

一. 現有某織布廠考慮兩種織布製程 I 與 II 對所織成的布匹上的瑕疵個數是否有影響, 故分別在此二製程下各隨機抽取  $n_1$  與  $n_2$  個樣本。其中樣本之定義是以 10 碼布為單位, 觀察這一塊 10 碼布上的瑕疵個數。所得數據如下:

製程 I: 0 0 2 1 0 1 3 0 2 4 3 0 0 0 1 0 0 0 2 0

製程 II: 0 2 0 0 1 2 3 4 2 2 0 2 3 0 1 1 2 1 2 0 0  
1 1 2 1

其中  $n_1 = 20$      $\bar{x}_1 = 0.95$      $s_1 = 1.276$   
 $n_2 = 25$      $\bar{x}_2 = 1.40$      $s_2 = 1.225$

(5%) (1) 試就平均數, 標準差, 變異係數 (c.v. 值) 三量做比較, 說明那個製程比較好。

(5%) (2) 你認為這二組數據應服從什麼分配? 為什麼?

(10%) (3) 試就製程 II 所得的 25 個數據, 以檢定的方法來判斷是否與 (2) 中的分配相符. ( $\alpha = 0.05$ )

二. 交通安全常識中一再強調機車騎士戴安全帽的重要性。由於台南氣候炎熱, 所以台南市機車騎士配戴安全帽的比例  $p$  成為我們關心的問題。今在各路口隨機觀察 400 位機車騎士行車時, 其中 50 位有配戴安全帽。

(5%) (1). 題中“在各路口隨機觀察”, 你認為詳細的步驟應如何呢? 換言之, 如何隨機觀察, 所取得的樣本才能對  $p$  值做出合理的統計推論。

(5%) (2) 試求“台南市機車騎士配戴安全帽的比例  $p$ ”的 95% 信賴區間。

(5%) (3) 請解釋 (2) 中信賴係數 95% 的涵義。

854

(5%) (4) 若希望 (2) 中的信賴區間寬度不要超過 0.05, 則題目中的樣本數是否足夠? 若不夠, 則需增加至多少以上?

(10%) (5) 若有人存疑台南市機車騎士配戴安全帽的比例不超過  $\frac{1}{6}$ , 你如何利用題目中的數據做成結論呢? ( $\alpha = 0.05$ ).  
(注意: 所有檢定均須陳述  $H_0$  與  $H_1$ , 檢定統計量及其分配, 並做成清楚的結論.)

三. 某工廠生產線採三班制 (8:00 A.M. ~ 4:00 P.M., 4:00 P.M. ~ 12:00 P.M., 12:00 P.M. ~ 8:00 A.M.), 並由兩位領班輪流監督此生產線。現在該廠經理懷疑各班產品的產量與各班時段及領班均有關係, 所以著手一個  $2 \times 3$  的因子實驗, 並在每一個組合下, 分別重複試驗三次。將所得  $3 \times 2 \times 3 = 18$  個數據利用 Minitab ANOVA 得到下列 ANOVA 表。

Source	DF	SS	MS	F	P
A (領班)	1	19012	19012	>6.43	0.000
B (三班)	2	>58	129	0.18	0.838
A*B	a	80908	40454	e	0.000
Error	b	c	d		
Total	17	108813			

(5%) (1) 請將表中 a, b, c, d, e 各項計算出來。

(10%) (2) 由於上表中, 對於因子 A, 領班的影响, 其 p-value = 0.000. 我們是否可結論如下: 兩位領班的監督會對產品產量有不同的影响? 為什麼?

(10%) (3) 如果在 (2) 中, 答案為否, 則我們應該再對原始數據做適當的統計分析, 以說明何者會影響產量。請陳述你的下一步驟將如何。

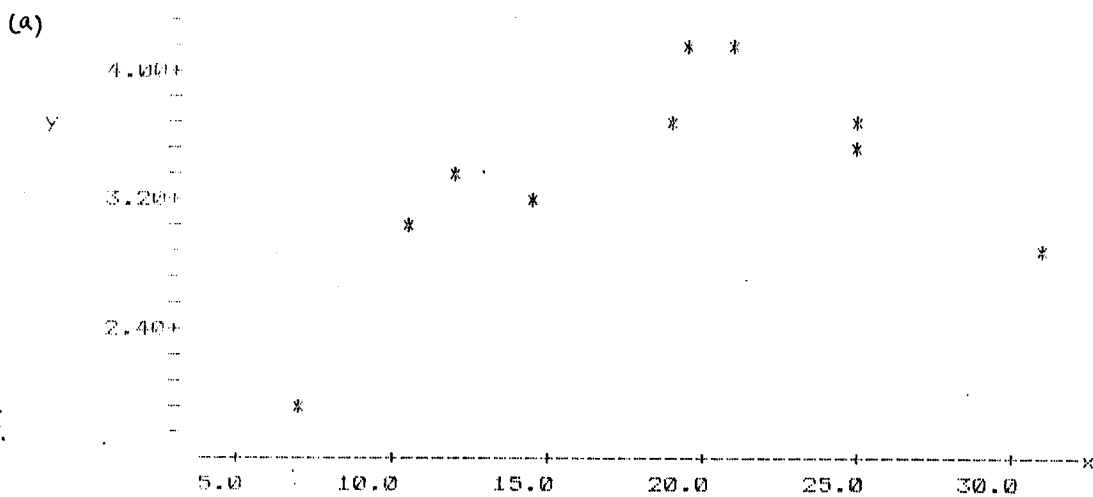
(5%) (4) 在 ANOVA 分析之前, 我們需要那些基本假設? 為什麼?

