}, p_{22}, \cdots, p_{kk})$, and others are nuisance parameter.

Let us denote $\mathbf{p}_{-i} = (p_{i1}, \cdots, p_{i,i-1}, p_{i,i+1}, \cdots, p_{ik})$, for all $i = 1, 2, \cdots, k$ and also $\mathbf{p}_{-} = (\mathbf{p}_{-1}^T, \cdots, \mathbf{p}_{-k}^T)^T$. Then we have the joint p.m.f. of n_{ij}'s, and for all $i, j = 1, 2, \cdots, k$ is

$$
\begin{aligned}
f(\mathbf{n}|p, \mathbf{p}_{-}) &= \prod_{i=1}^{k} f(\mathbf{n}_i|p, \mathbf{p}_{-i}) = \prod_{i=1}^{k} \left(\frac{n_{i.}!}{n_{i1}!n_{i2}!\cdots n_{ik}!} p_{i1}^{n_{i1}} p_{i2}^{n_{i2}} \cdots p_{i,i-1}^{n_{i,i-1}} p^{n_{ii}} p_{i,i+1}^{n_{i,i+1}} p_{ik}^{n_{ik}} \right) \\
&= \prod_{i=1}^{k} \left(\frac{n_{i.}!}{n_{i1}!n_{i2}!\cdots n_{ik}!} \right) \times p^{\sum_{i=1}^{k} n_{ii}} \prod_{i=1}^{k} \prod_{\substack{j=1 \\ j \neq i}}^{k} p_{ij}^{n_{ij}}.
\end{aligned}
\tag{2}
$$

Under $p_{11} = p_{22} = \cdots = p_{kk} = p$ (with p unknown), the prior distribution is taken as

$$
\begin{aligned}
\pi(p, \mathbf{p}_{-}) &= \prod_{i=1}^{k} \left[\frac{\Gamma(\alpha_{i.})}{\Gamma(\alpha_{i1})\Gamma(\alpha_{i2})\cdots\Gamma(\alpha_{ik})} \times p_{i1}^{\alpha_{i1}-1} \cdots p_{i,i-1}^{\alpha_{i,i-1}-1} p^{\alpha_{ii}-1} p_{i,i+1}^{\alpha_{i,i+1}-1} \cdots p_{ik}^{\alpha_{ik}-1} \right] \\
&= \prod_{i=1}^{k} \left[\frac{\Gamma(\alpha_{i.})}{\Gamma(\alpha_{i1})\Gamma(\alpha_{i2})\cdots\Gamma(\alpha_{ik})} \right] p^{\sum_{i=1}^{k} \alpha_{ii} - k} \prod_{i=1}^{k} \prod_{\substack{j=1 \\ j \neq i}}^{k} p_{ij}^{\alpha_{ij}-1},
\end{aligned}
\tag{3}
$$

where $\alpha_{i.} = \sum_{j=1}^{k} \alpha_{ij}$, for all $i = 1, \cdots, k$.

Result 3.1. *Bayes estimator of p_{ij}'s when $p_{11} = p_{22} = \cdots = p_{kk} = p$ (with p unknown) is*

$$
\hat{\delta}(n_{ii}) = \frac{\sum_{i=1}^{k}(n_{ii} + \alpha_{ii}) - (k-1)}{n + \alpha - 2(k-1)} \quad \text{for all } i = 1, 2, \cdots, k,
$$

and

$$
\hat{\delta}(n_{ij}) = \frac{n_{ij} + \alpha_{ij}}{\sum_{\substack{j=1 \\ j \neq i}}^{k}(n_{ij} + \alpha_{ij})} \frac{\sum_{i=1}^{k} \sum_{\substack{j=1 \\ j \neq i}}^{k}(n_{ij} + \alpha_{ij}) - (k-1)}{n + \alpha - 2(k-1)} \quad \text{for all } i, j(j \neq i) = 1, 2, \cdots, k.
$$

Proof. We have from (2) and (3) the joint distribution of n_{ij}'s and p, p_{ij}'s $(i \neq j)$ is

$$
f(\mathbf{n}|p, \mathbf{p}_{-})\pi(p, \mathbf{p}_{-}) = \prod_{i=1}^{k} \left[\frac{n_{i.}!}{n_{i1}!n_{i2}!\cdots n_{ik}!} \frac{\Gamma(\alpha_{i.})}{\Gamma(\alpha_{i1})\cdots\Gamma(\alpha_{ik})} \right] p^{\sum_{i=1}^{k}(n_{ii} + \alpha_{ii}) - k} \prod_{i=1}^{k} \prod_{\substack{j=1 \\ j \neq i}}^{k} p_{ij}^{n_{ij}+\alpha_{ij}-1}.
$$

