Assignment 7

(1) (a) False. $y = X\beta + \varepsilon$, and in general, $y \neq \hat{y}$

(b) True. $X\hat{\beta} + \varepsilon = \hat{y} + (y - \hat{y}) = y$

(c) False. Again, $y \neq \hat{y}$ in general

(d) False. $y = X\beta + \varepsilon$, and $P(\varepsilon = 0) = 0$

(e) False. If $X'\varepsilon = 0$ then $\operatorname{cov}(X'\varepsilon) = \operatorname{cov}(0) = 0$, but $\operatorname{cov}(X'\varepsilon) = X'\operatorname{cov}(\varepsilon)X = \sigma^2 X'X \neq 0$ because columns of X linearly independent guarantees $X'X$ positive definite.

(f) True, since $y = X\beta + \varepsilon$.

(g) False. $\hat{\varepsilon}'\hat{\varepsilon} = 0$ iff $\hat{\varepsilon} = y - \hat{y} = 0 \Rightarrow y = \hat{y}$, which is not true in general.

(h) True, though a bit surprising. $\hat{\varepsilon}'\hat{\varepsilon} = (y - \hat{y})'\hat{\varepsilon}$

$$= (y' - \hat{y}')\hat{\varepsilon} = y'\hat{\varepsilon} - \hat{y}'\hat{\varepsilon} = y'\hat{\varepsilon} - (X\hat{\beta})'\hat{\varepsilon}$$

$$= y'\hat{\varepsilon} - \hat{\beta}'\underbrace{X'\hat{\varepsilon}}_{0} = y'\hat{\varepsilon}.$$

(i) True: $\frac{\varepsilon_i - 0}{\sigma} \sim N(0,1)$, $\left(\frac{\varepsilon_i - 0}{\sigma}\right)^2 \sim \chi^2(1)$, independent,

so $$\frac{\varepsilon'\varepsilon}{\sigma^2} = \sum_{i=1}^{n} \left(\frac{\varepsilon_i - 0}{\sigma}\right)^2 \sim \chi^2(n)$$

(1j) False. By (i), $E\left(\frac{\varepsilon'\varepsilon}{\sigma^2}\right) = n$

$$\Rightarrow E(\varepsilon'\varepsilon) = n\sigma^2 \neq 0$$

(k) False. $\frac{\hat{\varepsilon}'\hat{\varepsilon}}{\sigma^2} \sim \chi^2(n-k-1) \Rightarrow E\left(\frac{\hat{\varepsilon}'\hat{\varepsilon}}{\sigma^2}\right) = n-k-1$

$$\Rightarrow E(\hat{\varepsilon}'\hat{\varepsilon}) = \sigma^2(n-k-1) \neq 0.$$

(2) (a) $y \sim N_n(X\beta, \sigma^2 I_n)$

(b) $\hat{\beta} = (X'X)^{-1}X'y = Ay$ so $\hat{\beta}$ is multivariate normal.

$$E(\hat{\beta}) = (X'X)^{-1}X'E(y) = (X'X)^{-1}X'X\beta = \beta$$

$$\begin{aligned}
\operatorname{cov}(\hat{\beta}) &= (X'X)^{-1}X'\operatorname{cov}(y)\left((X'X)^{-1}X'\right)' \\
&= (X'X)^{-1}X'\sigma^2 I_n X\left((X'X)^{-1}\right)' \\
&= \underbrace{(X'X)^{-1}X'X}_{I_{k+1}}\sigma^2(X'X)^{-1} = \sigma^2(X'X)^{-1}
\end{aligned}$$

so $\hat{\beta} \sim N_{k+1}\left(\beta, \sigma^2(X'X)^{-1}\right)$

(2c) $\hat{y} = Hy$. $E(\hat{y}) = HE(y) = HX\beta$

$$= X\underbrace{(X'X)^{-1}X'X}_{I}\beta = X\beta$$

$\text{cov}(\hat{y}) = H\,\text{cov}(y)H' = H\sigma^2 I_n H = \sigma^2 H$,

so

$$\hat{y} \sim N_n(X\beta, \sigma^2 H)$$

(d) $\hat{\varepsilon} = (I-H)y$. $E(\hat{\varepsilon}) = (I-H)E(y)$

$$= (I-H)X\beta = X\beta - HX\beta = X\beta - X\beta = 0$$

$\text{cov}(\hat{\varepsilon}) = (I-H)\text{cov}(y)(I-H)' = (I-H)\sigma^2 I_n (I-H)$

