期望和方差的性质：

1） $E(cX+b) = cE(X)+b$　　　特别地：$D(b)=0$

$D(cX+b) = c^2 D(X)$　　　$D(-X) = D(X)$

2) $E(\sum_{k=1}^{n} X_k) = \sum_{k=1}^{n} E(X_k)$,

$E(E(X)) = E(X)$

$D(E(X)) = 0$

$$\sum_{i=1}^{n} E\left\{\left[X_i - E(X_i)\right]^2\right\} + 2\sum_{1\le j\ne i\le n} E\left\{\left[X_i - E(X_i)\right]\left[X_j - E(X_j)\right]\right\}$$

特别:$D(X\pm Y) = D(X) + D(Y) \pm 2E\left\{\left[X - E(X)\right]\left[Y - E(Y)\right]\right\}$

若$X_1, X_2, \dots, X_n$相互独立，则

$D(\sum_{k=1}^{n} X_k) = \sum_{k=1}^{n} D(X_k)$,　　$E\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} E(X_i)$

3）$D(X)=0$ ⇔ $P\{X=E(X)\}=1$(方差为0的随机变量必为常数)

重要结论:

1. 随机变量X的$E(X), D(X)$存在, 且$D(X)>0$, 则$X^* = \dfrac{X-E(X)}{\sqrt{D(X)}}$

$$E(X^*) = 0,\ \ D(X^*) = 1$$

2. 设随机变量$X_1, X_2, \ldots, X_n$相互独立, 且$E(X_i)=a, D(X_i)=b$,

则: $E(\bar{X}) = E\left(\dfrac{1}{n}(X_1+X_2+\ldots+X_n)\right)=a, D(\bar{X})=\dfrac{1}{n}b.$

3. 设随机变量$X_1, X_2, \ldots, X_n$,相互独立且同分布,均值为a和方差为b,

则: $E(\bar{X}) = E\left(\dfrac{1}{n}(X_1+X_2+\ldots+X_n)\right)=a, D(\bar{X})=\dfrac{1}{n}b$

(回忆)频率具有稳定性: 在一定条件下, 频率稳定于概率.

另一个验证频率稳定性的著名试验是由英国生物统计学家高尔顿(Galton)设计的. 它的试验模型如图 1.1.1 所示.

自上端放入一小球,任其自由下落,在下落过程中当小球碰到钉子时,从左边落下与从右边落下的机会相等. 碰到下一排钉子时又是如此. 最后落入底板中的某一格子. 因此,任意放入一球,则此球落入哪一个格子,预先难以确定. 但是实验证明,如放入大量小球,则其最后所呈现的曲线,几乎总是一样的. 也就是说,小球落入各个格子的频率十分稳定. 这个试验模型称为高尔顿板. 试验中呈现出来的规律性,在学习第五章极限定理之后,就会有更深刻的理解.

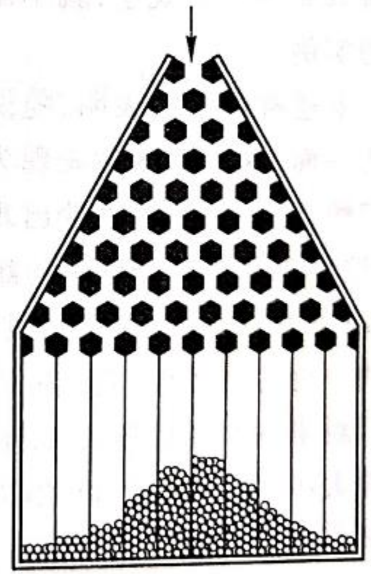

图 1.1.1 高尔顿板

· 5 ·

事实上, 高尔顿板可以看作是伯努利试验的一个实验模型.

如果我们把小球碰到钉子看作是一次试验, 而把从右边落下算是成功, 从左边落下就算失败, 这时就有了一次p=1/2的伯努利试验.

