104年特種考試地方政府公務人員考試試題

等 別：四等考試
類 科：統計
科 目：統計學概要

一、欲檢測某學院的先修課程的有效性，10 位學生在修課前與修課後參加商學概論考試，其成績如下：

學生	前	後
1	530	670
2	690	770
3	910	1,000
4	700	710
5	450	550
6	820	870
7	820	770
8	630	610
9	540	580
10	440	610

㈠就此資料寫出虛無假設與對立假設。
㈡就㈠之假設，做出相對應的檢定，並做出結論。
$t_{18,0.025}=2.101$ ，$t_{19,0.025}=2.093$ ，$t_{9,0.025}=2.262$

【擬答】：

㈠ $\begin{cases} H_0: \mu_x \le \mu_y \\ H_1: \mu_x > \mu_y \end{cases}$

其中 X 為修課後成績，Y 為修課前成績

㈡

前 (Y)	530	690	910	700	450	820	820	630	540	440
後 (X)	670	770	1000	710	550	870	770	610	580	610
$D=X-Y$	140	80	90	10	100	50	-50	-20	40	170

假設母體為常態，且為相依樣本，利用 t 檢定

令 $\alpha=0.025 \Rightarrow$ 拒絕域 $C=\{t \mid t>t_{0.025}(9)=2.262\}$

$\overline{D}=61 \quad S_D=69.03$

$$t=\frac{\overline{D}-\mu_D}{\frac{S_D}{\sqrt{n}}}=\frac{61-0}{\frac{69.03}{\sqrt{10}}}=2.794 \in C \Rightarrow \mathrm{Re}\,Ho$$

結論：在 $\alpha=0.025$ 條件下，有證據顯示先修課程有效
（註:此題應為無母數統計,但題目附 t 值,所以先假設母體為常態）

二、在某一生產過程中，金屬的硬度受溫度的影響。今蒐集 10 筆資料如下：

溫度	x	100	200	300	400	450	500	550	600	650	700
硬度	y	4	5	4	5	5.5	6	6	7	7.5	7

㈠寫出簡單線性迴歸模型式及假設。
㈡求斜率的最小平方估計值。
㈢計算此模型下的變異數分析表（ANOVA）並解釋其結果（包括假設檢定）。
㈣求 $\beta_0+500\beta_1$ 的估計值及 95%的信賴區間。
$t_{9,0.025}=2.262$ ，$t_{8,0.025}=2.306$ ，$t_{7,0.025}=2.365$

【擬答】：

(一) $Y_i=\beta_0+\beta_1 X_i+\varepsilon_i;\ i=1,2,\ldots\ldots,n$

為了對於迴歸參數做估計，所以迴歸模型有以下的基本假設：

1. 常態性假設：

依變項 Y 的母體為常態分配，即 $Y_i \sim N(\beta_0+\beta_1 X_i,\ \sigma^2)$

2. 線性關係假設：

自變項 X 與依變項 Y 必為限性的關係。若 X 與 Y 具有非線性關係，則必須先將資料作轉換。

3. ε_i 為隨機誤差（Random Error Term），且須滿足以下條件：

(1) ε_i 為隨機誤差，且 $\varepsilon_i \overset{iid}{\sim} N(0,\sigma^2)$，即 ε_i 為常態分配且 $E(\varepsilon_i)=0$，$Var(\varepsilon_i)=\sigma^2$。

(2) $Cov(\varepsilon_i,\varepsilon_j)=0;\ i\neq j$，即任二組隨機誤差無關。

(3) $Cov(\varepsilon_i,X)=0$，即隨機誤差與 X 變項無關。

（參考：吳柏林（2013 年 3 月）、林惠玲、陳正倉（2008 年 9 月））

(二) $\sum_{i=1}^{n}x_i=4450$，$\sum_{i=1}^{n}x_i^2=2327500$，$\sum_{i=1}^{n}y_i=57$，$\sum_{i=1}^{n}y_i^2=338.5$，$\sum_{i=1}^{n}x_iy_i=27350$

$$\hat\beta_1=\frac{\sum_{i=1}^{n}x_iy_i-\frac{\left(\sum_{i=1}^{n}x_i\right)\left(\sum_{i=1}^{n}y_i\right)}{n}}{\sum_{i=1}^{n}x_i^2-\frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n}}=\frac{27350-\frac{4450\times57}{10}}{2327500-\frac{4450^2}{10}}=0.0057$$

(三) $$SST=SS_Y=\sum_{i=1}^{n}y_i^2-\frac{\left(\sum_{i=1}^{n}y_i\right)^2}{n}=388.5-\frac{57^2}{10}=13.6$$

$$SSR=\hat\beta_1^2 SS_x=0.0057^2\times\left[2327500-\frac{4450^2}{10}\right]=11.28$$

$$SSE=SST-SSR=2.32$$

ANOVA 表

來源	SS	df	MS	F 值
迴歸	11.28	1	11.28	F =38.897
誤差	2.32	8	0.29	
總變異	13.6	9		

$$\begin{cases}H_0:\ \beta_1=0\\ H_1:\ \beta_1\neq0\end{cases}$$

$\alpha=0.05\Rightarrow$ 拒絕域 $C=\{t\mid t>2.306$ 或 $t<-2.306\}$

檢定統計量 $t=\sqrt{38.897}=6.237\in C\Rightarrow \text{Re Ho}$

結論：有證據顯示 $\beta_1\neq0$，即金屬的硬度會受溫度的影響。

(四) $$\hat\beta_0=\overline{Y}-\hat\beta_1\overline{X}=\frac{57}{10}-0.0057\times\frac{4450}{10}=3.1635$$

