《統計學概要》

一、目前核融合技術的重大突破，讓未來核融合發電可望成真。已知國內某大學的實驗室有三座核融合反應爐，令變數T_i為第i座反應爐的核融合實際反應時間與目標反應時間之間的差異，$i=1,2,3$。假設變數T_1, T_2, T_3彼此相互獨立，且都服從平均數為0，變異數為4之常態分配。

(一) 求出機率 $P[T_1^2+T_2^2\le 2]$。（10分）

(二) 令變數 $S=\sqrt{T_1^2+T_2^2+T_3^2}$，請求出變數S之機率密度函數f(s)。（10分）

(三) 令變數 $W=\dfrac{T_1^2}{T_1^2+T_2^2}$，請求出變數W之機率密度函數f(w)。（10分）

(四) 求出題(三)之變數W的期望值E(W)。（10分）

(五) 求出機率 $P[\mathrm{Min}\{\mathrm{Max}\{T_1,T_2\},T_3\}<0]$，此處Max{a,b}代表取a,b之最大值，Min{a,b}代表取a,b之最小值。（10分）

(六) 假設每一座反應爐每次點火成功的機率為0.2，且假設三座反應爐點火成功與否彼此相互獨立。令X_i為第i座反應爐直到第一次點火成功前，所需的點火（失敗）次數，$i=1,2,3$。請求出機率 $P[X_1\ge X_2]$。（10分）

試題評析	本題著重在機率論之範圍，考生需要對於分配間之關係非常熟悉才比較容易作答，難度算中上程度，考驗考生之數學能力。
考點命中	1.《高點・高上統計學講義》第一回，趙治勳編撰，第七章第二、三節，常態/卡方/指數分配之關係。 2.《高點・高上統計學講義》第一回，趙治勳編撰，第七章第四節，伽瑪/貝它分配之關係。 3.《高點・高上統計學講義》第一回，趙治勳編撰，第五章第四節，MIN/MAX之分配。

答：

$$T_i\overset{iid}{\sim}N(0,4)\qquad \frac{T_i}{\sqrt{4}}=\frac{T_i}{2}\overset{iid}{\sim}N(0,4)\qquad \frac{T_i^2}{4}\overset{iid}{\sim}\chi^2_{(1)}=Gamma(\alpha=\frac{1}{2},\lambda=\frac{1}{2})\ ,\ i=1,2,3$$

(一) 令 $Y=\dfrac{T_1^2}{4}+\dfrac{T_2^2}{4}=\dfrac{T_1^2+T_2^2}{4}\sim\chi^2_{(2)}=Gamma(\alpha=\dfrac{2}{2}=1,\lambda=\dfrac{1}{2})=Exp(\lambda=\dfrac{1}{2})$

$$P(T_1^2+T_2^2\le 2)=P(Y\le\frac{1}{2})=1-e^{-\frac{1}{2}(\frac{1}{2})}=1-e^{-\frac{1}{4}}$$

(二) 令 $H=\dfrac{T_1^2}{4}+\dfrac{T_2^2}{4}+\dfrac{T_3^2}{4}=\dfrac{T_1^2+T_2^2+T_3^2}{4}\sim\chi^2_{(3)}=Gamma(\alpha=\dfrac{3}{2},\lambda=\dfrac{1}{2})$

$$F_S(s)=P(S\le s)=P(\sqrt{T_1^2+T_2^2+T_3^2}\le s)=P(H\le\frac{s^2}{4})=\int_0^{\frac{s^2}{4}}\frac{(\frac{1}{2})^{\frac{3}{2}}}{\Gamma(\frac{3}{2})}h^{\frac{3}{2}-1}e^{-\frac{1}{2}h}dh$$

$$f_S(s)=\frac{dF_S(s)}{ds}=\frac{(\frac{1}{2})^{\frac{3}{2}}}{\Gamma(\frac{3}{2})}(\frac{s^2}{4})^{\frac{3}{2}-1}e^{-\frac{1}{2}(\frac{s^2}{4})}\times\frac{s}{2}=\frac{1}{4\sqrt{2\pi}}s^2e^{-\frac{s^2}{8}}\qquad ,0<s$$

(三)

$$\boxed{\begin{array}{l} \text{已知} \quad X_1 \sim Gamma(\alpha,\lambda) \perp X_2 \sim Gamma(\beta,\lambda) \\ \qquad \dfrac{X_1}{X_1+X_2} \sim Beta(\alpha,\beta) \end{array}}$$

$$\frac{T_i^2}{4} \overset{iid}{\sim} \chi^2_{(1)} = Gamma(\alpha=\frac{1}{2},\lambda=\frac{1}{2}) \ , \ i=1,2,3$$

$$W=\frac{T_1^2}{T_1^2+T_2^2}=\frac{\frac{T_1^2}{4}}{\frac{T_1^2}{4}+\frac{T_2^2}{4}} \sim Beta(\frac{1}{2},\frac{1}{2})$$

$$f_W(w)=\frac{\Gamma(\frac{1}{2}+\frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2})} w^{\frac{1}{2}-1}(1-w)^{\frac{1}{2}-1}=\frac{1}{\pi}\frac{1}{\sqrt{w(1-w)}} \quad ,0<w<1$$

