

國立中正大學
115 學年度碩士班招生考試
試題

[第 4 節]

科目名稱	統計學
系所組別	財務金融學系

—作答注意事項—

※作答前請先核對「試題」、「試卷」與「准考證」之系所組別、科目名稱是否相符。

1. 預備鈴響時即可入場，但至考試開始鈴響前，不得翻閱試題，並不得書寫、畫記、作答。
2. 考試開始鈴響時，即可開始作答；考試結束鈴響畢，應即停止作答。
3. 入場後於考試開始 40 分鐘內不得離場。
4. 全部答題均須在試卷（答案卷）作答區內完成。
5. 試卷作答限用藍色或黑色筆（含鉛筆）書寫。
6. 試題須隨試卷繳還。

國立中正大學 115 學年度碩士班招生考試試題

科目名稱：統計學
系所組別：財務金融學系

本科目共 6 頁 第 1 頁

Part I. Multiple-Choice Questions (75%):

Choose the **best** answer (單選題) from the 4 or 5 options provided.

There are 25 questions, worth 3 points each.

1. Which of the following statements is/are correct?
 - I. A sample is a subset of a population.
 - II. One can always draw conclusions about a population using descriptive statistics.
 - III. Inferential statistics enables us to study a sample drawn from a population in order to characterize the population.
 - IV. Hypothesis testing allows us to hypothesize the characteristics of a sample based on the population from which the sample is drawn.

A) I, II, III
B) I, II, III, IV
C) I, III
D) I, III, IV
2. A researcher intends to conduct an M&A event study by collecting data on companies that experienced mergers and acquisitions during the period 2000–2025. What type of data is being used?

A) Cross-sectional data
B) Time-series data
C) Panel data
D) Pooling data
3. 15 students in an economics class receive the following scores:
92 87 99 56 70 65 59 100 78 80 75 89 95 80 60
A 16th student who was absent the day of exam takes the same exam the next day. What score of the sixteenth student would not change the median score of this exam?

A) 44
B) 79
C) 96
D) None of the above
E) All of the above
4. An investor has a portfolio of stocks. The returns of the stocks are 15%, 12.5%, 5%, -3.2% and -6.9%. What is the geometric mean for the rates of return of her portfolio?

A) 7.298%
B) 4.129%
C) 4.480%
D) 9.536%

國立中正大學 115 學年度碩士班招生考試試題

科目名稱：統計學

本科目共 6 頁 第 2 頁

系所組別：財務金融學系

5. When a census reports that the distribution of household income is skewed to the right, this means
 - A) a small number earn very high incomes.
 - B) a long "tail" of extreme values stretching to the right on a graph.
 - C) an existence of income inequality, with the majority clustered at the high end.
 - D) A and B
 - E) All of the above are correct.
6. A correlation coefficient of -1.0 between x and y variables indicates that a correlation between x and y is:
 - A) incorrectly computed.
 - B) very poor.
 - C) perfect.
 - D) nonlinear.
7. A student is taking an exam and answer a multiple-choice question with 5 possible answers. The probability that the student knows the answer is 0.75. If the student answers correctly, what is the probability that the student guessed the answer?
 - A) 0.0625
 - B) 0.2
 - C) 0.05
 - D) None of the above is correct.
8. A 95% confidence interval means that:
 - A) 95% of the population values lie in this interval.
 - B) if we repeatedly take random samples from the population and construct a confidence interval from each sample using the same method, then about 95% of those intervals will contain the true population mean.
 - C) there is a 95% probability that the population mean is in this interval.
 - D) 95% of sample means fall in this interval.
9. In a stepwise regression analysis, what test is used to find the best fitting model?
 - A) t statistics
 - B) χ^2 statistics
 - C) F statistics
 - D) h statistics
10. Which of the following statement is correct?
 - A) A point estimate is preferred over interval estimate from a statistic point of view.
 - B) Statistical inference permits us to address conclusions concerning a sample from the population we drawn.

- C) A statistical test is performed under the assumption that the alternative hypothesis is true.
D) Increasing the sample size reduces the probability of making a type II error.
11. Which of the following statements is/are correct?
- Using the t distribution as opposed to the normal distribution for hypothesis testing under otherwise similar conditions results in more rejections of the null hypothesis.
 - Everything else being equal, a one-tailed test has greater power than a two-tailed test.
 - The researcher would usually prefer to reject as opposed to accept the null hypothesis because only then does one know the probability of making an error.
- A) II
B) I, II
C) I, III
D) I, II, III
12. A small computed χ^2 value indicates that:
- the observed distribution does not fit the theoretical distribution.
 - one classification variable is not related to another classification variable.
 - a variance is significantly different from a hypothesized value.
 - more than one of the above is true.
13. Which of the following statements is/are correct?
- The F distribution is symmetrical such that $F(v_1, v_2, 0.05) = 1/F(v_2, v_1, 0.95)$.
 - The F distribution can be used to test for differences among two variances.
 - The F distribution can be used to test for differences among two population means.
- A) I, II
B) II, III
C) I, III
D) I, II, III
14. The probability of making a Type I error is most directly affected by the:
- sample size.
 - difference between the true and the hypothesized parameter values.
 - population standard deviation.
 - significance level.
15. A city government is considering purchasing a large number of LED streetlights from a new supplier. The supplier claims that the streetlights will last, on average, at least 48 months before needing replacement. The appropriate hypothesis for the government to test would be:
- a two-tailed test.
 - a one-tailed test to the left.

