

《統計學概要》

一、一汽車生產工廠將某日生產的309汽車，依A、B、C、D四種缺點及大、中、小三種車型，分別檢驗並記錄其資料如下表：

車型\缺點	A	B	C	D
大	15	21	45	13
中	26	31	34	5
小	33	17	49	20

在 $\alpha=0.05$ 下，試以卡方檢定汽車缺點與車型大小是否有關（請詳細寫出a.虛無與對立假設、b.檢定統計量公式、c.拒絕域、d.檢定值之計算過程、e.檢定結果與結論）？（25分）

試題評析 卡方獨立性檢定，注意計算不要出錯並把題目要求寫清楚，題目本身並不困難。

答：

車型\缺點	A	B	C	D	Total
大	15 (22.5)	21 (21.0)	45 (38.9)	13 (11.6)	94
中	26 (23.0)	31 (21.4)	34 (39.8)	5 (11.8)	96
小	33 (28.5)	17 (26.6)	49 (49.3)	20 (14.6)	119
Total	74	69	128	38	309

H_0 : 汽車缺點與車型大小無關 against H_1 : 汽車缺點與車型大小有關

$$\varphi = \sum_{i=1}^3 \sum_{j=1}^4 \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi^2(6)$$

$$\alpha = 0.05, RR = \{\varphi | \varphi > \chi_{0.05}^2(6) = 12.592\}$$

$$\varphi^* = \sum_{i=1}^3 \sum_{j=1}^4 \frac{(O_{ij} - e_{ij})^2}{e_{ij}} = 19.26 \in RR$$

\therefore 拒絕虛無假設

我們有足夠證據推論汽車缺點與車型大小有關

二、已知函數 $f_X(x)$ 為

x	0	1	3
$f_X(x)$	1/3	1/2	1/6

試求：（每小題5分，共25分）

(一) 驗證 $f_X(x)$ 為一機率函數。

(二) $P(X \leq 2.5 | X > 1)$ 。【版權所有，重製必究！】

(三) $Var(X)$ 。

(四) $E[(X+1)^4]$ 。

(五) $Var[(X+1)^2]$ 。

試題評析 期望值與變異數的基本題，清楚定義即可拿分。

答：

(一)由於 $f_X(x) > 0, \forall x \in \{1, 2, 3\}$ 且 $\sum_x f_X(x) = \frac{1}{3} + \frac{1}{2} + \frac{1}{6} = 1$
 $f_X(x)$ 為一機率函數

$$(二) P(X \leq 2.5 | X > 1) = \frac{P(1 < X \leq 2.5)}{P(X > 1)} = \frac{0}{\frac{1}{6}} = 0$$

$$(三) E(X) = 0 \times \frac{1}{3} + 1 \times \frac{1}{2} + 3 \times \frac{1}{6} = 1$$

$$E(X^2) = 0^2 \times \frac{1}{3} + 1^2 \times \frac{1}{2} + 3^2 \times \frac{1}{6} = 2$$

$$Var(X) = E(X^2) - [E(X)]^2 = 2 - 1^2 = 1$$

$$(四) E[(X+1)^4] = (0+1)^4 \times \frac{1}{3} + (1+1)^4 \times \frac{1}{2} + (3+1)^4 \times \frac{1}{6} = 51$$

$$(五) Var[(X+1)^2] = E[(X+1)^4] - (E[(X+1)^2])^2 = 51 - [E(X^2 + 2X + 1)]^2 = 51 - 5^2 = 26$$

三、在簡單直線迴歸中，設自變數 X 為廣告費用，應變數 Y 為銷售量，樣本數 $n=10$ 。已知：

$$\sum_{i=1}^n x_i = 140, \quad \sum_{i=1}^n y_i = 1300, \quad \sum_{i=1}^n x_i y_i = 21040,$$

$$\sum_{i=1}^n x_i^2 = 2528, \quad \sum_{i=1}^n y_i^2 = 184730$$

(一)用最小平方法求迴歸直線。(10分)

(二)求斜率 β_1 之95%信賴區間。(10分)

(三)求判定係數 R^2 。(5分)

試題評析

本題為迴歸分析之範圍，計算上稍微複雜需特別注意；區間計算公式較為繁瑣，容易出錯務必要當心。

答：

$$(一) \hat{\beta}_1 = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} = \frac{10 \times 21040 - 140 \times 1300}{10 \times 2528 - 140^2} = 5$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = 130 - 5 \times 14 = 60$$

因此迴歸直線為 $\hat{Y} = 60 + 5X$

$$(二) SST = \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - \frac{(\sum_{i=1}^n y_i)^2}{n} = 184730 - \frac{1300^2}{10} = 15730$$

$$SSR = \hat{\beta}_1^2 \sum_{i=1}^n (x_i - \bar{x})^2 = \hat{\beta}_1^2 \left[\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n} \right] = 25 \times \left(2528 - \frac{140^2}{10} \right) = 14200$$

$$SSE = SST - SSR = 15730 - 14200 = 1530$$

$$\hat{\sigma}^2 = \frac{SSE}{n-2} = \frac{1530}{10-2} = 191.25$$

$$SE(\hat{\beta}_1) = \sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} = \sqrt{\frac{191.25}{568}} = 0.5803, \text{ 重製必究!}$$

β_1 之 95% 信賴區間為

$$[\hat{\beta}_1 \mp t_{0.025}(8) \times SE(\hat{\beta}_1)] = [5 \mp 2.306 \times 0.5803] = [3.6618, 6.3382]$$

$$(三) R^2 = \frac{SSR}{SST} = \frac{14200}{15730} = 0.9027$$

四、為了測定工廠安全訓練的效率，下列資料是12家工廠於安全訓練前之年平均意外事件數（ X ）和安全訓練後之年平均意外事件數（ Y ）：

X	50	87	37	141	59	65	24	88	25	36	50	35
Y	41	75	35	129	60	53	26	85	29	31	48	37

假設資料的差異具有常態分配。令 $d=X-Y$ ，請以顯著水準 $\alpha=0.05$ ，進行單尾檢定，以檢定安全訓練的效果（請詳細寫出a.虛無與對立假設、b.檢定統計量公式、c.拒絕域、d.檢定值之計算過程、e.檢定結果與結論）。（25分）

試題評析 本題考常見的假設檢定。看清楚題目要求，將步驟完整寫出便可得分。

答：

X	50	87	37	141	59	65	24	88	25	36	50	35
Y	41	75	35	129	60	53	26	85	29	31	48	37
D	9	12	2	12	-1	12	-2	3	-4	5	2	-2

$$\bar{d} = \frac{\sum_{i=1}^{12} d_i}{12} = 4, s_D = \sqrt{\frac{\sum_{i=1}^{12} d_i^2 - 12 \times \bar{d}^2}{11}} = \sqrt{\frac{580 - 12 \times 4^2}{11}} = \sqrt{\frac{388}{11}}$$

$$H_0: \mu_D \leq 0 \text{ against } H_1: \mu_D > 0$$

$$\varphi = \frac{\bar{D} - 0}{s_D/\sqrt{12}} \sim t(11)$$

$$\alpha = 0.05, RR = \{\varphi | \varphi > t_{0.05}(11) = 1.796\}$$

$$\varphi^* = \frac{\bar{d} - 0}{s_D/\sqrt{12}} = 2.333 \in RR$$

\therefore 拒絕虛無假設

我們有足夠證據推論安全訓練有效果

【版權所有，重製必究！】