小球从顶端到底层共需要经过n排钉子, 这就相当于一个n次伯努利试验, 剩下的只是说明为什么高度曲线会是正态分布密度函数曲线?

演示

应用背景

在实际应用中, 常遇到如下问题:

1. "频率是随机变量, 它的稳定性", 即它的极限将采用何种提法呢? 能否用: $\lim_{n\to\infty} Y_n = \boldsymbol{p}$?

在实际应用中有什么作用?

2 正态分布是日常生活中最常见、常用的一种分布, 如: 身高、体重、成绩、测量数据等.

为什么现实中大量随机变量服从正态分布?

依据什么来断定一个随机变量服从正态分布?

等等, 诸如此类的问题.

第五章 大数定律和中心极限定理

(Center Limit Theorem(CLT) and Law of Large Number(LLN))

§5.1 大数定律 (*LLN*)

一、 切比雪夫(*Chebyshev*)不等式

设随机变量*X*的数学期望 $E(X)$和方差$D(X)$都存在, 则对于任意给定的 $\varepsilon > 0$, 有

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

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

证 明

若随机变量X的方差$D(X)$存在，则$\forall\ \varepsilon>0$有

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

证明：仅就X是连续型的情况给予证明，设X的为$f(x)$

$$P\{|X-E(X)|\geq\varepsilon\}$$

$$=P\{X\geq E(X)+\varepsilon\text{或}X\leq E(X)-\varepsilon\}$$

$$=\int_{\left\{x\left|\frac{|x-E(X)|}{\varepsilon}\geq 1\right.\right\}} f(x)dx$$

$$\leq\int_{\{x||x-E(X)|\geq\varepsilon\}}\frac{|x-E(X)|^2}{\varepsilon^2}f(x)dx$$

$$\leq\frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}[x-E(X)]^2 f(x)dx=\frac{D(X)}{\varepsilon^2}$$

(*Chebyshev*不等式)若随机变量X的方差$D(X)$存在，则$\forall\varepsilon>0$,有

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

固定ε(一个较小的数),

若X的方差小，事件$\{|X-E(X)|\geq\varepsilon\}$发生的概率小,

即事件$\{|X-E(X)|<\varepsilon\}$发生的概率大.

切比雪夫(Chebyshev)不等式说明：

X与$E(X)$之间偏离程度由方差来控制

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

或者 $P\{|X-E(X)|<\varepsilon\}\geq 1-\dfrac{D(X)}{\varepsilon^2}$,

切比雪夫不等式的应用

1) 估计随机变量落在某个区间内的概率. (粗略估计)

估计概率

2) 估计 ε 的值, 使 $P(|X-E(X)|<\varepsilon)\geq a\ \ (0<a<1)$

方法: 这实际是1) 的反问题, 主要将 a 与 $1-\dfrac{D(X)}{\varepsilon^2}$ 比较

重复试验次数估计

3) 证明大数定律.

补充：方差性质

例：设随机变量$X_1, X_2, \cdots$，X_{100}相互独立且都服从参数为$\lambda=4$的泊松分布，令$Y=\sum_{i=1}^{100} X_i$，利用切比雪夫不等式估计概率$P\{360<Y<440\}$.

解：因为$E(X_i)=D(X_i)=4\ (i=1,2,\cdots,100)$

而且$X_1, X_2, \cdots, X_{100}$相互独立，有

$$E(Y)=\sum_{i=1}^{100} E(X_i)=400 \qquad D(Y)=\sum_{i=1}^{100} D(X_i)=400.$$

由切比雪夫不等式：$P\{|X-E(X)|<\varepsilon\} \geq 1-\dfrac{D(X)}{\varepsilon^2}$

$$P\{360<Y<440\}=P\{|Y-E(Y)|<40\} \geq 1-\frac{400}{40^2}=\frac{3}{4}=0.75$$

切比雪夫不等式只能对概率作出很粗略的估计.

$$P\{\underset{-\varepsilon}{\underline{a}}<Y-E(Y)<\underset{\varepsilon}{\underline{b}}\}$$

