概率论与数理统计第四次习题课题目解答

题 1 设 $(X_1, X_2, \ldots, X_r)$ 服从多项分布 $M(n, p_1, p_2, \ldots, p_r)$（$p_i > 0, i = 1, \ldots, r$ 且 $p_1 + \cdots + p_r = 1$）。求 X_i 与 X_j 的相关系数，$i, j = 1, \ldots, r$。

解法 1： 由相关系数的定义不难验证 $\mathrm{Corr}(X_i, X_i) = 1$。下面考虑 $i \neq j$ 的情形。

由多项分布的性质知，$X_i \sim B(n, p_i)$，$(X_i, X_j, n - X_i - X_j) \sim M(n, p_i, p_j, 1 - p_i - p_j)$，所以

$$EX_i = np_i, \qquad \mathrm{Var}X_i = np_i(1 - p_i).$$

$$\begin{aligned}
E(X_iX_j) &= \sum_{k\geq 0, l\geq 0, k+l\leq n} kl \cdot P(X_i = k, X_j = l) \\
&= \sum_{k\geq 0, l\geq 0, k+l\leq n} kl \cdot \frac{n!}{k!l!(n-k-l)!} p_i^k p_j^l (1 - p_i - p_j)^{n-k-l} \\
&= n(n-1)p_ip_j \sum_{k\geq 1, l\geq 1, k+l\leq n} \frac{(n-2)!}{(k-1)!(l-1)!(n-k-l)!} p_i^{k-1} p_j^{-1-1} (1 - p_i - p_j)^{n-k-l} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\text{（注意这个求和的概率含义）} \\
&= n(n-1)p_ip_j.
\end{aligned}$$

所以，

$$\mathrm{Cov}(X_i, X_j) = E(X_iX_j) - EX_i \cdot EX_j = n(n-1)p_ip_j - np_i \cdot np_j = -np_ip_j.$$

从而

$$\mathrm{Corr}(X_i, X_j) = \frac{\mathrm{Cov}(X_i, X_j)}{\sqrt{\mathrm{Var}X_i \mathrm{Var}X_j}} = -\sqrt{\frac{p_ip_j}{(1-p_i)(1-p_j)}}.$$

解法 2： 由相关系数的定义不难验证 $\mathrm{Corr}(X_i, X_i) = 1$。下面考虑 $i \neq j$ 的情形。

由多项分布的性质知，$X_i \sim B(n, p_i)$，$X_i + X_j \sim B(n, p_i + p_j)$，所以

$$EX_i = np_i, \qquad \mathrm{Var}X_i = np_i(1 - p_i).$$

由

$$\mathrm{Var}(X_i + X_j) = \mathrm{Var}X_i + \mathrm{Var}X_j + 2\mathrm{Cov}(X_i, X_j)$$

得

$$\mathrm{Cov}(X_i, X_j) = \frac{n(p_i + p_j)(1 - p_i - p_j) - np_i(1 - p_i) - np_j(1 - p_j)}{2} = -np_ip_j.$$

所以，

$$\mathrm{Corr}(X_i, X_j) = \frac{\mathrm{Cov}(X_i, X_j)}{\sqrt{\mathrm{Var}X_i \mathrm{Var}X_j}} = -\sqrt{\frac{p_ip_j}{(1-p_i)(1-p_j)}}.$$

解法 3： 由相关系数的定义不难验证 $\mathrm{Corr}(X_i, X_i) = 1$。下面考虑 $i \neq j$ 的情形。

记

$$Y_k^{(i)} = \begin{cases} 1, & \text{若第 } k \text{ 次试验得到第 } i \text{ 种结果;} \\ 0, & \text{否则} \end{cases}, \qquad 1 \leq i \leq r, 1 \leq k \leq n.$$

于是

$$X_i = \sum_{k=1}^{n} Y_k^{(i)}.$$

关于 $Y_k^{(i)}$，我们有以下结论：

- 任何一次试验不会有两种不同的结果同时发生，所以对任何 $i \neq j$，$Y_k^{(i)}$ 和 $Y_k^{(j)}$ 中至少有一个为零，从而 $Y_k^{(i)} Y_k^{(j)} = 0$。
- 各次试验是独立进行的，所以对任何 $i_1, \ldots, i_n \in \{1, \ldots, r\}$，$Y_1^{(i_1)}, \ldots, Y_n^{(i_n)}$ 独立。
- 对任何 i 和 k，$EY_k^{(i)} = p_i$，$\mathrm{Var} Y_k^{(i)} = P(Y_k^{(i)} = 1) P(Y_k^{(i)} = 0) = p_i(1 - p_i)$。

