BUEC 333: Some MORE Study Questions for the Midterm

1. Studenmund, Ch4, Q 2,7,11a,11c; Ch5 Q1,2,4,6

2. [Final Exam, Spring 2009] Suppose you want to test the following hypothesis at the 5% level of significance:

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

Which of the following statements is/are **true**?

a) the probability of a Type I error is 0.025
b) the probability of a Type I error is 0.05
c) the probability of a Type II error is 0.95
d) the probability of a Type II error is 0.05
e) none of the above

3. [Midterm, 2006] If the covariance between two random variables X and Y is zero then
   a. X and Y are independent
   b. Knowing the value of X provides no information about the value of Y
   c. $E(X) = E(Y) = 0$
   d. a and b are true
   e. none of the above

4. [Sample Question, 2009] To be useful for hypothesis testing, a test statistic must:
   a. be computable using sample data
   b. have a known sampling distribution when the null hypothesis is true
   c. have a known sampling distribution when the null hypothesis is false
   d. all of the above
   e. a and b only

5. [Sample Question 2009] Suppose you draw a random sample of n observations, $X_1, X_2, ..., X_n$, from a population with unknown mean μ. Which of the following estimators of μ is/are **biased**?
   f) the first observation you sample, X_1
   g) $\sqrt{\overline{X}^2}$
   h) $\sqrt{\overline{X}^2 - s^2/n}$
   i) b and c
   j) a, b, and c

6. [Midterm, Spring 2009] Which of the following is a linear regression model:
   a. $Y_i = \alpha + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i$
   b. $\log(Y_i) = \beta_0 + \beta_1 \log(X_i) + \varepsilon_i$
   c. $Y_i = \beta_0 + \beta_1 e^{X_i} + \varepsilon_i$
   d. all of the above
   e. none of the above

7. [Midterm , Spring 2009] Suppose you have a random sample of 100 SFU students. In response to a short survey, each student reported the usual number of hours per week that he/she spent working off-campus (X_i), and the usual number of hours per week that he/she spent engaged in social activities (Y_i). The university has hired you to analyze these data. Their main interest is the total number of hours per week that SFU students engage in non-academic activities. Define a new variable, $Z_i = X_i + Y_i$, to measure the number of hours that SFU students engage in non-academic activities. Suppose that in the population of SFU students:

$$E[X_i] = \mu_X, E[Y_i] = \mu_Y, Var[X_i] = \sigma_X^2, Var[Y_i] = \sigma_Y^2, Cov[X_i, Y_i] = \sigma_{XY}.$$

   a. The university has asked you to estimate the average number of hours that SFU students engage in non-academic activities, i.e., $E[Z_i] = \mu_Z$. Your roommate, who has already taken BUEC 333, says "That's easy! Just take the first person in your sample. Their value, Z_1, is an unbiased estimator of μ_Z." Is your roommate right? Explain.
   b. What is the variance of your roommate's estimator in part a?
   c. Give a more efficient, but still unbiased, estimator of μ_Z. Don't forget to **show** that it is unbiased, and that it is more efficient than your roommate's estimator!
   d. What is the variance of the sampling distribution of your estimator in part d? Explain what the variance of the sampling distribution measures.

8. Show that the sample mean is an unbiased estimator of the population mean. Be careful to state your assumptions and show the mathematical steps that are needed. Show that the sample mean of half of your sample is an unbiased estimator of the population mean.

9. Why does s-squared have *n-1* in its denominator rather than *n*?

10. What is the 5% critical value for a *t-test* when *n=100*. What about for a *z-test*, which is distributed standard normal? Why are these so close?

11. How are critical values, type I and type II errors, and confidence intervals related to each other?

12. Regression equations contain error terms. How is this related to type I and type II errors?

13. Regression equations contain error terms. What error is made by the regression? Do error terms matter? Why or why not? Qualify your answer if you need to.