四. 令  $X_1, X_2, \dots, X_n$  是一組獨立, 而且具有相同分配的自變利  
 (20%) 隨機變數, 亦即  $X_1, X_2, \dots, X_n \stackrel{i.i.d.}{\sim} B(1, \theta)$ . 試求出參數  
 $\theta$  的 U.M.V.U.E 估計值。(Uniformly minimum variance unbiased  
 estimator)

五. 在一個零售雜貨店人力配備研究中, 我們想瞭解每小時人力應得  
 多少報酬  $Y$  與雜貨店的面積大小 (每 1000 平方英尺為單位)  $X$  之間的  
 關係式. 今調查 10 家雜貨店, 並分別記錄  $(X, Y)$  如下:

X	21.0	12.0	25.2	10.4	30.9	6.8	19.6	14.5	25.0	19.1
Y	4.08	3.4	3.51	3.09	2.92	1.94	4.11	3.16	3.75	3.60

並經由 Minitab 得 (a) scatter-plot (b) linear regression (c) polynomial  
 regression 之結果如下:



(b)  
 The regression equation is  
 $y = 2.62 + 0.0396 x$

Predictor	Coef	Stdev	t-ratio	p
Constant	2.6247	0.5219	5.03	0.000
x	0.03964	0.02638	1.50	0.171

s = 0.5955      R-sq = 22.0%      R-sq(adj) = 12.3%

Analysis of Variance

SOURCE	DF	SS	MS	F	p
Regression	1	0.8003	0.8003	2.26	0.171
Error	8	2.8371	0.3546		
Total	9	3.6374			

(c)

The regression equation is  
 $y = -0.159 + 0.392x - 0.00949x^2$

Predictor	Coef	Stdev	t-ratio	P
Constant	-0.1594	0.5006	-0.32	0.760
x	0.39193	0.05801	6.76	0.000
x**2	-0.009495	0.001535	-6.19	0.000

s = 0.2503      R-sq = 87.9%      R-sq(adj) = 84.5%

Analysis of Variance

SOURCE	DF	SS	MS	F	P
Regression	2	3.1989	1.5994	25.53	0.001
Error	7	0.4385	0.0626		
Total	9	3.6374			

SOURCE	DF	SEQ SS
x	1	0.8003
x**2	1	2.3986

(1). 在 (c) 中分別對括弧 ① ② ③ ④ 的量所代表的意義做解釋:

(10%) ① t-ratio = -0.32 是想檢定什麼? 其結論又如何? t 之自由度=?

(10%) ② 在迴歸分析中, 需要對模式做什麼假設?  $s = 0.2503$  表示什麼呢?  $s$  與模式中有任何關係嗎?

(5%) ③  $R^2 = R\text{-sq} = 87.9\%$ , 其  $R^2$  之定義為何呢?  $R^2$  與  $s$  有關係嗎?

(10%) ④  $F = 25.53$  所欲檢定的  $H_0, H_1$ , 檢定統計量, 及其分配並做成統計推論。

(2) 在 (b) 中,  $F = 2.26$ ,  $p = 0.171 > \alpha = 0.05$ , 而 (c) 中  $F = 25.53$ ,  $p = 0.001$

(5%) 所以我們認為模式 (b) 比模式 (c) 好, 你認為這樣的結論合理嗎?

(3) 又在此題中, 若新觀查一個  $x = 20.0$ , 你認為在合理的模式下,

(10%) 每小時人力應得之報酬為多少美金? 請寫出 95% 的合理預估區間。

參考資料:

$$Z_{0.025} = 1.96$$

$$\chi^2_{0.05}(5) = 11.07$$

$$t_{0.025}(10) = 2.228$$

$$Z_{0.05} = 1.645$$

$$\chi^2_{0.05}(4) = 9.49$$

$$t_{0.025}(9) = 2.262$$

$$Z_{0.1} = 1.282$$

$$\chi^2_{0.05}(3) = 7.81$$

$$t_{0.05}(8) = 2.306$$