

《統計學概要》

一、為開發某新型流行病的檢測工具，研發單位分別招募經醫師診斷確認已得病與未受感染的自願受測者。檢測結果顯示，已得病者的陽性率為90%；未受感染者的陰性率為95%。過去的資料顯示這個新型疾病的盛行率（整體得病的比率）大約15%。根據這些檢測工具的資訊，請回答下列問題：

(一) 整體的陽性率為何？(8分)

(二) 檢測為陽性的民眾其得病的機率為何？檢測為陰性的民眾其未受感染的機率為何？(12分)

試題評析 本題涉及貝氏定理之計算題，難度不高，惟有理解題意就能夠獲得滿分。

考點命中 《高點·高上統計學講義》第一回，趙治勳編撰，第四章第四節。

答：

令A表實際上有得病

B表檢測結果為陽性

$$(一) P(B) = (0.15)(0.9) + (0.85)(0.05) = 0.1775$$

$$(二) P(A|B) = \frac{0.15 \times 0.9}{0.1775} = 0.7606$$

$$(三) P(A^c|B^c) = \frac{0.85 \times 0.95}{1 - 0.1775} = 0.9818$$

二、咖啡豆批發商販售母豆（或稱平豆，為雙核果咖啡豆）、公豆（或稱圓豆，為單核果咖啡豆）以及混合豆（公、母豆以特定比例混合）等三類型咖啡豆。某連鎖咖啡店欲購買一批標示為30%混合咖啡豆（公豆顆數約30%），透過重量推估，這批貨約N顆咖啡豆，為確保品質（公豆比例超過30%），採購經理先隨機取出100顆混合豆樣本，逐一檢視並記錄其中公豆的個數，記為x，作為下單決策的參考。假設真正的混合比例為p，請回答下列問題：

(一) 請以隨機變數的概念，寫出公豆個數X的分配，以及其對應的常態近似分配。(10分)

(二) 經理將透過統計檢定決定是否採購，並希望該決策的錯誤率控制在5%內。請寫出適當的檢定假說、檢定統計量的計算公式、臨界值，以及棄卻域，在x=38下，完成檢定假說並建議經理，是否採購。(15分)

試題評析 本題涉及假設檢定之計算題，但(二)中設立 H_0 H_1 時要小心，題意與樣本結果並不相符，這就是課堂上說的「用賤招會有1%的錯誤率」。

考點命中 《高點·高上統計學講義》第三回，趙治勳編撰，第十一章。

答：

$$(一) X \sim \text{Bin}(n = 100, p) \approx N(100p, 100p(1-p))$$

$$(二) H_0: p \geq 0.3 \text{ vs } H_1: p < 0.3$$

$$\text{T.S.: } Z = \frac{\hat{p} - 0.3}{\sqrt{\frac{0.3 \times 0.7}{100}}} \underset{\text{by C.L.T.}}{\sim} N(0, 1)$$

$$\text{R.R.: Reject } H_0 \text{ at } \alpha = 0.05 \text{ if } Z^* < z_{0.05} = -1.645$$

$$\therefore Z^* = \frac{\frac{38}{100} - 0.3}{\sqrt{\frac{0.3 \times 0.7}{100}}} = 1.7457 \quad \therefore \text{don't reject } H_0$$

結論：我們沒有足夠證據去推論咖啡豆不符合品質，故經理應該要向該咖啡豆供應商採購。

三、抽樣調查的結果顯示，10月份15家以個人旅遊為主的旅行社的月平均營收為2,930萬元，標準差為485萬元；10家以企業旅遊為主的旅行社的月平均營收為3,250萬元，標準差為245萬元。假設月營收資料服從常態，請定義適當的變數及參數，並根據上述資訊，回答下列問題：

(一)請寫出兩類型的旅行社營收的變異數是否相同的假說檢定、統計量的計算公式、棄卻域，以及臨界值，並根據上述資料，在5%的顯著水準下，完成檢定提出檢定結果。(12分)

(二)請根據(一)的結果寫出兩類型旅行社月平均營收差是否相同的假說檢定、統計量的計算公式、棄卻域，以及臨界值，並算出檢定統計量及檢定p-值。(18分)

試題評析	本題涉及兩獨立母體平均數與變異數的假說檢定之計算題，考古題中已經出現過很多次，大部份考生應該能夠獲得滿分。
考點命中	《高點·高上統計學講義》第三回，趙治勳編撰，第十一章。

