《統計學》

試題評析	今年高考統計的統計學題目是基本題型，很好把握！前三題考驗的是同學平時的穩定度與寫題的速度，大部分題型在我們的小考中都有出現；最後一題（第四題）考到實驗設計的範圍計算量有點多，同學只要多點耐心，拿分應該不是問題。若要上榜，同學至少要考到85分以上。

（附$F_{0.05}$、χ^2值表）

一、假設兩獨立樣本分別取自兩個常態母體，其母體變異數分別為σ_1^2和σ_2^2。令(S_1^2, n_1)和(S_2^2, n_2)分別為兩組樣本的樣本變異數和樣本大小（sample size）。

(一)請分別推導求得兩母體變異數比（σ_1^2/σ_2^2）與兩母體標準差比（σ_1/σ_2）之$100(1-\alpha)\%$信賴區間（confidence interval）估計式。α為顯著水準（the level of significance）。（14分）

(二)分別自兩個不同廠牌的汽水罐裝填機器隨機抽取樣本並測量其汽水罐容量（單位：ml）。樣本大小分別為（$n_1=15$，$n_2=17$），計算得樣本容量變異數分別為（$S_1^2=2$，$S_2^2=4$）。請分別計算σ_1^2/σ_2^2與σ_1/σ_2之90%信賴區間估計值。假設汽水容量呈常態分配。（6分）

考點命中	《高點・高上統計學講義》第五回，楊老師編撰，6-2區間估計，頁18-19。

答：

(一)假設兩母體平均數μ_1,μ_2未知

1.點估計 $\dfrac{S_1^2}{S_2^2}\xrightarrow{\text{估計}}\dfrac{\sigma_1^2}{\sigma_2^2}$，其中$S_1^2=\dfrac{\sum(X_i-\bar{X}_1)^2}{n_1-1}$

$$S_2^2=\frac{\sum(X_i-\bar{X}_2)^2}{n_2-1}$$

2.樞紐量 $F=\dfrac{S_1^2}{S_2^2}\dfrac{\sigma_2^2}{\sigma_1^2}\sim F(n_1-1,n_2-1)$

3.合理機率區間 $1-\alpha=P(F_{1-\alpha/2}(n_1-1,n_2-1)<F=\dfrac{S_1^2}{S_2^2}\dfrac{\sigma_2^2}{\sigma_1^2}<F_{\alpha/2}(n_1-1,n_2-1))$

4.兩個變異數比$\dfrac{\sigma_1^2}{\sigma_2^2}$之$100(1-\alpha)\%$信賴區間為$\left(\dfrac{S_1^2}{S_2^2}\dfrac{1}{F_{\alpha/2}(n_1-1,n_2-1)},\dfrac{S_1^2}{S_2^2}\dfrac{1}{F_{1-\alpha/2}(n_1-1,n_2-1)}\right)$

兩個標準差比$\dfrac{\sigma_1}{\sigma_2}$之$100(1-\alpha)\%$信賴區間為$\left(\sqrt{\dfrac{S_1^2}{S_2^2}\dfrac{1}{F_{\alpha/2}(n_1-1,n_2-1)}},\sqrt{\dfrac{S_1^2}{S_2^2}\dfrac{1}{F_{1-\alpha/2}(n_1-1,n_2-1)}}\right)$

(二)$n_1=15$，$s_1^2=2$；$n_2=17$，$s_2^2=4$

變異數比值估計值為$\dfrac{s_1^2}{s_2^2}=\dfrac{2}{4}=0.5$

查表值 $F_{\alpha/2}(n_1-1,n_2-1)=F_{0.05}(14,16)=2.37$

$$F_{1-\alpha/2}(n_A-1,n_B-1)=F_{0.95}(14,16)=\frac{1}{F_{0.05}(16,14)}=\frac{1}{2.44}$$

兩變異數比值 σ_1^2 / σ_2^2 90%信賴區間 $\left(\frac{s_1^2}{s_2^2}\frac{1}{F_{\alpha/2}(n_1-1,n_2-1)},\frac{s_1^2}{s_2^2}\frac{1}{F_{1-\alpha/2}(n_1-1,n_2-1)}\right)$

$$\Rightarrow (0.5\times\frac{1}{2.37},0.5\times 2.44)\Rightarrow(0.211,1.22)$$

兩標準差比值 σ_A / σ_B 95%信賴區間為

$$\left(\sqrt{\frac{s_A^2}{s_B^2}\frac{1}{F_{\alpha/2}(n_A-1,n_B-1)}},\sqrt{\frac{s_A^2}{s_B^2}\frac{1}{F_{1-\alpha/2}(n_A-1,n_B-1)}}\right)\quad\Rightarrow\ (0.4593,1.22)$$

