Homework 5

Due Thursday, 3/18/20 on Canvas.

1. Consider the one-way ANOVA model from class: $Y_{ij} = \mu_i + \epsilon_{ij}$ for $i = 1, \ldots, r$ and $j = 1, \ldots, n_i$, where $\text{Cov}\left(\epsilon_{ij}, \epsilon_{i'j'}\right) = \sigma^2 \delta_{ii'} \delta_{jj'}$. Let

$$\begin{aligned}
\boldsymbol{Y} &= \left(Y_{11}, \ldots, Y_{1n_1}, \ldots, Y_{r1}, \ldots, Y_{rn_r}\right)^T \\
\boldsymbol{\epsilon} &= \left(\epsilon_{11}, \ldots, \epsilon_{1n_1}, \ldots, \epsilon_{r1}, \ldots, \epsilon_{rn_r}\right)^T \\
\beta &= \left(\mu_1, \ldots, \mu_r\right)^T .
\end{aligned}$$

   (a) What is the design matrix $\boldsymbol{X}$?

   (b) Compute the corresponding hat matrix $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}$.

   (c) Show that $\boldsymbol{HY} = \left(\bar{Y}_{1\cdot} 1_{n_1}, \ldots, \bar{Y}_{r\cdot} 1_{n_r}\right)^T$.

   (d) Let $n_T = \sum_{i=1}^{r} n_i$, $\boldsymbol{L} = n_T^{-1} 1_{n_T} 1_{n_T}^T$ and $\bar{Y}_{\cdot\cdot} = n_T^{-1} \sum_{i=1}^{r} \sum_{j=1}^{n_i} Y_{ij}$.

      i. Show that

$$SSTR = \sum_{i=1}^{r} n_i \left(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot}\right)^2 = \boldsymbol{Y}^T \left(\boldsymbol{H} - \boldsymbol{L}\right) \boldsymbol{Y},$$

      where $\boldsymbol{H} - \boldsymbol{L}$ is itself an orthogonal projection operator (i.e. it is symmetric and idempotent).

      ii. What is the rank of $\boldsymbol{H} - \boldsymbol{L}$?

      iii. Show that $\text{Im}\left(\boldsymbol{H} - \boldsymbol{L}\right)$ is orthogonal to $\text{Im}\left(I_n - \boldsymbol{H}\right)$.

      iv. If $\boldsymbol{Y}$ is normally distributed, show that $SSTR$ is independent of SSE.

2. Let $Z \in \mathbb{R}^n$ be a random vector with $\mathbb{E}(Z) = \mu$ and $\text{Var}(Z) = V$.

   (a) Show that for any non-random matrix $A \in \mathbb{R}^{n\times n}$, $\mathbb{E}\left(Z^T A Z\right) = \text{Tr}\left(AV\right) + \mu^T A \mu$.

   (b) Using the notation from problem 1, show that $\mathbb{E}(MSE) = \sigma^2$ and $\mathbb{E}(MSTR) = \sigma^2 + \frac{\sum_{i=1}^{r} n_i(\mu_i - \mu_\cdot)^2}{r-1}$, where $\mu_\cdot = \frac{\sum_{i=1}^{r} n_i \mu_i}{\sum_{i=1}^{r} n_i}$.