从而利用独立和的方差性质，

$$\mathrm{Var} X_i = \mathrm{Var}\left(\sum_{k=1}^{n} Y_k^{(i)}\right) = \sum_{k=1}^{n} \mathrm{Var}(Y_k^{(i)}) = n p_i (1 - p_i).$$

利用协方差的双线性、对称性，独立性蕴含不相关性等性质，我们有

$$\begin{aligned}
\mathrm{Cov}(X_i, X_j) &= \mathrm{Cov}\left(\sum_{k=1}^{n} Y_k^{(i)}, \sum_{l=1}^{n} Y_l^{(j)}\right) \\
&= \sum_{k=1}^{n} \mathrm{Cov}(Y_k^{(i)}, Y_k^{(j)}) + 2 \sum_{1 \leq k < l \leq n} \mathrm{Cov}(Y_k^{(i)}, Y_l^{(j)}) \\
&= \sum_{k=1}^{n} \mathrm{Cov}(Y_k^{(i)}, Y_k^{(j)}) \\
&= \sum_{k=1}^{n} \left[E(Y_k^{(i)}, Y_k^{(j)}) - E(Y_k^{(i)}) E(Y_k^{(j)}) \right] \\
&= \sum_{k=1}^{n} \left[E(Y_k^{(i)}, Y_k^{(j)}) - E(Y_k^{(i)}) E(Y_k^{(j)}) \right] \\
&= \sum_{k=1}^{n} \left[- p_i p_j \right] = -n p_i p_j.
\end{aligned}$$

所以，

$$\mathrm{Corr}(X_i, X_j) = \frac{\mathrm{Cov}(X_i, X_j)}{\sqrt{\mathrm{Var} X_i \mathrm{Var} X_j}} = -\sqrt{\frac{p_i p_j}{(1 - p_i)(1 - p_j)}}.$$

题 2 设 (X, Λ) 的概率分布为：Λ 的边缘分布密度为 $\frac{1}{\lambda^2} e^{-\beta/\lambda}$，$\lambda > 0$，其中 β 为正常数；给定 $\Lambda = \lambda$ 时，X 服从期望为 $\frac{1}{\lambda}$ 的指数分布。求 EX。

解法 1： 由 X 关于 Λ 的条件分布知，

$$E(X | \Lambda = \lambda) = \frac{1}{\lambda}.$$

因此由全期望公式

$$EX = E(E(X|\Lambda)) = \int_0^{+\infty} E(X|\Lambda = \lambda) f_\Lambda(\lambda) d\lambda = \int_0^{+\infty} \frac{1}{\lambda^3} e^{-\frac{\beta}{\lambda}} d\lambda = \frac{1}{\beta^2},$$

最后的积分计算中利用了变量代换 $t=1/\lambda$。

而由 Λ 的概率密度的归一化

$$1=\int_0^{+\infty}\frac{1}{\lambda^2}e^{-\frac{\beta}{\lambda}}d\lambda$$

得到 $\beta=1$，从而 $EX=1$。

解法 2： 先求 (X,Λ) 的联合概率密度

$$f_{X,\Lambda}(x,\lambda)=f_\Lambda(\lambda)f_{X|\Lambda}(x|\lambda)=\frac{1}{\lambda^2}e^{-\frac{\beta}{\lambda}}I_{\lambda>0}\cdot\lambda e^{-\lambda x}I_{x>0}.$$

进而得到 X 的边缘概率密度

$$f_X(x)=\int_{-\infty}^{+\infty}f_{X,\Lambda}(x,\lambda)d\lambda=\int_0^{+\infty}\frac{1}{\lambda^3}e^{-\frac{\beta}{\lambda}}\lambda e^{-\lambda x}d\lambda,$$

