110 年公務人員高等考試三級考試試題

類　　科：統計
科　　目：統計學

參考值：
$Z_{0.025}=1.96$，$Z_{0.05}=1.645$，$z_{0.1}=1.28$，
$t_{0.025,8}=2.306$，$t_{0.025,9}=2.262$，$t_{0.025,10}=2.228$，$t_{0.05,8}=1.860$，$t_{0.05,9}=1.833$，$t_{0.05,10}=1.812$
$F_{0.025,2,8}=6.059$，$F_{0.025,4,8}=6.053$，$F_{0.05,2,8}=4.459$，$F_{0.05,4,8}=3.838$

一、令 X_1 與 X_2 為具獨立同分布、期望值 $1/\lambda$ 的指數(exponential)隨機變數。
令 $Y_1=X_1-X_2$ 以及=$Y_2=X_2$。(每小題 10 分，共 20 分)
㈠試求 Y_1 與 Y_2 之聯合機率密度函數。
㈡試求 Y_1 之邊際機率密度函數。

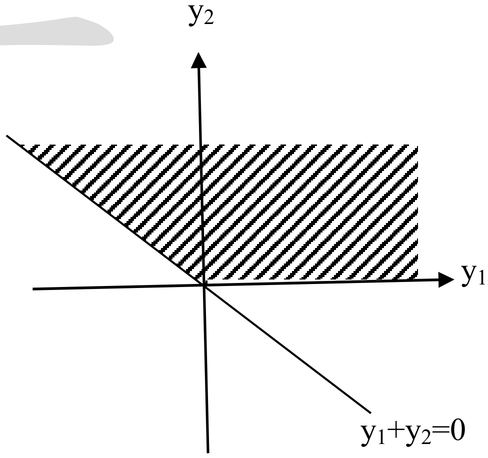
***1.*《考題難易》**：★★
***2.*《解題關鍵》**：雙變數函數之變數變換
***3.*《命中特區》**：吳迪著統計學 P4-21~P4-22

【擬答】：

㈠(X_1,X_2) iid $\exp(\lambda)$

$$f(X_1,X_2)=f(X_1)f(X_2)=\lambda e^{-\lambda X_1},\lambda e^{-\lambda X_2}$$
$$=\lambda^2 e^{-\lambda(X_1+X_2)},X_1>0,X_2>0$$
$$\begin{cases}Y_1=X_1-X_2\\Y_2=X_2\end{cases}\Rightarrow\begin{cases}X_1=Y_1+Y_2\\X_2=Y_2\end{cases}$$
$$J=\frac{\partial(X_1,X_2)}{\partial(y_1,y_2)}=\begin{vmatrix}\frac{\partial X_1}{\partial y_1}&\frac{\partial X_1}{\partial y_2}\\\frac{\partial X_2}{\partial y_1}&\frac{\partial X_2}{\partial y_2}\end{vmatrix}$$
$$=\begin{vmatrix}1&1\\0&1\end{vmatrix}=1$$
$$\Rightarrow f(y_1,y_2)=\lambda^2 e^{-\lambda(y_1+2y_2)},\begin{matrix}y_1+y_2>0\\y_2>0\end{matrix}$$

㈡



1. $y_1<0$
$$f(y_1)=\int_{-y_1}^{\infty}\lambda^2 e^{-\lambda y_1}.e^{-2\lambda y_2}dy_2$$
$$=\lambda^2 e^{-\lambda y_1}\left[\frac{1}{-2\lambda}e^{-2\lambda y_2}\right]\Big|_{-y_1}^{\infty}$$
$$=\frac{\lambda}{2}e^{\lambda y_1}$$

2. $y_1\geq 0$
$$f(y_1)=\int_0^{\infty}\lambda^2 e^{-\lambda y_1}e^{-2\lambda y_2}dy_2$$
$$=\lambda^2 e^{-\lambda y_1}\left[\frac{1}{-2\lambda}e^{-2\lambda y_2}\right]\Big|_0^{\infty}$$

$$=\frac{\lambda}{2}e^{-\lambda y_1}$$

$$\therefore f(y_1)=\begin{cases}\frac{\lambda}{2}e^{-\lambda y_1}, y_1\ge 0\\ \frac{\lambda}{2}e^{\lambda y_1}, y_1<0\end{cases}$$

二、令 $X_1,X_2,\cdots X_n$ 為一組隨機抽自常態分配 $N(0,\sigma^2)$ 之樣本。假設 $\sigma_1^2>\sigma_0^2$，在顯著水準 0.05 下：