(四) $\because W=\dfrac{T_1^2}{T_1^2+T_2^2} \sim Beta(\dfrac{1}{2},\dfrac{1}{2})$

$$\therefore E(W)=\frac{\frac{1}{2}}{\frac{1}{2}+\frac{1}{2}}=\frac{1}{2}$$

(五) $P[Min\{Max\{T_1,T_2\},T_3\}<0]$

$$\begin{aligned}
&=1-P[Min\{Max\{T_1,T_2\},T_3\}\geq 0]\\
&=1-P[Max\{T_1,T_2\}\geq 0]\times P[T_3\geq 0]\\
&=1-\{1-P[Max\{T_1,T_2\}<0]\}\times P[T_3\geq 0]\\
&=1-\{1-P(T_1<0)\times P(T_2<0)\}\times P(T_3\geq 0)\\
&=1-P(T_3\geq 0)+P(T_1<0)\times P(T_2<0)\times P(T_3\geq 0)\\
&=1-\frac{1}{2}+\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}=\frac{5}{8}
\end{aligned}$$

其中 $P(T_1<0)=P(T_2<0)=P(Z<0)=\dfrac{1}{2}$

$$P(T_3\geq 0)=P(Z\geq 0)=\frac{1}{2}$$

(六) 令 $Y_i=X_i+1 \overset{iid}{\sim} Geo(p=0.2) \quad ,i=1,2$

$$P(X_1\geq X_2)=P(Y_1\geq Y_2)=1-P(Y_1<Y_2)=1-\frac{1-P(Y_1=Y_2)}{2}$$

$$=1-\frac{1-[0.2^2+(0.2\times 0.8)^2+(0.2\times 0.8^2)^2+(0.2\times 0.8^3)^2+\cdots\cdots]}{2}$$

$$=1-\frac{1-[0.2^2+0.2^2\times0.8^2+0.2^2\times0.8^4+0.2^2\times0.8^6+\cdots\cdots]}{2}$$

$$=1-\frac{1-\frac{0.2^2}{1-0.8^2}}{2}=\frac{5}{9}$$

二、ChatGPT的問世帶動了AI商機的蓬勃發展，也促成了市場對GPU需求量的急遽增加。已知國內某生產GPU的工廠，所生產的GPU之壽命服從變異數為θ之指數分配。今由此公司之生產線隨機抽檢n筆GPU樣本並測驗其壽命，令$Y_1,Y_2,...,Y_n$表此n筆相互獨立樣本之觀測值。令$\text{Min}\{Y_1,Y_2,...,Y_n\}$代表取$Y_1,Y_2,...,Y_n$之最小值，$\text{Max}\{Y_1,Y_2,...,Y_n\}$代表取$Y_1,Y_2,...,Y_n$之最大值。

(一)求出此GPU壽命分配之中位數的均勻最小變異不偏估計量（uniformly minimum variance unbiased estimator）。（10分）

(二)求出機率$P[\text{Min}\{Y_1,Y_2,...,Y_n\}<1]$之最大概似估計量（maximum likelihood estimator）。（10分）

(三)求出機率$P[\text{Min}\{Y_1,Y_2,...,Y_n\}>1,\text{Max}\{Y_1,Y_2,...,Y_n\}>2]$。（10分）

(四)令F(y)為變數Y_i之累積分配函數（cumulative distribution function）。請求出機率$P[\text{Min}\{F(Y_1),F(Y_2),...,F(Y_n)\}>\frac{1}{4},\text{Max}\{F(Y_1),F(Y_2),...,F(Y_n)\}<\frac{1}{2}]$。（10分）

試題評析	本題有考到數統之UMVUE，還有MLE之不變性與機率積分轉換定理，在普考中難度算很高的。
考點命中	1.《高點・高上統計學講義》第二回，趙治勳編撰，第十章第三節，MLE之不變性。 2.《高點・高上統計學講義》第一回，趙治勳編撰，第七章第一節，機率積分轉換定理。 3.《迴歸分析申論題完全制霸》，高點文化出版，趙治勳編著，第一篇第五章，UMVUE。

答：

母體：$Y\sim Exp(\lambda=\frac{1}{\sqrt{\theta}})$　其中 $V(Y)=\frac{1}{(1/\sqrt{\theta})^2}=\theta$

