

Review Questions for the Final Exam

Note: Bluebooks will be used for the final exam. These questions are to help you review for the final. This part of the final is open-book and open-notes. Be sure to bring a calculator and your book. In order to get full credit, you must show the calculations used to arrive at your answers. Draw any diagrams clearly. Show at least two significant digits in final answers.

- (1) One stage of a production process requires a delicate adjustment. If the machinery for this operation has been properly adjusted, 10% of the units produced will be defective; but if not, 40% will be defective. In the past, the proper adjustment has been made 85% of the time.
- Suppose that a unit is tested and found defective. What is the chance the machinery has been properly adjusted?
 - Suppose that two units are tested, and one is found defective and the other is not. Now what is the chance that the machinery has been properly adjusted?
- (2) In a sales training program, each participant tries to make a simulated sale twice in a row. Each attempt can result in either success S or failure F. On the first try, the chance of S is 70%. If the first outcome was S, there is now a 90% chance of S on the second trial. If the first outcome was F, there is only a 50% chance of S on the second trial. In other words, "success encourages success."
- Find the probability distribution of X = the number of successes ($X = 0, 1, 2$).
 - Find the mean and standard deviation of X .
- (3) To reduce conflicts among its junior executives, a large corporation tried out two programs (A and B) that taught personal skills. From among their newly hired MBAs, they randomly selected 4 for Program A and 4 for Program B. One year later, each person was ranked by peers and superiors according to openness, cooperation, integrity, etc., to obtain an overall score. Ranked from best to worst, the scores turned out as follows:

<u>A</u>	<u>B</u>
60	72
50	66
46	52
44	50

- Calculate the appropriate 95% confidence interval and state in simple language what it means.
- Calculate the two-sided p-value for H_0 and state what it means. Are the two programs statistically discernible at level $\alpha = 5\%$?

- (4) A random sample of five boys had their heights (in inches) measured at age 4 and again at age 18, with the following results:

<u>Boy</u>	<u>age 4</u>	<u>age 18</u>
Tom	40	68
Dick	43	74
Harry	40	70
Chip	40	68
Dale	42	70

- a. Suppose your nephew Jack was 42 inches tall at age 4. Would you predict he would be 70 inches tall at age 18 (like the last boy in the sample)? Why or why not?
- b. Construct a 95% confidence interval for what Jack's height will be at age 18.
- (5) In the consumer price index (CPI), housing is an important ingredient. And as housing ages, its value and rent decrease, which should be allowed for in constructing the CPI. A researcher studying this issue sampled 10,000 rental housing units and estimated the following regression equation (where the standard error for each coefficient is shown below it in parentheses):

$$\text{LOGRENT} = 5.2 - .009*(\text{AGE}) + .18*(\text{BR}) - .0041*(\text{UNEMP}) + \dots$$

(.0014) (.010) (.0037)

where LOGRENT = log of the monthly rent (\$)
AGE = age of the unit (years)
BR = number of bedrooms in the unit
UNEMP = unemployment rate in the neighborhood (percent)

Estimate how much more or less the monthly rent would be (in percentage terms) for an apartment that is just like another except that:

- a. it has one more bedroom
- b. it is in a neighborhood with more unemployment (8 percent, instead of 5 percent)
- c. It is 10 years older
- d. it is 10 years older and also is in a neighborhood with more unemployment (8 percent, instead of 5 percent)
- e. Construct a 95% confidence interval for your answer to (a)

- (6) A random sample of 5 baseball players was drawn out of all the National League players who came to bat at least 100 times in both the 1979 and 1980 seasons. Their batting averages (rounded) were as follows:

<u>Player</u>	<u>1979 (X)</u>	<u>1980 (Y)</u>
Milner	.270	.300
Mazzilli	.300	.280
Schmidt	.250	.290
Kennedy	.280	.250
Cruz	.290	.300

- a. Calculate the correlation r between the 1979 and 1980 performances. Then find the 95% confidence interval for the population correlation ρ .
- b. Calculate the regression of Y against X . Then find the 95% confidence interval for the population regression coefficient. Is there a statistically discernible linear relation between Y and X at level $\alpha = 5\%$?
- (7) The multiple regression of average annual hay yield Y against temperature T and rainfall R was computed from 20 successive years of data (where the standard error for each coefficient is shown below it in parentheses):

$$Y = 9.14 + .036T + 3.4R$$

(.009) (.70)

- a. Assuming these 20 years form a random sample from the population of many, many years, what is the 95% confidence interval for the coefficient of temperature T ?
- b. What is the p -value for the null hypothesis that it is 0? Is the coefficient statistically discernible from 0 at level $\alpha = 5\%$?
- c. In what ways may your assumption in (a) be wrong?
- (8) An economist wishes to estimate the following model, in which there are 4 mutually dependent Y variables and 2 predetermined X variables (where each $*$ represents a coefficient to be estimated from the data):

$$Y_1 = *Y_3 + *Y_4 + *X_1$$

$$Y_2 = *Y_1 + *X_1 + *X_2$$

$$Y_3 = *Y_2 + *X_2$$

$$Y_4 = *Y_2 + *Y_3 + *X_1 + *X_2$$

Which equations are not identified?