104 年公務人員普通考試試題

類 科：經建行政、交通技術
科 目：統計學概要

註：本試題可能使用之查表值如下：
1. Z_α（標準常態分配之第 $100(1-\alpha)$ 分位數）：
$Z_{0.05}=1.6449, Z_{0.025}=1.9600$
2. $t_\alpha(n)$（具有自由度 n 之 t 分配之第 $100(1-\alpha)$ 分位數）：
$t_{0.05}(4)=2.1318, t_{0.05}(3)=2.3534, t_{0.05}(2)=2.9200,$
$t_{0.025}(4)=2.7764, t_{0.025}(3)=3.1824, t_{0.025}(2)=4.3027$

一、某家日本電器公司在臺灣設廠，臺灣廠的員工薪資均是以新臺幣支付。總公司為了瞭解臺灣廠員工薪資狀況，做了一些統計分析，考慮下列的統計量及分析。
㈠薪資之變異係數。
㈡薪資與年齡之相關係數。
㈢為了檢定平均薪資 μ 是否為新臺幣 50000 元，針對檢定 H_0：$\mu=50000$，計算 t 檢定統計量。
㈣為了建立員工年齡與薪資之簡單直線迴歸模型，以年齡為自變數，以薪資為應變數，計算迴歸線斜率估計量。
㈤為了瞭解員工薪資是否受教育程度之影響，以員工教育程度為因子作單因子變異數分析，計算 F 檢定統計量。
前一陣子日圓匯率降至低點，引發日本廠員工抱怨。總公司欲比較臺灣廠和日本廠之員工薪資狀況，故將臺灣廠員工薪資統一轉為以日圓計算。試問上述題㈠至題㈤統計量是否會受計算貨幣為新臺幣或日圓不同之影響？請分別就上述題㈠至題㈤之統計量，說明影響是變大、變小、或不變。（題㈠至題㈤每小題 4 分，共 20 分）

【擬答】：
設薪資新臺幣為 X，日圓為 Y，將新臺幣轉為日圓其關係為 $Y=aX$， 且 $a>1$，且年齡為 Z

㈠變異係數 $CV_Y=\dfrac{S_Y}{\bar{Y}}=\dfrac{aS_X}{a\bar{X}}=\dfrac{S_X}{\bar{X}}=CV_X$

所以變異係數不變

㈡ $\gamma_{YZ}=\dfrac{S_{YZ}}{S_Y S_Z}=\dfrac{aS_{XZ}}{aS_X S_Z}=\dfrac{S_{XZ}}{S_X S_Z}=\gamma_{XZ}$

所以薪資與年齡之相關係數不變

㈢ $t=\dfrac{\bar{Y}-\mu_Y}{\frac{S_Y}{\sqrt{n}}}=\dfrac{a\bar{X}-a\mu_X}{\frac{aS_X}{\sqrt{n}}}=\dfrac{\bar{X}-\mu_X}{\frac{S_X}{\sqrt{n}}}$

所以 t 檢定統計量不變

㈣迴歸線斜率估計量 $\hat{\beta}_Y=\gamma_{YZ}\dfrac{S_Y}{S_Z}=\gamma_{XZ}\dfrac{aS_X}{S_Z}=a\hat{\beta}_X$ ，且 $a>1$

所以迴歸線斜率估計量變大

㈤ $$F_Y=\frac{MSTR}{MSE}=\frac{SSTR/K-1}{SSE/N-K}=\frac{\sum_{i=1}^{K}\sum_{j=1}^{n_i}\left(\bar{Y}_{i\cdot}-\bar{Y}_{\cdot\cdot}\right)^2\Big/K-1}{\sum_{i=1}^{K}\sum_{j=1}^{n_i}\left(Y_{ij}-\bar{Y}_{i\cdot}\right)^2\Big/N-K}$$

$$=\frac{\sum_{i=1}^{K}\sum_{j=1}^{n_i}\left(a\bar{X}_{i\cdot}-a\bar{X}_{\cdot\cdot}\right)^2\Big/K-1}{\sum_{i=1}^{K}\sum_{j=1}^{n_i}\left(aX_{ij}-a\bar{X}_{i\cdot}\right)^2\Big/N-K}$$

$$=\frac{a^2\sum_{i=1}^{K}\sum_{j=1}^{n_i}\left(\bar{X}_{i\cdot}-\bar{X}_{\cdot\cdot}\right)^2\Big/K-1}{a^2\sum_{i=1}^{K}\sum_{j=1}^{n_i}\left(X_{ij}-\bar{X}_{i\cdot}\right)^2\Big/N-K}=\frac{MSTR}{MSE}=F_X$$