答：

令 X_1, X_2 分別表以個人旅遊為主與以企業旅遊為主旅行社之月營收（萬元）

假設 (1) $X_1 \perp X_2$ (2) 隨機樣本

母體： $X_1 \sim N(\mu_1, \sigma_1^2) \perp X_2 \sim N(\mu_2, \sigma_2^2)$

樣本： $X_{11}, \dots, X_{115} \stackrel{iid}{\sim} N(\mu_1, \sigma_1^2) \perp X_{21}, \dots, X_{210} \stackrel{iid}{\sim} N(\mu_2, \sigma_2^2)$

點： $\bar{X}_1 \sim N(\mu_1, \frac{\sigma_1^2}{15}) \quad \bar{X}_2 \sim N(\mu_2, \frac{\sigma_2^2}{10})$

$$S_1^2 = \frac{\sum (X_{1i} - \bar{X}_1)^2}{15-1} \quad S_2^2 = \frac{\sum (X_{2i} - \bar{X}_2)^2}{10-1}$$

(一) $H_0: \sigma_1^2 = \sigma_2^2$ vs $H_1: \sigma_1^2 \neq \sigma_2^2$

$$\text{T.S.: } F = \frac{S_1^2}{S_2^2} \sim F_{(14,9)}$$

R.R.: Reject H_0 at $\alpha = 0.1$ if

$$F^* < F_{0.95(14,9)} = \frac{1}{F_{0.05(9,14)}} = \frac{1}{2.65} = 0.377 \quad \text{或} \quad F^* > F_{0.05(14,9)} = 3.03$$

【版權所有，重製必究！】
[考卷只給 $F_{0.05(v_1, v_2)}$ 表，把 $\alpha = 0.05$ 改為 $\alpha = 0.1$]

$$\therefore F^* = \frac{485^2}{245^2} = 3.9188 \quad \therefore \text{reject } H_0$$

結論：我們有足夠證據去推論兩母體變異數不相同。

(二) $H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$

$$T.S.: T = \frac{(\bar{X}_1 - \bar{X}_2) - (0)}{\sqrt{\frac{S_1^2}{15} + \frac{S_2^2}{10}}} \sim t_{(v=22)} \quad \text{其中 } v = \frac{(\frac{S_1^2}{15} + \frac{S_2^2}{10})^2}{\frac{(\frac{S_1^2}{15})^2}{15-1} + \frac{(\frac{S_2^2}{10})^2}{10-1}} = 21.8 \approx 22$$

R.R.: Reject H_0 at $\alpha = 0.05$ if $|T^*| > t_{0.025(22)} = 2.0739$

$$\therefore |T^*| = 2.1731$$

$$p\text{-value} = 2P(T_{22} > 2.1731) = 2 \times 0.02157 = 0.04314$$

其中 $P(T_{22} > 2.1731) = a$

$$\frac{0.025 - a}{2.1731 - 2.0739} = \frac{a - 0.01}{2.5083 - 2.1731} \Rightarrow a = 0.02157$$

四、美國某貨運公司經理想要了解黑色星期五熱銷期間，運送距離（英里）與送達所需要的天數之間的關係。隨機挑選30筆批發運送業務的資料，其敘述性統計如下：平均運送距離為725.85、標準差為87.45；平均送達所需的天數為8.4、標準差為2.7；運送距離與送達天數的相關係數為0.69。請根據上述資訊，回答下列問題：

(一) 請算出運送距離與送達天數的共變異數。(6分)

(二) 請以適當符號，寫出以運送距離預測送達所需天數的迴歸預測模式，並以上述資料估計迴歸模式的係數。(13分)

(三) 請用(二)所算出的係數，寫出以運送距離預測送達所需天數的迴歸預測模式。假設有一批貨的運送距離是755英里，請運用上述迴歸預測模式，預測這批貨的送達天數。(6分)

試題評析 本題涉及簡迴歸的計算題，考古題中已經出現過很多次，大部份考生應該能夠獲得滿分。

考點命中 《迴歸分析熱門題庫》，高點文化出版，趙治勳編著，第二章。

答：

令 X 表運送距離（英里）

Y 表運送天數

$$(一) \text{ 共變異數 } S_{XY} = r_{XY} S_X S_Y = 0.69 \times 87.45 \times 2.7 = 162.919$$

$$(二) \text{ 模型: } Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \varepsilon_i \stackrel{iid}{\sim} N(0, \sigma^2), i = 1, 2, \dots, 30$$

$$\hat{\beta}_1 = r_{XY} \frac{S_Y}{S_X} = 0.69 \times \frac{2.7}{87.45} = 0.0213 \quad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = -7.061$$

$$(三) \hat{y} = -7.061 + 0.0213x$$

當運送距離 $x = 755$ （英里）時，這批貨之平均送達天數估計值為

$$\hat{y} = -7.061 + 0.0213(755) = 9.0205 \quad (\text{天})$$