《統計學概要》

一、假設有一個球箱中放置了8顆標記有數字的球，其中4顆標記數字"0"，2顆標記數字"-1"，2顆標記數字"1"。現在從這個球箱中，以取後放回的方式隨機抽出2顆球，球上標記的數字分別為X_1, X_2，並計算平均數$\overline{X}=(X_1+X_2)/2$。

(一)寫出此抽樣問題中的母體分配。(3分)

(二)求算母體平均數與母體變異數。(6分)

(三)說明X_1, X_2兩個隨機變數是否獨立及其理由。(5分)

(四)寫出$\overline{X}$的抽樣分配。(5分)

(五)求算$\overline{X}$的平均數與變異數。(6分)

(六)若同樣以取後放回的方式隨機抽出4顆球，求算$\overline{X_4}=(X_1+X_2+X_3+X_4)/4$的標準差。(5分)

試題評析	本題為樣本平均數之抽樣分配，與課本範例幾乎一模一樣。
考點命中	《高點・高上統計學講義》第二回，趙治勳編撰，第九章例題一。

答：

(一)

$X=x$	-1	0	1
$f_X(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

(二) $E(X)=(-1)\frac{1}{4}+(0)\frac{1}{2}+(1)\frac{1}{4}=0$

$V(X)=E(X^2)-[E(X)]^2=\frac{1}{2}$

其中 $E(X^2)=(-1)^2\frac{1}{4}+(0)^2\frac{1}{2}+(1)^2\frac{1}{4}=\frac{1}{2}$

(三) $X_1 \perp X_2$，因為實驗規則屬於抽後放回

(四)

樣本		$\overline{X}$	機率
-1	-1	-1	$\frac{1}{4}\times\frac{1}{4}=\frac{1}{16}$
0	0	0	$\frac{1}{2}\times\frac{1}{2}=\frac{1}{4}$
1	1	1	$\frac{1}{4}\times\frac{1}{4}=\frac{1}{16}$
-1	0	$-\frac{1}{2}$	$2\times\frac{1}{4}\times\frac{1}{2}=\frac{1}{4}$
-1	1	0	$2\times\frac{1}{4}\times\frac{1}{4}=\frac{1}{8}$
0	1	$\frac{1}{2}$	$2\times\frac{1}{2}\times\frac{1}{4}=\frac{1}{4}$

$\bar{X}=\bar{x}$	-1	$-\frac{1}{2}$	0	$\frac{1}{2}$	1
$f_{\bar{X}}(\bar{x})$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

(五) $E(\bar{X})=(-1)\frac{1}{16}+(-\frac{1}{2})\frac{1}{4}+(0)\frac{3}{8}+(\frac{1}{2})\frac{1}{4}+(1)\frac{1}{16}=0$

$V(\bar{X})=E(\bar{X}^2)-[E(\bar{X})]^2=\frac{1}{4}$

其中 $E(\bar{X}^2)=(-1)^2\frac{1}{16}+(-\frac{1}{2})^2\frac{1}{4}+(0)^2\frac{3}{8}+(\frac{1}{2})^2\frac{1}{4}+(1)^2\frac{1}{16}=\frac{1}{4}$

(六) $\sqrt{V(\bar{X}_4)}=\sqrt{\frac{V(X)}{4}}=\sqrt{\frac{\frac{1}{2}}{4}}=\sqrt{\frac{1}{8}}=0.3536$

二、大一統計學課程共有100位同學選修該課程，學期中共舉行期中考與期末考兩次考試。假設這兩次考試成績分別符合常態分配 $N(\mu_1,\sigma_1^2)$ 與 $N(\mu_2,\sigma_2^2)$。為了想了解該課程同學期中考平均成績與期末考平均成績是否有差異，我們從100位同學中隨機抽出4位同學並記錄他們的期中考(X)與期末考(Y)成績如下：

(X, Y)：(40, 60)(60, 50)(80, 50)(60, 80)

(一)請分別估計此課程100位同學期中考及期末考成績的平均數。(4分)

(二)請分別估計此課程100位同學期中考及期末考成績的變異數。(6分)

(三)請估計此課程100位同學期中考成績與期末考成績的相關係數。(6分)

(四)想了解該課程100位同學期中考平均成績與期末考平均成績是否有差異，請寫出檢定的虛無假設 H_0 與對立假設 H_1。(4分)

(五)寫出(四)小題中檢定統計量及其在虛無假設下的分配。(6分)

(六)計算(四)小題中檢定統計量的值。(4分)

試題評析	本題為兩相依母體平均數之假設檢定。
考點命中	《高點・高上統計學講義》第三回，趙治勳編撰，第十一章。

答：

母體: $(X,Y)\sim N_2(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$

樣本: $(X_1,Y_1),(X_2,Y_2),(X_3,Y_3),(X_4,Y_4)\overset{iid}{\sim}N_2(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$

