《統計學》

試題評析	今年命題內容包括：基本的機率論、成對母體t檢定、點估計。總共只有三大題，難度中等偏易，一般考生應該可以拿到70分，預估上榜的水準大概90分左右。
考點命中	第一題：《高點統計學講義》第一回，秦大成編撰，頁37《例題2》。 第二題：《高點統計學講義》第五回，秦大成編撰，頁48。 第三題：(一)《高點統計學講義》第四回，秦大成編撰，頁10《例題3》； (二)《高點統計學講義》第四回，秦大成編撰，頁16《例題10》； (三)(四)(五)《高點統計學講義》第四回，秦大成編撰，頁33《例題11》。

一、某一汽車經銷商計算其每賣一部新車的獲利（以十萬元為單位）為 $Y=X^2$，若 X 的機率密度函數如下：

$$g(x)=\begin{cases} c(1-x) & 0<x<1 \\ 0 & \text{其他} \end{cases}$$

(一)請計算求得 c 值。（5分）

(二)推導求得 Y 之機率密度函數 $f(y)$。（10分）

(三)請計算該經銷商在賣下一部新車時，其獲利小於一萬元的機率。（5分）

答：

(一) $\int_0^1 c(1-x)\,dx = c\left(1-\frac{1}{2}\right)=1 \quad \therefore c=2$

(二) $y=x^2 \Rightarrow x=\pm\sqrt{y}$ (取+)，$\left|\frac{dx}{dy}\right|=\frac{1}{2\sqrt{y}}$

$$f(y)=f_X\left(\sqrt{y}\right)\cdot\left|\frac{dx}{dy}\right|$$

$$=2\left(1-\sqrt{y}\right)\cdot\frac{1}{2\sqrt{y}}=\frac{1}{\sqrt{y}}-1 \quad ,\ 0<y<1$$

(三) $P(Y<0.1)=\int_0^{0.1}\frac{1}{\sqrt{y}}-1\,dy=2\sqrt{y}-y\Big|_0^{0.1}=0.5325$

二、若考慮成對比較試驗（paired-difference experiment），每一成對觀測值 j（$j=1,2,\ldots,n$），分別接受處理（treatment）i，$i=1$、2。以下模型描述此一試驗結果：

$$Y_{ij}=\mu_i+P_j+\varepsilon_{ij}$$

其中 μ_i 代表處理 i 的期望反應，P_j 為第 j 個成對實驗單位的隨機效應，ε_{ij} 為隨機誤差。對所有 $i=1,2$，$j=1,2,\ldots,n$，假設 P_j 為獨立的常態分配，其均數為 $E(P_j)=0$，變異數 $Var(P_j)=\sigma_P^2$；ε_{ij} 為獨立的常態分配，$E(\varepsilon_{ij})=0$，$Var(\varepsilon_{ij})=\sigma^2$；且 P_j 與 ε_{ij} 互相獨立。

(一)若 $\bar{Y}_i$ 為處理 i（$i=1$、2）的 n 個觀測值之均數，請推導求得 $E(\bar{Y}_i)$。（5分）

(二)請推導求得 $Var(\bar{Y}_i)$。（5分）

(三)若 $d_j=Y_{1j}-Y_{2j}$，$j=1,2,\ldots,n$，且 $\bar{d}$ 為其均數，s_d 為標準差。推導求得 $E(\bar{d})$。（5分）

(四)推導求得 $Var(\bar{d})$。（7分）

(五)在虛無假設 $H_0:\mu_1-\mu_2=0$ 為真的情況下，請證明 $\dfrac{\bar{d}\sqrt{n}}{s_d}$ 服從 t 分配；並說明其自由度。（13

分）

（六）若將原前述模型改為

$$Y_{ij} = \mu_i + P_{ij} + \varepsilon_{ij}$$

假設 P_{ij} 為獨立的常態分配，其均數為 $E(P_{ij}) = 0$，變異數 $Var(P_{ij}) = \sigma_P^2$；其餘符號之表達及假設與前述相同。推導求得此模型下的 $Var(\bar{d})$；並比較此結果與題(四)的差異。（10分）

答：

(一) $E[\bar{Y}_i] = E\left[\frac{\sum_{j=1}^{n} Y_{ij}}{n}\right] = E\left[\frac{\sum_{j=1}^{n}(\mu_i + P_j + \varepsilon_{ij})}{n}\right] = \frac{n(\mu_i+0+0)}{n} = \mu_i$ ，$i = 1, 2$

(二) $Var[\bar{Y}_i] = Var\left[\frac{\sum_{j=1}^{n} Y_{ij}}{n}\right] = Var\left[\frac{\sum_{j=1}^{n}(\mu_i + P_j + \varepsilon_{ij})}{n}\right] = \frac{n(\sigma_P^2+\sigma^2)}{n^2} = \frac{\sigma_P^2+\sigma^2}{n}$ ，$i = 1, 2$

(三) $\bar{d} = \frac{\sum_{j=1}^{n} d_j}{n} = \frac{\sum_{j=1}^{n}(Y_{1j} - Y_{2j})}{n} = \bar{Y}_1 - \bar{Y}_2$