例5.2: 将一枚均匀硬币连续抛 n 次, 试用切比雪夫不等式求出 n, 使下式成立:

$$P\{|f_n(A)-P(A)|<0.01\}\geq 0.99$$

其中 A ={ 出现正面 }. 随机变量

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

解： $P(A)$=1/2，令

$$X_i=\begin{cases}1, & \text{第}i\text{次出现正面;}\\ 0, & \text{else.}\end{cases} \qquad i=1,2,\cdots n$$

所以出现正面的总次数为 $\sum_{i=1}^{n}X_i$，从而有

$$f_n(A)=\frac{1}{n}\sum_{i=1}^{n}X_i \qquad E[f_n(A)]=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{2};$$

$$P\{|f_n(A)-P(A)|<0.01\}\geq 0.99 \qquad E[f_n(A)]=\frac{1}{2}$$

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

$$D[f_n(A)]=D(\frac{1}{n}\sum_{i=1}^{n}X_i)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n}\times\frac{1}{2}\times\frac{1}{2}=\frac{1}{4n}$$

由切比雪夫不等式:

$$P\{|f_n(A)-P(A)|<0.01\}=P\{|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{2}|<0.01\}$$

$$\geq 1-\frac{D[f_n(A)]}{0.01^2}=1-\frac{1}{0.01^2\times 4n}\geq 0.99$$

从而有 $4n\geq\dfrac{1}{0.01^3}\Rightarrow n\geq 250,000(\text{次})$

方差性质： $D(X)=0 \Leftrightarrow P\{X=E(X)\}=1.$

证明：充分性：$P\{X=E(X)\}=1 \Rightarrow D(X)=0$(显然成立)

必要性： $D(X)=0 \Rightarrow P\{X=E(X)\}=1.$

$$\{X \neq E(X)\}=\{|X-E(X)|>0\}$$

次可加性

$$=\bigcup_{n=1}^{\infty}\{|X-E(X)| \geq \frac{1}{n}\}$$ 非重点

$$P\{X \neq E(X)\} \leq \sum_{n=1}^{\infty} P\{|X-E(X)| \geq \frac{1}{n}\}$$

Chebyshev不等式

$$\leq \sum_{n=1}^{\infty} \frac{D(X)}{(1/n)^2}=0$$

二. 大数定律的定义

1 依概率收敛

设$X_1, \ldots, X_n, \ldots$是一个随机变量序列, 记为$\{X_n\}$

X 是一个随机变量或常数, 对于任意给定的$\varepsilon > 0$,

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

或者 $$\lim_{n\to\infty} P\{|X_n - X| \geq \varepsilon\} = 0$$

则称随机变量序列$\{X_n\}$依概率收敛于X，记为

$$X_n \xrightarrow{P} X \quad \text{或者} \quad \lim_{n\to\infty} X_n = X\ (P)$$

注: 随机变量序列依概率收敛的意义不同于微积分中数列收敛的意义. $\{P(|X_n - X| < \varepsilon),\ n = 1, \cdots\}$_看作数列

1、切比雪夫(Chebyshev)不等式

任意给定的$\varepsilon>0$, $P\{|X-E(X)|\geq\varepsilon\}\leq\dfrac{D(X)}{\varepsilon^2}$

或者$P\{|X-E(X)|<\varepsilon\}\geq 1-\dfrac{D(X)}{\varepsilon^2}$,

设$X_1,\ldots,X_n,\ldots$是一个随机变量序列,记为$\{X_n\}$

2、依概率收敛:

$X_n \xrightarrow{P} X$:

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

令$y_n=P\{|X_n-X|<\varepsilon\}$ 数列$\{y_n\}$收敛于1:

对任意给定的正数$\eta>0$, ∃正整数$N(\eta)$,

使得对$n>N$的任何n, $|y_n-1|<\eta$.