Then the posterior distribution is of the form

$$
\pi(p, \mathbf{p}_{-}|\mathbf{n}) = \frac{1}{C(\mathbf{n}, \alpha)} \times p^{\sum_{i=1}^{k}(n_{ii}+\alpha_{ii}) - k} \prod_{i=1}^{k} \prod_{\substack{j=1 \\ j \neq i}}^{k} p_{ij}^{n_{ij}+\alpha_{ij}-1},
\tag{4}
$$

where

$$C(\mathbf{n}, \alpha) = \int_0^1 \left(\int_{S_1'} \cdots \int_{S_k'} p^{\sum\limits_{i=1}^{k}(n_{ii}+\alpha_{ii})-k} \prod_{i=1}^{k}\prod_{\substack{j=1\\j\neq i}}^{k} p_{ij}^{n_{ij}+\alpha_{ij}-1} \prod_{i=1}^{k}\prod_{\substack{j=1\\j\neq i}}^{k} dp_{ij} \right) dp,$$

where the integration is carried out over the region

$$S_i' = \left\{ p_{ij} : p_{ij} \geq 0, p \geq 0, \text{ and } \sum_{\substack{j=1\\j\neq i}}^{k} p_{ij} = 1-p \right\} \text{ for all } i = 1, 2, \cdots, k.$$

Now, for a particular i, we have

$$\int_{S_i'} \prod_{j=1}^{i-1} p_{ij}^{n_{j1}+\alpha_{j1}-1} \prod_{l=i+1}^{k} p_{il}^{n_{il}+\alpha_{il}-1} \prod_{\substack{j=1\\j\neq i}}^{k} dp_{ij} = \frac{\prod\limits_{\substack{j=1\\j\neq i}}^{k}\Gamma(n_{ij}+\alpha_{ij})}{\Gamma\left(\sum\limits_{\substack{j=1\\j\neq i}}^{k}(n_{ij}+\alpha_{ij})\right)}(1-p)^{\sum\limits_{\substack{j=1\\j\neq i}}^{k}(n_{ij}+\alpha_{ij})-1}, \qquad (5)$$

for all $i = 1, 2, \cdots, k$.

So after integrating w.r.t. p_{ij}'s $(i \neq j)$ from (4) we have

$$
\begin{aligned}
C(\mathbf{n}, \alpha) &= \frac{\prod\limits_{i=1}^{k}\prod\limits_{\substack{j=1\\j\neq i}}^{k}\Gamma(n_{ij}+\alpha_{ij})}{\prod\limits_{i=1}^{k}\Gamma\left(\sum\limits_{\substack{j=1\\j\neq i}}^{k}(n_{ij}+\alpha_{ij})\right)} \int_0^1 p^{\sum\limits_{i=1}^{k}(n_{ii}+\alpha_{ii})-k}(1-p)^{\sum\limits_{i=1}^{k}\sum\limits_{\substack{j=1\\j\neq i}}^{k}(n_{ij}+\alpha_{ij})-k} dp \\[2ex]
&= \frac{\prod\limits_{i=1}^{k}\prod\limits_{\substack{j=1\\j\neq i}}^{k}\Gamma(n_{ij}+\alpha_{ij})}{\prod\limits_{i=1}^{k}\Gamma\left(\sum\limits_{\substack{j=1\\j\neq i}}^{k}(n_{ij}+\alpha_{ij})\right)} \frac{\Gamma\left(\sum\limits_{i=1}^{k}(n_{ii}+\alpha_{ii})-(k-1)\right)\Gamma\left(\sum\limits_{i=1}^{k}\sum\limits_{\substack{j=1\\j\neq i}}^{k}(n_{ij}+\alpha_{ij})-(k-1)\right)}{\Gamma(n+\alpha-2(k-1))},
\end{aligned}
$$

where $n = \sum\limits_{i=1}^{k}\sum\limits_{j=1}^{k} n_{ij}$ and $\alpha = \sum\limits_{i=1}^{k}\sum\limits_{j=1}^{k} \alpha_{ij}$.

