《統計學》

一、一盒中放有編號1，2，3，4，5之大小、形狀、重量完全相同的5顆球。今由此盒隨機抽一球，若抽出為n號球，則投擲n個正、反兩面出現機率均等之均勻的銅板。（每小題10分，共20分）

(一)試求最後投擲之銅板中剛好有3個正面的機率。

(二)試求投擲之銅板出現正面次數的期望值。

試題評析	本題是條件隨機變數與雙重期望值之相關考題，這類題型在課堂中也有練習過，本題難度僅在於破題而已，只要考生瞭解題意，再回想到課堂中練習過之題型，拿滿分不難。
考點命中	《高點統計學講義》第二回，趙治勳編撰，頁28。

答：

令N表從盒子中抽出之球號

X表最後投擲之n個銅板中出現正面之個數

$$f_N(n)=P(N=n)=\frac{1}{5}, n=1,2,3,4,5$$

$$X\mid N=n \sim Bin(n, p=\frac{1}{2})$$

(一) $$P(X=3)=\sum_{n=1}^{5}P(N=n)P(X=3\mid N=n)=\sum_{n=3}^{5}\frac{1}{5}P(X=3\mid N=n)$$

$$=\frac{1}{5}[\binom{3}{3}(\frac{1}{2})^3(\frac{1}{2})^{3-3}+\binom{4}{3}(\frac{1}{2})^3(\frac{1}{2})^{4-3}+\binom{5}{3}(\frac{1}{2})^3(\frac{1}{2})^{5-3}]=\frac{11}{80}$$

(二) $$E(X)=E[E(X\mid N)]=E[N\frac{1}{2}]=\frac{1}{2}E(N)=\frac{1}{2}(3)=\frac{3}{2}$$

二、投擲三個均勻銅板4次，令X表4次投擲中三個銅板都是反面的次數，Y表4次投擲中三個銅板只有出現一個反面的次數。（每小題10分，共20分）

(一)試求X與Y皆小於1的機率。

(二)試求$X-Y=1$的機率。

試題評析	本題是三項分配之相關考題，只要考生瞭解題意，知道可以利用三項分配答題，拿滿分不難。
考點命中	《高點統計學講義》第二回，趙治勳編撰，第五章第6節。

答：

每一次都投擲三個均勻銅板,故每次投擲有8種可能結果,如下:

{HHH}{HTT},{THT},{TTH},{THH},{HTH},{HHT},{TTT}

X表4次投擲中出現{TTT}之個數

Y表4次投擲中出現{THH},{HTH},{HHT}之個數

$$(X,Y)\sim 三項(n=4, p_X=\frac{1}{8}, p_Y=\frac{3}{8})$$

$$f_{XY}(x,y)=P(X=x\cap Y=y)=\frac{4!}{x!\,y!(4-x-y)!}(\frac{1}{8})^{x}(\frac{3}{8})^{y}(\frac{4}{8})^{4-x-y}\ ,0\le x\le 4,0\le y\le 4,0\le x+y\le 4$$

(一) $P(X=0\cap Y=0)=f_{XY}(x=0,y=0)=\frac{4!}{0!0!4!}(\frac{1}{8})^{0}(\frac{3}{8})^{0}(\frac{4}{8})^{4}=\frac{1}{16}$

(二) $P(X-Y=1)=P(X=1\cap Y=0)+P(X=2\cap Y=1)+P(X=3\cap Y=2)+P(X=4\cap Y=3)$

$$=f_{XY}(x=1,y=0)+f_{XY}(x=2,y=1)+0+0$$

$$=\frac{4!}{1!0!4!}(\frac{1}{8})^{1}(\frac{3}{8})^{0}(\frac{4}{8})^{3}+\frac{4!}{2!1!1!}(\frac{1}{8})^{2}(\frac{3}{8})^{1}(\frac{4}{8})^{1}=\frac{13}{256}$$

三、設隨機變數 X 與 Y 相互獨立，且各別具有自由度為2與3之卡方分配 $\chi^2(2)$ 與 $\chi^2(3)$。令隨機變數 $T=\frac{X}{X+Y}$。（每小題10分，共20分）

