

113 年公務人員高等考試三級考試試題

類科：統計

科目：抽樣方法與迴歸分析

一、假設某城市共有人口為 $N = 24000$ 人，為了研究此城市的失業人數，於是以簡單隨機取樣自此城市中抽取 500 人，結果發現有 10 人失業。試求：（每小題 5 分，共 25 分）

- (1) 此城市人口失業率的估計值。
- (2) 此城市人口失業率估計量的標準誤的估計值。
- (3) 此城市人口失業率估計量的 95% 近似誤差界限 B 。
- (4) 此城市人口失業率的 95% 近似信賴區間。
- (5) 若我們希望估計此城市人口失業率的誤差界限為 $B = 5\%$ ，則須要在此城市抽取多少人數才能達到我們的需求？（註： $Z_{0.025} = 1.96$ ）。

《考題難易》★☆☆☆☆

《解題關鍵》超級基本的簡單隨機抽樣估計比例值、標準物與樣本數，每隔幾年就有類似的考古題出現，109 年地特與 108 年身障有非常類似的考題。

《命中特區》王瑋，抽樣方法，志光出版，頁 2-51 與頁 2-56~頁 2-59。

解析：

$$\begin{aligned}
 (1) \quad p &= \frac{\sum_{i=1}^{500} y_i}{500} = \frac{10}{500} = 0.02 \xrightarrow{\text{估計}} P \\
 (2) \quad s_p &= \sqrt{(1-f) \frac{s^2}{n}} = \sqrt{(1-f) \frac{\frac{n}{n-1} pq}{n}} \\
 &= \sqrt{(1 - \frac{500}{24000}) \frac{0.02 \times 0.98}{500-1}} = 0.0062 \\
 (3) \quad B &= Z \cdot s_p = 1.96 \times 0.0062 = 0.0122 \\
 (4) \quad &\text{失業率之 95\% 信賴區間為} \\
 &p \pm B \Rightarrow (0.02 \pm 0.0122) \Rightarrow (0.0078, 0.0322) \\
 (5) \quad n_0 &= \frac{Z^2 s_p^2}{B^2} \approx \frac{Z^2 pq}{B^2} = \frac{1.96^2 \times 0.02 \times 0.98}{0.05^2} = 30.12 \approx 31 \\
 n &= \frac{n_0}{1 + \frac{n_0}{N}} = \frac{31}{1 + \frac{31}{24000}} = 30.96, \quad n \text{ 取 } 31
 \end{aligned}$$

二、給定一分層隨機樣本如下：

層 (h)	N_h	n_h	$\bar{y}_h$	s_h
1	200	20	150	30
2	200	40	100	20
3	100	20	200	50

其中 N_h 表示第 h 層的母體大小， n_h 表示第 h 層的樣本大小， $\bar{y}_h$ 表示第 h 層的樣本平均數， s_h 表示第 h 層的樣本標準差。試求：（每小題 5 分，共 25 分）

- (1) 每一層母體平均數的分層估計值。
- (2) 抽樣母體平均數的分層估計值。
- (3) 分層估計量標準誤的估計值。
- (4) 抽樣母體平均數的 95% 近似誤差界限 B 。
- (5) 抽樣母體平均數的 95% 近似信賴區間。（註： $Z_{0.025} = 1.96$ ）

《考題難易》★☆☆☆☆

《解題關鍵》超級基本的分層隨機抽樣估計平均數、標準誤與信賴區間，109 年高考亦曾出現這樣容易的考題。

《命中特區》王瑋，抽樣方法，志光出版，頁 3-9~頁 3-20。

解析：

(1) 第 1 層母體平均數的分層估計值 $\bar{y}_1 = 150$

第 2 層母體平均數的分層估計值 $\bar{y}_2 = 100$

第 3 層母體平均數的分層估計值 $\bar{y}_3 = 200$

$$(2) \bar{y}_{st} = \sum_{h=1}^3 W_h \bar{y}_h = \frac{200}{500} \times 150 + \frac{200}{500} \times 100 + \frac{100}{500} \times 200 = 140$$

