

《統計學概要》

試題評析

今年考題著重於計算，考題難度不高，但因為題目量多，非常考驗考生的解題速度。以普考來說，本卷難度中下，基本分應該要有90分。

第一題：範圍是二項分配，只要有看懂題目，加上對於二項分配隨機實驗之認知，考生就可以獲得滿分。

第二題：範圍是平均數之信賴區間及假設檢定，只要必要假設不忘補上，應該不難獲得滿分。

第三題：範圍是母體成功比例之假設檢定，只要必要假設不忘補上，應該不難獲得滿分。

一、愛美斯公司每星期需作報紙廣告費之決策，但事證顯示：每週花\$250或\$500報紙廣告費之機率分別為0.4與0.6。若欲研究每月4週花\$500報紙廣告費之週數，則請利用相關性質解答下列問題：

- (一)請定義必要之隨機變數 (The Required Random Variable)。(2分)
- (二)請列衍生的樣本空間 (The Induced Sample Space)。(3分)
- (三)試問花\$500報紙廣告費週數之機率函數 (Probability Function)，並敘明其名稱。(5分)
- (四)試求\$500報紙廣告費週數之期望值 (Expected Value) 與變異數 (Variance)。(5分) 若欲研究每月報紙廣告費之支出總額，則
- (五)請定義必要之隨機變數 (The Required Random Variable)。(2分)
- (六)請列衍生的樣本空間 (The Induced Sample Space)。(3分)
- (七)請定義月報紙廣告費支出總額與花\$500報紙廣告費週數之關係式。(3分)
- (八)試求每月報紙廣告費支出總額之機率函數 (Probability Function)，並敘明其名稱。(7分)
- (九)試求每月報紙廣告費支出總額之期望值 (Expected Value) 與變異數 (Variance)。(10分)

答：

(一)令 X 表每月花\$500報紙廣告費之週數

(二) $x = 0, 1, 2, 3, 4$

(三) $X \sim Bin(n = 4, p = 0.6)$

$$f_X(x) = \binom{4}{x} 0.6^x 0.4^{4-x}, x = 0, 1, 2, 3, 4$$

(四) $E(X) = (4)(0.6) = 2.4$ ， $V(X) = (4)(0.6)(0.4) = 0.96$

(五)令 Y 表每月報紙廣告費之支出總額

(六) $y = 1000, 1250, 1500, 1750, 2000$

(七) $Y = 500X + 250(4 - X) = 250(4 + X)$

(八) $Y = 250(4 + X) \Rightarrow X = \frac{Y}{250} - 4$

$$f_Y(y) = f_X\left(x = \frac{y}{250} - 4\right) = \binom{4}{\frac{y}{250} - 4} 0.6^{\frac{y}{250} - 4} 0.4^{4 - \frac{y}{250} + 4}$$

$$= \left(\frac{4}{\frac{y}{250} - 4} \right) 0.6^{\frac{y}{250} - 4} 0.4^{8 - \frac{y}{250}}, y = 1000, 1250, 1500, 1750, 2000$$

$$(九) E(Y) = E[250(4 + X)] = 250(4 + E(X)) = 250(4 + 2.4) = 1600$$

$$V(Y) = V[250(4 + X)] = 250^2 V(4 + X) = 250^2 V(X) = 250^2 (0.96) = 60000$$

【高分命中】趙治勳老師《統計學上課講義第一回》，5.2節。

二、茲為研究愛美詩地區夫妻身高之關係，遂於該區隨機抽出100對夫妻之身高（以公分為測量單位）表列如下：

夫 妻	165-170	170-175	175-180	180-185	185-190	190-195	195-200	200-205	205-210
190-195								1	2
185-190						1	2	3	1
180-185				2	3	5	3	2	
175-180			1	2	6	5	2		
170-175			1	6	8	4	1	2	
165-170		1	3	4	5	1	1		
160-165	1	2	4	4	2				
155-160		2	3	1		1			
150-155	1	1							

若夫妻身高之母體 (Population) 不知是否為常態 (Normal)，則請問

(一) 應如何處理？(2分)

