# 概率论与数理统计第四次习题课题目解答

**题1** 随机变量$Z$服从区间$[0,2\pi]$上的均匀分布。求$X=\sin Z, Y=\cos Z$的相关系数$\mathrm{Corr}(X,Y)$，并判断$X$与$Y$的独立性。

<u>解</u>:

$$EX=E(\sin Z)=\int_0^{2\pi}\sin z\cdot\frac{1}{2\pi}dz=0,$$

$$\mathrm{Var}X=EX^2-(EX)^2=EX^2=E(\sin^2 Z)=\int_0^{2\pi}\sin^2 z\cdot\frac{1}{2\pi}dz=\int_0^{2\pi}\frac{1-\cos 2z}{2}\cdot\frac{1}{2\pi}dz=\frac{1}{2},$$

同理可得

$$EY=0,\qquad \mathrm{Var}Y=\frac{1}{2}.$$

再由

$$E(XY)=E(\sin Z\cos Z)=\int_0^{2\pi}\sin z\cos z\cdot\frac{1}{2\pi}dz=0$$

知$\mathrm{Cov}(X,Y)=0$，从而

$$\mathrm{Corr}(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}X\mathrm{Var}Y}}=0,$$

因此$X,Y$不相关。

但是$X,Y$不独立，这是因为

$$P(|X|<0.5,|Y|<0.5)\le P(X^2+Y^2<0.5)=0<P(|X|<0.5)\cdot P(|Y|<0.5).$$

在这个问题中，虽然$X,Y$各自有概率密度，但由于$(X,Y)$只能出现在单位圆周

$$S^1=\{(x,y)|x^2+y^2=1\}$$

上，所以它们没有联合概率密度。而如果$X,Y$独立，则$p_X(x)p_Y(y)$就是$(X,Y)$的联合概率密度函数，这与$(X,Y)$无联合概率密度的事实矛盾，所以$X,Y$不独立。

于是，我们得到不相关但不独立的一个实例。

**题2** 设$(X_1,X_2,\dots,X_r)$服从多项分布$M(n,p_1,p_2,\dots,p_r)$（$p_i>0, i=1,\dots,r$且$p_1+\dots+p_r=1$）。求$X_i$与$X_j$的相关系数，$i,j=1,\dots,r$。

**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

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

$$\begin{aligned}
E(X_iX_j) &= \sum_{k\geq 0, l\geq 0, k+l\leq n} kl \cdot P(X_i = k, X_j = l) \\
&= \sum_{k\geq 0, l\geq 0, k+l\leq n} kl \cdot \frac{n!}{k!l!(n-k-l)!} p_i^k p_j^l (1 - p_i - p_j)^{n-k-l} \\
&= n(n-1)p_ip_j \sum_{k\geq 1, l\geq 1, k+l\leq n} \frac{(n-2)!}{(k-1)!(l-1)!(n-k-l)!} p_i^{k-1} p_j^{-1-1} (1 - p_i - p_j)^{n-k-l}
\end{aligned}$$

（注意这个求和的概率含义）

$$= n(n-1)p_ip_j.$$

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

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

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$i_1, \dots, i_n \in \{1, \dots, r\}$，$Y_1^{(i_1)}, \dots, Y_n^{(i_n)}$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，

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

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

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

所以，

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

**解法4：** 首先，$X_i\sim B(n,p_i)$，$X_j\sim B(n,p_j)$。

在已知$X_j=k$的条件下，$X_i\sim B\left(n-k,\frac{p_i}{1-p_j}\right)$。所以

$$E(X_i|X_j=k)=\frac{(n-k)p_i}{1-p_j}.$$

所以，

$$E(X_i|X_j)=\frac{(n-X_j)p_i}{1-p_j},$$

从而

$$\begin{aligned}E(X_iX_j)&=E[E(X_iX_j|X_j)]=E[X_jE(X_i|X_j)]\\&=E\left[\frac{X_j(n-X_j)p_i}{1-p_j}\right]\\&=\frac{np_i}{1-p_j}EX_j-\frac{p_i}{1-p_j}EX_j^2\\&=\frac{n^2p_ip_j}{1-p_j}-\frac{p_i(np_j-np_j^2+n^2p_j^2)}{1-p_j}\\&=n^2p_ip_j+np_ip_j.\end{aligned}$$

