Finding correlation for acmPRS

November 9, 2016

Goals:

- Find $Var(\text{acmPRS})$
- Find $Cov(Y, \text{acmPRS})$

0.1 Finding Var(acmPRS):

Define

$$\hat{S}_i = \sum_{j=1}^{m_i} \hat{\beta}\mathbf{G_j}$$

$$\hat{S}_{acm} = \sum_{i=1}^{c} n_i r_i \sqrt{\frac{h_i}{h_1}} \hat{S}_i$$

Theorem 1. $E[\hat{S}_{acm}] = 0$

Proof.

$$\begin{aligned} E[\hat{S}_{acm}] &= E[\sum_{i=1}^{c} n_i r_i \sqrt{\frac{h_i}{h_1}} \hat{S}_i] \\ &= \sum_{i=1}^{c} E[n_i r_i \sqrt{\frac{h_i}{h_1}} \hat{S}_i] \\ &= \sum_{i=1}^{c} n_i r_i \sqrt{\frac{h_i}{h_1}} E[\hat{S}_i] \end{aligned}$$

$$\because E[\hat{S}_i] = 0$$

$$\therefore E[\hat{S}_{acm}] = 0$$

□

Theorem 2. $Var(\hat{S}_{acm}) = \sum_{i=1}^{c} \frac{m_i n_i^2 r_i^2 h_i}{h_1} Var(\hat{\beta}_i) + \sum_{i \neq j}^{c} n_i n_j r_i r_j \sqrt{\frac{h_i h_j}{h_1^2}} Cov(\hat{S}_i, \hat{S}_j)]$

Proof.

$$Var(\sum_{i=1}^{c} n_i r_i \sqrt{\frac{h_i}{h_1}} \hat{S}_i) = \sum_{i=1}^{c} Var\left(n_i r_i \sqrt{\frac{h_i}{h_1}} \hat{S}_i\right) + \sum_{i \neq j} Cov\left(n_i r_i \sqrt{\frac{h_i}{h_1}} \hat{S}_i, n_j r_j \sqrt{\frac{h_j}{h_1}} \hat{S}_j\right)$$

$$\begin{aligned} \sum_{i=1}^{c} Var\left(n_i r_i \sqrt{\frac{h_i}{h_1}} \hat{S}_i\right) &= \sum_{i=1}^{c} n_i r_i \sqrt{\frac{h_i}{h_1}} Var(\hat{S}_i) \\ &= \sum_{i=1}^{c} n_i r_i \sqrt{\frac{h_i}{h_1}} m_i Var(\hat{\beta}_i) \end{aligned} \tag{1}$$

$$\begin{aligned} \sum_{i \neq j} Cov\left(n_i r_i \sqrt{\frac{h_i}{h_1}} \hat{S}_i, n_j r_j \sqrt{\frac{h_j}{h_1}} \hat{S}_j\right) &= \textstyle\sum_{i \neq j} E\left[\left(n_i r_i \sqrt{\frac{h_i}{h_1}} \hat{S}_i\right), \left(n_j r_j \sqrt{\frac{h_j}{h_1}} \hat{S}_j\right)\right] + \cancelto{0}{E\left[\textstyle\sum_{i=1}^{c} n_i r_i \sqrt{\frac{h_i}{h_1}}] \hat{S}_i\right]} \cancelto{0}{E\left[\textstyle\sum_{j=1}^{c} n_j r_j \sqrt{\frac{h_j}{h_1}} \hat{S}_j\right]} \\ &= \sum_{i \neq j} n_i n_j r_i r_j \sqrt{\frac{h_i h_j}{h_1^2}} E[\hat{S}_i \hat{S}_j] \\ &= \sum_{i \neq j} n_i n_j r_i r_j \sqrt{\frac{h_i h_j}{h_1^2}} Cov\left(\hat{S}_i, \hat{S}_j\right) \end{aligned} \tag{2}$$

The above is because

$$\begin{aligned} Cov(\hat{S}_i, \hat{S}_j) &= E\left[\hat{S}_i \hat{S}_j\right] + \cancelto{0}{E[\hat{S}_i]} + \cancelto{0}{E[\hat{S}_j]} \\ &= E\left[\hat{S}_i \hat{S}_j\right] \end{aligned}$$

Combining (1) and (2):

$$Var(\hat{S}_{acm}) = \sum_{i=1}^{c} \frac{m_i n_i^2 r_i^2 h_i}{h_1} Var(\hat{\beta}_i) + \sum_{i \neq j}^{c} n_i n_j r_i r_j \sqrt{\frac{h_i h_j}{h_1^2}} Cov(\hat{S}_i, \hat{S}_j)]$$

□

Theorem 3. $Cov(Y, acmPRS) = \frac{1}{c}\sum_{i=1}^{c} n_i r_i \sqrt{\frac{h_i}{h_1}} m_i Cov(\hat{\beta}_i, \beta_Y)$

Proof.

$$\begin{aligned}\mathrm{Cov}(Y, \mathrm{acmPRS}) &= E[Y\mathrm{acmPRS}] + \cancelto{0}{E[Y]} + \cancelto{0}{E[\mathrm{acmPRS}]} \\ &= E\left[\left(\sum_{j=1}^{m} \beta_{jY} G_i\right)\left(\sum_{i=1}^{c} n_i r_i \sqrt{\frac{h_i}{h_1}} \sum_{j=1}^{m} \hat{\beta}_{ji} G_j\right)\right]\end{aligned}$$

I'm really not sure if I'm allowed to do this.

$$E\left[\left(\sum_{j=1}^{m} \beta_{jY} G_i\right)\left(\sum_{i=1}^{c} n_i r_i \sqrt{\frac{h_i}{h_1}} \sum_{j=1}^{m} \hat{\beta}_{ji} G_j\right)\right] = E\left[\sum_{i=1}^{c} n_i r_i \sqrt{\frac{h_i}{h_1}} \left[\sum^{m_i} \beta_j G_j\right] \frac{1}{c}\left[\sum^{m_Y} \beta_{jY} G_j\right]\right]$$

Here I've brought the Y score inside the summation with a $\frac{1}{c}$ coefficient.

$$= \frac{1}{c}\sum_{i=1}^{c}\left[n_i r_i \sqrt{\frac{h_i}{h_1}} E\left[\sum^{m_i} \beta_j G_j \sum^{m_Y} \beta_{jY} G_j\right]\right]$$

And since by equation (7) in power and predicitve accuracy: $E\left[\sum^{m_i} \beta_j G_j \sum^{m_Y} \beta_{jY} G_j\right] = m\mathrm{Cov}(\hat{\beta}_i, \beta_Y)$

Then

$$Cov(Y, \mathrm{acmPRS}) = \frac{1}{c}\sum_{i=1}^{c} n_i r_i \sqrt{\frac{h_i}{h_1}} m_i Cov(\hat{\beta}_i, \beta_Y)$$

□

Theorem 4. $R^2_{Y,acmPRS} = \frac{\frac{1}{c}\sum_{i=1}^{c} n_i r_i \sqrt{\frac{h_i}{h_1}} m_i Cov(\hat{\beta}_i, \beta_Y)}{\left(\sum_{i=1}^{c} \frac{m_i n_i^2 r_i^2 h_i}{h_1} Var(\hat{\beta}_i) + \sum_{i \neq j}^{c} n_i n_j r_i r_j \sqrt{\frac{h_i h_j}{h_1^2}} Cov(\hat{S}_i, \hat{S}_j)]\right)\left(\sigma_g^2 + \sigma_e^2\right)}$