

112 年公務人員高等考試三級考試試題

類科：衛生行政、漁業技術、養殖技術、食品衛生檢驗

科目：生物統計學

王瑋老師解題

一、某研究人員欲評估有無吸菸與第一秒用力呼氣量 (FEV1) 的相關性，以肺功能量計測量 340 名個案的第一秒用力呼氣量，並依其有無吸菸分成兩組，以比較吸菸組與非吸菸組間的第一秒用力呼氣量平均值的差異，得到下列結果：

組別	人數	第一秒用力呼氣量平均值 (公升)	第一秒用力呼氣量標準差 (公升)
非吸菸組	116	2.92	0.71
吸菸組	224	3.30	0.68

(一)請計算吸菸組之第一秒用力呼氣量平均值的百分之九十五信賴區間。(5分)

(二)請檢定兩組第一秒用力呼氣量之母群體平均值的差異是否有統計顯著性？並列出虛無假說及對立假說，設顯著水準 $\alpha = 0.05$ 。(20分)

1. 《考題難易》：★★☆☆☆

2. 《破題關鍵》本題結合了單母體平均數之信賴區間與雙母體平均差的假設檢定，因為試卷並無合適的 t 表，所以建議皆可採用大樣本 Z 分配的做法即可，111 年高考二級衛技出了類似考題。

3. 出版社書籍命中特區：王瑋，生物統計學，志光出版，頁 175~176；王瑋，生物統計學精選 500 題全解，志光出版，頁 199~200。

【擬答】：

(一)假設 X 代表吸菸組 FEV1， Y 代表非吸菸組 FEV1

吸菸組 FEV1 平均值的 95% 信賴區間為

$$\begin{aligned} \bar{X} \pm Z_{0.025} \frac{s_x}{\sqrt{n_x}} \\ \Rightarrow 3.3 \pm 1.96 \cdot \frac{0.68}{\sqrt{224}} \\ \Rightarrow [3.2109, 3.3891] \end{aligned}$$

(二) $H_0: \mu_x = \mu_y$ $H_1: \mu_x \neq \mu_y$

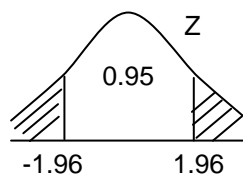
$\alpha = 0.05$

$$Z^* = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}} = \frac{3.3 - 2.92}{\sqrt{\frac{0.68^2}{224} + \frac{0.71^2}{116}}} = 4.75 \in C$$

$C: \{|Z^*| > Z_{0.025} = 1.96\}$

拒絕 H_0 ，有顯著的證據說

兩組第一秒用力呼氣量平均值有差異



公職王歷屆試題 (112 高考三級)

二、某流行病學家欲研究有無肥胖與脂肪肝嚴重程度的關係，自某醫院收集 430 名個案，得到下列結果：

肥胖情況	無脂肪肝組 (n=158)	輕度脂肪肝組 (n=180)	中度脂肪肝組 (n=92)	總和
無	142	145	47	334
有	16	35	45	96
總和	158	180	92	430

(一)請以統計檢定方法檢定有無肥胖與脂肪肝嚴重程度間是否有統計顯著相關？並列出虛無假說及對立假說，設顯著水準 $\alpha = 0.05$ 。(15 分)

(二)請計算中度脂肪肝組與無脂肪肝組兩組間肥胖比率的差異的百分之九十五信賴區間。(10 分)

- 《考題難易》：★★☆☆☆
- 《破題關鍵》本題結合了卡方獨立性檢定與雙母體比例差信賴區間，兩者皆為課內的基本內容，但早年很少結合兩者，而本題型式如同去年 111 年高考二級衛技試題，預期將來會是重要的考點。
- 出版社書籍命中特區：王瑋，生物統計學精選 500 題全解，志光出版，頁 330 完全命中試題。

【擬答】：

(一)計算各格子的期望值 $E_{ij} = \frac{R_i \times C_j}{n}$ ，並將資料整理如下，並且斜線的左上方為觀察值，右下角為

期望值：

肥胖情況	無脂肪肝組 (n=158)	輕度脂肪肝組 (n=180)	中度脂肪肝組 (n=92)	總和
無	142 122.73	145 139.81	47 71.46	334
有	16 35.27	35 40.19	45 20.54	96
總和	158	180	92	430

H_0 有無肥胖與脂肪肝嚴重程度無關

H_1 ：有無肥胖與脂肪肝嚴重程度有關

$\alpha = 0.05$

$df = (2 - 1) \times (3 - 1) = 2$

$C: \{\chi^2 > \chi_{0.05}^2(2) = 5.99\}$

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{(142 - 122.73)^2}{122.73} + \dots + \frac{(45 - 20.54)^2}{20.54} = 51.92 \in C$$

拒絕 H_0 ，有顯著證據說有無肥胖與脂肪肝嚴重程度有關

(二)中度脂肪肝組與無脂肪肝組肥胖比率差異的 95% 信賴區間為

$$(\hat{p}_3 - \hat{p}_1) \pm Z_{0.025} \sqrt{\frac{\hat{p}_3(1 - \hat{p}_3)}{n_3} + \frac{\hat{p}_1(1 - \hat{p}_1)}{n_1}}$$

$$\Rightarrow \left(\frac{45}{92} - \frac{16}{158}\right) \pm 1.96 \sqrt{\frac{\frac{45}{92} \times \frac{47}{92}}{92} + \frac{\frac{16}{158} \times \frac{142}{158}}{158}}$$

$$\Rightarrow [0.2754, 0.5003]$$

三、某社區醫學研究團隊欲探討吸菸情形與高密度脂蛋白膽固醇的相關性，該研究團隊收集 575 名個案，抽血檢測其高密度脂蛋白膽固醇，並依其吸菸情形分成：現在吸菸組、過去吸菸組、從未吸菸組三組，以比較三組間的高密度脂蛋白膽固醇平均值的差異，得到下列結果：

