

1 Casella: An Introduction to Empirical Bayes Data Analysis

1.1 Introduction

- In a parametric approach to empirical Bayesian inference, the parametric family of prior distributions is specified. In a non-parametric approach, no prior is specified.
- Efron and Morris in the 1970s laid the groundwork for modern empirical Bayesian analysis.

1.2 Empirical Bayes Estimators for the Normal Case

Say we are trying to estimate the sampling distribution of a population from which we have drawn p random variables with cell means X_i , each with a normal distribution with different means and known variance. Let θ_i be the true means of each of these cells.:

- The Classical school argues that the observed mean of the data is the best estimator of the mean, that X_i best estimates θ_i .
- The Bayesian school assumes a normal prior, and computes a Bayesian estimate which is the mean of the normal posterior. This estimate is a weighted function of the observed and prior means, favoring one or the other depending on how large the sample and prior variance is. As sample variance grows larger, more weight is put upon the prior mean in determining this estimate.

The empirical Bayesian disagrees in some way with both of these approaches. The empirical Bayesian does not specify any values of the mean or variance for the posterior normal. Instead, these values are estimated using parameters from the data. This information is contained in the marginal distribution, $f(X_i)$, that is unconditional on θ_i , the parameter of interest.

By substituting in these estimates, a new empirical Bayesian estimator of θ_i is constructed. This new estimate takes into account information from each of the X_i s when approximating θ_i . This is good because of the **Stein effect**, which states that all estimates can be improved by using information from all coordinates when estimating each coordinate.

The Stein effect is illustrated in how the empirical Bayesian estimator has smaller Mean Squared Error in estimating θ_i than the classical estimator for all values of X_i .

1.3 Empirical Bayes Intuition

For five cells, say we wish to test if the cells have equivalent means. Let X_i denote the observed means of these cells and θ_i denote the true means of these

cells. We can construct H_0 such that if all means are equal, each cell's mean should be best estimated by the average mean of all cells. Alternatively, if the cells do not have equal means, the true mean of each individual cell would be best estimated by simply using the sample mean of that cell.

The empirical Bayesian approach allows for one to remain non-committal in anticipation of future evidence. It assumes the plausibility of both strategies, and uses the T-statistic calculated from testing H_0 to decide whether or not to weight X_i or \bar{X} more in estimating θ_i . For example, for larger T statistics, H_0 seems increasingly implausible, and so the Bayesian empirical estimate weights the \bar{X}_i estimation strategy less in its estimates.

1.4 More Properties of Empirical Bayesian Estimates

- One can use the observed distribution obtained by the Empirical Bayesian estimates of the Bayes estimates to try and determine a true prior distribution. For example, if one finds that a distribution of happy and sad feelings in a population is u-shaped, with the majority of people either very happy or very sad, one can use this to update priors in future studies.

2 Gelman: Objections to Bayesian Statistics

2.1 A Bayesian's attempt to see the other side

According to Gelman:

- **Bayesian inference** represents statistical estimation as the conditional distribution of parameters and unobserved data, given observed data.
- **Bayesian statisticians** are those who would apply Bayesian methods to *all* problems.
- "Anti-Bayesians" avoid using Bayesian methods and object to their use by others

2.2 Overview of the objections

Gelman outlines a few main objections:

1. Bayesian inference is an "automatic inference machine" as opposed to a more careful process-specific approach. The goal of inference should be to provide reasonable answers with a minimal number of assumptions.
2. Bayesian analysis relies too much on subjectivity, as there exists few ways to properly assess the validity of subjective knowledge.
3. The "key" topics of modern statistics are being shirked in favor of the computational Bayes approach.

2.3 Further possible objections

Writing in the voice of the "Anti-Bayesian", Gelman argues:

- There exist no objective best practices to the subjective process of prior selection.
- Making Bayesian analysis accessible to natural scientists and political scientists is only furthering its misuse.
- Unbiasedness and convergence are more important than the main results of Bayesian analysis.
- Multilevel and hierarchical models only work to confirm existing biases in the Bayesian's initial assumptions.
- Assumptions about "exchangeability" in Bayesian work on political science don't work