西北工業大學
NORTHWESTERN POLYTECHNICAL UNIVERSITY

概率论与数理统计

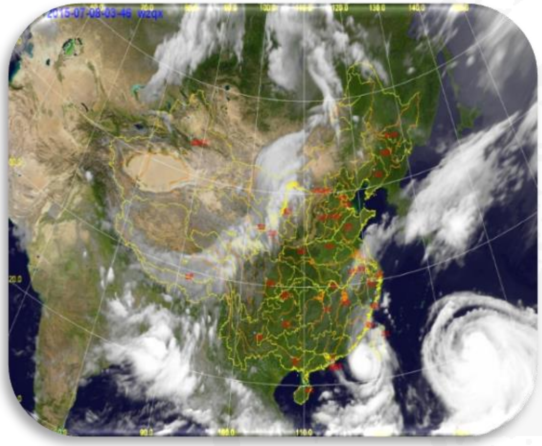
徐 爽

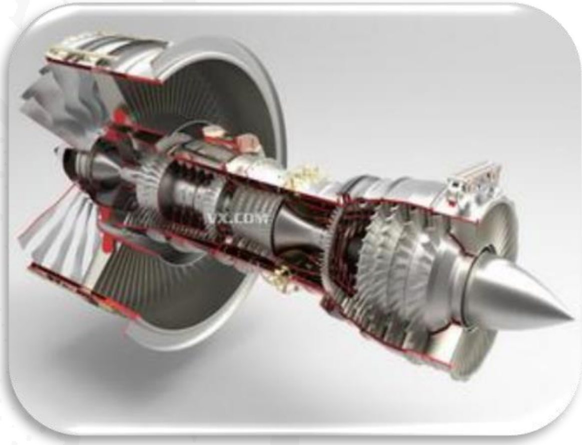
西北工业大学理学院概率统计教研室

第一节 大数定律

- 一、问题的提出
- 二、随机变量序列的收敛性
- 三、伯努利大数定律
- 四、常用的四种大数定律

一、问题的提出





万物看似随机，但有其统计的宿命。

随机现象 → **统计规律性**

引例 相同条件下射击， 命中率？

A：射击射中 $P(A) = ?$

随机现象 → 随机试验 → 揭示内在规律

n：射击次数

μ_n：命中次数

$$f = \frac{\mu_n}{n} \approx P(A)$$

命中频率≈ 命中率

理论依据 ？

二、随机变量序列的收敛性

定义4.2 设**随机变量序列** $\{Y_n\}$ 和随机变量Y，若对任意实数 $\varepsilon > 0$，有

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

或

$$\lim_{n\to\infty} P\{|Y_n - Y| \geq \varepsilon\} = 0$$