因此，

$$\mathrm{Cov}(X_i,X_j)=E(X_iX_j)-EX_i\cdot EX_j=np_ip_j,$$
$$\mathrm{Corr}(X_i,X_j)=\frac{\mathrm{Cov}(X_i,X_j)}{\sqrt{\mathrm{Var}X_i\mathrm{Var}X_j}}=-\sqrt{\frac{p_ip_j}{(1-p_i)(1-p_j)}}.$$

**题3** 设$(X,\Lambda)$的概率分布为：$\Lambda$的边缘分布密度为$\frac{1}{\lambda^2}e^{-\beta/\lambda}$，$\lambda>0$，其中$\beta$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$\Lambda$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$\Lambda$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$(X, \Lambda)$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$EX$

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

这里交换累次积分的顺序是关键。

而由$\Lambda$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4** 设$(X, Y)$服从二维正态分布$N(0, 0, 1, 1, \rho)$。求$\max(X, Y)$的数学期望和方差。

解：令

$$U = \frac{X+Y}{2}, \qquad V = \frac{X-Y}{2},$$

则

$$\max(X, Y) = U + |V|,$$

而

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}$$

于是，$(U,V)$服从二维正态，并且

$$
\begin{aligned}
EU &= \frac{EX+EY}{2} = 0, \\
EV &= \frac{EX-EY}{2} = 0, \\
\mathrm{Var}U &= \frac{1}{4}\left[\mathrm{Var}X + \mathrm{Var}Y + 2\mathrm{Cov}(X,Y)\right] = \frac{1+\rho}{2}, \\
\mathrm{Var}V &= \frac{1-\rho}{2}, \\
\mathrm{Cov}(U,V) &= \mathrm{Cov}\left(\frac{X+Y}{2}, \frac{X-Y}{2}\right) = \frac{1}{4}\left(\mathrm{Var}X - \mathrm{Var}Y\right) = 0,
\end{aligned}
$$

因此$U,V$相互独立，而且$U \sim N(0,(1+\rho)/2)$，$V \sim N(0,(1-\rho)/2)$。于是

$$Z = \sqrt{\frac{2}{1-\rho}}V \sim N(0,1),$$

所以，

$$
\begin{aligned}
E[\max(X,Y)] &= EU + E|V| = 0 + \sqrt{\frac{1-\rho}{2}}E|Z| \\
&= \sqrt{\frac{1-\rho}{2}}\int_{-\infty}^{\infty}|t|\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt = \sqrt{\frac{1-\rho}{\pi}}, \\
\mathrm{Var}[\max(X,Y)] &= \mathrm{Var}U + \mathrm{Var}|V| = \frac{1+\rho}{2} + \frac{1-\rho}{2}\mathrm{Var}|Z| \\
&= \frac{1+\rho}{2} + \frac{1-\rho}{2}\left(1-\frac{2}{\pi}\right) = \frac{1}{2} - \frac{1-\rho}{\pi},
\end{aligned}
$$

**题5** 设一对夫妻购买了一项生命年金，支付方式为：当夫妻两人中有一人去世时开始支付，而当另一人也去世时停止支付。记年金开始支付时间为$X$，停止支付时间为$Y$。设这对夫妻在购买年金后的存活年限相互独立，都服从指数分布$\mathrm{Exp}(\lambda)$。