$E[\bar{d}] = E[\bar{Y}_1 - \bar{Y}_2] = \mu_1 - \mu_2$

(四) $Var(d_j) = Var(Y_{1j} - Y_{2j}) = Var[(\mu_1 - \mu_2) + (\varepsilon_{1j} - \varepsilon_{2j})]$

$= Var(\varepsilon_{1j}) + Var(\varepsilon_{2j}) = 2\sigma^2$

$Var[\bar{d}] = Var\left[\frac{\sum_{j=1}^{n} d_j}{n}\right] = \frac{2\sigma^2}{n}$

(五) $\bar{d} = \bar{Y}_1 - \bar{Y}_2 \sim N\left(\mu_1 - \mu_2, \frac{2\sigma^2}{n}\right) \overset{令}{=\!=} N(\mu_1 - \mu_2, \sigma_d^2)$

$$\text{indep.}\begin{cases} Z = \dfrac{\bar{d} - (\mu_1 - \mu_2)}{\sigma_d/\sqrt{n}} \sim N(0,1) \\ \chi^2 = \dfrac{(n-1)S_d^2}{\sigma_d^2} \sim \chi^2(n-1) \end{cases}$$

$$T = \frac{Z}{\sqrt{\chi^2/(n-1)}} = \frac{\frac{\bar{d}-(\mu_1-\mu_2)}{\sigma_d/\sqrt{n}}}{\sqrt{\frac{(n-1)S_d^2}{\sigma_d^2}\Big/(n-1)}} = \frac{[\bar{d}-(\mu_1-\mu_2)]\sqrt{n}}{S_d}$$

$$\overset{H_0真}{=\!=} \frac{\bar{d}\sqrt{n}}{S_d} \sim t(n-1)$$

(六)(1) $Var(d_j) = Var(Y_{1j} - Y_{2j}) = Var[(\mu_1 + P_{1j} + \varepsilon_{1j}) - (\mu_2 + P_{2j} + \varepsilon_{2j})]$

$= Var[(P_{1j} - P_{2j}) + (\varepsilon_{1j} - \varepsilon_{2j})]$

$= Var[(P_{1j} - P_{2j})] + Var[(\varepsilon_{1j} - \varepsilon_{2j})] = 2(\sigma_P^2 + \sigma^2)$

$\therefore Var[\bar{d}] = Var\left[\frac{\sum_{j=1}^{n} d_j}{n}\right] = \frac{2(\sigma_P^2 + \sigma^2)}{n}$

(2)差異值 $= \frac{2\sigma_P^2}{n}$

三、若 $Y_1, Y_2, \ldots, Y_n$ 為互相獨立且其機率密度函數如下：

$$f(y;\theta) = \begin{cases} \frac{1}{\theta} e^{-y/\theta} & y > 0 \\ 0 & 其他 \end{cases}$$

(一)推導求得θ的最大概似估計(maximum likelihood estimator)。(10分)
(二)推導求得$P(Y \le 2)$的最大概似估計。(5分)
(三)若$Y_{(1)} = \min(Y_1, Y_2, \ldots, Y_n)$為最小順序統計量,推導求得$Y_{(1)}$之分配。(10分)
(四)證明$\hat{\theta}_1 = nY_{(1)}$為$\theta$的不偏估計(unbiased estimator)。(4分)
(五)推導求得$MSE(\hat{\theta}_1)$(mean square error of $\hat{\theta}_1$)。(6分)

答:

(一)$L(\theta) = (\frac{1}{\theta})^n e^{-\frac{n\bar{y}}{\theta}} \Rightarrow \ln L(\theta) = -n\ln\theta - \frac{n\bar{y}}{\theta}$

$\frac{d\ln L}{d\theta} = -\frac{n}{\theta} + \frac{n\bar{y}}{\theta^2} = 0 \Rightarrow \theta = \bar{y}$

$\because \left.\frac{d^2\ln L}{d\theta^2}\right|_{\theta=\bar{y}} < 0 \qquad \therefore \hat{\theta} = \bar{Y}$ 為 θ 之 MLE

(二)$P(Y \le 2) = F_Y(2) = 1 - e^{-\frac{2}{\theta}}$,且 $g(\theta) = 1 - e^{-\frac{2}{\theta}}$為 $1-1$ 函數

$\therefore$ 根據 MLE 不變性:$g(\bar{Y}) = 1 - e^{-\frac{2}{\bar{Y}}}$為 $P(Y \le 2)$之 MLE

(三)$f_{Y_{(1)}}(y) = n[1-F(y)]^{n-1}f(y) = n\left(e^{-\frac{y}{\theta}}\right)^{n-1}(\frac{1}{\theta}e^{-\frac{y}{\theta}})$

$= \frac{n}{\theta}e^{-\frac{ny}{\theta}}$, $y > 0$

$\therefore Y_{(1)} \sim \text{Exp}(\lambda = \frac{n}{\theta})$

(四)$E[\hat{\theta}_1] = E[nY_{(1)}] = n \cdot \frac{1}{\lambda} = n \cdot \frac{\theta}{n} = \theta$ (unbiased)

(五)$MSE(\hat{\theta}_1) = Var(\hat{\theta}_1) + (E\hat{\theta}_1 - \theta)^2 = Var(\hat{\theta}_1)$

$= Var(nY_{(1)}) = n^2 \cdot \frac{1}{\lambda^2} = \theta^2$