# 《統計學》

| | |
|---|---|
| **試題評析** | 今年考卷難度屬於中等，第一大題數學推導較多，其中(二)要求聯合c.d.f.，不過這都在課堂中舉例說明過定義與計算方法的題目，只是因為以往高考很少考過，擔心同學們是否有認真準備這一種題型。另外三大題則屬於基礎計算題型，考生只要小心計算，獲得高分應該不難。 |
| **考點命中** | 一、《高點・高上統計學講義第二回》，趙治勳編撰，第五章。<br>二、《高點・高上統計學講義第四回》，趙治勳編撰，第九章。<br>三、《高點・高上統計學講義第四回》，趙治勳編撰，第十四章例13。<br>四、《高點・高上統計學講義第六回》，趙治勳編撰，第十三章。 |

一、設X及Y為兩個連續隨機變數，其聯合機率密度函數為當 $0\le x\le 2$ 且 $0\le y\le 2$，則 $f(x,y)=cy^2x(1+x)$；否則 $f(x,y)=0$，試問：

(一)c值為何？（10分）

(二)找出X與Y的聯合累積機率密度函數（$0\le x\le a$ 及 $0\le y\le b$），a及b介於0及2之間。（10分）

(三)X與Y是否獨立？（5分）

**答**：

(一)

根據機率公理假設 $\int_0^2\int_0^2 cy^2x(1+x)dydx=1$

$$\int_0^2\int_0^2 cy^2x(1+x)dydx=\frac{112}{9}c=1 \quad \Rightarrow c=\frac{9}{112}$$

(二)

注意：題目只要求 $0\le x<2$ 與 $0\le y<2$ 範圍內之聯合累積分配函數

$$F_{XY}(x,y)=\int_0^x\int_0^y\frac{9}{112}t_2^2t_1(1+t_1)dt_2dt_1=\frac{(3x^2+2x^3)y^3}{224},\quad 0\le x<2,0\le y<2$$

(三)

$$f_X(x)=\int_0^2\frac{9}{112}y^2x(1+x)dy=\frac{3}{14}x(1+x),0\le x\le 2$$

$$f_Y(y)=\int_0^2\frac{9}{112}y^2x(1+x)dx=\frac{3}{8}y^2,0\le y\le 2$$

$$\because f_{XY}(x,y)=f_X(x)f_Y(y) \qquad \therefore X\perp Y$$

二、你需要付出10元才能加入投擲一個六面均勻骰子的遊戲。如果骰子出現2、4、5及6朝上，你會損失所付出的那10元；如果骰子出現3，可拿回所付出的那10元；如果骰子出現1，則除了拿回所付出的那10元外，還可多得到20元。試問：



(一)此遊戲玩1次的收益之期望值及標準差為何？（5分）

(二)此遊戲玩100次的平均收益之期望值及標準差為何？（5分）

(三)如果你玩此遊戲100次，平均收益少於多少以下的機率為95%？（10分）

**答**：

(一) 令 $X$ 表收益

| $X=x$ | -10 | 0 | 20 |
|---|---|---|---|
| $f_X(x)$ | 2/3 | 1/6 | 1/6 |

$$E(X)=-10\times\frac{2}{3}+0\times\frac{1}{6}+20\times\frac{1}{6}=-\frac{10}{3}$$

$$E(X^2)=-10^2\times\frac{2}{3}+0^2\times\frac{1}{6}+20^2\times\frac{1}{6}=\frac{400}{3}$$

$$\sqrt{V(X)}=\sqrt{E(X^2)-[E(X)]^2}=11.0554$$

(二) 令 $X_i$ 表第i次遊戲之收益，$i=1,2,\cdots,100$

假設 $X_1,X_2,\ldots,X_{100}\overset{iid}{\sim}f_X(x)$

$\bar{X}$ 為100次之平均收益

$$E(\bar{X})=E(X)=-\frac{10}{3}$$

$$\sqrt{V(\bar{X})}=\sqrt{\frac{V(X)}{n}}=\frac{11.0554}{\sqrt{100}}=1.1055$$

(三) $\bar{X}\underset{byC.L.T.}{\sim}N(-\frac{10}{3},1.1055^2)$

$$P(\bar{X}\le a)=P(Z\le\frac{a-(-\frac{10}{3})}{1.1055})=0.95$$

$$\Rightarrow\frac{a-(-\frac{10}{3})}{1.1055}=z_{0.05}=1.645\quad\Rightarrow a=5.1486$$

故遊戲100次之平均收益少於5.1486之機率為95%。



三、某光電公司人資部門有三種針對組裝作業員的教育訓練方法，目的在不降低組裝品質的情況下減少作業員組裝產品所需時間。現在隨機獨立選取25位組裝作業員，將其分成三組分別以此三種教育訓練方法進行訓練，訓練結束後對組裝時間進行測試，所得結果如下表（單位：秒）：

| 方法一 | 方法二 | 方法三 |
|---|---|---|
| 148 | 82 | 15 |
| 75 | 144 | 106 |
| 170 | 32 | 13 |
| 59 | 19 | 60 |
| 95 | 114 | 200 |
| 276 | 55 | 41 |
| 71 | 62 | 51 |
| 85 | | 5 |
| 347 | | 46 |



(一)試問在信賴水準95%的情況下，根據以上所提供資訊，選擇適當的統計方法分析以此三種教

育訓練方法訓練後，其組裝作業員每人組裝時間的分配是否相同？（15分）

(二)說明為何使用小題(一)之統計方法進行分析？（10分）

**答**：

(一) K-W等級檢定法(Kruskal-Wallis rank test)

