統計分析結果的報導方式

Chen-Pan Liao

December 8, 2019

目錄

[1]詳細授權內容請見 `http://creativecommons.org/licenses/by-sa/4.0/deed.zh_TW`。
[2]`https://github.com/chenpanliao/report-statistical-results-TC`。

1 前言

一般而言，在收集樣本後必須報導描述性統計，包括中央趨勢 (如平均值或中位數) 、樣本數及變異程度 (如標準偏差或標準誤差)；這些敘述性統計若內容太多可以改以圖或表的方式呈現。對於特別感到興趣的參數應計算其信賴區間。進行檢驗後應報導檢定統計量 (如 T、F、χ^2 等) 、自由度與 p-value，並報導合適的效果量 (如 Cohen's D、r、R^2 等)。在撰寫統計結果時，必須報導上述重要的統計結果，重新編排成有意義的圖或表，最終正確解讀結果。然而，在筆者經驗中，初學統計的大學生往往不能掌握這一連串撰寫統計結果的技巧。

因此，我收集了一些常見統計例題，並以學術報告口吻示範如何報導上述統計結果供學生模仿。以下我將按不同的分析情況示範報導分析結果，包括結果的文字撰寫與製作合適的圖表。因課程訓練需求，我刻意報導較多細

節而看來十分繁瑣冗長。學生可以先模仿我的內容以撰寫統計學報告與作業，但未來其它課程或學術報告時應有所取捨。最末一併附上計算及繪圖之 R code。本文內容將隨課程進度持續增加內容。

本文關於效果量在筆者主持之課堂中並未多加說明，且不同的效果量適合不同的統計方法，學生可按自己的能力決定是否報導效果量。

2 單樣本均值檢驗

2.1 常態情況

檢驗 8.8, 10.3, 11.1, 7.7, 10.4, 10.5, 9.4, 9.5, 9.4, 9.1 之中央趨勢是否顯著不同於 9。

結果指出，樣本平均 ± 標準差為 9.62 ± 0.987 ($n = 10$)。由 Shapiro-Wilk test 檢驗常態性發現不能拒絕常態之虛無假設 ($W = 0.960$, $P = 0.790$)，故以 one-sample two-tailed Student-t test 進行檢驗 $H_0: \mu = 9$。結果指出，平均值的 95% 信賴區間為 $[8.914, 10.236]$，無法拒絕 $\mu = 9$ 的虛無假設 ($T = 1.986$, DF = 9, $P = 0.078$)。此外，Cohen's $D = 0.627$ 顯示中度效果量。結論是，母體平均不顯著偏離 9，但由中度效果量推測，不顯著可能是因樣本數不足造成的。

2.2 非常態情況

檢驗 2.5, 0.25, 0.01, 1.74, 0.39, 0.09, 0.82, 0.2, 0.84, 0.76 之中央趨勢是否顯著不同於 2。

結果指出，樣本平均 ± 標準差為 0.76 ± 0.797 ($n = 10$)。由 Shapiro-Wilk test 檢驗常態性發現拒絕常態之虛無假設 ($W = 0.841, P = 0.045$)，故以 Wilcoxon signed rank sum test 進行檢驗 H_0 : 中位數 = 2。結果指出，應拒絕中位數 = 2 的虛無假設 (樣本中位數 = 0.575, $V = 2$, $P = 0.006$)。此外，多達 90% 的樣本小於 2，顯示高度的效果量。結論是，母體中位數顯著不等於 2 且小於 2。[3]

[3] 在雙尾檢驗後若顯著可以藉樣本平均或中位數的大小直接解釋為顯著大於或小於。

3 配對兩樣本均值檢驗

3.1 常態情況

檢驗以下配對樣本

x_1	8.8	10.3	11.1	7.7	10.4	10.5	9.4	9.5
x_2	9.2	10.4	11.6	7.7	10.6	11.6	11.4	10.4

之差值 $(x_1 - x_2)$ 中央趨勢是否顯著小於 0.1。

結果指出，x_1 與 x_2 之平均 ± 標準差分別為 9.71 ± 1.095 及 10.36 ± 1.344 ($n_{\text{pair}} = 8$；圖 1a)。差值平均 ± 標準差為 -0.65 ± 0.665 (圖 1b)。由 Shapiro-Wilk test 檢驗差值之常態性發現不能拒絕常態之虛無假設 ($W = 0.883$，$P = 0.202$)，故以 two-sample paired t-test 檢驗 $H_0: \mu_1 - \mu_2 \geq 0.1$。結果指出，應拒絕虛無假設 ($T = -3.188$，DF = 7，$P = 0.015$)。此外，差值平均之 95% 信賴區間為 $[-1.206, -0.0936]$，且 Cohen's $D = 1.728$ 顯示高度效果量。結論是：差值平均顯著小於 0 且差距之效果量甚高。

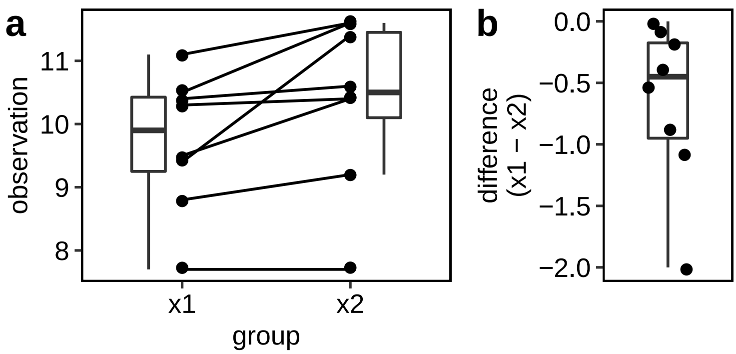

圖 1: 配對兩樣本的觀測值盒形圖 (a) 及差值盒形圖 (b)。

3.2 非常態情況

檢驗以下配對樣本

x_1	5.1	6.9	7.2	6.5	7.2	6.4	5.3	7.7
x_2	5.6	6.2	6.6	6.7	6.7	5.8	4.8	8.1

之差值 $(x_1 - x_2)$ 中央趨勢是否顯著偏離 1。

結果指出，x_1 與 x_2 之平均 ± 標準差分別為 6.538 ± 0.924 及 6.313 ± 0.975 ($n_{\text{pair}} = 8$；圖 2a)。差值平均 ± 標準差為 0.225 ± 0.501 (圖 2b)。由 Shapiro-Wilk test 檢驗差值之常態性，結果顯示應拒絕常態之虛無假設 ($W = 0.797$，$P = 0.026$)，故以 Wilcoxon signed rank sum test 進行檢驗 H_0 : 差值中位數 = 1。結果顯示，差值中位數顯著偏離 1 ($V = 0$, $P = 0.014$) 而是小於 1。此外，100% 的樣本差值小於 1，具極高的效果量。結論是，差值母體中位數顯著小於 1 且效果量高。

