

1. 選擇題 (各4分, 共20分)

(i) Which of the following measurements randomly selected from a population could be considered a rare event from any population distribution?

Choose one of the following:

- a) $x = 0$ where $\mu = -3$ and $\sigma = 2$
- b) $x = -5$ where $\mu = 1$ and $\sigma = 4$
- c) $x = 7$ where $\mu = 3$ and $\sigma = 2$
- d) $x = 4$ where $\mu = 0$ and $\sigma = 1$
- e) $x = 10$ where $\mu = 0$ and $\sigma = 5$

(ii) An experiment was conducted in an attempt to prove that fluoride toothpaste produces a lower average number of cavities than non-fluoride toothpaste. The experimenter concluded that there was no difference in the average number of cavities for fluoride and non-fluoride toothpaste when in reality the fluoride toothpaste was better. This is:

Choose one of the following:

- a) a type I error.
- b) a type II error.
- c) a correct decision.

(iii) The term SSE is calculated by considering (在迴歸上探討)

Choose one of the following:

- a) The difference between the observed value y_i and \bar{y} , the overall mean of the y observations.
- b) The difference between the observed y_i and the fitted \hat{y}_i .
- c) An observation which is available but not incorporated into the regression analysis.
- d) The difference between the observed value of the dependent variable y_i and the observed value of the independent variable x_i .

(iv) Any experimenter testing hypotheses should know that

Choose one of the following:

- a) statistical significance and practical significance are the same thing.
- b) statistical significance does not dictate practical significance.
- c) revealing p-values to an experimenter prevents the experimenter from drawing his or her own conclusions.
- d) it is not necessary to reveal how the sample was obtained.

(v) Multicollinearity refers to a conditions where

Choose one of the following:

- a) several independent variables can be used to predict a response, y .
- b) more than one indicator variable is required to represent a qualitative variable having several classes.
- c) more than one response variable needs to be predicted from the same set of independent variables.
- d) the observations of the independent variables are highly correlated.

2. 由常態 $N(\mu, 3^2)$, 抽出四個隨機樣本: 3.6 2.4 -3.2 5.1 時
求在 90% 信賴度下之 μ 的信賴區間。(5分)

$$(z_{0.1} = 1.282 \quad z_{0.05} = 1.645 \quad t_{0.1}(3) = 1.638 \quad t_{0.05}(3) = 2.353)$$

3. USING THE COMPUTER PRINTOUT PROVIDED, ANSWER THE FOLLOWING QUESTIONS:

(各 5 分, 共 25 分, 算到或取至小數第 4 位即可)

General Linear Models Procedure

X'X Inverse Matrix

	INTERCEPT	X1	X2	X2*X2
INTERCEPT	5.5573931854	-0.188337249	-1.387787309	0.1000193631
X1	-0.188337249	0.0703566842	-0.06133794	0.0025800415
X2	-1.387787309	-0.06133794	0.5552235216	-0.037943206
X2*X2	0.1000193631	0.0025800415	-0.037943206	0.0027027568

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	99.74389129	33.24796376	26.50	0.0001
Error	9	11.29338563	1.25482063		
Corrected Total	12	111.03727692			
	R-Square	C.V.	Root MSE		Y Mean
	0.898292	14.44114	1.120188		7.75692308

Source	DF	Type I SS	Mean Square	F Value	Pr > F
X1	1	1.50295073	1.50295073	1.20	0.3022
X2	1	95.83116837	95.83116837	76.37	0.0001
X2*X2	1	2.40977220	2.40977220	1.92	0.1992

Source	DF	Type III SS	Mean Square	F Value	Pr > F
X1	1	30.69307800	30.69307800	24.46	0.0008
X2	1	12.20341267	12.20341267	9.73	0.0123
X2*X2	1	2.40977220	2.40977220	1.92	0.1992

Parameter	Estimate	T for H0: Parameter=0	Pr > T	Std Error of Estimate
INTERCEPT	4.206815646	1.59	0.1456	2.64074452
X1	-1.469511210		0.0008	
X2	2.603002451	3.12	0.0123	0.83468912
X2*X2	-0.080703334	-1.39	0.1992	0.05823637

- (i) Find R^2 for this model.
- (ii) Write the prediction equation for the fitted model.
- (iii) Find the standard error of the estimate of the coefficient of X1.
- (iv) To test $H_0: \beta_2 = 2$ versus $H_A: \beta_2 \neq 2$, the numerical value of the test statistic would be what? (β_2 is the coefficient of X2 in the model.)
- (v) If X2 were dropped from this model, find R^2 for the new model with just X1 and X2.

4. A drug manufacturer wishes to assess the effect of three drugs on the motor ability of rats. He obtains four litters of rats, each litter containing three rats, and places each litter of three rats in separate cages. He randomly assigns and administers either drug A, B, or C to the three rats in each cage, each receiving a different drug. He repeats the procedure with each cage and then measures the motor ability of each rat. (10分)

- (i) What type of experiment is this? (4分)
(ii) Identify the factors and the number of levels for each factor. (6分)

5. 小孩玩某種比賽所花時間 t (分鐘) 服從下面分配 (Distribution):

t	3	4	5	6	7	8
$P(t)$	0.1	0.1	0.3	0.2	0.2	0.1

若小孩比賽所花時間比 7 分鐘快 1 分鐘 (即花 6 分鐘), 就給餅乾一塊, 4 快 2 分鐘, 則給 2 塊, 依次類推, 但達 7 分鐘或以上時, 則無餅乾, 試求其平均報酬為何? (10分)

6. 由 A, B, C, 三條製造同一產品的生產線, 分別抽取 4, 5, 6 件產品, 並求得產品重量的平均數 \bar{X}_i 及變異數 S_i^2 依次為:

平均數 \bar{X}_i	: 21.75	23.30	19.50
變異數 S_i^2	: 1.58	2.70	2.30

試以上面資料做一變異數分析表 (ANOVA 表) (註: ANOVA 表即可, 不必檢定) (15分)

7. (a) 每天 9:00 - 9:10, 總機平均接到 5 通電話, 在未來兩天同一時間共接到 3 通電話的機率? (5分)

(b) 一產品之不良率為 0.001, 今自一批中抽取 200 件, 試問樣本中含有 2 個及 2 個以上不良品的機率為何? (註: 當然您不會用二項分配求, 而會用其近似分配來解, 所以, 請問您會用什麼近似分配? 並列出您的計算式子). (5分)

(c) X 分配為 $N(20, 16)$, Y 分配為 $N(18, 9)$, 令 \bar{X}, \bar{Y} 分別是樣本大小為 25 的樣本平均數, 求 $P(\bar{X} > \bar{Y})$ 之機率. (註: N 是常態之意) (5分)