DSC 330 – Business Statistics

Practice Mid-term Exam

Note that these questions are based on different situations to those in the actual exam.

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question. Write the letter corresponding to your choice.

1) In the construction of confidence intervals, if all other quantities are unchanged, an increase in the sample size will lead to a _____ interval.

A) wider B) less significant C) biased D) narrower

2) For simple linear regression, which of the following four assumptions concerning the probability distribution of the random error term is stated incorrectly?

A) The errors are independent.
B) The distribution is normal for each value of X.
C) The variance of the distribution increases as X increases.
D) The mean of the distribution is 0 for each value of X.

(Situation 1) Suppose a 95% confidence interval for the population mean, E(Y), turns out to be (1000, 2100).

3) Give a definition of what it means to be "95% confident" here.

A) 95% of the observations in the entire population fall in the given interval.
B) In repeated sampling, 95% of the intervals constructed would contain the population mean.
C) 95% of the observations in the sample fall in the given interval.
D) In repeated sampling, the population mean would fall in the interval (1000, 2100) 95% of the time.

(Situation 2) To help consumers assess the risks they are taking, the Food and Drug Administration (FDA) publishes the amount of nicotine found in all commercial brands of cigarettes. A new cigarette has recently been marketed. The FDA tests on this cigarette gave a mean nicotine content of 26.4 milligrams and standard deviation of 2.0 milligrams for a sample of n = 9 cigarettes. You may find some of the following information useful: the 90th percentile of the t-distribution with 8 degrees of freedom is 1.397; the 95th percentile of the t-distribution with 8 degrees of freedom is 1.860.

4) Construct a 90% confidence interval for the mean nicotine content of the brand of cigarette described above.

A) 26.4 ± 1.240 B) 26.4 ± 2.794 C) 26.4 ± 0.931 D) 26.4 ± 3.720

5) The FDA claims that the mean nicotine content exceeds 25 milligrams for the brand of cigarette described above and their stated reliability is 95%. Do you agree?

A) No, since t = (26.4-25)/(2/3) = 2.10 is greater than 1.860.
B) No, since t = (26.4-25)/(2/3) = 2.10 is less than 5%.
C) Yes, since t = (26.4-25)/(2/3) = 2.10 is greater than 1.860.
D) Yes, since t = (26.4-25)/(2/3) = 2.10 is less than 5%.

(Situation 3) Researchers have claimed that the average number of headaches during a quarter of Statistics is 14. Statistics professors dispute this claim vehemently. Statistics professors believe the average is much more than this. They sample n = 13 students and find the sample mean is 16 and the sample standard deviation is 2.0.

6) Which of the following represent the null and alternative hypothesis that the professors wish to test?

A) NH: $E(Y) > 14$ vs. AH: $E(Y) = 14$ B) NH: $E(Y) = 14$ vs. AH: $E(Y) > 14$
C) NH: $E(Y) = 14$ vs. AH: $E(Y) < 14$ D) NH: $E(Y) = 14$ vs. AH: $E(Y) \neq 14$

7) The test statistic for this test is:

A) t = 3.606 B) t = –1.000 C) t = –3.606 D) t = 1.000

(Situation 4) Consider the following printout.

NULL HYPOTHESIS: Pop. MEAN of Y = 3

Y	**=**	**gpa**
SAMPLE MEAN OF Y	**=**	**2.943**
SAMPLE VARIANCE OF Y	**=**	**0.227**
SAMPLE SIZE OF Y	**=**	**167**
t-statistic	**=**	**–1.546**

----

8) Suppose a two-tailed test for the alternative hypothesis that the population mean is not equal to 3 is desired. Find upper and lower limits for the p-value for the test. You may find some of the following information useful: the 90th percentile of the t-distribution with 166 degrees of freedom is 1.28; the 95th percentile of the t-distribution with 166 degrees of freedom is 1.65.

A) $.9 < p < .95$ B) $.05 < p < .1$ C) $.025 < p < .05$ D) $.1 < p < .2$

(Situation 5) A goal of a neighborhood association was to discover if the average sale price of the homes in the neighborhood changed in the last 18 months. A study conducted 18 months ago indicated that the average sale price of neighborhood homes was $80,000. Data was collected and the following printout generated.

NULL HYPOTHESIS: Pop. MEAN of Y = ?

Y	**=**	**Sale_Price**
SAMPLE MEAN OF Y	**=**	**89,280**
SAMPLE SIZE OF Y	**=**	**28**
t-statistic	**=**	**2.263**
TWO-TAILED P-VALUE	**=**	**0.032**

----

9) Identify the missing value in the null hypothesis statement in this problem.

A) 0.016 B) 89,280 C) 80,000 D) 2.263

10) Specify the rejection region for conducting a two-tailed test at significance level 10%. You may find some of the following information useful: the 90th percentile of the t-distribution with 27 degrees of freedom is 1.314; the 95th percentile of the t-distribution with 27 degrees of freedom is 1.703.