最后这个积分虽无初等函数的表达式，但并不妨碍我们从定义计算 EX

$$EX=\int_{-\infty}^{+\infty}xf_X(xdx=\int_0^{+\infty}x\int_0^{+\infty}\frac{1}{\lambda^2}e^{-\frac{\beta}{\lambda}-\lambda x}d\lambda dx=\int_0^{+\infty}\int_0^{+\infty}\frac{x}{\lambda^2}e^{-\frac{\beta}{\lambda}-\lambda x}dxd\lambda=\frac{1}{\beta^2},$$

这里交换累次积分的顺序是关键。

而由 Λ 的概率密度的归一化

$$1=\int_0^{+\infty}\frac{1}{\lambda^2}e^{-\frac{\beta}{\lambda}}d\lambda$$

得到 $\beta=1$，从而 $EX=1$。

题 3 设 $X,Y\overset{i.i.d}{\sim}\mathrm{Exp}(\lambda)$。求

$$Z=\begin{cases}3X+1, & 若X\geq Y\\ 6Y, & 若X<Y\end{cases}.$$

的数学期望。（这是教材习题 3.5.12。）

解法 1： 记

$$\begin{aligned}h(x,y)&=\begin{cases}3x+1, & x\geq y\\ 6y, & x<y.\end{cases}\\&=(3x+1)\cdot I_{x\geq y}+6y\cdot I_{x<y}\\&=(3x+1)\cdot(1-I_{x<y})+6y\cdot I_{x<y}\\&=(3x+1)+(6y-3x-1)\cdot I_{x<y},\end{aligned}$$

所以

$$Z=h(X,Y)=(3X+1)+(6Y-3X-1)\cdot I_{X<Y}.$$

于是

$$
\begin{aligned}
E[Z] &= E[h(X,Y)] = \iint\limits_{\mathbb{R}^2} h(x,y) f_{X,Y}(x,y)dxdy \\
&= \int_0^{+\infty}\int_0^{+\infty} \big[(3x+1)+(6y-3x-1)I_{x<y}\big]\lambda e^{-\lambda x}\lambda e^{-\lambda y}dxdy \\
&= \int_0^{+\infty}(3x+1)\lambda e^{-\lambda x}dx\int_0^{+\infty}\lambda e^{-\lambda y}dy + \int_0^{+\infty}\left(\int_x^{+\infty}(6y-3x-1)\lambda e^{-\lambda y}dy\right)\lambda e^{-\lambda x}dx \\
&= \left(\frac{3}{\lambda}+1\right)\cdot 1 + \int_0^{+\infty}\left(\int_x^{+\infty}\big[6(y-x)+3x-1\big]\lambda e^{-\lambda(y-x)}dy\right)\lambda e^{-2\lambda x}dx \\
&= \frac{3}{\lambda}+1+\int_0^{+\infty}\left(\frac{6}{\lambda}+(3x-1)\right)\lambda e^{-2\lambda x}dx = \frac{3}{\lambda}+1+\frac{3}{\lambda}+\frac{3}{4\lambda}-\frac{1}{2} = \frac{27}{4\lambda}+\frac{1}{2}.
\end{aligned}
$$

解法 2： 与解法 1 类似，得到

$$
Z = h(X,Y) = (3X+1)+(6Y-3X-1)\cdot I_{X<Y}.
$$

我们希望通过全期望公式 $EZ = E[E(Z|X)]$ 来计算 Z 的数学期望。

$$
\begin{aligned}
E[Z|X=x] &= E(h(X,Y)|X=x) = E(h(x,Y)|X=x) \\
&= E(h(x,Y)) \qquad (h(x<Y) \text{ 是 } Y \text{ 的函数与 } X \text{ 独立}) \\
&= \int_0^{+\infty}\big[(3x+1)+(6y-3x-1)\cdot I_{x<y}\big]\lambda e^{-\lambda y}dy \\
&= 3x+1+e^{-\lambda x}\int_x^{+\infty}\big[6(y-x)+3x-1\big]\lambda e^{-\lambda(y-x)}dy \\
&= 3x+1+e^{-\lambda x}\left(\frac{6}{\lambda}+(3x-1)\right).
\end{aligned}
$$