则称随机变量序列 $\{Y_n\}$ **依概率收敛**于随机变量Y，

简记为

$$Y_n \xrightarrow{P} Y$$

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

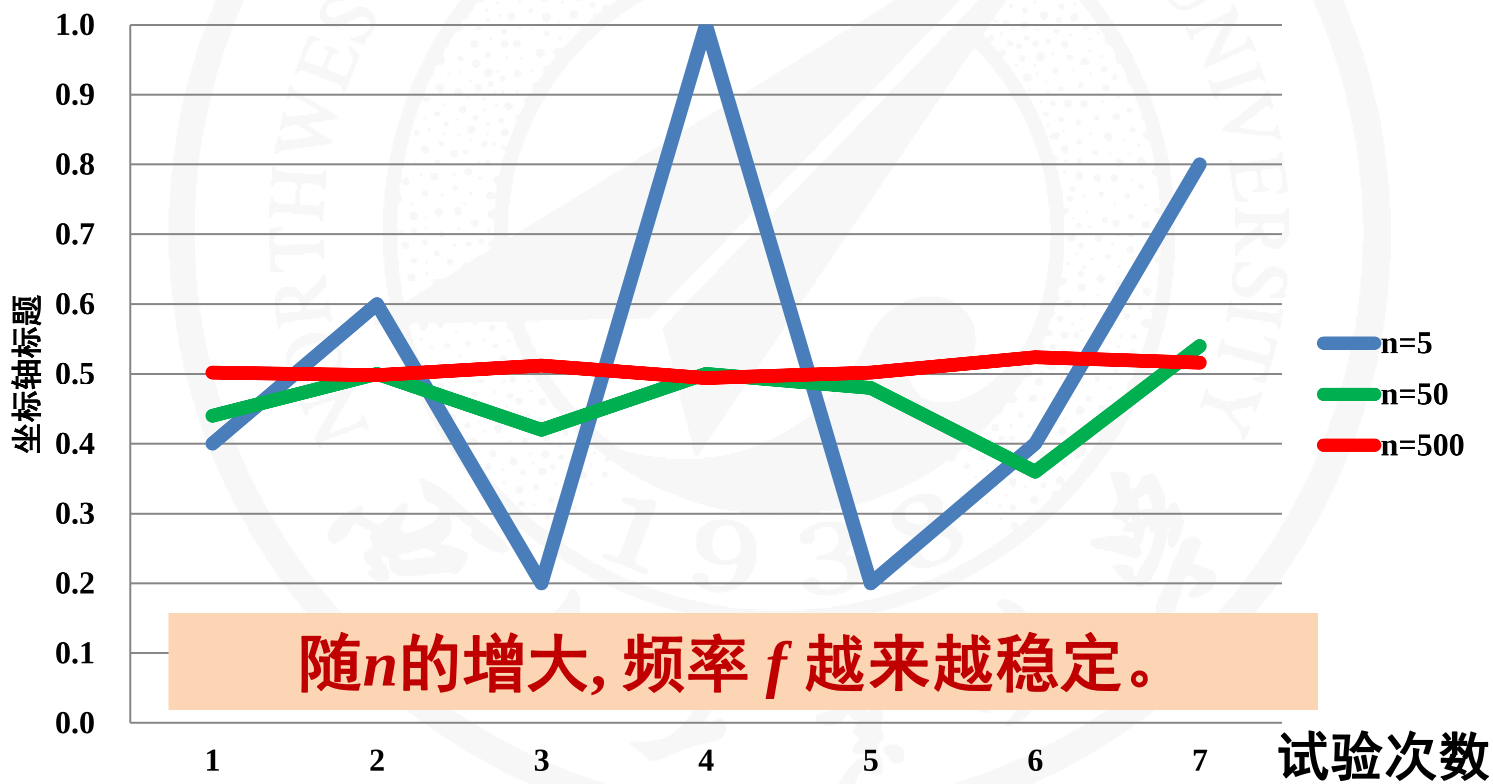
依概率收敛表示：Y_n 与 Y 的绝对误差小于任意小的正数 ε 的**可能性(即概率)**将随着n增大而愈来愈大，直至趋于1.

定理4.1 设 $\{Y_n\}$ 为一随机变量序列，$Y_n \xrightarrow{P} C$ (常数)，且函数$g(\cdot)$ 在点C处连续，则有

$$g(Y_n) \xrightarrow{P} g(C).$$

三、伯努利大数定律

1. 频率的稳定性 将一枚硬币抛掷 5 次、50 次、500 次, 各做 7 遍, 观察正面出现的次数及频率.



试验者	n	μ_n	f
德.摩根	2048	1061	0.5181
蒲丰	4040	2048	0.5069
费勒	10000	4979	0.4979
皮尔逊	12000	6019	0.5016
杰万斯	20480	10379	0.5068
罗曼.诺夫斯基	80640	39699	0.4932