Thus we get

$$\frac{1}{C(\mathbf{n},\alpha)} = \frac{\Gamma(n+\alpha-2(k-1))}{\Gamma\left(\sum_{i=1}^{k}(n_{ii}+\alpha_{ii})-(k-1)\right)\Gamma\left(\sum_{i=1}^{k}\sum_{\substack{j=1\\j\neq i}}^{k}(n_{ij}+\alpha_{ij})-(k-1)\right)} \cdot \frac{\prod_{i=1}^{k}\Gamma\left(\sum_{\substack{j=1\\j\neq i}}^{k}(n_{ij}+\alpha_{ij})\right)}{\prod_{i=1}^{k}\prod_{\substack{j=1\\j\neq i}}^{k}\Gamma(n_{ij}+\alpha_{ij})}. \quad (6)$$

From (4) we see that the marginal distribution of p follows Beta distribution with parameters

$$\sum_{i=1}^{k}(n_{ii}+\alpha_{ii})-(k-1) \quad \text{and} \quad \sum_{i=1}^{k}\sum_{\substack{j=1\\j\neq i}}^{k}(n_{ij}+\alpha_{ij})-(k-1).$$

Therefore we have

$$E(p) = \frac{\sum_{i=1}^{k}(n_{ii}+\alpha_{ii})-(k-1)}{n+\alpha-2(k-1)}.$$

Thus Bayes estimator of p_{ii} is

$$\hat{\delta}(n_{ii}) = \frac{\sum_{i=1}^{k}(n_{ii}+\alpha_{ii})-(k-1)}{n+\alpha-2(k-1)} \quad \text{for all } i=1,2,\cdots,k.$$

Now we want to find out the posterior mean of p_{ij}'s for $(i\neq j)$

In particular

$$E(p_{12}) = \frac{1}{C(\mathbf{n},\alpha)}\int_{0}^{1}p^{\sum_{i=1}^{k}(n_{ii}+\alpha_{ii})-k}\int_{S_1'}p_{12}^{n_{12}+\alpha_{12}}\prod_{j=3}^{k}p_{1j}^{n_{1j}+\alpha_{1j}-1}\prod_{j=2}^{k}dp_{1j}\times$$

$$\int_{S_2'}\prod_{\substack{j=1\\j\neq 2}}^{k}p_{2j}^{n_{2j}+\alpha_{2j}-1}\prod_{\substack{j=1\\j\neq 2}}^{k}dp_{2j}\times\cdots\times\int_{S_k'}\prod_{\substack{j=1\\j\neq k}}^{k}p_{kj}^{n_{kj}+\alpha_{kj}-1}\prod_{\substack{j=1\\j\neq k}}^{k}dp_{kj}dp. \quad (7)$$

Now,

$$\int_{S_1'}p_{12}^{n_{12}+\alpha_{12}}p_{13}^{n_{13}+\alpha_{13}-1}\cdots p_{1k}^{n_{1k}+\alpha_{1k}-1}\prod_{j=2}^{k}dp_{1j} = \frac{\Gamma(n_{12}+\alpha_{12}+1)\prod_{j=3}^{k}\Gamma(n_{1j}+\alpha_{1j})}{\Gamma\left(\sum_{j=2}^{k}(n_{1j}+\alpha_{1j})+1\right)}(1-p)^{\sum_{j=2}^{k}(n_{1j}+\alpha_{1j})}. \quad (8)$$

Thus we have, from (7) using (8) and (5) for $i=2,3,\cdots,k$

$$E(p_{12}) = C'(\mathbf{n},\alpha)\int_{0}^{1}p^{\sum_{i=1}^{k}(n_{ii}+\alpha_{ii})-k}(1-p)^{\sum_{i=1}^{k}\sum_{\substack{j=1\\j\neq i}}^{k}(n_{ij}+\alpha_{ij})-(k-1)}dp$$

where

$$C'(\mathbf{n}, \alpha) = \frac{1}{C(\mathbf{n}, \alpha)} \times \frac{\Gamma(n_{12} + \alpha_{12} + 1)\prod\limits_{j=3}^{k}\Gamma(n_{1j} + \alpha_{1j})}{\Gamma\left(\sum\limits_{j=2}^{k}(n_{1j} + \alpha_{1j}) + 1\right)} \times \frac{\prod\limits_{\substack{i=2 \\ }}^{k}\prod\limits_{\substack{j=1 \\ j\neq i}}^{k}\Gamma(n_{ij} + \alpha_{ij})}{\prod\limits_{i=2}^{k}\Gamma\left(\sum\limits_{\substack{j=1 \\ j\neq i}}^{k}(n_{ij} + \alpha_{ij})\right)}$$

$$= \frac{n_{12} + \alpha_{12}}{\sum\limits_{j=2}^{k}(n_{1j} + \alpha_{1j})} \frac{\Gamma(n + \alpha - 2(k-1))}{\Gamma\left(\sum\limits_{i=1}^{k}(n_{ii} + \alpha_{ii}) - (k-1)\right)\Gamma\left(\sum\limits_{i=1}^{k}\sum\limits_{\substack{j=1 \\ j\neq i}}^{k}(n_{ij} + \alpha_{ij}) - (k-1)\right)} \quad \text{[using (6)].} \quad (9)$$

