第四章 大数定律与中心极限定理

1.(Ch4-1) 试利用切比雪夫不等式证明:能以 0.97 的概率断言,将一枚均匀硬币连续抛 1000 次，其出现正面H的次数在 400 至 600 次之间。

分析：将一枚均匀硬币连续抛 1000 次可看成是 1000 重贝努利试验，因此 1000 次试验中出现正面H的次数服从二项分布。

解：设X表示 1000 次试验中出现正面H的次数，则X是一个随机变量，且 $X \sim B$(1000，1/2)。因此

$$EX = np = 1000 \times \frac{1}{2} = 500,$$

$$DX = np(1-p) = 1000 \times \frac{1}{2} \times (1-\frac{1}{2}) = 250,$$

而所求的概率为

$$\begin{aligned} P\{400 < X < 600\} &= P\{400-500 < X < 600-500\} \\ &= P\{-100 < X - EX < 100\} \\ &= P\{|X - EX| < 100\} \\ &\geq 1 - \frac{DX}{100^2} = 0.975。\end{aligned}$$

2.(Ch4-2) 已知随机变量X的概率分布为

X	1	2	3
P	0.2	0.3	0.5

试利用切比雪夫不等式估计事件$\{|X - E(X)| < 1.5\}$的概率。

分析：要利用切比雪夫不等式，需先根据给出的随机变量分布列求得相应的期望和方差。

解：由题设知，

$$EX = 1\times 0.2 + 2\times 0.3 + 3\times 0.5 = 2.3,$$

$$EX^2 = 1^2 \times 0.2 + 2^2 \times 0.3 + 3^2 \times 0.5 = 5.9。$$

从而 $$DX = EX^2 - (EX)^2 = 5.9 - 2.3^2 = 0.61。$$

由切比雪夫不等式得

$$P\{|X - EX| < 1.5\} \geq 1 - \frac{DX}{1.5^2} \approx 0.729。$$

3.(Ch4-3) 设X为非负随机变量，试证；当$t > 0$时，

$$P(X < t) \geq 1 - \frac{EX}{t}。$$

分析：$P\{X < t\} = F(t) = \int_{-\infty}^{t} f(x)dx$，而$EX = \int_{-\infty}^{+\infty} xf(x)dx$，代入要证的不等

式的两侧比较，会发现证明实质上是对积分限的放大或缩小，以及变量间暗含的大小关系，很容易就联系到对切比雪夫不等式的证明技巧。

证明：设随机变量 X 的分布密度函数为 $f(x)$，则当 $t>0$ 时，

$$\begin{aligned} P\{X<t\} &= \int_{-\infty}^{t} f(x)dx = 1-\int_{t}^{+\infty} f(x)dx \\ &\geq 1-\int_{t}^{+\infty}\frac{x}{t} f(x)dx = 1-\frac{1}{t}\int_{t}^{+\infty} xf(x)dx \\ &\geq 1-\frac{1}{t}\int_{-\infty}^{+\infty} xf(x)dx = 1-\frac{1}{t}EX \text{。} \end{aligned}$$

4.（Ch4-4）设 $X_1, X_2, \cdots, X_n, \cdots$ 为一列独立同分布的随机变量，且 k 阶原点矩存在，记作 $EX^k=\mu_k$，。试证明：$\frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow{p} \mu_k$。

分析：由题设条件 $X_1, X_2, \cdots, X_n, \cdots$ 为一列独立同分布的随机变量，以及 $E(\frac{1}{n}\sum_{i=1}^{n} X_i^k)=\frac{1}{n}\sum_{i=1}^{n} EX_i^k=\frac{1}{n}\cdot n\mu_k=\mu_k$，可见所证结论与辛钦大数定律的结论非常类似，即知证明应用独立同分布的辛钦大数定律。

