

《統計學概要》

試題評析

今年考題著重在觀念解析，並沒有複雜的計算，相信上過總複習班的同學都應該覺得很熟悉，例如，第一題在P9《例題5》，第二題在P22《例題1》經驗法則與P23《例題3》單邊柴比雪夫不等式，第三題在P2整理完備，第四題paired-t-test對照無母數sign test與Wilcoxon test在P41《例題1》，第五題Normal求機率與C.L.T.分別在P7《例題2》與P26《例題1》均為雷同。

一、欲瞭解食品近一年價格波動情況，而至市場實地訪查，得下列兩類食品資料：

食品	平均數	中位數	標準差	最小值	最大值	樣本數
雞蛋(X)	29.16	30.00	2.72	24.00	33.00	12
牛肉(Y)	109.33	110.00	6.63	98.00	120.00	12

$\sum_{i=1}^{12} X_i Y_i = 38314$

(一)求兩者價格的全距和變異係數 (coefficient of variation) (8分)

(二)何者價格波動大？兩者價格波動的相關程度 (相關係數) 為何？ (12分)

答：

(一)(1)雞蛋 (X)：

$$\text{全距} = 33 - 24 = 9$$

$$C.V. = \frac{2.72}{29.16} \times 100\% \approx 9.33\%$$

(2)牛肉 (Y)：

$$\text{全距} = 120 - 98 = 22$$

$$C.V. = \frac{6.63}{109.33} \times 100\% \approx 6.06\%$$

(二)(1) $\because C.V.(X) > C.V.(Y)$

\therefore 雞蛋波動大

$$r = \frac{S_{XY}}{\sqrt{S_{XX}} \sqrt{S_{YY}}} = \frac{\sum XY - n\bar{X}\bar{Y}}{\sqrt{(n-1)S_X^2} \sqrt{(n-1)S_Y^2}}$$

$$(2) = \frac{38314 - 12 \times 29.16 \times 109.33}{\sqrt{11 \times 2.72^2} \sqrt{11 \times 6.63^2}} \approx 0.2886$$

二、某單位主管欲採購每棵至少高三公尺高的行道樹150棵。甲林場負責人報告目前樹木生長概況：林場有600棵樹，平均高度是2.5公尺，標準差是0.5公尺，樹木高度的分布接近對稱。乙林場負責人報告其林場有500棵樹，平均高度是2.6公尺，標準差是0.2公尺，樹木高度的分布並不對稱。

(一)甲林場大約有幾棵樹符合條件？(提示：柴比雪夫不等式或經驗法則) (8分)

(二)乙林場大約有幾棵樹符合條件？(提示：柴比雪夫不等式或經驗法則) (8分)

(三)該主管應如何採購？(4分)

答：

(一)經驗法則 $P(|X - \mu| \leq \sigma) \doteq 0.68$ (對稱)

$$P(X \geq 3) = P(X - \mu \geq 3 - 2.5)$$

$$= P(X - \mu \geq \sigma) \doteq 0.16 \text{ (經驗法則)}$$

$$600 \times 0.16 = 96$$

∴ 甲林場大約有96棵符合條件

(二)柴比雪夫單邊不等式：

$$P(X - \mu \geq k) \leq \frac{\sigma^2}{k^2 + \sigma^2}, \quad k > 0$$

$$P(X \geq 3) = P(X - \mu \geq 3 - 2.6) = P(X - \mu \geq 0.4)$$

$$\leq \frac{0.2^2}{0.4^2 + 0.2^2} \text{ (柴比雪夫單邊不等式)}$$

$$= 0.2$$

$$500 \times 0.2 = 100$$

∴ 乙林場不超過100棵符合條件

(三)甲、乙兩林場估出符合條件的機率接近

∴ 採購上差異不大

三、進行統計檢定時，檢定統計量可能會發生誤判：

(一)請問這些誤判是那些？(5分)

(二)請問進行統計檢定時，那些誤判發生之機率可以被控制？為什麼？(5分)

