國立中山大學114學年度
碩士班考試入學招生考試試題

科目名稱：商用統計學【企管系企管甲班碩士班甲組選考、乙組選考、丙組選考】

—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，請衡酌作答。
- 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114學年度碩士班考試入學招生考試試題

科目名稱：商用統計學【企管系企管甲班碩士班甲組選考、乙組選考、丙組選考】題號：441002

※本科目依簡章規定「可以」使用計算機（廠牌、功能不拘）（選擇題）

一、複選題（共20題，每題五分，合計100分。每題有五個選項，答錯k個選項者，得該題全部分數的（5-2k）/5，得分低於零分或所有選項均未作答者，該題以零分計算。

Use the tables of probability distributions in Appendix as needed

1. A sample of 20 packages of ground pork (in pounds) is listed in ascending order as follows:

0.70	0.75	0.80	0.84	0.86	0.89	0.92	0.95	0.98	1.01
1.01	1.07	1.10	1.12	1.15	1.18	1.22	1.27	1.30	1.34

Based on the data above, which of the following statements is/are correct?
(A)The mean of the sample is approximately 1.02 pounds.
(B)The standard deviation of the sample is approximately 0.16 pounds.
(C)The 40th percentile is 1.30 pounds.
(D)The variance of the sample is approximately 0.03 pounds.
(E)The upper quartile is 0.80 pounds.

2. A fair six-sided dice is rolled three times. Based on classical probability principles, which of the following statements is/are correct?
(A) The probability of rolling at least one 6 is 1-125/216.
(B) The probability of rolling three even numbers is 1/8
(C) The probability of rolling a total of 18 is 1/216
(D) The probability of rolling exactly two 6 is 5/216
(E) The probability of rolling a total greater than 4 is 215/216

3. Which of the following descriptions of variables is/are INCORRECT?
(A) "Student ID number" is a nominal variable.
(B) "Bank account balance" is an interval variable.
(C) "Customer satisfaction rating (1 to 5)" is a ordinal variable.
(D) "Number of children in a household" is a ratio variable.
(E) "Type of transportation (car, bus, train)" is a categorical variable.

4. At a university, 15% of undergraduate students study abroad. Among the students who study abroad, 55% are female, while 45% of those who do not study abroad are female. Based on this information, which of the following statements is/are INCORRECT?
(A) The overall percentage of female undergraduate students is 48.00%.
(B) Given that an undergraduate student is female, the probability that she studies abroad is 12.98%.
(C) Given that an undergraduate student studies abroad, the probability that the student is male is 9.45%.
(D) The overall percentage of male undergraduate students is 53.50%.
(E) Given that an undergraduate student is male, the probability that he studies abroad is 6.75%.

5. Which of the following is/are NOT discrete probability distributions?
(A) Poisson probability distribution
(B) Normal probability distribution
(C) Binomial probability distribution
(D) Exponential probability distribution
(E) Chi-square probability distribution

6.Which of the following statements about hypothesis testing is/are correct?
(A) Type I error occurs when the null hypothesis is rejected even though it is true.
(B) Type II error occurs when the null hypothesis is not rejected even though it is false.
(C) The power of a test is the probability of rejecting the null hypothesis when it is false.
(D) Confidence limit refers to the range within which the true population parameter is expected to fall.
(E) Type III error occurs when the null hypothesis is correctly rejected, but for the wrong reason.

7. Customers arrive at a service desk randomly and independently. The probability of an arrival is the same for any interval of equal length. The average time interval between customer arrivals is 2 minutes. Which of the following statements is/are correct?
(A) The probability that no customers will arrive in the next 4 minutes is e^{-2}.
(B) The probability that exactly 2 customers will arrive in the next 6 minutes is $9*e^{-3}/2$.
(C) The average number of customers arriving in the next 10 minutes is 5.
(D) The standard deviation of the arrival rate per 10 minutes is 5.
(E) The average number of customers arriving in the next 8 minutes is 4.

8. Suppose X and Y are two random variables. Which of the following statements about correlation coefficient and covariance is/are correct?
(A) The correlation coefficient between X and Y is always non-negative.
(B) If the covariance between X and Y is zero, then X and Y are independent.
(C) The covariance between X and Y is equal to the product of their standard deviations and the correlation coefficient.
(D) A correlation coefficient of +1 indicates a perfect positive linear relationship between X and Y.
(E) The covariance between X and Y is always greater than or equal to zero.