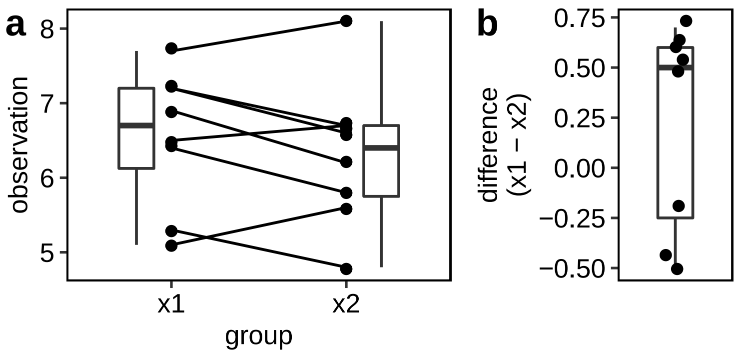

圖 2: 配對兩樣本的觀測值盒形圖 (a) 及差值盒形圖 (b)。

4 獨立兩樣本均值檢驗

4.1 常態情況

檢驗以下兩獨立樣本

x_1	8.6	10	9.2	10.2	11.4	10.7	
x_2	9.7	8.8	9.2	10.2	9.3	7.6	8.6

之中央趨勢是否顯著偏離 0。

結果指出，x_1 與 x_2 之平均 ± 標準差分別為 10.02 ± 1.095 及 9.057 ± 0.836 ($n_1 = 6$，$n_2 = 7$；圖 3)。由 Shapiro-Wilk test 檢驗差值之常態性發現二樣本皆不能拒絕常態之虛無假設 (x_1，$W = 0.985$，$P = 0.975$；x_2, $W = 0.976$, $P = 0.938$), 故以 Welch two-Sample t-test 檢驗 $H_0 : \mu_1 - \mu_2 = 0$。結果指出不應拒絕虛無假設 ($T = 1.848$，DF = 9.794,

$P = 0.095$)。此外，差值平均之 95% 信賴區間為 $[-0.201, 2.120]$，且 Cohen's $D = 1.044$ 顯示高度效果量。結論是二樣本平均無顯著差異，但效果量甚高，可能因樣本數不足而發生型二錯誤。

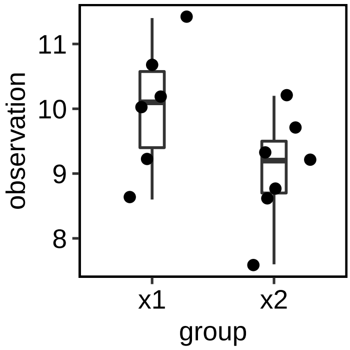

圖 3: 獨立兩樣本的觀測值盒形圖。

4.2 非常態情況

檢驗以下兩獨立樣本

x_1	0	0.1	0.7	0.7	0.9	0.7	0	0.9
x_2	0.7	1.6	0.6	0.4	1.7	0.2	1.4	

之中央趨勢是否顯著偏離 0。

結果指出，x_1 與 x_2 之平均 $\pm$ 標準差分別為 0.5 ± 0.396 及 0.943 ± 0.611 ($n_1 = 8$；$n_2 = 7$；圖 4)。由 Shapiro-Wilk test 檢驗差值之常態性發現 x_1 拒絕常態之虛無假設 (x_1，$W = 0.794$，$P = 0.024$；x_2，$W = 0.890$，$P = 0.276$)，故以 Mann-Whitney U test 檢驗 $H_0 : \text{Median}_1 - \text{Median}_2 = 0$。結果指出不應拒絕虛無假設 ($W = 18.5$，$P = 0.292$)。此外，Cliff's $d = 0.339$ 顯示中等程度效果量。結論是，二樣本之中位數無顯著差異，但效果量程中度，可能因樣本數不足而發生型二錯誤。

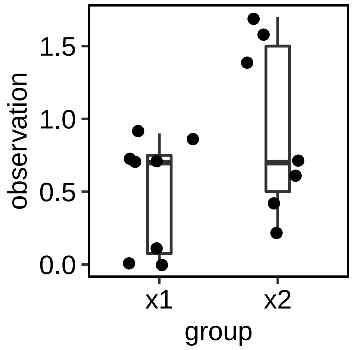

圖 4: 獨立兩樣本的觀測值盒形圖。

5 多樣本單因子均值檢驗

5.1 常態且變方同質情況

檢驗以下三獨立樣本

x_1	5.16	4.24	4.7	4.58	6.06	5.99	
x_2	4.91	5.65	5.58	5.12	4.32		
x_3	7.65	6.64	7	5.57	5.84	8.48	7.07

之中央趨勢是否相等。

三樣本的描述性統計如表 1。由於三組樣本分布並不顯著偏離常態 (Shapiro-Wilk test, $P_1 = 0.709$, $P_2 = 0.85$, $P_3 = 0.925$), 且變異數未顯著異質 (Bartlett test, $\chi^2 = 0.469$, DF = 2, $P = 0.791$), 故以 one-way ANOVA 檢驗 $H_0 : \mu_1 = \mu_2 = \mu_3$。結果顯示, x 為顯著因子 ($F = 9.297$, DF = $[2, 15]$, $P = 0.0024$), 且 η^2 顯示有 55.3% 的變異量可由 x 因子解釋。接下來以 Tukey's range test 進行多重比較, 結果顯示, x_1 與 x_3 存在顯著差異, 而 x_2 與另二組皆無顯著差異 (表 2;圖 5)。

表 1: 獨立三樣本的描述性統計。

Group	Mean	SD	n
x_1	4.94	0.77	6
x_2	5.86	0.79	5
x_3	6.65	0.59	7

表 2: 獨立三樣本的事後多重比較。

Comparison	Cohen's D	Estimate	95% CI	P_{adj}
$x_2 - x_1$	1.177	0.917	[−0.200, 2.034]	0.117
$x_3 - x_1$	2.505	1.704	[0.677, 2.730]	0.002
$x_3 - x_2$	1.163	0.787	[−0.293, 1.867]	0.175

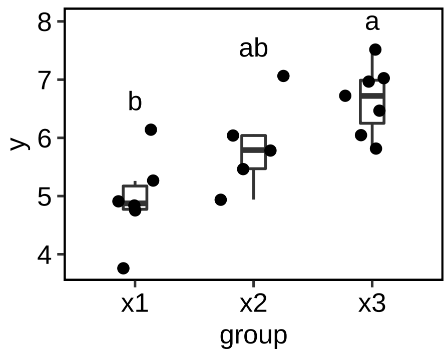

圖 5: 獨立三樣本的觀測值盒形圖。上方字母為多重比較的分群結果；若任二組存在相同字母則表示不存在顯著差異，反則反之。

5.2 常態且變方異質情況

檢驗以下三獨立樣本

x_1	3.18	4.12	3.52	3.29	5.13	5.2	
x_2	5.7	5.21	7.62	8.19	6.26		
x_3	7.12	7.4	8.19	3.66	3.78	11.9	6.87

之中央趨勢是否相等。

三樣本的描述性統計如表 3。由於三組樣本分布並不顯著偏離常態 (Shapiro-Wilk test, $P_1 = 0.158$, $P_2 = 0.593$, $P_3 = 0.388$), 且變異數顯著不相等 (Bartlett test, $\chi^2 = 6.340$, DF = 2, $P = 0.042$), 故以 Welch one-way ANOVA 檢驗 $H_0 : \mu_1 = \mu_2 = \mu_3$。結果顯示, x 為顯著因子 ($F = 8.248$, DF = $(2, 8.953)$, $P = 0.009$) 即三組母體平均不全相等, 且 η^2 顯示有 34.77% 的變異量可由 x 因子解釋。接下來以 Games-Howell method 進行多重比較, 結果顯示, x_1 與 x_3 存在顯著差異, 而 x_2 與另二組皆無顯著差異 (表 4；圖 6)。

