《統計學概要》

試題評析	今年普考統計人員統計學考題甚為簡單，一般考生應該皆有80分以上，第一題若沒計算錯誤，考生幾乎皆能滿分。此次考試包含二維機率分配、常態分配機率、假設檢定、變異數分析及無母數統計
考點命中	第一題：《統計學(概要)最新版》，高點文化出版，程大器編撰，頁4-34習題8。 第二題：《高點‧高上統計學總複習講義》第一回，程大器編撰，頁62範題58。 第三題：《高點‧高上統計學總複習講義》第一回，程大器編撰，頁117範題101。 第四題：《高點‧高上統計學總複習講義》第一回，程大器編撰，頁131範題113。 第五題：《高點‧高上統計學總複習講義》第一回，程大器編撰，頁140範題120。 《統計學(概要)最新版》，高點文化出版，程大器編撰，頁13-15、16例8、例9。

一、設隨機變數 X 的機率分配為

x	0	1
$f(x)$	0.4	0.6

隨機變數 Y 的機率分配為

y	1	2	3
$f(y)$	0.25	0.5	0.25

今已知 $P(X=1|Y=1)=0.6$，$P(X=0|Y=2)=0.4$，則：

(一)試求隨機變數 X 與 Y 之聯合機率分配 $f_{X,Y}(x,\ y)$ 為何？（6分）

(二)試求 $P(X+Y\le 2)=$？（6分）

(三)試求隨機變數 X 與 Y 的相關係數 $\rho_{X,Y}=$？（6分）

(四)隨機變數 X 與 Y 是否獨立？請說明之。（6分）

答：

(一) $\because P(X=1|Y=1)=0.6 \Rightarrow \dfrac{P(X=1,Y=1)}{P(Y=1)}=0.6 \Rightarrow P(X=1,Y=1)=0.6P(Y=1)$

$$\Rightarrow P(X=1,Y=1)=0.6\times 0.25=0.15$$

$P(X=0|Y=2)=0.4 \Rightarrow \dfrac{P(X=0,Y=2)}{P(Y=2)}=0.4 \Rightarrow P(X=0,Y=2)=0.4P(Y=2)$

$$\Rightarrow P(X=0,Y=2)=0.4\times 0.5=0.2$$

又 $P(Y=1)=P(X=0,Y=1)+P(X=1,Y=1)=0.25$

$\Rightarrow P(X=0,Y=1)=0.25-P(X=1,Y=1)=0.25-0.15=0.1$

$P(Y=2)=P(X=0,Y=2)+P(X=1,Y=2)=0.5$

$\Rightarrow P(X=1,Y=2)=0.5-P(X=0,Y=2)=0.5-0.2=0.3$

$P(X=0)=P(X=0,Y=1)+P(X=0,Y=2)+P(X=0,Y=3)=0.4$

$\Rightarrow P(X=0,Y=3)=0.4-P(X=0,Y=1)-P(X=0,Y=2)=0.4-0.1-0.2=0.1$

$P(X=1)=P(X=1,Y=1)+P(X=1,Y=2)+P(X=1,Y=3)=0.6$

$\Rightarrow P(X=1,Y=3)=0.6-P(X=1,Y=1)-P(X=1,Y=2)=0.6-0.15-0.3=0.15$

故 X 與 Y 之聯合機率分配為

x \ y	1	2	3
0	0.1	0.2	0.1
1	0.15	0.3	0.15

(二) $P(X+Y\le 2)=P(X=0,Y=1)+P(X=0,Y=2)+P(X=1,Y=1)$

$$=0.1+0.2+0.15=0.45$$

(三)$\because E(X)=\sum_x x\cdot f(x)=0\times 0.4+1\times 0.6=0.6$

$$E(X^2)=\sum_x x^2\cdot f(x)=(0)^2\times 0.4+(1)^2\times 0.6=0.6$$

$$E(Y)=\sum_y y\cdot f(y)=1\times 0.25+2\times 0.5+3\times 0.25=2$$

$$E(Y^2)=\sum_y y^2\cdot f(y)=(1)^2\times 0.25+(2)^2\times 0.5+(3)^2\times 0.25=4.5$$

