10/2統計學實習課

114-1統計學實習課

Chapter 3 review

114-1統計學實習課

題目 1：位置與離散量

某公司 10 位新進員工的起薪（單位：千元）如下：

`38, 42, 45, 45, 47, 48, 50, 50, 52, 60`

(1) 計算平均數。

(2) 計算中位數。

(3) 找出眾數。

題目1解答：

- 平均數 $= \frac{38+42+45+45+47+48+50+50+52+60}{10} = 47.7$ 千元
- 中位數 = 第 5 與第 6 個值平均 = $(47+48)/2 = 47.5$ 千元
- 眾數 = 45、50（兩個眾數）

題目 2：百分位數與四分位數

某 8 位學生的統計學分數如下：

```
55, 60, 68, 70, 72, 75, 78, 90
```

請計算：

(1) 第 25 百分位（Q1）

(2) 第 75 百分位（Q3）

(3) 第 80 百分位（P80）

題目2解答：

公式：位置 $L_p = \frac{p}{100}(n+1)$，其中 $n = 8$。

(1) Q1（25 百分位）：

$L_{25} = 0.25 \times (8+1) = 2.25$。

第 2 個值是 60，第 3 個值是 68。

在第 2 和第 3 個之間 0.25 的距離：

$Q1 = 60 + 0.25 \times (68 - 60) = 62$。

(2) Q3（75 百分位）：

$L_{75} = 0.75 \times 9 = 6.75$。

第 6 個值是 75，第 7 個值是 78。

在第 6 和第 7 個之間 0.75 的距離：

$Q3 = 75 + 0.75 \times (78 - 75) = 77.25$。

(3) P80（80 百分位）：

$L_{80} = 0.80 \times 9 = 7.2$。

第 7 個值是 78，第 8 個值是 90。

在第 7 和第 8 個之間 0.2 的距離：

$P_{80} = 78 + 0.2 \times (90 - 78) = 80.4$。

題目 3：變異數與標準差

某小組五名員工的專案完成天數如下：

```
12, 15, 11, 18, 14
```

請計算：

(1) 平均數

(2) 樣本變異數

(3) 樣本標準差

題目3解答：

- 平均數 $= 14$ 天
- 偏差平方和 $= (12-14)^2 + (15-14)^2 + (11-14)^2 + (18-14)^2 + (14-14)^2 = 4 + 1 + 9 + 16 + 0 = 30$
- 樣本變異數 $s^2 = \frac{30}{5-1} = 7.5$
- 樣本標準差 $s = \sqrt{7.5} \approx 2.74$

題目 4：z 分數與離群值偵測

研究某班研究生的第一份工作月薪（單位：元）。已知這組資料的樣本平均數 $\bar{x} = 7200$，樣本標準差 $s \approx 180$。其中最大值為 7800，最小值為 6900。已知 Q1 = 7050、Q3 = 7350。

1. 計算最大值 7800 的 z-score 與最小值 6900 的 z-score。
2. 用 z-score ±3 規則判斷是否為離群值。
3. 用 IQR 法判斷哪個值為離群值。

題目4解答：

(1) z-score 計算

公式：$z = \frac{x-\bar{x}}{s}$

最大值 7800 的 z-score：

$$z = \frac{7800-7200}{180} = \frac{600}{180} \approx 3.33$$

最小值 6900 的 z-score：

$$z = \frac{6900-7200}{180} = \frac{-300}{180} \approx -1.67$$

(2) ±3 規則判斷

z-score ±3 規則：若 z 值小於 -3 或大於 3，即為離群值。

觀察：

最大值 z ≈ 3.33 → 超過 3 → **離群值**

最小值 z ≈ -1.67 → 在範圍內 → 不是離群值

(3) IQR 法判斷

計算 IQR：

$$IQR = Q3 - Q1 = 7350 - 7050 = 300$$

下限：

$$Q1 - 1.5 \cdot IQR = 7050 - 450 = 6600$$

上限：

$$Q3 + 1.5 \cdot IQR = 7350 + 450 = 7800$$

判斷：

最大值 7800 = 上限 → 不算離群值

最小值 6900 > 下限 → 不算離群值

結論：

- z-score ±3 規則：最大值 7800 為離群值
- IQR 法：無離群值
- 本例中兩種方法判斷結果 **不一致**，可用來討論不同離群值判定標準的差異。

題目 5：共變異數與相關係數

某咖啡店記錄了 10 週週末促銷活動數量 x 與下週平均客流量 y（單位：人數）如下：

x（週末促銷活動數）：1, 3, 2, 4, 3, 2, 5, 3, 4, 2

y（下週客流量，人）：45, 52, 48, 55, 53, 50, 60, 51, 57, 49

請計算：

1. 樣本平均 $\bar{x}, \bar{y}$
2. 樣本共變異數 s_{xy}（以 $n-1$ 計算）
3. x 與 y 的樣本標準差 s_x, s_y
4. 皮爾森相關係數 r_{xy}