$$(3) s_{\bar{y}_{st}} = \frac{1}{N} \sqrt{\sum N_h (N_h - n_h) \frac{s_h^2}{n_h}}$$

$$= \frac{1}{500} \sqrt{200(200-20) \frac{30^2}{20} + 200(200-40) \frac{20^2}{40} + 100(100-20) \frac{50^2}{20}}$$

$$= 3.4293$$

$$(4) B = Z \cdot s_{\bar{y}_{st}} = 1.96 \times 3.4293 = 6.7214。$$

(5) 母體平均數的 95% 近似信賴區間為

$$\bar{y}_{st} \pm B \Rightarrow (140 \pm 6.7214) \Rightarrow (133.2786, 146.7214)$$

志光 學儒 保成

效率考取關鍵

上榜生唯一指定

全國
狀元雙料
金榜

劉○岑

112 高考經建行政狀元
112 普考經建行政榜眼

正規班的每一堂課老師都很用心講解，使我當下都能立即吸收各種新知識。題庫班就是講解各類型的歷屆試題，可以省去找試題及解答的困擾，對於我有幫助。參加奪榜特訓班是想練習如何撰寫申論題，以及我想知道自己的寫作可以拿到幾分，再加上老師會提供完美的擬答供參考。

一年
考取雙料
金榜

池○恆

112 高考會計
112 普考會計

奪榜特訓班藉大量題目練習和反覆複習，讓我可以把自己不擅長和不熟悉的部分，靠題目和老師講解把自己的困難處一一破解。而最大的好處就是可以跟好夥伴一起解題，遇到不會的題目可以一起思考要如何破題，可以讓彼此的實力更上一層樓。

一年
考取雙料
金榜

陳○玟

112 高考統計
112 普考統計

奪榜特訓班有一個很好的環境，讓我能專心讀書，最重要是不會的題目現場有人可以問。每次想休息時，抬頭看到大家都在認真讀書，可以讓我打起精神，覺得我還能再加油。小班導和老師也會跟我們聊天，關心我們的學習情況，每天也幾乎都有考試，成為我學習重要的鞭子。

一年
考取雙料
金榜

趙○育

112 高考經建行政
112 普考經建行政

選擇面授課程的原因是因為我是一個很喜歡問問題的人，如果選擇面授課程，就能夠即時得到老師的回覆，透過詢問的過程中找出自己的盲點。也因此，我很慶幸報名題庫班，因為它的上課方式是老師透過題目講解該題的重點，以及申論題的寫作方向，來將之前上課的內容融會貫通。

三、對一組樣本大小 $n = 100$ 的資料配適一複迴歸模型

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$$

其中 ε_i 為 iid $N(0, \sigma^2)$ 。進行分析後，得到如下結果：

$$\hat{\sigma} = 5, \quad s = \sqrt{\frac{1}{n-1} \sum (Y_i - \bar{Y})^2} = 10$$

(1) 依據前述資訊，完成下列變異數分析表（將此表繪製於試卷上，並寫出詳細計算過程，再將結果填入表中）：（9 分）

變異 來源	平 方 和	自 由 度	均 方	F
迴歸	(a)	(e)	(g)	(i)
誤差	(b)	(d)	(h)	
總和	(c)	(f)		

(2) 若要檢定 $\beta_1, \beta_2, \beta_3$ 是否同時為 0，請列出虛無假設、對立假設、F 值的分配（需標明自由度）、以及在顯著水準 $\alpha = 0.05$ 下拒絕虛無假設的條件。（8 分）