9. Suppose X is a continuous random variable that follows a uniform distribution on the interval [a,b]. Which of the following statements about the uniform distribution is/are correct?
(A) The probability density function (pdf) of X is constant over the interval [a,b].
(B) The cumulative distribution function (cdf) of X increases linearly from 0 to 1 over the interval [a,b].
(C) The expected value of X is the midpoint of a and b.
(D) The range of X is equal to the difference between b and a.
(E) The probability that X falls outside the interval [a,b] is greater than zero.

10. A researcher repeatedly draws random samples of size n from a population with mean μ and standard deviation σ. Which of the following statements about the sampling distribution of the sample mean $\bar{X}$ is/are correct?
(A) The sampling distribution of $\bar{X}$ is always normal, regardless of the sample size n.
(B) The standard deviation of the sampling distribution of $\bar{X}$ decreases as the sample size n increases.
(C) The mean of the sampling distribution of $\bar{X}$ is equal to $\frac{\sigma}{\sqrt{n}}$
(D) The sampling distribution of $\bar{X}$ has a variance equal to σ^2.
(E) The shape of the sampling distribution of $\bar{X}$ depends only on the population distribution.

11. A company records the monthly electricity usage of its offices, which follows a normal distribution with a mean of 500 kWh and a standard deviation of 50 kWh. Based on this information, which of the following statements about electricity usage is/are correct?
(A) Approximately 68% of the monthly electricity usage falls between 450 kWh and 550 kWh.
(B) Approximately 95% of the monthly electricity usage falls between 450 kWh and 550 kWh.
(C) Approximately 99.7% of the monthly electricity usage falls between 350 kWh and 650 kWh.
(D) About 16% of the offices use more than 550 kWh per month.
(E) The probability of usage being exactly 500 kWh is zero.

12. The following data are from a simple random sample:

5, 8, 10, 7, 10, 14

Based on the data above, which of the following statements are correct?
(A) The point estimate of the population mean is 9.00.
(B) The point estimate of the population standard deviation is approximately 4.10.
(C) The point estimate of the population mean is 8.67.
(D) The range of the sample is 9.
(E) The median of the sample is 8.50.

13. The state of California has a mean annual rainfall of 60 cm, while New York has a mean annual rainfall of 110 cm. The standard deviation for both states is 12 cm. A sample of 25 years of rainfall for California and a sample of 40 years of rainfall for New York has been taken.
Based on the information above, which of the following statements is correct?
(A) The standard error of the sample mean rainfall for California is greater than that for New York.
(B) The difference between the standard error of California and New York is 0.50.
(C) The probability that the sample mean is within 3 cm of the population mean is the same for both states.
(D) The sample size does not affect the standard error of the sample mean.
(E) The larger the sample size, the smaller the standard error of the sample mean.

14. A simple random sample of 60 items resulted in a sample mean of 80. The population standard deviation is σ=15.
Based on this information, which of the following statements are correct?
(A) The confidence interval for the population mean becomes narrower as the sample size increases.
(B) The confidence interval for a sample size of 120 will be wider than that for a sample size of 60.
(C) Increasing the sample size decreases the margin of error for the population mean.
(D) The confidence interval for the population mean is unaffected by changes in sample size.
(E) The larger the sample size, the greater the uncertainty in estimating the population mean.

15. A sample of 100 items is drawn from a population with a standard deviation of σ=25. The sample mean is $\bar{X}$=150. The null hypothesis is that the population mean μ=155. Based on this information, which of the following statements is correct?
(A) The test statistic is −2.0.
(B) The standard error of the mean for this sample is 2.5.
(C) The null hypothesis states that the sample mean is greater than 155.
(D) A negative test statistic indicates that the sample mean is less than the hypothesized population mean.
(E) The null hypothesis does not make any claim about the sample mean.

16. Which of the following statements about ANOVA is correct?
(A) One assumption of ANOVA is that each population has the same mean.
(B) One assumption of ANOVA is that the response variable is normally distributed for each population.
(C) The null hypothesis in ANOVA states that at least one population mean differs from the others.
(D) If ANOVA results indicate no significant difference, a post-hoc test must still be conducted to confirm.
(E) One-way ANOVA is a one-tailed test, while two-way ANOVA is a two-tailed test.

17. Below is the estimated linear regression equation based on 12 observations, where x_1 and x_2 are numerical variables:

$$\hat{y} = 35 + 0.75x_1 + 0.4x_2$$

Where: SST(Total Sum of Squares)=8500; SSR(Regression Sum of Squares)=7650; S_{b1}=0.10; S_{b2}=0.09. The b_1 and b_2 are the estimated values of β_1 and β_1 for the variables x_1 and x_2 respectively; S_{b1}, S_{b2} are the standard errors of the estimated coefficients. Which of the following is/are correct?
(A) R-square =7650/8500
(B) MSR (regression mean square) =7650/2
(C) MSE (mean square error) =850/9
(D) The t-ratio for b_1 is 8.33.
(E) The standard error of the estimate decreases if the model fits the data better.