㈠試求檢定 H_0: $\sigma^2=\sigma_0^2$ VS. H_1: $\sigma^2=\sigma_1^2$ 的最強力檢定(mostpowerfultest)。(10 分)

㈡試求檢定 H_0: $\sigma^2=\sigma_0^2$ VS. H_1 ：$\sigma^2>\sigma_0^2$ 的齊一最強力檢定(uniformly most powerful test)。(5 分)

㈢說明檢定 H_0: $\sigma^2=\sigma_0^2$ VS. H_1: $\sigma^2\ne\sigma_0^2$ 的齊一最強力檢定是否存在。(5 分)

***1.*《考題難易》：★★★**
***2.*《解題關鍵》：最強力檢定與概度比檢定**
***3.*《命中特區》：吳迪著統計學 P8-51~P8-53**

【擬答】：

㈠ $\begin{cases}H_0: \sigma^2=\sigma_0{}^2\\ H_1: \sigma^2=\sigma_1{}^2\end{cases}$

$(X_1,X_2,\ldots\ldots,X_n)\overset{\text{iid}}{\sim}N(0,\sigma^2)$

$$L(\sigma^2)=\prod_{i=1}^{n}f(x_1;\sigma^2)=(\frac{1}{\sqrt{2\pi\sigma}})^n e^{-\frac{\sum_{i=1}^{n}x_1^2}{2\sigma^2}}$$

$$\Rightarrow\frac{L(\sigma_0^2)}{L(\sigma_1^2)}=\frac{(\frac{1}{\sqrt{2\pi}\sigma_0})^n e^{-\frac{\sum_{i=1}^{n}x_1^2}{2\sigma_1^2}}}{(\frac{1}{\sqrt{2\pi}\sigma_1})^n e^{-\frac{\sum_{i=1}^{2}x_1^2}{2\sigma_1^2}}}=(\frac{\sigma_1}{\sigma_0})^n e^{\frac{(\sigma_1^2-\sigma_0^2)\sum_{i=1}^{n}x_1^2}{2\sigma_0^2\sigma_1^2}}$$

$$令\frac{L(\sigma_0^2)}{L(\sigma_1^2)}\le K\Rightarrow(\frac{\sigma_1}{\sigma_0})^n e^{\frac{(\sigma_1^2-\sigma_0^2)\sum_{i=1}^{n}x_1^2}{2\sigma_1^2\sigma_1^2}}\le K$$

$$\Rightarrow e^{-\frac{(\sigma_1^2-\sigma_0^2)\sum_{i=1}^{2}x_1^2}{2\sigma_0^2\sigma_1^2}}\le(\frac{\sigma_0}{\sigma_1})^n K$$

$$\Rightarrow-\frac{(\sigma_1^2-\sigma_0^2)\sum_{i=1}^{n}x_1^2}{2\sigma_0^2\sigma_1^2}\le\ell n〔(\frac{\sigma_0}{\sigma_1})^n K〕$$

$$\Rightarrow\frac{\sum_{i=1}^{n}x_1^2}{\sigma_0^2}\ge(-\frac{2\sigma_0^2}{\sigma_1^2-\sigma_0^2})\ell n[(\frac{\sigma_0}{\sigma_1})^n K]=k'$$

由 Neyman-Pearson Lemma

此檢定最佳拒絕域為 $C=\{\frac{\sum_{i=1}^{n}x_1^2}{\sigma_0^2}\mid\frac{\sum_{i=1}^{n}x_1^2}{\sigma_0^2}\ge k'\}$

其中 $K'=(-\frac{2\sigma_0^2}{\sigma_1^2-\sigma_0^2})\ell n〔(\frac{\sigma_0}{\sigma_1})^n K〕$ 由 α 決定

又 $\frac{\sum_{i=1}^{n}x_1^2}{\sigma_0^2}=\frac{\sum_{i=1}^{n}(x_i-0)^2}{\sigma_0^2}\sim x^2(n)$

且 $\alpha=0.05$

由查表得 $K'=\chi^2{}_{0.05}(n)$

由此拒絕域所做的檢定稱為最強力檢定

㈡ $\begin{cases} H_0: \sigma^2 = {\sigma_0}^2 \\ H_1: \sigma^2 > {\sigma_0}^2 \end{cases}$