$= \sigma^2 (I-H)$, so

$$\hat{\varepsilon} \sim N_n(0, \sigma^2(I-H))$$

(3) Using $\text{cov}(Ay, By) = A\,\text{cov}(y)B'$,

$$\text{cov}(\hat{\varepsilon}, \hat{\beta}) = \text{cov}\left((I-H)y, (X'X)^{-1}X'y\right)$$
$$= (I-H)\sigma^2 I_n \left((X'X)^{-1}X'\right)' = \sigma^2 (I-H)X(X'X)^{-1}$$
$$= \sigma^2\left(X(X'X)^{-1} - X(X'X)^{-1}\underline{X'X(X'X)^{-1}}\right)$$
$$= \sigma^2\left(X(X'X)^{-1} - X(X'X)^{-1}\right) = 0$$

This means $\hat{\varepsilon}$ is independent of $\hat{\beta}$, since zero covariance implies independence for the multivariate normal. Because $SSE = \hat{\varepsilon}'\hat{\varepsilon}$ is a function of $\hat{\varepsilon}$, SSE is independent of $\hat{\beta}$.

(4) $$\text{cov}(\hat{\varepsilon}, \hat{y}) = \text{cov}\left((I-H)y, Hy\right)$$
$$= (I-H)\,\text{cov}(y)H' = (I-H)\sigma^2 I\, H$$
$$= \sigma^2(H - H^2) = \sigma^2(H-H) = 0$$

By question (3), $\hat{\varepsilon}$ is independent of $\hat{\beta}$ if the error terms are normal, and so $\hat{y} = X\hat{\beta}$ is also independent of $\hat{\varepsilon}$ & hence $\text{cov}(\hat{y}, \hat{\varepsilon}) = 0$.

Independence only follows if the errors are normal but this problem shows by direct calculation that $\text{cov}(\hat{\varepsilon}, \hat{y}) = 0$ without any distributional assumptions.

(5) Since ε is MVN, $s_1 = X'\varepsilon$ is MVN.

$E(X'\varepsilon) = X'E(\varepsilon) = \underset{\sim}{0}$

$cov(X'\varepsilon) = X' cov(\varepsilon) X = X'\sigma^2 I_n X = \sigma^2 X'X$

Positive definite, since columns of X are linearly independent.

(6) (a) $s_2 = X'\hat{\varepsilon} \sim N(\underset{\sim}{0}, \underset{\sim}{0})$

$(k+1)\times 1$ $\quad$ $(k+1)\times(k+1)$

(b) $E(s_2) = E(X'\hat{\varepsilon}) = E(\underset{\sim}{0}) = 0$

$$\begin{aligned} cov(s_2) &= X' cov(\hat{\varepsilon}) X = X'\sigma^2(I-H)X \\ &= \sigma^2(X'X - X'HX) \\ &= \sigma^2(X'X - X'\underbrace{X(X'X)^{-1}X'X}_{I}) \\ &= \sigma^2(X'X - X'X) = \underset{\sim}{0} \end{aligned}$$

(c) No surprise. $X'\hat{\varepsilon} = \underset{\sim}{0}$, a vector of constants

(d) $P(s_2 = \underset{\sim}{0}) = 1$

7 (a) $(y-X\beta)'(y-X\beta)=(y-\hat{y}+\hat{y}-X\beta)'(y-\hat{y}+\hat{y}-X\beta)$

$$=(\hat{\varepsilon}+X\hat{\beta}-X\beta)'(\hat{\varepsilon}+X(\hat{\beta}-\beta))$$

$$=\left(\hat{\varepsilon}'+(X(\hat{\beta}-\beta))'\right)(\hat{\varepsilon}+X(\hat{\beta}-\beta))$$

$$=\hat{\varepsilon}'\hat{\varepsilon}+\underbrace{\hat{\varepsilon}'X}_{0}(\hat{\beta}-\beta)+(\hat{\beta}-\beta)'\underbrace{X'\hat{\varepsilon}}_{0}$$
$$+(\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta)$$

$$=\hat{\varepsilon}'\hat{\varepsilon}+0+0+(\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta)$$