$$f_n \approx 0.5 = P(\text{正面})$$

频率的稳定性

频率的稳定性： 概率统计定义的理论基础

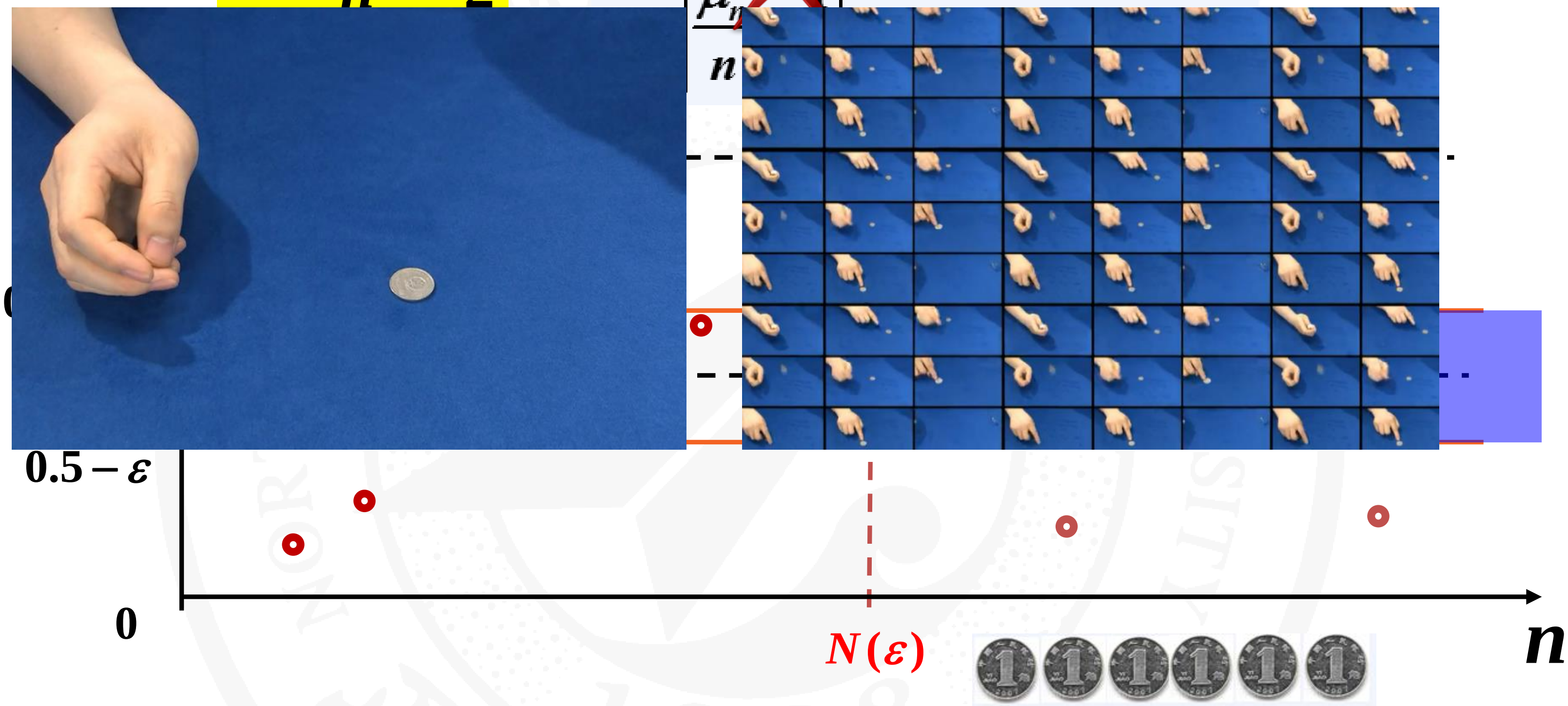
随机现象在大量重复试验中呈现明显的统计规律性，即事件发生的频率具有稳定性.

频率 $\xrightarrow{n\to\infty}$ 概率

数学上如何准确刻画频率的稳定性？ 数列极限？

$$\lim_{n\to\infty}\frac{\mu_n}{n}=\frac{1}{2}$$

$\Leftrightarrow$ 任意实数 $\varepsilon>0$，存在N，当$n>N$，有



极端情况：$\dfrac{\mu_n}{n}=1,\quad \dfrac{\mu_n}{n}=0 \Rightarrow \qquad \left|\dfrac{\mu_n}{n}-\dfrac{1}{2}\right|=\dfrac{1}{2}>\varepsilon$


$$P\left\{\frac{\mu_n}{n}=1\right\}=\frac{1}{2^n},\quad P\left\{\frac{\mu_n}{n}=0\right\}=\frac{1}{2^n}$$

$$P\left\{\left|\frac{\mu_n}{n}-\frac{1}{2}\right|=\frac{1}{2}\right\}=P\left\{\frac{\mu_n}{n}=0\right\}+P\left\{\frac{\mu_n}{n}=1\right\}$$