1. 求$X$对$Y$的线性回归（即用$X$的一次函数对$Y$作最小二乘最佳逼近）；
   **解：** 设夫妻二人的存活年限分别为$U,V$，则$U,V \overset{i.i.d.}{\sim} \mathrm{Exp}(\lambda)$，$X = \min\{U,V\}$，$Y = \max\{U,V\}$。于是$X \sim \mathrm{Exp}(2\lambda)$，从而

$$EX = \frac{1}{2\lambda}, \quad \mathrm{Var}X = \frac{1}{4\lambda^2}.$$

从而

$$EY = E(X+Y) - EX = E(U+V) - EX = EU + EV - EX = \frac{1}{\lambda} + \frac{1}{\lambda} - \frac{1}{2\lambda} = \frac{3}{2\lambda},$$

$$E(XY) = E(UV) = EU \cdot EV = \frac{1}{\lambda^2},$$

$$\mathrm{Cov}(X,Y) = E(XY) - EX \cdot EY = \frac{1}{\lambda^2} - \frac{1}{2\lambda} \cdot \frac{3}{2\lambda} = \frac{1}{4\lambda^2},$$

所以$X$对$Y$线性回归方程为

$$\hat{Y} = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}X}(X - EX) + EY = \left(X - \frac{1}{2\lambda}\right) + \frac{3}{2\lambda} = X + \frac{1}{\lambda}.$$

2. 求$Y$对$X$的线性回归；
**解：** 在第一问计算的基础上，我们只需再求$Y$的方差。利用

$$\mathrm{Var}(X+Y) = \mathrm{Var}(U+V),$$

我们得到

$$\mathrm{Var}X + 2\mathrm{Cov}(X,Y) + \mathrm{Var}Y = \mathrm{Var}U + \mathrm{Var}V,$$

从而

$$\mathrm{Var}Y = \frac{1}{\lambda^2} + \frac{1}{\lambda^2} - \frac{1}{4\lambda^2} - \frac{2}{4\lambda^2} = \frac{5}{4\lambda^2},$$

所以$Y$对$X$线性回归方程为

$$\hat{X} = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}Y}(Y - EY) + EX = \frac{1}{5}\left(Y - \frac{3}{2\lambda}\right) + \frac{1}{2\lambda} = \frac{Y}{5} + \frac{1}{5\lambda}.$$

3. 求$X$的函数对$Y$作最小二乘最佳逼近。
**解：** 我们知道$E(Y|X)$就是最小二乘最佳逼近。我们先计算$X,Y$的联合概率分布函数

$$\begin{aligned} f_{X,Y}(x,y) &= \left( f_{U,V}(x,y)\frac{1}{\left|\det \frac{\partial(x,y)}{\partial(x,y)}\right|} + f_{U,V}(y,x)\frac{1}{\left|\det \frac{\partial(x,y)}{\partial(y,x)}\right|} \right) I_{x\le y} \\ &= 2\lambda^2 e^{-\lambda(x+y)} I_{0<x\le y}, \end{aligned}$$

从而得到$X$的边缘概率密度

$$f_X(x) = I_{x>0}\int_x^{+\infty} 2\lambda^2 e^{-\lambda(x+y)}dy = 2\lambda e^{-2\lambda x} I_{x>0},$$

及条件概率密度函数

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \lambda e^{-\lambda(y-x)} I_{y \geq x > 0}$$

因此

$$E(Y|X=x) = I_{x>0} \int_x^{+\infty} y\lambda e^{-\lambda(y-x)} dy = \left(x + \frac{1}{\lambda}\right) I_{x>0},$$