(b) From (a),

$$\frac{1}{\sigma^2}(y-X\beta)'(y-X\beta)=\frac{\hat{\varepsilon}'\hat{\varepsilon}}{\sigma^2}+(\hat{\beta}-\beta)'\frac{1}{\sigma^2}X'X(\hat{\beta}-\beta)$$

$$\Rightarrow (y-X\beta)'(\sigma^2 I_n)^{-1}(y-X\beta)=\frac{SSE}{\sigma^2}$$
$$+(\hat{\beta}-\beta)\left(\sigma^2(X'X)^{-1}\right)^{-1}(\hat{\beta}-\beta)$$

That is, $w=w_1+w_2$

Know that if $v\sim N(\mu,\Sigma)$ then $(v-\mu)'\Sigma^{-1}(v-\mu)\sim\chi^2$

So $w\sim\chi^2(n)$, $w_2\sim\chi^2(k+1)$, w_1 & w_2 are independent because $\hat{\varepsilon}$ & $\hat{\beta}$ are independent, and

$$w_1=\frac{SSE}{\sigma^2}\sim\chi^2(n-(k+1))$$

(8) (a) $\hat{\beta} \sim N_{k+1}(\beta, \sigma^2 (X'X)^{-1})$ so $a'\hat{\beta}$ is normal.

$$E(a'\hat{\beta}) = a'E(\hat{\beta}) = a'\beta, \quad cov(a'\hat{\beta}) = a'cov(\hat{\beta})a$$

$$= \sigma^2 a'(X'X)^{-1}a, \text{ so}$$

$$a'\hat{\beta} \sim N(a'\beta, \sigma^2 a'(X'X)^{-1}a)$$

(b)

$$z = \frac{a'\hat{\beta} - a'\beta}{\sqrt{\sigma^2 a'(X'X)^{-1}a}}$$

(c)

$$t = \frac{z}{\sqrt{w/r}} = \frac{\dfrac{a'\hat{\beta} - a'\beta}{\sqrt{\cancel{\sigma}^2 a'(X'X)^{-1}a}}}{\sqrt{\dfrac{SSE}{\cancel{\sigma}^2} \Big/ (n-k-1)}}$$

$$= \frac{a'\hat{\beta} - a'\beta}{\sqrt{MSE\, a'(X'X)^{-1}a}} \sim t(n-k-1)$$

(d) Numerator is a function of $\hat{\beta}$. Denominator is a function of SSE. Functions of independent random variables are independent.

(8e)

$$t^* = \frac{a'\hat{\beta} - c}{\sqrt{MSE\ a'(X'X)^{-1}a}}$$

(f) $a'\beta = (0\ 0\ 1\ 0\ 0)\begin{pmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\\ \beta_4\end{pmatrix} = 0$

$$a = \begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix}$$

(g)

$$a'\beta = \left(0\ 1\ -\tfrac{1}{2}\ -\tfrac{1}{2}\ 0\right)\begin{pmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\\ \beta_4\end{pmatrix} = 0$$

$$a = \begin{pmatrix}0\\1\\-\frac{1}{2}\\-\frac{1}{2}\\0\end{pmatrix}$$

(8 h)

$$a = \begin{pmatrix} 1 \\ 91 \\ 83 \\ 24 \end{pmatrix}$$

(i) $1-\alpha = P\left(-t_{\alpha/2} < t < t_{\alpha/2}\right)$

$$= P\left(-t_{\alpha/2} < \frac{a'\hat{\beta} - a'\beta}{\sqrt{MSE\; a'(X'X)^{-1}a}} < t_{\alpha/2}\right)$$