若相關母體為常態，且顯著水準為5%，則

(二) 請列出夫妻平均身高差之估計式。(1分)

(三) 請計算夫妻平均身高差之估計值。(2分)

(四) 請列出夫妻平均身高差之信賴區間估計式。(1分)

(五) 請計算夫妻平均身高差之信賴區間估計值。(4分)

(六) 請問夫妻平均身高是否不同？(5分)

(七) 請問夫平均身高是否高於妻平均身高5公分但不高於6.3公分？(10分)

(八) 試問夫妻身高之相關係數是否超過0.8？(5分)

答：

令 X_1, X_2 分別表夫與妻之身高 (公分)

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

樣本: $(X_{11}, X_{21}), (X_{12}, X_{22}), \dots, (X_{1100}, X_{2100}) \stackrel{iid}{\sim} (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

設 $D_i = X_i - Y_i$

(一) 利用中央極限定理 $\bar{D} \stackrel{by C.L.T.}{\sim} N(\mu_D, \sigma_D^2)$

其中 $\mu_D = \mu_1 - \mu_2$, $\sigma_D^2 = \sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2\rho$

(二) $\bar{D} = \bar{X}_1 - \bar{X}_2$

(三) 以各組之組中點表示組內之觀察值，如：(190-195)以192.5代表之，可見本題計算量超大，考生要

小心計算

$$\bar{D} = \bar{X} - \bar{Y} = 187.6 - 172.95 = 14.65$$

(四) 母體: $D \sim N(\mu_D, \sigma_D^2)$

樣本: $D_1, D_2, \dots, D_{100} \stackrel{iid}{\sim} N(\mu_D, \sigma_D^2)$

點估計: $\bar{D} \sim N(\mu_D, \frac{\sigma_D^2}{100})$

樞紐量: $T = \frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{100}}} \sim t_{(99)}$

機率區間: $P(-t_{0.025(99)} \leq \frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{100}}} \leq t_{0.025(99)}) = 0.95$

信賴區間: $P(\bar{D} - t_{0.025(99)} \frac{S_D}{\sqrt{100}} \leq \mu_D \leq \bar{D} + t_{0.025(99)} \frac{S_D}{\sqrt{100}}) = 0.95$

結論: $\mu_1 - \mu_2$ 之 95% 信賴區間估計式為

$$(\bar{D} - t_{0.025(99)} \frac{S_D}{\sqrt{100}}, \bar{D} + t_{0.025(99)} \frac{S_D}{\sqrt{100}})$$

(五) $(14.65 - 1.96 \frac{6.6}{\sqrt{100}}, 14.65 + 1.96 \frac{6.6}{\sqrt{100}}) = (13.3564, 15.9436)$

(六) $H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2 \Rightarrow H_0: \mu_D = 0$ vs $H_1: \mu_D \neq 0$

T.S.: $T = \frac{\bar{D} - (0)}{\frac{S_D}{\sqrt{100}}} \sim t_{(99)}$

R.R.: Reject H_0 at α if $T^* > t_{0.025(99)} = 1.96$

$$\therefore T^* = \frac{14.65 - (0)}{\sqrt{\frac{6.6^2}{100}}} = 22.197 \quad \therefore \text{reject } H_0$$

我們有足夠證據去推論夫妻平均身高是有差異的

(七) $H_0: \mu_1 - \mu_2 \leq 5$ vs $H_1: \mu_1 - \mu_2 > 5 \Rightarrow H_0: \mu_D \leq 5$ vs $H_1: \mu_D > 5$

T.S.: $T = \frac{\bar{D} - (5)}{\frac{S_D}{\sqrt{100}}} \sim t_{(99)}$

R.R.: Reject H_0 at α if $T^* > t_{0.025(99)} = 1.96$

$$\therefore T^* = \frac{14.65 - (5)}{\sqrt{\frac{6.6^2}{100}}} = 14.6212 \quad \therefore \text{reject } H_0$$