从而

$$
\begin{aligned}
EZ = E[E(Z|X)] &= \int_0^{+\infty} E(Z|X=x)\lambda e^{-\lambda x}dx \\
&= \int_0^{+\infty}\left[3x+1+e^{-\lambda x}\left(\frac{6}{\lambda}-1+3x\right)\right]\lambda e^{-\lambda x}dx \\
&= \frac{3}{\lambda}+1+\frac{3}{\lambda}-\frac{1}{2}+\frac{3}{4\lambda} = \frac{27}{4\lambda}+\frac{1}{2}.
\end{aligned}
$$

解法 3: 用 X,Y 的大小关系来划分，这相当于 $E[E(Z|I_{X\geq Y})]$，

$$
\begin{aligned}
EZ &= E(Z|X\geq Y)P(X\geq Y)+E(Z|X<Y)P(X<Y) \\
&= E(3X+1|X\geq Y)P(X\geq Y)+E(6Y|X<Y)P(X<Y) \\
&= \big[3E(X|X\geq Y)+1\big]P(X\geq Y)+6E(Y|X<Y)P(X<Y).
\end{aligned}
$$

由于 X,Y 独立同分布，所以 X,Y 的联合概率密度函数与 Y,X 的联合概率密度函数相同，因此我们可以对调 X,Y 的角色，

$$
P(X<Y) = P(Y<X) = P(Y\leq X),
$$

而

$$P(X<Y)+P(X\geq Y)=1,$$

所以 $P(X\geq Y)=P(X<Y)=1/2$，另外，

$$E(X|X\geq Y)=E(Y|X<Y),$$

所以

$$EZ=\frac{9}{2}E(X|X\geq Y)+\frac{1}{2}.$$

为计算 $E(X|X\geq Y)$，我们先计算在已知 $X\geq Y$ 的条件下，X 的条件概率分布函数和条件概率密度函数。

$$\begin{aligned}F_{X|X\geq Y}(x)&=P(X\leq x|X\geq Y)\\&=\frac{P(Y\leq X\leq x)}{P(X\geq Y)}=2P(Y\leq X\leq x)\\&=2\int_0^x\int_v^x\lambda e^{-\lambda u}\lambda e^{-\lambda v}dudv\\&=2\int_0^x\left[e^{-\lambda v}-e^{-\lambda x}\right]\lambda e^{-\lambda v}dv,\end{aligned}$$

所以

$$\begin{aligned}f_{X|X\geq Y}(x)&=\frac{d}{dx}F_{X|X\geq Y}(x)\\&=2\int_0^x\lambda e^{-\lambda x}\lambda e^{-\lambda v}dv=2\lambda e^{-\lambda x}\left(1-e^{-\lambda x}\right)I_{x>0},\end{aligned}$$

于是

$$E(X|X\geq Y)=\int_0^{+\infty}x\cdot 2\lambda e^{-\lambda x}\left(1-e^{-\lambda x}\right)dx=\frac{2}{\lambda}-\frac{1}{2\lambda}=\frac{3}{2\lambda},$$

从而

$$EZ=\frac{27}{4\lambda}+\frac{1}{2}.$$

附注：

- 《概率论与数理统计教程》书后答案是错的，《概率论与数理统计教程——习题与解答》中给出的解法也是错的！！！作者错误地认为当 X,Y 独立时，

  $$EZ=E(3X+1)P(X\geq Y)+E(6Y)P(X<Y),$$

  这对初学的人是个常见的错误，因为他们只关心随机变量 Z 在某些特定条件下的取值（比如当 $X\geq Y$，$Z=3X+1$），而忽略了随机变量的概率分布受这些特定条件的影响。所以这个错误的解答对初学者更具迷惑性，而且作者还特别指出这个等式“的证明这里略去”，这就更加让学生认为这个“直观的结论”是对的，只不过证明需要费些口舌。
- 在解法 3 中我们可以更多地利用 X,Y 的对称性。比如

$$\begin{aligned}P(Y\leq X\leq x)=P(X\leq Y\leq x)&=\frac{P(X\leq Y\leq x)+P(Y\leq X\leq x)}{2}\\&=\frac{P(X\leq x,Y\leq x)}{2}\\&=\frac{\left[P(X\leq x)\right]^2}{2}=\frac{(1-e^{-\lambda x})^2}{2}I_{x>0},\end{aligned}$$