$$= P\left(-t_{\alpha/2}\sqrt{MSE\; a'(X'X)^{-1}a} < a'\hat{\beta} - a'\beta < t_{\alpha/2}\sqrt{MSE\; a'(X'X)^{-1}a}\right)$$

$$= P\Big(-a'\hat{\beta} - t_{\alpha/2}\sqrt{MSE\; a'(X'X)^{-1}a} < -a'\beta$$
$$< -a'\hat{\beta} + t_{\alpha/2}\sqrt{MSE\; a'(X'X)^{-1}a}\Big)$$

$$= P\Big(a'\hat{\beta} + t_{\alpha/2}\sqrt{MSE\; a'(X'X)^{-1}a} > a'\beta$$
$$> a'\hat{\beta} - t_{\alpha/2}\sqrt{MSE\; a'(X'X)^{-1}a}\Big)$$

$$= P\Big(a'\hat{\beta} - t_{\alpha/2}\sqrt{MSE\; a'(X'X)^{-1}a} < a'\beta$$
$$< a'\hat{\beta} - t_{\alpha/2}\sqrt{MSE\; a'(X'X)^{-1}a}\Big), \text{ or}$$

$$a'\hat{\beta} \pm t_{\alpha/2}\sqrt{MSE\; a'(X'X)^{-1}a}$$

9) Here $\hat\beta = \bar y$, $\hat y = X\hat\beta$,

(a)

$$\hat\varepsilon = y - X\hat\beta = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \bar y = \begin{pmatrix} y_1 - \bar y \\ y_2 - \bar y \\ \vdots \\ y_n - \bar y \end{pmatrix}$$

$SSE = \hat\varepsilon'\hat\varepsilon = \sum_{i=1}^{n} (y_i - \bar y)^2$, and $MSE = \dfrac{SSE}{n-0-1}$

$$= \frac{\sum_{i=1}^{n} (y_i - \bar y)^2}{n-1} = s^2,$$ the usual sample variance

(b) Confidence interval for β_0: $a = 1$, $(X'X)^{-1} = \frac{1}{n}$, and

$$a'\hat\beta \pm t_{\alpha/2} \sqrt{MSE\, a'(X'X)^{-1}a}$$

$$= \bar y \pm t_{\alpha/2} \sqrt{s^2/n} = \bar y \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

the usual confidence interval

10 (a) Because $\hat{\beta} \sim N_{k+1}(\beta, \sigma^2(X'X)^{-1})$,

$C\hat{\beta} \sim N_q(C\beta, \sigma^2 C(X'X)^{-1}C')$

(b) Formula sheet says if $y \sim N_p(\mu, \Sigma)$, then $w = (y-\mu)'\Sigma^{-1}(y-\mu) \sim \chi^2(p)$.

$$\frac{1}{\sigma^2}(C\hat{\beta}-t)'(C(X'X)^{-1}C')^{-1}(C\hat{\beta}-t)$$

$$= (C\hat{\beta}-t)'(\sigma^2 C(X'X)^{-1}C')^{-1}(C\hat{\beta}-t) \sim \chi^2(q)$$

c) $\frac{SSE}{\sigma^2} \sim \chi^2(n-k-1)$, whether H_0 is true or false.

d) i)

$$F^* = \frac{\cancel{\frac{1}{\sigma^2}}(C\hat{\beta}-t)(C(X'X)^{-1}C')^{-1}(C\hat{\beta}-t)/q}{\frac{SSE}{\cancel{\sigma^2}}/(n-k-1)}$$

$$= \frac{(C\hat{\beta}-t)'(C(X'X)^{-1}C')^{-1}(C\hat{\beta}-t)}{q\,MSE}$$

ii) The numerator is a function of $\hat{\beta}$, the denominator is a function of SSE, $\hat{\beta}$ & SSE are independent, and functions of independent random variables are independent.

