Stat 511 HW 4 Solution

1. (a) Set $\underset{\sim}{W} = [W_1 \ W_2 \ W_3]'$, $Y_1 = W_1 + W_2 - W_3$, $Y_2 = W_1 + W_2 + W_3$, $\underset{\sim}{Y} = [Y_1, Y_2]'$.

Then, $\underset{\sim}{Y} = A\underset{\sim}{W}$, where $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$.

Therefore, the distribution of $\underset{\sim}{Y}$ is $N(A\underset{\sim}{\mu}, A\Sigma A')$, where

$$A\underset{\sim}{\mu} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

and $$A\Sigma A' = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 & 1 \\ -1 & 4 & 2 \\ 1 & 2 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 16 \end{bmatrix}$$

(b) Let $U \equiv \dfrac{W_1^2 + W_2^2 + W_3^2 + 2W_1W_2 + 2W_1W_3 + 2W_2W_3}{16}$, then we have

$$U = \frac{(W_1+W_2+W_3)^2}{16} = \left(\frac{Y_2}{4}\right)^2$$

By part (a), $Y_2 \sim N(0, 16)$, therefore $\frac{Y_2}{4} \sim N(0, 1)$.

This implies that $U = \left(\frac{Y_2}{4}\right)^2 \sim \chi_1^2$

(C) $$\frac{4W_1^2+4W_2^2+4W_3^2+8W_1W_2-8W_1W_3-8W_2W_3}{W_1^2+W_2^2+W_3^2+2W_1W_2+2W_1W_3+2W_2W_3}$$

$$=\frac{4(W_1+W_2-W_3)^2}{(W_1+W_2+W_3)^2}$$

$$=\frac{4Y_1^2}{Y_2^2}=\frac{\left(\frac{Y_1}{2}\right)^2}{\left(\frac{Y_2}{4}\right)^2}$$, where $Y_1=W_1+W_2-W_3$, $Y_2=W_1+W_2+W_3$.

By Part (a) , $Y_1\sim N(-2,4)$, $Y_2\sim N(0,16)$. therefore,

$$\frac{Y_1}{2}\sim N(-1,1),\quad \frac{Y_2}{4}\sim N(0,1)$$

$$\therefore\ \left(\frac{Y_1}{2}\right)^2\sim \chi_1^2(1),\quad \left(\frac{Y_2}{4}\right)^2\sim \chi_1^2(0) \qquad (*)$$

Because $\mathrm{cov}(Y_1,Y_2)=0$, and Y_1, Y_2 normally distributed,

we know that Y_1, Y_2 are independent.

Thus, $\left(\frac{Y_1}{2}\right)^2$ and $\left(\frac{Y_2}{4}\right)^2$ are independent.

Also by $(*)$, $\frac{\left(\frac{Y_1}{2}\right)^2}{\left(\frac{Y_2}{4}\right)^2}$ has $F_{1,1}(1)$ distribution.

2.(a) Let T be $\dfrac{\underset{\sim}{c}'\hat{\beta} - \underset{\sim}{d}}{\sqrt{\hat{\sigma}^2 \underset{\sim}{c}'(X'X)^- \underset{\sim}{c}}}$,

i.e., $T = \dfrac{\underset{\sim}{c}'\hat{\beta} - \underset{\sim}{d}}{\sqrt{\hat{\sigma}^2 \underset{\sim}{c}'(X'X)^- \underset{\sim}{c}}} = \dfrac{\dfrac{\underset{\sim}{c}'\hat{\beta} - \underset{\sim}{d}}{\sqrt{\sigma^2 \underset{\sim}{c}'(X'X)^- \underset{\sim}{c}}}}{\sqrt{\left[(n-\text{rank}(X))\dfrac{\hat{\sigma}^2}{\sigma^2}\right] \Big/ (n-\text{rank}(X))}}$

$$= \frac{w}{\sqrt{\frac{u}{n-\text{rank}(X)}}},$$

where $w = \dfrac{\underset{\sim}{c}'\hat{\beta} - \underset{\sim}{d}}{\sqrt{\sigma^2 \underset{\sim}{c}'(X'X)^- \underset{\sim}{c}}}$ $\quad u = (n-\text{rank}(X))\dfrac{\hat{\sigma}^2}{\sigma^2}$.

Normal Theory Gauss-Markov model says $\underset{\sim}{Y} \sim N(X\underset{\sim}{\beta}, \sigma^2 I)$, where N stands for mutivariate normal distribution, thus,

$\underset{\sim}{c}'\hat{\beta} - \underset{\sim}{d} = \underset{\sim}{c}'(X'X)^- X'\underset{\sim}{Y} - \underset{\sim}{d} \sim N(\underset{\sim}{c}'\underset{\sim}{\beta} - \underset{\sim}{d}, V)$,

where $V = \sigma^2 \underset{\sim}{c}'(X'X)^- \underset{\sim}{c}$ by the Lecture notes.

This implies that $w = \dfrac{\underset{\sim}{c}'\hat{\beta} - \underset{\sim}{d}}{\sqrt{\sigma^2 \underset{\sim}{c}'(X'X)^- \underset{\sim}{c}}} \sim N\left(\dfrac{\underset{\sim}{c}'\underset{\sim}{\beta} - \underset{\sim}{d}}{\sqrt{\sigma^2 \underset{\sim}{c}'(X'X)^- \underset{\sim}{c}}}, 1\right)$.