所以

$$F_{X|X\geq Y}(x)=(1-e^{-\lambda x})^2 I_{x>0},$$

从而得到

$$f_{X|X\geq Y}(x)=2\lambda e^{-\lambda x}(1-e^{-\lambda x})I_{x>0}.$$

- 另外一种利用对称性的方法是

$$\begin{aligned}EZ&=\frac{9}{2}E(X|X\geq Y)+\frac{1}{2}\\&=\frac{9}{2}\left[P(X\geq Y)E(X|X\geq Y)+P(X<Y)E(Y|X<Y)\right]+\frac{1}{2}\\&=\frac{9}{2}E(\max\{X,Y\})+\frac{1}{2}\\&=\frac{9}{2}\big[EX+EY-E(\min\{X,Y\})\big]+\frac{1}{2}.\end{aligned}$$

由于 X,Y 服从指数分布 $\mathrm{Exp}(\lambda)$ 并且相互独立，所以 $\min\{X,Y\}$ 服从指数分布 $\mathrm{Exp}(\lambda+\lambda)$，这可以利用指数分布是唯一具有“无记忆性”的连续分布这一事实得到。所以，

$$EZ=\frac{9}{2}\left[\frac{1}{\lambda}+\frac{1}{\lambda}-\frac{1}{2\lambda}\right]+\frac{1}{2}=\frac{27}{4\lambda}+\frac{1}{2}.$$

题4 投掷一个公平的硬币，及正面为H、反面为T，

1. 直至首次出现HH时停止，请计算投掷次数的期望与方差；
2. 如果停止准则变为首次出现HT，此时投掷次数的期望和方差分别是多少；
3. 假设甲、乙进行一场比赛，投掷硬币直至首次出现HH或HT停止。如果以HH结束，则甲胜，HT结束为乙胜，请问甲、乙的获胜概率; 若改为HH先出现甲胜，HT先出现乙胜，结果怎样。

解:

1. 设X为投掷次数，Y为随机变量，其样本为:首次掷出反面（记为T），前两次掷出正反（记为HT），前两次掷出反正（记为TH）。则

$$\begin{aligned}E(X)&=E(E(X|Y))\\&=P(Y=T)E(X|Y=T)+P(Y=HT)E(X|Y=HT)+P(Y=HH)E(X|Y=HH)\\&=\tfrac{1}{2}(1+E(X))+\tfrac{1}{4}(2+E(X))+\tfrac{1}{4}\cdot 2\end{aligned}$$

解得$E(X)=6$。

$$\begin{aligned}E(X^2)&=E(E(X^2|Y))\\&=P(Y=T)E(X^2|Y=T)+P(Y=HT)E(X^2|Y=HT)+P(Y=HH)E(X^2|Y=HH)\\&=\tfrac{1}{2}E((X+1)^2)+\tfrac{1}{4}E((2+X)^2)+\tfrac{1}{4}\cdot 4\end{aligned}$$

解得$E(X^2)=58$，$Var(X)=22$。

2.

$$\begin{aligned}E(X|H)&=P(Y=H)E(X|HH)+P(Y=T)E(X|HT)\\&=\tfrac{1}{2}(1+E(X|H))+\tfrac{1}{2}\cdot 2\\&=\tfrac{1}{2}(E(X|H))+\tfrac{3}{2}\end{aligned}$$

解得$E(X|H)=3$。
又由

$$\begin{aligned} E(X) &= P(T)E(X|T)+P(H)E(X|H) \\ &= \tfrac{1}{2}E(X)+2 \end{aligned}$$

解得$E(X)=4$。

$$\begin{aligned} E(X^2|H) &= P(T)E(X^2|HT)+P(H)E(X^2|HH) \\ &= \tfrac{1}{2}E((X+1)^2|H)+\tfrac{1}{2}\cdot 4 \\ &= \tfrac{1}{2}E(X^2|H)+\tfrac{11}{2} \end{aligned}$$

解得$E(X^2|H)=11$。
又由

$$\begin{aligned} E(X^2) &= P(T)E(X^2|T)+P(H)E(X^2|H) \\ &= P(T)(E(X)^2+9)+P(H)E(X^2|H) \\ &= \tfrac{1}{2}E(X^2)+10 \end{aligned}$$