Therefore

$$E(p_{12}) = C'(\mathbf{n}, \alpha)\frac{\Gamma\left(\sum\limits_{i=1}^{k}(n_{ii} + \alpha_{ii}) - (k-1)\right)\Gamma\left(\sum\limits_{i=1}^{k}\sum\limits_{\substack{j=1 \\ j\neq i}}^{k}(n_{ij} + \alpha_{ij}) - (k-2)\right)}{\Gamma(n + \alpha - (2k-3))}$$

$$= \frac{n_{12} + \alpha_{12}}{\sum\limits_{j=2}^{k}(n_{1j} + \alpha_{1j})} \frac{\sum\limits_{i=1}^{k}\sum\limits_{\substack{j=1 \\ j\neq i}}^{k}(n_{ij} + \alpha_{ij}) - (k-1)}{n + \alpha - 2(k-1)} \quad \text{[using (9)].}$$

Therefore we get, Bayes estimate of $p_{ij}, (j \neq i)$ to be

$$\hat{\delta}(n_{ij}) = \frac{n_{ij} + \alpha_{ij}}{\sum\limits_{\substack{j=1 \\ j\neq i}}^{k}(n_{ij} + \alpha_{ij})} \frac{\sum\limits_{i=1}^{k}\sum\limits_{\substack{j=1 \\ j\neq i}}^{k}(n_{ij} + \alpha_{ij}) - (k-1)}{n + \alpha - 2(k-1)} , \text{ for all } i, j(j \neq i) = 1, 2, \cdots, k,$$

and that completes the proof. $\square$

4. Minimax Estimator

In this section, we develop minimax estimator for p_{ij}'s , for all i, j with respect to the sum of squared error loss. Here we use the result that a Bayes estimator with constant risk is minimax. Now the risk

function for ith state is

$$
\begin{aligned}
E_{\mathbf{p_i}}\left[\sum_{j=1}^{k}\{\delta(n_{ij})-p_{ij}\}^2\right] &= \sum_{j=1}^{k}[\text{Bias}(\delta(n_{ij}))]^2 + \sum_{j=1}^{k}\text{Var}[\delta(n_{ij})] \\
&= \sum_{j=1}^{k}\left[p_{ij}-E\left(\frac{\alpha_{ij}+n_{ij}}{\alpha_{i\cdot}+n_{i\cdot}}\right)\right]^2 + \sum_{j=1}^{k}\text{Var}\left[\frac{\alpha_{ij}+n_{ij}}{\alpha_{i\cdot}+n_{i\cdot}}\right] \\
&= \sum_{j=1}^{k}\left[p_{ij}-\left(\frac{\alpha_{ij}+n_{i\cdot}p_{ij}}{\alpha_{i\cdot}+n_{i\cdot}}\right)\right]^2 + \sum_{j=1}^{k}\left[\frac{n_{i\cdot}p_{ij}(1-p_{ij})}{(\alpha_{i\cdot}+n_{i\cdot})^2}\right] \\
&= \frac{\displaystyle\sum_{j=1}^{k}(p_{ij}\alpha_{i\cdot}-\alpha_{ij})^2 + \sum_{j=1}^{k}n_{i\cdot}p_{ij}(1-p_{ij})}{(\alpha_{i\cdot}+n_{i\cdot})^2} \\
&= \frac{(\alpha_{i\cdot}^2-n_{i\cdot})\displaystyle\sum_{j=1}^{k}p_{ij}^2 + \sum_{j=1}^{k}\alpha_{ij}^2 + \sum_{j=1}^{k}(n_{i\cdot}-2\alpha_{i\cdot}\alpha_{ij})p_{ij}}{(\alpha_{i\cdot}+n_{i\cdot})^2} \\
&= \frac{1}{(\alpha_{i\cdot}+n_{i\cdot})^2}\left[(\alpha_{i\cdot}^2-n_{i\cdot})\left(\sum_{j=1}^{k-1}p_{ij}^2 + (1-\sum_{j=1}^{k-1}p_{ij})^2\right) + \sum_{j=1}^{k}\alpha_{ij}^2 \right.\\
&\qquad\left. + \sum_{j=1}^{k-1}(n_{i\cdot}-2\alpha_{i\cdot}\alpha_{ij})p_{ij} + (n_{i\cdot}-2\alpha_{i\cdot}\alpha_{ik})(1-\sum_{j=1}^{k-1}p_{ij})\right] \\
&= \frac{1}{(\alpha_{i\cdot}+n_{i\cdot})^2}\left[(\alpha_{i\cdot}^2-n_{i\cdot})\left(\sum_{j=1}^{k-1}p_{ij}^2 + (1-\sum_{j=1}^{k-1}p_{ij})^2\right) + \sum_{j=1}^{k}\alpha_{ij}^2 \right.\\
&\qquad\left. + \sum_{j=1}^{k-1}\left((n_{i\cdot}-2\alpha_{i\cdot}\alpha_{ij})-(n_{i\cdot}-2\alpha_{i\cdot}\alpha_{ik})\right)p_{ij} + (n_{i\cdot}-2\alpha_{i\cdot}\alpha_{ik})\right]. \quad (10)
\end{aligned}
$$