依概率收敛与数列收敛的不同：(补充)

1)数列$\{x_n\}$是$N=\{1,2,\ldots\}$上的一个函数序列

随机变量序列$\{X_n\}$是概率空间上的一个函数序列.

2)数列$\{x_n\}$收敛：

常数A，对任意给定的正数$\varepsilon>0$,

$\exists$正整数$N(\varepsilon)$，使得对$n>N$的任何n, $|x_n - A| < \varepsilon$

随机变量序列$\{X_n\}$依概率收敛随机变量或常数X：

无论多么小的正数ε, 当n很大时,

X_n与X的偏差不超过ε的可能性很大, 接近于1,

即$\lim\limits_{n\to\infty} P\{|X_n - X| < \varepsilon\} = 1$　　但并不是一定发生！

例子： 依概率收敛

抛一枚均匀硬币，设：μ_n：正面向上的次数

随着投掷次数n的增大，出现正面的频率

$\frac{\mu_n}{n}$"稳定于"$\frac{1}{2}$　能否用：$\lim\limits_{n\to\infty} Y_n = \frac{1}{2}$?

记：$X_i = \begin{cases} 1, & 第i次掷出正面； \\ 0, & 第i次掷出反面。 \end{cases}$　$i = 1,2,\cdots,n$

则有：$\frac{\mu_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i$

问题归结为随机变量序列 $Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$，$n = 1,2,\cdots$

是否"收敛于"极限0.5。

$\forall \varepsilon > 0, \exists N$, 使得当 $n > N$ 时，有 $|Y_n - \frac{1}{2}| < \varepsilon$?

能否用：$\lim\limits_{n\to\infty} Y_n = \frac{1}{2}$?

$\forall \varepsilon > 0, \exists N$，使得当$n > N$时，有$|Y_n - \frac{1}{2}| < \varepsilon$?

Y_n是随着试验结果而变的。

$$X_i = \begin{cases} 1, & \text{第}i\text{次掷出正面;} \\ 0, & \text{第}i\text{次掷出反面。} \end{cases} \quad i = 1,2,\cdots,n$$

有可能每次抛硬币都是正面向上。

$$Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad n = 1,2,\cdots$$

当$Y_n = 1, |Y_n - \frac{1}{2}| = \frac{1}{2}$

则有：$\frac{\mu_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i$

$P\{Y_n = 1\} = P\{X_1 = 1, \ldots X_n = 1\} = \frac{1}{2^n}$ (当n很大时,$\{Y=1\}$的可能性很小)

当n很大时，事件$\{|Y_n - \frac{1}{2}| \geq \varepsilon\}$发生的可能性很小，

即事件$\{|Y_n - \frac{1}{2}| < \varepsilon\}$发生的可能性很大。

即：频率“稳定于”概率不是：$\lim_{n\to\infty} \frac{\mu_n}{n} = \frac{1}{2}$

2 大数定律的定义

设$X_1 \ldots, X_n \ldots$是一个随机变量序列, 记为$\{X_n\}$ 其期望$E(X_n)$都存在, 若对于任意给定的$\varepsilon > 0$, 有

$$\lim_{n\to\infty} P\{|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)| < \varepsilon\} = 1$$

或者 $\lim_{n\to\infty} P\{|\overline{X} - E(\overline{X})| < \varepsilon\} = 1$

$$E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X_i)$$

则称随机变量序列$\{X_n\}$服从大数定律(或大数法则).

注: 1)$\{X_n\}$服从大数定律, 指 $\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i) \xrightarrow{P} 0$

2) 服从大数定律即: $\{X_k\}$, k=1,2…的前n 项算术平均紧密地聚集在其数学期望的附近;

3) 期望$E(X_n)$都存在是 $\{X_n\}$服从大数定律必要条件.