- C) a one-tailed test to the right.
 D) determined by the sample result.
16. A film studio investigated audience response to a new movie trailer. The studio wants to know whether the trailer is equally appealing to college students and non-students. A sample of 200 college students watched the trailer and 60% said they liked it. Of 300 non-students who watched it, 50% indicated a favorable response.
- A) The null hypothesis can be rejected at $\alpha = 0.1$.
 B) The null hypothesis can be rejected at $\alpha = 0.05$.
 C) The null hypothesis can be rejected at $\alpha = 0.01$.
 D) A and B.
 E) A, B and C.
17. In statistics, a result is said to be significant if it is:
- A) unlikely to have occurred by chance under the null hypothesis.
 B) likely to have occurred by chance under the null hypothesis.
 C) important.
 D) real.
18. The larger the absolute value of the t statistic, the more likely we are to:
- A) reject the null hypothesis.
 B) fail to reject the null hypothesis.
 C) conclude the result is not statistically significant.
 D) conclude the result depends on whether the test is one-sided or two-sided.
19. In a simple linear regression, if there is no correlation between x and y then:
- A) $SSR = SST$. (SSR: Sum of Squares Regression, SST: Sum of Squares Total)
 B) the regression line has a negative slope.
 C) the standard error of the estimation is zero.
 D) none of the above are true.
20. To depict a trend using an exponential equation of $\hat{y} = (100) \cdot (0.9)^x$, where x is the year trend variable, indicates:
- A) 90 percent increase in y each year
 B) 10 percent increase in y each year
 C) 90 percent decrease in y each year
 D) 10 percent decrease in y each year
21. Which of the following statements is/are correct?
- I. If all independent variables are uncorrelated, the variance inflation factor will be equal to zero.

- II. Imperfect multicollinearity may not always cause a problem in regression analysis.
- III. The limitation of simply examining the correlation coefficient between any two independent variables is that it ignores the possibility of linear relationships involving more than two variables.
- A) I and II
B) I and III
C) II and III
D) I, II and III
22. OLS is unbiased but not BLUE when:
- A) errors are non-normal.
B) there is heteroskedasticity.
C) there is endogeneity.
D) there is perfect multicollinearity.
23. Using time-series data from 2005 to 2024, consider the following regression model: $DPS = \beta_0 + \beta_1 CRISIS + \beta_2 COVID + \beta_3 EPS + \varepsilon$, where DPS denotes dividends per share; CRISIS equals 1 if the observation is from 2008 and 0 otherwise; COVID equals 1 if the observation is from 2019 and 0 otherwise; and EPS denotes earnings per share. The coefficient β_1 , ceteris paribus, represents:
- A) the expected DPS in 2008.
B) the expected difference in DPS between year 2008 and non-2008 years in the sample.
C) the expected difference in DPS between 2008 and all sample periods except 2019.
D) how the expected DPS changes before and after 2008.
24. To test for the existence of heteroskedasticity in a regression using cross-sectional data, which of the following is used as a proxy for the variance of each observation?
- A) The residuals from the regression.
B) The squares of the residuals from the regression.
C) The variance of the residuals from the regression.
D) Modified data obtained by assigning different weights to data points.
25. Suppose the true population model is $y = \beta_0 + \beta_1 x + \beta_2 z + \varepsilon$ but the researcher estimates the regression as $y = \beta_0 + \beta_1 x + \varepsilon$. Which of the following conditions must hold for the OLS estimator of β_1 to be biased due to an omitted variable?
- A) The omitted variable z affects y and is correlated with x .
B) The omitted variable z affects y but is uncorrelated with x .
C) The omitted variable z is correlated with x but does not affect y .
D) The omitted variable z is neither correlated with x nor affects y .

Part II. Essay Questions (25%):

Show your work clearly and explain your reasoning where appropriate.

If the result is not an integer, round it to **three decimal places**.

1. Consider an urn with 8 balls: 4 of them are white, 3 are black and 1 is red. Now two balls are drawn from the urn. The random variables X and Y are defined as follows:

$$X = \begin{cases} 1 & \text{black ball} \\ 2 & \text{red ball in the first draw} \\ 3 & \text{white ball} \end{cases} \quad Y = \begin{cases} 1 & \text{black ball} \\ 2 & \text{red ball in the second draw} \\ 3 & \text{white ball} \end{cases}$$

- (1) When are X and Y independent – when the two balls are drawn with replacement or without replacement? Give a numerical example to illustrate. (2%)
- (2) What is the correlation coefficient between variables X and Y ? (8%)
2. The following regression is estimated using firm-year data from 2020–2024 (sample size = 465):

$$INS_{it} = -102.12 + 0.47ESG_{it} + 5.51SIZE_{it} - 0.01RD_{it}$$

(5.477) (0.082) (0.460) (0.029)

where INS represents institutional ownership, ESG represents ESG scores, $SIZE$ represents firm size (total assets), and RD represents research and development expenditures. The total sum of squares (SST) is 168,863 and the R-squared is 0.56.

- (1) What is the adjusted R-squared? (2%)
- (2) To test the overall significance of the regression, specify the null and alternative hypotheses using notation (not words) and report the value of the F statistic. (8%)
- (3) With year fixed effects included, the R-squared rises to 0.64. What is the F statistic for testing whether the year dummies are jointly redundant? (5%)