統　計

# 《統計學》

**試題評析**

**今年的試題命題範疇平均，難易也適中，沒有特別艱難的題目。**

**但由於今年統計學禁止使用計算機，嚴重考驗考生的計算能力（尤其是長位數及開平方根的計算）；再加上有關變異數分析的試題，同學也可能因不熟而錯漏一些分數。**

**整體來說，諸如獨立性、不偏性、配對樣本t檢定、迴歸分析等部分考題皆為基本題型，考生只要計算小心，拿70分以上不成問題，程度好的考生，拿滿分也不意外。**

**一、假設$X_1,X_2,X_3$為一組取自二項分布$B(3,\frac{1}{3})$的隨機樣本（即獨立同分布）。計算$E(2X_1|X_3)$，$E(X_1^2+X_2X_3)$。（二十分）**

**答：**

$\because X_i \overset{iid}{\sim} B(3,\frac{1}{3})$，$E(X_i)=3\cdot\frac{1}{3}=1$，$Var(X_i)=3\cdot\frac{1}{3}\cdot\frac{2}{3}=\frac{2}{3}$

$\therefore E(2X_1|X_3)=2E(X_1|X_3)=2E(X_1)=2\cdot 1=2$ ($X_1,X_3$獨立)

$$E(X_1^2+X_2X_3)=E(X_1^2)+E(X_2)+E(X_3)=Var(X_1)+[E(X_1)]^2+E(X_2)E(X_3)$$

$$=\frac{2}{3}+1^2+1\cdot 1=\frac{8}{3}$$

**二、設$X_1, X_2, ..,.X_n$，$n>1$，為來自普阿松母體Poisson($\lambda$)的隨機樣本。**

**(一)試舉出兩個$\lambda$的不偏估計量。（十分）**

**(二)若不偏估計量之形式為$\sum_{i=1}^{n} a_i X_i$，試找出$a_i$，$i=1, ..,.n$，使此估計量之變異數為最小。（十分）**

**答：**

(一)$\because X_i \overset{iid}{\sim} Poisson(\lambda)\ \ i=1,\ldots,n, E(X_i)=\lambda$，$Var(X_i)=\lambda$

$$\therefore E(\overline{X})=E(\frac{\sum_{i=1}^{n}X_i}{n})=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\sum_{i=1}^{n}\lambda=\frac{1}{n}\cdot n\lambda=\lambda$$

$$E(S^2)=E\left[\frac{\sum(X_i-\overline{X})^2}{N-1}\right]=\frac{1}{N-1}\sum_{i}E(X_i-\overline{X})^2=\frac{1}{N-1}\sum_{i=1}^{n}Var(X_i-\overline{x})$$

$$=\frac{1}{n-1}\left\{\sum_{i=1}^{n}Var[(1-\frac{1}{n})X_i-\frac{1}{n}\sum_{j\neq i}X_j]\right\}$$

$$=\frac{1}{n-1}\left\{\sum_{i=1}^{n}[(1-\frac{1}{n})^2Var(X_i)+\frac{1}{n^2}\sum_{j\neq i}Var(X_j)]\right\}$$

$$=\frac{1}{n-1}\left\{\sum_{i=1}^{n}[(1-\frac{1}{n})^2\lambda+\frac{(n-1)}{n^2}\lambda]\right\}$$

