KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DEPARTMENT OF MATHEMATICAL SCIENCES
DHAHRAN, SAUDI ARABIA

STAT 201: INTRODUCTORY STATISTICS

Final Exam, Term 121
Sunday January 6, 2013

Student Name: ID#

You are allowed to use electronic calculators and other reasonable writing accessories that help write the exam. Try to define events, formulate problem and solve.

Do not keep your **mobile** with you during the exam, turn off your mobile and leave it aside.

Question No	Full Marks	Marks Obtained
1	30	
2	20	
3	7	
4	6	
5	15	
6	5	
7	7	
8	10	
Total	100	

Note: You may assume $\alpha = 0.05$ for testing and 95% for confidence interval estimation if not otherwise stated.

Q1. The owner of restaurant **A** wants to study the characteristics of his customers. He decides to focus on two variables:

- The amount of money spent by customers
- Whether the customer's order dessert.

The results from a sample of 62 customers are as follows:

- Amount spent: the mean = $40.11 and the standard deviation = $ 7.41
- 20 customers purchased dessert

a. (*7 points*) Is there evidence that the population mean amount spent per customer in the restaurant is less than $42.11?

1. The hypothesis:

2. The test statistics

3. The decision rule and the critical value (s)

4. The decision

5. The conclusion

b. (*1 point*) Do you need any assumptions? If yes, what? If no, why?

c. *(5 points)* Construct and interpret a 90% confidence interval estimate of the population proportion of customers who purchase dessert.

The owner of a competing restaurant **B** wants to conduct a similar survey in his restaurant.

d. (*2 points*) Based on his competitors information he tries to find the sample size he needs to have 95% confidence to estimate the population mean amount spent in his restaurant to within $\pm$\$2.7. What is the sample size?

He decides to sample the number he found above, and the results are

- Amount spent: the mean = \$35.32 and the standard deviation = \$ 15.65
- 20 customers purchased dessert

e. (*8 points*) Do you think that there is no difference between the percentages of the customers who purchased dessert in both restaurants? Test using the p – value approach

1. The hypothesis:

2. The test statistics

3. The decision rule

4. The decision

5. The conclusion

f. (*7 points*) Construct a 99% confidence interval for the difference in the mean spent in both restaurants. Based on this confidence interval, what conclusion can you draw about the mean amount spent in restaurant **B** compared to the mean amount spent in restaurant **A**?

Q2. A candy bar manufacturer is interested in trying to estimate how sales are influenced by the price of their product. To do this, the company randomly chooses 6 small cities and offers the candy bar at different prices. Using candy bar sales as the dependent variable, the company will conduct a simple linear regression on the data below:

City	River Falls	Hudson	Ellsworth	Prescott	Rock Elm	Stillwater
Price ($)	1.3	1.6	1.8	2	2.4	2.9
Sales	100	90	90	40	38	32

You may use the following information's

$n = 6, \sum x = 12, \ \sum x^2 = 25.66, \ \sum y = 390, \sum y^2 = 30268, \sum xy = 700$

a) (*8 points*) If the price of the candy bar is set at $2, the estimated average sales will be?

b) (*3 points*) What is the coefficient of correlation for these data? What does this tell you about the relation between the price and the sales

c) (*9 points*) Find a 99% confidence interval for the true slop. Use you answer to test whether a change in price will have any impact on average sales

Q3: 80% of all customers applying for a loan at a bank are accepted. Suppose that 50 new loan applications are selected at random.

a) (*2 points*) Find the expected value and the standard deviation of the number of loans that will be accepted by the bank.

b) (*5 points*) What is the probability that at least 42 loans will be accepted? Justify your method.

Q4: A survey indicates that 50% of Khobar residents own a home, 80% own a car, and 90% of the homeowners also own a car. What proportion of residents

a) (*4 points*) own both a car and a house?

b) (*2 points*) own a car or a house, or both?

Q5: The following data give the number of cars owned for a population of 4 families

Family	A	B	C	D
Number of Cars Owned	1	2	3	4

a) (*3 points*) Find the mean and the standard deviation for the population.

b) (*6 points*) List all possible samples of 2 families that can be selected without replacement from this population, and compute the sample mean $\bar{X}$ for each sample.

c) (*3 points*) Find the sampling distribution of $\bar{X}$.

d) (*3 points*) Use the sampling distribution and calculate the mean and the standard deviation of $\bar{X}$.

Q6: An official from the securities commission estimates that 65% of all investment bankers have profited from the use of insider information. If 10 investment bankers are selected at random from the commission's registry, find the probability that:

a) (*3 points*) at most 2 have profited from insider information.

b) *(2 points*) at least 9 have profited from insider information.