所以 F 檢定統計量不變

二、兄弟三人依老大、老二、老三，大小順序由大到小，先後輪流投擲三個銅板，看誰先投出剛好兩個正面誰就獲勝。假設兄弟三人約定一定要分出勝負遊戲才停。

㈠試求老大獲勝的機率。

㈡試求老三獲勝的機率。

【擬答】：

投擲三個銅板剛好兩個正面之機率為

$$\frac{C_2^3}{2^3}=\frac{3}{8}$$

㈠老大獲勝之機率

$$P=\frac{3}{8}+(\frac{5}{8})^3(\frac{3}{8})+(\frac{5}{8})^6(\frac{3}{8})+\cdots\cdots=\frac{\frac{3}{8}}{1-(\frac{5}{8})^3}=\frac{64}{129}$$

㈡老三獲勝之機率

$$P=(\frac{5}{8})^2(\frac{3}{8})+(\frac{5}{8})^5(\frac{3}{8})+(\frac{5}{8})^8(\frac{3}{8})\cdots\cdots=\frac{(\frac{5}{8})^2(\frac{3}{8})}{1-(\frac{5}{8})^3}=\frac{25}{129}$$

三、一袋中放入編號為 1、2、3、4 的大小、形狀、重量完全相同的 4 顆球。若以不歸還方式（取出不放回）由此袋中抽出 3 顆球為一樣本，令 S 表所抽 3 球中最大球號和最小球號的差（大減小），令 T 表所抽 3 球中最大的 2 個球號的和。（每小題 10 分，共 20 分）

㈠試求 S 與 T 的相關係數。

㈡試求給定 $T=7$ 之下 S 的條件變異數 $Var(S|T=7)$。

【擬答】：

㈠

(x_1,x_2,x_3)	S	T
$(1,2,3)$	2	5
$(1,2,4)$	3	6
$(1,3,4)$	3	7
$(2,3,4)$	2	7

$$\sum_{i=1}^{n}S_i=2+3+3+2=10$$

$$\sum_{i=1}^{n}S_i^2=2^2+3^2+3^2+2^2=26$$

$$\sum_{i=1}^{n} T_i = 5+6+7+7=25$$

$$\sum_{i=1}^{n} T_i^2 = 5^2+6^2+7^2+7^2=159$$

$$\sum_{i=1}^{n} S_i T_i = 2\times5+3\times6+3\times7+2\times7=63$$

$$\gamma_{ST}=\frac{\sum_{i=1}^{n}S_iT_i-\frac{(\sum_{i=1}^{n}S_i)(\sum_{i=1}^{n}T_i)}{n}}{\sqrt{\sum_{i=1}^{n}S_i^{\,2}-\frac{(\sum_{i=1}^{n}S_i)^2}{n}}\sqrt{\sum_{i=1}^{n}T_i^{\,2}-\frac{(\sum_{i=1}^{n}T_i)^2}{n}}}=\frac{63-\frac{10\times25}{4}}{\sqrt{26-\frac{10^2}{4}}\sqrt{159-\frac{25^2}{4}}}=0.3015$$

㈡

S	2	3
$f(S\|T=7)$	0.5	0.5

$$E(S|T=7)=\sum_{S=-\infty}^{\infty} S\,f(S|T=7)=2\times0.5+3\times0.5=2.5$$

$$E(S^2|T=7)=\sum_{S=-\infty}^{\infty} S^2\,f(S|T=7)=2^2\times0.5+3^2\times0.5=6.5$$

$$\Rightarrow Var(S|T=7)=E(S^2|T=7)-\left[E(S|T=7)\right]^2=6.5-2.5^2=0.25$$

四、令隨機變數 X 具有機率分配 $f(x)=\begin{cases}4x^2e^{-2x}, 0<x<\infty\\ 0, \quad 其他\end{cases}$，且令 $F(x)$ 為 X 之累積機率分配函數(cumulative distribution function)（每小題 10 分，共 20 分）