結果保留到小數第 3 或 4 位。

題目5解答：

1. 計算樣本平均

$$\bar{x} = \frac{1+3+2+4+3+2+5+3+4+2}{10} = \frac{29}{10} = 2.9$$

$$\bar{y} = \frac{45+52+48+55+53+50+60+51+57+49}{10} = \frac{520}{10} = 52.0$$

2. 計算樣本共變異數

公式：$s_{xy} = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{n-1}$

逐項計算 $(x_i - \bar{x})(y_i - \bar{y})$ 的總和為 47.0

$$s_{xy} = \frac{47.0}{10-1} = \frac{47.0}{9} \approx 5.222$$

3. 計算樣本標準差

$$\sum(x_i - \bar{x})^2 = 13.9$$

$$s_x^2 = \frac{13.9}{9} \approx 1.544$$

$$s_x = \sqrt{1.544} \approx 1.243$$

$$\sum(y_i - \bar{y})^2 = 165$$

$$s_y^2 = \frac{165}{9} \approx 18.333$$

$$s_y = \sqrt{18.333} \approx 4.282$$

4. 計算皮爾森相關係數

$$r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{5.222}{1.243 \times 4.282} \approx \frac{5.222}{5.321} \approx 0.981$$

Chapter 4 review

114-1統計學實習課

題目 1：樣本空間與計數

某公司要從三位候選人 (A, B, C) 中選出 **董事長** 和 **副董事長**（不重複）。

(1) 寫出所有可能的選法樣本空間。

(2) 一共有多少種可能結果？

(3) 若假設每一種選法等可能，選到 A 當董事長的機率是多少？

題目1解答：

(1) 樣本空間（列出所有可能）

(A, B), (A, C), (B, A), (B, C), (C, A), (C, B)

(2) 種數

排列數 $P^3_2 = \dfrac{3!}{(3-2)!} = \dfrac{6}{1} = 6$ 種。

(3) 選到 A 當董事長的機率

「A 當董事長」的結果為 (A,B) 與 (A,C)，共 2 種。

機率 $= \dfrac{2}{6} = \dfrac{1}{3} \approx 0.3333$。

題目 2：加法法則

某家店的顧客調查顯示：

- 20% 的顧客買了咖啡
- 35% 的顧客買了點心
- 10% 的顧客同時買了咖啡和點心

請問：

(1) 隨機選一位顧客，他至少買了一項產品的機率是多少？

(2) 買咖啡與買點心的事件是否互斥？

題目2解答：

- $P(C \cup D) = P(C) + P(D) - P(C \cap D) = 0.20 + 0.35 - 0.10 = 0.45$
- 因為 $P(C \cap D) = 0.10 \neq 0$，所以兩事件 **不是互斥**。

題目 3：條件機率

某公司 500 名員工中：

- 男性 300 人，其中 90 人是主管
- 女性 200 人，其中 30 人是主管

請計算：

(1) 隨機選一人是主管的機率。

(2) 若選到男性，該員工是主管的條件機率。

(3) 若選到女性，該員工是主管的條件機率。

題目3解答：

- 總主管人數 = 90+30=120 → $P(\text{主管}) = 120/500 = 0.24$
- 男性主管率 = $90/300 = 0.30$
- 女性主管率 = $30/200 = 0.15$

題目 4：獨立事件判斷

投擲一顆公平骰子兩次，定義事件：

- A：第一次投擲結果是偶數
- B：兩次點數和大於 8

請問事件 A 與 B 是否獨立？（需計算 $P(A)$、$P(B)$、$P(A \cap B)$，並比較 $P(A \cap B)$ 與 $P(A)P(B)$）

題目4解答：

(1) 計算 $P(A)$

第一次為偶數的機率：有 3 個偶數 / 6 = $P(A) = \frac{3}{6} = \frac{1}{2} = 0.5$。

(2) 計算 $P(B)$（列出和 > 8 的組合）

和為 9 的組合：(3,6),(4,5),(5,4),(6,3) → 4 個

和為 10 的組合：(4,6),(5,5),(6,4) → 3 個

和為 11 的組合：(5,6),(6,5) → 2 個

和為 12 的組合：(6,6) → 1 個

總數 = $4 + 3 + 2 + 1 = 10$。

$P(B) = \frac{10}{36} = \frac{5}{18} \approx 0.277777$。

(3) 計算 $P(A \cap B)$：同時「第一次為偶數」且「和>8」

把第一個點數分別代入看第二個點數允不允許：

- 若第一 = 2：需要第二 > 6（因為 2 + x > 8 → x >6）→ 不可能 → 0 組合
- 若第一 = 4：需要第二 > 4 → 第二可為 5 或 6 → 2 組合：(4,5),(4,6)
- 若第一 = 6：需要第二 > 2 → 第二可為 3,4,5,6 → 4 組合：(6,3),(6,4),(6,5),(6,6)

交集總數 = $2 + 4 = 6$ 組。

$P(A \cap B) = \frac{6}{36} = \frac{1}{6} \approx 0.1666667$。

(4) 比較 $P(A \cap B)$ 與 $P(A)P(B)$：

$P(A)P(B) = \frac{1}{2} \cdot \frac{5}{18} = \frac{5}{36} \approx 0.1388889$。

但 $P(A \cap B) = \frac{6}{36} = \frac{1}{6} = 0.1666667 \neq \frac{5}{36}$。

結論：**A 與 B 並非獨立事件**。

題目 5：貝氏定理應用

某疾病的檢測有以下數據：

- 有病者 1%
- 有病時檢測陽性率（靈敏度）= 0.95
- 無病時檢測陽性率（假陽性）= 0.05

(1) 隨機一人檢測陽性的機率是多少？

(2) 若某人檢測陽性，他真的有病的機率是多少？

題目5解答：

- $P(+) = 0.95(0.01) + 0.05(0.99) = 0.0095 + 0.0495 = 0.059$
- $P(D \mid +) = \frac{0.95(0.01)}{0.059} = \frac{0.0095}{0.059} \approx 0.161$

  → 約 16.1%