Econ 3040 - 2020 Midterm Answer Key

Ryan T. Godwin

1. a) $\text{E}[Y] = 0.8 \times 1 + 0.2 \times 6 = 2$.
   b) $\text{Var}[Y] = 0.8 \times (1-2)^2 + 0.2 \times (6-2)^2 = 4$.
   c) $\text{E}[Z] = 4 \times \text{E}[Y] = 8$. $\text{Var}[Z] = 16 \times \text{Var}[Y] = 64$.
   d) There is an exact, positive linear relationship between the two variables ($Z = 4Y$), so the correlation is 1.

2. $\bar{y}$, b_0, and b_1 are *random* estimators because they are calculated from a sample drawn randomly from the population. The CLT says that the sum of random variables tend to be Normally distributed. Since these estimators involve *summing* random variables, the CLT theorem implies that these estimators should follow a Normal distribution.

3.

$$
\begin{aligned}
\text{Var}\left[\bar{y}\right] &= \text{Var}\left[\frac{1}{n}\sum_{i=1}^{n} y_i\right] \\
&= \frac{1}{n^2}\text{Var}\left[\sum_{i=1}^{n} y_i\right] \\
&= \frac{1}{n^2}\text{Var}\left[y_1 + y_2 + \cdots + y_n\right] \\
&= \frac{1}{n^2}\left(\text{Var}\left[y_1\right] + \text{Var}\left[y_2\right] + \cdots + \text{Var}\left[y_n\right]\right) \\
&= \frac{1}{n^2}\left(\sigma_y^2 + \sigma_y^2 + \cdots + \sigma_y^2\right) \\
&= \frac{n\sigma_y^2}{n^2} = \frac{\sigma_y^2}{n}
\end{aligned}
$$

   The Gauss-Markov theorem shows that this variance is the smallest among all linear unbiased estimators for the population mean (provided the sample is i.i.d.). This variance is used in the formulas for z- and t-statistics, and for confidence intervals, which can all be used for hypothesis testing.

4. The z-test requires that the population variance is known. The z-statistic follows the standard Normal distribution (provided the null hypothesis is true). The t-statistic replaces the unknown population variance with an estimator (s^2). This changes the distribution of the test statistic so that it is no longer Normal; it know follows the t-distribution.

5. By minimizing the sum of squared residuals, we end up with an estimator that is unbiased, efficient, and consistent (under certain assumptions made about the model and data).

6. R^2 would equal 0 if the slope estimator would be calculated to be exactly 0. R^2 would equal 1 if the data points were on a straight line so that the LS fitted line passed exactly through each data point.

7. a)

$$\bar{Y} = 2\tfrac{1}{3}, \quad \bar{X} = -1\tfrac{1}{3}$$

$$b_1 = \frac{(0-2\frac{1}{3})(1--1\frac{1}{3}) + (2-2\frac{1}{3})(-1--1\frac{1}{3}) + (5-2\frac{1}{3})(-4--1\frac{1}{3})}{(1--1\frac{1}{3})^2 + (-1--1\frac{1}{3})^2 + (-4--1\frac{1}{3})^2}$$

$$= -1$$

$$b_0 = 2\tfrac{1}{3} - -1 \times -1\tfrac{1}{3} = 1$$

   b) $\hat{Y}_2 = 1 + -1(-1) = 2$. $e_2 = 2 - 2 = 0$.

8. $\hat{\sigma}_y^2$ is a biased estimator, whereas s_y^2 is unbiased.

9. a) $\mathrm{E}[\tilde{\mu}_y] = \mathrm{E}\left[\frac{y_1+y_n}{2}\right] = \frac{\mu_y+\mu_y}{2} = \mu$. The estimator is unbiased.

   b) $\mathrm{Var}[\tilde{\mu}_y] = \mathrm{Var}\left[\frac{y_1+y_n}{2}\right] = \frac{\sigma_y^2+sigma_y^2}{4} = \frac{\sigma_y^2}{2}$. This variance is greater than the variance of $\bar{y}$, so it is not efficient. Only $\bar{y}$ is efficient - we know this from the Gauss-Markov theorem.

   c) This estimator is not consistent. The variance does not go to 0 as $n \to \infty$.

10. a) Quantity demanded decreases by 0.74 on average.

   b) The model predicts that quantity demanded will be $12.34 - 0.74(5) = 8.64$.

   c) Price explains 93% of the variation in quantity demanded. Some other factors that might determine quantity demanded are incomes, expectations, advertising, prices of complements and substitutes, etc.

   d) $b_1 \pm 1.96 \times s.e.(b_1) = -0.74 \pm 1.96 \times 0.024 = [-0.79, -0.69]$. The null hypothesis that $\beta_1 = 0$ is rejected at the 5% significance level; it does not lie inside the 95% confidence interval.

11. a) $b_1 = -0.85$ is the estimated difference. In this sample, women make \$0.85 less per hour than men, on average.

   b) $b_0 = 11.24$ is the sample mean wage for men, $b_0 + b_1 = 10.39$ is the sample mean wage for women.

   c) This amounts to testing $H_0 : \beta_1 = 0$ vs. $H_A : \beta_1 \neq 0$. The t-statistic for this test is $t = \frac{-0.85}{0.56} = -1.52$. Using the standard Normal table, the p-value is $2 \times 0.0643 = 0.1286$. We fail to reject the null hypothesis that there is no difference in the wages of men and women, at the 10% significance level.

   d) $b_0 = 10.39$, $b_1 = 0.85$.

END