第五章
大数定律及中心极限定理

大数定律的客观背景

大量的随机现象中平均结果的稳定性

大量抛掷硬币
正面出现频率

生产过程中
的废品率

字母使用频率

……

思考：频率是概率的反映，随着观察的次数增加，频率将会“逐渐稳定”或“靠近”到概率，“逐渐稳定”或“靠近”到概率的原因是什么？

§5.1 大数定律

定义1 设$Y_1, Y_2, \ldots, Y_n, \ldots$为一随机变量序列，$a$是常数，若对任意正数ε，有

$$\lim_{n\to\infty} P\{|Y_n - a| < \varepsilon\} = 1$$

则称随机变量序列$Y_1, Y_2, \ldots, Y_n, \ldots$ **依概率收敛于a**，记为：

$$Y_n \xrightarrow{P} a$$

请注意:

$\{Y_n\}$依概率收敛于Y，意味着对任意给定的$\varepsilon>0$，当n充分大时，事件$|Y_n-Y|<\varepsilon$ 的概率很大，接近于1；并不排除事件$|Y_n-Y|\geq\varepsilon$ 的发生，而只是说它发生的可能性很小.

依概率收敛比高等数学中的普通意义下的收敛弱些，它具有某种不确定性.

依概率收敛序列的性质:

设 $X_n \xrightarrow{P} a$，$Y_n \xrightarrow{P} b$,

又设函数 $g(x,y)$在点(a,b)连续，则

$$g(X_n,Y_n) \xrightarrow{P} g(a,b)$$

证明　因为$g(x,y)$在(a,b)连续,

$\forall \varepsilon > 0$, $\exists \delta > 0$,

使得当$|x-a|+|y-b|<\delta$时,

$|g(x,y)-g(a,b)|<\varepsilon$,

于是 $\{|g(X_n,Y_n)-g(a,b)|\geq\varepsilon\}$

$$\subset\{|X_n-a|+|Y_n-b|\geq\delta\}$$

$$\subset\left\{|X_n-a|\geq\frac{\delta}{2}\right\}\cup\left\{|Y_n-b|\geq\frac{\delta}{2}\right\},$$

因此 $P\{|g(X_n,Y_n)-g(a,b)|\geq\varepsilon\}$

$$\leq P\left\{|X_n-a|\geq\frac{\delta}{2}\right\}+P\left\{|Y_n-b|\geq\frac{\delta}{2}\right\}\xrightarrow{n\to\infty}0,$$

故 $\lim\limits_{n\to\infty}P\{|g(X_n,Y_n)-g(a,b)|<\varepsilon\}=1.$ [证毕]

定理1（切比雪夫定理的特殊情况）设随机变量序列 $X_1, X_2, \ldots, X_n, \ldots$相互独立，且具有相同的数学期望和方差：$E(X_k)=\mu$，$D(X_k)=\sigma^2$ $(k=1,2,\ldots)$，则对任意的 $\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| < \varepsilon\right\} = 1$$

即

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu$$

提示：利用切比雪夫不等式证.

此定理表明: 相互独立具有相同期望和方差的随机变量 $X_1, X_2, \ldots, X_n$ 的算术平均值依概率收敛于其数学期望值 μ .

证 $E(\overline{X})=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\mu$

$$D(\overline{X})=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{\sigma^2}{n}$$

由切比雪夫不等式

$$P\{|\overline{X}-E(\overline{X})|<\varepsilon\}\geq 1-\frac{D(\overline{X})}{\varepsilon^2}$$

即 $P\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right|<\varepsilon\}\geq 1-\frac{\sigma^2}{n\varepsilon^2}$

$$\Longrightarrow \lim_{n\to\infty}P\{|\overline{X}-\mu|<\varepsilon\}=1$$

定理2 （贝努力大数定律）设n_A是n 次独立重复试验中A发生的次数. p 是事件A在每次试验中发生的概率, 则对任意$\varepsilon > 0$,有

$$\lim_{n\to\infty} \mathrm{P}\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1 \Longleftrightarrow \lim_{n\to\infty} \mathrm{P}\left\{\left|\frac{n_A}{n} - p\right| \geq \varepsilon\right\} = 0$$