(一)試求 T 之機率密度函數 $f(t)$。

(二)試求 T 之變異數 $Var(T)$。

試題評析	本題是貝他分配之隨機實驗，課堂中已經討論過了，只要考生把熟讀課本，拿滿分不難。
考點命中	《高點統計學講義》第二回，趙治勳編撰，頁57(註2)。

答：

$$X\sim\chi^2_{(2)}=Gamma(\alpha=1,\lambda=\frac{1}{2})\perp Y\sim\chi^2_{(3)}=Gamma(\alpha=\frac{3}{2},\lambda=\frac{1}{2})$$

(一) $T=\frac{X}{X+Y}\sim Beta(\alpha=1,\beta=\frac{3}{2})$

$$f_T(t)=\frac{\Gamma(1+\frac{3}{2})}{\Gamma(1)\Gamma(\frac{3}{2})}t^{1-1}(1-t)^{\frac{3}{2}-1}=\frac{\frac{3}{2}\Gamma(\frac{3}{2})}{\Gamma(\frac{3}{2})}\sqrt{1-t}=\frac{3}{2}\sqrt{1-t}\ ,0<t<1$$

(二) $V(T)=\dfrac{(1)(\frac{3}{2})}{(1+\frac{3}{2})^{2}(1+\frac{3}{2}+1)}=\dfrac{12}{175}$

四、假設 $X_1,X_2,\ldots,X_n$ 為抽自具有常態分配 $N(\mu,\sigma^2)$ 之隨機樣本。（每小題10分，共20分）

(一)若 μ 和 σ^2 為未知的參數，且設 c 滿足 $P[X_1\le c]=0.90$，試求 c 之最大概似估計量（maximum likelihood estimator）。

(二)若 μ 已知為1，σ^2 為未知的參數，試求滿足 $P[X_1\le c]=0.90$ 之 c 的 $100(1-\alpha)\%$ 信賴水準的信賴區間。

（已知若隨機變數 Z 具有標準常態分配，則 $P[Z<1.2816]=0.90$）

試題評析	本題是MLE不變性與建立信賴區間之考題，只要對信賴區間之觀念清楚，拿滿分也不難。

考點命中	《高點統計學講義》第三回，趙治勳編撰，頁34、48。

答：

$X_1, X_2, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ 可得 $\hat{\mu}_{MLE} = \overline{X}, \hat{\sigma}^2_{MLE} = S_n^2 = \dfrac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}$

(一) $P(X_1 \le c) = P(Z \le \dfrac{c-\mu}{\sigma}) = 0.9 \Rightarrow \dfrac{c-\mu}{\sigma} = 1.2816 \Rightarrow c = \mu + 1.2816\sigma$

由MLE之不變性，

$\hat{c}_{MLE} = \hat{\mu}_{MLE} + 1.2816\hat{\sigma}_{MLE} = \overline{X} + 1.2816 S_n$

(二)由於 $\mu = 1$，故 $\hat{\sigma}^2_{MLE} = S_n^2 = \dfrac{\sum_{i=1}^{n}(X_i - 1)^2}{n}, c = 1 + 1.2816\sigma$

已知 σ 之 $100(1-\alpha)\%$ 之信賴區間為

$$P(\sqrt{\frac{nS_n^2}{\chi^2_{\frac{\alpha}{2}(n)}}} \le \sigma \le \sqrt{\frac{nS_n^2}{\chi^2_{1-\frac{\alpha}{2}(n)}}}) = 1-\alpha$$

可得 $P(1 + 1.2816\sqrt{\dfrac{nS_n^2}{\chi^2_{\frac{\alpha}{2}(n)}}} \le 1 + 1.2816\sigma \le 1 + 1.2816\sqrt{\dfrac{nS_n^2}{\chi^2_{1-\frac{\alpha}{2}(n)}}}) = 1-\alpha$

$$\Rightarrow P(1 + 1.2816\sqrt{\frac{nS_n^2}{\chi^2_{\frac{\alpha}{2}(n)}}} \le c \le 1 + 1.2816\sqrt{\frac{nS_n^2}{\chi^2_{1-\frac{\alpha}{2}(n)}}}) = 1-\alpha$$

因此，c 之 $100(1-\alpha)\%$ 之信賴區間為

$$(1 + 1.2816\sqrt{\frac{nS_n^2}{\chi^2_{\frac{\alpha}{2}(n)}}}, 1 + 1.2816\sqrt{\frac{nS_n^2}{\chi^2_{1-\frac{\alpha}{2}(n)}}})$$