Now in order to make it constant over all p_{ij}, we need coefficient of p_{ij} and that of p_{ij}^2 to be zero. Which implies $\alpha_{i\cdot}^2-n_{i\cdot}=0$, for all $i=1,2,\cdots,k$, and $(n_{i\cdot}-2\alpha_{i\cdot}\alpha_{ij})-(n_{i\cdot}-2\alpha_{i\cdot}\alpha_{ik})=0$, for all $j=1,2,\cdots,k-1$. Which gives $\alpha_{i\cdot}=\sqrt{n_{i\cdot}}$, for all $i=1,2,\cdots,k$, and $\alpha_{i1}=\alpha_{i1}=\cdots=\alpha_{ik}=\frac{1}{k}\sqrt{n_{i\cdot}}$, for all $i=1,2,\cdots,k$. So that minimax estimator of p_{ij}, using sum of squared error loss, is

$$
\delta^*(n_{ij}) = \frac{n_{ij}+\frac{1}{k}\sqrt{n_{i\cdot}}}{n_{i\cdot}+\sqrt{n_{i\cdot}}}, \text{ for all } i,j=1,2,\cdots,k.
$$

From (10) we have the minimax risk

$$
r_{\delta_i^*} = \frac{k-1}{k}\frac{1}{(1+\sqrt{n_{i\cdot}})^2}, \text{ for all } i=1,2,\cdots,k
$$

and risk of ML estimator

$$
r_{\hat{\mathbf{p}}_i} = E\left[\sum_{j=1}^{k}(p_{ij}-\hat{p}_{ij})^2\right] = \sum_{j=1}^{k}\frac{p_{ij}(1-p_{ij})}{n_{i\cdot}} = \frac{1-\displaystyle\sum_{i=1}^{k}p_{ij}^2}{n_{i\cdot}}, \text{ for all } i=1,2,\cdots,k.
$$

This shows that minimax risk is less than the maximum possible risk of m.l.e, for all $i = 1, 2, \cdots, k$. This holds when in particular $p_{i1} = p_{i2} = \cdots = p_{ik} = \frac{1}{k}$ for all $i = 1, 2, \cdots, k$. Then risk of MLE $r_{\hat{\mathbf{p}}_i} = \frac{(k-1)}{k} \frac{1}{n_{i.}}$ for all $i = 1, 2, \cdots, k$, which is larger than minimax risk. Thus minimax estimator is better than maximum likelihood estimator in this particular case also.

5. Concluding Remarks

More importantly we have worked out in Section 2 and 3 for a class of priors. But from practical point of view if there are other guesses regarding transitions, then that kind of prior should be chosen by imposing conditions on hyperparameters of the prior model. The spirit of the thinking behind the investigation and development reported here came from (Shryoack et al., 1976) and (United Nations, 1970). A good reference for Bayes and Minimax procedures is (Robbins, 1955).

References

MacPherson, D. W. (2004) Irregular migration and health, Global Commission on International Migration, Geneva.

Robbins, H. (1955) An Empirical Bayes Approach to Statistics, in Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability 1, Berkeley, University of California Press, 157–164.

Seal, B., Hossain, S.J. (2013a) Testing of hypotheses: LRT test of some important hypothesis for internal migration model", Slovak Statistics and Demography - Scientific Journal, Vol. 23, Issue 2, 35–47.

Seal, B., Hossain, S.J. (2013b) An Internal Migration Model and Estimation of its Parameters, International Journal of Mathematics and Statistics, Vol. 14, Issue 2, 66–77.

Seal, B., Hossain, S.J. (2013c) Measure of state's importance and its estimation from migration data", ProbStat Forum, Vol. 6, April 2013, pages 10–17.

Shryock, H.S., Siegel, S. et al. (1976) The Methods and Materials of Demography, Academic Press, New York.

United Nations (1970) Methods of Measuring Internal Migration, Manuals on Methods of Estimating Population, Manual VI.