Midterm 1 - 2015 Answer Key

1. C
2. D
3. C
4. A
5. C

6.) $e = y - \hat{y} = y - Xb = (X\beta + \varepsilon) - X(X'X)^{-1}X'(X\beta + \varepsilon)$

$= X\beta + \varepsilon - X\beta - X(X'X)^{-1}X'\varepsilon$

$= \varepsilon - X(X'X)^{-1}X'\varepsilon \;\; = M\varepsilon$

So, the residuals are a linear function of the (unobservable) error term. Hence, if ε is Normally distributed, so is e. Need to assume that X is non-stochastic.

7.) a) $R^2 = 0$ when the estimated regression has "no fit", i.e. when X has no predictive power for y.

This could occur if all of the estimated slope coefficients are equal to zero. In this case, the estimate for the intercept is $\bar{y}$, and SSR $= 0$, since $\hat{y} = \bar{y}$.

b) $R^2 = 1$ if there is "perfect fit", i.e. if the regression "line" (or "plane") passes through each data point. In this case, each OLS predicted value is equal to the actual value ($\hat{y} = y$), and SSR $=$ SST.

8.) a) $b_1 = (X_1'X_1)^{-1}X_1'y$

$$= (X_1'X_1)^{-1}X_1'(X_1\beta_1 + X_2\beta_2 + u)$$

$$= \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2 + (X_1'X_1)^{-1}X_1'u,$$

and

$$E(b_1) = \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2,$$

by A.3. Since $E(b_1) \neq \beta_1$ in general, b_1 is biased.

b) There are two situations in which b_1 will be unbiased, however. The term:

$$(X_1'X_1)^{-1}X_1'X_2\beta_2$$

will disappear if X_1 and X_2 are orthogonal ($X_1'X_2 = 0$), or if $\beta_2 = 0$.

9.) In order to get the squared sum of deviations-from-means, we can make use of the "residual maker" matrix "M", but where the "X" variable in "M" is a column of 1's. That is,

$$M_0 = I - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}', \text{ where } \mathbf{1} \text{ is an } n\times 1 \text{ column of 1's.}$$

So, $M_0 = I - \frac{1}{n}\mathbf{1}\mathbf{1}'$,

and $M_0 y = y - \bar{y}_{n\times 1}$

In order to get the sum of squares: $y'M_0'M_0y = y'M_0y$.

M_0, and $y'M_0y$, are in the formula sheet provided on the exam.

10.) a) $\tilde{\beta} = (X'C'CX)^{-1}X'C'Cy$

$$= (X'C'CX)^{-1}X'C'C(X\beta + \varepsilon)$$

$$= \beta + (X'C'CX)^{-1}X'C'C\varepsilon,$$

If X is assumed to be non-random, then

$E(\tilde{\beta}) = \beta + 0$, and $\tilde{\beta}$ is unbiased.

b) The C matrix effectively eliminates half of the data from the data-set. For example, if $n=4$ and $K=2$:

$$CX = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \\ X_{31} & X_{32} \\ X_{41} & X_{42} \end{bmatrix} = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and similarly

$$Cy = \begin{bmatrix} y_1 \\ y_2 \\ 0 \\ 0 \end{bmatrix},$$

Hence, $\tilde{\beta}$ is essentially OLS, using half of the data. Since OLS is unbiased regardless of sample size, $\hat{\beta}$ should be unbiased.

c) We would expect $\tilde{\beta}$ to have higher variance than b (in a matrix sense) for similar arguments as in part (b).

We Know that the variability of our estimator, b, decreases as the sample size grows. $\tilde{\beta}$ is like OLS, but with a smaller data set, so we would expect it to have higher variance.