《統計學概要》

一、甲、乙進行擲骰子比賽，兩人同時各擲一個公正骰子，若擲出的點數相差 1 或 2 時，則甲贏得比賽，比賽結束；若擲出的點數相差 3、4 或 5 時，則乙贏得比賽，比賽結束；若點數相同時，則兩人再擲一次，直至分出勝負才停止比賽。假設 X 表示兩人擲出的點數差。

(一)請列出 X 的機率分配。（10 分）

(二)計算 X 的期望值。（5 分）

(三)試求比賽只進行一次投擲就結束的機率。（5 分）

(四)試求甲在第二次投擲時贏得比賽的機率。（5 分）

試題評析	本題是考驗考生如何產生一個隨機變數，然而題目的遊戲並不複雜，且隨機變數的意義也很明瞭，故要拿到滿分不難。
考點命中	《高點・高上統計學講義》第二回，趙治勳編撰。

答：

令 X_1, X_2 分別表甲與乙投出之點數，$X_1, X_2 \overset{iid}{\sim} DU(1,6)$

$X = |X_1 - X_2|$

(一)

$X = \|X_1 - X_2\|$		X_2					
		1	2	3	4	5	6
X_1	1	0	1	2	3	4	5
	2	1	0	1	2	3	4
	3	2	1	0	1	2	3
	4	3	2	1	0	1	2
	5	4	3	2	1	0	1
	6	5	4	3	2	1	0

由上表可得 X 之機率分配如下：

$X = x$	0	1	2	3	4	5
$f_X(x)$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{8}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{2}{36}$

(二) $E(X) = 0\times\frac{6}{36} + 1\times\frac{10}{36} + 2\times\frac{8}{36} + 3\times\frac{6}{36} + 4\times\frac{4}{36} + 5\times\frac{2}{36} = \frac{35}{18}$

(三) 根據題意，比賽只進行一次就結束了，表示第一次投擲時兩人之點數差為1或2或3或4或5(點數差不是0)

$$P(\text{比賽只進行一次}) = 1 - P(X=0) = 1 - \frac{6}{36} = \frac{5}{6}$$

(四) 根據題意，甲於第二次投擲中贏得比賽，就表示第一次投出之點數差為0且第二次投出之點數差為1或2

$$P(\text{甲於第二次投擲中贏得比賽}) = \frac{6}{36}\times(\frac{10}{36} + \frac{8}{36}) = \frac{1}{12}$$

二、甲公司品管檢驗員抽驗該公司生產之燈泡 20 盒，得各盒不良品件數 X 的分配如下表所示：

X	0	1	2	3	4	5
盒數	1	8	5	3	2	1

試求 不良品件數 的下列項目：（每小題 5 分，共 25 分）
(一)平均數
(二)中位數
(三)眾數
(四)標準差
(五)四分位距（Interquartile Range）

試題評析	本題是考敘述統計學，四等常考範圍，且也屬於基本計算，要拿到滿分並不難。
考點命中	《高點・高上統計學講義》第一回，趙治勳編撰，第三章。

答：

(一) 平均數 $=0\times\frac{1}{20}+1\times\frac{8}{20}+2\times\frac{5}{20}+3\times\frac{3}{20}+4\times\frac{2}{20}+5\times\frac{1}{20}=2$

(二) 中位數 $=(X_{(\frac{N}{2})},X_{(\frac{N}{2}+1)})=(X_{(10)},X_{(11)})=(2,2)=2$

(三) 眾數 $=1$ (出現次數最多)

(四) 標準差 $=\sqrt{\frac{\sum x_i f_i-n\bar{X}^2}{n-1}}=\sqrt{\frac{112-20\times 2^2}{20-1}}=1.2978$

其中 $\sum x_i^2 f_i=0^2\times1+1^2\times8+2^2\times5+3^2\times3+4^2\times2+5^2\times1=112$

(五) $1\times\frac{20}{4}=5$ $\quad Q_1=\frac{X_{(5)}+X_{(6)}}{2}=\frac{1+1}{2}=1$

$3\times\frac{20}{4}=15$ $\quad Q_3=\frac{X_{(15)}+X_{(16)}}{2}=\frac{3+3}{2}=3$

四分位距 $=Q_3-Q_1=3-1=2$

三、捷運公司想了解民眾搭乘捷運上下班的情形，故隨機訪問了 160 位男士及 140 位女士，得知有 88 位男士及 84 位女士搭乘捷運上下班。（每小題 5 分，共 25 分）
(一)估計民眾搭乘捷運上下班的比例。
(二)在 95%信心水準下，求男士搭乘捷運上下班比例的抽樣誤差。
(三)求女士搭乘捷運上下班比例的 95%信賴區間。
(四)求男士與女士搭乘捷運上下班比例差的 95%信賴區間。
(五)在顯著水準 $\alpha=0.05$ 下，利用(四)的結果，判斷男士與女士搭乘捷運上下班的比例是否有顯著差異。

試題評析	本題是考兩獨立母體成功比例之信賴區間與假設檢定，屬於基本計算考題，要拿到高分並不難。
考點命中	《高點・高上統計學講義》第四回，趙治勳編撰，第十一、十二章。

答：

令 X_1,X_2 分別表男士與女士為搭乘捷運上下班

母體： $X_1\sim Ber(p_1)\perp X_2\sim Ber(p_2)$ 　假設 $X_1\perp X_2$

樣本： $X_{11},X_{12},\cdots,X_{1160}\overset{iid}{\sim}Ber(p_1)$

$X_{21}, X_{22}, \cdots, X_{2140} \overset{iid}{\sim} Ber(p_2)$

點估計： $\hat{p}_1 = \dfrac{\sum X_{1i}}{160} \underset{byC.L.T.}{\sim} N(p_1, \dfrac{p_1(1-p_1)}{160}) \perp \hat{p}_2 = \dfrac{\sum X_{2i}}{140} \underset{byC.L.T.}{\sim} N(p_2, \dfrac{p_2(1-p_2)}{140})$