$$=\frac{1}{n-1}\cdot n\cdot\left\{\frac{(n-1)^2+(n-1)}{n^2}\right\}\lambda$$

$$=\frac{1}{n-1}\cdot n\cdot\frac{n(n-1)}{n^2}\lambda$$

$$=\lambda$$

**故**$\overline{x}$，$S^2$**均為**$\lambda$**之不偏估計量**

(二)**令**$Y=\sum_{i=1}^{n}a_i x_i$

1.$E(Y)=E(\sum a_i x_i)=\sum a_i E(X_i)=\sum a_i\lambda=\lambda$

$\Rightarrow \sum_{i=1}^{n} a_i = 1$

2. $Var(Y)= Var(\sum a_i x_i) = \sum a_i^2 Var(X_i) = \sum a_i^2 \lambda$

**由柯西不等式知**

$(\sum a_i^2)(\sum 1^2) \geq (\sum a_i)^2$

$\Rightarrow (\sum a_i^2) \cdot n \geq 1^2 = 1$

$\Rightarrow Var(Y) \geq \frac{\lambda}{n}$

**等號於** $\frac{1}{a_1} = \frac{1}{a_2} = \frac{1}{a_3} = \ldots = \frac{1}{an}$**時成立，亦即是**

$a_1=a_2=\cdots=a_n=\frac{1}{n} (\because \sum a_i = 1)$**時，此估計量變異數為最小**

**三、某瘦身中心30位會員瘦身前X及瘦身後Y體重之樣本統計資料如下：**

**瘦身前樣本平均體重為55.00；瘦身後樣本平均體重為53.46**

**瘦身前及瘦身後體重之樣本共變異數矩陣為**

| | X | Y |
|---|---|---|
| X | 7.01227 | 4.50309 |
| Y | 4.50309 | 12.41188 |

**(一)令D=X–Y，請算出D之樣本變異數。（十分）**

**(二)請寫出檢定假設，以檢定瘦身中心瘦身是否有效？（二分）**

**(三)請檢定瘦身中心瘦身是否有效？（顯著水準5%）（八分）**

（$t_{28,0.05}$=1.701，$t_{29,0.05}$=1.699，$t_{30,0.05}$=1.697）

**答：**

(一) Var(D)=Var(X–Y)=Var(X)+Var(Y)–2Cov(X,Y) =7.01227+12.4118-2(4.50309)=10.41797

(二) 1. $H_0: \mu_X - \mu_Y \leq 0$

2. $H_1: \mu_X - \mu_Y > 0$

**其中**$\mu_X$，$\mu_Y$**各為瘦身前及瘦身後母體體重平均數**

(三) 1. $\partial = 0.05$

2. $C = \{t \mid t > t_{29,\,0.05} = 1.699\}$

3. $t = \frac{D}{\sqrt{\frac{1}{n}Var(D)}} = \frac{\overline{X}-\overline{Y}}{\sqrt{\frac{1}{n}Var(D)}} = \frac{55-53.46}{\sqrt{\frac{1}{30}(10.41797)}}$ $\quad 2.61 \in C$

4. **在顯著水準**$\alpha = 0.05$**下，我們拒絕虛無假設**$H_0$**，亦即瘦身中心瘦身顯著有效。**

**四、考慮簡單線性迴歸模型：**$y_i = \alpha + \beta x_i + \varepsilon_i, i = 1,.,.,n$**，其中**$\varepsilon_i \sim N(0,\sigma^2)$**且相互獨立。假定就一組n=7的樣本，已知**

$\bar{x} = 4, \bar{y} = 9, \sum x_i^2 = 140, \sum y_i^2 = 667, \sum x_i y_i = 304$

**(一)試求y對x的預測迴歸式。**$\hat{y} = a + bx$ **（六分）**

**(二)試求x,y的樣本相關係數r及判定係數（coefficient of determination）**$R^2$**，並說明二者的關係。（八分）**

**(三)請檢定x,y間的相關係數是否等於0？（令**$\alpha = 0.05$**）（六分）**

（$t_{6,0.05}$=1.943，$t_{5,0.05}$=2.015，$t_{6,0.025}$=2.447，$t_{5,0.025}$=2.571）

**答：**

(一) 1. $SS_x = \Sigma(x_i - \bar{x})^2 = \Sigma x_i^2 - n(\bar{x})^2 = 140 - 7(4)^2 = 28$

2. $SS_{xy} = \Sigma(x_i - \bar{x})(y_i - \bar{y}) = \Sigma x_i y_i - n(\bar{x})(\bar{y}) = 304 - 7(4)(9) = 52$

3. $b = \frac{SS_{xy}}{SS_x} = \frac{52}{28} = 1.857$

4. $a = \bar{y} - b\bar{x} = 9 - 1.857 \cdot 4 = 1.572$

$\therefore \hat{y} = 1.572 + 1.857x$

(二) 1. $SS_y = \Sigma(y_i - \bar{y})^2 = \Sigma y_i^2 - n(\bar{y})^2 = 667 - 7(9)^2 = 100$

$r = \dfrac{SS_{xy}}{\sqrt{SS_x}\sqrt{SS_y}} = \dfrac{52}{\sqrt{28}\sqrt{100}} \quad 0.98$

2. $R^2 = r^2 = (0.98)^2 = 0.9604$

(三)1. $H_0：\rho_{xy} = 0$

2. $H_1：\rho_{xy} \neq 0$

3. $\alpha = 0.05$

4. $C = \{t \mid |t| > t_{5,\,0.025} = 2.571\}$

5. $t = \dfrac{r}{\sqrt{1-r^2/(n-2)}} = \dfrac{0.98}{\sqrt{1-0.9604/(7-2)}} = |1.0| \in C$

