Answer Keys to Problem Set 1

1. Show that $E(\epsilon|X) = 0$ implies $E(\epsilon) = 0$.

By the Law of Iterated Expectations,

$$E(\varepsilon) = E(E(\varepsilon|X)) = 0$$

2. Show $E(\epsilon|X) = 0$ implies $E(X'\epsilon) = 0$.

By the Law of Iterated Expectations.

$$E(X'\varepsilon) = E(E(X'\varepsilon|X))$$

$$E(X'\varepsilon) = E(X'E(\varepsilon|X)) = 0 \text{ , since } E[\varepsilon|X] = 0$$

3. Show $Var(X) = E(Var(X|Y)) + Var(E(X|Y))$.

By definition

$$Var(X) = E(X^2) - (E(X))^2$$

$$Var(X|Y) = E(X^2|Y) - (E[X|Y])^2$$

Take expectation on both side

$$E[Var(X|Y)] = E[X^2] - E[(E(X|Y))^2]$$

$$Var(E(X|Y)) = E[(E(X|Y))^2] - [E[E(X|Y)]]^2 = E[(E(X|Y))^2] - (E(X))^2$$

Combine the above 2 terms together, we obtain the desired result.

4. Show $Cov(X, Y) = Cov(X, E(Y|X))$.

$$cov(X \cdot Y) = E(X \cdot Y) - E(X)E(Y)$$

$$cov(X \cdot E(Y|X)) = E(X\, E(Y|X)) - E(X)\, E(E(Y|X))$$

$$= E(XY) - E(X)E(Y) = cov(X \cdot Y)$$

5. Show that the matrix $M = (I - X(X'X)^{-1}X')$ is idempotent.

$$P = X(X'X)^{-1}X'$$

$$M = I - P$$

$$M^2 = (I-P)(I-P) = I - P - P + P^2 = I - P$$

$$\text{Since } P^2 = X(X'X)^{-1}X'X(X'X)^{-1}X' = X(X'X)^{-1}X' = P$$

6. If $x \sim N(\mu, \sigma^2)$, show $\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\sigma^2} \sim \chi^2(n-1)$.

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n}(x_i - \mu + \mu - \bar{x})^2 = \sum_{i=1}^{n}(x_i - \mu)^2 + 2\sum_{i=1}^{n}(x_i - \mu)(\mu - \bar{x}) + \sum_{i=1}^{n}(\mu - \bar{x})^2$$

$$= \sum_{i=1}^{n}(x_i - \mu)^2 - 2n(\bar{x} - \mu)^2 + n(\bar{x} - \mu)^2$$

$$= \sum_{i=1}^{n}(x_i - \mu)^2 - n(\bar{x} - \mu)^2$$

By definition

$$\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{\sigma^2} \sim \chi^2(n) \; ; \; \frac{\sum_{i=1}^{n}(\bar{x} - \mu)^2}{\sigma^2} \sim \chi^2(1) \text{, Thus } \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\sigma^2} \sim \chi^2(n-1)$$

7. Let X and Y be random variables with finite means. Show

$$\min_{g(X)} E(Y - g(X))^2 = E(Y - E(Y|X))^2,$$

where $g(X)$ ranges over all functions. This implies that conditional mean $(E(Y|X))$ is the best predictor of Y.

$$E[(y - g(x))^2] = E[((y - E(y|x)) - (g(x) - E(y|x)))^2]$$

$$= E[(y - E(y|x))^2 - 2[y - E(y|x)][g(x) - E(y|x)] + (g(x) - E(y|x))^2]$$

$$= E[e^2] - 2E[e(g(x) - E(y|x))] + E[(g(x) - E(y|x))^2], \quad E[e] = 0$$

$$= E[e^2] + E[(g(x) - E[y|x])^2]$$

$$\geq E[e^2] = E[(y - E(y|x))^2]$$

If $g(x) = E[y|x]$,

$E[(y - g(x))^2]$ will be minimized.

Question II. The OLS estimator for the linear regression $Y = X\beta + \epsilon$ is $\hat{\beta}_{OLS} = (X'X)^{-1}X'Y$:

1. Derive the OLS estimator $\hat{\beta}_{OLS}$

$\min (Y - X\beta)'(Y - X\beta)$

F.O.C. $X'X\beta = X'Y$.

$\hat{\beta}_{OLS} = (X'X)^{-1}(X'Y)$. The full rank condition is required.