解得$E(X^2)=20$，$Var(X)=4$。

3. 设甲胜为事件X，Y为随机变量，其样本为:掷出反面（记为T），与掷出正面（记为H）。

$P(X)=P(Y=H)P(X|Y=H)+P(Y=T)P(X|Y=T)=\frac{1}{2}P(X|Y=H)+\frac{1}{2}P(X|Y=T)$

$P(X|Y=H)=P(Y=H)P(X|HH)+P(Y=T)P(X|HT)=\frac{1}{2}\cdot 1+\frac{1}{2}P(X|T)$

$P(X|Y=T)=P(Y=H)P(X|TH)+P(Y=T)P(X|TT)=\frac{1}{2}\cdot 0+\frac{1}{2}P(X|T)$

$\Rightarrow P(X|Y=T)=0$。

所以$P(X)=1/4$。

思考：如果更一般的情况，有$1\sim n$编号的n个筹码，随机抽取筹码，如果相连的两个11先出现甲胜，连的两个n1先出现乙胜，求甲的获胜概率。

题 5 将编号为 1 至 n 的 n 个球随机投入编号为 1 至 n 的 n 个盒子中，并限制每一个盒子中只能放入一个球，设球与盒子的号码一致的个数为 S_n ，求证：

$$\frac{S_n-ES_n}{n}\xrightarrow{P}0,\quad n\to\infty$$

证法 1. 记

$$X_k=\begin{cases}1, & \text{如果编号为 } k \text{ 的球被放入编号为 } k \text{ 的盒子} \\ 0, & \text{否则}\end{cases}$$

则

$$S_n=X_1+\cdots+X_n,$$

于是对任意 $\varepsilon>0$

$$P\left(\left|\frac{S_n-ES_n}{n}\right|>\varepsilon\right)\le\frac{\mathrm{Var}S_n}{n^2\varepsilon^2}=\frac{1}{n^2\varepsilon^2}\left(\sum_{i=1}^{n}\mathrm{Var}(X_i)+\sum_{1\le i\ne j\le n}\mathrm{Cov}(X_i,X_j)\right).$$

而

$$P(X_k=1)=\frac{1}{n},\quad P(X_i=1,X_j=1)=\frac{1}{n(n-1)},\quad \forall 1\le k\le n,\quad \forall 1\le i\ne j\le n,$$

故

$$EX_k = \frac{1}{n}, \quad \mathrm{Var}X_k = \frac{n-1}{n^2}, \quad \mathrm{Cov}(X_k, X_l) = \frac{1}{n(n-1)} - \frac{1}{n^2} = \frac{1}{n^2(n-1)}.$$

所以,

$$\frac{1}{n^2\varepsilon^2}\left(\sum_{i=1}^{n}\mathrm{Var}(X_i) + \sum_{1\le i\ne j\le n}\mathrm{Cov}(X_i, X_j)\right) = \frac{1}{n^2\varepsilon^2}\left(\frac{n-1}{n} + \frac{1}{n}\right) \to 0, \quad n\to\infty.$$

因此 $\frac{S_n - ES_n}{n} \xrightarrow{P} 0, \quad n\to\infty$。 □

证法 2. 符号同上，因为

$$\mathrm{Cov}(X_k, X_l) \le \sqrt{\mathrm{Var}X_k}\cdot\sqrt{\mathrm{Var}X_l} = \frac{n-1}{n^2}$$

由证法 1

$$\begin{aligned} P\left(\left|\frac{S_n - ES_n}{n}\right| > \varepsilon\right) &\le \frac{1}{n^2\varepsilon^2}\left(\sum_{i=1}^{n}\mathrm{Var}(X_i) + \sum_{1\le i\ne j\le n}\mathrm{Cov}(X_i, X_j)\right) \\ &\le \frac{1}{n^2\varepsilon^2}\left(n\cdot\frac{n-1}{n^2} + n(n-1)\cdot\frac{n-1}{n^2}\right) \\ &= \frac{n-1}{n^2\varepsilon^2} \to 0, \quad n\to\infty. \end{aligned}$$