表 3: 獨立三樣本的描述性統計。

Comparison	Cohen's D	Estimate	95% CI	P
$x_2 - x_1$	1.176	0.917	[−0.200, 2.034]	0.117
$x_3 - x_1$	2.475	1.704	[0.677, 2.730]	0.002
$x_3 - x_2$	1.130	0.787	[−0.293, 1.867]	0.175

表 4: 獨立三樣本的事後多重比較。

comparison	Cohen's D	Estimate	95% CI	P
$x_2 - x_1$	2.290	2.523	[0.536, 4.509]	0.017
$x_3 - x_1$	1.400	2.915	[−0.345, 6.175]	0.077
$x_3 - x_2$	0.180	0.393	[−2.973, 3.758]	0.943

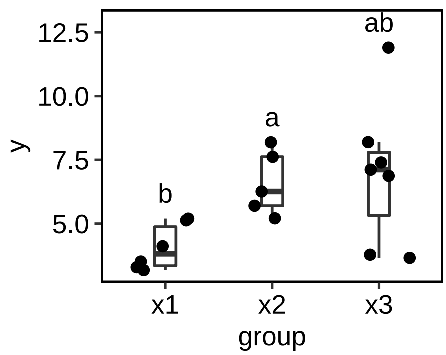

圖 6: 獨立三樣本的觀測值盒形圖。上方字母為多重比較的分群結果；若任二組存在相同字母則表示不存在顯著差異，反則反之。

5.3 非常態情況

檢驗以下三獨立樣本

x_1	11.2	0.5	4.5	2.8	16.5	2.2	2.1	2
x_2	2.8	1.8	1.7	3.6	3.5	1.4	0.5	
x_3	0.5	2	0.8	0.3	0.4	0.4		

之中央趨勢是否相等。

三樣本的描述性統計如表 5。由於 x_1 與 x_3 分布顯著偏離常態 (Shapiro-Wilk test，$P_1 = 0.014$，$P_2 = 0.577$，$P_3 = 0.008$)，故以 Kruskal-Wallis rank sum test 檢驗 H_0 : 中位數$_1$ = 中位數$_2$ = 中位數$_3$。結果顯示，x 為顯著因子 ($\chi^2 = 9.041$, DF = 2, $P = 0.011$)，且 $\eta^2 = 0.391$ 顯示高度效果量。Dunn's Kruskal-Wallis multiple comparisons 之多重比較結果顯示，x_1 與 x_3 存在顯著差異，而 x_2 與另二組皆無顯著差異 (表 6；圖 7)。

表 5: 獨立三樣本的描述性統計。

group	Mean	SD	n
x_1	5.225	5.617	8
x_2	2.186	1.151	7
x_3	0.733	0.644	6

表 6: Dunn's Kruskal-Wallis 多重比較之結果。

comparison	Cliff's d	Z	P	P_{adj}
x_1 - x_2	0.357	0.964	0.335	0.335
x_1 - x_3	0.875	2.979	0.003	0.009
x_2 - x_3	0.738	1.994	0.046	0.092

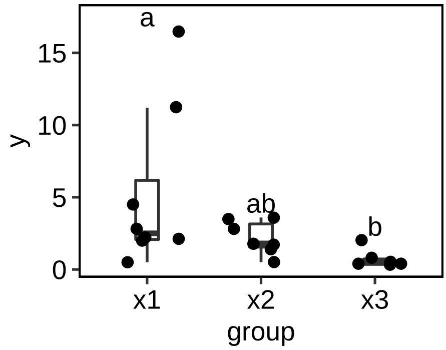

圖 7: 獨立三樣本的觀測值盒形圖。上方字母為多重比較的分群結果；若任二組存在相同字母則表示不存在顯著差異，反則反之。

6 多樣本多因子均值檢驗

6.1 二因子交互作用因子實驗

完全隨機 3×2 雙因子之因子實驗結果如下,

A	B	Y
A_1	B_1	$4.7, 3.4, 4, 4.7$
A_1	B_2	$3.8, 4.2, 5.1, 5.9$
A_1	B_3	$6.7, 5.9, 6.7, 5.5, 5.9$
A_2	B_1	$5.4, 4.4, 5.9, 6.6$
A_2	B_2	$7.4, 8, 5.9, 6.8$
A_2	B_3	$9.5, 9.5, 11, 10.3, 10$

分析二因子或交互作用對 Y 中央趨勢之影響。

各組平均及其信賴區間如圖 8 所示。Two-way ANOVA 結果指出 $A \times B$ 交互作用顯著 ($P = 0.006$；表 7)，故將二因子合併進行 one-way ANOVA 以估計簡單主效應。結果指出，六組間不全相等 ($F = 35.33$，DF = $[5, 20]$，$P < 0.001$，$\eta^2 = 89.8\%$)。以 Tukey's range test 進行事後兩兩比較，結果指出，在 A 因子相同的情況下，B_3 皆顯著高於 B_1 與 B_2 且 B_1 與 B_2 間沒有顯著差異；在 B_2 與 B_3 之內，A_2 皆顯著高於 A_1 (圖 8)。上述二點即為 $A \times B$ 交互作用顯著之原因。在模型診斷方面，六組皆無顯著偏離常態且變方無顯著差異，符合 ANOVA 及 Tukey's range test 之前題。[4]

[4]此例中的簡單主效應其實與第 5 節所述的單因子均值比較相同，可一併參考該節。

表 7: Two-way TYPE III ANOVA 的 variance 拆解。

	SS	DF	*F*	*P*
A	40.93	1	71.01	< 0.001
B	49.28	2	42.74	< 0.001
A × *B*	7.56	2	6.56	0.006
Residuals	11.53	20		

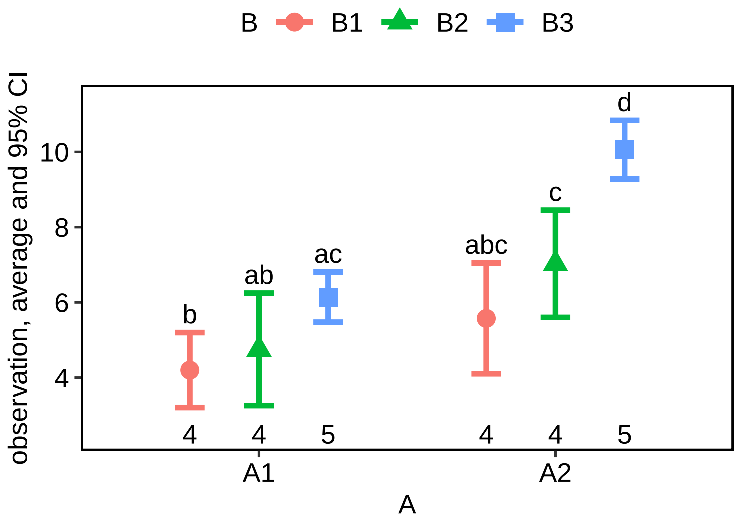

圖 8: 雙因子之因子實驗於各組的觀測結果、平均及 95% 信賴區間。上方字母為多重比較的分群結果；若任二組存在相同字母則表示不存在顯著差異，反則反之。下方數字表示各組之重覆數。