证: 因为$n_A \sim b(n, p)$, 有 $n_A = X_1 + X_2 + \cdots + X_n$

因而 $E(X_k)=p$，$D(X_k)=p(1-p)$, (k=1,2,...),由定理1,

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - p\right| < \varepsilon\right\} = 1 \quad \text{即} \quad \lim_{n\to\infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1$$

定理2 （贝努力大数定律）设n_A是n 次独立重复试验中A发生的次数. p 是事件A在每次试验中发生的概率, 则对任意$\varepsilon > 0$,有

$$\lim_{n\to\infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1 \Longleftrightarrow \lim_{n\to\infty} P\left\{\left|\frac{n_A}{n} - p\right| \geq \varepsilon\right\} = 0$$

此定理表明： $\frac{n_A}{n} \xrightarrow{P} P(A),(n \to \infty)$

即：事件A发生的频率依概率收敛于事件的概率 p.

这个定理以严格的数学形式表达了频率的稳定性.

定理3（辛钦定理，或称弱大数定理）设随机变量序列$X_1,X_2,\ldots,X_n,\ldots$相互独立且同分布，数学期望：$E(X_k)=\mu$，则对任意正数ε，有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| < \varepsilon\right\} = 1$$

(利用特征函数证明，略)

[注]

- 贝努力大数定律就是频率稳定性的理论依据．因而在实际应用中，当试验次数很大时，往往用事件发生的频率来代替事件的概率．
- 贝努力大数定律是辛钦定理的特殊情况，辛钦定理不要求随机变量的方差存在．

其它一些一般情况：

1) $X_1, X_2, \cdots, X_n, \cdots$ 相互独立但不一定有相同的数学期望与方差，可设

$$E(X_k) = \mu_k, \quad D(X_k) = \sigma_k^2 \le c < \infty, \quad k = 1, 2, \cdots$$

有

$$\lim_{n\to\infty} P\left(\left| \frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} \mu_k \right| \ge \varepsilon \right) = 0$$

2) $X_1, X_2, \cdots, X_n, \cdots$ 相互独立的条件可以去掉，代之以

$$\frac{1}{n^2} D\left(\sum_{k=1}^{n} X_k \right) \xrightarrow{n\to\infty} 0$$

例1 设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立，具有如下分布律：

X_n	$-na$	0	na
P	$\frac{1}{2n^2}$	$1-\frac{1}{n^2}$	$\frac{1}{2n^2}$

问是否满足切比雪夫条件下的大数定理?

解 独立性依题意可知, 检验是否具有数学期望?

$$E(X_n) = -na \cdot \frac{1}{2n^2} + 0 \cdot (1 - \frac{1}{n^2}) + na \cdot \frac{1}{2n^2} = 0,$$

说明每一个随机变量都有数学期望,

检验是否具有有限方差?

因为

X_n^2	$(na)^2$	0	$(na)^2$
P	$\dfrac{1}{2n^2}$	$1-\dfrac{1}{n^2}$	$\dfrac{1}{2n^2}$

所以 $E(X_n^2) = 2(na)^2 \cdot \dfrac{1}{2n^2} = a^2,$

所以 $D(X_n) = E(X_n^2) - [E(X_n)]^2 = a^2.$

说明离散型随机变量有有限方差,

故满足切比雪夫定理的条件.

例2 设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 独立同分布，且 $E(X_k)=0$, $D(X_k)=\sigma^2$, $k=1,2,\cdots$, 证明对任意正数 ε 有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k{}^2 - \sigma^2\right| < \varepsilon\right\} = 1.$$

解 因为 $X_1, X_2, \cdots, X_n, \cdots$ 是相互独立的，

所以 $X_1{}^2, X_2{}^2, \cdots, X_n{}^2, \cdots$ 也是相互独立的，

由 $E(X_k)=0$, 得 $E(X_k{}^2) = D(X_k) + [E(X_k)]^2 = \sigma^2$,

由辛钦定理知

对于任意正数 ε, 有 $\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k{}^2 - \sigma^2\right| < \varepsilon\right\} = 1.$

小结

三个大数定理

- 切比雪夫定理的特殊情况
- 伯努利大数定理
- 辛钦定理

频率的稳定性是概率定义的客观基础，而伯努利大数定理以严密的数学形式论证了频率的稳定性.