答：

(一) $P(\text{type I error}) = P(\text{reject } H_0 | H_0 \text{ is true})$

$P(\text{type II error}) = P(\text{accept } H_0 | H_1 \text{ is true})$

(二)型I錯誤控制：控制最大型I錯誤機率不得超過 α 。

型II錯誤控制：在型I錯誤控制下，尋找一個最佳危險域使得power最大，

即 $\beta = 1 - \text{power}$ 最小。

四、一位研究人員欲瞭解兩種肥料對韓國草的生長影響。他隨機選擇10所學校區域並將每個區域劃分為二塊，分別施灑兩種肥料進行觀察，紀錄韓國草的高度如下：

肥料	1	2	3	4	5	6	7	8	9	10
A	3.1	2.6	2.9	3.5	3.4	2.8	3.2	2.6	2.8	2.4
B	3.0	2.4	2.1	3.4	3.0	2.7	3.3	2.7	2.6	2.3

(一)請問此種蒐集資料方式的優點或缺點？(10分)

(二)請協助此位研究人員進行單尾統計檢定和推論，顯著水準 $\alpha = 0.05$ (10分)

答：

(一)採成對抽樣：

優點：在相同的區域內種植（在相同的條件下試驗），比較不同的肥料對韓國草生長影響較為客觀，所

以檢定較為精確。

缺點：因為 (X_i, Y_i) 內部相關，所以兩種肥料可能相互影響生長情形。

(二) 採sign test：

A	3.1	2.6	2.9	3.5	3.4	2.8	3.2	2.6	2.8	2.4
B	3.0	2.4	2.1	3.4	3.0	2.7	3.3	2.7	2.6	2.3
A - B 符號	+	+	+	+	+	+	-	-	+	+

$$n^+ = 8, n^- = 2, n' = 10$$

$$S \sim \text{Bin}(n' = 10, p = \frac{1}{2})$$

$$\textcircled{1} \begin{cases} H_0: \eta_A = \eta_B \\ H_1: \eta_A > \eta_B \quad (\because n^+ > n^-) \end{cases}$$

$$\textcircled{2} \text{p-value} = P(S \leq \min(n^+, n^-) | H_0)$$

$$= \sum_{s=0}^2 C_S^{10} \left(\frac{1}{2}\right)^{10} = 0.0546 > \alpha = 0.05$$

\therefore Not reject H_0 ，兩種肥料對韓國草成長影響無顯著差異。

五、若監理站申請汽車新牌照辦理時間為一平均數是6分鐘和標準差是1.5分鐘的常態分布

(一) 請問一份牌照申請需花7分鐘以上才能取得的可能性？(5分)

(二) 若有1000份牌照申請，請問至少有270份牌照辦理需花7分鐘以上才能取得的可能性(寫出真正機率之算式，再求近似機率即可)？(15分)

(三) 單位主管簡化牌照辦理手續，並實施辦理數月。單位主管欲知簡化後平均辦理時間是否小於6分鐘，採用系統抽樣，得25位牌照辦理時間之平均數為5.2分鐘，標準差是2分鐘。請問在顯著水準 $\alpha = 0.05$ 下，牌照辦理時間是否縮短了？(10分)

答：辦理時間 $X \sim N(6, 1.5^2)$

$$(一) P(X > 7) = P(Z > \frac{7-6}{1.5}) = P(Z > 0.6666) \doteq 0.2525$$

(二) 令 Y ：1000份申請中，需花7分鐘以上的份數

$$Y \sim \text{Bin}(1000, p = 0.2525)$$

$$EY = 1000 \times 0.2525 = 252.5$$

$$V(Y) = 1000 \times 0.2525 \times (1 - 0.2525) = 188.7438$$

$$P(Y \geq 270) = \sum_{y=270}^{1000} C_y^{1000} (0.2525)^y (1 - 0.2525)^{1000-y}$$

(真正機率算式)

$$\doteq P(Z \geq \frac{270 - 252.5 - 0.5}{\sqrt{188.7438}})$$

(中央極限定理)

$$= P(Z \geq 1.237) \doteq 0.108$$

$$(三) \textcircled{1} \begin{cases} H_0: \mu \geq 6 \\ H_1: \mu < 6 \quad (\because \bar{X} = 5.2 < 6) \end{cases}$$

$$\textcircled{2} C = \{T \mid T < -t_{0.05}(25-1) = -1.71\}$$

$$\textcircled{3} T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{5.2 - 6}{2/\sqrt{25}} = -2 \in C$$

\therefore reject H_0 ，有充分證據顯示牌照辦理時間縮短了。