6.2 二因子無交互作用因子實驗

(todo)

7 簡單線性迴歸

以下樣本

x	8.8	10.3	11.1	7.7	10.4	10.5	9.4	9.5
y	17	19.7	21.7	14.4	20	21.1	19.8	18.9

中，y 為應變數，x 為自變數，建立 $y = \beta_0 + \beta_1 x + \varepsilon$ 的簡單線性迴歸。

簡單線性迴歸之結果如表 8 及 圖 9a。結果顯示，每 x 增加 1 單位使 y 平均顯著增加 2.069 單位，應拒絕 $H_0 : \beta_1 = 0$ (表 8a)。就效果量而言，自變數可解釋 $R^2 = 92.4\%$ 之變異量，屬高效果量。就迴歸診斷而言，存在單一觀測值之標準化殘差 > 2 (圖 9b)，且殘差之 Q-Q 圖顯示殘差呈輕微右偏態 (圖 9c)，但 Shapiro-Wilk test 顯示殘差並未顯著偏離常態分布 ($W = 0.848$，$P = 0.090$)，模型配適尚可。結論是，自變數顯著地增加應變數且效果明顯。

表 8: 簡單線性迴歸之結果。

Variable	Estimate ± Std. Error	*T* (DF = 6)	*P*	95% CI
Intercept	−1.016 ± 2.408	−0.422	0.688	[−6.909, 4.877]
x	2.069 ± 0.247	8.389	< 0.001	[1.465, 2.672]

8 簡單相關

8.1 雙常態分布情況

以下樣本

x_1	8.8	10.3	11.1	7.7	10.4	10.5	9.4	9.5
x_2	17	19.7	21.7	14.4	20	21.1	19.8	18.9

中，分析二變數之間的相關性。

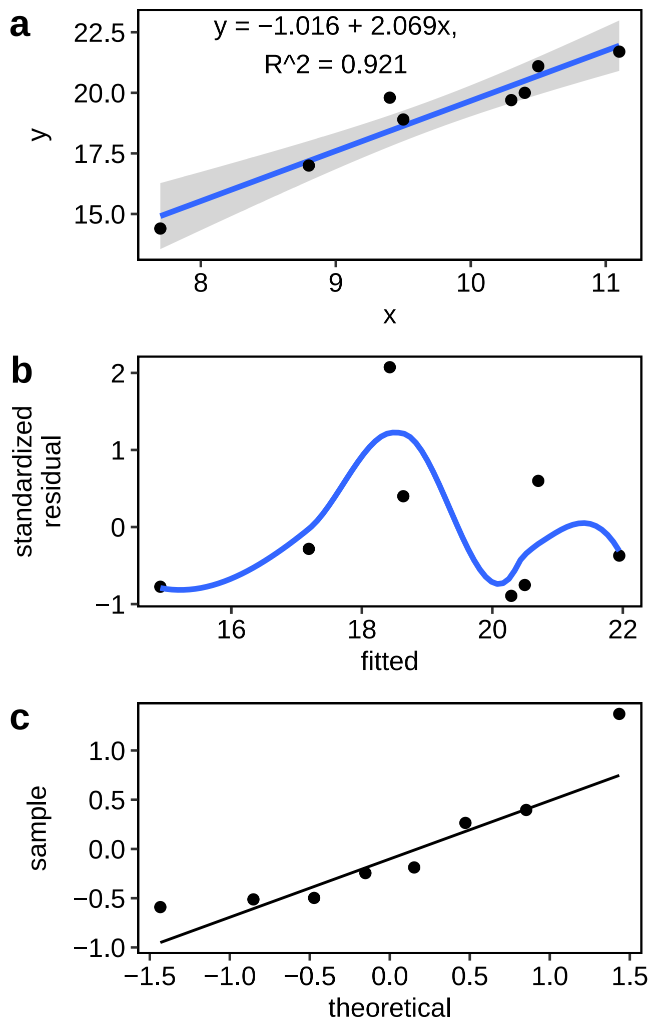

圖 9: 簡單線性迴歸之散布圖及迴歸線 (a)、標準化殘差圖 (b) 及殘差 Q-Q plot (c)。圖中灰色區域為 95% confidence pointwise band。

首先以 Shapiro-Wilk multivariate normality test 檢驗 x_1 與 x_2 是否偏離雙變量常態分布，結果顯示不能拒絕 H_0: x_1 與 x_2 之母體聯合分配為常態 ($W = 0.860$, $P = 0.120$)，故可計算 Pearson correlation。結果指出，$r = 0.960$ 屬高度正相關且應拒絕 $H_0 : \rho = 0$ (95% CI = $[0.789, 0.993]$, $T = 8.389$, DF = 6, $P < 0.001$；圖 10)。結論是，x_1 與 x_2 間存在顯著的高度正向線性相關性。

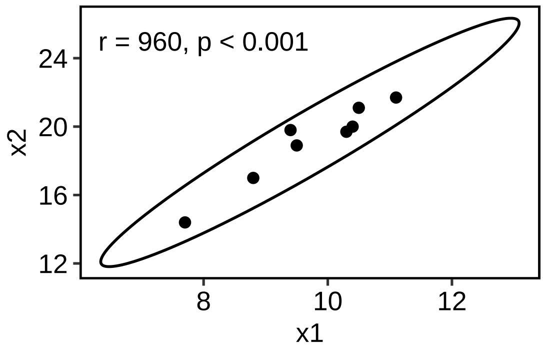

圖 10: x_1 與 x_2 散布圖。圖中橢圓區域表示相關性之 95% confidence ellipse。

8.2 次序相關

以下樣本

x_1	7.5	5.4	5.9	6.1	7.9	7.9	6.6	6
x_2	0.3	0.2	4.4	2.7	0.1	1	0.3	0.5

中，分析二變數之間的相關性。

首先以 Shapiro-Wilk multivariate normality test 檢驗 x_1 與 x_2 是否偏離雙變量常態分布，結果顯示應拒絕 $H_0 : x_1$與x_2之母體聯合分配為常態 ($W = 0.739$, $P = 0.006$), 故計算 Spearman's rank correlation coefficient。結果指出，$r_s = -0.241$ 屬低度負相關且無法拒絕 $H_0 : \rho_s = 0$ ($S = 104.24$, $P = 0.565$；圖 11)。結論是，x_1 與 x_2 間不存在顯著次序相關性。

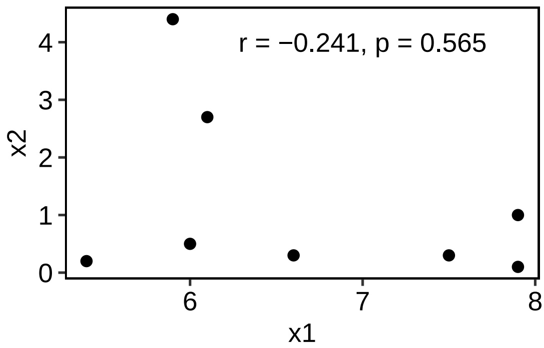

圖 11: x_1 與 x_2 散布圖。

9 卡方適合度檢驗

隨機抽樣 40 人，發現各血型次數為 $[n_{\mathrm{O}}, n_{\mathrm{A}}, n_{\mathrm{B}}, n_{\mathrm{AB}}] = [20, 15, 3, 2]$。檢驗 $H_0 : [\pi_{\mathrm{O}}, \pi_{\mathrm{A}}, \pi_{\mathrm{B}}, \pi_{\mathrm{AB}}] = [0.4, 0.3, 0.2, 0.1]$。