结论成立。 □

证法 3. 仍然是对 S_n 用 Chebyshev 不等式。由于 $0\le S_n\le n$ 恒成立，所以

$$\mathrm{Var}(S_n) \le ES_n^2 \le nE(S_n) = n,$$

因此

$$\frac{\mathrm{Var}(S_n)}{n^2} \le \frac{1}{n} \to 0, \quad n\to+\infty.$$

□

证法 4. 由于 $ES_n = 1$，所以

$$\frac{ES_n}{n} = \frac{1}{n} \to 0, \quad n\to\infty.$$

由于 $S_n\ge 0$，所以对任何 $\varepsilon>0$，由 Markov 不等式

$$P\left(\left|\frac{S_n}{n}\right| > \varepsilon\right) = P\left(\frac{S_n}{n} > \varepsilon\right) \le \frac{1}{\varepsilon}E\left(\frac{S_n}{n}\right) = \frac{1}{n\varepsilon} \to 0, \quad n\to\infty.$$

所以

$$\frac{S_n - ES_n}{n} \xrightarrow{P} 0, \quad n\to+\infty.$$

□

题 6 一报贩发现每个路过他的报摊的行人向他买报的概率为 $\frac{1}{3}$ 。求在他出售了 100 份报纸时的过路人的数目在 280 人到 320 人之间的概率。（用两种不同的估计方法，并比较它们的优劣）

解法 1. 记

$$Y_n = \begin{cases} 1, & \text{若第 } n \text{ 个人买报纸,} \\ 0, & \text{否则} \end{cases}$$

于是 $Y_1, Y_2, \ldots, Y_n, \ldots$ 独立同分布，服从 Bernoulli 分布 $B(1, 1/3)$。
如果第 n 个行人买了第 100 份报纸，则

$$Y_1 + Y_2 + \cdots + Y_n = 100.$$

显然 $n \geq 100$（因为每个行人只买一份报纸）。根据中心极限定理，

$$\frac{Y_1 + \cdots + Y_n - \frac{n}{3}}{\sqrt{n \cdot \frac{1}{3}\left(1 - \frac{1}{3}\right)}} = \frac{3(Y_1 + \cdots + Y_n) - n}{\sqrt{2n}}$$

近似服从标准正态分布 $N(0, 1)$。
我们关心 $280 \leq n \leq 320$ 的概率。易见 $\frac{300-n}{\sqrt{2n}}$ 关于 n 是严格减函数，所以

$$\begin{aligned} P\left(\frac{300-320}{\sqrt{640}} \leq \frac{3(Y_1 + \cdots + Y_n) - n}{\sqrt{2n}} = \frac{300-n}{\sqrt{2n}} \leq \frac{300-280}{\sqrt{560}}\right) &\approx \Phi\left(\frac{5}{\sqrt{35}}\right) - \Phi\left(-\frac{5}{\sqrt{40}}\right) \\ &\approx \Phi(0.845) + \Phi(0.791) - 1 \\ &= 0.8009 + 0.7852 - 1 \\ &= 0.5930 \end{aligned}$$

曾有不少学生是这样解答的，但是它是有问题的！

问题在哪里呢？
首先，

$$Y_1 + \cdots + Y_n = 100$$

并不能说明第 n 个人买了第 100 份报纸，因为 n 可以是从卖出第 100 份报纸开始到卖出第 101 份报纸之前任何一个行人的序号。实际上，买到第 100 份报纸的行人的序号应该是

$$N = \min\{n : Y_1 + \cdots + Y_n = 100\},$$

它是个随机变量。
其次，中心极限定理说对任何充分大但是确定的整数 n，$\frac{Y_1+\cdots+Y_n-nEY}{\sqrt{n\mathrm{Var}Y}}$ 近似服从标准正态分布。但当 N 是随机变量时，即使我们知道 $N \geq 100$，中心极限定理并不能告诉我们 $\frac{Y_1+\cdots+Y_N-N\cdot EY}{\sqrt{N\mathrm{Var}Y}}$ 是否具有同样的性质。
好，我们来看看正确的解答。由 N 的定义，我们知道，事件 $\{N > m\}$ 就是“前 m 个人都没有买到第 100 份报纸”，也就是

$$Y_1 + \cdots + Y_m < 100.$$

由中心极限定理，

$$U_n = \frac{Y_1 + \cdots + Y_n - \frac{n}{3}}{\sqrt{\frac{2n}{9}}} = \frac{3(Y_1 + \ldots + Y_n) - n}{\sqrt{2n}} \dot{\sim} N(0, 1),$$

