HIERARCHICAL AND EMPIRICAL BAYES MULTIVARIATE ESTIMATION

Malay Ghosh*, Department of Statistics, University of Florida

Abstract

This article reviews and unifies the hierarchical and empirical Bayes approach for estimating the multivariate normal mean. Both the ANOVA and the regression models are considered.

Introduction

Empirical and hierarchical Bayes methods are becoming increasingly popular in statistics, especially in the context of simultaneous estimation of several parameters. For example, agencies of the Federal Government have been involved in obtaining estimates of per capita income, unemployment rates, crop yields and so forth simultaneously for several state and local government areas. In such situations, quite often estimates of certain area means, or simultaneous estimates of several area means can be improved by incorporating information from similar neighboring areas. Examples of this type are especially suitable for empirical Bayes (EB) analysis. As described in Berger (1985), an EB scenario is one in which known relationships among the coordinates of a parameter vector, say $\underset{\sim}{\theta} = (\theta_1,\ldots,\theta_p)^T$ allow use of the data to estimate some features of the prior distribution. For example, one may have reason to believe that the θ_i's are iid from a prior $\pi_0(\lambda)$, where π_0 is structurally known except possibly for some unknown parameter λ. A *parametric empirical Bayes* (EB) procedure is one where λ is estimated from the marginal distribution of the observations.

Closely related to the EB procedure is the hierarchical Bayes (HB) procedure which models the prior distribution in stages. In the first stage, conditional on $\Lambda = \lambda$, θ_i's are iid with a prior $\pi_0(\lambda)$. In the second stage, a prior distribution (often improper) is assigned to Λ. This is an example of a two stage prior. The idea can be generalized to multistage priors, but that will not be pursued in this article.

It is apparent that both the EB and the HB procedures recognize the uncertainty in the prior information, but whereas the HB procedure models the uncertainty in the prior information by assigning a distribution (often *noninformative* or *improper*) to the prior parameters (usually called *hyperparameters*), the EB procedure attempts to estimate the unknown hyperparameters, typically by some classical method like the method of moments, method of maximum likelihood etc., and use the resulting estimated priors for inferential purposes. It turns out that the two methods can quite often

---

*This paper is dedicated to Professor D. Basu on the occasion of his 65th birthday. The research is partially supported by NSF Grant Numbers DMS 8701814 and DMS 8901334.