

國立中山大學 115 學年度 碩士班考試入學招生考試試題

科目名稱：統計學【數金所碩士班乙組、國資所碩士班乙組】

— 作答注意事項 —

考試時間：100 分鐘

- 考試開始鈴響前不得翻閱試題，並不得書寫、劃記、作答。請先檢查答案卷（卡）之應考證號碼、桌角號碼、應試科目是否正確，如有不同立即請監試人員處理。
- 答案卷限用藍、黑色筆(含鉛筆)書寫、繪圖或標示，可攜帶橡皮擦、無色透明無文字墊板、尺規、修正液（帶）、手錶(未附計算器者)。每人每節限使用一份答案卷，請斟酌作答。
- 答案卡請以 2B 鉛筆劃記，不可使用修正液（帶）塗改，未使用 2B 鉛筆、劃記太輕或污損致光學閱讀機無法辨識答案者，後果由考生自負。
- 答案卷（卡）應保持清潔完整，不得折疊、破壞或塗改應考證號碼及條碼，亦不得書寫考生姓名、應考證號碼或與答案無關之任何文字或符號。
- 可否使用計算機請依試題資訊內標註為準，如「可以」使用，廠牌、功能不拘，唯不得攜帶書籍、紙張（應考證不得做計算紙書寫）、具有通訊、記憶、傳輸或收發等功能之相關電子產品或其他有礙試場安寧、考試公平之各類器材入場。
- 試題及答案卷（卡）請務必繳回，未繳回者該科成績以零分計算。
- 試題採雙面列印，考生應注意試題頁數確實作答。
- 違規者依本校招生考試試場規則及違規處理辦法處理。

國立中山大學 115 學年度碩士班考試入學招生考試試題

科目名稱：統計學【數金所碩士班乙組、國資所碩士班乙組】
※本科目依簡章規定「不可以」使用計算機(混合題)

題號：490001
共 4 頁第 1 頁

PART I: Answer each question by selecting one of options from A-E (64%, Each question 4%)

1. Consider the linear regression model: $Y_i = \beta_0 + \beta_1 X_i + u_i$, What does BLUE stand for?
 - A. Best Linear Unrestricted Estimator
 - B. Best Linear Unbiased Estimator
 - C. Best Likelihood Unbiased Estimator
 - D. Biased Linear Unbiased Estimator
 - E. Bayesian Linear Unbiased Estimator
2. In the term Best Linear Unbiased Estimator (BLUE), what does “Best” specifically mean?
 - A. The estimator has the smallest bias among all estimators.
 - B. The estimator has the smallest variance among all estimators.
 - C. The estimator has the smallest variance among all linear and unbiased estimators.
 - D. The estimator always equals the true parameter value.
 - E. The estimator performs best when the sample size is large.
3. Which of the following statements correctly describes the meanings of “Linear” and “Unbiased” in BLUE?
 - A. Linear means the regression function is linear in parameters, and unbiased means the estimator converges to the true value as sample size increases.
 - B. Linear means the estimator is a nonlinear transformation of the data, and unbiased means it has minimum variance.
 - C. Linear means the estimator is a linear combination of the observed data, and unbiased means its expected value equals the true parameter.
 - D. Linear means the dependent variable is linear, and unbiased means there is no omitted variable bias.
 - E. Linear means the sample mean is used, and unbiased means the variance is minimized.
4. What is the main objective of Ordinary Least Squares (OLS)?
 - A. To minimize the sum of absolute residuals
 - B. To minimize the sum of squared residuals
 - C. To maximize the likelihood function
 - D. To minimize the variance of the dependent variable
 - E. To eliminate all estimation errors
5. What does R-squared measure?
 - A. The causal effect of X on Y
 - B. The goodness of fit of the regression model
 - C. The unbiasedness of the estimator
 - D. The variance of the error term
 - E. The significance of regression coefficients

國立中山大學 115 學年度碩士班考試入學招生考試試題

科目名稱：統計學【數金所碩士班乙組、國資所碩士班乙組】

題號：490001

※本科目依簡章規定「不可以」使用計算機(混合題)