二、令隨機變數 X 和 Y 為獨立的卜瓦松分配（Poisson distribution），其參數各為 λ_1 和 λ_2。令 $Z=X+Y$。（每小題10分，共20分）

(一)推導求得隨機變數 Z 的機率分配（probability distribution）。

(二)若 $\lambda_1=3$ 和 $\lambda_2=5$，試求隨機變數 Z 之變異數（variance）與 $Z<3$ 的機率（$P(Z<3)=?$）。

考點命中	1.《高點・高上統計學講義》第二回，楊老師編撰，3-2 mgf 變數變換，頁13、49。 2.《高點・高上統計學講義》第三回，楊老師編撰，4-1一維常用分配離散型，頁10。

答：

(一) $Z=X+Y$ and $M_X(t)=e^{\lambda_1(e^t-1)}$，$t\in\mathbb{R}$，$M_Y(t)=e^{\lambda_2(e^t-1)}$，$t\in\mathbb{R}$

$$M_Z(t)=E(e^{tZ})=E(e^{tX+tY})=E(e^{tX})E(e^{tY})=M_X(t)M_Y(t)=e^{(\lambda_1+\lambda_2)(e^t-1)},\ t\in\mathbb{R}$$

由mgf唯一性得知 $Z\sim poi(\lambda_1+\lambda_2)$

即 $f_Z(z)=\dfrac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^z}{z!}$，$z=0,1,2,\ldots\ldots$

(二) $\lambda_1=3$，$\lambda_2=5\ \Rightarrow\ Z\sim poi(8)$，$Var(Z)=8$

$$P(Z<3)=P(Z=0)+P(Z=1)+P(Z=2)=\frac{e^{-8}8^0}{0!}+\frac{e^{-8}8^1}{1!}+\frac{e^{-8}8^2}{2!}$$

$$=e^{-8}(1+8+32)=0.0138$$

三、假設隨機變數 X_j，$j=1,2,3,4$，為互相獨立的同分配，期望值為 μ，且變異數為 σ^2。令 $Y_1=(X_2-X_1)^2$，$Y_2=(X_4-X_3)^2$。

(一)推導求得 Y_i 之期望值 $(\mathrm{E}(Y_i))$，$i=1,2$。（10分）

(二)若 $Y_i>\mathrm{E}(Y_i)$，則令 $I_i=1$；否則 $I_i=0$。令 $\mathrm{M}=\sum_{i=1}^{2}I_i$，推導求得 M 之分配。（5分）

(三)若 X_j 為常態分配，推導求得 $Y_i>\mathrm{E}(Y_i)$ 之機率（列出計算的過程即可，不必寫出機率值）。（10分）

考點命中	1.《高點・高上統計學講義》第二回，楊老師編撰，3-2期望值變異數，頁48。 2.《高點・高上統計學講義》第四回，楊老師編撰，5-4常態母體下抽樣分配，頁22。

答：

(一) $Y_i=(X_{2i}-X_{2i-1})^2$，$i=1,2$

$$E(Y_i)=E[(X_{2i}-X_{2i-1})^2]=Var(X_{2i}-X_{2i-1})+[E(X_{2i}-X_{2i-1})]^2$$
$$=Var(X_{2i})-Var(X_{2i-1})+[E(X_{2i})-E(X_{2i-1})]^2=2\sigma^2\ ,\ i=1,2$$

(二) $I_i=\begin{cases}1\ ,\ if\ \ Y_i>2\sigma^2\\ 0\ ,\ o.w.\end{cases}\quad i=1,2$ ， $M=I_1+I_2$ 為 $Y_i>2\sigma^2$ 之個數

故 M 為二項分配參數 $n=2$ 與 $p=P(Y_i>2\sigma^2)$

(三)若 X_j 為常態分配，則 $X_{2i}-X_{2i-1}\sim N(0,2\sigma^2)$

$$\frac{X_{2i}-X_{2i-1}}{\sqrt{2\sigma^2}}\sim N(0,1)\ \Rightarrow\ \frac{(X_{2i}-X_{2i-1})^2}{2\sigma^2}\sim\chi^2_{(1)}\ ,\ i=1,2$$

$$P(Y_i>2\sigma^2)=P[(X_{2i}-X_{2i-1})^2>2\sigma^2]=P[\frac{(X_{2i}-X_{2i-1})^2}{2\sigma^2}>1]$$