㈠令隨機變數 $T=F(X)$，試求 T 之機率分配。

㈡令 $F^{-1}(x)$ 為 $F(x)$ 之反函數，且設 U 為具有連續型均等分配 $U(0,1)$ 之隨機變數。令隨機變數 $Y=F^{-1}(U)$，試求 Y 之機率分配。

【擬答】：

㈠令 $r.v.T=F(X)$

$$F_T(t)=P(T\le t)=P(F(X)\le t)=P(X\le F_X^{-1}(t))=F_X(F_X^{-1}(t))=t, 0\le t\le 1$$

$$\Rightarrow f_T(t)=\frac{d}{dt}F_T(t)=1, 0\le t\le 1$$

即 $T\sim U(0,1)$

㈡ $U\sim U(0,1)\Rightarrow f(x)=1, 0\le x\le 1$

又 $Y=F^{-1}(U)\Rightarrow U=F(Y)\sim U(0,1)$

由㈠得 Y 之機率分配為

$$f_Y(y)=\begin{cases}4y^2e^{-2y}\ ,\ 0<y<\infty\\ 0\quad ,\ 其它\end{cases}$$

五、已知兩母體資料相依且服從常態分配。今蒐集兩母體之成對變數資料如下：

母體 I（X_i）	0	4	2	1	3
母體 II（Y_i）	2	4	3	1	5

（每小題 10 分，共 20 分）

㈠試求顯著水準 $\alpha=0.05$ 檢定兩變數之母體相關係數是否為零。

㈡試求出 Y 對 X 的迴歸模型 $Y_i = \beta X_i + \varepsilon_i$ 之參數 β 的最小平方估計值。

【擬答】：

㈠

$$\sum_{i=1}^{n} X_i = 0+4+2+1+3 = 10$$

$$\sum_{i=1}^{n} X_i^2 = 0^2+4^2+2^2+1^2+3^2 = 30$$

$$\sum_{i=1}^{n} Y_i = 2+4+3+1+5 = 15$$

$$\sum_{i=1}^{n} Y^2{}_i = 2^2+4^2+3^2+1^2+5^2 = 55$$

$$\sum_{i=1}^{n} X_i Y_i = 0\times2+4\times4+2\times3+1\times1+3\times5 = 38$$

$$\gamma_{XY} = \frac{\sum_{i=1}^{n} X_i Y_i - \frac{(\sum_{i=1}^{n} X_i)(\sum_{i=1}^{n} Y_i)}{n}}{\sqrt{\sum_{i=1}^{n} X_i^2 - \frac{(\sum_{i=1}^{n} X_i)^2}{n}}\sqrt{\sum_{i=1}^{n} Y_i^2 - \frac{(\sum_{i=1}^{n} Y_i)^2}{n}}} = \frac{38 - \frac{10\times15}{5}}{\sqrt{30-\frac{10^2}{5}}\sqrt{55-\frac{15^2}{5}}} = 0.8$$

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

$\alpha = 0.05$

拒絕域 $C = \{t \mid t > t_{0.025}(3)$ 或 $t < -t_{0.025}(3)\}$

$= \{t \mid t > 3.1824$ 或 $t < -3.1824\}$

檢定統計量

$$t = \frac{\gamma\sqrt{n-2}}{\sqrt{1-\gamma^2}} = \frac{0.8\sqrt{5-2}}{\sqrt{1-0.8^2}} = 2.3094 \notin C$$

$\Rightarrow not\ \mathrm{Re}\, H_0$

結論：沒有證據顯示母體相關係數 $\rho \neq 0$

㈡$Y_i = \beta X_i$，利用最小平方法

$$SSE = \sum_{i=1}^{n}(Y_i - Y_i)^2 = \sum_{i=1}^{n}(Y_i - \beta X_i)^2$$

$$\Rightarrow \frac{\partial SSE}{\partial \beta} = 2\sum_{i=1}^{n}(Y_i - \beta X_i)(-X_i) = 0$$

$$\Rightarrow \sum_{i=1}^{n}(Y_i - \beta X_i)(X_i) = 0$$

$$\Rightarrow \sum_{i=1}^{n} X_i Y_i - \beta\sum_{i=1}^{n} X_i^2 = 0$$

$$\Rightarrow \beta = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2} = \frac{38}{30} = 1.2667$$