§5.2 中心极限定理

实例: **考察射击命中点与靶心距离的偏差.**

这种偏差是大量微小的偶然因素造成的微小误差的总和, 这些因素包括: 瞄准误差、测量误差、子弹制造过程方面 (如外形、重量等) 的误差以及射击时武器的振动、气象因素(如风速、风向、能见度、温度等) 的作用, 所有这些不同因素所引起的微小误差是相互独立的, 并且它们中每一个对总和产生的影响不大.

问题:

某个随机变量是由大量相互独立且均匀小的随机变量相加而成的, 其概率分布情况如何呢?

这种随机变量一般都近似服从正态分布. 在概率论中，习惯于把和的分布收敛于正态分布这一类定理都叫做中心极限定理.

➢独立同分布的（林德贝格-列维）中心极限定理

定理1 设随机变量$X_1,X_2,\ldots,X_n,\ldots$ 相互独立，服从同一分布, 且 $E(X_k)=\mu$，$D(X_k)=\sigma^2\neq 0$ $(k=1,2,\ldots)$， 则

$$Y_n=\frac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}$$

和的标准化随机变量

的分布函数$F_n(x)$满足:对任意实数x，有

$$\lim_{n\to\infty}F_n(x)=\lim_{n\to\infty}P\left\{\frac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}\leq x\right\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\Phi(x).$$

(利用特征函数证明，略)

□在一般情况下，很难求出Y_n的分布的确切形式

□定理表明, 当n充分大时, Y_n近似服从标准正态分布

➢独立同分布的（林德贝格-列维）中心极限定理

定理1 设随机变量$X_1,X_2,\ldots,X_n,\ldots$ 相互独立，服从同一分布, 且 $E(X_k)=\mu$，$D(X_k)=\sigma^2\neq 0$ $(k=1,2,\ldots)$， 则

$$Y_n=\frac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}$$

和的标准化随机变量

的分布函数$F_n(x)$满足:对任意实数x，有

$$Y_n=\frac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}\overset{\text{近似}}{\sim}N(0,1)$$

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i\overset{\text{近似}}{\sim}N(\mu,\frac{\sigma^2}{n})$$

依分布收敛

例1 一盒同型号螺丝钉共100个，已知该型号的螺丝钉的重量是一个随机变量，期望值是100g，标准差是10g，求一盒螺丝钉的重量超过10.2kg的概率.

解：设X_i 为第i个螺丝钉的重量, i=1,2,…,100, 且相互独立,

于是, 一盒螺丝钉的重量为 $X=\sum_{i=1}^{100} X_i$

且 $\mu=E(X_i)=100,\sigma=\sqrt{D(X_i)}=10,n=100$

由中心极限定理

$$P\{X>10200\}=P\{\sum_{i=1}^{100} X_i>10200\}=P\left\{\frac{\sum_{i=1}^{100} X_i-n\mu}{\sigma\sqrt{n}}>\frac{10200-n\mu}{\sigma\sqrt{n}}\right\}$$

$$=P\left\{\frac{X-10000}{100}>\frac{10200-10000}{100}\right\}=P\left\{\frac{X-10000}{100}>2\right\}$$

$$\approx 1-\Phi(2)=1-0.97725=0.02275$$

定理2 (李雅普诺夫定理)设随机变量$X_1,X_2,\ldots,X_n,\ldots$相互独立，且具有数学期望和方差： $E(X_k)=\mu_k$， $D(X_k)=\sigma_k^2\neq 0$ (k=1,2,...)，

记 $B_n^2=\sum_{k=1}^{n}\sigma_k^2$,若存在$\delta>0$,使得 $\frac{1}{B_n^{2+\delta}}\sum_{k=1}^{n}E\{|X_k-\mu_k|^{2+\delta}\}\to 0,$
$(n\to\infty)$

则随机变量

$$Z_n=\frac{\sum_{k=1}^{n}X_k-E(\sum_{k=1}^{n}X_k)}{\sqrt{D(\sum_{k=1}^{n}X_k)}}=\frac{\sum_{k=1}^{n}X_k-\sum_{k=1}^{n}\mu_k}{B_n}$$

的分布函数$F_n(x)$对任意x，有

(证明略)

$$\lim_{n\to\infty}F_n(x)=\lim_{n\to\infty}P\left\{\frac{\sum_{i=1}^{n}X_i-\sum_{i=1}^{n}\mu_i}{B_n}\le x\right\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\Phi(x).$$

❑**此定理表明，当n充分大时，Z_n的分布近似于标准正态分布.**

请注意:

1、定理中随机变量之和$\sum_{k=1}^{n} X_k$及其标准化变量在n很大时,分别近似服从

$$\sum_{k=1}^{n} X_k \overset{\text{近似地}}{\sim} N(\sum_{k=1}^{n} \mu_k, B_n^2); \qquad \frac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \mu_k}{B_n} \overset{\text{近似地}}{\sim} N(0,1)$$

2、随机变量X_k无论服从什么分布,只要满足定理条件,当n很大时,随机变量之和$\sum_{k=1}^{n} X_k$就近似服从**正态分布**,这就是为什么正态分布在概率论中占有重要地位的一个基本原因.(如实例中射击偏差服从正态分布)

定理一的特殊情况.

定理3 (德莫佛−拉普拉斯定理)设随机变量η_n(n=1,2,…)服从参数为n，p(0< p < 1)的二项分布，则对任意 x，恒有

$$\lim_{n\to\infty} \mathrm{P}\left\{\frac{\eta_n - np}{\sqrt{np(1-p)}} \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x).$$

证 由§4.2例知, η_n可以看成n个相互独立的服从同一(0-1)分布的随机变量$X_1,\ldots,X_n$之和，即 $\eta_n = X_1 + X_2 + \cdots + X_n$

$$E(X_i) = p, D(X_i) = p(1-p), \qquad i = 1,2,\cdots,n$$

由定理1知，$\lim\limits_{n\to\infty} \mathrm{P}\left\{\frac{\eta_n - np}{\sqrt{np(1-p)}} \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$

□此定理表明，<u>正态分布是二项分布的极限分布</u>，所以当n充分大时，我们可以用标准正态分布近似二项分布.

定理一的特殊情况.

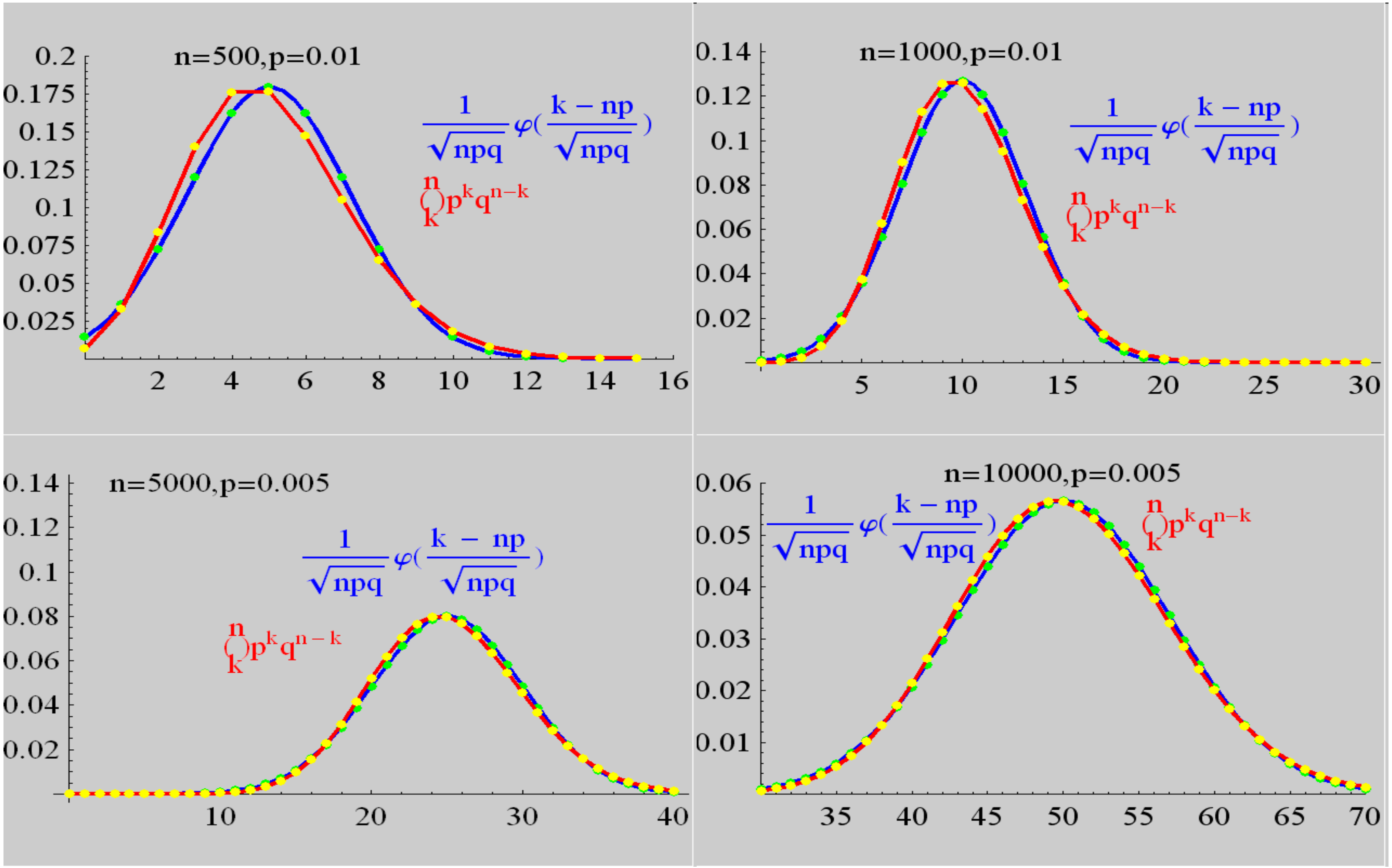
定理3 (德莫佛–拉普拉斯定理)设随机变量$\eta_n(n=1,2,\ldots)$服从参数为n，$p(0< p < 1)$的二项分布，则对任意 x，恒有