樣本：$Y_1,Y_2,\ldots,Y_n\overset{iid}{\sim}Exp(\lambda=\frac{1}{\sqrt{\theta}})$

(一) 令η為Y之中位數

$$F_Y(\eta)=P(Y\le\eta)=1-e^{-\frac{1}{\sqrt{\theta}}\eta}=\frac{1}{2}\Rightarrow\eta=-\ln(\frac{1}{2})\sqrt{\theta}$$

$\sum_{i=1}^{n}Y_i$為θ之c.s.s.　，　$\sum_{i=1}^{n}Y_i\sim\text{Gamma}(\alpha=n,\lambda=\frac{1}{\theta})$

已知　$X\sim Gamma(\alpha,\lambda)$

$$E(X^c)=\frac{\Gamma(\alpha+c)}{\Gamma(\alpha)\lambda^c},c>-\alpha$$

$$E(\sqrt{\sum_{i=1}^{n}Y_i})=E[\left(\sum_{i=1}^{n}Y_i\right)^{\frac{1}{2}}]=\frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)(\frac{1}{\theta})^{\frac{1}{2}}}=\frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)}\sqrt{\theta}$$

$$\Rightarrow E(-\ln(\frac{1}{2})\frac{\Gamma(n)}{\Gamma(n+\frac{1}{2})}\sqrt{\sum_{i=1}^{n}Y_i}) = -\ln(\frac{1}{2})\sqrt{\theta} = \eta$$

由Lehmann Scheffe, $-\ln(\frac{1}{2})\frac{\Gamma(n)}{\Gamma(n+\frac{1}{2})}\sqrt{\sum_{i=1}^{n}Y_i}$ 為 η 之UMVUE

(二) 令 $Y_{(1)} = Min\{Y_1, Y_2, \ldots, Y_n\}$

$$g_{Y_{(1)}}(y_{(1)}) = \frac{n!}{0!(n-1)!}[1-(1-e^{-\frac{1}{\sqrt{\theta}}y_{(1)}})]^{n-1}\frac{1}{\sqrt{\theta}}e^{-\frac{1}{\sqrt{\theta}}y_{(1)}} = \frac{n}{\sqrt{\theta}}e^{-\frac{n}{\sqrt{\theta}}y_{(1)}} \qquad , 0 < y_{(1)}$$

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

$$P[Min\{Y_1, Y_2, \ldots, Y_n\} < 1] = P(Y_{(1)} < 1) = 1 - e^{-\frac{n}{\sqrt{\theta}}}$$

已知 $\hat{\theta} = \overline{Y}^2$ 為 θ 之MLE

由MLE不變性， $1 - e^{-\frac{n}{\sqrt{\overline{Y}^2}}} = 1 - e^{-\frac{n}{\overline{Y}}}$ 為 $P[Min\{Y_1, Y_2, \ldots, Y_n\} < 1]$ 之MLE

(三) 令 $Y_{(1)} = Min\{Y_1, Y_2, \ldots, Y_n\}$ ， $Y_{(n)} = Max\{Y_1, Y_2, \ldots, Y_n\}$

$$P[Min\{Y_1, Y_2, \ldots, Y_n\} > 1 \cap Max\{Y_1, Y_2, \ldots, Y_n\} > 2] = P[Y_{(1)} > 1 \cap Y_{(n)} > 2]$$

$$\overset{\text{差集公式}}{=} P[Y_{(1)} > 1] - P[Y_{(1)} > 1 \cap Y_{(n)} < 2]$$

$$= \{1 - P[Y_{(1)} \le 1]\} - P[1 < Y_1 < 2 \cap 1 < Y_2 < 2 \cap \cdots \cap 1 < Y_n < 2]$$

$$= \{1 - [1 - e^{-\frac{n}{\sqrt{\theta}}}]\} - \{P(1 < Y < 2)\}^n$$

$$= e^{-\frac{n}{\sqrt{\theta}}} - [e^{-\frac{2}{\sqrt{\theta}}} - e^{-\frac{1}{\sqrt{\theta}}}]^n$$

(四) 由機率積分轉換定理，$U_i = F(Y_i) \overset{iid}{\sim} U(0,1)$

$$P[Min\{F(Y_1), F(Y_2), \ldots, F(Y_n)\} > \frac{1}{4} \cap Max\{F(Y_1), F(Y_2), \ldots, F(Y_n)\} < \frac{1}{2}]$$

$$= P[Min\{U_1, U_2, \ldots, U_n\} > \frac{1}{4} \cap Max\{U_1, U_2, \ldots, U_n\} < \frac{1}{2}]$$

$$= P[\frac{1}{4} < U_1 < \frac{1}{2} \cap \frac{1}{4} < U_2 < \frac{1}{2} \cap \cdots \cap \frac{1}{4} < U_n < \frac{1}{2}]$$

$$= \{P[\frac{1}{4} < U_1 < \frac{1}{2}]\}^n = (\frac{1}{4})^n$$