6.**在顯著水準**$\partial = 0.05$ **下，我們拒絕**$H_0$**，亦即**$x, y$**間的相關係數顯著不為**0。

**五、隨機區集設計的統計模型為：** $Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$， $i = 1,2,\ldots,a$， $j = 1,2,\ldots,k$， $\sum_{i=1}^{a}\alpha_i = 0 = \sum_{j=1}^{b}\beta_j$， $\varepsilon_{ij} \sim N(0,\sigma^2)$， $\varepsilon_{ij}$, $i = 1,2,\ldots,a$， $j = 1,2,\ldots,k$ **獨立。設** i**為A因子之第i個水準的效果，** $\beta_j$ **為第j個區集的效果，**SST**表總平方和，**SSA**表A因子的平方和，**SSB**為區集平方和。**

(一)**寫下**SSA**及**SSB**的公式，並求**E(SSA)**及**E(SSB)。**（十分）**

(二)**若** $Y_{ij}$ **的觀測值為** $y_{ij}$**，及定義** $h = \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}_{\cdot\cdot})^2$**，其中** $\bar{y}_{i\cdot} = \dfrac{\sum_{j=1}^{k} y_{ij}}{k}$， $\bar{y}_{\cdot j} = \dfrac{\sum_{i=1}^{a} y_{ij}}{a}$ **及** $\bar{y}_{\cdot\cdot} = \dfrac{\sum_{i=1}^{a}\sum_{j=1}^{k} y_{ij}}{ak}$ 。**若已知** $\sum_{i=1}^{a}\sum_{j=1}^{k} y_{ij}^2 = 43862$ ， $\sum_{i=1}^{a}\bar{y}_{i\cdot}^2 = 10814$ ， $\sum_{j=1}^{k}\bar{y}_{\cdot j}^2 = 14566$ ， $\bar{y}_{\cdot\cdot}^2 = 3600$**及**a=3**，**k=4**，試求**h**及**SSA**的值。（十分）**

**答：**

(一)1. $SSA = \sum_{i=1}^{a}\sum_{j=1}^{k}(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 = k\sum_{i=1}^{a}(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2$

$SSB = \sum_{i=1}^{a}\sum_{j=1}^{k}(\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})^2 = a\sum_{j=1}^{k}(\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})^2$

2.(1) $y_{ij} \overset{iid}{\ } N(M + \alpha_i + \beta_j, \sigma^2), \forall_i, j$

$$\Rightarrow \begin{cases} \bar{y}_i \overset{iid}{\ } N(\mu + \alpha_i + \beta_j, \dfrac{\sigma^2}{k}), \forall_i \\ \bar{y}_j \overset{iid}{\ } N(\mu + \alpha_i + \beta_j, \dfrac{\sigma^2}{a}), \forall_j \end{cases}$$

(2)**令** $S_A^2 = \dfrac{\sum_{i=1}^{a}(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2}{a-1}$， **則** $E(S_A^2) = \dfrac{\sigma^2}{k}$

$S_B^2 = \dfrac{\sum_{j=1}^{a}(\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})^2}{k-1}$， **則** $E(S_B^2) = \dfrac{\sigma^2}{a}$

**其中** $\bar{y}_{\cdot\cdot} = \dfrac{1}{a}\sum_{i=1}^{a}\bar{y}_{i\cdot} = \dfrac{1}{k}\sum_{j=1}^{k}\bar{y}_{\cdot j}$

(3) $E(SSA) = kE\left[\sum_{i=1}^{a}(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2\right] = kE[(a-1)S_A^2] = k(a-1)\cdot\dfrac{\sigma^2}{k} = (a-1)\sigma^2$

$E(SSB) = aE\left[\sum_{j=1}^{k}(\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})^2\right] = aE[(k-1)S_B^2] = a(k-1)\cdot\dfrac{\sigma^2}{a} = (k-1)\sigma^2$

(二)1. $SSA = k\sum(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 = k[\sum\bar{y}_{i\cdot}^2 - a\bar{y}_{\cdot\cdot}^2] = 4[10814 - 3.3600] = 56$

2. $SSB = a\sum(\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})^2 = a[\sum\bar{y}_{\cdot j}^2 - k\bar{y}_{\cdot\cdot}^2] = 3[14566 - 4\cdot 3600] = 498$

3. $SST = \sum_i\sum_j(y_{ij} - \bar{y}_{\cdot\cdot})^2 = \sum_i\sum_j y_{ij}^2 - a\cdot k(\bar{y}_{\cdot\cdot})^2 = 43862 - 3\cdot 4\cdot 3600 = 662$

4. $h = SST - SSA - SSB = 662 - 56 - 498 = 108$