四種血型的母體比例估計如表 9。卡方適合度檢驗結果指出 $\chi^2 = 5.875$ (DF = 3)。由於血型 B 和 AB 之期望值極低 (< 1)，故採用 4,999 次 permutation 得 $P = 0.112$ 而不能拒絕 H_0。

表 9: 血型頻率與比例估計。

Blood type	Frequency	Proportion	95% CI
A	15	0.375	$[0.216, 0.598]$
AB	2	0.050	$[0.008, 0.154]$
B	3	0.075	$[0.019, 0.194]$
O	20	0.500	$[0.312, 0.752]$

10　卡方獨立性檢驗

隨機抽樣 350 人，發現居住地與收入分級的次數為

	income		
location	low	median	high
A	2	10	68
B	30	10	50
C	60	19	101

試檢驗居住地與收入分級之機率相互獨立 $H_0 : \pi_{ij} = \pi_i \pi_j$。

居住地與收入的邊際比例如圖 12。卡方適合度檢驗結果指出應拒絕 H_0 表示居住地與收入並不獨立而存在關聯性 ($\chi^2 = 30.849$, DF $= 4$, $P < 0.001$)。Cohen's $W = 0.297$ 顯示居住地與收入之關聯性幾乎達到中等程度 (0.3 即達中度關聯)。若以居住地為多重比較的目標因子，再以卡方適合度比較居住地間比例並以 Holm-Bonferroni method 控制整體型一錯誤率，結果顯示居住地 A 顯著不同於居住地 B 及居住地 C 但 B 與 C 間無顯著差異 (圖 12a)。綜合上述，居住地與收入並不獨立而存在顯著關聯性，主要由居住地 A 與另外二地點的差異 (若以居住地為比較對象) 造成。

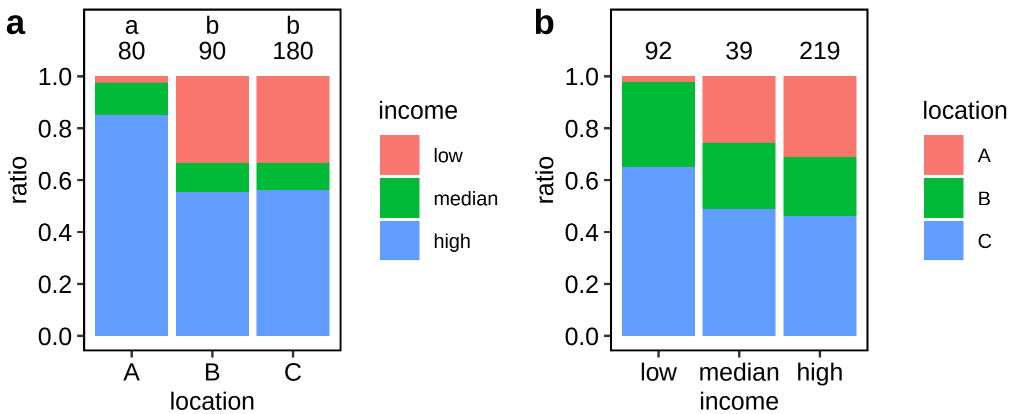

圖 12: 按居住地 (a) 與收入 (b) 的邊際比例。柱上方的數字表示該居住地或收入的樣本數。柱上方的小寫字母表示居住地間多重比較之結果；若任二組存在相同字母則表示不存在顯著差異，反則反之。

A R code

以下為本文中所有產生資料、進行分析、製作表格與繪圖之 R code，亦可至 Github 下載。[5]

```r
library(lsr)
library(coin)
library(ggpubr)
library(data.table)
library(multcomp)
library(multcompView)
library(xtable)
library(userfriendlyscience)
library(mvnormtest)
library(car)
library(FSA)
library(rstatix)
library(rcompanion)
library(effsize)
sessionInfo()
# R version 3.6.1 (2019-07-05)
# Platform: x86_64-w64-mingw32/x64 (64-bit)
# Running under: Windows 10 x64 (build 17763)
#
# Matrix products: default
#
# locale:
# [1] LC_COLLATE=English_United States.1252  LC_CTYPE=English_United States.1252
# [3] LC_MONETARY=English_United States.1252 LC_NUMERIC=C
# [5] LC_TIME=English_United States.1252
#
# attached base packages:
# [1] stats     graphics  grDevices utils     datasets  methods   base
#
# other attached packages:
#  [1] effsize_0.7.6             rcompanion_2.3.7          rstatix_0.3.0
#  [4] FSA_0.8.26                car_3.0-5                 carData_3.0-3
#  [7] mvnormtest_0.1-9          userfriendlyscience_0.7.2 xtable_1.8-4
# [10] multcompView_0.1-7        multcomp_1.4-10           TH.data_1.0-10
# [13] MASS_7.3-51.4             mvtnorm_1.0-11            data.table_1.12.6
# [16] ggpubr_0.2.4              magrittr_1.5              ggplot2_3.2.1
# [19] coin_1.3-1                survival_3.1-7            lsr_0.5



## normal one-sample test
set.seed(1234)
```

