Accepted Manuscript

Bayesian simultaneous estimation for means in k-sample problems

Ryo Imai, Tatsuya Kubokawa, Malay Ghosh

PII: S0047-259X(17)30679-6
DOI: https://doi.org/10.1016/j.jmva.2018.08.013
Reference: YJMVA 4408

To appear in: *Journal of Multivariate Analysis*

Received date : 15 November 2017

Bayesian simultaneous estimation for means in k-sample problems

Ryo Imai[a], Tatsuya Kubokawa[b], Malay Ghosh[c]

[a]*Graduate School of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[b]*Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[c]*Department of Statistics, University of Florida, 102 Griffin-Floyd Hall, Gainesville, FL 32611, USA*

Abstract

This paper is concerned with the simultaneous estimation of k population means when one suspects that the k means are nearly equal. As an alternative to the preliminary test estimator based on the test statistics for testing hypothesis of equal means, we derive Bayesian and minimax estimators which shrink individual sample means toward a pooled mean estimator given under the hypothesis. It is shown that both the preliminary test estimator and the Bayesian minimax shrinkage estimators are further improved by shrinking the pooled mean estimator. The performance of the proposed shrinkage estimators is investigated by simulation.

Keywords: Bayes estimator, Empirical Bayes, k sample problem, Minimaxity, Quadratic loss, Shrinkage estimator
2010 MSC: 62C20, 62F15

1. Introduction

Consider the multivariate k-sample problem expressed in the following canonical form: p-variate random vectors $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k$ and a positive scalar random variable S are mutually independent and distributed as

$$\boldsymbol{X}_1 \sim \mathcal{N}_p(\boldsymbol{\mu}_1, \sigma^2 \boldsymbol{V}_1), \quad \ldots, \quad \boldsymbol{X}_k \sim \mathcal{N}_p(\boldsymbol{\mu}_k, \sigma^2 \boldsymbol{V}_k), \quad S/\sigma^2 \sim \chi^2_{(n)}, \tag{1}$$

where the $p\times 1$ means $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k$ and the scale parameter σ^2 are unknown, and $\boldsymbol{V}_1, \ldots, \boldsymbol{V}_k$ are $p\times p$ known and positive definite symmetric matrices. In this model, we want to estimate $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k$ simultaneously relative to the quadratic loss function

$$L(\boldsymbol{\delta}, \boldsymbol{\omega}) = \frac{1}{\sigma^2}\sum_{i=1}^k \|\boldsymbol{\delta}_i - \boldsymbol{\mu}_i\|^2_{\boldsymbol{Q}_i} = \frac{1}{\sigma^2}\sum_{i=1}^k (\boldsymbol{\delta}_i - \boldsymbol{\mu}_i)^\top \boldsymbol{Q}_i (\boldsymbol{\delta}_i - \boldsymbol{\mu}_i), \tag{2}$$

where $\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_k$ are $p \times p$ known and positive definite symmetric matrices, $\|\boldsymbol{a}\|^2_{\boldsymbol{A}} = \boldsymbol{a}^\top \boldsymbol{A}\boldsymbol{a}$ with $\boldsymbol{a}^\top$ standing for the transpose of $\boldsymbol{a}$, $\boldsymbol{\omega} = (\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k, \sigma^2)$ is a vector of unknown parameters and $\boldsymbol{\delta} = (\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_k)$ esimates $(\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k)$.

A typical example of Model (1) is a k-sample problem. For $i \in \{1, \ldots, k\}$, a random sample of size n_i is drawn from the ith population, say $\boldsymbol{X}_{i1}, \ldots, \boldsymbol{X}_{in_i} \sim \mathcal{N}_p(\boldsymbol{\mu}_i, \sigma^2 \boldsymbol{V}_{i,0})$, where $\boldsymbol{V}_{i,0}$ is a known matrix. In this case, $\boldsymbol{X}_i$, $\boldsymbol{V}_i$, S, and n in (1) correspond to

$$\overline{\boldsymbol{X}}_i = \sum_{j=1}^{n_i} \boldsymbol{X}_{ij}/n_i, \quad \boldsymbol{V}_{i,0}/n_i, \quad \mathrm{tr}\left\{\sum_{i=1}^k \boldsymbol{V}_{i,0}^{-1}\sum_{j=1}^{n_i}(\boldsymbol{X}_{ij} - \overline{\boldsymbol{X}}_i)(\boldsymbol{X}_{ij} - \overline{\boldsymbol{X}}_i)^\top\right\}, \quad \sum_{i=1}^k (n_i - 1)p,$$

respectively. Another example of (1) is k linear regression models such that $\boldsymbol{y}_1 = \boldsymbol{Z}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{y}_k = \boldsymbol{Z}_k\boldsymbol{\beta}_k + \boldsymbol{\varepsilon}_k$, where $\boldsymbol{y}_i$ is an $n_i \times 1$ vector of observations, $\boldsymbol{\beta}_i$ is a $p \times 1$ vector of regression coefficients, $\boldsymbol{Z}_i$ is an $n_i \times p$ matrix of explanatory variables and $\boldsymbol{\varepsilon}_i$ is an $n_i \times 1$ vector having $\mathcal{N}_{n_i}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_{n_i})$. In this case, $\boldsymbol{X}_i$, $\boldsymbol{V}_i$, S and n in (1) respectively correspond to

$$\widehat{\boldsymbol{\beta}}_i = (\boldsymbol{Z}_i^\top \boldsymbol{Z}_i)^{-1}\boldsymbol{Z}_i^\top \boldsymbol{y}_i, \quad (\boldsymbol{Z}_i^\top \boldsymbol{Z}_i)^{-1}, \quad \sum_{i=1}^k \|\boldsymbol{y}_i - \boldsymbol{Z}_i\widehat{\boldsymbol{\beta}}_i\|^2_{\boldsymbol{I}_{n_i}}, \quad \sum_{i=1}^k (n_i - p).$$