§ 5.2 中心极限定理

定理一

林德伯格-列维中心极限定理
(Lindberg-levi)

[独立同分布的中心极限定理]

定理二

棣莫弗-拉普拉斯中心极限定理
(De Moivre-Laplace)

[二项分布以正态分布为极限分布]

定理 1 独立同分布的中心极限定理

设随机变量序列 $X_1, X_2, \cdots, X_n, \cdots$

独立同一分布, 且有期望和方差:

$$E(X_k) = \mu\,,\ D(X_k) = \sigma^2 > 0\,,\ k = 1,2,\cdots$$

则对于任意实数 x,

$$\lim_{n\to\infty} P\left(\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \le x \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt = \Phi(x)$$

注　记 $Y_n = \dfrac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma}$

则 Y_n 为 $\sum_{k=1}^{n} X_k$ 的标准化随机变量.

$$\lim_{n\to\infty} P(Y_n \le x) = \Phi(x)$$

即 n 足够大时，Y_n 的分布函数近似于标准正态随机变量的分布函数

$$Y_n \overset{\text{近似}}{\sim} N(0,1)$$

$\sum_{k=1}^{n} X_k = \sqrt{n}\sigma Y_n + n\mu$ 近似服从$N(n\mu, n\sigma^2)$

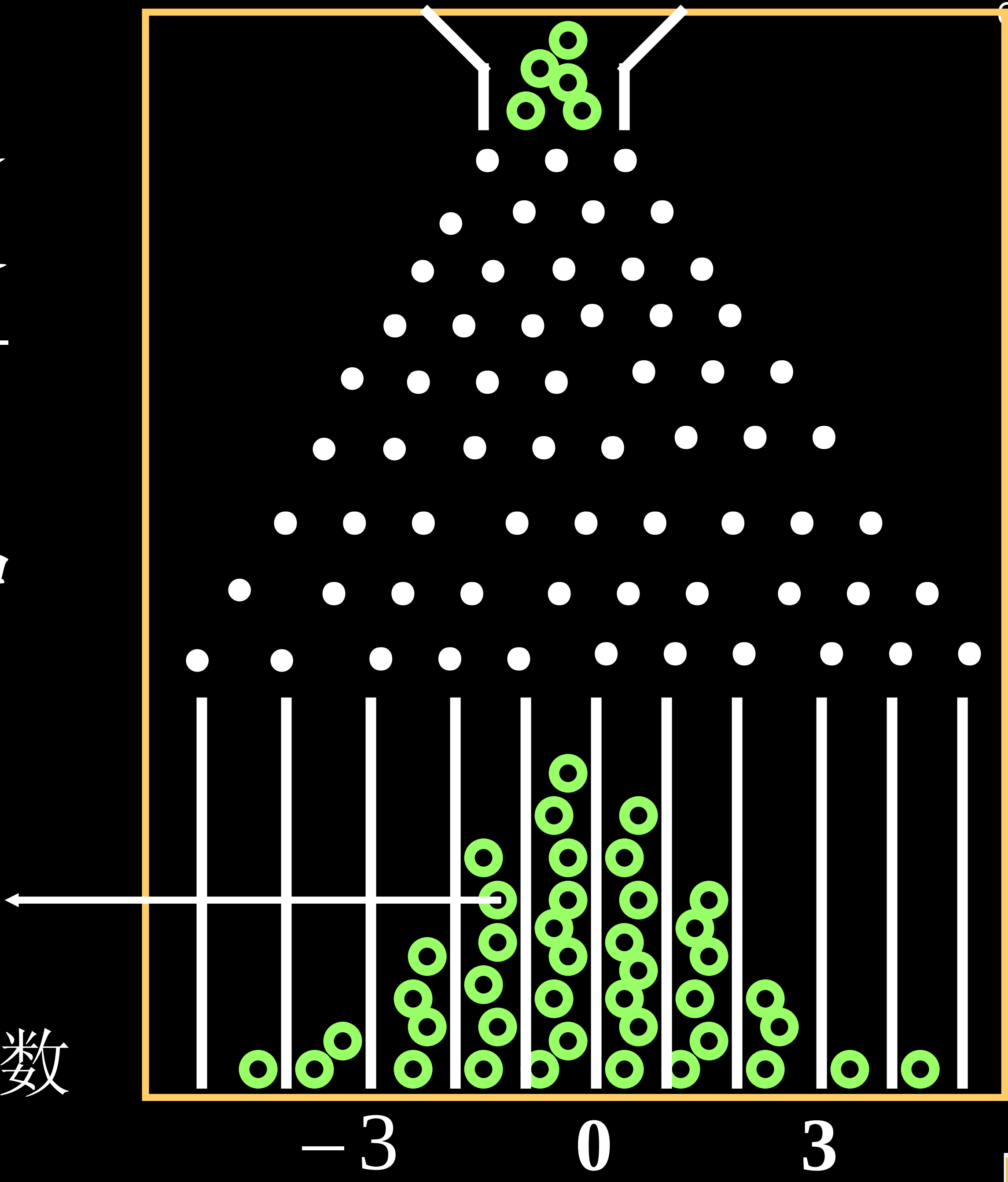
中心极限定理的意义

在第二章曾讲过有许多随机现象服从正态分布 是由于许多彼次没有什么相依关系、对随机现象谁也不能起突出影响，而均匀地起到微小作用的随机因素共同作用(即这些因素的叠加)的结果.

