《統計學》

試題評析
本年度考題都屬於基本觀念題，可惜第二、三題語意不清，容易造成考生混淆，高低分主要差在語意的斟酌上。

一、設以2, 4, 6, 8, 10, 12, 14構成一個大小為7之母體，自此母體中以還原方式隨機抽出一個 $n=4$ 的樣本，並求算其平均數 $\bar{x}=4.5$，$s^2=12$：
(一)此母體變異數之值為何？（5分）
(二)$\bar{X}$ 之平均數、變異數、標準誤（standard error）分別為何？（15分）

答：

(一) $\mu=\dfrac{2+4+6+8+10+12+14}{7}=8$

$$\sigma^2=\frac{(2-8)^2+(4-8)^2+\cdots+(14-8)^2}{7}=16$$

(二) $E\bar{X}=EX=8$

$$V(\bar{X})=\frac{\sigma^2}{n}=\frac{16}{4}=4$$

$$\sigma_{\bar{X}}=\sqrt{4}=2$$

【高分閱讀】
1. 參見秦大成老師，《統計學》第一回講義，P.34例題1相同題型(基本實例題)。
2. 參見秦大成老師，《統計學》第四回講義，P.4預備知識。

二、某政府單位想瞭解A城市與B城市之家庭平均收入的差異性，抽樣調查的結果如下：

	樣本數	平均收入（千元）	標準差
A城市	35	401	20
B城市	40	398	22

(一)兩樣本平均差的標準誤為何？（5分）
(二)若想檢定 $H_0：\mu_1=\mu_2$ 時，檢定統計量為何？（5分）
(三)當標準差相等時，若想檢定 $H_0：\mu_1=\mu_2$ 時，檢定統計量為何？檢定統計量之值為何？（10分）
(四)當標準差相等時，在 $H_0：\mu_1=\mu_2$ 時，檢定統計量的分配為何？（5分）
(五)當標準差相等時，若想檢定 $H_0：\mu_1=\mu_2$ vs. $H_1：\mu_1\neq\mu_2$，在顯著水準 $\alpha=0.05$ 下，你如何下結論？（5分）

答：

(一) $V(\bar{X}_A-\bar{X}_B)=V(\bar{X}_A)+V(\bar{X}_B)=20^2+22^2=884$ (假設兩母體indep.)

$$\sigma_{\bar{X}_A-\bar{X}_B}=\sqrt{884}=29.7321$$

(二) $n_A=35\geq 30$，$n_B=40\geq 30$ 採大樣本，查Z表

檢定統計量 $Z=\dfrac{\bar{X}_A-\bar{X}_B}{\sqrt{\dfrac{S_A^2}{n_A}+\dfrac{S_B^2}{n_B}}}\sim N(0,1)$

(三)本題採大樣本：根據Slutsky定理

sample variance 各別用 S_A^2，S_B^2 ，與用 S_p^2

檢定統計量 $Z=\dfrac{\bar{X}_A-\bar{X}_B}{\sqrt{\dfrac{S_A^2}{n_A}+\dfrac{S_B^2}{n_B}}}\xrightarrow[n_A,\,n_B\to\infty]{\mathcal{L}}N(0,1)$

$$Z=\frac{\bar{X}_A-\bar{X}_B}{\sqrt{\dfrac{S_p^2}{n_A}+\dfrac{S_p^2}{n_B}}}\xrightarrow[n_A,\,n_B\to\infty]{\mathcal{L}}N(0,1)$$

∵ 題意：$\sigma_A=\sigma_B$ ∴ **用**S_p^2**取代**σ^2

(1) $S_p^2=\dfrac{(n_A-1)S_A^2+(n_B-1)S_B^2}{n_A+n_{B-2}}$

檢定統計量 $Z=\dfrac{\bar{X}_A-\bar{X}_B}{\sqrt{\dfrac{S_p^2}{n_A}+\dfrac{S_p^2}{n_B}}}\sim N(0,1)$

(2) $S_p^2=\dfrac{(35-1)\times20^2+(40-1)\times22^2}{35+40-2}=444.8767$

統計值 $Z=\dfrac{401-398}{\sqrt{444.8767(\frac{1}{35}+\frac{1}{40})}}=\dfrac{3}{4.8819}=0.6145$

(四) $N(0,1)$

(五) ① $\begin{cases}H_0: \mu_1=\mu_2\\ H_1: \mu_1\neq\mu_2\end{cases}$

② $C=\{Z\,\big|\,|Z|>Z_{\frac{0.05}{2}}=1.96\}$

③ $Z=0.16145\notin C$

∴ Not reject，μ_1**與**μ_2**無顯著差異**

【高分閱讀】
本題未給條件：兩家庭收入為獨立Normal母體
若條件給足：採用pooled-t-test。因條件不足，而且 $n_1\geq30, n_2\geq30$ ，∴採用C.L.T. ：Z-test。
(一)(二)(三)(四)：參見秦大成老師，《統計學》第五回講義P.61(基本觀念)完全含蓋。
(五)：參見秦大成老師，《統計學》第五回講義，P.62例題1相同題型。