因此

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

**附注：**类似计算可得

$$E(X|Y) = \frac{-e^{-\lambda Y}}{1-e^{-\lambda Y}} Y + \frac{1}{\lambda}.$$

这说明所谓“非线性最小二乘回归”有可能是真正非线性的，也有可能是线性的。

**题6** 一个家族第$n$代男性子孙有$X_n$个人，$X_0 = 1$。假设这个家族中每个男性成员的儿子的个数是独立同分布的随机变量，求$\lim\limits_{n\to+\infty} P(X_n = 0)$。

**解：** 记$Y_i^{(n)}$为第$n$代中第$i$个人的儿子个数，所以

$$X_{n+1} = \sum_{i=1}^{X_n} Y_i^{(n)},$$

从而由全概率公式，

$$\begin{aligned}
P(X_{n+1} = 0) &= \sum_{j\geq 0} P(X_n = j) P(X_{n+1} = 0 | X_n = j) \\
&= \sum_{j\geq 0} P(X_n = j) P(Y_1^{(n)} + \cdots + Y_j^{(n)} = 0 | X_n = j) \\
&= \sum_{j\geq 0} P(X_n = j) P(Y_1^{(n)} = \cdots = Y_j^{(n)} = 0 | X_n = j) \\
&= \sum_{j\geq 0} P(X_n = j) P(Y_1^{(n)} = 0) \cdots P(Y_j^{(n)} = 0) \\
&= \sum_{j\geq 0} P(X_n = j) \big[P(X_1 = 0)\big]^j \\
&= g_{X_n}\big(P(X_1 = 0)\big).
\end{aligned}$$

其中$g_{X_n}(z)$是$X_n$的概率母函数，所以$P(X_{n+1} = 0)$涉及$X_n$的整个概率分布。因此我们考虑$X_n$的概率母函数的递归关系，利用全期望公式

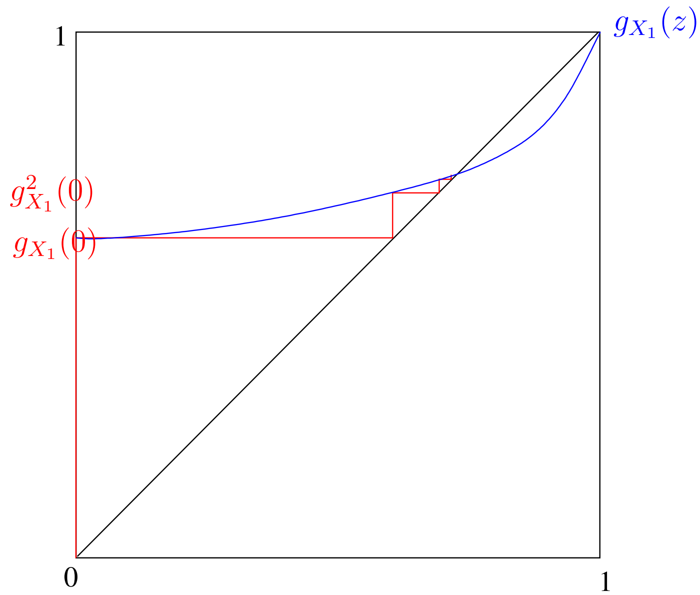
$$
\begin{aligned}
g_{X_{n+1}}(z) &= E\left(z^{X_{n+1}}\right) \\
&= E\left[E\left(z^{X_{n+1}}|X_n\right)\right] \\
&= \sum_{j\geq 0} P(X_n = j) E\left(z^{X_{n+1}}|X_n = j\right) \\
&= \sum_{j\geq 0} P(X_n = j) E\left(z^{Y_1^{(n)}+\cdots+Y_j^{(n)}}|X_n = j\right) \\
&= \sum_{j\geq 0} P(X_n = j) E\left(z^{Y_1^{(n)}+\cdots+Y_j^{(n)}}\right) \\
&= \sum_{j\geq 0} P(X_n = j) E\left(z^{Y_1^{(n)}}\right)\cdots E\left(z^{Y_j^{(n)}}\right) \\
&= \sum_{j\geq 0} P(X_n = j)\left[E\left(z^{X_1}\right)\right]^n \\
&= g_{X_n}\big(g_{X_1}(z)\big),
\end{aligned}
$$