組別	人數	高密度脂蛋白膽固醇平均值 (毫克/分升)	標準差
現在吸菸組	277	44.5	11.1
過去吸菸組	150	46.2	13.6
從未吸菸組	148	50.8	12.9

請檢定三組間的高密度脂蛋白膽固醇平均值的差異是否有統計上顯著不同？並請列出虛無假說及對立假說（設顯著水準 $\alpha = 0.05$ ， $F_{2,100,0.975} = 3.83$ ， $F_{2,1000,0.975} = 3.70$ ）。（25 分）

- 《考題難易》：★★☆☆☆
- 《破題關鍵》單因子變異數分析是課內基本內容，但近年通常以填表為主，較少利用綜合統計量實際計算各種平方和，108 年轉任衛行有類似考題。
- 出版社書籍命中特區：王瑋，生物統計學，志光出版，頁 203~206；王瑋，生物統計學精選 500 題全解，志光出版，頁 227~236。

【擬答】：

假設 1st 組代表現在吸菸組，2nd 組為過去吸菸組，3rd 組是從未吸菸組

$$\text{已知 } n_1 = 277, \bar{X}_{1.} = 44.5, s_1 = 11.1$$

$$n_2 = 150, \bar{X}_{2.} = 46.2, s_2 = 13.6$$

$$n_3 = 148, \bar{X}_{3.} = 50.8, s_3 = 12.9$$

$$SST = \sum \sum (\bar{X}_{i.} - \bar{X}_{..})^2 = \sum n_i (\bar{X}_{i.} - \bar{X}_{..})^2 = 3855.5874$$

$$SSE = \sum \sum (X_{ij} - \bar{X}_{i.})^2 = \sum (n_i - 1) s_i^2$$

$$= (277 - 1) \times 11.1^2 + (150 - 1) \times 13.6^2 + (148 - 1) \times 12.9^2 = 86027.27$$

$$SSTO = SST + SSE = 89882.8574$$

ANOVA 表

	SS	df	MS	F
Treatment	3855.5874	2	1927.7937	12.818
Error	86027.27	572	150.3973	
Total	89882.8574	574		

$$H_0: \mu_1 = \mu_2 = \mu_3 \quad H_1: \mu_i \text{ 不全相等}$$

$$\alpha = 0.05$$

$$C: \{F^* > F_{2,572,0.95} \approx 3.83\}$$

$$F^* = 12.818 \in C$$

拒絕 H_0 ，有顯著的證據說三組高密度脂蛋白膽固醇有差異。

註：本卷未附合適的 F 查表值

四、某研究人員想要評估尿酸 (mg/dL) 與血清肌酸酐 (mg/dL) 間的關係，於某健康管理中心收集 415 名個案的資料，建立以尿酸預測血清肌酸酐的迴歸模式，得到下列結果：

變異數分析表

變異來源	平方和	自由度	均方	F 值	P 值
迴歸模型	1.85	1	1.85	87.19	< 0.001
殘差	8.76	413	0.02		
總和	10.61	414			

模式	迴歸係數	標準誤	標準化迴歸係數	t 值	P 值
尿酸	0.604	0.005	0.418	9.337	< 0.001
常數	0.048	0.036	-	16.728	

請寫出此線性迴歸方程式、解釋此迴歸模式、說明此迴歸模式是否達統計顯著性？並解釋其理由。設顯著水準 $\alpha = 0.05$ 。另請計算此線性迴歸模式的決定係數，並解釋此決定係數的意義。(25 分)

- 《考題難易》：★☆☆☆☆
- 《破題關鍵》課內已強調迴歸分析屬考試基本必出題，本題難度並不高，迴歸係數表與變異數分析表解讀，以及迴歸係數與判定係數的解釋本就是必考題，應可完整拿下分數。近年在 111 年高考衛行與 111 年地特三等衛生皆有類似考題。
- 出版社書籍命中特區：王瑋，生物統計學，志光出版，頁 300~301；王瑋，生物統計學精選 500 題全解，志光出版，頁 379~380 與頁 390~394。

公職王歷屆試題 (112 高考三級)

【擬答】：

(一) 設尿酸為 X_i ，血清肌酸酐為 Y_i

線性迴歸方程式為

$$\hat{Y} = 0.048 + 0.604X$$

代表每增加 1 mg/dL 尿酸，則血清肌酸酐增加 0.604 mg/dL

(二) $H_0: \beta_1 = 0$ $H_1: \beta_1 \neq 0$

$$\alpha = 0.05$$

$T^* = 9.337$ ，對應之 $p\text{-value} < 0.001 < \alpha = 0.05$

拒絕 H_0 ，代表迴歸模式有達到統計上的顯著性

(三) 判定係數 $R^2 = \frac{SSR}{SSTO} \times 100\% = \frac{1.85}{10.61} = 17.44\%$

代表以尿酸來預測血清肌酸酐有 17.44% 解釋度

公
職
王