(一) $\hat{p} = \dfrac{\sum X_{1i} + \sum X_{2i}}{160+140} = \dfrac{88+84}{160+140} = \dfrac{43}{75} = 0.5733$

(二) B 表抽樣誤差

$$P(|\hat{p}_1 - p_1| \le B) = 0.95 \quad \Rightarrow \quad P(|Z| \le \frac{B}{\sqrt{\frac{p_1(1-p_1)}{160}}}) = 0.95$$

$$\Rightarrow \frac{B}{\sqrt{\frac{p_1(1-p_1)}{160}}} = z_{0.025}$$

$$\Rightarrow B = z_{0.025}\sqrt{\frac{p_1(1-p_1)}{160}} \overset{p_1代\hat{p}_1=\frac{88}{160}}{=} 1.96\sqrt{\frac{\frac{88}{160}(1-\frac{88}{160})}{160}}=0.0771$$

(三) 樞紐量： $\dfrac{\hat{p}_2 - p_2}{\sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{140}}} \underset{byC.L.T.}{\sim} N(0,1)$

機率區間： $P(-z_{0.025} \le \dfrac{\hat{p}_2 - p_2}{\sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{140}}} \le z_{0.025})=0.95$

結論： p_2 之95%信賴區間為

$$(\hat{p}_2 \mp z_{0.025}\sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{140}})=(\frac{84}{140} \mp 1.96\sqrt{\frac{\frac{84}{140}(1-\frac{84}{140})}{140}})$$
$=(0.5188,0.6812)$

(四) 樞紐量： $\dfrac{(\hat{p}_1 - \hat{p}_2)-(p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{160} + \frac{\hat{p}_2(1-\hat{p}_2)}{140}}} \underset{byC.L.T.}{\sim} N(0,1)$

機率區間： $P(-z_{0.025} \le \dfrac{(\hat{p}_1 - \hat{p}_2)-(p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{160} + \frac{\hat{p}_2(1-\hat{p}_2)}{140}}} \le z_{0.025})=0.95$

結論： $p_1 - p_2$ 之95%信賴區間為

$$((\hat{p}_1 - \hat{p}_2) \mp z_{0.025}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{160} + \frac{\hat{p}_2(1-\hat{p}_2)}{140}})$$

$$=((\frac{88}{160}-\frac{84}{140})\mp 1.96\sqrt{\frac{\frac{88}{160}(1-\frac{88}{160})}{160}+\frac{\frac{84}{140}(1-\frac{84}{140})}{140}})=(-0.1619,0.0619)$$

(五) $H_0: p_1 = p_2 \quad vs \quad H_1: p_1 \neq p_2$

由於 $p_1 - p_2$ 之95%信賴區間有包含0，故在顯著水準為5%下，採不拒絕 H_0 之決策。我們沒有足夠證據去推論男士與女士搭乘捷運上下班之比例有差異。

四、已知一有限母體所含的數字為{0,1,2,3,4}，以抽出不放回方式隨機抽取 2 個為一樣本，設 Xi 為第 i 次抽出的數字，$i=1,2$，$\bar{X}=(X_1+X_2)/2$。

(一)試求 $\bar{X}$ 的抽樣分配。（10 分）

(二)利用(一)的結果，驗證 $E(\bar{X})=\mu$，其中μ為母體平均數。（5 分）

(三)利用(一)的結果，驗證$V(\bar{X})=\frac{N-n}{N-1}\frac{\sigma^2}{n}$，其中 N 為母體觀察值個數，n 為樣本數，σ^2為母體變異數。（10 分）

試題評析	本題是考抽樣分配，此類題型在考古題中也多次出現，課本例題也幾乎相同，要拿到滿分並不難。
考點命中	《高點・高上統計學講義》第四回，趙治勳編撰，第九章，例1(2)。

答：

樣本	$\bar{X}$	樣本	$\bar{X}$
0,1	0.5	1,3	2
0,2	1	1,4	2.5
0,3	1.5	2,3	2.5
0,4	2	2,4	3
1,2	1.5	3,4	3.5

(一)

$\bar{X}=\bar{x}$	0.5	1	1.5	2	2.5	3	3.5
$f_{\bar{X}}(\bar{x})$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

(二)

$$\mu=\frac{0+1+2+3+4}{5}=2, \qquad E(\bar{X})=0.5\times\frac{1}{10}+1\times\frac{1}{10}+\cdots+3.5\times\frac{1}{10}=2$$

可得 $E(\bar{X})=\mu$

(三)

$$\frac{N-n}{N-1}\frac{\sigma^2}{n}=\frac{5-2}{5-1}\frac{2}{2}=0.75 \quad 其中\ \sigma^2=\frac{(0-2)^2+(1-2)^2+\cdots+(4-2)^2}{5}=2$$

$$V(\bar{X})=E(\bar{X}^2)-[E(\bar{X})]^2=4.75-(2)^2=0.75$$

其中 $E(\bar{X}^2)=0.5^2\times\frac{1}{10}+1^2\times\frac{1}{10}+\cdots+3.5^2\times\frac{1}{10}=4.75$

可得 $V(\bar{X})=\frac{N-n}{N-1}\frac{\sigma^2}{n}$