Next, by the lecture notes, we have

$$u = (n-\text{rank}(X))\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-\text{rank}(X)}.$$

Next, we want to show w and u are independent.

Because $\underset{\sim}{c}'\hat{\underset{\sim}{\beta}} = \underset{\sim}{c}'(X'X)^- X'\underset{\sim}{Y} = \underset{\sim}{a}'X(X'X)^- X'\underset{\sim}{Y} = \underset{\sim}{a}' P_X \underset{\sim}{Y}$, w is a function of $P_X \underset{\sim}{Y}$.

(a) (cont'd) ⑤

Because
$$u = (n-\text{rank}(X))\frac{\hat{\sigma}^2}{\sigma^2}$$
$$= (n-\text{rank}(X))\frac{1}{(n-\text{rank}(X))}\, \underset{\sim}{Y}'\frac{(I-P_X)}{\sigma^2}\underset{\sim}{Y}$$
$$= \frac{\underset{\sim}{Y}'(I-P_X)\underset{\sim}{Y}}{\sigma^2} = \frac{[(I-P_X)\underset{\sim}{Y}]'(I-P_X)\underset{\sim}{Y}}{\sigma^2},$$

we know u is a function of $(I-P_X)\underset{\sim}{Y}$.

Because $P_X\underset{\sim}{Y}$ are $(I-P_X)\underset{\sim}{Y}$ are both multivariate normally distributed,

and $\text{cov}(P_X\underset{\sim}{Y}, (I-P_X)\underset{\sim}{Y}) = P_X\,\text{var}(\underset{\sim}{Y})(I-P_X) = P_X\,\sigma^2 I\,(I-P_X) = 0$,

we get w and u are independent.

Therefore, by the definition of noncentral t-distribution,

$$T = \frac{w}{\sqrt{\frac{u}{n-\text{rank}(X)}}}$$ has noncentral t distribution with

$$ncp = \frac{\underset{\sim}{c}'\underset{\sim}{\beta} - d}{\sqrt{\sigma^2\underset{\sim}{c}'(X'X)^-\underset{\sim}{c}}}$$ and $df = n-\text{rank}(X)$.

(b) A test statistic for testing the null hypothesis

$$H_0 : \underset{\sim}{c}'\underset{\sim}{\beta} = \underset{\sim}{d} \quad \text{v.s.} \quad H_A : \underset{\sim}{c}'\underset{\sim}{\beta} \neq \underset{\sim}{d} \qquad \forall \text{ fixed } \underset{\sim}{d} \in \mathbb{R}$$

is $$T = \frac{\underset{\sim}{c}'\hat{\underset{\sim}{\beta}} - \underset{\sim}{d}}{\sqrt{\hat{\sigma}^2 \underset{\sim}{c}'(X'X)^{-}\underset{\sim}{c}}}$$

(c) Under H_0, $\underset{\sim}{c}'\beta - \underset{\sim}{d} = \underset{\sim}{0}$, therefore $ncp = 0$. This means that the test statistic T will follow a central-t distribution with $n - \text{rank}(X)$ degrees of freedom.

Therefore, using significance level α, we would reject the null hypothesis if $|T| \geq t_{n-\text{rank}(X),\, 1-\frac{\alpha}{2}}$

(d) Let $t = t_{n-\text{rank}(X),\, 1-\frac{\alpha}{2}}$ and $se = \sqrt{\hat{\sigma}^2 \underset{\sim}{c}'(X'X)^- \underset{\sim}{c}}$

By part (a),

$$1-\alpha = P\left(-t \le \frac{\underset{\sim}{c}'\hat{\beta} - \underset{\sim}{c}'\beta}{se} \le t\right)$$

$$= P(-t \cdot se \le \underset{\sim}{c}'\hat{\beta} - \underset{\sim}{c}'\beta \le t \cdot se)$$

$$= P(-\underset{\sim}{c}'\hat{\beta} - t \cdot se \le -\underset{\sim}{c}'\beta \le -\underset{\sim}{c}'\hat{\beta} + t \cdot se)$$

$$= P(\underset{\sim}{c}'\hat{\beta} + t \cdot se \ge \underset{\sim}{c}'\beta \ge \underset{\sim}{c}'\hat{\beta} - t \cdot se)$$

$$= P(\underset{\sim}{c}'\hat{\beta} - t \cdot se \le \underset{\sim}{c}'\beta \le \underset{\sim}{c}'\hat{\beta} + t \cdot se)$$

Thus, a $100(1-\alpha)\%$ confidence interval for $\underset{\sim}{c}'\beta$ is

$$(\underset{\sim}{c}'\hat{\beta} - t \cdot se,\ \underset{\sim}{c}'\hat{\beta} + t \cdot se)$$

3. Proof: Since A is symmetric, by Result 1.11 (ii) and (v),

we have $\text{rank}(A) =$ number of nonzero eigenvalues,

$$\text{trace}(A) = \sum_{i=1}^{k} \lambda_i .$$

In addition, since A is also idempotent, by Result 1.11 (vii),

$$\lambda_i = 0 \text{ or } 1 , \quad \forall\, i = 1, \cdots, k.$$

Thus, $\text{rank}(A) =$ number of eigenvalues that are 1's,

$\text{trace}(A) = \sum_{i=1}^{k} \lambda_i =$ sum of eigenvalues that are 1's

$=$ number of eigenvalues that are 1's.

$\therefore \quad \text{rank}(A) = \text{trace}(A)$