由㈠

又在 H_1: $\sigma_1 > \sigma_0$ 的條件下，此最佳拒絕域只有跟 σ_0 有關，與 σ_1 無關，所以在 H_1: $\sigma_1 > \sigma_0$ 中的每一點均以此拒絕域為最佳拒絕域，即以此拒絕域 $C = \left\{ \frac{\sum_{i=1}^{n} x_i^2}{\sigma_0^2} \mid \frac{\sum_{i=1}^{n} x_i^2}{\sigma_0^2} \geq k' \right\}$ 所做的檢定稱為齊一最強檢定(UMPT)。

㈢ $\begin{cases} H_0: \sigma^2 = {\sigma_0}^2 \\ H_1: \sigma^2 \neq {\sigma_0}^2 \end{cases}$

在 H_0 條件下 $\hat{\sigma}^2 = {\sigma_0}^2$

在 $H_0 \cup H_1$ 條件下 $\hat{\sigma}^2$ 的 MLE 為 $\frac{1}{n}\left[\sum_{i=1}^{n} (x_i - \bar{x})^2\right]$ 存在

且 $\lambda = \dfrac{L(\sigma^2 = {\sigma_0}^2)}{L(\sigma^2 = {\hat{\sigma}^2}_{MLE})}$

令 $\lambda \leq K$，

由㈠㈡得知可以得到一個最佳拒絕域 C

由此最佳拒絕域所做的檢定稱為概度比檢定(LRT)

三、以下是(X,Y)兩變數之觀測資料：

X	11	9	14	10	12	15	7	5	13	8	6
Y	7.46	6.77	12.74	7.11	7.81	8.84	6.08	5.39	8.15	6.42	5.73

以下考慮皮爾森相關係數(Pearson'scorrelationcoefficient r)與皮爾曼等級相關係數(Spearman's rank correlationcoefficient r_s)。

㈠試畫出(X,Y)之散布圖，並試計算 r 與 r_s。(10 分)

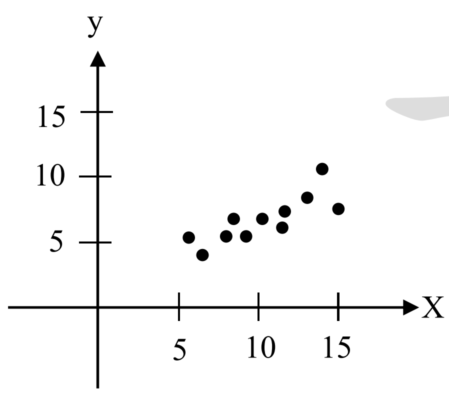
㈡試刪去本數據中之離群子後，重新計算 r 與 r_s。(5 分)

㈢試問 r 與 r_s 何者容易受離群子影響?(5 分)

1.《考題難易》：★★
2.《解題關鍵》：簡單迴歸分析與 spearman 等級相關
3.《命中特區》：吳迪著統計學 P12-33 範例 21

【擬答】：

㈠



1. $\Sigma X = 110, \Sigma X^2 = 1210$
   $\Sigma Y = 82.5, \Sigma Y^2 = 659.9762$
   $\Sigma XY = 879.97$

$$r_{XY}=\frac{\Sigma XY-\frac{(\Sigma X)(\Sigma Y)}{n}}{\sqrt{\Sigma X^2-\frac{(\Sigma X)^2}{n}}\sqrt{\Sigma Y^2-\frac{(\Sigma Y)^2}{n}}}$$

$$=\frac{879.97-\frac{110\times 82.5}{11}}{\sqrt{1210-\frac{110^2}{11}}\sqrt{659.9762-\frac{82.5^2}{11}}}$$

=0.8163

2.

R	11 (7)	9 (5)	14 (10)	10 (6)	12 (8)	15 (11)	7 (3)	5 (1)	13 (9)	8 (4)	6 (2)
S	7.46 (7)	6.77 (5)	12.74 (11)	7.11 (6)	7.81 (8)	8.84 (10)	6.08 (3)	5.39 (1)	8.15 (9)	6.42 (4)	5.73 (2)
d-R-S	0	0	-1	0	0	1	0	0	0	0	0

$$r_s=1-\frac{6\Sigma d^2}{n(n^2-1)}$$

$$=1-\frac{6[(-1)^2+1^2]}{11(11^2-1)}=0.9909$$

㈡離群值(14,12.74)刪除後資料如下：

X (R)	11 (7)	9 (5)	10 (6)	12 (8)	15 (11)	7 (3)	5 (1)	13 (9)	8 (4)	6 (2)
Y (S)	7.46 (7)	6.77 (5)	7.11 (6)	7.81 (8)	8.84 (10)	6.08 (3)	5.39 (1)	8.15 (9)	6.42 (4)	5.73 (2)
d-R-S	0	0	0	0	1	0	0	0	0	0