证明：由 $X_1, X_2, \cdots, X_n$ 为一列独立同分布的随机变量，以及 $y=x^k$ 是连续函数知，$X_1^k, X_2^k, \cdots, X_n^k$ 相互独立。再由 $EX^k=\mu_k$，得

$$E(\frac{1}{n}\sum_{i=1}^{n} X_i^k)=\frac{1}{n}\sum_{i=1}^{n} EX_i^k=\frac{1}{n}\cdot n\mu_k=\mu_k \text{，}$$

则由辛钦大数定律知：$\frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow{p} \mu_k$。

5.（Ch4-5）在一家保险公司里 10000 个人参加保险，每人每年付 12 元保险费，在一年内一个人死亡的概率为 0.006，死亡者家属可向保险公司领得 1000 元。问：

（1）保险公司亏本的概率多大？

（2）保险公司一年的利润不少于 40000 元的概率多大？

分析：对于每个人，在一年内要么死亡，要么不死亡，只有这两种可能性，因此考虑 10000 个人在一年中是否死亡可看成 10000 重贝努利试验，故死亡人数服从二项分布。因此应用棣莫弗-拉普拉斯极限定理解决该问题。

解：设一年中死亡的人数为 X，每人的死亡概率就为 $p=0.006$，从而

$$X \sim B(10000,\quad 0.006) \text{，}$$

保险公司每年收入 $10000\times 12=120000$ 元，需支付 $1000X$ 元。

（1）设 A：“保险公司亏本”，则有

$P(A)=P\{1000X>120000\}=P\{X>120\}$

$$=1-P\{0<X<120\}\approx 1-\left[\Phi(\frac{120-np}{\sqrt{npq}})-\Phi(\frac{0-np}{\sqrt{npq}})\right]$$

$$=1-\left[\Phi(\frac{120-10000\times 0.006}{\sqrt{10000\times 0.006\times(1-0.006)}})-\Phi(\frac{0-10000\times 0.006}{\sqrt{10000\times 0.006\times(1-0.006)}})\right]$$

$$=1-\left[\Phi(7.7693)-\Phi(-7.7693)\right]=2-2\Phi(7.7693)$$

$$\approx 2\text{-}2=0$$

可见保险公司亏本的概率近似为零。

（2）设 B ："保险公司一年中获利不少于 40000 元"，则

$$P(B)=P\{120000-1000X\geq 40000\}=P\{0\leq X\leq 80\}$$

$$\approx\Phi(\frac{80-np}{\sqrt{npq}})-\Phi(\frac{0-np}{\sqrt{pq}})$$

$$\approx\Phi(\frac{80-10000\times 0.006}{\sqrt{10000\times 0.006\times(1-0.006)}})-\Phi(\frac{0-10000\times 0.006}{\sqrt{10000\times 0.006\times(1-0.006)}})$$

$$=\Phi(2.59)-\Phi(-7.7693)=\Phi(2.59)-(1-\Phi(7.7693))$$

$$\approx 0.9952\text{。}$$

即一年中保险公司以近 99.52%的概率获利 40000 元以上。

6.(Ch4-6) 100 道单项选择题，每题 1 分，考生每次从四个答案中选一个正确答案。若一考生全为乱猜，试用切比雪夫不等式和正态逼近两种方法计算其成绩 15 分至 35 分之间的概率约为多少？

解：设 X 表示考生成绩（选对个数），则 X 服从二项分布 $B(100,1/4)$，由切比雪夫不等式

$$P\{15<X<35\}=P\{|X-25|<10\}\geq 1-\frac{DX}{100}\text{，}$$

由于 $EX=25$， $DX=75/4$，所以

$$P\{15<X<35\}\geq 1-\frac{DX}{100}=1-\frac{75/4}{100}=0.8125\text{。}$$

正态逼近法

$$P\{15<X<35\}=\Phi(\frac{35-25}{\sqrt{75/4}})-\Phi(\frac{15-25}{\sqrt{75/4}})$$

$$\approx\Phi(2.31)-\Phi(-2.31)$$

$$=2\Phi(2.31)-1=0.9792\text{。}$$

7．设在开关电路的试验中，每次试验开或关的概率为 1/2，欲使开的频率 $f_n(A)=n(A)/n$ 与概率 1/2 的差的绝对值小于 0.01 有 99%的可靠性，试问试验次数 n 应取多少？