iii) $F^* \overset{H_0}{\sim} F(q, n-k-1)$

(11) To test $H_0: a'\beta = t_0$, a' plays the role of C, and

$$F^* = \frac{(a'\hat{\beta} - t_0)'(a'(X'X)^{-1}a)^{-1}(a'\hat{\beta} - t_0)}{1 \cdot MSE}$$

$$= \frac{(a'\hat{\beta} - t_0)^2}{MSE\ a'(X'X)^{-1}a}$$, since both $a'\hat{\beta} - t_0$ and $a'(X'X)^{-1}a$ are scalars

For the t-test,

$$t^{*2} = \left(\frac{a'\hat{\beta} - t_0}{\sqrt{MSE\ a'(X'X)^{-1}a}} \right)^2$$

$$= \frac{(a'\hat{\beta} - t_0)^2}{MSE\ a'(X'X)^{-1}a} = F^*$$

(12) For $H_0: AC\beta = At$, the test statistic is

$$\frac{(AC\hat{\beta} - At)'\left(AC(X'X)^{-1}(AC)'\right)^{-1}(AC\hat{\beta} - At)}{q\,MSE}$$

$$= \frac{\left(A(C\hat{\beta} - t)\right)'\left(AC(X'X)^{-1}C'A'\right)^{-1}\left(A(C\hat{\beta} - t)\right)}{q\,MSE}$$

$$= \frac{(C\hat{\beta} - t)'\underbrace{A'A'^{-1}}_{I}(C(X'X)^{-1}C')^{-1}\underbrace{A^{-1}A}_{I}(C\hat{\beta} - t)}{q\,MSE}$$

$$= \frac{(C\hat{\beta} - t)(C(X'X)^{-1}C')^{-1}(C\hat{\beta} - t)}{q\,MSE} = F^*,$$

The test statistic for $H_0: C\beta = t$

13

(a) $$\frac{\dfrac{SSR(full)}{\cancel{SST}} - \dfrac{SSR(Reduced)}{\cancel{SST}}}{q\left(\dfrac{SST - SSR(full)}{\cancel{SST}}\right)/(n-k-1)}$$

$$= \frac{SSR(full) - SSR(reduced)}{q\ SSE(full)/(n-k-1)} = \frac{SSR(full) - SSR(reduced)}{q\ MSE(full)}$$

(b) $$\frac{SSR(full) - SST + SST - SSR(reduced)}{q\ MSE(full)} = \frac{-SSE(full) + SSE(reduced)}{q\ MSE(full)}$$

$$= \frac{SSE(reduced) - SSE(full)}{q\ MSE(full)}$$

(c) $$\left(\frac{n-k-1}{q}\right)\left(\frac{p}{1-p}\right) = \left(\frac{n-k-1}{q}\right)\left(\frac{\dfrac{R^2(full) - R^2(reduced)}{1 - R^2(reduced)}}{1 - \dfrac{R^2(full) - R^2(reduced)}{1 - R^2(reduced)}}\right)$$

$$= \left(\frac{n-k-1}{q}\right)\left(\frac{\dfrac{R^2(full) - R^2(reduced)}{\cancel{1 - R^2(reduced)}}}{\dfrac{1 - \cancel{R^2(reduced)} - R^2(full) + \cancel{R^2(reduced)}}{\cancel{1 - R^2(reduced)}}}\right)$$

$$= \frac{(R^2(full) - R^2(reduced))/q}{(1 - R^2(full))/(n-k-1)} = \frac{SSR(full) - SSR(reduced)}{q\ MSE(full)}$$

by part (a)

(13d)

$$F^* = \left(\frac{n-k-1}{q}\right)\frac{p}{1-p} \iff F^* - pF^* = \frac{n-k-1}{q}\, p$$

$$\iff qF^* - pqF^* = (n-k-1)p$$

$$\iff qF^* = (n-k-1)p + qF^*p = p(qF^* + n-k-1)$$

$$\iff p = \frac{qF^*}{qF^* + n-k-1}$$

(14) (a) The first element of $X'\hat{\varepsilon}$ is the inner product of a row of ones with $\hat{\varepsilon}$, so $\sum_{i=1}^{n}\hat{\varepsilon}_i = 0$

(b) $$0 = \sum_{i=1}^{n}\hat{\varepsilon}_i = \sum_{i=1}^{n}(y_i - \hat{y}_i) = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n}\hat{y}_i$$

$$\implies \sum_{i=1}^{n} y_i = \sum_{i=1}^{n}\hat{y}_i$$

(c) $$\bar{y} = \frac{1}{n}\sum_{i=1}^{n}\hat{y}_i = \frac{1}{n}\underset{\text{all ones}}{j'}\hat{y} = \frac{1}{n}j'X\hat{\beta}, \text{ so}$$

$\bar{y}$ is a function of $\hat{\beta}$.