$H_0$:三種訓練方法之組裝時間相同 vs

$H_1$:三種訓練方法之組裝時間不完全相同

T.S.: $H=\dfrac{12}{n(n+1)}\sum_{j=1}^{k}\dfrac{R_{\bullet j}^2}{n_j}-3(n+1)\overset{d}{\sim}\chi^2_{(3-1=2)}$

R.R.: Reject $H_0$ at $\alpha=0.05$ if $H>\chi^2_{0.05(2)}=5.99147$

$\because \mathrm{H}^*=\dfrac{12}{25(25+1)}[\dfrac{162^2}{9}+\dfrac{84^2}{7}+\dfrac{79^2}{9}]-3(25+1)=7.2451$

$\therefore$ reject $H_0$

我們有足夠證據去推論三種訓練方法之組裝時間分配有顯著之差異。

(二)因為未知母體分配而無法使得一因子變異數分析，替代方案為無母數統計學中之K-W等級檢定法。

四、業務經理想瞭解其部門內業務代表每個月拜訪客戶的交通里程數（X，單位：仟公里）與每月總銷售額（Y，單位：仟元）間的關係。此業務經理蒐集9個月的資料，初步樣本資料的分析產生以下的資訊：

$n=9$，$\sum X=60$，$\sum Y=553$，$\sum X^2=582$，$\sum Y^2=39653$，$\sum XY=4329$，

$\sum(X-\bar{X})^2=182$，$\sum(Y-\bar{Y})^2=5674.2222$，$\sum(Y-\hat{Y})^2=3407.233$。

使用以上資料回答下列問題，請詳細將所使用之公式及計算過程列出：（每小題10分，共30分）

(一)求最小平方直線，並在散布圖中繪出此直線。

(二)以顯著水準為0.05，檢定迴歸斜率是否顯著異於0。

(三)迴歸判定係數為何？

**答**：

由題意，$SS_X=182$, $SS_Y=5674.2222$, $SSE=3407.233$

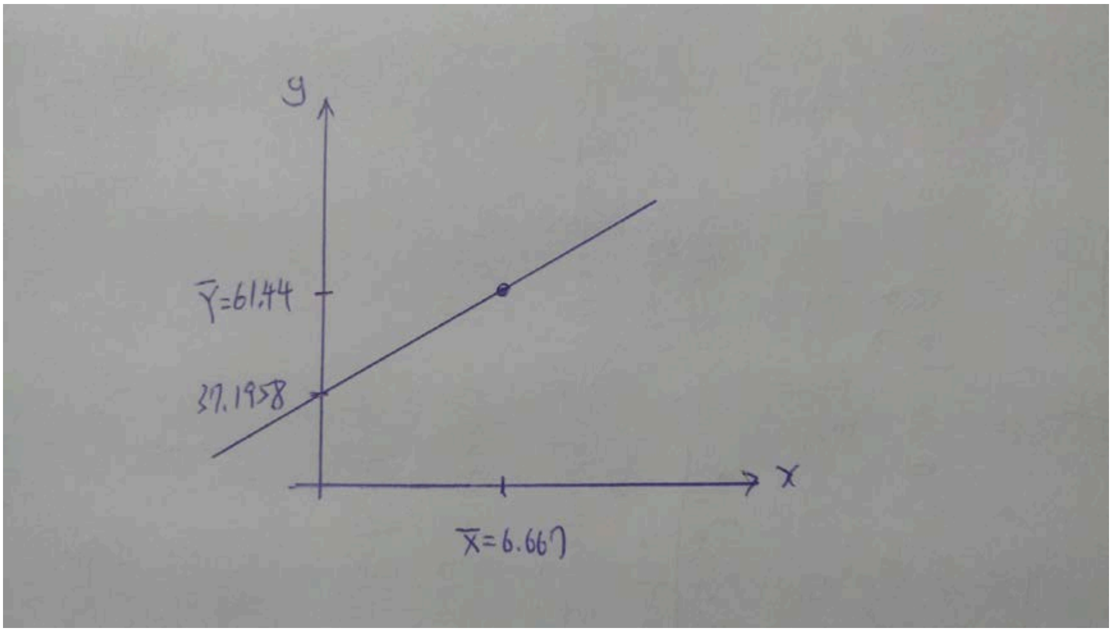
$$SS_{XY}=\sum XY-\frac{(\sum X)(\sum Y)}{n}=642.3333$$

假設模型 $Y_i=\beta_0+\beta_1X_i+\varepsilon_i,\varepsilon_i\overset{iid}{\sim}N(0,\sigma^2),i=1,2,\cdots,9$

(一) $\hat{\beta}_1=\dfrac{SS_{XY}}{SS_X}=3.5293$ $\hat{\beta}_0=\bar{Y}-\hat{\beta}_1\bar{X}=37.9158$

$\therefore\ \hat{y}=37.9158+3.5293x$

圖形:




(二) $H_0:\beta_1=0$ vs $H_1:\beta_1\neq 0$

T.S.: $T=\dfrac{\hat{\beta}_1-0}{\sqrt{\dfrac{MSE}{SS_X}}}\sim t_{(9-2=7)}$

R.R.: Reject $H_0$ at $\alpha=0.05$ if $|T^*|>t_{(7)0.025}=2.365$

$\because T=\dfrac{3.5293-0}{\sqrt{\dfrac{486.7476}{182}}}=2.1581$

其中 $MSE=\dfrac{SSE}{n-2}=\dfrac{3407.233}{7}=486.7476$

$\therefore$ don't reject $H_0$

我們沒有足夠證據去推論斜率項不為零。

(三)

$$R^2=1-\frac{SSE}{SST}=1-\frac{SSE}{SS_Y}=1-\frac{3407.233}{5674.2222}=0.3995$$