[5] https://github.com/chenpanliao/report-statisti!cal-results-TC/blob/master/plot/report-results.R。

```r
x <- rnorm(10, 10, 1) %>% round(1)
paste0(x, collapse = ", ")
mean(x)
sd(x)
shapiro.test(x)
t.test(x)
t.test(x, mu = 9)
cohensD(x, mu = 9)


## non-normal one-sample test
set.seed(1234)
x <- rexp(10, 1) %>% round(2)
paste0(x, collapse = ", ")
mean(x)
sd(x)
shapiro.test(x)
wilcox.test(x, mu = 2)


## normal paired test
set.seed(1234)
x1 <- rnorm(8, 10, 1) %>% round(1)
x2 <- round(x1 + rnorm(8, 1), 1)
shapiro.test(x1 - x2)
t.test(x1, x2, paired = T, mu = 0.1)
paste0(x1, collapse = " & ")
paste0(x2, collapse = " & ")
mean(x1)
sd(x1)
mean(x2)
sd(x2)
mean(x1 - x2)
sd(x1 - x2)
cohensD(x1 - x2, mu = 0.5)
d.plot <-
  data.table(
    observation = c(x1, x2),
    group = gl(2, 8, labels = c("x1", "x2")),
    block = gl(8, 1, 16)
  )
f1 <-
  ggplot(d.plot, aes(group, observation)) +
  geom_boxplot(width = 0.2,
               position = position_nudge(x = c(-0.2, 0.2)),
               outlier.shape = NA) +
  geom_jitter(width = 0) +
  geom_segment(
    aes(
```

```
      x = 1,
      xend = 2,
      y = `x1`,
      yend = `x2`
    ),
    dcast(d.plot, block ~ group, value.var = "observation")
  ) +
  theme_pubr(10, border = T)
f2 <-
  dcast(d.plot, block ~ group, value.var = "observation") %>%
  .[, .(difference = x1 - x2), by = block] %>%
  ggplot(aes(x = 0, y = difference)) +
  geom_boxplot(width = 0.1) +
  geom_jitter(width = 0.05) +
  theme_pubr(10, border = T) +
  theme(
    axis.ticks.x = element_blank(),
    axis.text.x = element_blank(),
    axis.title.x = element_blank()
  ) +
  ylab("difference\n(x1 - x2)") +
  xlim(c(-0.15, 0.15))
windows(4, 2.0)
ggarrange(
  f1,
  f2,
  nrow = 1,
  labels = "auto",
  align = "h",
  widths = c(2, 1.2)
)
ggsave("normal_paired_test.pdf")


## non-normal paired test
set.seed(1212314)
x1 <- runif(8, 5, 8) %>% round(1)
x2 <- round(x1 + runif(8,-1, 1), 1)
shapiro.test(x1 - x2)
wilcox.test(x1,
            x2,
            paired = T,
            mu = 1,
            exact = F)
t.test(x1, x2, paired = T, mu = 0.1)
paste0(x1, collapse = " & ")
paste0(x2, collapse = " & ")
mean(x1)
sd(x1)
```

```
mean(x2)
sd(x2)
mean(x1 - x2)
sd(x1 - x2)
cohensD(x1 - x2, mu = 0.5)
table(x1 - x2 < 1)
d.plot <-
  data.table(
    observation = c(x1, x2),
    group = gl(2, 8, labels = c("x1", "x2")),
    block = gl(8, 1, 16)
  )
f1 <-
  ggplot(d.plot, aes(group, observation)) +
  geom_boxplot(width = 0.2,
               position = position_nudge(x = c(-0.2, 0.2)),
               outlier.shape = NA) +
  geom_jitter(width = 0) +
  geom_segment(
    aes(
      x = 1,
      xend = 2,
      y = `x1`,
      yend = `x2`
    ),
    dcast(d.plot, block ~ group, value.var = "observation")
  ) +
  theme_pubr(10, border = T)
f2 <-
  dcast(d.plot, block ~ group, value.var = "observation") %>%
  .[, .(difference = x1 - x2), by = block] %>%
  ggplot(aes(x = 0, y = difference)) +
  geom_boxplot(width = 0.1) +
  geom_jitter(width = 0.05) +
  theme_pubr(10, border = T) +
  theme(
    axis.ticks.x = element_blank(),
    axis.text.x = element_blank(),
    axis.title.x = element_blank()
  ) +
  ylab("difference\n(x1 - x2)") +
  xlim(c(-0.15, 0.15))
windows(4, 2.0)
ggarrange(
  f1,
  f2,
  nrow = 1,
  labels = "auto",
  align = "h",
```

```
  widths = c(2, 1.2)
)
ggsave("non-normal_paired_test.pdf")


## normal independent two-sample test
set.seed(124)
x1 <- rnorm(6, 10, 1) %>% round(1)
x2 <- rnorm(7, 9, 1) %>% round(1)
shapiro.test(x1)
shapiro.test(x2)
t.test(x1, x2)
paste0(x1, collapse = " & ")
paste0(x2, collapse = " & ")
mean(x1)
sd(x1)
mean(x2)
sd(x2)
cohensD(x1, x2)
d.plot <-
  data.table(
    observation = c(x1, x2),
    group = c(rep("x1", 6), rep("x2", 7))
  )
windows(2, 2.0)
ggplot(d.plot, aes(group, observation)) +
  geom_boxplot(width = 0.2,
               outlier.shape = NA) +
  geom_jitter(width = 0.3) +
  theme_pubr(10, border = T)
ggsave("normal_independent_test.pdf")


## non-normal independent two-sample test
set.seed(6324)
x1 <- runif(8, 0, 1) %>% round(1)
x2 <- runif(7, 0, 2) %>% round(1)
shapiro.test(x1)
shapiro.test(x2)
paste0(x1, collapse = " & ")
paste0(x2, collapse = " & ")
mean(x1)
sd(x1)
mean(x2)
sd(x2)
cliff.delta(x1, x2)
d.plot <-
  data.table(
    observation = c(x1, x2),
```

```
    group = c(rep("x1", 8), rep("x2", 7))
  )
windows(2, 2.0)
ggplot(d.plot, aes(group, observation)) +
  geom_boxplot(width = 0.2,
               outlier.shape = NA) +
  geom_jitter(width = 0.3) +
  theme_pubr(10, border = T)
ggsave("non-normal_independent_test.pdf")


## oneway ANOVA
set.seed(364)
d <-
  data.table(y = round(c(rep(5, 6), rep(6, 5), rep(7, 7)) + rnorm(18), 2),
             group = factor(c(rep("x1", 6), rep("x2", 5), rep("x3", 7))))
tapply(d$y, d$group, shapiro.test)
bartlett.test(y ~ group, d)
d[, paste0(y, collapse = " & "), by = group]
d[, .(Mean = mean(y),
      SD = sd(y),
      n = length(y)), by = group] %>%
  as.data.frame %>%
  xtable(
    .,
    digits = 3,
    auto = T,
    label = "table:oneway_ANOVA",
    caption = "獨立三樣本的描述性統計。"
  )
fit <- aov(y ~ group, d)
summary(fit)
cohens_d(d, y ~ group, var.equal = T) %>%
  as.data.table %>%
  .[, comparison := paste0(group2, "-", group1)] %>%
  .[, .(comparison, effsize)] %>%
  merge(.,
        TukeyHSD(fit, "group")$group %>%
          as.data.table(keep.rownames = "comparison"),
        by = "comparison") %T>%
  print %>%
  xtable(
    digits = 3,
    auto = T,
    label = "table:oneway_ANOVA_post",
    caption = "獨立三樣本的事後多重比較。"
  )
fit.mult <-
  TukeyHSD(fit, "group")$group[, "p adj"] %>%
```

```r
  multcompLetters %>%
  .$Letters %>%
  data.table(group = names(.), rank = .) %>%
  merge(., d[, .(max.val = max(y)), by = group], by = "group")
windows(2.5, 2.0)
ggplot(d, aes(group, y)) +
