

試利用心血管疾病資料 (CVD ALL) 分析沒有心血管疾病成人之腰圍與收縮壓的關係。(顯著水準設定為 0.05)

1. 請問腰圍與收縮壓的皮爾生相關係數為何？兩者是否存在顯著的線性關係？

答：

(1) 資料檔整理：因題目中分析的對象為沒有心血管疾病的成人，故須將符合此條件的資料篩出。

R-web 分析步驟：

資料處理→資料篩選(篩選條件：心血管疾病=0；建議另存新檔 CVD_noHD)

分析結果

- 資料名稱：CVD_All
- 篩選規則：心血管疾病 == "0"
- 保留變數：ID, 心血管疾病, 年齡, 性別, 追蹤時間, 腰圍, 收縮壓, 舒張壓, 空腹血糖, 高密度脂蛋白, 三酸甘油酯, 檳榔, 飲酒, 家族病史, 抽菸, 抽菸量
- 儲存位置：[使用者個人資料檔](#) - CVD_noHD

(2)腰圍與收縮壓變項皆為連續型(或稱數值變項)，此處要求計算皮爾生相關係數

R-web 分析步驟：

分析方法→相關暨列聯表分析→皮爾生相關係數

→步驟一：資料匯入(CVD_noHD)

→步驟二：參數設定(檢定變數：腰圍、收縮壓)

→進階選項(可免設定) →開始分析

分析結果

- 分析方法：皮爾生相關係數
- 資料名稱：CVD_noHD
- 變數名稱：腰圍, 收縮壓
- 虛無假設：相關係數 $\rho = 0$ (雙尾檢定)
- 計算時間：0.034 秒
- 皮爾生相關係數矩陣^I：

	腰圍	收縮壓
腰圍	1.000	0.428
	0.000	0.000
	57055	56612
收縮壓	0.428	1.000
	0.000	0.000
	56612	57387

- I：表格內容為皮爾生相關係數 / P-值 / 樣本數

結論：計算出的皮爾生相關係數為 0.428，檢定 P 值為 0，此值小於顯著水準 0.05，表示此相關係數為顯著有意義；但相關係數值僅有 0.428 表示腰圍與收縮壓之間的線性關係程度不高。

2. 請問腰圍與收縮壓的斯皮爾曼等級相關為何？兩者是否存在顯著的等級相關？

答：

腰圍與收縮壓變項皆為連續型(或稱數值變項)，此處要求計算斯皮爾曼等級相關係數

R-web 分析步驟：

分析方法→相關暨列聯表分析→斯皮爾曼等級相關係數

→步驟一：資料匯入(CVD_noHD)

→步驟二：參數設定(檢定變數：腰圍、收縮壓)

→進階選項(可免設定) →開始分析

分析結果

- 分析方法：斯皮爾曼等級相關係數
- 資料名稱：CVD_noHD
- 變數名稱：腰圍, 收縮壓
- 虛無假設：相關係數 $\rho = 0$ (雙尾檢定)
- 計算時間：0.131 秒
- 斯皮爾曼相關係數矩陣^I：

	腰圍	收縮壓
腰圍	1.000	0.458
	0.000	0.000
	57055	56612
收縮壓	0.458	1.000
	0.000	0.000
	56612	57387

- I：表格內容為斯皮爾曼等級相關係數 / P-值 / 樣本數

結論：計算出的斯皮爾曼等級相關係數為 0.458，檢定 P 值為 0，此值小於顯著水準 0.05，表示此相關係數為顯著有意義；但相關係數值僅有 0.458 表示腰圍與收縮壓之間的線性關係程度不高。

3. 試利用簡單線性迴歸模型建立腰圍預測收縮壓之模型。

I. 請問此模型為何？

II. 腰圍是否與收縮壓有顯著相關？此模型腰圍解釋了收縮壓變異的多少百分比？

III. 若有一人腰圍為 100 公分，請預測此人平均而言收縮壓何？

答：

I. 題目中須以腰圍預測收縮壓，故收縮壓為依變數，腰圍為自變數，可寫出迴歸模型

$Y_i = \alpha + \beta X_i + \varepsilon_i$ ， $\varepsilon_i \sim N(0, \sigma^2)$ ，其中

Y_i ：收縮壓

X_i ：腰圍

II. 此處需利用 R-web 的迴歸模式來解釋腰圍與收縮壓之間的關係

R-web 分析步驟：

分析方法→迴歸模式→迴歸分析

→步驟一：資料匯入(CVD_noHD)

→步驟二：參數設定(依變數：收縮壓、自變數：腰圍)

→進階選項(可免設定) →開始分析

分析結果

- 分析方法：迴歸分析
- 資料名稱：CVD_noHD
- 依變數名稱：收縮壓
- 自變數名稱：腰圍
- 顯著水準：0.05
- 計算時間：20.84 秒
- 迴歸模式的變異數分析：

虛無假設：迴歸模式不顯著						
來源	平方和	自由度	均方和	F 檢定統計量	臨界值	p-值 I
source	sum of squares	d.f.	mean square	F-statistic	F(d.f.1,d.f.2,1- α)	p-value
迴歸	4264673.0255	1	4264673.0255	12682.9367	3.8416	< 1e-04
誤差	19035271.1745	56610	336.2528			***
總和	23299944.2	56611				
total						
判定係數(R-square)：18.3 %						
調整判定係數(adjusted R-square)：18.3 %						

- I：顯著性代碼： '***' : < 0.001, '**' : < 0.01, '*' : < 0.05, '#': < 0.1

- 迴歸係數估計^I：

係數 coefficient	估計值 estimation	標準差 std. err.	t 檢定統計量 t-statistic	p 值 II p-value	參數的 95% 信賴區間 95% C.I. for estimations	
					下界 lower	上界 upper
(截距項)	58.5096	0.5717	102.342	< 2.22e-16 ***	57.389	59.6301
腰圍	0.8189	0.0073	112.6185	< 2.22e-16 ***	0.8047	0.8332

- I：依變數為收縮壓，模式包含常數項

II：顯著性代碼： '***' : < 0.001, '**' : < 0.01, '*' : < 0.05, '#': < 0.1

結論：分析結果列出變異數分析表及迴歸係數估計表，可整理出結果為

- (1) 變異數分析表中的 P 值(<e-4)小於顯著水準，拒絕虛無假設，此迴歸模式為顯著
- (2) 迴歸係數估計表中的腰圍檢定 P 值(< 2.22e-16)小於顯著水準，表示自變項腰圍對於依變項是有影響的
- (3) 此迴歸模式的判定係數 R^2 僅有 18.3%，解釋能力很低，表示此模式僅能解釋收縮壓 18.3%的變異百分比

III. 由前面分析得到迴歸式，收縮壓=58.5096+0.8189 腰圍
當腰圍為 100 時，可得該人的預測收縮壓為
 $58.5096+0.8189*100=140.3996$