三. 常见大数定律

借助切比雪夫不等式:

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

$\{X_n\}$服从大数定律证明目的: 任意给定的$\varepsilon > 0$,

$$\lim_{n\to\infty} P\{|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)| < \varepsilon\} = 1$$

下面在$X_1, \ldots, X_n, \ldots,$ 相互独立且每个期望

$E(X_n)$均存在的条件下, 说明4个大数定律之间的关系.

借助切比雪夫不等式: $P\{|X-E(X)|<\varepsilon\}\geq 1-\frac{D(X)}{\varepsilon^2}$

Th5.2.2(切比雪夫大数定律) X_i 独立,

$E(X_i)$存在, $D(X_i)<C$ (一致有界), $i=1,2,3,\ldots$

令 $X=\frac{1}{n}\sum_{i=1}^{n}X_i$

$$\lim_{n\to\infty}P\{|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E(X_i)|<\varepsilon\}=1\text{(任给定的}\varepsilon>0)$$

$$D(\frac{1}{n}\sum_{i=1}^{n}X_i)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)\leq\frac{nC}{n^2}=\frac{C}{n}$$

$$E(\frac{1}{n}\sum_{i=1}^{n}X_i)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)$$

$$1\geq P\{|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E(X_i)|<\varepsilon\}$$

$$\geq 1-\frac{D(\frac{1}{n}\sum_{i=1}^{n}X_i)}{\varepsilon^2}\geq 1-\frac{C}{n\varepsilon^2}\xrightarrow{n\to\infty}1$$

借助切比雪夫不等式：$P\{|X-E(X)|<\varepsilon\}\geq 1-\dfrac{D(X)}{\varepsilon^2}$

Th5.2.2(切比雪夫大数定律) X_i 独立，

$E(X_i)$存在，$D(X_i)<C$ (一致有界) $i=1,2,3,\ldots$，

$$\lim_{n\to\infty}P\{|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E(X_i)|<\varepsilon\}=1$$

令 $X=\dfrac{1}{n}\sum_{i=1}^{n}X_i$

(任意给定的 $\varepsilon>0$)

Th5.2.3(独立同分布大数定律)

$E(X_i)=\mu, D(X_i)=\sigma^2, i=1,2,\ldots$

X_i 独立有相同分布

令 $\dfrac{1}{n}\sum_{i=1}^{n}E(X_i)=\mu$

$$\lim_{n\to\infty}P\{|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu|<\varepsilon\}=1 \quad (\varepsilon>0)$$

注：1) 此定律为切比雪夫大数定律的一个推论；

2) 为在实际应用中用大量重复测量值的算术平均值作为精确值的估计提供了理论依据.

(大数定律是由“频率依概率收敛于概率值”引申而来)

借助切比雪夫不等式: $P\{|X-E(X)|<\varepsilon\}\geq 1-\frac{D(X)}{\varepsilon^2}$

Th5.2.2(切比雪夫大数定律) X_i 独立, $E(X_i)$存在, $D(X_i)<C$ (一致有界), $i=1,2,3,\ldots$

令 $X=\frac{1}{n}\sum_{i=1}^{n}X_i$

$$\lim_{n\to\infty}P\{|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E(X_i)|<\varepsilon\}=1$$

(任意给定的$\varepsilon>0$)

Th5.2.3(独立同分布大数定律) X_i 独立, $E(X_i)=\mu, D(X_i)=\sigma^2, i=1,2,\ldots$

$\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\mu$

$$\lim_{n\to\infty}P\{|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu|<\varepsilon\}=1\ \ (\varepsilon>0)$$

$X_i=\begin{cases}1, & 第i次A发生;\\ 0, & 第i次A不发生.\end{cases}$

Th5.2.4(伯努里大数定律)
m为 n 重伯努里试验中事件A出现的次数,
p为A在每次试验中发生的概率.