$$\lim_{n\to\infty} \mathrm{P}\left\{\frac{\eta_n - np}{\sqrt{np(1-p)}} \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x).$$

$$\frac{\eta_n - np}{\sqrt{np(1-p)}} \overset{\text{近似}}{\sim} N(0,1)$$

□此定理表明，正态分布是二项分布的极限分布，所以当n充分大时，我们可以用标准正态分布近似二项分布.

下面的图形表明:正态分布是二项分布的逼近.



例2 某车间有200台车床独立工作，设每台车床的开工率为0.6，开工时耗电1千瓦，问供电所至少要供多少电才能以不小于99.9%的概率保证该车间不会因供电不足而影响生产？

解 记X为200台车床中工作着的车床台数，则$X\sim b(200, 0.6)$.
按题意，要求最小的 k，使$P\{X\le k\}\ge 0.999$

即 $\sum_{i=0}^{k} C_{200}^{i}\, 0.6^{i}(1-0.6)^{200-i} \ge 0.999$ ，由定理3

$$P\{0\le X\le k\} = P\left\{\frac{0-120}{48} \le \frac{X-200\times 0.6}{\sqrt{200\times 0.6\times 0.4}} \le \frac{k-120}{\sqrt{48}}\right\}$$

$$= P\left\{\frac{-120}{\sqrt{48}} \le \frac{X-120}{\sqrt{48}} \le \frac{k-120}{\sqrt{48}}\right\} \quad \Longrightarrow \quad \frac{k-120}{\sqrt{48}} \ge 3.1$$

$$\approx \Phi\left(\frac{k-120}{\sqrt{48}}\right) - \Phi\left(\frac{-120}{\sqrt{48}}\right) \ge 0.999 \qquad k \ge 141.48,$$

至少供电142千瓦,才能保证以不小于99.9%的概率正常工作.

思考：能否使用泊松分布逼近？

例3 在人寿保险公司里,有3000个同一年龄的人参加保险.设在一年内这些人的死亡率为0.1%, 参加保险的人在一年的头一天交付保险费10元，死亡时，家属可从保险公司领取2000元.
求（1）保险公司一年中获利不小于10000元的概率；

（2）保险公司亏本的概率是多少？

解 设一年中死亡人数为X，X=0,1,⋯,3000, 死亡率=0.001, 则 $X\sim b(3000, 0.001)$.而保险公司每年获利=3000×10-2000X (元)

(1) P{保险公司获利不小于10000元} =$P\{30000-2000X\geq 10000\}$ =$P\{0\leq X\leq 10\}$， 而由拉普拉斯定理，有

$$P\{0 \leq X \leq 10\} = P\left\{\frac{0-3}{1.7312} \leq \frac{X-3}{1.7312} \leq \frac{10-3}{1.7312}\right\}$$

$$\approx \Phi(4.043) - [1 - \Phi(1.733)] = 0.96$$

即一年中保险公司获利10000元以上的概率为96%.