注意到

$$\{280 \leq N \leq 320\} = \{N > 279\} - \{N > 320\},$$

所以

$$\begin{aligned}
&P(280 \le N \le 320)\\
=&P\left(Y_1+Y_2+\cdots+Y_{279}<100\right)-P\left(Y_1+\cdots+Y_{320}<100\right)\\
=&P\left(U_{279}<\frac{99.5-93}{\sqrt{62}}\right)-P\left(U_{320}<-\frac{3\times 99.5-320}{8\sqrt{10}}\right)\quad \text{这里为什么做 0.5 修正？}\\
\approx&\Phi\left(0.826\right)-\Phi\left(-0.850\right)\\
\approx&0.5981.
\end{aligned}$$

□

解法 2. “买，还是不买，这的确是个问题。”每个路过报摊的行人都在**相互独立地重复着**这样的抉择，至少在报贩看来大致是这样的。因此和每个早上一样，他面临着这样的 Bernoulli 试验。

他今天突然有了雅兴，想了解第 k 张报纸到底是卖给了第几个路经报摊的行人，而 Pascal 先生骄傲地告诉他，那个行人的序号 X_k 服从的正是用自己的名字命名的概率分布，而那些 Newton 迷们却宁愿叫它负二项分布。报贩知道，在这样的 Bernoulli 试验中相继卖出任何两份报纸之间，路过的行人数 $Z_1, Z_2, \ldots, Z_k, \ldots$ 相互独立，都服从参数为 1/3 的几何分布，并且 $X_k=\sum_{j=1}^{k} Z_j$。于是 X_{100} 服从参数为 100、1/3 的 Pascal 分布（也就是 Newton 迷们称呼的负二项分布）。

忙于照看生意的报贩显然没有时间去计算 $P(280 \le Z_1+\cdots+Z_{100} \le 320)$ 所涉及的 Pascal 分布中 41 项概率值的和（更何况他 压根$_{\text{儿}}$ 就忘记了 Pascal 写给他的那张纸条被放在哪了）。于是他想到了上回买报的 Chebyshev 先生对他讲的那个不等式，据说不知道分布也可以估计概率的大小。由于

$$E(Z_1+\cdots+Z_{100})=100\times 3=300,\quad \mathrm{Var}(Z_1+\cdots+Z_{100})=100\times 3^2\times\frac{2}{3}=600,$$

所以用 Chebyshev 不等式估计上述概率，那么

$$\begin{aligned}
&P\left(280 \le Z_1+Z_2+\cdots+Z_{100} \le 320\right)\\
=&1-P\left(\left|Z_1+\cdots+Z_{100}-E(Z_1+\cdots+Z_{100})\right| \ge 20\right)\\
\ge&1-\frac{\mathrm{Var}(Z_1+\cdots+Z_{100})}{20^2}=1-\frac{6}{4}=-\frac{1}{2},
\end{aligned}$$

这的确太让人失望了：即使不做任何计算也不至于得出这么糟糕的估计吧。报贩有些心灰意冷。突然，他想起了上周报纸上的头条新闻“Laplace 爵士发现了中心极限定理”，而自己遇到的问题刚好符合了爵士那个定理的要求，于是

$$\begin{aligned}
&P\left(280 \le Z_1+Z_2+\cdots+Z_{100} \le 320\right)\\
=&P\left(\frac{280-300}{\sqrt{600}} \le \frac{Z_1+\cdots+Z_{100}-E(Z_1+\cdots+Z_{100})}{\sqrt{\mathrm{Var}(X_1+\cdots+X_{100})}} \le \frac{320-300}{\sqrt{600}}\right)\\
=&2\Phi\left(\frac{20.5}{\sqrt{600}}\right)-1=2\Phi(0.837)-1=0.5973.
\end{aligned}$$

（实际上，当他料理完一天的生意，却在零钱袋里看到了 Pascal 留给他的那张字条，他静下心来拿着自己捡来的那个计算器用 Pascal 分布进行了计算，费了半天的劲才发现上述概率值为 0.59786，对爵士的钦佩在内心里油然而生。） □