《統計學》

試題評析	今年度試題以統計學之基礎運算及應用爲主，只要充分準備課本中的內容範圍，必能拿到高分。本卷基本分80分。
考點命中	第一題：《高點統計學講義第一回》，趙治勳老師編撰，頁40。 第二題：《高點統計學講義第一回》，趙治勳老師編撰，頁78。 第三題：1.《高點統計學講義第二回》，趙治勳老師編撰，頁7。 2.《高點統計學講義第二回》，趙治勳老師編撰，頁6。 第四題：《高點統計學講義第一回》，趙治勳老師編撰，頁21。 第五題：《高點統計學講義第二回》，趙治勳老師編撰，頁72。 第六題：《高點統計學講義第二回》，趙治勳老師編撰，頁100～101。

註：$t_{.05}(9)=1.833$, $t_{.025}(9)=2.262$, $t_{.05}(10)=1.812$, $t_{.025}(10)=2.228$, $F_{.05}(2,6)=5.14$
$F_{.05}(2,9)=4.26$, $F_{.025}(2,6)=7.26$, $F_{.025}(2,9)=5.71$, $F_{.05}(3,6)=4.76$, $F_{.05}(3,9)=3.86$
$F_{.025}(3,6)=6.6$, $F_{.025}(3,9)=5.08$
所有假設檢定問題，皆須正確寫出虛無假設、對立假設、檢定統計量、拒絕域、檢定結果與結論。

一、請證明：若事件A與事件B獨立，則事件A與事件B^C也獨立。（10分）

答：

$$P(A\cap B^c)=P(A)-P(A\cap B)\overset{A\perp B}{=}P(A)-P(A)P(B)=P(A)[1-P(B)]=P(A)P(B^c)$$

二、一隨機變數X的動差母函數為$M_X(t)=(0.3+0.7e^t)^n$，請利用動差母函數與動差的關係求算E(X)與V(X)。（10分）

答：

$$E(X)=M_X'(t)|_{t=0}=n(0.3+0.7e^t)^{n-1}(0.7e^t)|_{t=0}=n(0.7)$$

$$E(X^2)=M_X''(t)|_{t=0}=n(n-1)(0.3+0.7e^t)^{n-2}(0.7e^t)^2+n(0.3+0.7e^t)^{n-1}(0.7e^t)|_{t=0}$$

$$=n(n-1)(0.7)^2+n(0.7)$$

$$V(X)=E(X^2)-[E(X)]^2=n(n-1)(0.7)^2+n(0.7)-[n(0.7)]^2$$

$$=n(0.7)(0.3)=n(0.21)$$

三、(一)何謂中央極限定理？試詳述之。（10分）
(二)何謂弱大數法則？試詳述之。（5分）

答：

(一)不論母體分配爲何，只要隨機樣本之樣本數夠大且$E(X_i)=\mu, V(X_i)=\sigma^2<\infty$的話，由樣本線性組合下統計量的抽樣分配會迫近於常態分配。其中樣本數何謂夠大，則要視真正母體分配偏離常態分配之程度。

例：母體：$X\sim(\mu,\sigma^2)$

樣本：$X_1,X_2,\cdots,X_n\overset{iid}{\sim}(\mu,\sigma^2)$

當n夠大時，$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \underset{byC.L.T.}{\sim} N(0,1)$

(二)設 $X_1, X_2, \cdots, X_n$ ($n \ge 1$)爲 iid 之隨機變數序列，且 $E(X_i)=\mu, V(X_i)=\sigma^2<\infty$ ，則 $\overline{X_n}=\dfrac{\sum X_i}{n} \xrightarrow{p} \mu$ ，此稱爲弱大數法則(weak law of large numbers, WLLN)。

證明： $E(\overline{X_n})=E(\dfrac{\sum X_i}{n})=\mu$ ， $V(\overline{X_n})=V(\dfrac{\sum X_i}{n})=\dfrac{\sigma^2}{n}$

由馬可夫不等式，可得

$\lim\limits_{n\to\infty} P(|\overline{X_n}-\mu| \ge \varepsilon) \le \lim\limits_{n\to\infty} \dfrac{\sigma^2/n}{\varepsilon^2}=0$ ，故 $\overline{X_n}=\dfrac{\sum X_i}{n} \xrightarrow{p} \mu$