若联系于此随机现象的随机变量为X，则它可被看成为许多相互独立的起微小作用的因素X_k的总和 $\sum_k X_k$，而这个总和服从或近似服从正态分布.

对此现象还可举个有趣的例子——

高尔顿钉板试验——加以说明.



n — 钉子层数

定理2 德莫佛—拉普拉斯中心极限定理（DeMoivre-Laplace ）

设 $Y_n \sim B(n, p)$，$0 < p < 1$，$n = 1,2,\ldots$

则对任一实数 x，有

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

即对任意的 $a < b$，

$$\lim_{n\to\infty} P\left(a < \frac{Y_n - np}{\sqrt{np(1-p)}} \le b\right) = \frac{1}{\sqrt{2\pi}}\int_{a}^{b} e^{-\frac{t^2}{2}}\,dt$$

$Y_n \sim N(np, np(1-p))$ (近似)

中心极限定理的应用

例1 炮火轰击敌方防御工事 100 次, 每次轰击命中的炮弹数服从同一分布, 其数学期望为 2 , 均方差为1.5. 若各次轰击命中的炮弹数是相互独立的, 求100 次轰击

(1) 至少命中180发炮弹的概率;
(2) 命中的炮弹数不到200发的概率.

解 设 X_k 表示第 k 次轰击命中的炮弹数

$$E(X_k)=2,\quad D(X_k)=1.5^2,\quad k=1,2,\cdots,100$$

$X_1,X_2,\cdots,X_{100}$ 相互独立,

设 X 表示100次轰击命中的炮弹数, 则

$$X=\sum_{k=1}^{100}X_k,\ E(X)=200,\ D(X)=225,$$

由独立同分布中心极限定理, 有

$$X \overset{\text{近似}}{\sim} N(200\ ,225\ )$$

$$(1)\quad P(X \geq 180) \approx 1 - \Phi\left(\frac{180 - 200}{15}\right)$$

$$= 1 - \Phi(-1.3) = \Phi(1.3) = 0.91$$

$$(2)\quad P(0 \leq X < 200) \approx \Phi\left(\frac{200 - 200}{15}\right) - \Phi\left(\frac{0 - 200}{15}\right)$$

$$= \Phi(0) - \Phi(-13.33) = 0.5$$

例2 售报员在报摊上卖报，已知每个过路人在报摊上买报的概率为1/3. 令X是出售了100份报时过路人的数目，求

$P(280 \le X \le 320)$.

解 令X_i为售出了第 $i-1$ 份报纸后到售出第i份报纸时的过路人数，$i=1,2,\ldots,100$

$$P(X_i = k) = p(1-p)^{k-1}\Big|_{p=1/3}, \quad k = 1,2,\cdots$$

(几何分布)

$$E(X_i) = \frac{1}{p}\Big|_{p=1/3} = 3, \; D(X_i) = \frac{1-p}{p^2}\Big|_{p=1/3} = 6$$

$X_1, X_2, \cdots, X_{100}$ 相互独立, $X = \sum_{k=1}^{100} X_k$

$$E(X) = 300, \quad D(X) = 600$$

由独立同分布中心极限定理, 有

$$X \sim N(300, 600) \text{ (近似)}$$

$$P(280 \le X \le 320) \approx \Phi\left(\frac{320-300}{\sqrt{600}}\right) - \Phi\left(\frac{280-300}{\sqrt{600}}\right)$$

$$\approx 2\Phi\left(\frac{20}{\sqrt{600}}\right) - 1 \approx 2\Phi(0.8165) - 1 \approx 0.5878$$

例3 检验员逐个检查某产品,每查一个需用10秒钟. 但有的产品需重复检查一次,再用去10秒钟. 若产品需重复检查的概率为 0.5, 求检验员在 8 小时内检查的产品多于1900个的概率.

解 若在 8 小时内检查的产品多于1900个,即检查1900个产品所用的时间小于 8 小时.