A) Reject NH if $t < -1.703$ or $t > 1.703$. B) Reject NH if $t > 1.314$.
C) Reject NH if $t < -1.314$ or $t > 1.314$. D) Reject NH if $t < -2.263$ or $t > 2.263$.

11) Based on the information presented in this printout, which of the following conclusions would be correct if conducting a two-tailed test at significance level 10%?

A) Accept NH B) Reject NH C) Fail to reject NH D) Reject AH

(Situation 6) Civil engineers often use the straight-line equation

$$\mathbf{E(Y) = b_0 + b_1 X}$$

to model the relationship between mean shear strength, E(Y), of masonry joints and precompression stress, X. To test this theory, a series of stress tests were performed on solid bricks arranged in triplets and joined with mortar. The precompression stress was varied for each triplet and the ultimate shear load just before failure (called the shear strength) was recorded. The stress results for n = 7 triplet tests is shown in the accompanying table followed by a SPSS printout of the regression analysis.

Triplet Test	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Shear Strength, Y (tons)	**1.00**	**2.18**	**2.24**	**2.41**	**2.59**	**2.82**	**3.06**
Precomp. Stress, X (tons)	**0**	**0.60**	**1.20**	**1.33**	**1.43**	**1.75**	**1.75**

R Square = 0.905 Std. Error of the Estimate (s) = 0.224

Unstandardized Coefficients

	B	**Std. Error**	**t**	**Sig.**
(Constant)	**1.192**	**0.185**	**6.442**	**0.0013**
X	**0.987**	**0.143**	**6.909**	**0.0010**

---

12) Give a practical interpretation of the estimate of the slope of the least squares line in the situation above.

A) For every .987 ton increase in precompression stress, we estimate the shear strength of the joint to increase by 1 ton.
B) For every 1 ton increase in precompression stress, we estimate the shear strength of the joint to increase by .987 ton.
C) For a triplet test with a precompression stress of 0 tons, we estimate the shear strength of the joint to be 1.192 tons.
D) For a triplet test with a precompression stress of 1 ton, we estimate the shear strength of the joint to be .987 ton.

13) Give a practical interpretation of the estimate of the Y-intercept of the least squares line in the situation above.

A) For a triplet test with a precompression stress of 0 tons, we estimate the shear strength of the joint to be 1.192 tons.
B) The point at which the line crosses the X-axis.
C) No practical interpretation since a triplet test with a precompression stress of 0 tons is outside the range of the sample data.
D) For every 1 ton increase in precompression stress, we estimate the shear strength of the joint to increase by .987 ton.

14) For the situation above, give a practical interpretation of s, the estimate of the standard deviation of the random error term in the model.

A) We expect about 95% of the observed shear strength values to lie within .448 ton of their least squares predicted values.
B) We expect about 95% of the observed shear strength values to lie within .224 ton of their least squares predicted values.
C) About 44% of the total variation in the sample of Y-values can be explained by (or attributed to) the linear relationship between shear strength and precompression stress.
D) We expect about 95% of the observed shear strength values to lie on the least squares line.

15) For the situation above, give a practical interpretation of R^2, the coefficient of determination for the least squares model.

A) About 90.5% of the total variation in the sample of Y-values can be explained by (or attributed to) the linear relationship between shear strength and precompression stress.
B) We expect to predict shear strength of a triplet test to within about .905 ton of its true value.
C) We expect about 90.5% of the observed shear strength values to lie on the least squares line.
D) In repeated sampling, approximately 90.5% of all similarly constructed regression lines will accurately predict shear strength.

16) What set of hypotheses would you test to determine whether shear strength is positively linearly related to precompression stress in the situation above?

A) NH: $b_0 = 0$ vs. AH: $b_0 > 0$
B) NH: $b_1 = 0$ vs. AH: $b_1 < 0$
C) NH: $b_1 = .987$ vs. AH: $b_1 > .987$
D) NH: $b_1 = 0$ vs. AH: $b_1 > 0$

17) In the situation above, is there sufficient evidence of a positive linear relationship between shear strength and precompression stress? Use significance level 5%.

A) No, since the coefficient of determination exceeds 90%.
B) Yes, since the p-value of the test is less than .05.
C) No, since the p-value of the test exceeds .05.
D) No, since the p-value of the test is less than 1.

TRUE/FALSE. Write 'T' if the statement is true and 'F' if the statement is false.

18) The smaller the p-value, the less likely you are to reject the null hypothesis.

19) The simple linear regression model allows the E(Y) values (expected or predicted Y-values) to fall around the regression line while the actual values of Y must fall on the line.

20) The coefficient of correlation is a useful measure of the linear relationship between two variables.