Otherwise, the inverse $(X'X)$ would not exist.

2. Derive the variance of the OLS estimator, $var(\hat{\beta}_{OLS}|X)$.

$$
\begin{aligned}
Var(\hat{\beta}_{OLS}|X) &= E[(\hat{\beta}_{OLS} - E(\hat{\beta}_{OLS}|X))'(\hat{\beta}_{OLS} - E(\hat{\beta}_{OLS}|X))|X] \\
&= E[(X'X)^{-1}X'\varepsilon\varepsilon'X(X'X)^{-1}|X] \\
&= (X'X)^{-1}X'E[\varepsilon\varepsilon'|X]X(X'X)^{-1} \\
&= (X'X)^{-1}X' var(\varepsilon|X)X(X'X)^{-1} \underset{\text{if homoskedasticity}}{=} \sigma^2(X'X)^{-1}
\end{aligned}
$$

3. Show the unbiasedness of $\hat{\beta}_{OLS}$, i.e., $E(\hat{\beta}_{OLS}|X) = \beta$.

$$E[\hat{\beta}_{OLS}|X] = E[(X'X)^{-1}(X'Y)|X] = (X'X)^{-1}X'E[X\beta + \varepsilon|X]$$

$$= \beta + (X'X)^{-1}X'E[\varepsilon|X] = \beta$$

if $E[\varepsilon|X]=0$. We can see that the unbiasedness of the OLS requires the conditional independence (exogeneous) assumption.

4. Given that the error term $\epsilon \sim N(0, \sigma^2 I)$, write down the distribution of $\hat{\beta}$ conditional on X, i.e., $P(\hat{\beta}_{OLS}|X)$.

From 2. and 3. we know that

$E(\hat{\beta}_{OLS}|X) = \beta$ and $Var(\hat{\beta}_{OLS}|X) = (X'X)^{-1}X' Var(\varepsilon|X) X (X'X)^{-1}$

under proper assumption. With normality and homoskedasticity assumptions.

$\varepsilon \overset{iid}{\sim} N(0, \sigma^2 I)$, we have

$$Var(\hat{\beta}_{OLS}|X) = (X'X)^{-1}\sigma^2$$

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

5. Write down the asymptotic distribution of $\hat{\beta}_{OLS}$.

$$\hat{\beta}_{OLS} = (X'X)^{-1}(X'Y) = \beta + (X'X)^{-1}(X'\varepsilon) = \beta + \left(\frac{X'X}{n}\right)^{-1}\left(\frac{X'\varepsilon}{n}\right)$$

$$Q_{xx} = \operatorname*{plim}_{n\to\infty} \frac{X'X}{n} = \operatorname*{plim}_{n\to\infty} \frac{\sum_{i=1}^{n} x_i x_i'}{n} = E[x_i x_i'] = \frac{1}{n} E[X'X]$$

$$(\hat{\beta}_{OLS} - \beta) = \left(\frac{X'X}{n}\right)^{-1}\left(\frac{X'\varepsilon}{n}\right)$$

$$\sqrt{n}(\hat{\beta}_{OLS} - \beta) = \left(\frac{X'X}{n}\right)^{-1}\left(\frac{X'\varepsilon}{\sqrt{n}}\right)$$

$$\operatorname*{plim}_{n\to\infty}\left(\frac{X'X}{n}\right) = Q_{xx} \quad ; \quad \operatorname*{plim}_{n\to\infty}\left(\frac{X'\varepsilon}{\sqrt{n}}\right) = E[X'\varepsilon] = 0 \text{ by WLLN}$$

$$\left(\frac{X'\varepsilon}{\sqrt{n}}\right) \xrightarrow{d} N(0, \Omega) \quad , \quad \Omega = E[x_i x_i' \varepsilon_i^2]$$

$$\sqrt{n}(\hat{\beta}_{OLS} - \beta) \xrightarrow{d} N(0. Q_{xx}^{-1} \Omega Q_{xx}^{-1})$$

Note, when ε_i is Homoskedastic Projection error

$$\operatorname{cov}(x_i x_i' \varepsilon_i^2) = 0 \quad , \quad \Omega = E[x_i x_i'] E[\varepsilon_i^2] = Q_{xx}\sigma^2$$

$$\sqrt{n}(\hat{\beta}_{OLS} - \beta) \xrightarrow{d} N(0. \sigma^2 Q_{xx}^{-1})$$

