Relationship between empirical predictors (Equal within cluster variances)

Silvina San Martino

The predictor under each model can be written as \( \hat{T}_i = \bar{Y} + k_{\text{model}} (\bar{Y} - \bar{Y}) \).

Alternatively, we the predictors can be expressed as \( \hat{T}_i = a_{\text{model}} \bar{Y} + b_{\text{model}} \bar{Y} \), a form that has been used in the simulation studies. We describe the coefficients and their possible estimators here. Notice that \( a_{\text{model}} = 1 - k_{\text{model}} \) and \( b_{\text{model}} = k_{\text{model}} \).

1. Mixed Model (ME)

For the ME model,

\[
\hat{k}_{\text{ME}} = \frac{m \sigma^2}{m \sigma^2 + \sigma_{v}^2 + \sigma_{r}^2}.
\]

Since \( E(MSE) = \sigma_{v}^2 + \sigma_{r}^2 \) and \( E(MSB) = m \sigma^2 + \sigma_{v}^2 + \sigma_{r}^2 \), we may estimate \( \hat{k}_{\text{ME}} \) by

\[
\hat{k}_{\text{ME}} = \max \left( 0, \frac{\text{MSB} - \text{MSE}}{\text{MSB}} \right).
\]

For the simulation studies, \( \hat{a}_{\text{ME}} = 1 - \hat{k}_{\text{ME}} \) and \( \hat{b}_{\text{ME}} = \hat{k}_{\text{ME}} \).

2. Scott and Smith’s Model (SS)

For the SS model

\[
k_{\text{SS}} = f + (1 - f) \hat{k}_{\text{ME}} = \frac{m \sigma^2 + f \left( \sigma_{v}^2 + \sigma_{r}^2 \right)}{m \sigma^2 + \sigma_{v}^2 + \sigma_{r}^2}.
\]

Since \( E(MSE) = \sigma_{v}^2 + \sigma_{r}^2 \) and \( E(MSB) = m \sigma^2 + \sigma_{v}^2 + \sigma_{r}^2 \), we may estimate \( k_{\text{SS}} \) by

\[
\hat{k}_{\text{SS1}} = f + (1 - f) \hat{k}_{\text{ME}}.
\]

where \( \hat{k}_{\text{ME}} = \max \left( 0, \frac{\text{MSB} - \text{MSE}}{\text{MSB}} \right) \). For simulation studies, \( \hat{a}_{\text{SS1}} = (1 - f) \left( 1 - \hat{k}_{\text{ME}} \right) \) and \( \hat{b}_{\text{SS1}} = f + (1 - f) \hat{k}_{\text{ME}} \).

We can also express

\[
k_{\text{SS}} = k_{\text{ME}} + f \left( 1 - k_{\text{ME}} \right) = \frac{m \sigma^2 + f \left( \sigma_{v}^2 + \sigma_{r}^2 \right)}{m \sigma^2 + \sigma_{v}^2 + \sigma_{r}^2}.
\]

Since \( E(MSE) = \sigma_{v}^2 + \sigma_{r}^2 \) and \( E(MSB) = m \sigma^2 + \sigma_{v}^2 + \sigma_{r}^2 \), we may estimate \( k_{\text{SS}} \) by

\[
\hat{k}_{\text{SS2}} = \max \left( 0, \frac{\text{MSB} - (1 - f) \text{MSE}}{\text{MSB}} \right).
\]

For simulation studies, \( \hat{a}_{\text{SS2}} = (1 - \hat{k}_{\text{SS2}}) \) and \( \hat{b}_{\text{SS2}} = \hat{k}_{\text{SS2}} \).

3. Random Permutation Model (RP)
For the RP model

$$k_{rp} = k^* + f(k^*_r - k^*)$$

where $k^* = \frac{m\sigma^2}{m\sigma^2 + (\sigma^2_r + \sigma^2_i)}$ and $k^*_r = \frac{m\sigma^2 + \sigma^2}{m\sigma^2 + (\sigma^2_r + \sigma^2_i)}$. Now $E(MSE) = \sigma^2 + \sigma^2$ and

$$E(MSB) = m\sigma^2 + \sigma^2 + \sigma^2$$

As a result,

$$k^* = \frac{m\sigma^2 - f\sigma^2}{m\sigma^2 + (1-f)\sigma^2 + \sigma^2_i}$$

and

$$k^*_r = \frac{m\sigma^2 + (1-f)\sigma^2}{m\sigma^2 + (1-f)\sigma^2 + \sigma^2_i}.$$ This implies that minimum value of $k^*$ is $\frac{-f}{1-f} = \frac{-m}{M-m}$, and $k^*_r$ must be positive. We may estimate $k^*$ and $k^*_r$ by

$$\hat{k}^* = \max\left(-\frac{f}{1-f}, \frac{MSB - MSE}{MSB}\right),$$

and

$$\hat{k}^*_r = \max\left(0, \frac{MSB - \sigma^2}{MSB}\right),$$

resulting in

$$\hat{k}_{rp} = \hat{k}^* + f(\hat{k}^*_r - \hat{k}^*).$$

For simulation studies, $\hat{a}_{rp} = 1 - (1-f)\hat{k}^* - f\hat{k}^*_r$ and $\hat{b}_{rp} = (1-f)\hat{k}^* + f\hat{k}^*_r$.

We can also express

$$k_{rp} = k^* + f(k^*_r - k^*)$$

$$= \frac{m\sigma^2}{m\sigma^2 + \sigma^2 + \sigma^2_i} + f\left(\frac{m\sigma^2 + \sigma^2}{m\sigma^2 + \sigma^2 + \sigma^2_i} - \frac{m\sigma^2}{m\sigma^2 + \sigma^2 + \sigma^2_i}\right)$$

$$= \frac{m\sigma^2 + \sigma^2}{m\sigma^2 + \sigma^2 + \sigma^2_i} + f\left(\frac{m\sigma^2 + \sigma^2}{m\sigma^2 + \sigma^2 + \sigma^2_i} - \frac{m\sigma^2}{m\sigma^2 + \sigma^2 + \sigma^2_i}\right)$$

$$= \frac{(m\sigma^2 + \sigma^2) - (\sigma^2_r + \sigma^2)}{m\sigma^2 + \sigma^2 + \sigma^2_i} + f\sigma^2_i$$

Now $\sigma^2 = \rho(\sigma^2_r + \sigma^2_i)$ so that $f\sigma^2 = f\rho(\sigma^2_r + \sigma^2_i)$. Then

$$k_{rp} = \frac{m\sigma^2 + \sigma^2 + \sigma^2_i}{m\sigma^2 + \sigma^2 + \sigma^2_i} - (1-f\rho)(\sigma^2_r + \sigma^2)$$

which may be estimated by

$$\hat{k}_{rp} = \max\left(0, \frac{MSB - (1-f\rho)MSE}{MSB}\right).$$

For simulation studies, $\hat{a}_{rp} = 1 - \hat{k}_{rp}$ and $\hat{b}_{rp} = \hat{k}_{rp}$. This may help to clarify the relationship between the empirical predictors.