四、下列是5位學生的身高與體重資料：

身高（x） 172 168 164 170 176 體重（y） 62 54 58 64 62

請問那一個變數的離勢（Dispersion）較大？（15分）

答：

$$C.V._X=\frac{\sigma}{\mu}\times 100\%=\frac{4}{170}\times 100\%=2.353\%$$

$$C.V._Y=\frac{\sigma}{\mu}\times 100\%=\frac{3.578}{60}\times 100\%=5.963\%$$

$\because C.V._X < C.V._Y$

$\therefore$ 變數體重Y之離散趨勢較變數身高X的大。

五、下表是隨機選出的10個家庭的父母與小孩觀賞電視時間（分鐘）：

家庭	1	2	3	4	5	6	7	8	9	10
小孩（X_1）	45	56	73	53	27	34	76	21	54	43
父母（X_2）	23	25	43	26	21	29	32	23	25	21

(一)請問在.05顯著水準下，我們是否可以推論父母觀賞電視的平均時間比小孩短？（15分）

(二)在執行(一)的檢定時，我們需要那些假設條件？（5分）

答：

(一)設 $D_i=X_{1i}-X_{2i}$

母體： $D \sim N(\mu_D, \sigma_D^2)$

樣本： $D_1, D_2, \cdots, D_{10} \overset{iid}{\sim} N(\mu_D, \sigma_D^2)$

點估計： $\overline{D} \sim N(\mu_D, \dfrac{\sigma_D^2}{n})$

$H_0: \mu_D \le 0$ vs $H_0: \mu_D > 0$

T. S.： $T=\dfrac{\overline{D}-0}{\frac{S_D}{\sqrt{10}}}\sim t_{(9)}$

R. R.： Reject H_0 at $\alpha=0.05$ if $T^*>t_{0.05(9)}=1.833$

$\because T^*=\dfrac{21.4-0}{\frac{14.222}{\sqrt{10}}}=4.758$ $\therefore reject H_0$

我們有足夠證據去推論父母觀賞電視平均時間比小孩短。

(二)假設(1) $D\sim N$ (2)隨機樣本

六、為瞭解測量誤差，一位統計學教授請4位同學測量他自己、一位男學生及一位女學生的身高，下表是正確值與測量值的差（以公分計）。

測量誤差

同學	教授（1）	男學生（2）	女學生（3）
1	1.4	1.5	1.3
2	3.1	2.6	2.4
3	2.8	2.1	1.5
4	3.4	3.6	2.9

(一)請問在.05顯著水準下，我們是否可以推論3位被測量者的平均測量誤差間存在差異？（20分）

(二)在執行(一)的檢定時，我們需要那些假設條件？（10分）

答：

假設模型：$y_{ij}=\mu+\alpha_i+\beta_j+\varepsilon_{ij}$ ，$\varepsilon_{ij}\overset{iid}{\sim}N(0,\sigma^2)$

(一)

ANOVA TABLE

Source	SS	d. f.	MS	F
被測量者	0.872	2	0.436	$F_1^*=5.0816$
測量者	5.91	3	1.97	$F_2^*=22.96$
Error	0.515	6	0.0858	
Total	7.297	11		

$H_0:\mu_1=\mu_2=\mu_3$ vs H_1:至少一個$\mu_i\neq\mu_j$

T. S.： $F_1=\dfrac{MSA}{MSE}\sim F_{(2,6)}$

R. R.： Reject H_0 at $\alpha=0.05$ if $F_1^*>F_{0.05(2,6)}=5.14$

$\because F_1^*=5.0816$ $\therefore don't reject H_0$

我們沒有足夠證據去推論3位被測量者的平均測量誤差間存在差異。

(二)假設模型：$y_{ij}=\mu+\alpha_i+\beta_j+\varepsilon_{ij}$ ，$\varepsilon_{ij}\overset{iid}{\sim}N(0,\sigma^2)$， $i=1,2,3$， $j=1,2,3,4$

1.常態性假設

2.變異數齊一性假設
3.獨立性假設
4. $E(\varepsilon_i)=0$
5.模型之正確性