

國立臺灣科技大學
九十一學年度碩士班招生考試試題

系所組別：資訊管理系乙組
科目：統計學

總分100分

1. Randomly select one student from the class consisting of 30 female and 70 male students. If the selected student is weighted 40 kg then the probability that the selected student is a female student is larger than 3/10. Explain what Theorem is used to get this argument.(10%)
2. When observing sample mean $\bar{X} = 7.2$, do you believe that the population mean $\mu = 8$? What probability would you calculate? How to get this probability and draw the conclusion? State the assumptions required.(10%)
3. When observing sample variance $S^2 = 1.2$, do you believe that the population variance $\sigma^2 = 1$? What probability would you calculate? How to get this probability and draw the conclusion? State the assumptions required.(10%)
4. Consider the *simple linear regression* model

$$Y_{ij} = \alpha + \beta x_i + \varepsilon_{ij}, \quad j = 1, 2, \dots, n_i, \quad i = 1, 2, \dots, k$$

$$\varepsilon_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma^2),$$

where Y_{ij} is the j th observation with *independent variable* at level x_i .

- (a) What is the *null hypothesis* H_0 in *lack-of-fit test*?(5%)
- (b) Give the *test statistic* and its *sampling distribution*.(5%)
- (c) State how and why the *test statistic* is formed.(10%)

(Caution: Correct short answer beats long answer. Nonsense answers receive penalty.)



國立臺灣科技大學
九十一學年度碩士班招生考試試題

系所組別：資訊管理系乙組
科 目：統計學

5、(1)在點估計中，何謂估計量(Estimator)的不偏性(Unbiased Property)? (3%)

(2) 假設 $X_1, X_2, \dots, X_n$ 來自常態分配 $N(\mu, \sigma^2)$ 之隨機樣本

$$(f(X_i; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{X_i - \mu}{\sigma}\right)^2}), \text{ 請求取參數 } \mu \text{ 與 } \sigma^2 \text{ 之最大概似估計量}$$

(MLE, Maximum Likelihood Estimators)?並判斷其是否具有不偏性? (12%)

6、(1)何謂型 I 誤差(Type I Error)之顯著水準(Significant Level) α ? 為什麼在進行假設檢定前必須先決定 α 值? (5%)

(2)某工廠宣稱其產品之不良率最多只有 5%，但消費者團體卻不以為然，因此從產品中抽出 100 個樣本，結果有 10 個樣本為不良品，試問如何制訂虛無假設(Null Hypothesis)與對立假設(Alternative Hypothesis)? 且在顯著水準為 0.01 時的檢定結果為何? 【 $\Pr(Z \leq 2.326) = 0.99$ 】 (5%)

(3)承(2)，假設我們要控制型 I 誤差 α 值為 0.05，型 II 誤差 β 值為 0.01 時，則樣本數至少應該為多少? 【 $\Pr(Z \leq 1.645) = 0.95$ 】 (5%)

7、以下為統計軟體所跑出之結果，因變數與自變數分別代表如下：Y 代表軟體開發所需的時間(月)、 X_1 代表軟體程式碼行數(千行)、 X_2 代表專案經理的工作年資(年)、 X_3 代表舊軟體程式碼的可再用程度(百分比)。

參數估計表：

Variable	Parameter Estimate	Standard Error	t for H_0 : Parameter=0	Prob. > T
INTERCEPT	.533	1.495	.357	.726
X_1	.112	.038	2.919	.010
X_2	.758	.286	2.652	.017
X_3	-0.052	.020	-2.550	.021



國立臺灣科技大學
九十一學年度碩士班招生考試試題

系所組別：資訊管理系乙組
科 目：統計學

變異數分析(ANOVA)表：

Model	DF	Sum of Squares	Mean Square	F Value	Prob.>F
Regression	3	283.048	94.349	15.311	.001
Residual	16	98.594	6.162		
Total	19	381.642			

試回答下列問題：

- (1) 此迴歸模式的樣本數為何？(2%)
- (2) 請寫出複迴歸方程式？(2%)
- (3) 請計算判定係數(R Square)與修正後判定係數(Adjusted R Square)？(3%)
- (4) 請問軟體程式碼行數、專案經理的工作年資與舊軟體程式碼的可再用程度和軟體開發所需的時間在顯著水準 0.05 下是否呈線性關係？(3%)
- (5) 若想進行逐步迴歸分析，在自變數程式設計師的人數(X_2)已引入迴歸模式中，且已知 $SSR(X_1|X_2) = 32.895$ 、 $SSE(X_1, X_2) = 138.675$ 、 $SSE(X_2, X_3) = 151.098$ 。請計算軟體開發所需的時間(Y)與專案經理的工作年資(X_2)迴歸模式的判定係數與修正後判定係數之值？並說明軟體程式碼行數(X_1)與舊軟體程式碼的可再用程度(X_3)中，那個自變數宜先引入(請用偏判定係數來說明)？(10%)