Q7: A certain brand of flood lamps has a lifetime that is normally distributed with a mean of 3,750 hours and a standard deviation of 300 hours.

a) *(4 points*) What proportion of these lamps will last between 3000 and 4000 hours?

b) (*3 points*) What lifetime should the manufacturer advertise for these lamps in order that only 2% of the lamps will burn out before the advertised lifetime?

Q8. (*10 points)* Select the true answer

1. If, as a result of a hypothesis test, you reject the null hypothesis when it is false, then you have committed
   a. No error.
   b. An acceptance error
   c. A Type I error
   d. A Type II error
   e. You can't tell depend the information given.

2. You have created a 95% confidence interval for μ with result $(10, 15)$. What decision will you make if you test $H_0 : \mu = 16$ *vs* $H_A : \mu \neq 16$ at $\alpha = 0.1$?
   a. Don't reject H_0 in favor of H_A
   b. Reject H_0 in favor of H_A
   c. Reject H_A in favor of H_0
   d. Fail to reject H_A in favor of H_0
   e. We can't tell what our decision will be from the information given

3. Suppose you wish to test $H_0 : \mu \leq 47$ *vs* $H_A : \mu > 47$. What will the result if we conclude that the mean is greater than 47 when the true value is really 52?
   a. We have made a Type II error.
   b. We have made a Type I error.
   c. We have made a correct decision.
   d. You can't tell depend the information given.
   e. None of the above true.

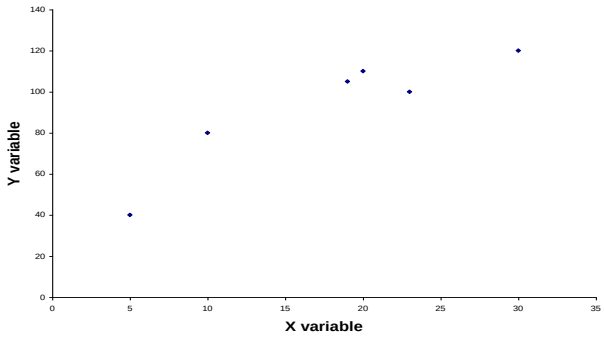
4. When testing $H_0 : \mu_1 - \mu_2 = 0$ *vs* $H_A : \mu_1 - \mu_2 \neq 0$, the observed value of z – score was – 2.13. The p – value for the test would be
   a. 0.9834
   b. 0.0166
   c. 0.9668
   d. 0.0332
   e. None of the above

5. The sample correlation coefficient between X and Y is 0.375. It has been found out that the p – value is 0.256 when testing $H_0 : \beta_1 = 0$ *vs* $H_A : \beta_1 \neq 0$. To test $H_0 : \beta_1 \leq 0$ *vs* $H_A : \beta_1 > 0$ at significance level of 20%, the p – value is?
   a. We can't find the p – value depend the information that we have.
   b. 0.256.
   c. 0.872.
   d. 0.375.
   e. 0.128.

6. If the dependent variable increases as the independent variable increases in an estimating equation, the coefficient of correlation will be in the range:
   a. – 0.05 to 0
   b. – 1 to 0
   c. – 0.05 to + 0.5
   d. 0 to +1
   e. None of the above.

7. After taking a sample and computing $\overline{x}$, a statistician says," I am 95 percent confident that the population mean is between 106 and 122." What does he really mean? :
   a. The probability is 0.95 that μ is between 106 and 122.
   b. The probability is 0.95 that $\mu = 114$, the midpoint of the interval.
   c. 95 percent of the intervals calculated from samples of this size will contain the population mean.
   d. None of the above.
   e. a and b but not c

8. Consider the following chart. Which of the following statements is most correct?
   a. There is a negative linear relationship between the two variables.
   b. There is a positive linear relationship between the two variables.
   c. There is no apparent relationship between the two variables.
   d. There is a perfect linear relationship between the two variables.
   e. None of the above.



9. A major retail store has studied customer behavior and found that the distribution of time customers spend in a store per visit is symmetric with a mean equal to 17.3 minutes. Based on this information, which of the following is true?
   a. The median is approximately 17.3 minutes.
   b. The distribution is bell-shaped.
   c. The median is to the right of the mean.
   d. None of the above.
   e. a and c but not b

10. if the coefficient of determination is equal to $R^2 = 64\%$ and the best fit equation is $\hat{y} = 0.4 - 2x$, then the correlation equal to
    a. 0.8
    b. -0.8
    c. 064
    d. -0.64
    e. $|0.64|$