(2)　P{保险公司亏本}=P{2000X>30000}=P{X>15}

$$= 1 - P\{0 \le X \le 15\}$$

$$= 1 - P\left\{\frac{0-3}{1.7312} \le \frac{X-3}{1.7312} \le \frac{15-3}{1.7312}\right\}$$

$$\approx 1 - \left[\Phi\left(\frac{12}{1.7312}\right) - \Phi\left(\frac{-3}{1.7312}\right)\right]$$

$$= 1 - 0.9582 = 0.0418$$

由此可见保险公司亏本的概率是很小的.

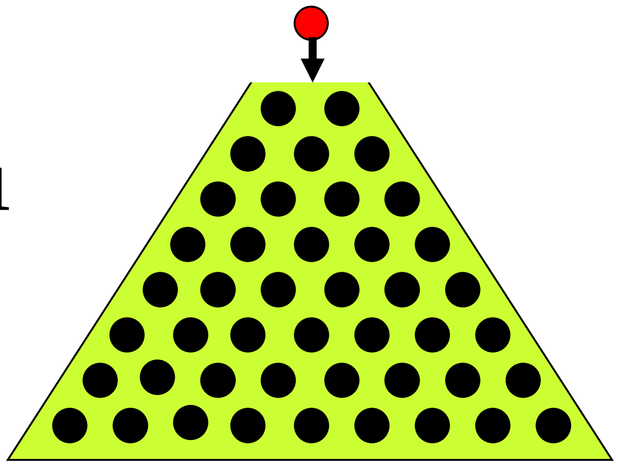
思考：能否使用泊松分布逼近？

例3 高尔顿钉板试验

如图是高尔顿钉板，常常在赌博游戏中见到，现在可用中心极限定理来揭穿这个赌博中的奥秘.

设n为钉子的排数,Y_n表示第n次碰钉后小球的位置，

$$X_i=\begin{cases}1, & \text{第i次碰钉后小球从左落下，}\\ -1, & \text{第i次碰钉后小球从右落下.}\end{cases}$$

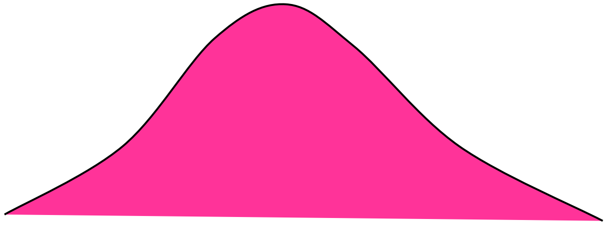


则X_i服从两点分布， $E(X_i)=0$, $D(X_i)=1$

$$Y_n=X_1+X_2+\dots+X_n$$

由中心极限定理知，

$$Y_n\sim N(0,n)$$



由正态分布的特征知，小球落在中间的概率远远大于落在两边的概率.

大数定律与中心极限定理的区别与联系：

设 $\{X_n\}$ 为独立同分布随机变量序列，且

$$E(X_i)=\mu \qquad D(X_i)=\sigma^2>0$$

则由大数定理，对于任意的$\varepsilon>0$有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i-\mu\right|<\varepsilon\right\}=1.$$

大数定律并未给出 $P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i-\mu\right|<\varepsilon\right\}$ 的表达式, 但保证了其极限是1.

而在以上同一条件下，中心极限定理(林德贝格—列维）亦成立，这时，对于任意的ε＞0及充分大的n，有

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right|<\varepsilon\right\}=P\left\{\left|\frac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}\right|<\frac{\sqrt{n}\varepsilon}{\sigma}\right\}\approx 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sigma}\right)-1$$

由于

$$2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sigma}\right)-1\xrightarrow[n\to\infty]{}1,$$

因此，在所给条件下，中心极限定理不仅给出了概率的近似表达式，而且也能保证其极限是1，可见中心极限定理的结论更为深入。

小结

三个中心极限定理

- 独立同分布的中心极限定理
- 李雅普诺夫定理
- 棣莫弗—拉普拉斯定理

中心极限定理表明, 在相当一般的条件下, 当独立随机变量的个数增加时, 其和的分布趋于正态分布.

作业

- **第五章习题3, 6, 12, 14**