**UBC, ECON 326 - 033**

**Solution to 2010 Midterm Examination**

## Question 1

**(a)** True. $\hat{U}_i$'s are the fitted residuals from a regression with an intercept, and $\sum_i \hat{U}_i = 0$ is the normal equation obtained from the OLS first-order conditions for the intercept.

**(b)** False (in general). $\tilde{U}_i$'s are the fitted residuals from a regression without an intercept, and therefore the OLS first-order condition corresponding to the intercept does not have to hold.

**(c)** False (in general). $EU_i = 0$, however, a sample average of (finitely many) $U_i$'s does not have to be zero.

**(d)** True. By the law of iterated expectation (LIE),

$$\begin{aligned} E\left(U_i X_i^4\right) &= E\left(E\left(U_i X_i^4 | X_i\right)\right) \\ &= E\left(X_i^4 E\left(U_i | X_i\right)\right) \\ &= E\left(X_i^4 \cdot 0\right) \\ &= 0. \end{aligned}$$

**(e)** False. Since $\tilde{\beta}_1$ is a biased estimator, the Gauss-Markov Theorem does not apply in this case. In fact, one can show that $\tilde{\beta}_1$ has a smaller conditional variance given $X_1, \ldots, X_n$.

## Question 2

**(a)** Since $n = 29$ and the regression has an intercept, the number of df is 27.

**A.** We can use the formula for the lower bound of a confidence interval:

$$\begin{aligned} 0.2026 &= \hat{\beta}_1 - t_{27,1-0.05/2} \times \text{SE} \\ &= \hat{\beta}_1 - 2.052 \times 0.3739. \end{aligned}$$

It follows that $\hat{\beta}_1 = 0.9698$.

**B.** $t = (\hat{\beta}_1 - 0)/\text{SE} = 0.9698/0.3739 = 2.5939$.

**C.** Let $\xi$ be a random variable with $t_{27}$ distribution. According to the $t$-table, $P(\xi > 2.5939) \approx P(\xi > 2.473) \approx 0.01$. Hence, the $p$-value is $\approx 0.02$.

**D.** $\hat{\beta}_1 + t_{27,1-0.05/2} \times \text{SE} = 0.9698 + 2.052 \times 0.3739 = 1.7371$.

**(b)** The $p$-value for the test of $H_0 : \beta_1 = 0$ against $H_1 : \beta_1 \neq 0$ is $\approx 0.02$. Hence, we can conclude that x affects y if the significance level $\alpha > 0.02$.

**Question 3** Let $E_X$ denote the conditional expectation given $X_1, \ldots, X_n$, and $Var_X$ denote the conditional variance given $X_1, \ldots, X_n$, i.e. $E_X(\cdot) = E(\cdot|X_1, \ldots, X_n)$, and $Var_X(\cdot) = Var(\cdot|X_1, \ldots, X_n)$.

**(a)** To show unbiasedness, first, write:

$$\begin{aligned} \tilde{\beta} &= \frac{\bar{Y}}{\bar{X}} \\ &= \frac{\beta \bar{X} + \bar{U}}{\bar{X}} \\ &= \beta + \frac{\bar{U}}{\bar{X}}. \end{aligned}$$

\Next,

$$\begin{aligned} E_X(\tilde{\beta}) &= E_X\left(\beta + \frac{\bar{U}}{\bar{X}}\right) \\ &= \beta + \frac{E_X(\bar{U})}{\bar{X}} \\ &= \beta + \frac{n^{-1}\sum_{i=1}^n E_X(U_i)}{\bar{X}} \\ &= \beta + \frac{0}{\bar{X}} \\ &= \beta. \end{aligned}$$

The unbiasedness follows by the LIE.

**(b)** $Var_X(\tilde{\beta}) = Var_X(\beta + \bar{U}/\bar{X}) = Var_X(\bar{U})/(\bar{X})^2$. Next, since $E_X(\bar{U}) = 0$,

$$\begin{aligned} Var_X(\bar{U}) &= E_X(\bar{U})^2 \\ &= E_X\left(n^{-1}\sum_{i=1}^n U_i\right)^2 \\ &= n^{-2}\sum_{i=1}^n\sum_{j=1}^n E_X(U_iU_j) \\ &= n^{-2}\sum_{i=1}^n E_X(U_i^2) \text{ (no serial correlation)} \\ &= n^{-2}\sum_{i=1}^n \sigma^2 \text{ (homoskedasticity)} \\ &= n^{-1}\sigma^2. \end{aligned}$$

Thus,

$$Var_X(\tilde{\beta}) = \frac{\sigma^2}{n(\bar{X})^2}.$$

**(c)** The result follows if we can show that $\sum_{i=1}^n X_i^2 \geq n(\bar{X})^2$. Note, however, that

$$\sum_{i=1}^n X_i^2 - n(\bar{X})^2 = \sum_{i=1}^n (X_i - \bar{X})^2 \geq 0.$$