18. A company conducted an experiment to test whether different training programs have a significant impact on employee productivity. Three training programs (Program A, Program B, and Program C) were assigned to employees, and their productivity scores (out of 100) were recorded after completing the training. The following table summarizes the data collected from the experiment:

Training Program	Productivity Scores
Program A	63, 47, 54, 40
Program B	82, 72, 88, 66
Program C	69, 54, 61, 48

Based on the data above, which of the following statements are correct?
(A) The F-value from the ANOVA test is approximately 7.87.
(B) The p-value for the ANOVA test is approximately 0.011, which indicates that there are no significant differences between the training programs.
(C) The mean productivity scores for Program A and Program B differ significantly.
(D) The mean productivity scores for Program A and Program C differ significantly.
(E) The pooled standard deviation for the data can be used to compute the F-value directly without additional post-hoc tests.

19. Tom is a fitness enthusiast who owns a treadmill at home. He is considering whether to purchase a maintenance contract for his treadmill. Tom believes that the maintenance costs are directly related to the weekly usage of the treadmill. To help him make a decision, Tom collected data on the weekly hours of treadmill usage and corresponding annual maintenance costs (in hundreds of dollars) from other treadmill users in his fitness group:

Weekly Treadmill Usage (hours)	13	10	20	28	32	17	24	31	40
Annual Maintenance Cost (hundreds of dollars)	17.0	22.0	30.0	37.0	47.0	30.5	32.5	39.0	51.5

Based on the data above, which of the following statements are correct?
(A) The intercept of the regression equation is 10.53.
(B) The slope of the regression equation is 1.05.
(C) For a treadmill used 30 hours per week, the predicted annual maintenance cost is \$3913, and the 95% prediction interval is approximately \$(2874, 4952), assuming $t_{0.025}(7)$=2.365.
(D) Tom should purchase a maintenance contract costing \$3000 per year because the predicted annual maintenance cost exceeds \$3000.
(E) The intercept of the regression equation represents the annual maintenance cost when the treadmill is used for 0 hours per week.

20. A coffee shop owner is analyzing the sales of coffee on different days of the week to determine if the sales are evenly distributed across all days. A total of 420 cups of coffee were sold in one week, and the number of cups sold on each day is summarized in the table below:

Day	Cups of Coffee Sold
Sunday	66
Monday	50
Tuesday	53
Wednesday	47
Thursday	55
Friday	69
Saturday	80

Based on the data above, which of the following statements are correct?

$(x^2_{0.05}(6) = 12.59)$

(A) The expected number of cups of coffee sold per day, assuming equal distribution, is 60.

(B) The Chi-square statistic for this test is approximately 14.33.

(C) The p-value for this test is approximately 0.026, indicating that there is no significant difference in coffee sales across the days.

(D) Coffee sales on Saturday are significantly higher compared to the expected value, making it the day with the highest percentage of sales.

(E) At a significance level of 0.05, the null hypothesis that coffee sales are equally distributed across all days is rejected.

Appendix:

The Critical values for F-distribution at (df1, df2) degree of freedom, given alpha=0.05

		Numerator degrees of freedom(df1)				
Denominator degrees of freedom(df2)		1	2	3	4	5
	1	161.45	199.50	215.71	224.58	230.16
	2	18.513	19.000	19.164	19.247	19.296
	3	10.128	9.5521	9.2766	9.1172	9.0135
	4	7.7086	6.9943	6.5914	6.3382	6.2561
	5	6.6079	5.7861	5.4095	5.1922	5.0503
	6	5.9874	5.1433	4.7571	4.5337	4.3874
	7	5.5914	4.7374	4.3468	4.1203	3.9715
	8	5.3177	4.4590	4.0662	3.8379	3.6875
	9	5.1174	4.2565	3.8625	3.6331	3.4817
	10	4.9646	4.1028	3.7083	3.4780	3.3258
	11	4.8443	3.9823	3.5874	3.3567	3.2039
	12	4.7472	3.8853	3.4903	3.2592	3.1059
	13	4.6672	3.8056	3.4105	3.1791	3.0254
	14	4.6001	3.7389	3.3439	3.1122	2.9582
	15	4.5431	3.6823	3.2784	3.0556	2.9013