Week 10

4.8) $D_j = \sum_{i \in B} (y_i - \bar{y}_B)^2$

$D_j^* = \sum_{i \in A} (y_i - \bar{y}_A)^2 + \sum_{i \in B} (y_i - \bar{y}_B)^2$

$D_j - D_j^* = \sum_{i} (y_i - \bar{y})^2 - \sum_{i \in A} (y_i - \bar{y}_A)^2 - \sum_{i \in B} (y_i - \bar{y}_B)^2$

$= \underbrace{\sum_{i \in A} (y_i - \bar{y})^2 - \sum_{i \in A} (y_i - \bar{y}_A)^2}_{\geq 0} + \underbrace{\sum_{i \in B} (y_i - \bar{y})^2 - \sum_{i \in B} (y_i - \bar{y}_B)^2}_{\geq 0}$

$= 0$ when $\bar{y} = \bar{y}_B = \bar{y}_A$

5.11)

$D_j = -n_j (p_{j1} \log p_{j1} + p_{j0} \log p_{j0})$

$D_j^* = -n_A (p_{A1} \log p_{A1} + p_{A0} \log p_{A0}) - n_B (p_{B1} \log p_{B1} + p_{B0} \log p_{B0})$

$D_j^* - D_j$