五、考慮建立簡單直線迴歸模型 $Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i$, $i = 1,2,\ldots,n$ ，且假設 $\varepsilon_i, \varepsilon_2, \varepsilon_3, \ldots \varepsilon_n$ 為相互獨立且具常態分配 $N(0,\sigma^2)$ 之隨機誤差。（每小題10分，共20分）

(一)試求 β_1 與 β_2 之最大概似估計量，$\hat{\beta}_1$ 與 $\hat{\beta}_2$ 。

(二)試求題(一)中，$\hat{\beta}_1$ 與 $\hat{\beta}_2$ 之相關係數。

試題評析	本題是簡迴歸參數之MLE及迴歸係數估計值間之相關係數，只要有準備過簡迴歸，拿滿分是一定的。
考點命中	《迴歸分析熱門題庫》，高點文化出版，趙治勳著，頁2-10、2-19。

答：

(一)樣本：$\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$, i=1,2,⋯,n

$$f(\varepsilon_i)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(Y_i-\beta_1-\beta_2X_i)^2}{2\sigma^2}},-\infty<\varepsilon_i<\infty$$

概似函數：$L(\beta_1,\beta_2,\sigma^2)=\prod_{i=1}^{n}f(\varepsilon_i)=(\frac{1}{\sqrt{2\pi}\sigma})^n e^{-\frac{\sum(Y_i-\beta_1-\beta_2X_i)^2}{2\sigma^2}}$

$$\ln L(\beta_1,\beta_2,\sigma^2)=-\frac{n}{2}\ln(2\pi\sigma^2)-\frac{\sum(Y_i-\beta_1-\beta_2X_i)^2}{2\sigma^2}$$

令 $\begin{cases}\dfrac{\partial\ln L(\beta_1,\beta_2,\sigma^2)}{\partial\beta_1}=-\dfrac{2\sum(Y_i-\beta_1-\beta_2X_i)^1(-1)}{2\sigma^2}=0\\ \dfrac{\partial\ln L(\beta_1,\beta_2,\sigma^2)}{\partial\beta_2}=-\dfrac{2\sum(Y_i-\beta_1-\beta_2X_i)^1(-X_i)}{2\sigma^2}=0\\ \dfrac{\partial\ln L(\beta_1,\beta_2,\sigma^2)}{\partial\sigma^2}=-\dfrac{n}{2}\dfrac{1}{2\pi\sigma^2}2\pi+\dfrac{\sum(Y_i-\beta_1-\beta_2X_i)^2}{2\sigma^4}=0\end{cases}$

=> $\begin{cases}\hat{\beta}_2=\dfrac{SS_{XY}}{SS_X}\\ \hat{\beta}_1=\overline{Y}-\hat{\beta}_2\overline{X}\end{cases}$

(二)已知 $\hat{\beta}_2\sim N(\beta_2,\frac{\sigma^2}{SS_X})$，故 $V(\hat{\beta}_2)=\frac{\sigma^2}{SS_X}$

$\hat{\beta}_1\sim N(\beta_1,(\frac{1}{n}+\frac{\overline{X}^2}{SS_X})\sigma^2)$　故 $V(\hat{\beta}_1)=(\frac{1}{n}+\frac{\overline{X}^2}{SS_X})\sigma^2$

$Cov(\hat{\beta}_1,\hat{\beta}_2)=-\frac{\overline{X}}{SS_X}\sigma^2$

因此，$\rho_{\hat{\beta}_1\hat{\beta}_2}=\dfrac{Cov(\hat{\beta}_1,\hat{\beta}_2)}{\sqrt{V(\hat{\beta}_1)}\sqrt{V(\hat{\beta}_2)}}=\dfrac{-\frac{\overline{X}}{SS_X}\sigma^2}{\sqrt{(\frac{1}{n}+\frac{\overline{X}^2}{SS_X})\sigma^2}\sqrt{\frac{\sigma^2}{SS_X}}}=-\dfrac{\overline{X}}{\sqrt{SS_X(\frac{1}{n}+\frac{\overline{X}^2}{SS_X})}}$