Example 2(a): (One-way classification, 1 observation per group)

\[ X_i | \mu_i \overset{\text{indep}}{\sim} N(\mu_i, \sigma^2), \quad i = 1, \ldots, m \]

\[ \mu_i \overset{\text{iid}}{\sim} N(\phi, \tau^2), \quad \sigma^2, \tau^2 \text{ known.} \]

Then it can be shown

\[ q(x|\phi) = [2\pi(\sigma^2 + \tau^2)]^{-m/2} e^{-\frac{1}{2(\sigma^2 + \tau^2)} \sum_{i=1}^{m} (x_i - \phi)^2} \]

Hence the MLE of \( \phi \) is clearly \( \hat{\phi} = \bar{X} \).

The empirical Bayes estimator turns out to be

\[ E[\mu_i|x, \hat{\phi}] = \frac{\tau^2}{\sigma^2 + \tau^2} x_i + \frac{\sigma^2}{\sigma^2 + \tau^2} \bar{X}. \]
Example 2(b):

If we have a one-way classification with $m$ groups and $n$ observations per group, the previous example extends to

\[ X_{ij} | \mu_i \overset{\text{indep}}{\sim} N(\mu_i, \sigma^2), \quad i = 1, \ldots, m, \; j = 1, \ldots, n \]

\[ \mu_i \overset{\text{iid}}{\sim} N(\phi, \tau^2), \quad \sigma^2, \tau^2 \text{ known}. \]

Then note that

\[ \bar{X}_i \sim N\left(\phi, \frac{\sigma^2}{n} + \tau^2\right) \]

Hence the empirical Bayes estimate of $\mu_i (i = 1, \ldots, m)$ is

\[ \hat{\mu}_i = \frac{n\tau^2}{\sigma^2 + n\tau^2} \bar{X}_i + \frac{\sigma^2}{\sigma^2 + n\tau^2} \bar{X} \]
If $\tau^2$ is unknown, note that $\frac{m-3}{\sum(\bar{X}_i - \bar{X})^2}$ is an unbiased estimator of $\frac{1}{\sigma^2 + n\tau^2}$, so we can use

$$\hat{\mu}_i = \left[ 1 - \frac{(m-3)\sigma^2}{\sum(\bar{X}_i - \bar{X})^2} \right] \bar{X}_i + \left[ \frac{(m-3)\sigma^2}{\sum(\bar{X}_i - \bar{X})^2} \right] \bar{X}$$

as the empirical Bayes estimator.

On the other hand, if $\sigma^2$ is unknown, we can use

$$\hat{\mu}_i = \left[ \frac{(m-3)n\tau^2}{\sum(\bar{X}_i - \bar{X})^2} \right] \bar{X}_i + \left[ 1 - \frac{(m-3)n\tau^2}{\sum(\bar{X}_i - \bar{X})^2} \right] \bar{X}$$

as the empirical Bayes estimator.
Hierarchical Bayes (HB) and Empirical Bayes (EB) estimators both typically involve shrinkage.

Some Bayesians feel EB is “less honest” since EB plugs in estimates of the hyperparameters without accounting for the variability associated with the estimate.

HB places a distribution on the hyperparameters, and thus models the uncertainty in the hyperparameter values.

See HB/EB Comparison for the Italian Marriage Data example on course web page.