分析：依题意，问题化为 n 取多少才能使 $P\{\left|\frac{n(A)}{n}-\frac{1}{2}\right|<0.01\}=0.99$ 成立。

解：由切比雪夫不等式知 $P\{\left|\frac{n(A)}{n}-p\right|<\varepsilon\}\ge 1-\frac{D(\frac{n(A)}{n})}{\varepsilon^2}$，

而 $D(\frac{n(A)}{n})=\frac{D(n(A))}{n^2}=\frac{np(1-p)}{n^2}=\frac{p(1-p)}{n}$，取 $\varepsilon=0.01$，$p=\frac{1}{2}$ 代入上式，得

$$P\{\left|\frac{n(A)}{n}-p\right|<\varepsilon\}=P\{\left|\frac{n(A)}{n}-\frac{1}{2}\right|<0.01\}\ge 1-\frac{1}{4\times 10^{-4}\cdot n},$$

欲使上式左边大于 0.99，只须使 $1-\frac{1}{4\times 10^{-4}\cdot n}>0.99$。从而解得

$$n>\frac{1}{4\times 10^{-6}}=250000。$$

所以至少要做 250000 次试验才能满足要求。

8．(Ch4-7) 某厂有 400 台同类机器，各台机器发生故障的概率均为 0.02，假设各台机器工作是相互独立的，试求机器发生故障的台数不小于 2 的概率。

解：设 X 为机器发生故障的台数，则由题意知 $X\sim B(400，0.02)$，问题化为求 $P\{X\ge 2\}$。以下用三种方法来求解：

(1) 利用二项分布

$$P\{X\ge 2\}=1-P\{X<2\}=1-0.98^{400}-C_{400}^{1}\times 0.02\times 0.98^{399}\approx 0.9972。$$

(2) 用泊松分布作近似计算（此时 $\lambda=np=400\times 0.02=8$）

$$P\{X\ge 2\}=1-P\{X<2\}\approx 1-e^{-8}(1+8)=1-9e^{-8}=0.9970。$$

(3) 用正态分布作近似计算（利用定理 4-5 及 4-4 的推论 1）

由于 $X\sim B(400，0.02)$，则由定理 4-4 的推论 1 知

$$X\overset{\text{近似}}{\sim}N(n\mu，n\sigma^2)=N(8，400\times 0.98\times 0.02)=N(8，2.8^2)$$

于是

$$P\{X\ge 2\}=1-P\{0\le X<2\}=1-P\{\frac{0-8}{2.8}\le\frac{X-8}{2.8}<\frac{2-8}{2.8}\}$$

$$\approx 1-[\Phi(\frac{-6}{2.8})-\Phi(\frac{-8}{2.8})]=0.9859。$$

9．(Ch4-8) 假设 $X_1，X_2，\cdots，X_n$ 是来自总体 X 的简单随机抽样，已知 $EX^k=\alpha_k\ (k=1，2，3，4)$，证明当 n 充分大时，随机变量 $Z_n=\frac{1}{n}\sum_{i=1}^{n}X_i^2$ 近似服从正态分布，并指出其分布参数。

证明：由假设条件可知，$X_1^{\,2}，X_2^{\,2}，\cdots，X_n^{\,2}$ 为来自总体 X^2 的简单随机抽

样，则 $X_1{}^2$，$X_2{}^2$，…，$X_n{}^2$ 相互独立且与 X^2 同分布，即 $EX_i^2=\alpha_2(i=1,2,\cdots,n)$，$DX_i^2=E(X_i^2)^2-(EX_i^2)^2=\alpha_4-\alpha_2^2$，则由独立同分布的中心极限定理 $\forall x\in R$ 有

$$\lim_{n\to\infty}P\{\frac{\sum_{i=1}^{n}X_i^2-n\alpha_2}{\sqrt{n}\sqrt{\alpha_4-\alpha_2^2}}\le x\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}dt$$