$$E(XY)=\sum_x\sum_y xy\cdot f(x,y)=0\times 1\times 0.1+0\times 2\times 0.2+0\times 3\times 0.1+1\times 1+1.5+1\times 2\times 0.3+1\times 3\times 0.15=1.2$$

$$Var(X)=E(X^2)-\left(E(X)\right)^2=0.6-(0.6)^2=0.24$$

$$Var(Y)=E(Y^2)-\left(E(Y)\right)^2=4.5-(2)^2=0.5$$

$$Cov(X,Y)=E(XY)-E(X)\cdot E(Y)=1.2-0.6\times 2=0$$

故 X 與 Y 之相關係數爲

$$\rho_{X,Y}=\frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}=\frac{0}{\sqrt{0.24}\sqrt{0.5}}=0$$

(四)$\because f_{x,y}(0,1)=0.1=0.4\times 0.25=f_x(0)\cdot f_y(1)$

$$f_{x,y}(0,2)=0.2=0.4\times 0.5=f_x(0)\cdot f_y(2)$$

$$f_{x,y}(0,3)=0.1=0.4\times 0.25=f_x(0)\cdot f_y(3)$$

$$f_{x,y}(1,1)=0.15=0.6\times 0.25=f_x(1)\cdot f_y(1)$$

$$f_{x,y}(1,2)=0.3=0.6\times 0.5=f_x(1)\cdot f_y(2)$$

$$f_{x,y}(1,3)=0.15=0.6\times 0.25=f_x(1)\cdot f_y(3)$$

故知 $f_{x,y}(x,y)=f_x(x)\cdot f_y(y)$ 即 X 與 Y 爲獨立隨機變數

二、臺北第一超市欲向梨山向陽蘋果園採購1000箱蘋果，向陽蘋果園出產之蘋果重量（個）為一常態分配，其平均重量為360公克，標準差為30公克。今臺北第一超市派出採購部王經理，進行蘋果品質抽樣檢驗工作。試問：

（一）王經理自向陽蘋果園隨機抽取一個蘋果，則該顆蘋果重量少於330公克的機率為何？（8分）

（二）若向陽蘋果園以50個蘋果裝成一箱裝運，王經理自該園隨機抽取一箱檢驗，則該箱蘋果的重量在18±0.5公斤的機率為何？（8分）

答：

(一)設隨機變數 X_i 表示蘋果重量，則 $X_i\sim N(360, 30^2)$，故

$$P(X_i<330)=P\left(\frac{X_i-\mu}{\sigma}<\frac{330-360}{30}\right)=P(Z<-1)=0.1587$$

(二)令 $W=\sum_{i=1}^{50}X_i$，表示一箱重量，則 $W\sim N(18000, 45000)$，故知

$$P(17500\le W\le 18500)=P\left(\frac{17500-18000}{\sqrt{45000}}\le Z\le\frac{18500-18000}{\sqrt{45000}}\right)$$
$$=P(-2.36\le Z\le 2.36)=0.9818$$

三、設（$X_1,X_2,\cdots,X_{100}$）為抽自母體變異數為100，平均數 μ 為未知的常態分配之一組大小為 $n=100$ 之隨機樣本，今欲利用此組隨機樣本來檢定 H_0：$\mu=75$ 對 H_1：$\mu=78$，且訂定檢定規則（危險

域）為 $C=\{\bar{X}\geq k\}$，則：

（一）若取顯著水準為 $\alpha=0.025$，試決定上述檢定規則（危險域）C中之常數k之值為何？（8分）

（二）試求上述(一)中所訂定之檢定規則C的檢定力（power）為何？（8分）

（三）試求上述(一)中所訂定之檢定規則C會產生多大的型二誤差（typeⅡ error）之機率，即 $\beta=$？（8分）

答：

(一) $P(\bar{X}\geq k\mid\mu=75)=0.025\Rightarrow P\left(Z\geq\dfrac{k-75}{10/\sqrt{100}}\right)-0.025$

$\therefore\dfrac{k-75}{10/\sqrt{100}}=1.96\Rightarrow k=76.96$

(二) $1-\beta=P(\bar{X}\geq 76.96\mid\mu=78)=P\left(Z\geq\dfrac{76.96-78}{10/\sqrt{100}}\right)=P(Z\geq -1.04)=0.8508$

