

- [5] 一、爲了瞭解某種殺蟲劑在不同濃度下對某種害蟲的殺蟲效果，阿聰做了下列的試驗：共有八個不同濃度 (10, 20, 30, 40, 50, 60, 70, 80ppm/mL)，每個濃度噴灑於 100 隻害蟲。經過一小時後，阿聰得到下列八個濃度下害蟲的死亡率分別爲 6.38%，43.78%，68.79%，72.37%，91.53%，93.33%，96.67%，98.73%。如果沒有任何遺失值，請解釋上述八個濃度下之害蟲死亡率是否合理？請說明理由。
- [5] 二、教授請阿勇分析兩個獨立樣本 (A 組及 B 組) 的資料。下表爲阿勇計算兩組樣本平均值 (sample mean) 及樣本標準偏差 (sample standard deviation) 及兩組樣本合併後的樣本平均值與樣本標準偏差。

	A 組	B 組	兩組合併
樣本數	10	20	30
樣本平均值	100	150	200
樣本標準偏差	30	40	20

請問上表中有哪些錯誤或不合理的地方？並說明其理由。

- [10] 三、針對兩個獨立族群趨中性參數差異 (difference in parameters of central tendency) 的假說 (hypothesis) 之檢定方法有：非成對學生氏 t 檢定方法 (unpaired student t -test) 或無母數 (nonparametric) 的魏克森等級和檢定方法 (Wilcoxon rank sum test)。請說明這兩種檢定方法之異同。
- [10] 四、假設生男生女的機率是一樣的，如果一對夫婦計畫要一直生小孩直到生出一男一女，或是生了 3 個就停。
- [5] (1) 請計算此對夫婦的小孩中至少有一男及一女的機率。
- [5] (2) 爲了鼓勵生產，如果每生一個小孩政府就發給獎金十萬，請問他們平均可以拿到多少錢？
- [10] 五、阿梅將某殺蟲劑對於 n 隻某害蟲之致死效果整理爲下表：

死亡	存活	總和
O_1	O_2	$n = O_1 + O_2$

她利用下面假說，檢定此殺蟲劑的致死率 (p) 是否爲 p_0 ，

$$H_0: p = p_0 \quad \text{vs} \quad H_a: p \neq p_0$$

見背面

針對此一假說，阿梅在統計教科書上找到三種檢定統計量(test statistics)

$$(1) |Z_1| = \left| \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} \right|$$

$$(2) |Z_2| = \left| \frac{\hat{p} - p_0}{\sqrt{\hat{p}(1-\hat{p})/n}} \right|, \text{ 及}$$

$$(3) \chi^2 = \sum_{i=1}^2 \frac{(O_i - E_i)^2}{E_i}$$

其中 $\hat{p} = \frac{O_1}{n}$, E_i 為第 i 個類別之期望次數 (expected frequency), $i = 1, 2$, $E_1 = np_0$, $E_2 = n(1-p_0)$ 。針對上述假說，請解釋這三種檢定統計量之異同。並說明理由。

- [10] 六、為了要瞭解某種保育動物在人工扶養成年後放生於野外的存活情形，阿珠調查了 15 隻此種保育動物放生野外之死亡時間。因為經費及人力的限制，對有些動物只能觀察到最後存活時間而無法觀察到真正死亡時間 [此類資料為設限資料 (censored data)]。下面為 15 隻保育動物在放生後的死亡時間 (月)。

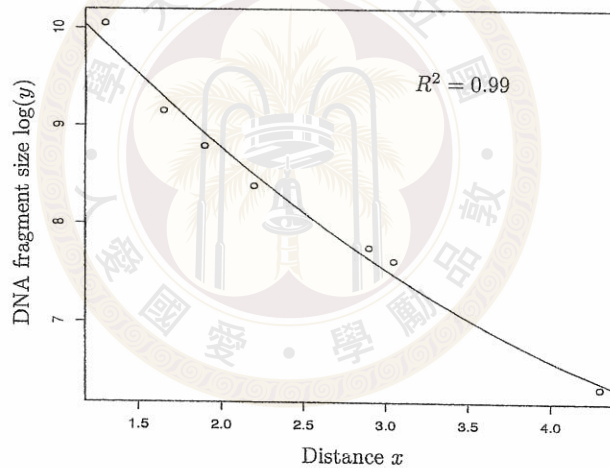
0.92 2.93 5.76 6.41 10.16 12.40⁺ 12.93⁺ 13.85⁺
14.70⁺ 15.20 23.32⁺ 24.47⁺ 25.33⁺ 36.38⁺ 39.67⁺

請計算此種保育動物放生野外 1 年 (12 個月) 的存活機率 (survival probability)。
註：資料符號說明：0.92 (第一橫排第一個觀測值) 表示此隻保育動物在放生野外後 0.92 個月時死亡。12.40⁺ (第一橫排第 6 個觀測值) 表示設限資料，也就是此隻保育動物在放生野外的真正死亡時間未知，但在 12.40 月時仍然存活。以此類推其他資料之意義。