1. $\Sigma X=96, \Sigma X^2=1014$

$\Sigma Y=69.76, \Sigma Y^2=497.6686$

$\Sigma XY=701.61$

$$r_{XY}=\frac{\Sigma XY-\frac{(\Sigma X)(\Sigma Y)}{n}}{\sqrt{\Sigma X^2-\frac{(\Sigma X)^2}{n}}\sqrt{\Sigma Y^2-\frac{(\Sigma Y)^2}{n}}}$$

$$=\frac{701.61-\frac{96\times 69.76}{10}}{\sqrt{1014-\frac{96^2}{10}\sqrt{497.6686-\frac{69.76^2}{10}}}}$$

=0.9992

2. $r_S=1-\frac{6\Sigma d^2}{n(n^2-1)}=1-\frac{6\times 0}{10(10^2-1)}$

=1

3. r 比 r_s 容易受離群值的影響

因為 r 相差 0.9992-0.8163=0.1892

r_s 相差 1-0.9992=0.0008

四、甲公司之零件製造部門有三台機器，輪流由五名員工(ABCDE)負責操作。李主任擬研究不同機器以及不同員工之生產量是否不同。以下是隨機抽取之生產量資料：

	機器一	機器二	機器三
員工 A	31	25	35
員工 B	33	26	33
員工 C	28	24	30
員工 D	30	29	28
員工 E	28	26	27

㈠試寫出 ANOVA 表(Analysis of VarianceTable。(5 分)

㈡在顯著水準 0.05 下，試檢定不同機器之生產量是否不同。(5 分)

㈢在顯著水準 0.05 下，試檢定不同員工之生產量是否不同。(5 分)

㈣試寫出模型假設。(5 分)

1.《考題難易》：★
2.《解題關鍵》：變異數分析隨機集區設計
3.《命中特區》：吳迪著統計學 P9-38 範例 17

【擬答】：

㈠

員工＼機器	一	二	三	和	
A	31	25	35	91	T1.
B	33	26	33	92	T2.
C	28	24	30	82	T3.
D	30	29	28	87	T4.
E	28	26	27	81	T5.
和	150 T.1	130 T.2	153 T.3	433 T..	

$$\sum\sum X_{ij}{}^2 = 31^2+25^2+\ldots+27^2 = 12639$$

r=5，c=3，N=15

1. SST=$\sum\sum X_{ij}{}^2 - \frac{T..^2}{N} = 12639 - \frac{433^2}{15} = 139.73$

2. SSR=$\frac{1}{c}\sum_{i=1}^{r} T_{i.}^2 - \frac{T..^2}{N}$

$$=\frac{1}{3}(91^2+92^2+82^2+87^2+81^2)-\frac{433^2}{15}=33.73$$

3. SSC=$\frac{1}{r}\sum_{j=1}^{C} T_{.j}^2 - \frac{T..^2}{N}$

$=\frac{1}{5}(150^2+130^2+153^2)-\frac{433^2}{15}=62.53$

4. SSE=SST-SSR-SSC=43.47

ANOVA 表

來源	SS	df	MS	F 值
員工	33.73	4	8.4325	F1=1.552
機器	62.53	2	31.265	F2=5.754
誤差	43.47	8	5.43375	
總變異	139.73	14		

(二)

$\begin{cases} H_0：不同機器生產量相同 \\ H_1：不同機器生產量不同 \end{cases}$

$\alpha = 0.05$

拒絕域$C = \{F|F > F_{0.05}(2,8) = 4.459\}$

檢定統計量$F = 5.754 \in C \Rightarrow ReH_0$

結論：有證據顯示不同機器生產量不同

(三)

$\begin{cases} H_0：不同員工生產量相同 \\ H_1：不同員工生產量不同 \end{cases}$

$\alpha = 0.05$

拒絕域$C = \{F|F > F_{0.05}(4,8) = 3.838\}$

檢定統計量$F = 1.552 \notin C \Rightarrow not\ ReH_0$

結論：沒有證據顯示不同員工生產量不同

(四) $X_{ij}=\mu+\alpha_i+\beta_j+\varepsilon_{ij}$,　　$\varepsilon_{ij} \sim N(0,\sigma^2)$

