國立成功大學 107 學年度碩士班招生考試試題

系　　所：電信管理研究所

考試科目：統計學

考試日期：0206，節次：2

※ 考生請注意：本試題不可使用計算機。　請於答案卷(卡)作答，於本試題紙上作答者，不予計分。

1. (20%) Consider a new computer system package developed to help systems analysts reduce the time required to design, develop, and implement an information system. To evaluate benefits of the new software package, a random sample of 24 systems analysts is selected. Each analyst is given specifications for a hypothetical information system. Then 12 of the analysts are instructed to produce the information system by using current technology. The other 12 analysts are trained in the use of new software package and then instructed to use it to produce the information system. Let μ_1 refer to the mean project completion time for system analysts by using the current technology and μ_2 to the mean project completion time for system analysts by using the new software package. The researcher is looking for evidence to conclude that μ_2 is less than μ_1. Suppose that the 24 analysts complete the study with the results shown in the following table. Please list the calculation of test statistic.

Summary Statistics	Current Technology	New software
Sample size	$n_1 = 12$	$n_2 = 12$
Sample mean	$\overline{x}_1 = 325$ hours	$\overline{x}_2 = 286$ hours
Sample standard deviation	$s_1 = 40$	$s_2 = 44$

2. (20%) The TAINAN Transit Authority is expanding bus service from the suburb of AnnPin into the central business district of TAINAN. There are four routes being considered from AnnPin to downtown TAINAN: (1) via Red Line, (2) via Green Line, (3) via Blue Line, and (4) via Orange Line. The TAINAN Transit Authority conducted several tests to determine whether there was a difference in the mean travel times along the four routes. Because there will be many different drivers, the test was set up so each driver drove along each of the four routes. Below is the travel time, in minutes, for each driver-route combination. At the 0.05 significance level, is there a difference of the drivers in the mean travel time along the four routes? If we remove the effect of the drivers, is there a difference in the mean travel time? With the 0.05 significance level, the critical value of F is 3.24.

Travel Time from AnnPin to downtown TAINAN (minutes)						
Driver	**Red Line**	**Green Line**	**Blue Line**	**Orange Line**	sum	mean
Mr. Chang	18	20	20	22	80	20
Mr. Chen	21	22	24	24	91	22.75
Mr. Lin	20	23	25	23	91	22.75
Mr. Song	25	21	28	25	99	24.75
Mr. Young	26	24	28	25	103	25.75
sum	110	110	125	119		
mean	22	22	25	23.8		23.2

3. (20%) There are three bags, each contains r red balls and b black balls. Tom picks a ball from the first bag and puts it into the second bag (without seeing the color, of course), and then picks a ball from the second bag and puts it into the third in the same way. He claims that the probability that he picks a black ball from the third bag now would be the same as if he did not move the balls at all. Do you agree? Explain your assertion.

4. (20%) A random sample of 100 freshman college students revealed 20 liked the HTC U11+ cellphone well enough to purchase it. Similarly, a sample of 200 senior college students revealed 100 liked the cellphone well enough to make a purchase. We let p_1 refer to the young students and p_2 to the older students. The question is whether the difference of 0.30 in the two sample proportions is due to chance or whether there is difference in the proportion of younger and older college students who like the HTC U11+ cellphone. The z value for a 95 (99) percent level of confidence is 1.96 (2.58).

5. (20%) Consider the following normal error simple linear regression model:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad i = 1 \ldots n \qquad \text{where}$$

Y_i is the observed response in the i th trial

X_i is a known constant, the level of the predictor variable in the i th trial

β_0 and β_1 are parameters

ε_i are independent $N(0, \sigma^2)$

Let X_h denote the level of independent variable for which we wish to estimate the mean response $E\{Y_h\}$. Let $\hat{Y}_h$ denote the point estimator of $E\{Y_h\}$; $\hat{Y}_{h(new)}$ denote the new observation on response variable at a given level X_h. $\bar{X} = \sum_i X_i / n$; $\bar{Y} = \sum_i Y_i / n$. Please indicate if the following statement is True or False. (4% for each question)

(a) The maximum likelihood estimators of β_1 and β_0 are unbiased and have minimum variance among all unbiased linear estimators.

(b) Estimators of the parameters β_1, β_0, and σ^2 can be obtained by the method of maximum likelihood when the probability distribution of the error terms is not specified.

(c) The variability of the sampling distribution of $\hat{Y}_h$ is affected by how far X_h is from $\bar{X}$.

(d) The 95 percent prediction interval for the mean of 3 new observations for given X_h is narrower than that obtained for a single new observation $Y_{h(new)}$.

(e) β_1 indicates the change in the mean of the probability distribution of response variable per unit increase in predictor variable.