《統計學》

試題評析	今年考卷難度屬於中等，但計算量很大，考生容易時間不夠，第四大題的皮爾曼等級相關係數再度出題，可見上課講義中第十三章無母數統計方法慢慢變的很重要。預期平均分數為70分，程度較好之考生應該可以得到80分以上。
考點命中	一、《高點・高上統計學講義》第一回，趙治勳編撰，第三章。 二、《高點・高上統計學講義》第二回，趙治勳編撰，第十一章。 三、《高點・高上迴歸分析講義》第一回，趙治勳編撰，第一章。 四、《高點・高上統計學講義》第三回，趙治勳編撰，第十三章。

一、某班級同學同時參加微積分及管理學兩科目考試，兩科目成績平均數及標準差如下表所示：

	微積分	管理學
平均數	56	68
標準差	14	10

為了比較結果，我們將微積分的分數(x)進行線性轉換（f(x)＝ax＋b，此處a及b為常數），使轉換後的分數之平均數及標準差與管理學分數的平均數及標準差相同。

(一)a及b分別為何？（5分）

(二)相對而言，此班級同學考試原始成績那一個科目分數之離散程度較高？說明理由。（10分）

(三)某同學微積分及管理學原始分數分別為44分及54分，相對而言，此同學那一科考得比較好？為什麼？（10分）

答：

(一)

$$V(aX+b)=a^2V(X)=14^2a^2=10^2 \Rightarrow a=\frac{5}{7} \quad or \quad -\frac{5}{7}$$

$$E(aX+b)=aE(X)+b=56a+b=68$$

當 $a=\frac{5}{7}, b=28$

當 $a=-\frac{5}{7}, b=108$

(二)因為 $C.V._{微積分}=\frac{14}{56}=0.25> C.V._{管理學}=\frac{10}{68}=0.147$

故微積分之相對離散程度較高

(三)因為 $z_{微積分}=\frac{44-56}{14}=-0.857 > z_{管理學}=\frac{54-68}{10}=-1.4$

故微積分考得比較好

二、下列兩個有關於X隨機變數的機率密度函數：

$$H_0: f(X)=8X \qquad 0<X<\frac{1}{2}$$
$$=0 \qquad 其他$$
$$H_a: f(X)=4-8X \qquad 0<X<\frac{1}{2}$$
$$=0 \qquad 其他$$

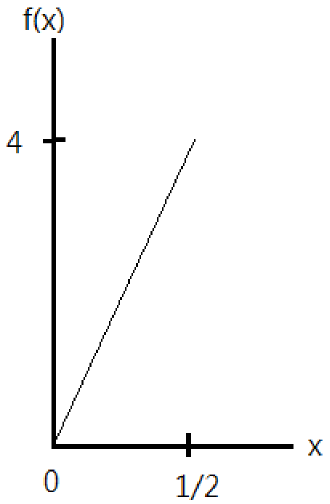
(一)請畫出此兩假說之機率密度函數圖。(5分)

(二)檢定之決策法則如下：某單一X值被觀察到，如果此X值超過某一特定值a $(0<a<\frac{1}{2})$，則接受 H_0；否則接受 H_a。如果已知 H_0 是對的，卻接受 H_a 的機率為 $\frac{1}{16}$，請計算出a值。(10分)

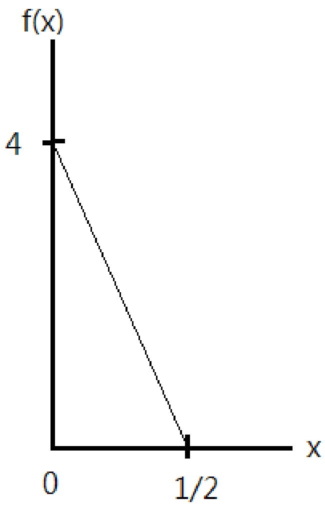
(三)承(二)，計算出此題之統計檢定力。(10分)

答：

(一) H0



H1



(二) $\alpha = P(X \le a \mid H_0) = \int_0^a 8x dx = 4a^2 = \frac{1}{16} \Rightarrow a = \frac{1}{8}$

(三) $power = P(X \le a \mid H_1) = \int_0^{\frac{1}{8}} 4 - 8x dx = \frac{7}{16}$

三、兒童福利機構想了解10歲兒童每週看電視時間(X)與其體重過重程度(Y)間之關係，隨機蒐集6位10歲兒童每週看電視時間及體重過重程度(實際體重減理想體重，以公斤為單位)的資料，如下表：

X	34	18	38	33	29	x_6
Y	6	-7	14	7	8	y_6

以最小平方法估計迴歸線為 $\hat{Y} = -22.609 + 0.937X$。然而，在存檔過程中，分析人員不小心將第6位兒童的資料刪除，試問：

(一)第6位兒童每週看電視時間(x_6)及其體重過重程度(y_6)各為何？(15分)

(二)顯著水準為0.01情況下，檢定此迴歸模型是否具有配適能力？(10分)