$=P(\mathrm{W}>1)$ 其中 $\mathrm{W}\sim\chi^2_{(1)}$ 。

四、一個完全隨機設計之二因子實驗之因子、因子水準和8個反應值如表所示：

X1	X2	反應值（y）	
–1	–1	20	30
–1	1	40	50
1	–1	50	60
1	1	12	16

(一)8個實驗之實驗順序如何決定？（3分）

(二)寫出變異數分析的固定效應模式（fixed effects model）（須考慮因子之交互作用）及其假設。假設因子是固定的（fixed factors）。（10分）

(三)列出變異數分析表並檢定X1、X2和X1X2（交互作用）之效應是否顯著。顯著水準皆為0.05。（寫出虛無假設和對立假設，並說明檢定統計量之分配。）（10分）

(四)請依(三)之檢定結果，寫出此因子實驗之配適後迴歸模型或反應曲面（response surface）。假設 $-1\le \mathrm{X1},\ \mathrm{X2}\le 1$。（12分）

考點命中	《高點・高上統計學講義》第七回，楊老師編撰，CH8變異數分析，頁1、14、15。

答：

(一) 8個實驗順序隨機選取決定。

(二)固定因子模型假設：

$Y_{ijk}=\mu_{ij}+\varepsilon_{ijk}=\mu+\alpha_i+\beta_j+(\alpha\beta)_{ij}+\varepsilon_{ijk}\ \ ,i=1,2,\ \ j=1,2,\ \ k=1,2$

Y_{ij} ：因子1第 i 組與因子2第 j 組下的第 k 個反應值

μ ：表示總平均數。

μ_{ij} ：因子1第 i 組與因子2第 j 組下平均數

α_i ：因子1第 i 組處理效果

β_j ：因子2第 j 組處理效果

$(\alpha\beta)_{ij}$ ：因子1第 i 組與因子2第 j 組下之交互影響

ε_{ijk} ：因子1第 i 組與因子2第 j 組下的第 k 個誤差項，其中 $\varepsilon_{ijk}\overset{iid}{\sim}N(0,\sigma^2)$

固定效應模式 α_i，β_j, $(\alpha\beta)_{ij}$ 為常數，且需假設 $\sum_{i=1}^{2}\alpha_i = 0$，$\sum_{j=1}^{2}\beta_j = 0$，$\sum_{i=1}^{2}(\alpha\beta)_{ij} = 0$，$\sum_{j=1}^{2}(\alpha\beta)_{ij} = 0$

(三)先將資料轉換如下表

X2 \ X1	-1	1	Total
-1	20,30	50,60	160
1	40,50	12,16	118
Total	140	138	278

$SST = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk} - \bar{Y})^2 = 2239.5$

$SSX1 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{i..} - \bar{Y})^2 = 0.5$　　$SSX2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{.j.} - \bar{Y})^2 = 220.5$

$$SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk} - \bar{Y}_{ij})^2 = 50+50+50+8 = 158$$

SSI=SST- SSX1- SX21-SSE=1860.5

ANOVA表

Source	*SS*	*df*	*MS*	*F*
X1	0.5	1	0.5	0.0127
X2	220.5	1	220.5	5.5823
交互	1860.5	1	1860.5	47.1013
誤差	158	4	39.5	
總合	2239.5	7		

①：X1和X2無交互作用

H_1：X1和X2有交互作用

H_0：因子X1效應不存在　　H_0：因子X2效應不存在

H_1：因子X1效應存在　　H_1：因子X2效應存在

②由ANOVA表可知檢定統計量 $F_I^* = 47.1013$，$F_1^* = 0.127$，$F_2^* = 5.5823$

③ $\alpha = 0.05$，$C = \{F^* \mid F^* > F_{0.05}(1,4) = 7.71\}$，$F_I^* = 47.1013 \in C$

④故在 $\alpha = 0.05$ 情況下，根據樣本資料顯示，X1與X2交互作用顯著。

(四)反應曲面 $\hat{y} = 34.75 - 0.25x_1 - 5.25x_2 - 15.25x_1x_2$