$i=1,2,\ldots,r$　　$j=1,2\ldots,c$

其中α_i為列效果

β_j為行效果

ε_{ij}為隨機誤差

五、乙公司從 2018 年第一季至 2020 年第四季之銷售量如下表所示：

	第一季	第二季	第三季	第四季
2018	1,600	2,500	2,800	2,970
2019	2,100	3,100	3,650	3,350
2020	2,250	3,250	3,840	3,860

假設該公司近年第一至四季的季節指數分別為 74.720、103.978、123.761、97.540。

(一)試計算去除季節因子之銷售量。(5 分)

(二)考慮簡單線性模型 $y_t=\alpha+\beta t+e_t$，試求出去除季節因子之銷售量的趨勢估計式，並在 0.05 顯著水準下檢定斜率是否為零。(10 分)

(三)試預測 2021 年第一至四季之銷售量。(5 分)

1.《考題難易》：★★
2.《解題關鍵》：時間序列及季節指數
3.《命中特區》：吳迪著統計學 P13-4~P13-7

【擬答】：

(一)去除季節因子之銷售量

$$Y_t^* = \frac{Y_t}{S_t} \times 100$$

其中 S_t 為季節指數，則銷售量如下

	第一季	第二季	第三季	第四季
2018	2141	2404	2262	3045
2019	2810	2981	2949	3434
2020	3011	3126	3103	3957

(二)$\widehat{Y_t} = \hat{\alpha} + \hat{\beta}t$

t	1	2	3	4	5	6	7	8	9	10	11	12
Y	2141	2404	2262	3045	2810	2981	2949	3434	3011	3126	3103	3957

$\Sigma t = 78$，$\Sigma t^2 = 650$

$\Sigma Y = 35223$，$\Sigma Y^2 = 106147639$

$\Sigma tY = 245942$

1. $\hat{\beta} = \dfrac{\Sigma tY - \frac{(\Sigma t)(\Sigma Y)}{n}}{\Sigma t^2 - \frac{(\Sigma t)^2}{n}}$

$= \dfrac{245942 - \frac{78\times 35223}{12}}{650 - \frac{78^2}{12}} = 118.829$

2. $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{t}$

$= \dfrac{35223}{12} - 118.829 \times \dfrac{78}{12} = 2162.8615$

$\widehat{Y_t} = 2162.8615 + 118.829t$

$\begin{cases} H_0: \beta = 0 \\ H_1: \beta \neq 0 \end{cases}$

$\alpha = 0.05$

拒絕域 C=$\{t \mid t > 2.228$ 或 $t < -2.228\}$

$SS_t = \Sigma t^2 - \dfrac{(\Sigma t)^2}{n} = 650 - \dfrac{78^2}{12} = 143$

MSE=$\dfrac{SSE}{n-2} = \dfrac{\Sigma Y^2 - \hat{\alpha}\Sigma Y - \hat{\beta}\Sigma tY}{n-2}$

$= \dfrac{106147639 - 2162.8615 \times 35223 - 118.829 \times 245942}{12-2}$

=74012.64675

檢定統計量

$t = \dfrac{\hat{\beta} - \beta}{\sqrt{\frac{MSE}{SSx}}} = \dfrac{118.829 - 0}{\sqrt{\frac{74012.64675}{143}}}$

$=5.223 \in C \Rightarrow ReH_0$

結論：有證據顯示斜率不為零

(三) $\widehat{Y_t} = 2162.8615 + 118.829t$

1. 2021 年第一季 t=13 代入

$\hat{Y} = 2162.8615 + 118.829 \times 13 = 3707.6385$

銷售量為$3707.6385 \times \frac{74.720}{100} = 2770$

2. 2021 年第二季 t=14 代入

$\hat{Y} = 2162.8615 + 118.829 \times 14 = 3826.4675$

銷售量為$3826.4675 \times \frac{103.978}{100} = 3979$

3. 2021 年第三季 t=15 代入

$\hat{Y} = 2162.8615 + 118.829 \times 15 = 3945.2965$

銷售量為$3945.2965 \times \frac{123.761}{100} = 4883$

4. 2021 年第四季 t=16 代入

$\hat{Y} = 2162.8615 + 118.829 \times 16 = 4064.1255$

銷售量為$4064.1255 \times \frac{97.540}{100} = 3964$