  geom_boxplot(width = 0.2,
               outlier.shape = NA) +
  geom_jitter(width = 0.3) +
  geom_text(aes(group, max.val + 0.5, label = rank),
            fit.mult,
            size = 10 * 0.352777778) +
  theme_pubr(10, border = T)
ggsave("oneway_ANOVA.pdf")


## Welch's ANOVA
set.seed(12234)
d <-
  data.table(y = round(rnorm(
    18,
    mean = c(rep(4, 6), rep(6, 5), rep(7, 7)),
    sd = c(rep(1, 6), rep(2, 5), rep(3, 7))
  ), 2),
  group = factor(c(rep("x1", 6), rep("x2", 5), rep("x3", 7))))
tapply(d$y, d$group, shapiro.test)
bartlett.test(y ~ group, d)
d[, paste0(y, collapse = " & "), by = group]
d[, .(Mean = mean(y),
      SD = sd(y),
      n = length(y)), by = group] %>%
  as.data.table %>%
  xtable(
    .,
    digits = 3,
    auto = T,
    label = "table:Welch_ANOVA",
    caption = "獨立三樣本的描述性統計。"
  )
aov(y ~ group, data = d) %>% summary
oneway.test(y ~ group, data = d)
cohens_d(d, y ~ group, var.equal = F) %>%
  as.data.table %>%
  .[, comparison := paste0(group2, "-", group1)] %>%
  .[, .(comparison, effsize)] %>%
  merge(
    .,
    posthocTGH(d$y, d$group, digits = 3)$output$games.howell %>%
      as.data.table(keep.rownames = "comparison")
```

```r
  ) %T>%
  print %>%
  xtable(
    digits = 3,
    auto = T,
    label = "table:Welch_ANOVA_post",
    caption = "獨立三樣本的事後多重比較。"
  )
fit.mult <-
  mc$p %>%
  set_names(rownames(mc)) %>%
  multcompLetters %>%
  .$Letters %>%
  data.table(group = names(.), rank = .) %>%
  merge(., d[, .(max.val = max(y)), by = group], by = "group")
windows(2.5, 2.0)
ggplot(d, aes(group, y)) +
  geom_boxplot(width = 0.2,
               outlier.shape = NA) +
  geom_jitter(width = 0.3) +
  geom_text(aes(group, max.val + 1, label = rank),
            fit.mult,
            size = 10 * 0.352777778) +
  theme_pubr(10, border = T)
ggsave("Welch_ANOVA.pdf")


## Kruskal-Wallis Rank Sum Test
set.seed(1132234)
d <-
  data.table(y = c(rexp(8, 0.2), rexp(7, 0.5), rexp(6, 1)) %>% round(1),
             group = c(rep("x1", 8), rep("x2", 7), rep("x3", 6)))
d[, shapiro.test(y), by = group]
d[, paste0(y, collapse = " & "), by = group]
d[, .(Mean = mean(y),
      SD = sd(y),
      n = length(y)), by = group] %>%
  as.data.frame %>%
  xtable(
    .,
    digits = 3,
    auto = T,
    label = "table:rank_oneway",
    caption = "獨立三樣本的描述性統計。"
  )
kruskal.test(y ~ group, d)
kruskal_effsize(d, y ~ group)
combn(d$group %>% unique, 2) %>%
  apply(., 2, function(x){
```

```
    cliff.delta(y ~ group, d[group == x[1] | group == x[2]])$estimate
  }) %>%
  data.table(`Cliff's d` = .) %>%
  .[, Comparison := combn(d$group %>% unique, 2, paste0, collapse = " - ")] %>%
  merge(., dunnTest(y ~ group, d)$res %>% as.data.table, by = "Comparison") %T>%
  print %>%
  xtable(caption = "Dunn's Kruskal-Wallis多重比較之結果。",
         label = "table:rank_oneway_post",
         digits = 3)
fit.mult <-
  dunnTest(y ~ group, d)$res[, "P.adj"] %>%
  set_names(dunnTest(y ~ group, d)$res[, "Comparison"] %>% gsub(" ", "", .)) %>%
  multcompLetters %>%
  .$Letters %>%
  data.table(group = names(.), rank = .) %>%
  merge(., d[, .(max.val = max(y)), by = group], by = "group")
windows(2.5, 2.0)
ggplot(d, aes(group, y)) +
  geom_boxplot(width = 0.2,
               outlier.shape = NA) +
  geom_jitter(width = 0.3) +
  geom_text(aes(group, max.val + 1, label = rank),
            fit.mult,
            size = 10 * 0.352777778) +
  theme_pubr(10, border = T)
ggsave("rank_oneway.pdf")


## twoway ANOVA factorial design
set.seed(1224)
d <- data.table(
  A = c(rep("A1", 13), rep("A2", 13)),
  B = c(rep("B1", 4), rep("B2", 4), rep("B3", 5)) %>% rep(2),
  Y = rnorm(26, mean = c(
    rep(5, 4), rep(5, 4), rep(6, 5), rep(6, 4), rep(6, 4), rep(10, 5)
  )) %>% round(1)
)
d[, shapiro.test(Y), by = .(A, B)]
bartlett.test(Y ~ interaction(A, B), data = d)
fit.full <-
  lm(Y ~ A * B,
     data = d,
     contrasts = list(A = contr.sum, B = contr.sum))
summary(fit.full)
Anova(fit.full, type = 3) %>%
  xtable(caption = "Twoway ANOVA之變方分析表。", label = "table:twowayANOVA")
drop1(fit.full, test = "F")
d[, group := paste0(A, B)]
fit <- aov(Y ~ group, data = d)
```

```r
summary(fit)
TukeyHSD(fit, "group")$group
TukeyHSD(fit, "group")$group %>%
  xtable(
    digits = 3,
    auto = T,
    label = "table:twoway_ANOVA_post",
    caption = "$3\tiems2$因子實驗之簡單主效應事後多重比較。"
  )
fit.mult <-
  TukeyHSD(fit, "group")$group[, "p adj"] %>%
  multcompLetters %>%
  .$Letters %>%
  data.table(group = names(.), rank = .) %>%
  merge(., d[, .(max.val = max(Y)), by = group], by = "group")
d.summary <-
  d[, .(
    average = mean(Y),
    SD = sd(Y),
    CI.lower = t.test(Y)$conf.int[1],
    CI.upper = t.test(Y)$conf.int[2],
    n = .N
  ), by = .(A, B, group)] %>%
  merge(., fit.mult, by = "group") %T>%
  print
ggplot(fortify(fit.full), aes(sample = .resid)) +
  stat_qq() +
  stat_qq_line() +
  theme_pubr(10, border = T)

pd <- position_dodge(0.7)
windows(4, 3)
ggplot(d.summary, aes(A, average)) +
  geom_errorbar(
    aes(ymin = CI.lower, ymax = CI.upper, color = B),
    width = 0.3,
    size = 1,
    position = pd
  ) +
  geom_point(aes(shape = B, color = B), position = pd, size = 3) +
  geom_text(
    aes(
      y = CI.upper + 0.5,
      label = rank,
      group = B
    ),
    position = pd,
    size = 10 * 0.352777778
  ) +
```

```
  geom_text(aes(y = 2.5, label = n, group = B),
            position = pd,
            size = 10 * 0.352777778) +
  geom_jitter(
    aes(A, Y, shape = B),
    data = d,
    position = position_jitterdodge(jitter.width = 0.15, dodge.width = 0.7),
    size = 1.2,
    color = 8
  ) +
  theme_pubr(10, border = T, legend = "top") +
  ylab("observation, average and 95% CI") +
  theme(legend.text = element_text(size = 10))
ggsave("twoway_ANOVA.pdf")


## simple linear regression
set.seed(1234)
