

國立高雄第一科技大學 97 學年度 碩士班 招生考試 試題紙

系所別：資訊管理系

組別：管理組

考科代碼：3222

考科：統計學

注意事項：

- 1、本科目可使用本校提供之電子計算器。
- 2、請於答案卷上規定之範圍作答，違者該題不予計分。

I、Multiple choice questions (40 points in total)

You would award 4 points for a correct answer in this part.

(10 題選擇題，請將答案寫在答案卷上，答對一題得 4 分，總分 40 分)

1. The t distribution is used when the population distribution is normal and
 - A) The t distribution is used when the population distribution is normal and
 - B) When the sample size is more than 30.
 - C) When sample size is less than 30 and the population standard deviation is not known.
 - D) Both A and B are correct.
2. A package delivery company purchased five trucks from manufacturer A, four from B, and five from C. They then recorded and compared the maintenance cost of each truck. To apply the F test, how many degrees of freedom are in the denominator?
 - A) 2
 - B) 3
 - C) 11
 - D) 14
3. Samples of the wires coming off the production line were tested for tensile strength. The statistical results (in PSI) were:

Arithmetic mean	500	Median	500
Mode	500	Standard deviation	40
Mean deviation	32	Quartile deviation	25
Range	240	Number is sample	100

According to the Empirical Rule, the middle 95 percent of the wires tested between approximately what two values?
 - A) 450 and 550
 - B) 460 and 540
 - C) 420 and 580

- D) 380 and 620
E) None of the above
4. What must be the least scale of measurement for the median?
A) Nominal
B) Ordinal
C) Interval
D) Ratio
E) None of the above
5. If two events A and B are mutually exclusive, what does the special rule of addition state?
A) $P(A \text{ or } B) = P(A) + P(B)$
B) $P(A \text{ and } B) = P(A) + P(B)$
C) $P(A \text{ and/or } B) = P(A) + P(B)$
D) $P(A \text{ or } B) = P(A) - P(B)$
E) None of the above
6. If the decision is to reject the null hypothesis at the 5% level of significance, what are the acceptable alternate hypothesis and rejection region?
A) $\pi_1 \neq \pi_2; z > 1.65 \text{ and } z < -1.65$
B) $\pi_1 \neq \pi_2; z > 1.96 \text{ and } z < -1.96$
C) $\pi_1 > \pi_2; z < -1.65$
D) $\pi_1 > \pi_2; z < -1.96$
E) None of the above
7. If two samples are used in a hypothesis test for which the combined degrees of freedom is 27, which one of the following is true about the two sample sizes?
A) Sample A = 14; sample B = 13
B) Sample A = 12; sample B = 13
C) Sample A = 15; sample B = 14
D) Sample A = 20; sample B = 9
E) Cannot determine from the above information
8. The simple linear regression (least squares method) minimizes:
A) The explained variation
B) SS_{yy}
C) Total variation
D) SS_{xx}
E) SSE

9. In testing the population for significance at a significance level of .05, what is the rejection point condition for the two-sided test?
- A) Reject H_0 if $|t| > 2.571$
 - B) Reject H_0 if $t > 2.571$
 - C) Reject H_0 if $|t| < 2.571$
 - D) Reject H_0 if $|t| > 2.051$
 - E) Reject H_0 if $t > 2.051$
10. In a completely randomized (one-way) analysis of variance problem with c groups and a total of n observations in all groups, the variance between groups is equal to:
- A) (Total sum of squares) - (Sum of squares within columns)
 - B) (Sum of squares between columns) / ($c-1$)
 - C) (Total sum of squares) - [(Sum of squares within columns) / ($n-c$)]
 - D) [(Total sum of squares) / ($n-1$)] - [(Sum of squares between columns) / ($c-1$)]

II. Problems (60 points in total)

(6 題問答題，請將答案寫在答案卷上，各題後面顯示該題分數，總分 60 分)

11. Here are descriptive statistics from Excel for annual per-pupil expenditures in 94 Ohio cities and home sizes in a certain neighborhood. Very briefly compare the *dispersion and shape* of the two data sets. (5 points)