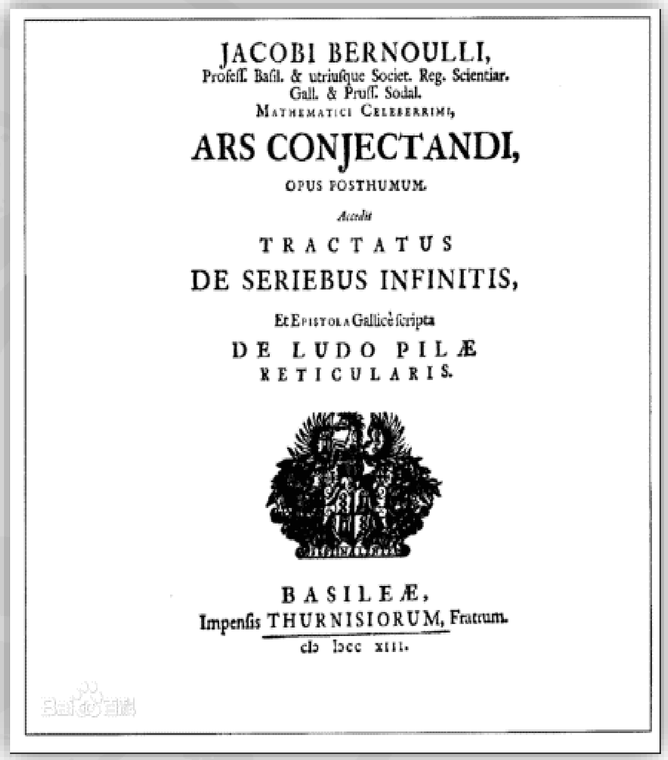
$$=\frac{1}{2^{n-1}}\xrightarrow{n\to\infty}0$$

一般情况：

$$\forall\varepsilon>0,\quad P\left\{\left|\frac{\mu_n}{n}-\frac{1}{2}\right|\geq\varepsilon\right\}\xrightarrow{n\to\infty}0$$

瑞士数学家 伯努利

Jakob Bernoulli (1654-1705)



JACOBI BERNOULLI,
Profess. Basil. & utriusque Societ. Reg. Scientiar.
Gall. & Pruss. Sodal.
MATHEMATICI CELEBERRIMI,

ARS CONJECTANDI,

OPUS POSTHUMUM.

Accedit

TRACTATUS
DE SERIEBUS INFINITIS,

Et EPISTOLA Gallicè scripta
DE LUDO PILÆ
RETICULARIS.

BASILEÆ,
Impensis THURNISIORUM, Fratrum.
cIↄ Iↄcc xiii.

伯努利大数定律

2. 伯努利大数定律

设 μ_n 是 n 重伯努利试验中事件 A 发生的次数，

p 为每次试验中事件 A 发生的概率，则对于任意 $\varepsilon \geq 0$，

有 $$\lim_{n\to\infty} P\left\{ \left|\frac{\mu_n}{n} - p\right| \geq \varepsilon \right\} = 0$$

依概率收敛性： 即 $$\frac{\mu_n}{n} \xrightarrow{P} p$$

当 n 很大，频率与概率有较大偏差的可能性很小

证明:

$\because \mu_n$ 是 n 重伯努利试验中事件A发生的次数,且 $P(A)=p$

$\therefore \mu_n \sim B(n,p)$ 且有 $E(\mu_n)=np;\ D(\mu_n)=np(1-p);$

令 $X=\dfrac{\mu_n}{n} \Rightarrow E(X)=E(\dfrac{\mu_n}{n})=p;\ D(X)=D(\dfrac{\mu_n}{n})$

$$=\frac{D(\mu_n)}{n^2}=\frac{p(1-p)}{n}$$

由Chebyshev不等式

$$0\le P\left\{\left|\frac{\mu_n}{n}-p\right|\ge\varepsilon\right\}\le\frac{p(1-p)}{n\varepsilon^2}\xrightarrow{n\to\infty}0$$

$$\therefore \lim_{n\to\infty}P\left\{\left|\frac{\mu_n}{n}-p\right|\ge\varepsilon\right\}=0$$

证明完毕.