(三) $\beta=P(\bar{X}<76.96\mid\mu=78)=P\left(Z<\dfrac{76.96-78}{10/\sqrt{100}}\right)=P(Z<-1.04)=0.1492$

四、以下是一項調查1000位學生，有關學生是否有吸菸習慣與其父母吸不吸菸的調查研究，得資料如下：

	父母皆吸菸	父或母親吸菸	父母皆不吸菸
學生不吸菸	150	300	450
學生吸菸	20	30	50

（一）試以顯著水準 $\alpha=0.05$之下，檢定學生是否吸菸與其父母吸不吸菸間是否有關？請寫出相關檢定的所有步驟與最後檢定之結果。（10分）

（二）請說明你（妳）所用的統計檢定方法其名稱為何？（6分）

答：

(一)1. H_0：學生是否吸菸與其父母吸不吸菸間無關

2. H_1：學生是否吸菸與其父母吸不吸菸間相關

3. $\alpha=0.05$

4. $C=\left\{\chi^2\mid\chi^2>\chi^2_{0.05}(2)=5.9915\right\}$

5.計算：若 H_0 爲真，則各格子之期望次數 $e_{ij}=\dfrac{R_i\cdot C_j}{n}$ 計算於下表括號中，即

	父母皆吸菸	父或母親吸菸	父母皆不吸菸	合計
學生不吸菸	150(153)	300(297)	450(450)	900
學生吸菸	20(17)	30(33)	50(50)	100
合計	170	330	500	1000

故檢定統計量之值爲

$$\chi^2=\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(o_{ij}-e_{ij})^2}{e_{ij}}=\frac{(150-153)^2}{153}+\frac{(300-2997)^2}{297}+\cdots+\frac{(50-50)^2}{50}=0.8913\notin C$$

6.結論：不拒絕 H_0，亦即沒有證據顯示學生是否吸菸與其父母吸不吸菸間有關。

(二)獨立性檢定

五、已知自三個具有相同變異數 σ^2 之常態母體，分別獨立的隨機抽出樣本，經整理得樣本資料訊息如下表所示：

母體	樣本數（n_i）	平均數（$\overline{x}_i$）	變異數（s_i^2）
1	3	13	25
2	5	14	16
3	7	15	9

（一）試依此資料訊息，取顯著水準 $\alpha = 0.05$，檢定此三個母體的平均數是否全相等？假設此資料適合做變異數分析。（註：參考統計值 $F_{0.05}(2,12) = 3.89$，$F_{0.05}(3,12) = 3.49$）（12分）

（二）變異數 σ^2 的估計值為何？（8分）

答：

(一)1. $H_0 : \mu_1 = \mu_2 = \mu_3$

2. $H_1 : \mu_i$ 不全相等，$i = 1,2,3$

3. $\alpha = 0.05$

4. $C = \left\{F \mid F > F_{0.05}(2,12) = 3.89\right\}$

5.計算：因 $SSE = \sum_{i=1}^{k}(n_i - 1)S_i^2 = (3-1)\times 25 + (5-1)\times 16 + (7-1)\times 9 = 168$

又 $\bar{X}_{\cdot\cdot} = \dfrac{\sum_{i=1}^{k} n_i \cdot \bar{X}_{i\cdot}}{N} = \dfrac{3\times 13 + 5\times 14 + 7\times 15}{15} = \dfrac{214}{15}$

$\therefore SSTR = \sum_{i=1}^{k} n_i(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2 = 3\times\left(13 - \dfrac{214}{15}\right)^2 + 5\times\left(14 - \dfrac{214}{15}\right)^2 + 7\times\left(15 - \dfrac{214}{15}\right)^2 = 8.933$

又 $F = \dfrac{MSTR}{MSE} = \dfrac{8.933/2}{168/12} = 0.319 \notin C$

6.結論：不拒絕 H_0，亦即無證據顯示三個母體平均數有顯著差異。

(二) σ^2之估計值爲 $MSE = \dfrac{SSE}{N-k} = \dfrac{168}{12} = 14$