<i>Expenditure Per Pupil (dollars)</i>		<i>Home Size (sq ft)</i>	
Mean	2724.606383	Mean	2231.40857
Median	2508.5	Median	2217
Mode	2506	Mode	2117
Standard Deviation	1095.217916	Standard Deviation	249.3151
Sample Variance	1199502.284	Sample Variance	62158.0191
Kurtosis	43.9989606	Kurtosis	0.52824942
Skewness	6.260249912	Skewness	0.35668642
Range	9310	Range	1315.4
Minimum	1916	Minimum	1593
Maximum	11226	Maximum	2908
Sum	256113	Sum	234297.9
Count	94	Count	105

12. Oakland University has 16,059 students. In a sample of 200 students, 12 were born outside the U.S. Construct a 95% confidence interval for the true population proportion. How large a sample is needed to estimate the true proportion of Oakland students who were born outside the U.S.

with an error of ± 2.5 percent and 95 percent confidence? Show your work and explain fully. (5 points)

13. Suppose that a day's production of 100 manufactured parts contains 10 parts that do not conform to customer requirements. Three parts are selected at random, without replacement from the batch. Let the random variable X equal the number of nonconforming parts in the sample. What is the cumulative distribution function of X ? (Hint: find $F(0)$, $F(1)$, $F(2)$, and $F(3)$) (10 points)

14. What are the key differences between the standard deviation (SD) and the standard error of the mean (SEM), and when each should be used? (Find at least three key differences) (10 points)

15. In hypothesis testing analysis, people would report insignificant ($p > \alpha$) using the terminology like "Accept H_0 ", "Conclude H_0 ", "Fail to reject H_0 ", " H_0 cannot be rejected" or even "The test is insignificant". Which terminology would you use while conducting a hypothesis testing? In addition, which one or ones would not be used by you? What is the logic behind your selection? Please, explain your viewpoints in detail. (10 points)

16. 某經濟學家認為個人年所得(Y)是受教育程度(X)及工作年資(Z)的影響，他隨機抽取 5 個人為樣本，以實證他的看法。試根據該經濟學家所做的檢定方法，以他所得的資訊(如下所示)，回答下列問題：(20 points)

Dependent Variable: Y

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	(<u>1</u>)	114.89464	25.50	0.0377
Error	2	(<u>2</u>)	(<u>3</u>)		
Corrected Total	4	(<u>4</u>)			

Root MSE	2.12258	R-Square	0.9623
Dependent Mean	47.20000	Adj R-Sq	0.9245
Coeff Var	4.49700		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	23.21267	3.61565	6.42	0.0234
X	1	1.36411	0.21594	6.32	0.0242
Z	1	0.91815	0.30046	3.06	0.0925

Obs	Dep Var	Predicted Value	Std Error Mean Predict	95% CL Predict	Residual
1	Y	37.8244	1.6580	26.2356 49.4132	0.1756

第 4 頁，合計 5 頁【尚有試題】

2	40.0000	40.0804	1.4620	28.9910	51.1698	-0.0804
3	53.0000	52.4361	2.0395	39.7707	65.1015	0.5639
4	50.0000	52.3836	1.2548	41.7745	62.9927	-2.3836
5	55.0000	53.2755	1.7017	41.5702	64.9808	1.7245

Sum of Residuals 0
Sum of Squared Residuals 9.01072
Predicted Residual SS (PRESS) 90.92462

Obs	Dep Var Y	Predicted Value	Std Error Mean Predict	95% CL Mean		Residual
1	38.0000	37.8244	1.6580	30.6904	44.9584	0.1756
2	40.0000	40.0804	1.4620	33.7901	46.3707	-0.0804
3	53.0000	52.4361	2.0395	43.6608	61.2114	0.5639
4	50.0000	52.3836	1.2548	46.9848	57.7824	-2.3836
5	55.0000	53.2755	1.7017	45.9538	60.5972	1.7245

Test of First and Second Moment Specification (heteroscedasticity)

DF	Chi-Square	Pr > ChiSq
5	4.75	0.4470

- 根據上述資訊，請問他的統計檢定方法為何？
- 請將空格(1)、(2)、(3)及(4)的數字填上。
- 請根據 Analysis of Variance 報表上的資訊，請詳細寫其虛無假設為何？
- 請寫出其所有參數的估計值，並指出該值是否顯著($\alpha=0.05$)，同時寫出其數學式。
- 請問其 5 個觀測值的預測值各為何，並指其 95%的信賴區間。
- 試寫出其 5 個觀測值的預測平均值之 95%的信賴區間。
- 請寫出其變異數齊一性之虛無假設，並指出是否顯著。
- 其模式的判定係數為何？修正判定係數為何？並請評論該模型的適配度(fit)。