《統計學概要》

試題評析	108地方特考經建行政組別統計學題目的五大題，全部都是基本題型，第一題為離散型的隨機變數計算；第二題是超幾何與二項分配的差別；第三、四題則是假設檢定直來直往的考法，沒有變化；最後一大題簡單線性迴歸的內容也是基本例題，這應該是近幾年來國家考試考得最簡單的一次，所以請同學一定要詳細檢查是否有計算錯誤。

註：I.對應右尾機率值a的標準常態分配臨界值 z_a：

$z_{0.05}=1.645$；$z_{0.025}=1.96$；$z_{0.35}=0.385$；$z_{0.55}=-0.126$

II.對應自由度df且右尾機率值a的t分配臨界值 $t_a(df)$：

$t_{0.025}(3)=3.182$；$t_{0.05}(3)=2.353$；$t_{0.025}(4)=2.776$；$t_{0.05}(4)=2.132$；

$t_{0.025}(7)=2.365$；$t_{0.05}(7)=1.895$；$t_{0.025}(8)=2.306$；$t_{0.05}(8)=1.860$

III.對應自由度df且累積機率值a的卡方分配臨界值 $x_a^2(df)$：

$x_{0.025}^2(7)=1.69$；$x_{0.05}^2(7)=2.167$；$x_{0.025}^2(8)=2.18$；$x_{0.05}^2(8)=2.733$；

$x_{0.975}^2(7)=16.013$；$x_{0.95}^2(7)=14.067$；$x_{0.975}^2(8)=17.535$；$x_{0.95}^2(8)=15.507$

所有假設檢定問題，皆需正確寫出虛無假設、對立假設、檢定統計量、拒絕域、檢定結果與結論。

一、假設隨機變數X之機率分配如下：（每小題5分，共20分）

x	1	2	3	4
f(x)	0.35	0.25	0.15	0.25

(一)試求X的期望值。

(二)試求X的標準差。

(三)試求機率P(X>2)。

(四)令 $Y=X^2+2X+1$，則Y的期望值為何？

考點命中	《高點‧高上統計學講義》第二回，蘇建郎編撰，頁11，ch3-2 動差與不等式。

答：

(一) $E(X)=\sum_{x=1}^{4} x\cdot f_X(x)=1\times0.35+2\times0.25+3\times0.15+4\times0.25=2.3$

$E(X^2)=\sum_{x=1}^{4} x^2\cdot f_X(x)=1^2\times0.35+2^2\times0.25+3^2\times0.15+4^2\times0.25=6.7$

(二) $Var(X)=E(X^2)-(EX)^2=6.7-2.3^2=1.41$

$\sigma_X=\sqrt{1.41}=1.874$

(三) $P(X>2)=P(X=3)+P(X=4)=0.15+0.25=0.4$

(四) $E(Y)=E(X^2+2X+1)=E(X^2)+2E(X)+1)=6.7+2\times2.3+1=12.7$

二、已知10個零件中有2個瑕疵，若任取3個來檢驗，求：（每小題5分，共15分）

(一)若採不歸還抽樣，則3個零件中有1個瑕疵之機率為多少？

(二)承題(一)，3個零件中沒有瑕疵之機率為多少？

(三)若採歸還抽樣，則3個零件中至少有1個瑕疵之機率為多少？

考點命中	《高點‧高上統計學講義》第二回，蘇建郎編撰，頁58，ch4-1 一維常用分配離散型。

答：

(一) 由題意得知X代表抽出不放回之瑕疵數，即$X \sim Hyp(N=10, n=3, D=2)$

$$f_X(x)=\frac{C_x^2 C_{3-x}^8}{C_3^{10}}\ ,\ x=0,1,2 \qquad P(X=1)=\frac{C_1^2 C_2^8}{C_3^{10}}=\frac{7}{15}$$

