國　立　清　華　大　學　命　題　紙

九十二學年度 科技管理 （所） 科管乙 組碩士班研究生招生考試

科目 統計學 科號 5801 共 4 頁第 1 頁 *請在試卷【答案卷】內作答

一、問答題(50%)

1. 請說明何謂中央極限定理(Central Limit Theorem)。(10 分)

2. 說明何謂自由度(Degree of Freedom)。及爲什麼樣本變異數 $\sum_{i=1}^{n}(x_i - \bar{x})^2/n-1$ 的自由度爲 n-1?(10 分)

3. 常態分配(Normal Distribution)的可能值範圍爲(-∞，∞)。身高明明不可能爲負值，爲何可以用常態分配作爲身高的機率模型？(10 分)

4. A psychologist speaking to a meeting of American Association of University Professors recently said, "The evidence suggests that there is nearly correlation zero between teaching ability of a faculty member and his or her research productivity." The student newspaper reported this as "Professor McDaniel said that good teachers tend to be poor researchers and good researchers tend to be poor teachers."

   Explain what (if anything) is wrong with the newspaper s report. If the report is not accurate write your own plain-language account of what the speaker meant. (10 分)

5. Measurements in large samples show that the correlation

   (a) between father's height and son's adult height is about _____

   (b) between husband's height and wife's height is about _____

   (c) between a male's height at age 4 and his height at age 18 is about _____

   The answers (in scrambled order) are

   $$\gamma = 0.25,\ \gamma = 0.5,\ \gamma = 0.8.$$

   Match the answers to the statements and explain your choice. (10 分)

二、選擇題(50%)：共 10 題，每題 5 分

1. In a one-way ANOVA, if the null hypothesis, $u_i=u$, i=1..n, is rejected, we may conclude that
   A) the population means are all different
   B) the population means are all equal
   C) most of the population means are different
   D) the variation among sample means is greater than the variation within sample means

2. If variable Y is a function of variable X (that is, Y = f(X)), then which one of the following statements is correct?
   A) the correlation coefficient between X and Y is equal to 1 or –1
   B) the correlation coefficient between X and Y is equal to 0
   C) the covariance between X and Y can be positive or negative
   D) the covariance between X and Y is greater than 0

3. A student has n different books, $n = 2m$, and $m \subset N$ (natural number). This student would like to separate these n different books into two groups such that the same volumes of books are assigned to each group (that is, the number of books in each group is equal to m). How many different separations can this student make?
   A) n!/m!
   B) n!/(m!m!)
   C) n!/(m!m!2!)
   D) n!/(m!m!4!)

4. A student would like to evaluate the means of two populations. This student finds that the sample mean for population A is 20, and the variance for population A is 5. While the sample mean for population B is 30, and the variance for population B is 10. Which one of the following statements is correct?
   A) a one-way ANOVA can be used, and the population means are statistically equal
   B) a one-way ANOVA can be used, and the population means are statistically unequal
   C) a one-way ANOVA can be used, and this student needs more information such as F-table and degrees of freedom
   D) a one-way ANOVA can not be used

5. A student would like to study the relationship between two variables, X and Y. If for any value of X, there is one and only one correspondent value of Y. For example, when $X = 1$, Y can only be 10, when $X = 2$, Y can only be 20, and when $X = 3$, Y can only be 30, and so on. Which one of the following statements is correct?
   A) a regression, such as $Y = B_0 + B_1X$, can be used, and B_1 is positive
   B) a regression, such as $Y = B_0 + B_1X$, can be used, and B_1 is negative
   C) a regression, such as $Y = B_0 + B_1X$, can be used, and this student needs to collect more data to determine B_0 and B_1
   D) a regression can not be used

6. For a regression, $Y = B_0 + B_1X + \varepsilon$, ε is a residual. Which one of the following statements about assumptions of regression is correct?
   A) residual is not related to X
   B) residual is not related to Y
   C) residual is not related to X nor Y
   D) the correlation coefficient between X and ε can be positive

7. In a study of the relationship between income and age, a sample regression equation is obtained, in which age is used as the predictor variable (i.e., income $= B_0 + B_1(\text{age}) + \varepsilon$). This sample regression equation is then used to make inferences. Which of the following statements is true?
   A) A 95% confidence interval for the mean income of all subjects at 30 years old will be narrower than a 95% prediction interval for the income of an individual at 30 years old
   B) A 95% confidence interval for the mean income of all subjects at 30 years old will be equal to a 95% prediction interval for the income of an individual at 30 years old
   C) A 95% confidence interval for the mean income of all subjects at 30 years old will be wider than a 95% prediction interval for the income of an individual at 30 years old
   D) A 95% confidence interval for the mean income of all subjects at 30 years old may be wider or narrower than a 95% prediction interval for the income of an individual at 30 years old

8. In a regression analysis, the least-squares method minimizes which of the followings?
   A) total sum of squares
   B) regression sum of squares
   C) error sum of squares
   D) average sum of squares

9. Which of the following statements is correct?
   A) when performing regression analysis, different samples will all yield the same sample regression line
   B) when performing chi-square test of independence. Two variables under consideration are sex and blood type. If the two variables are not associated, we would expect that the number of women with a given blood type would be roughly equal to the number of men with the same blood type.
   C) in the context of regression analysis, the sum of the residuals is always zero
   D) in the context of one-way ANOVA, the expectation value of MSTR, treatment mean square is always greater than the expectation value of MSE, error mean square

10. Which of the following statements is correct?

A) when performing regression analysis, different samples will all yield the same sample regression line

B) when performing chi-square test of independence, two variables under consideration are sex and blood type. If the two variables are not associated, we would expect that the number of women with a given blood type would be roughly equal to the number of men with the same blood type

C) in the context of regression analysis, the sum of the residuals is always zero

D) in the context of one-way ANOVA, the expectation value of MSTR, treatment mean square is always greater than the expectation value of MSE, error mean square