我們有足夠證據去推論夫妻平均身高差大於5公分

$$H_0 : \mu_1 - \mu_2 \geq 6.3 \text{ vs } H_1 : \mu_1 - \mu_2 < 6.3 \Rightarrow H_0 : \mu_D \geq 6.3 \text{ vs } H_1 : \mu_D < 6.3$$

$$\text{T.S.: } T = \frac{\bar{D} - (6.3)}{\frac{S_D}{\sqrt{100}}} \sim t_{(99)}$$

$$\text{R.R.: Reject } H_0 \text{ at } \alpha \text{ if } T^* > t_{0.025(99)} = 1.96$$

$$\therefore T^* = \frac{14.65 - (6.3)}{\sqrt{\frac{6.6^2}{100}}} = 12.6515 \quad \therefore \text{reject } H_0$$

我們有足夠證據去推論夫之平均身高高於妻之平均身高5公分

$$\text{(八) } H_0 : \rho \leq 0.8 \text{ vs } H_1 : \rho > 0.8$$

$$\text{T.S.: } Z = \frac{\frac{1}{2} \ln\left(\frac{1+r}{1-r}\right) - \frac{1}{2} \ln\left(\frac{1+0.8}{1-0.8}\right)}{\sqrt{\frac{1}{100-3}}} \sim N(0,1)$$

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

$$\therefore Z^* = \frac{\frac{1}{2} \ln\left(\frac{1+0.745}{1-0.745}\right) - \frac{1}{2} \ln\left(\frac{1+0.8}{1-0.8}\right)}{\sqrt{\frac{1}{100-3}}} = -1.3492 \quad \therefore \text{don't reject } H_0$$

我們沒有足夠證據去推論夫妻身高相關係數超過0.8

【高分命中】

1. 趙治勳老師《統計學上課講義第二回》，頁69。
2. 《迴歸分析熱門題庫》，高點出版，趙治勳編著，頁2-25。

三、若為瞭解某區適婚男女未婚之詳情，遂於該區之適婚男女各隨機抽出300人為樣本，發現適婚男性未婚者81人而適婚女性未婚者99人，則

- (一)請列適婚男女未婚比率之估計式 (Estimator)。(1分)
- (二)請分別計算適婚男女未婚比率之估計值 (Estimates)。(2分)

若相關母體 (Population) 不知是否為常態 (Normal)，則請問

- (三)應如何處理？(2分)

若相關母體為常態，且顯著水準為 (The Level of Significance) 5%，則

- (四)請列適婚男女未婚比率差之估計式。(1分)
- (五)請分別計算適婚男女未婚比率差之估計值。(2分)
- (六)請問適婚女性未婚之比率是否低於0.4？(5分)
- (七)請問適婚男性未婚之比率是否低於0.3？(5分)
- (八)請問適婚男女未婚之比率是否相同？(5分)
- (九)請問適婚女性未婚之比率是否比適婚男性未婚之比率高5%？(5分)
- (十)請簡要說明你的發現與理由。(2分)

答：

令 X_1, X_2 分別表適婚男性未婚與適婚女性未婚，假設 $X_1 \perp X_2$

母體： $X_1 \sim \text{Ber}(p_1) \perp X_2 \sim \text{Ber}(p_2)$

樣本： $X_{11}, X_{12}, \dots, X_{1300} \stackrel{iid}{\sim} \text{Ber}(p_1), X_{21}, X_{22}, \dots, X_{2300} \stackrel{iid}{\sim} \text{Ber}(p_2)$

$$(一) \hat{p}_1 = \frac{\sum_{i=1}^{300} X_{1i}}{300}, \hat{p}_2 = \frac{\sum_{i=1}^{300} X_{2i}}{300}$$

$$(二) \hat{p}_1 = \frac{\sum_{i=1}^{300} X_{1i}}{300} = \frac{81}{300} = 0.27, \hat{p}_2 = \frac{\sum_{i=1}^{300} X_{2i}}{300} = \frac{99}{300} = 0.33$$