即 $\frac{\sum_{i=1}^{n}X_i^2-n\alpha_2}{\sqrt{n}\sqrt{\alpha_4-\alpha_2^2}}\overset{近似}{\sim}N$（0，1），$\frac{\frac{1}{n}\sum_{i=1}^{n}X_i^2-\alpha_2}{\sqrt{(\alpha_4-\alpha_2^2)/n}}\overset{近似}{\sim}N$（0，1），所以当 n 充分大时，$Z_n=\frac{1}{n}\sum_{i=1}^{n}X_i^2$ 近似服从参数为（α_2，$\frac{\alpha_4-\alpha_2^2}{n}$）的正态分布。

10．假设 X_1，X_2，…，X_n 是来自总体 X 的简单随机抽样，已知 $EX^k=\alpha_k$ $(k=1，2，3，4)$，证明当 n 充分大时，随机变量 $Z_n=\frac{1}{n}\sum_{i=1}^{n}X_i^2$ 近似服从正态分布，并指出其分布参数。

证明：由假设条件可知，$X_1{}^2$，$X_2{}^2$，…，$X_n{}^2$ 为来自总体 X^2 的简单随机抽样，则 $X_1{}^2$，$X_2{}^2$，…，$X_n{}^2$ 相互独立且与 X^2 同分布，即 $EX_i^2=\alpha_2(i=1,2,\cdots,n)$，$DX_i^2=E(X_i^2)^2-(EX_i^2)^2=\alpha_4-\alpha_2^2$，则由独立同分布的中心极限定理 $\forall x\in R$ 有

$$\lim_{n\to\infty}P\{\frac{\sum_{i=1}^{n}X_i^2-n\alpha_2}{\sqrt{n}\sqrt{\alpha_4-\alpha_2^2}}\le x\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}dt$$

即 $\frac{\sum_{i=1}^{n}X_i^2-n\alpha_2}{\sqrt{n}\sqrt{\alpha_4-\alpha_2^2}}\overset{近似}{\sim}N$（0，1），$\frac{\frac{1}{n}\sum_{i=1}^{n}X_i^2-\alpha_2}{\sqrt{(\alpha_4-\alpha_2^2)/n}}\overset{近似}{\sim}N$（0，1），所以当 n 充分大时，$Z_n=\frac{1}{n}\sum_{i=1}^{n}X_i^2$ 近似服从参数为（α_2，$\frac{\alpha_4-\alpha_2^2}{n}$）的正态分布。

11．一生产线生产的产品成箱包装，每箱的重量是随机的。假设每箱平均重 50 千克，标准差为 5 千克。若用最大载重量为 5 吨的汽车承运，试利用中心极限定理说明每辆车最多可以装多少箱，才能保障不超载的概率大于 0.977。（$\Phi(2)=0.977$，其中 $\Phi(x)$ 是标准正态分布函数）

解：设 $X_i(i=1,2,\cdots,n)$ 是装运的第 i 箱的重量（单位：千克），n 是所求箱数。由条件可以把 $X_1,X_2,\cdots,X_n$ 视为独立同分布随机变量，而 n 箱的总重量

$$T_n=X_1+X_2+\cdots+X_n$$

是独立同分布随机变量之和。

由条件知 $EX_i=50$，$\sqrt{DX_i}=5$，$ET_n=50n$，$\sqrt{DT_n}=5\sqrt{n}$（单位：千克）。

根据列维—林德伯格中心极限定理，T_n 近似服从正态分布 $N(50n,\quad 25n)$。箱数 n 决定于条件

$$P\{T_n \le 5000\} = P\{\frac{T_n - 50n}{5\sqrt{n}} \le \frac{5000 - 50n}{5\sqrt{n}}\}$$

$$\approx \Phi(\frac{1000 - 10n}{\sqrt{n}}) > 0.977 = \Phi(2)\text{。}$$

由此可见 $\frac{1000 - 10n}{\sqrt{n}} > 2$，从而 $n < 98.0199$，即最多可以装 98 箱。