令$\sum_{i=1}^{n}X_i=m$

$E(X_i)=\mu=p$

$$\lim_{n\to\infty}P\{|\frac{m}{n}-p|<\varepsilon\}=1\ \ (\varepsilon>0)$$

Th5.2.4.伯努里(Bernulli)大数定律

设 $\frac{m}{n}$ 是n次独立重复试验中事件A发生的频率，

p是事件A在每次试验中发生的概率, 则对于

$$\forall \varepsilon > 0,\ \lim_{n\to\infty} P\{|\frac{m}{n} - p| < \varepsilon\} = 1$$

证明： $$m = \sum_{i=1}^{n} X_i,\ X_i = \begin{cases} 1, & \text{第}i\text{次试验中事件}A\text{发生;} \\ 0, & \text{第}i\text{次试验中事件}A\text{不发生.} \end{cases}$$

$X_1, X_2, \cdots, X_n, \cdots$独立都服从$(0-1)$分布，

$$E(X_i) = p, D(X_i) = p(1-p) \le \frac{1}{4}, i = 1, 2, \cdots$$

由切比雪夫大数律或独立同分布大数定律,

$$\forall \varepsilon > 0,\ \lim_{n\to\infty} P\{|\frac{m}{n} - p| < \varepsilon\} = 1$$

Th5.2.4. 伯努里(*Bernulli*)大数定律

设 $\frac{m}{n}$ 是n次独立重复试验中事件A发生的频率，p是事件A在每次试验中发生的概率，则对于

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} P\{|\frac{m}{n} - p| < \varepsilon\} = 1$$

注: 1) 此定律为切比雪夫大数或独立同分布定律的一个推论.

2) 此定律以严格的数学形式描述了频率的稳定性.

正因为这种稳定性, 概率的概念才有客观意义.

3) 可得小概率事件原理:

概率很小的事件, 在一次试验中几乎是不可能发生的, 从而在实际中可看成不可能事件.

二战后期，一批从事原子弹研究的美国科学家，与第一架数字电子计算机的诞生几乎同时提出了一系列随机模拟的计算方案:一大类是关于粒子运动的模拟，另一大类则是以积分计算为代表的新计算方法. 他们还以赌城的名字命名这类方法，这就是著名的蒙特卡罗**(Monte Carlo)**方法.下面介绍定积分计算的蒙特卡罗法.

设在 $[a,b]$ 上，函数 $0 \leq f(x) \leq M$，要求计算

$$I = \int_a^b f(x)\,\mathrm{d}x$$

在图 1.4.3 中，积分 I 等于阴影所标的曲边梯形的面积. 它被以 $\overline{ab}$ 为边，高为 M 的矩形 G 所包围，假如我们在 G 中随机地取点，则该点落在阴影部分的概率为

$$P = \frac{I}{(b-a)M}$$

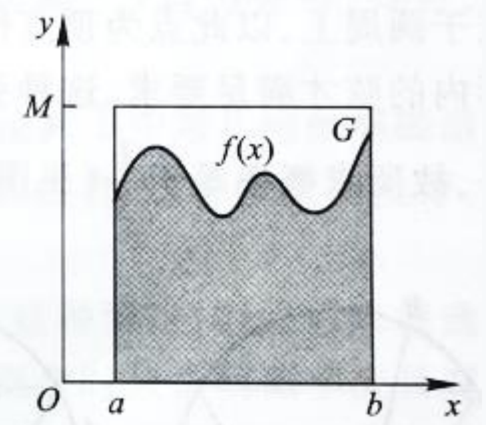

图 1.4.3 积分计算

仿照蒲丰投针问题的做法，投点 N 次，计算落入阴影部分的次数 n，并以频率值 $\frac{n}{N}$ 作为概率 P 之值代入上式，求得

$$I = \frac{nM(b-a)}{N}$$

这个做法不难推广到多重积分与任意边界的场

蒙特卡洛方法

Monte Carlo方法是计算机模拟的基础，它的名字来源于世界著名的赌城——摩纳哥的蒙特卡洛， 其历史起源于 1777 年法国科学家蒲丰提出的一种计算圆周π 的方法——随机投针法，即著名的蒲丰投针问题。

“在平面上画有一组间距为d的平行线，将一根长度为l(l<d)的针任意掷在这个平面上，求此针与平行线中任一条相交的概率。”

Monte Carlo方法的基本思想是首先建立一个概率模型，使所求问题的解正好是该模型的参数或其他有关的特征量. 然后通过模拟一统计试验， 即多次随机抽样试验 (确定 m和 n)，统计出某事件发生的百分比. 只要试验次数很大，该百分比便近似于事件发生的概率.