(三) 利用中央極限定理

$$\hat{p}_1 \underset{\text{by C.L.T.}}{\sim} N\left(p_1, \frac{p_1(1-p_1)}{300}\right), \hat{p}_2 \underset{\text{by C.L.T.}}{\sim} N\left(p_2, \frac{p_2(1-p_2)}{300}\right)$$

$$(四) \hat{p}_1 - \hat{p}_2 = \frac{\sum_{i=1}^{300} X_{1i}}{300} - \frac{\sum_{i=1}^{300} X_{2i}}{300}$$

$$(五) \hat{p}_1 - \hat{p}_2 = \frac{\sum_{i=1}^{300} X_{1i}}{300} - \frac{\sum_{i=1}^{300} X_{2i}}{300} = \frac{81}{300} - \frac{99}{300} = -0.06$$

(六) $H_0: p_2 \geq 0.4$ vs $H_1: p_2 < 0.4$

$$\text{T.S.: } Z = \frac{\hat{p}_2 - 0.4}{\sqrt{\frac{0.4(1-0.4)}{300}}} \underset{\text{by C.L.T.}}{\sim} N(0,1)$$

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

$$\therefore Z^* = \frac{0.33 - 0.4}{\sqrt{\frac{0.4(1-0.4)}{300}}} = -2.475 \quad \therefore \text{reject } H_0$$

我們有足夠證據去推論適婚女性未婚之比率低於0.4

(七) $H_0: p_1 \geq 0.3$ vs $H_1: p_1 < 0.3$

$$\text{T.S.: } Z = \frac{\hat{p}_1 - 0.3}{\sqrt{\frac{0.3(1-0.3)}{300}}} \underset{\text{by C.L.T.}}{\sim} N(0,1)$$

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

$$\therefore Z^* = \frac{0.27 - 0.3}{\sqrt{\frac{0.3(1-0.3)}{300}}} = -1.134 \quad \therefore \text{don't reject } H_0$$

我們沒有足夠證據去推論適婚男性未婚之比率低於0.3

(八) $H_0: p_1 = p_2$ vs $H_1: p_1 \neq p_2$

$$\text{T.S.: } Z = \frac{\hat{p}_1 - \hat{p}_2 - (0)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{300} + \frac{1}{300}\right)}} \underset{\text{by C.L.T.}}{\sim} N(0,1) \quad \text{其中 } \hat{p} = \frac{81+99}{300+300} = 0.3$$

R.R.: Reject H_0 at $\alpha = 0.05$ if $|Z^*| > z_{0.025} = 1.96$

$$\therefore |Z^*| = \left| \frac{(-0.06) - (0)}{\sqrt{0.3(1-0.3)\left(\frac{1}{300} + \frac{1}{300}\right)}} \right| = 1.604 \quad \therefore \text{don't reject } H_0$$

我們沒有足夠證據去推論適婚男女未婚之比率不相同

(九) $H_0: p_1 - p_2 \geq -0.05$ vs $H_1: p_1 - p_2 < -0.05$

$$\text{T.S.: } Z = \frac{\hat{p}_1 - \hat{p}_2 - (-0.05)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{300} + \frac{\hat{p}_2(1-\hat{p}_2)}{300}}} \underset{\text{by C.L.T.}}{\sim} N(0,1)$$

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

$$\therefore Z^* = \frac{(-0.06) - (-0.05)}{\sqrt{\frac{0.27(1-0.27)}{300} + \frac{0.33(1-0.33)}{300}}} = -0.268 \quad \therefore \text{don't reject } H_0$$

我們沒有足夠證據去推論適婚女性未婚之比率是比適婚男性未婚之比率高5%

(十) 由(八)可推論適婚男女未婚之比率相同，但由(九)可推論女性未婚之比率是比適婚男性未婚之比率低5%，看似兩個小題結論不一致，其實是一致的，由於(九) $p_1 - p_2 \geq -0.05$ 已經包含了(八) $p_1 - p_2 = 0$ 了，因此，我們可以透過以上檢定去推論適婚男女未婚之比率相同的。

【高分命中】趙治勳老師《統計學上課講義第二回》，頁76-78。