(3) 計算 R^2 並說明其意義。（8 分）

【解題關鍵】

《考題難易》★☆☆☆☆

《破題關鍵》相當容易的多元迴歸分析考題，本題與 89 年高考迴歸試題完全 u 相同，連數字也沒有改，僅需留意 $\hat{\sigma}$ 與 s 的差別即可。

《命中特區》王瑋，迴歸分析，志光出版，頁 5-26~頁 5-27。

解析：

(1) (d) $4 - 1 = 3$

(e) $100 - 4 = 96$

(f) $100 - 1 = 99$

(h) $MSE = \hat{\sigma}^2 = 25$

(c) $SSTO = SS_Y = (100 - 1) \times 10^2 = 9900$

(b) $SSE = (n - p)MSE = (100 - 4) \times 25 = 2400$

(a) $SSR = SSTO - SSE = 9900 - 2400 = 7500$

(g) $MSR = 7500 \div 3 = 2500$

(i) $F^* = 2500 \div 25 = 100$

變異數分析表整理如下

變異 來源	平 方 和	自 由 度	均 方	F
迴歸	7500	3	2500	100
誤差	2400	96	25	
總和	9900	99		

(2) $H_0: \beta_1 = \beta_2 = \beta_3 = 0$ $H_1: \beta_i$ 不全為 0, $i = 1, 2, 3$

 $F \sim F(3, 96)$ ，分子與分母自由度分別為 3 與 96拒絕域 $C: \{F^* > F_{0.05}(3, 96)\}$ 若 $F^* = 100$ 超過臨界值 $F_{0.05}(3, 96)$ 則拒絕 H_0 ，有顯著證據說 β_i 不全為 0

(3) $R^2 = \frac{SSR}{SSTO} = \frac{7500}{9900} \times 100\% = 75.76\%$

代表以 X_1, X_2, X_3 共同解釋應變數 Y ，具有 75.76% 解釋力。



公職超強班

面授 + 視訊 + 函授

開啟上榜三效模式

★12期分期0利率

★面授 / 視訊 / 雲端函授 自由選

★優惠最低85折 (持金卡&尊榮優惠可再享折扣)

★提供 正規班+總複習 CP值最高

自選面授or視訊 or雲端函授課程		返班選擇適合學習模式	
第一年	超強 ▼ 第一年考取退學費	方案一 ▶ 到班學習	方案二 ▶ 雲端學習
	扣除第一年學費& 第二年已使用教材費	升級 面授or視訊考取班 安心專注 一次繳費輔導至考取 隔年起 僅繳交教材換證費	函授 年度正規班 便利自主 輔考至該年度考試前 享有申論批改與 超級解惑王APP上榜資源
第二年			

四、考慮配適一簡單線性迴歸模型： $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ ， $i = 1, 2, \dots, n$ ，並假設 ε_i 為 iid $N(0, \sigma^2)$ 。

(1)請寫出模型中的應變數與自變數。(5分)

(2)請問 iid 是那三個英文字縮寫，是代表什麼假設？請詳細說明。(5分)

(3)若以最小平方估計式 $\hat{\beta}_0$ 及 $\hat{\beta}_1$ 得到 $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ ，且令 $e_i = Y_i - \hat{Y}_i$ 。請問 $\sum_{i=1}^n X_i e_i = ?$ 請詳細列出推導過程。(15分)

【解題關鍵】

《考題難易》★☆☆☆☆

《破題關鍵》本題(1)(2)是直接送分題，過去高普考皆不會有這麼基本的試題；題(3)則是利用最小平方方法即可輕鬆獲得推導結果。

《命中特區》王瑋，迴歸分析，志光出版，頁 2-3 與頁 2-25。

解析：

(1)應變數為 Y_i ；自變數是 X_i 。

(2) iid 代表 Independent and Identical Distribution，即自母體隨機抽出一組樣本，假設他們會獨立且同分配。

(3)令 $Q = \sum \varepsilon_i^2 = \sum (Y_i - \beta_0 - \beta_1 X_i)^2$

$$\frac{\partial Q}{\partial \beta_0} = -2 \sum (Y_i - \beta_0 - \beta_1 X_i) = 0 \Rightarrow \sum (Y_i - \beta_0 - \beta_1 X_i) = 0 \cdots \text{①}$$

$$\frac{\partial Q}{\partial \beta_1} = -2 \sum (Y_i - \beta_0 - \beta_1 X_i) X_i = 0 \Rightarrow \sum (Y_i - \beta_0 - \beta_1 X_i) X_i = 0 \cdots \text{②}$$

聯立解出 $\hat{\beta}_1 = \frac{SS_{XY}}{SS_X}$ ， $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$ ，並帶入②

可得 $\sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) X_i = \sum e_i X_i = 0$