(一) $\hat{\mu}_1=\bar{x}=60 \qquad \hat{\mu}_2=\bar{y}=60$

(二) $\hat{\sigma}_1^2=s_1^2=266.6667 \qquad \hat{\sigma}_2^2=s_2^2=200$

(三) $\hat{\rho}=r=\frac{SS_{XY}}{\sqrt{SS_X}\sqrt{SS_Y}}=\frac{-200}{\sqrt{800}\sqrt{600}}=-0.2887$

其中 $SS_X=\sum X_i^2-\frac{(\sum X_i)^2}{n}=15200-\frac{(240)^2}{4}=800$

$SS_Y=\sum Y_i^2-\frac{(\sum Y_i)^2}{n}=15000-\frac{(240)^2}{4}=600$

$$SS_{XY}=\sum X_i Y_i-\frac{(\sum X_i)(\sum Y_i)}{n}=14200-\frac{(240)(240)}{4}=-200$$

(四) $H_0:\mu_1=\mu_2$ vs $H_1:\mu_1\neq\mu_2$

(五) T.S.: $T=\dfrac{\bar{D}-(0)}{\dfrac{S_D}{\sqrt{n}}}\sim t_{(4-1=3)}$ 其中 $D_i=X_i-Y_i$

(六) $\bar{d}=\bar{x}-\bar{y}=60-60=0$ $s_D^2=\dfrac{\sum(d_i-\bar{d})^2}{n-1}=600$

$$T^*=\frac{0-(0)}{\dfrac{\sqrt{600}}{\sqrt{4}}}=0$$

三、承前一題中的資料，

(一)如果想了解該課程100位同學，期中考成績與期末考成績的變異數是否一樣，這組4位同學期中考及期末考成績的變異數是否可用來推導F分配？請說明理由。(5分)

(二)若進一步考慮簡單線性迴歸模型：

$Y=\alpha+\beta X+\varepsilon$，其中X為期中考成績，Y為期末考成績，求算判定係數 r^2（coefficient of determination），迴歸線截距項估計值 $\hat{\alpha}$ 及迴歸線斜率估計值 $\hat{\beta}$。(15分)

試題評析	本題為簡迴歸。
考點命中	《迴歸分析熱門題庫》，高點文化出版，趙治勳編著，第二章。

答：

(一) 不行。因為兩個母體並非獨立。$H_0:\sigma_1^2=\sigma_2^2$ 為真下，$\dfrac{S_1^2}{S_2^2}$ 也不會服從F分配。

(二) 判定係數 $r^2=(-0.2887)^2=0.0833$

$$\hat{\beta}=r\frac{\sqrt{SS_Y}}{\sqrt{SS_X}}=-0.2887\frac{\sqrt{600}}{\sqrt{800}}=-0.25$$

$$\hat{\alpha}=\bar{Y}-\hat{\beta}\bar{X}=\frac{240}{4}-(-0.25)\frac{240}{4}=75$$

四、(一)假設 $X_i\sim N(\mu,\sigma^2), i=1,\ldots,n,$ 為n個獨立的隨機變數，$\bar{X}$ 為其平均數，$S^2=\sum_{i=1}^{n}(X_i-\bar{X})^2/(n-1)$ 為變異數。請寫出 $\sum_{i=1}^{n}(\frac{X_i-\mu}{\sigma})^2$ 及 $\sum_{i=1}^{n}(\frac{X_i-\bar{X}}{\sigma})^2$ 的分配。(10分)

(二)自某母體分配中抽取一組樣本數n=100的隨機樣本，得樣本平均數 $\bar{x}=50.12$，樣本標準差s=6.76。若以樣本平均數為母體平均數的估計值，令e為其95%誤差界限，即 $P(|\bar{X}-\mu|\leq e)=0.95$。請概算誤差界限e之值。(10分)

參考值：$Z_{0.025}=1.96$

試題評析	本題為抽樣分配，相關內容都在課本中。
考點命中	《高點・高上統計學講義》第二回，趙治勳編撰，第九章。

答：

(一) $\dfrac{\sum(X_i-\mu)^2}{\sigma^2}\sim\chi^2_{(n)}$　　$\dfrac{\sum(X_i-\bar{X})^2}{\sigma^2}\sim\chi^2_{(n-1)}$

(二) 母體: $X\sim(\mu,\sigma^2)$

樣本: $X_1,X_2,\ldots,X_{100}\overset{iid}{\sim}(\mu,\sigma^2)$

點估計: $\bar{X}\underset{by\,C.L.T.}{\sim}N(\mu,\dfrac{\sigma^2}{100})$

$P(|\bar{X}-\mu|\le e)=0.95\quad\Rightarrow P(|Z|\le\dfrac{e}{\sigma/\sqrt{100}})=0.95\quad\Rightarrow\dfrac{e}{\sigma/\sqrt{100}}=z_{0.025}=1.96$

$\Rightarrow e=1.96\dfrac{\sigma}{\sqrt{100}}\overset{\sigma\text{以}s\text{代替}}{=}1.96\dfrac{6.76}{\sqrt{100}}=1.32496$