$$\Rightarrow\hat\beta_0+500\hat\beta_1=3.1635+500\times0.0057=6.0135$$

又 $E(\hat\beta_0+500\hat\beta_1)=\beta_0+500\beta_1$

即 $\beta_0+500\beta_1$ 的估計值為 6.0135

所以 $\beta_0+500\beta_1$ 信賴度 95%的信賴區間為

$$\left(\hat{\beta}_0+500\,\hat{\beta}_1-t_{0.025}(8)\sqrt{MSE\left[\frac{1}{n}+\frac{(X-\overline{X})^2}{SS_X}\right]},\hat{\beta}_0+500\,\hat{\beta}_1+t_{0.025}(8)\sqrt{MSE\left[\frac{1}{n}+\frac{(X-\overline{X})^2}{SS_X}\right]}\right)$$

$$\Rightarrow\left(6.0135-2.306\sqrt{0.29\left[\frac{1}{10}+\frac{(500-445)^2}{347250}\right]},6.0135+2.306\sqrt{0.29\left[\frac{1}{10}+\frac{(500-445)^2}{347250}\right]}\right)$$

$\Rightarrow$（5.6041，6.4229）

三、某書店在每 30 分鐘內顧客出現的人次數服從$\lambda=5$的波松（poisson）分配。回答下列問題：

㈠在 60 分鐘內有 12 位顧客光臨的機率？

㈡在兩小時內的顧客平均數，變異數為何？

㈢估計在多長的時間，平均顧客數可達到 100 位。

【擬答】：

㈠$r.v.X$：顧客的人數，$\lambda=\frac{1}{6}$人/分鐘

$X\sim Poi\left(\lambda=\frac{1}{6}\right)$，$t=60\Rightarrow\lambda t=10$

$$\Rightarrow f(x)=\frac{e^{-\lambda t}(\lambda t)^x}{x!}=\frac{e^{-10}10^x}{x!}\quad,\ x=0,1,2,\ldots\ldots,\infty$$

$$\Rightarrow p(x=12)=\frac{e^{-10}10^{12}}{12!}=0.0948$$

㈡$\lambda=\frac{1}{6}$，$t=120\Rightarrow\lambda t=20$

1. $E(X)=\lambda t=20$
2. $Var(X)=\lambda t=20$

㈢設$r.v.Y$：顧客數可達到 100 位的時間

$Y\sim Gamma\left(\alpha=100,\lambda=\frac{1}{6}\right)$

$$\Rightarrow E(Y)=\frac{\alpha}{\lambda}=\frac{100}{\frac{1}{6}}=600\ \text{分鐘}$$

四、某學院院長想要比較會計系和經濟系的研究生在畢業時就取得工作的比例。隨機抽取各 100 位會計系和經濟系的畢業生，其中經濟系畢業生有 52 位，會計系畢業生有 65 位已經在畢業時就獲得就業機會。回答下列問題：

㈠寫出虛無假設與對立假設。

㈡完整地計算㈠的檢定，在顯著水準為 5%下陳述結論。

【擬答】：

㈠$\begin{cases}H_0: P_1=P_2\\ H_1: P_1\neq P_2\end{cases}$

其中P_1為會計系畢業時取得工作的比例

P_2為經濟系畢業時取得工作的比例

㈡因為$n_1=n_2=100$為大樣本，利用 Z 檢定

$\alpha=0.05\Rightarrow$拒絕域$C=\{Z\mid Z>1.96\text{或}Z<-1.96\}$

$\hat{P}_1=\frac{65}{100}=0.65$，$\hat{P}_2=\frac{52}{100}=0.52$，$\hat{P}=\frac{65+52}{100+100}=0.585$

檢定統計量

$$Z=\frac{\left(\hat{P}_1-\hat{P}_2\right)-\left(P_1-P_2\right)}{\sqrt{\frac{\hat{P}\left(1-\hat{P}\right)}{n_1}+\frac{\hat{P}\left(1-\hat{P}\right)}{n_2}}}=\frac{(0.65-0.52)-0}{\sqrt{\frac{0.585\times0.415}{100}+\frac{0.585\times0.415}{100}}}=1.866\notin C\Rightarrow not\quad \text{Re}\ Ho$$

結論：沒有證據顯示會計系和經濟系的研究生在畢業時就取得工作的比例不同。

五、回答下列問題：

㈠何謂統計量？

㈡何謂抽樣分配？

㈢敘明中央極限定理。

【擬答】：

㈠一組隨機樣本$(X_1, X_2, \ldots\ldots\ldots, Xn)$的實數函數，且不包含未知的母數, 稱為樣本統計量。

㈡從母體中抽無限多組樣本數n的隨機樣本可得到無限多個樣本統計量，則其樣本統計量所形成的機率分配稱為抽樣分配。

㈢若母體平均數為μ，變異數為σ^2。從母體中抽出無限多組樣本數為n的隨機樣本。若n趨近於無限大$(n\geq30)$，不論母體為何種分配，則樣本平均數$\overline{X}$之抽樣分配會趨近於常態分配。

（參考：林惠玲，陳正倉著統計學方法與應用三版（上，下）；2008 年 9 月;雙葉書廊）