x <- rnorm(8, 10) %>% round(1)
y <- (x * 2 + rnorm(8)) %>% round(1)
d <- data.table(x, y)
fit <- lm(y ~ x, d)
summary(fit)
confint(fit)
shapiro.test(fit$residuals)
d[, paste0(x, collapse = " & ")]
d[, paste0(y, collapse = " & ")]
fit %>% {
  cbind(
    (.) %>% summary %>% .$coefficients ,
    confint(.)
  )
} %>%
  xtable(caption = "簡單線性迴歸之結果。",
         label = "table:simple_regression",
         digits = 3)
f1 <-
  ggplot(d, aes(x, y)) +
  geom_smooth(method = "lm") +
  geom_point() +
  annotate(
    "text",
    label = "y = -1.016 + 2.069x,\nR^2 = 0.921",
    x = 9,
    y = 22,
    size = 10 * 0.352777778
  ) +
  theme_pubr(10, border = T)
f2 <-
```

```
  fortify(fit) %>%
  .[order(.$.fitted)] %>%
  ggplot(., aes(.fitted, scale(.resid))) +
  geom_point() +
  geom_smooth(se = F) +
  xlab("fitted") +
  ylab("standardized\nresidual") +
  theme_pubr(10, border = T)
f3 <-
  ggplot(d, aes(sample = fit$residuals)) +
  stat_qq() +
  stat_qq_line() +
  theme_pubr(10, border = T)
windows(3.5, 5.5)
ggarrange(
  f1,
  f2,
  f3,
  nrow = 3,
  labels = "auto",
  align = "hv",
  widths = c(1, 1)
)
ggsave("simple_regression.pdf")


## simple linear cor
set.seed(1234)
x1 <- rnorm(8, 10) %>% round(1)
x2 <- (x1 * 2 + rnorm(8)) %>% round(1)
d <- data.table(x1, x2)
d[, paste0(x1, collapse = " & ")]
d[, paste0(x2, collapse = " & ")]
mshapiro.test(d %>% as.matrix %>% t)
cor.test(d$x1, d$x2)
fit <- lm(x2 ~ x1, d)
windows(3, 2.0)
ggplot(d, aes(x1, x2)) +
  geom_path(data =
              dataEllipse(
                x1,
                x2,
                draw = F,
                levels = 0.95,
                segments = 500
              ) %>%
              as.data.table,
            aes(x, y)) +
  geom_point() +
```

```
  annotate(
    "text",
    label = "r = 960, p < 0.001",
    x = 8,
    y = 25,
    size = 10 * 0.352777778
  ) +
  theme_pubr(10, border = T)
ggsave("simple_cor.pdf")


## spearman correlation
set.seed(125)
x1 <- runif(8, 5, 8) %>% round(1)
x2 <- rexp(8) %>% round(1)
d <- data.table(x1, x2)
mshapiro.test(d %>% as.matrix %>% t)
d[, paste0(x1, collapse = " & ")]
d[, paste0(x2, collapse = " & ")]
cor.test(d$x1, d$x2, method = "spearman")
windows(3, 2.0)
ggplot(d, aes(x1, x2)) +
  geom_point() +
  annotate(
    "text",
    label = "r = -0.241, p = 0.565",
    x = 7,
    y = 4,
    size = 10 * 0.352777778
  ) +
  theme_pubr(10, border = T)
ggsave("spearman_cor.pdf")


## chi-squared goodness of fit
obs.val <- c(20, 15, 3, 2)
exp.p <- c(4, 3, 2, 1) %>% divide_by(sum(.))
d <-
  data.table(observation = obs.val,
             expectation = exp.p * sum(obs.val),
             blood = c("O", "A", "B", "AB")) %>%
  melt(measure.vars = c("observation", "expectation"),
       value.name = "frequency") %>%
  .[, proportion := frequency / sum(frequency), by = variable] %T>%
  print
glm(
  frequency ~
    -1 + blood +
    offset(sum(d[variable == "observation"]$frequency) %>% log %>% rep(4)),
```

```
  family = poisson,
  data = d[variable == "observation"]
) %>%
  confint %>%
  exp %>%
  as.data.table(keep.rownames = "blood") %>%
  .[, blood := gsub("blood", "", blood)] %>%
  merge(d[variable == "observation"], .) %>%
  .[, variable := NULL] %>%
  xtable(caption = "血型頻率與比例估計。",
         label = "table:chisq_goodness",
         digits = 3)
chisq.test(
  obs.val,
  p = exp.p,
  rescale.p = T,
  simulate.p.value = T,
  B = 4999
)


## chi-squared independent test
d <-
  matrix(c(2, 10, 68, 30, 10, 50, 60, 19, 101), 3, byrow = T) %>%
  set_rownames(c("A", "B", "C")) %>%
  set_colnames(c("low", "median", "high"))
chisq.test(d)
chisq.test(d, simulate.p.value = T, B = 4999)
# ufs::cramersV(d)
# ufs::confIntV(d)
# rcompanion::cramerV(d, ci = T, bias.correct = T)
chisq.val <- chisq.test(d)$statistic
C <- sqrt(chisq.val/(chisq.val + sum(d)))
W <- sqrt(C^2 / (1-C^2))
W # Cohan's W
mc <-
  combn(1:nrow(d), 2, function(x) {
    chisq.test(d[, c(x[1], x[2])], simulate.p.value = T, B = 4999)
  }) %>%
  .[c(1, 3), ] %>%
  as.matrix %>%
  set_rownames(c("chisq", "p")) %>%
  set_colnames(combn(rownames(d), 2, function(x) {
    paste0(x[1], "-", x[2])
  })) %T>%
  print %>%
  .["p",] %>%
  unlist %>%
  p.adjust %>%
```

```r
  multcompLetters %>%
  .$Letters %T>%
  print
d %>% as.table %>% prop.table
d %>% as.table %>% plot
d.long <-
  merge(
    d %>%
      reshape2::melt(
        varnames = c("location", "income"),
        value.name = "frequency"
      ) %>%
      as.data.table,
    d %>% as.table %>% prop.table(margin = 1) %>%
      reshape2::melt(
        varnames = c("location", "income"),
        value.name = "ratio by location"
      ) %>%
      as.data.table,
    by = c("location", "income")
  ) %>%
  merge(
        d %>% as.table %>% prop.table(margin = 2) %>%
      reshape2::melt(
        varnames = c("location", "income"),
        value.name = "ratio by income"
      ) %>%
      as.data.table
  ) %T>%
  print
f1 <-
  ggplot(d.long, aes(location, `ratio by location`)) +
  geom_col(aes(fill = income)) +
  theme_pubr(10, border = T) +
  annotate(
    "text",
    label = rowSums(d),
    x = 1:3,
    y = 1.1,
    size = 10 * 0.352777778
  ) +
  annotate(
    "text",
    label = mc,
    x = 1:3,
    y = 1.2,
    size = 10 * 0.352777778
  ) +
  scale_y_continuous("ratio", breaks  = seq(0, 1, 0.2), limits = c(0, 1.2))
```

```r
f2 <-
  ggplot(d.long, aes(income, `ratio by income`)) +
  geom_col(aes(fill = location)) +
  theme_pubr(10, border = T) +
  annotate(
    "text",
    label = colSums(d),
    x = 1:3,
    y = 1.1,
    size = 10 * 0.352777778
  ) +
  scale_y_continuous("ratio", breaks  = seq(0, 1, 0.2), limits = c(0, 1.2))
windows(6, 2.5)
ggarrange(f1, f2, nrow = 1, align = "hv", labels = "auto", legend = "right")
ggsave("chisq_independent.pdf")
```