这实际上就是概率的统计定义.

利用建立的概率模型，求出要估计的参数.

蒙特卡洛方法适用范围很广泛，它既能求解确定性的问题，也能求解随机性的问题以及 科学研究中的理论问题.

例如利用蒙特卡洛方法可以近似地计算定积分，即产生数值积分问题.

4. 辛钦大数定律(补充)

设X_k, k=1,2…是相互独立且分布相同的随机变量序列, 且 $E(X_i)=\mu$存在, $i=1,2,\ldots,$

则 $\{X_k\}$ 服从大数定律.

即对任意的$\varepsilon>0$,有

$$\lim_{n\to\infty} P\{|\frac{1}{n}\sum_{k=1}^{n} X_k-\mu|<\varepsilon\}=1$$

注: 1) 独立同分布大数定律更一般的结论.

2) 对方差没有限制, 即方差可以不存在.

3) 证明不要求.(不同于前面3个定律证明)

大数定律的应用:

实际工作中, 以大量测量值的平均值作为精确值的估计值 → 以独立同分布大数定律 为理论依据

频率的稳定性：试验次数充分多，则A发生的频率 $\frac{m}{n}$ 趋于A发生的概率p → 以伯努里大数定律为理论依据

小概率事件原理: 概率很小的事件, 在一次试验中几乎不可能发生, 实际中看作不可能事件→

以伯努里大数定律 为理论依据

三倍标准差原理(3σ原理)： 若 $X \sim N(\mu, \sigma^2)$，则当X取值在$(\mu - 3\sigma, \mu + 3\sigma)$之外时，看作小概率事件 →

$$P\{|X - \mu| > 3\sigma\} = 0.0026$$

大数定律：(“频率依概率收敛于概率值”引申而来)

只是对随机变量序列前n项算术平均值与它的期望值之差依概率收敛于0, 给出了<u>定性</u>的说明:

$$\lim_{n\to\infty} P\{|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)| < \varepsilon\} = 1$$

切比雪夫不等式只能对概率作出很粗略的估计.

$$P\{|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)| < \varepsilon\} \geq 1 - \frac{D(\sum_{i=1}^{n} X_i)}{\varepsilon^2}$$

需要较精确的概率值 $P\{\underline{\overset{-\varepsilon}{a}} < \frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i) < \underline{\overset{\varepsilon}{b}}\}$

中心极限定理给出<u>定量</u>的说明

以保险业为例来说明人类如何利用大数定律以增进社会福祉。

天有不测风云,人有旦夕祸福,自然界和社会生活都充满着不确定性.一场灾害或事故可使一家人立即陷入经济困境,个人对此无能为力,但社会却可建立损失分担机制,保险业与精算学应运而生。

作为精算科学基础的保险原理,即损失分担原理,可以认为是大数定律用保险术语的重述.精算学的研究主题是分析巨大的、不可预测的损失的各种金融后果,并设计某种机制以缓冲这类损失的有害的金融效应。

例如在财产保险中每月支付一笔小钱以求在一场大火或失窃后得到高额赔偿,便是人类对抗偶然性的有效方法.为规避某类风险,人们找中介即保险公司,保险公司则创造满足下面4个条件的群体:1.损失是不可预测的;2.风险是独立的;3.风险是齐性的;4.这个群体相当大,使得各个个体要求赔偿的整个损失额变成相对确定。

把上述4个条件与前面讲过的或后面将要讲到的大数定律成立的条件作对比.因此可以说,保险业是人类利用大数定律的范例。