Stat873 HW09

1. Suppose $Y \sim N(X\beta, \sigma^2\Sigma)$ where $X = (X_I, X_{II}) \in R^{n\times p}$ with $\text{rank}(X) = r$ and $X_I \in R^{n\times p_1}$ with $\text{rank}(X_I) = r_1$.

   (1) Show that $(\Sigma^{-1/2}X)(\Sigma^{-1/2}X)^+(\Sigma^{-1/2}X_I)(\Sigma^{-1/2}X_I)^+ = (\Sigma^{-1/2}X_I)(\Sigma^{-1/2}X_I)^+$

   Note that $X_I = (X_I, X_{II})\begin{pmatrix} I \\ 0 \end{pmatrix} = XH$ where $H = \begin{pmatrix} I \\ 0 \end{pmatrix}$. So

$$\begin{aligned} & (\Sigma^{-1/2}X)(\Sigma^{-1/2}X)^+(\Sigma^{-1/2}X_I)(\Sigma^{-1/2}X_I)^+ \\ = \; & (\Sigma^{-1/2}X)(\Sigma^{-1/2}X)^+(\Sigma^{-1/2}XH)(\Sigma^{-1/2}X_I)^+ = (\Sigma^{-1/2}XH)(\Sigma^{-1/2}X_I)^+ \\ = \; & (\Sigma^{-1/2}X_I)(\Sigma^{-1/2}X_I)^+. \end{aligned}$$

   (2) Show that $(\Sigma^{-1/2}X_I)(\Sigma^{-1/2}X_I)^+(\Sigma^{-1/2}X)(\Sigma^{-1/2}X)^+ = (\Sigma^{-1/2}X_I)(\Sigma^{-1/2}X_I)^+$

   By (1), $(\Sigma^{-1/2}X)(\Sigma^{-1/2}X)^+(\Sigma^{-1/2}X_I)(\Sigma^{-1/2}X_I)^+ = (\Sigma^{-1/2}X_I)(\Sigma^{-1/2}X_I)^+$.
   So $\left[(\Sigma^{-1/2}X)(\Sigma^{-1/2}X)^+(\Sigma^{-1/2}X_I)(\Sigma^{-1/2}X_I)^+\right]' = \left[(\Sigma^{-1/2}X_I)(\Sigma^{-1/2}X_I)^+\right]'$.
   Thus $(\Sigma^{-1/2}X_I)(\Sigma^{-1/2}X_I)^+(\Sigma^{-1/2}X)(\Sigma^{-1/2}X)^+ = (\Sigma^{-1/2}X_I)(\Sigma^{-1/2}X_I)^+$.

   (3) Let $A = \frac{(\Sigma^{-1/2}X)(\Sigma^{-1/2}X)^+ - (\Sigma^{-1/2}X_I)(\Sigma^{-1/2}X_I)^+}{\sigma^2}$.
   Find the distribution for $Z^2 = \left[\Sigma^{-1/2}(Y - X\beta)\right]' A \left[\Sigma^{-1/2}(Y - X\beta)\right]$.

   $Y \sin N(X\beta, \sigma^2\Sigma)$ implies that $\Sigma^{-1/2}(Y - X\beta) \sim N(0, \sigma^2 I_n)$.
   But by (1) and (2) $A(\sigma^2 I_n)A = A = A'$. So Z^2 has a χ^2-distribution.
   With $0'A0 = 0$ and $\text{tr}(A\sigma^2 I_n) = \text{rank}(X) - \text{rank}(X_I) = r - r_1$, $Z^2 \sim \chi^2(r - r_1)$.

2. File mydata.dat in HW07 contains variable y and character variable Sid that identifies 4 levels of a factor in one-way ANOVA. Find ANOVA table for this ANOVA model.
   Keep 4 digits after decimal point.

Source	DF	SS	MS	F	$Pr > F$
Model	3	2782.2394	927.4131	80.83	< 0.0001
Error	18	206.5333	11.4741		
C.Total	21	2988.7727			

3. File T6-10.dat in HW08 contains variables y, x_1, x_2 and a character variable type.
   Find ANOVA table for regression $y = \beta_1 x_1 + \beta_2 x_2 + \epsilon$. Keep 4 digits after decimal point.

Source	DF	SS	MS	F	$Pr > F$
Model	2	6951.4245	3475.7122	118.39	< 0.0001
Error	57	1673.4713	29.3592		
U.Total	59	8624.8958			