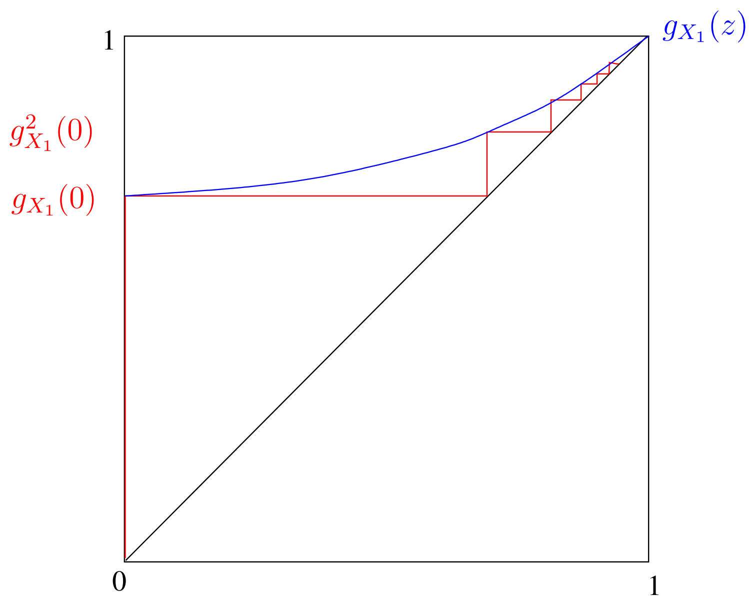
因此

$$
g_{X_n}(z) = g_{X_1}^n(z), \qquad g_{X_1}^n = \underbrace{g_{X_1} \circ \cdots \circ g_{X_1}}_{n},
$$

而且

$$
P(X_n = 0) = g_{X_n}(0) = g_{X_1}^n(0).
$$




由于$g_{X_1}(z) = \sum\limits_{k\geq 0} P(X_1 = k)z^k$在区间$[0,1]$上是单调不减函数列$z^n$的非负系数线性组合，所以它也是单调不减函数，于是对任何$z \in [0,1]$，

$$0 \leq P(X_1 = 0) = g_{X_1}(0) \leq g_{X_1}(z) \leq g_{X_1}(1) = 1,$$

所以$\{g^n_{X_1}(0)\}$是单调不减有上界的实数列，它收敛于$g_{X_1}$在区间$[0,1]$中最左边的那个不动点。

又由于$g_{X_1}(z)=\sum\limits_{k\geq 0}P(X_1=k)z^k$是凸函数序列$z^n$的凸组合，所以$g_{X_1}(z)$是凸函数，它在区间$[0,1]$上有且仅有以下三种可能

- $[0,1]$上到处都是$g_{X_1}$的不动点，即$g_{X_1}(z)\equiv z$。这当且仅当$P(X_1=1)=1$。在这种情形下，$P(X_n=0)=0$，所以这个家族任何一代中都会有男性成员。
- 恰好有两个不动点$x^*$（$0\leq x^*<1$）和1。这当且仅当$P(X_1=1)<1$且$EX_1>1$（$EX_1$是$g_{X_1}$在$z=1$处的切线斜率）。在这种情形下，$\lim\limits_{n\to+\infty}P(X_n=0)=x^*$，这时这个家族有$1-x^*$的概率可以保证任何一代中都有男性成员。
- 1是唯一不动点。这当且仅当$P(X_1=1)<1$且$EX_1\leq 1$。在这种情形下，$\lim\limits_{n\to+\infty}P(X_n=0)=1$，即这个家族几乎注定会断了香火。

**题7** 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：

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

<u>证法1</u>. 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$\varepsilon>0$

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

而

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

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

<u>证法2</u>. 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。 □

<u>证法3</u>. 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

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

□

<u>证法4</u>. 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8** 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，但是它是有缺陷的！