6. Show $\hat{\beta}_{OLS}$ is a consistent estimator of β, i.e., $\text{plim}_{n\to\infty}\hat{\beta}_{OLS} = \beta$.

$$\hat{\beta}_{OLS} = (X'X)^{-1}(X'Y) = \beta + (X'X)^{-1}(X'\varepsilon) = \beta + \left(\frac{X'X}{n}\right)^{-1}\left(\frac{X'\varepsilon}{n}\right)$$

$$\text{plim}_{n\to\infty}\left(\frac{X'X}{n}\right) = Q_{xx} \quad ; \quad \text{plim}_{n\to\infty}\left(\frac{X'\varepsilon}{\sqrt{n}}\right) = E[X'\varepsilon] = 0 \text{ by WLLN}$$

$$\therefore \text{plim}_{n\to\infty}\hat{\beta}_{OLS} = \beta.$$

$\hat{\beta}_{OLS}$ is a consistent estimator of β under suitable regularity conditions.

7. Show $S^2 = \frac{e'e}{n-K}$ is an unbiased estimator of σ^2, i.e., $E(S^2|X) = \sigma^2$.

$$E(S^2|X) = \frac{1}{n-K}E[e'e|X] = \frac{1}{n-K}\text{trace}[E(\varepsilon'MM\varepsilon)|X]$$

under Homoskedasticity assumption

$$= \frac{\sigma^2}{n-K}\text{trace } M$$

$$= \frac{\sigma^2}{n-K}\text{trace }(I_n - X'(XX')^{-1}X)$$

$$= \frac{\sigma^2}{n-K}\text{trace }(I_n - I_K)$$

$$= \frac{n-K}{n-K}\sigma^2 = \sigma^2$$

Note

$e = Y - X\hat{\beta}_{OLS}$, is residuals of the OLS estimator

$e = Y - X\hat{\beta}_{OLS}$

$= (I-P)Y = MY = M\varepsilon$

$y = X\beta + \varepsilon$

8. $\hat{\sigma}^2 = \frac{e'e}{n}$ is a consistent estimator of σ^2, true or false? Explain.

True. By WLLN, we know the sample moment $\hat{\sigma}^2 = \frac{e'e}{n}$ converge to population moment $E(\varepsilon'\varepsilon)$. Therefore, $\hat{\sigma}^2$ is a consistent estimator of σ^2

9. If the error term $\epsilon \sim N(0, \sigma^2\Omega)$, where Ω is a $n \times n$ symmetric, positive definite matrix, what should be the correct variance of $\hat{\beta}_{OLS}$, i.e., $var(\hat{\beta}_{OLS}|X)$?

$$Var(\hat{\beta}_{OLS}|X) = \sigma^2 (X'X)^{-1} X'\Omega X (X'X)^{-1}$$

Question III. The data matrix is $(Y, \mathbf{X})$ with $\mathbf{X} = [\mathbf{X_1}, \mathbf{X_2}]$. Consider the transformed regressor matrix $\mathbf{Z} = [\mathbf{X_1}, \mathbf{X_2} - \mathbf{X_1}]$. Suppose you do a least square regression of Y on $\mathbf{X}$ and a least square regression of Y on $\mathbf{Z}$. Let $\hat{\sigma}^2$ and $\tilde{\sigma}^2$ denote the residual variance estimate from the two regressions. Give a formula relating $\hat{\sigma}^2$ and $\tilde{\sigma}^2$.

$$\hat{e} = y - x\hat{\beta} = (I - P)y = My = Me$$

$$XB = Z$$

$$\because [X_1, X_2]B = [X_1, X_2 - X_1] \qquad \therefore B = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

$$Var(\hat{e} \mid X) = E[Mee'M' \mid X]$$
$$= M\,E[ee' \mid X]\,M' = M\sigma^2$$

$$M = I - P = I - X(X'X)^{-1}X'$$

$$\because P_Z = Z(Z'Z)^{-1}Z' = XB(B'X'XB)^{-1}B'X$$
$$= XBB^{-1}(X'X)^{-1}(B')^{-1}B'X$$
$$= X(X'X)^{-1}X = P_X$$

$$M_Z = M_X$$

$$\therefore M_X\sigma^2 = \hat{\sigma}^2 = \tilde{\sigma}^2 = M_Z\sigma^2$$