三、具有自由度a和b的F分配的$(1-\alpha)100^{th}$百分位數以$F_\alpha(a,b)$表示：

（一）百分位數$F_\alpha(a,b)$與$F_{1-\alpha}(b,a)$的關係式為何？（5分）

（二）若兩組樣本數分別為n_1與n_2，樣本變異數分別為S_1^2與S_2^2，則兩個常態母體樣本變異數比S_1^2/S_2^2的實際分配為何？（5分）

（三）若兩組樣本數分別為$n_1=21$與$n_2=25$，$S_1^2=96$與$S_2^2=100$，則兩個常態母體變異數比σ_1^2/σ_2^2的90%信賴區間為何？（10分）

（四）若想檢定H_0：$\sigma_1=\sigma_2$ vs. H_0：$\sigma_1\neq\sigma_2$，在$\alpha=0.1$下，你如何下結論？（5分）

答：

(一) $F_{1-\alpha}(b,a)=\frac{1}{F_\alpha(a,b)}$

(二)(1)當μ_1,μ_2已知

則 $\frac{S_1^2}{S_2^2}\times\frac{\sigma_2^2}{\sigma_1^2}\sim F(n_1,n_2)$ $\quad\therefore \frac{S_1^2}{S_2^2}$ under $\sigma_1^2=\sigma_2^2$ $\sim F(n_1,n_2)$

(2) 當μ_1,μ_2未知

則 $\frac{S_1^2}{S_2^2}\times\frac{\sigma_2^2}{\sigma_1^2}\sim F(n_1-1,n_2-1)$ $\quad\therefore \frac{S_1^2}{S_2^2}$ under $\sigma_1^2=\sigma_2^2$ $\sim F(n_1-1,n_2-1)$

(三) $\frac{\sigma_1^2}{\sigma_2^2}$ 之90% C.I. $=\left(\frac{S_1^2}{S_2^2}\times\frac{1}{F_{\frac{\alpha}{2}}(n_1-1,n_2-1)},\ \frac{S_1^2}{S_2^2}\times\frac{1}{F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)}\right)$

$$=\left(\frac{96}{100}\times\frac{1}{2.0267},\ \frac{96}{100}\times 2.0825\right)$$

$$=(0.4737,1.9992)$$

(四)根據(三)

$\because$ C.I. **包含** 1

$\therefore$ Not reject H_0，**無充分證據顯示**$\sigma_1^2\neq\sigma_2^2$

【高分閱讀】

(一)(二)：參見秦大成老師，《統計學》第三回P.60完全相同，若(二)沒有出錯：解法 $\frac{S_1^2}{S_2^2}\times\frac{\sigma_2^2}{\sigma_1^2}\sim F(n_1,n_2)$

再經由變數變換，求出 $\frac{S_1^2}{S_2^2}$ 之分配，不過才5分，似乎不太合理。

(三)：參見秦大成老師，《統計學》與第四回，P.85例題1相同題型。

(四)：參見秦大成老師，《統計學》與第五回，P.35例題1相同題型。

四、在單因子變異數分析（One-way ANOVA）問題中，包含有3個處理（treatment），而每個處理分別有18、10、15個觀測值。已知MSE＝6與F＝3：

（一）請寫出完整的變異數分析表。（12分）

（二）請說明進行單因子變異數分析需要之基本假設為何？（8分）

（三）若想檢定3個處理效果是否一致，在α＝0.05下，你如何下結論？（5分）

答：

$SSE = (N-k)\cdot MSE = (43-3)\times 6 = 240$

$MSTR = F \times MSE = 3 \times 6 = 18$

$SSTR = (k-1)\times MSTR = (3-1)\times 18 = 36$

(一) ANOVA table

Source	SS	df	MS	F − value
Treatment	36	2	18	3
Error	240	40	6	
Total	276	42		

(二) (1)每個母體都是 Normal 母體

(2) $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$

(3)每個樣本都是隨機樣本，而且母體之間彼此獨立

(三) ① $\begin{cases} H_0: \mu_1 = \mu_2 = \mu_3 \\ H_1: \text{不全等} \end{cases}$

② $C = \{F \mid F > F_{0.05}(3-1, 43-3) = 3.2317\}$

③ $F = 3 \notin C$

【高分閱讀】

送分題，參見秦大成老師，《統計學》第六回講義，P. 11例題1相同題型。