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$N\geq 100$，但是我们并不能直接有中心极限定理知道$\frac{Y_1+\cdots+Y_N-N\cdot EY}{\sqrt{N\mathrm{Var}Y}}$是否近似服从正态分布$N(0,1)$。

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

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

□

思考题：设$0<p<1$，$r$是正整数，$N_r$服从负二项分布NB$(r,p)$。问：当$r\to+\infty$时，随机变量

$$\frac{r-N_rp}{\sqrt{N_rp(1-p)}}$$

是否依分布收敛到正态分布$N(0,1)$？

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

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

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

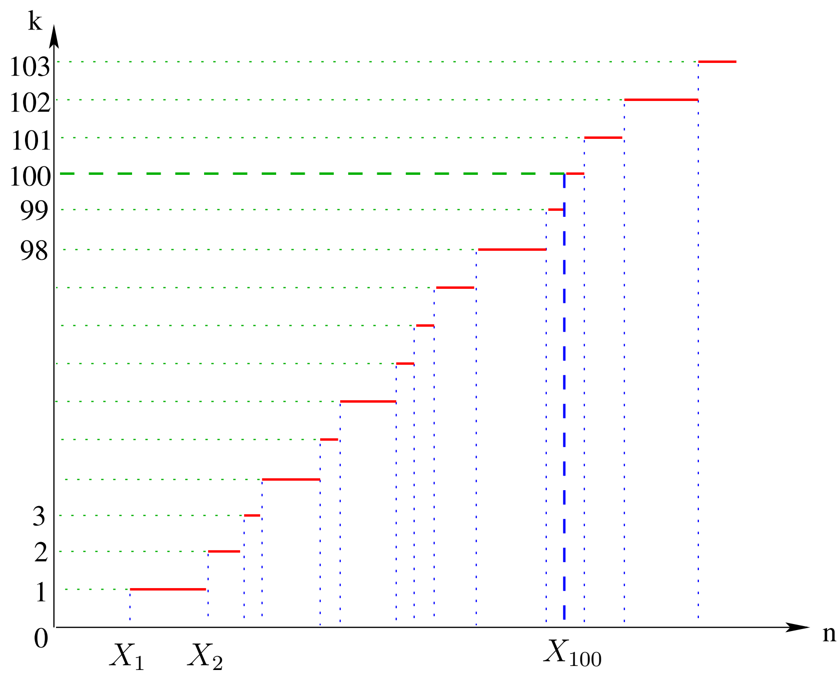
$$\begin{aligned}
&P\left(280 \le Z_1 + Z_2 + \cdots + Z_{100} \le 320\right) \\
=&P\left(\frac{280 - 300}{\sqrt{600}} \le \frac{Z_1 + \cdots + Z_{100} - E(Z_1 + \cdots + Z_{100})}{\sqrt{\mathrm{Var}(X_1 + \cdots + X_{100})}} \le \frac{320 - 300}{\sqrt{600}}\right) \\
=&2\Phi\left(\frac{20.5}{\sqrt{600}}\right)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

（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。） $\square$

**解法1和解法2的关系：** 下图是函数

$$n \mapsto \sum_{j=1}^{n} Y_j$$

的图像，其中横坐标是路过人数，纵坐标是卖出报纸份数。




这是个随机的、单调不减的阶梯函数，它的函数值在$X_k$处从$k-1$跳跃到$k$。于是

$$X_k = \min\left\{ n \middle| \sum_{j=1}^{n} Y_j = k \right\}$$

因此$X_{100} \geq 280$（也就是$X_{100} > 279$）等价于$\sum_{j=1}^{279} Y_j < 100$。对$X_{100} \leq 320$，我们考虑其对立事件$X_{100} > 320$，它等价于$\sum_{j=1}^{320} Y_j < 100$。

由于$Y_j$的非负性，$\sum_{j=1}^{320} Y_j < 100$是$\sum_{j=1}^{279} Y_j < 100$的一个子事件。