设 X 为检查1900 个产品所用的时间(秒)

设 X_k 为检查第 k 个产品所用的时间(单位：秒)，$k = 1,2,\ldots,1900$

X_k	10	20
P	0.5	0.5

$$E(X_k)=15,\quad D(X_k)=25$$

$X_1, X_2, \cdots, X_{1900}$ 相互独立同分布，$X=\sum_{k=1}^{1900} X_k$

$$E(X)=1900\times 15=28500$$

$$D(X)=1900\times 25=47500$$

$$X \overset{\text{近似}}{\sim} N(28500\ ,47500\ )$$

$$P(10\times 1900 \le X \le 3600\times 8)$$

$$= p(19000 \le X \le 28800)$$

$$\approx \Phi\left(\frac{28800-28500}{\sqrt{47500}}\right)-\Phi\left(\frac{19000-28500}{\sqrt{47500}}\right)$$

$$\approx \Phi(1.376)-\Phi(-43.589)$$

$$\approx 0.9162$$

例4 某车间有200台车床，每台独立工作，开工率为0.6. 开工时每台耗电量为 r 千瓦. 问供 电所至少要供给这个车间多少电力，才能以 99.9% 的概率保证这个车间不会因供电不足而影响生产？

解 设至少要供给这个车间 a 千瓦的电力，

X 为开工的车床数，则 $X \sim B(200, 0.6)$，

由德莫佛—拉普拉斯中心极限定理，有

$X \sim N(120, 48)$ (近似)

问题转化为求 a，使

$$P(0 \le rX \le a) = 99.9\%$$

$$P(0 \le rX \le a) = \Phi\left(\frac{a/r - 120}{\sqrt{48}}\right) - \Phi\left(\frac{0 - 120}{\sqrt{48}}\right)$$

$$= \Phi\left(\frac{a/r - 120}{\sqrt{48}}\right)$$

$= \Phi(-17.32)$

≈ 0

反查标准正态函数分布表，得

$$\Phi(3.09) = 99.9\%$$

令

$$\frac{\frac{a}{r} - 120}{\sqrt{48}} = 3.09$$

解得

$$a = (3.09\sqrt{48} + 120)r$$

$$\approx 141r \ (\text{千瓦})$$

例5 设有一批种子，其中良种占1/6. 试估计在任选的6000粒种子中，良种比例与 1/6 比较上下不超过1%的概率.

解 设 X 表示6000粒种子中的良种数，

则 $$X \sim B(6000, 1/6)$$

由德莫佛—拉普拉斯中心极限定理,

有 $$X \overset{\text{近似}}{\sim} N\left(1000, \frac{5000}{6}\right)$$

$$P\left(\left|\frac{X}{6000}-\frac{1}{6}\right|<0.01\right)=P\left(\left|X-1000\right|<60\right)$$

$$\approx \Phi\left(\frac{1060-1000}{\sqrt{5000/6}}\right)-\Phi\left(\frac{940-1000}{\sqrt{5000/6}}\right)$$

$$=\Phi\left(\frac{60}{\sqrt{5000/6}}\right)-\Phi\left(\frac{-60}{\sqrt{5000/6}}\right)$$

$$=2\Phi\left(\frac{60}{\sqrt{5000/6}}\right)-1\approx 0.9624$$

比较几个近似计算的结果

二项分布(精确结果) $P\left(\left|\frac{X}{6000}-\frac{1}{6}\right|<0.01\right)\approx 0.9590$

中心极限定理 $P\left(\left|\frac{X}{6000}-\frac{1}{6}\right|<0.01\right)\approx 0.9624$

Poisson 分布 $P\left(\left|\frac{X}{6000}-\frac{1}{6}\right|<0.01\right)\approx 0.9379$

Chebyshev 不等式 $P\left(\left|\frac{X}{6000}-\frac{1}{6}\right|<0.01\right)\geq 0.7685$

作业 P.186 习题五

8 10

12 15

*补充作业

设某农贸市场某种商品每日的价格的变化是个相互独立且均值为0，方差为$\sigma^2 = 2$的随机变量Y_n，并满足

$$X_n = X_{n-1} + Y_n \quad (n \geq 1)$$

其中X_n是第n天该商品的价格.如果今天的价格为100，求18天后该商品的价格在 96 与 104 之间的概率.

解 设 X_0 表示今天该商品的价格, X_{18}为18天后该商品的价格, 则

$$X_{18}=X_{17}+Y_{18}=X_{16}+Y_{17}+Y_{18}=X_0+\sum_{i=1}^{18}Y_i$$

得 $$P(96\le X_{18}\le 104)=P(-4\le\sum_{i=1}^{18}Y_i\le 4)$$

$$=P(-\frac{4}{\sqrt{36}}\le\frac{1}{\sqrt{36}}\sum_{i=1}^{18}Y_i\le\frac{4}{\sqrt{36}})$$

$$\approx\Phi(2/3)-\Phi(-2/3)=2\Phi(2/3)-1$$

$$=2\times 0.747-1=0.494.$$

第12周　问　题

一本书有 1 000 000 个印刷符号，排版时每个符号被排错的概率为千分之一. 校对时，每个排版错误被改正的概率为0.99.　求在校对后错误不多于15 个的概率.