(二) $P(X=0)=\frac{C_0^2 C_3^8}{C_3^{10}}=\frac{7}{15}$

(三) 由題意得知X代表抽出放回之瑕疵數，即$X \sim B(n=3, p=0.2)$

$$f_X(x)=C_x^3 0.2^x (0.8)^{3-x}\ ,\ x=0,1,2,3$$

至少有一個瑕疵數的機率為$P(X\ge 1)=1-P(X=0)=1-0.8^3=0.488$

三、下列是青少年對於所聽音樂類型喜好的意見調查：

音樂類型	調查的青少年數	喜愛該類型的青少年數
流行樂	400	204
饒舌樂	400	250

(一)請根據資料估計青少年喜歡流行樂的比例和喜歡饒舌樂的比例的差異。（5分）

(二)在0.05顯著水準下，是否可推論青少年喜歡流行樂的比例和喜歡饒舌樂的比例有差異？（10分）

考點命中	《高點・高上統計學講義》第三回，蘇建郎編撰，頁58，ch7-2 檢定統計量法決策準則。

答：

（一）p_1為青少年喜歡流行樂的比例；p_2青少年喜歡饒舌樂的比例

比例差異估計為$\hat{p}_1-\hat{p}_2=\frac{204}{400}-\frac{250}{400}=-0.115$

（二）

① $\begin{cases} H_0: p_1 = p_2 \\ H_1: p_1 \ne p_2 \end{cases}$

$\alpha=0.05$，$\hat{p}_1=0.51$，$\hat{p}_2=0.625$，$\hat{p}=\frac{x_1+x_2}{n_1+n_2}=\frac{454}{400+400}=0.5675$

② $$z^*=\frac{(\hat{p}_1-\hat{p}_2)-0}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1}+\frac{1}{n_2})}}=\frac{0.51-0.625}{\sqrt{0.5675\times 0.4235(\frac{1}{400}+\frac{1}{400})}}=\frac{-0.115}{0.0347}=-3.31$$

③ $\alpha=0.05$，$C=\{z^*:|z^*|>z_{\alpha/2}=z_{0.025}=1.96\}$

$z^*=-3.31\in C$，$reject\ H_0$

④ 故在顯著水準$\alpha=0.05$的情況下，根據樣本資料顯示，我們有證據說$p_1\ne p_2$，即青少年喜歡流行樂的比例與喜歡饒舌樂的的比例有差異。

四、某公司產出手提袋的耐重強度呈常態分配，經試驗8次得強度分別如下：（單位：公斤）（每小題10分，共20分）

6.0、5.9、5.8、6.5、6.6、6.9、5.9、6.3

(一)在0.05的顯著水準下，檢定母體平均強度是否超過6公斤。

(二)在0.05的顯著水準下，檢定母體變異數是否為0.3。

考點命中	《高點・高上統計學講義》第三回，蘇建郎編撰，頁66~68，ch7-2 檢定統計量法決策準則。

答：

(一)

① $\begin{cases} H_0: \mu \le 6 \\ H_1: \mu > 6 \end{cases}$ $\alpha = 0.05$，$n = 8$，$\bar{x} = 6.2375$，$s^2 = 0.1598$

② $t^* = \dfrac{\bar{x} - \mu_0}{\sqrt{s^2/n}} = \dfrac{6.2375 - 6}{\sqrt{0.1598/8}} = 1.68$

$\alpha = 0.05$，$n = 8$，$\bar{x} = 6.2375$，$s^2 = 0.1598$

③ $\alpha = 0.05$，$C = \{t^* \mid t^* > t_{\alpha/2}(n-1) = t_{0.05}(7) = 1.895\}$

$t^* = 1.68 \notin C$，$don't\ reject\ H_0$

④ 故在顯著水準$\alpha = 0.05$的情況下，根據樣本資料顯示，我們沒有證據說$\mu > 6$。

(二)