- [25] 七、以 Agarose gel 做電泳是分子生物學中最廣泛使用的技術。微小的 DNA 片段受到電流的作用後會在 gel matrix 上移動，其移動之速率受到 DNA 片段長度 (DNA fragment length) 的影響，DNA 片段泳動後的位置可以藉由 ethidium bromide 標示出來。現有一組數據顯示 DNA 泳動的距離 x (distance migrated) 與其大小 y (DNA fragment size),

Distance Migrated x cm	DNA Size y bp	$\log(\text{DNA Size})$ $\log(y)$ log(bp)
1.30	23,130	10.05
1.65	9,416	9.15
1.90	6,557	8.79
2.20	4,361	8.38
2.90	2,322	7.75
3.05	2,027	7.61
4.30	564	6.34

通常泳動距離 x 與 DNA 片段長度的對數值 $\log(y)$ 會呈現線性的關係 (直線或曲線), 下圖即顯示此種關係。圖中之 \circ 代表觀測值, — 代表配適的回歸線。



下列是根據統計計算軟體 R 所得的統計分析結果

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	12.24764	0.55039	22.253	2.41e-05 ***
x (linear)	-2.06066	0.43133	-4.777	0.0088 **
x2(quadratic)	0.16271	0.07624	2.134	0.0997 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1793 on 4 degrees of freedom

Multiple R-Squared: 0.9852, Adjusted R-squared: 0.9777

F-statistic: 132.7 on 2 and 4 DF, p-value: 0.0002205

見背面

Analysis of Variance Table

Response: log(y)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
x (linear)	1	8.3846	8.3846	260.8452	8.597e-05 ***
x2 (quadratic)	1	0.1464	0.1464	4.5549	0.09973 .
Residuals	4	0.1286	0.0321		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

請根據分析結果回答下列問題

- [5] (1) 請根據上述分析結果，寫出配適曲線之回歸方程式。
- [5] (2) 若有一個 DNA 片段泳動距離為 2.5，請推算此 DNA 片段之長度 (以 $\log(bp)$ 表示) 是多少？已知此推算值之標準誤差為 $SE = 0.0975$, $t_{0.025;4} = 2.78$ ，請計算當泳動距離為 $x = 2.5$ 時，DNA 片段長度的 95% 的信賴區間。
- [5] (3) 圖的右上方標示 $R^2 = 0.99$ ，請說明其意義。
- [5] (4) 請說明真的需要二次方程式 ($y = \beta_0 + \beta_1x + \beta_2x^2 + \epsilon$) 才能描述此資料嗎？或是一次方程式 ($y = \beta_0 + \beta_1x + \epsilon$) 就已經足夠描述此資料？若是以一次方程式來描述此資料，則請問此時 $R^2 = ?$
- [5] (5) DNA 的泳動距離 x 與其長度 $\log(y)$ 兩者的直線相關係數 $r = -0.984$ 請以檢定統計量

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

檢定相關係數 ρ 是否與 0 有顯著差異，虛無假說與對立假說如下：

$$H_0: \rho = 0 \quad \text{vs} \quad H_a: \rho \neq 0$$

並請敘述檢定結果。(請選擇正確的臨界值 $t_{0.025;4} = 2.78$, $t_{0.025;5} = 2.570$, $t_{0.025;6} = 2.447$)

- [25] 八、一位研究老年學的科學家，希望能瞭解攝食量與壽命 (life span) 之關係，她懷疑攝食量越大者壽命越短，她將實驗室的小白鼠分成三組，給予不同的食物量，第一組：食物不限量，第二組：食物僅有正常的 90%，第三組：食物僅有正常的 80%，每組有 5 隻小白鼠參與試驗。下表中的數字是各組的平均壽命與樣本均方

Group	I	II	III
Size (n_i)	5	5	5
Average (\bar{y}_i)	2.44	3.18	3.54
Sample variance (s_i^2)	0.19	0.17	0.25

請回答下列問題

- [5] (1) 請根據上表中之 s_i^2 求出共同均方 (pooled mean square) s_p^2 之數值。
- [5] (2) 資料經變方分析結果得到的 $F = 7.6966$ 以及隨伴的 $p = 0.007067$ (請見下列變方分析表)。請問如何解讀此分析結果。

Analysis of Variance Table

Response: y (life span)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
treatment	2	3.14533	1.57267	7.6966	0.007067 **
Residuals	12	2.45200	0.20433		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- [5] (3) 對於三組平均壽命值做 mean separation (鄧肯氏的多變域比較 Duncan's multiple range test) 的結果列於下表

Group	I	II	III
Average life span (year)	2.44 ^a	3.18 ^b	3.54 ^b

請問如何解讀此分析結果。

- [5] (4) 請計算 $\mu_3 - \mu_1$ 的 95% 信賴區間, 此處 μ_1, μ_3 分別是第一組與第三組的平均壽命。計算此信賴區間的目的是想瞭解少吃 20% 的食物可以多活多久。已知 $s = 0.452, \sqrt{2} = 1.414, \sqrt{5} = 2.236, t_{0.025, 12} = 2.18$ 。
- [5] (5) 請根據變方分析表推測: 此試驗所採用的設計是哪一種設計? 其英文名稱為何? 其統計模式為何?