Since $\hat{\beta}$ & $\hat{\varepsilon}$ are independent and functions of independent random vectors are independent, this means $\bar{y}$ & $\hat{\varepsilon}$ are independent.

(15) SSE is a function of $\hat{\varepsilon}$. $SSR = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$ is a function of $\hat{\beta}$ if the model has an intercept because, by problem 14, $\bar{y}$ is a function of $\hat{\beta}$. Since functions of independent random vectors are independent and $\hat{\varepsilon}$ & $\hat{\beta}$ are independent, this means SSE & SSR are independent.

(16) (a) $y_i \overset{iid}{\sim} N(\beta_0, \sigma^2)$

(b) $$\frac{SST}{\sigma^2} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{\sigma^2} \sim \chi^2(n-1)$$

(c) Have
$$\frac{SST}{\sigma^2} = \frac{SSR}{\sigma^2} + \frac{SSE}{\sigma^2}$$
$$w = w_1 + w_2$$

- w_1 & w_2 are independent because SSR & SSE are independent.
- $\frac{SST}{\sigma^2} \sim \chi^2(n-1)$ by part (b)
- $\frac{SSE}{\sigma^2} \sim \chi^2(n-k-1)$

So by the formula sheet, w_1 is chi-squared, with degrees of freedom $(n-1)-(n-k-1)$
$= n-1-n+k+1 = k$ $\quad \frac{SSR}{\sigma^2} \sim \chi^2(k)$

(16 d.) $\frac{SSE}{\sigma^2} \sim \chi^2(n-k-1)$ as usual

$\frac{SSR}{\sigma^2} \sim \chi^2(k)$ by part (c) If H_0 is true

$w_1 = \frac{SSR}{\sigma^2}$ & $w_2 = \frac{SSE}{\sigma^2}$ are independent because by problem 15, SSR & SSE are independent.

Then by the definition of the F distribution,

$$F^* = \frac{\frac{SSR}{\cancel{\sigma^2}}/k}{\frac{SSE}{\cancel{\sigma^2}}/(n-k-1)} \overset{H_0}{\sim} F(k, n-k-1)$$

(e) For the reduced model with $\beta_1 = \cdots = \beta_k = 0$, $\hat{y}_i = \bar{y}$, and $SSR(\text{reduced}) = \sum_{i=1}^{n}(\bar{y}-\bar{y})^2 = 0$ so

$$F^* = \frac{SSR(\text{full}) - SSR(\text{reduced})}{q\ MSE(\text{full})}$$

$$= \frac{SSR(\text{full}) - 0}{k\ MSE(\text{full})}$$

$$= \frac{SSR/k}{SSE/(n-k-1)}$$, same as 15(d)

(16f)

$$F^* = \frac{SSR/k}{SSE/(n-k-1)} = \left(\frac{n-k-1}{k}\right)\left(\frac{SSR}{SST-SSR}\right)$$

$$= \left(\frac{n-k-1}{k}\right) \frac{SSR/SST}{1 - \frac{SSR}{SST}}$$

$$= \left(\frac{n-k-1}{k}\right) \frac{R^2}{1-R^2}$$

$$\frac{\partial \ln F^*}{\partial R^2} = \frac{\partial}{\partial R^2}\left(\ln\left(\frac{n-k-1}{k}\right) + \ln R^2 - \ln(1-R^2)\right)$$

$$= 0 + \frac{1}{R^2} - \frac{1}{1-R^2}(-1)$$

$$= \frac{1}{R^2} + \frac{1}{1-R^2} > 0 \quad \text{increasing in } R^2$$