① $\begin{cases} H_0: \sigma^2 = 0.3 \\ H_1: \sigma^2 \ne 0.3 \end{cases}$，$\alpha = 0.05$，$n = 8$，$s^2 = 0.1598$

② $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma_0^2} = \dfrac{7 \times 0.1598}{0.3} = 3.7292$

③ $$\begin{aligned} \alpha = 0.05，C &= \{\chi_0^2 \mid \chi_0^2 < \chi_{0.025}^2(7)\ or\ \chi_0^2 > \chi_{0.975}^2(7)\} \\ &= \{\chi_0^2 \mid \chi_0^2 < 1.69\ or\ \chi_0^2 > 16.013\} \end{aligned}$$

$\chi_0^2 = 3.7292 \notin C$,$don't\ reject\ H_0$

④ 故在顯著水準$\alpha = 0.05$的情況下，根據樣本資料顯示，我們沒有證據說$\sigma^2 \ne 0.3$，即變異數為0.3

五、市場研究員為探討廠商投入的廣告費（X）對銷售額（Y）之影響，乃建立迴歸模型：$Y = \beta_0 + \beta_1 X + \varepsilon$。今隨機抽取5家廠商，得其廣告費與銷售額的關係如表所述：（單位：百萬元）（每小題10分，共30分）

	廠商1	廠商2	廠商3	廠商4	廠商5
廣告費X	12	15	8	10	8
銷售額Y	7	12	4	6	5

(一)試求廣告費與銷售額之相關係數。

(二)試用最小平方估計法（least squares method）求出Y對X的直線迴歸係數的估計值b_0和b_1。

(三)在0.05的顯著水準下檢定此迴歸線是否顯著。

考點命中	《高點・高上統計學講義》第五回，蘇建郎編撰，頁13、18，ch9-2 簡單線性迴歸與相關分析。

答：

(一)

$n = 5$，$\sum x_i = 44$，$\sum x_i^2 = 498$，$\sum y_i = 34$，$\sum y_i^2 = 270$，$\sum x_i y_i = 342$

$\bar{x} = 8.8$，$\bar{y} = 6.8$ ，$SS_{xx} = \sum (x_i - \bar{x})^2 = \sum x_i^2 - n\bar{x}^2 = 110.8$

$$SS_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - n\bar{x}\bar{y} = 42.8$$

$$SS_{yy}=\sum(y_i-\overline{y})^2=\sum y_i^2-n\overline{y}^2=38.8$$

廣告與銷售額之樣本相關係數 $r_{x,y}=\dfrac{SS_{xy}}{\sqrt{SS_{xx}SS_{yy}}}=\dfrac{42.8}{\sqrt{110.8\times 38.8}}=0.6528$

(二)

係數估計值 $b_1=\dfrac{SS_{xy}}{SS_{xx}}=\dfrac{42.8}{110.8}=0.3863$，$b_0=\overline{y}-\hat{\beta}_1\overline{x}=3.4007$

(三)

① $\begin{cases}H_0:\beta_1=0\\H_1:\beta_1\neq 0\end{cases}$ $\alpha=0.05$

$SSE=SST-\hat{\beta}_1^2SS_{xx}=38.8-16.5329=26.2671$

$MSE=\dfrac{SSE}{n-2}=7.4224$

② $t^*=\dfrac{b_1-\beta_1}{\sqrt{\dfrac{MSE}{SS_{xx}}}}=\dfrac{0.3863-0}{\sqrt{\dfrac{7.4224}{110.8}}}=1.49$

③ $\alpha=0.05$，$C=\{t^*:|t^*|>t_{\alpha/2}(n-2)=t_{0.025}(3)=3.182\}$

$t^*=1.49\notin C$，$don't\ reject\ H_0$

④ 故在顯著水準 $\alpha=0.05$ 的情況下，根據樣本資料顯示，我們沒有證據說 $\beta_1\neq 0$，即此迴歸線不顯著。