答：

(一) $\sum_{i=1}^{5} x_i = 152$，$\sum_{i=1}^{5} y_i = 28$，$\sum_{i=1}^{5} x_i y_i = 1073$

$\hat{\beta}_0 = \frac{28 + y_6}{6} - 0.937 \times \frac{152 + x_6}{6} = -22.609 \quad \Rightarrow y_6 = -21.23 + 0.937 x_6$ -----(1)

$$\hat{\beta}_1=\frac{(1073+x_6y_6)-\frac{(152+x_6)(28+y_6)}{6}}{(4854+x_6^2)-\frac{(152+x_6)^2}{6}}=0.937 \text{ -----(2)}$$

(1)代入(2)

$$\frac{(1073+x_6(-21.23+0.937x_6))-\frac{(152+x_6)(28+(-21.23+0.937x_6))}{6}}{(4854+x_6^2)-\frac{(152+x_6)^2}{6}}=0.937$$

$\Rightarrow 8.274x_6-231.78=0$

$\Rightarrow x_6=28.013$，$y_6=5.018$

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

$\sum_{i=1}^{6}x_i=180.013$，$\sum_{i=1}^{6}x_i^2=5638.728$，$\sum_{i=1}^{6}y_i=33.018$，$\sum_{i=1}^{6}y_i^2=419.18$，$\sum_{i=1}^{6}x_iy_i=1213.569$

$SS_X=237.948$，$SS_Y=237.482$，$SS_{XY}=222.957$

$SSR=\hat{\beta}_1^2SS_X=208.911$，$SSE=SS_Y-\hat{\beta}_1^2SS_X=28.571$

H_0:模型不具配適能力 vs H_1:模型具配適能力

T.S.: $F=\frac{MSR}{MSE}\sim F_{(1,4)}$

R.R.: Reject H_0 at $\alpha=0.01$ if $F^*>F_{(1,4)0.01}=21.2$

$\because F^*=\frac{208.911/1}{28.571/4}=29.248$ $\therefore$ reject H_0

我們有足夠證據去推論模型具配適能力。

四、下列為10台同品牌二手汽車車齡及賣價資料（α=0.05）：

車齡	8	12	9	11	6	7	10	8	6	13
賣價	23	12	24	11	26	30	24	15	24	12

(一)請計算二手汽車車齡與賣價之皮爾森相關係數（Pearson correlation coefficient）並檢定其顯著性。（10分）

(二)請計算二手汽車車齡與賣價之斯皮爾曼等級相關係數（Spearman's rank correlation coefficient）並檢定其顯著性。（10分）

(三)皮爾森相關係數與斯皮爾曼等級相關係數間之差異為何？（5分）

答：令X,Y分別表二手車之車齡與賣價

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

$\sum_{i=1}^{10}x_i=90$，$\sum_{i=1}^{10}x_i^2=864$，$\sum_{i=1}^{10}y_i=201$，$\sum_{i=1}^{10}y_i^2=4467$，$\sum_{i=1}^{10}x_iy_i=1691$

$SS_X=54$，$SS_Y=426.9$，$SS_{XY}=-118$

(一) $r=\frac{SS_{XY}}{\sqrt{SS_X}\sqrt{SS_Y}}=\frac{-118}{\sqrt{54}\sqrt{426.9}}=-0.777$

$H_0:\rho=0$ vs $H_1:\rho\neq0$

T.S.: $T=\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}\sim t_{(8)}$

R.R.: Reject H_0 at $\alpha = 0.05$ if $|T^*| > t_{(8)0.025} = 2.306$

$\because T^* = \dfrac{-0.777\sqrt{10-2}}{\sqrt{1-(-0.777)^2}} = -4.654$ $\therefore$ reject H_0

我們有足夠證據去推論 $\rho \neq 0$ 相關程度具有顯著性。

(二)計算 rank 如下:

車齡	4.5	9	6	8	1.5	3	7	4.5	1.5	10
賣價	5	2.5	7	1	9	10	7	4	7	2.5
d_i	-0.5	6.5	-1	7	-7.5	-7	0	0.5	-5.5	7.5

$r_s = 1 - \dfrac{6\sum d_i^2}{n(n^2-1)} = 1 - \dfrac{6 \times 284.5}{10(10^2-1)} = -0.724$

$H_0 : \rho_s = 0$ vs $H_1 : \rho_s \neq 0$

T.S.: $Z = r_s\sqrt{n-1} \sim N(0,1)$

R.R.: Reject H_0 at $\alpha = 0.05$ if $|Z^*| > z_{0.025} = 1.96$

$\because Z^* = -0.724\sqrt{10-1} = -2.172$ $\therefore$ reject H_0

我們有足夠證據去推論 $\rho_s \neq 0$ 相關程度具有顯著性。

(三)兩者間之差異處在於資料類型的不同，皮爾森相關係數針對屬量變數資料進行計算，但皮爾曼等級相關係數是針對於等級 rank 資料進行計算的。