共 4 頁第 2 頁

6. Let Y_1, Y_2, \dots, Y_n be a simple random sample from a population with variance σ^2 . What is the variance of the sample mean \bar{Y} ?
- A. σ^2
 - B. σ^2 / n
 - C. $\sigma^2 / (n - 1)$
 - D. $n \times \sigma^2$
 - E. $\sigma^2 + n$
7. Which of the following statements about the Central Limit Theorem (CLT) is correct?
- A. The population must be normally distributed.
 - B. The sample size must be infinite.
 - C. The sampling distribution of the sample mean approaches a normal distribution as n increases.
 - D. The sample mean is always normally distributed.
 - E. The variance of the sample mean increases with sample size.
8. Which of the following distributions is NOT a continuous distribution?
- A. Normal distribution
 - B. Exponential distribution
 - C. Uniform distribution
 - D. Poisson distribution
 - E. Gamma distribution
9. Which of the following distributions is NOT a discrete distribution?
- A. Binomial distribution
 - B. Poisson distribution
 - C. Geometric distribution
 - D. Uniform distribution on $(0, 1)$
 - E. Negative binomial distribution
10. Which of the following measures of central tendency is most affected by extreme values (outliers)?
- A. Median
 - B. Mode
 - C. Mean
 - D. Interquartile range
 - E. Range
11. Suppose an extremely large value is added to a dataset. Which of the following statements is correct?
- A. The mean, median, and mode will all remain unchanged.
 - B. Only the median will be affected.
 - C. Only the mode will be affected.
 - D. The mean will definitely change, while the median and the mode may remain unchanged.
 - E. All three measures will definitely change.

國立中山大學 115 學年度碩士班考試入學招生考試試題

科目名稱：統計學【數金所碩士班乙組、國資所碩士班乙組】

題號：490001

※本科目依簡章規定「不可以」使用計算機(混合題)

共 4 頁第 3 頁

12. Which of the following statements about the normal distribution is correct?
A. The normal distribution is skewed to the right.
B. The mean, median, and mode are equal.
C. Most observations lie far from the mean.
D. The normal distribution is defined only for positive values.
E. The variance of a normal distribution is always equal to 1.
13. Let A and B be events with $P(A) = 0.5$, $P(B) = 0.4$, and $P(A \cap B) = 0.2$.
What is $P(A \cup B)$?
A. 0.2
B. 0.5
C. 0.6
D. 0.7
E. 0.9
14. Which of the following correctly defines the conditional probability of A given B?
A. $P(A | B) = P(A) / P(B)$
B. $P(A | B) = P(A \cap B) / P(B)$
C. $P(A | B) = P(A \cup B)$
D. $P(A | B) = P(B | A)$
E. $P(A | B) = 1 - P(A)$
15. Which of the following distributions is commonly used to model the number of occurrences of an event in a fixed interval of time or space, assuming events occur independently?
A. Binomial distribution
B. Hypergeometric distribution
C. Poisson distribution
D. Normal distribution
E. Exponential distribution
16. Which of the following statements is correct?
A. A high R-squared implies all regression coefficients are statistically significant.
B. Individual coefficient significance is tested using an F-test.
C. Dummy variables are used to represent qualitative characteristics in regression models.
D. Predicted values must always lie within the observed range of the dependent variable.
E. Including more regressors always improves the quality of a regression model.

PART II: Short Answer Question (36%, Each question 4%)

Q1. A software company wants to examine the relationship between its online advertising expenditure (measured in units of NT\$100,000) (X) and the number of app downloads (measured in thousands) (Y). From the company's historical data, 10 observations are randomly selected, yielding the following summary statistics:

$$\sum X = 28, \sum X^2 = 303.4, \sum Y = 75, \sum Y^2 = 598.5, \sum XY = 237$$

Find the regression line $Y_i = \beta_0 + \beta_1 X_i + u_i$:

國立中山大學 115 學年度碩士班考試入學招生考試試題

科目名稱：統計學【數金所碩士班乙組、國資所碩士班乙組】

題號：490001

※本科目依簡章規定「不可以」使用計算機(混合題)

共 4 頁第 4 頁

17. What is the value of β_1 ?

18. What is the value of β_0 ?

Q2. Let X be a continuous random variable with probability density function:

$$f(x) = 1/a, \quad 0 < x < a.$$

$$f(x) = 0, \quad \text{otherwise}$$

19. Find its cumulative distribution function.

20. If $E(X) = 0.25$, find the value of a

Q3. A financial institution is evaluating the performance of four different investment strategies. The institution wants to determine whether the average monthly return differs among these four strategies. A random sample of investment portfolios is selected, and the monthly return is recorded. The following analysis of variance (ANOVA) table is obtained:

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F
Investment Strategy	(a)	(c)	104	(e)
Random Error	(b)	(d)		
Total	43	1908		

What is the value of each cell from the table above.

21. What is the value of cell (a) ?

22. What is the value of cell (b) ?

23. What is the value of cell (c) ?

24. What is the value of cell (d) ?

25. What is the value of cell (e) ?