历史上第一个大数定律，奠定了概率论的理论基础！

意义：以严格的数学形式证明了频率的稳定性，揭示了随机现象中的统计规律性。 $\frac{\mu_n}{n} \xrightarrow{P} p$

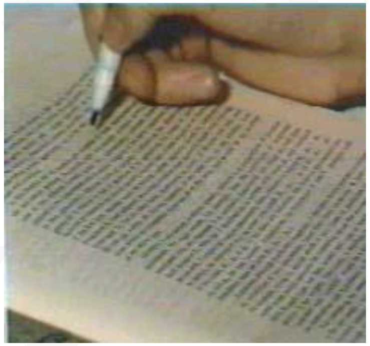
当$n \to \infty$时，

射中的频率$f_n \approx$ 命中率$P(A)$

理论依据：伯努利大数定律

四、常用的四种大数定律

在实践中，人们认识到单个随机现象没有规律，但大量随机现象**的算术平均值**具有稳定性. 大数定律就是用于研究**大量随机现象中平均结果的稳定性**的理论.



$$X \sim B(100, 0.3) \Rightarrow \quad Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$n = 100$

30.58, 29.91, 29.79……

$n = 10000$

30.0019, 30.0494, 29.9655……

$n = 1000000$

30.0065, 30.0115, 29.9991

$n \to \infty$

$$Y_n \approx 30 = EX = EY_n$$

$X \sim N(4,3) \Rightarrow$ $Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

$n = 100$

3.9481, 3.9846, 4.5133……

$n = 10000$

4.0177, 3.9876……

$n = 1000000$

4.0007, 4.0028,……

$n \to \infty$

$Y_n \approx 4$ $= EX = EY_n$

定义4.5 大数定律

设$X_1,X_2,\cdots,X_n,\cdots$是随机变量序列，

令 $$Y_n=\frac{1}{n}\sum_{i=1}^{n}X_i$$

如果存在这样一个常数序列$a_1,a_2,\cdots,a_n,\cdots$，

对任意的$\varepsilon>0$，恒有

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

则称随机变量序列$\{X_n\}$服从大数定律.

描述了随机变量序列均值的收敛性

定理4.3 切比谢夫大数定律

设$X_1, X_2, \cdots, X_n, \cdots$是两两不相关的随机变量序列,每一随机变量都有有限的方差,并有公共的上界

$$D(X_1) \le C, D(X_2) \le C, \cdots, D(X_n) \le C, \cdots$$

则对任意的$\varepsilon > 0$,恒有

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

即 $$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \frac{1}{n}\sum_{i=1}^{n} EX_i$$

注1° 当 n 很大时,随机变量 $X_1, X_2, \cdots, X_n$ 的算术平均值 $\frac{1}{n}\sum_{i=1}^{n} X_i$ 接近于它们的数学期望的算术平均值 $\frac{1}{n}\sum_{i=1}^{n} E(X_i)$.

这种接近是概率意义下的!

通俗地说, 在定理条件下, n 个随机变量的算术平均值, 当 n无限增加时, 几乎变成一个常数.

定理4.4 伯努利大数定律

设 μ_n 是 n 次独立重复伯努利试验中事件 A 发生的次数，p 是事件 A 在每次试验中发生的概率，则对任意的 $\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right\} = 1$$

即 $\dfrac{\mu_n}{n} \xrightarrow{P} p$

证 引入随机变量

$$X_k=\begin{cases}0, & \text{在第}k\text{次试验中事件}A\text{不发生}\\ 1, & \text{在第}k\text{次试验中事件}A\text{发生}\end{cases}$$

$k=1,2,\cdots,n.$

显然,由于$X_1,X_2,\cdots,X_n$是相互独立的,且同服从$B(1,p)$分布,故有

$$E(X_k)=p,D(X_k)=p(1-p)\le\frac{1}{4}\quad(k=1,2,\cdots,n)$$

由**定理4.3**对任意的$\varepsilon>0$,有

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

即

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right\} = 1$$

证毕.

定理4.4 推广 泊松大数定律*

如果在一个独立试验序列中,事件A在第k次试验中出现的概率等于p_k, 以μ_n记在前n次试验中事件A出现的次数, 则对任意的 $\varepsilon > 0$, 都有

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu_n}{n} - \frac{1}{n}\sum_{k=1}^{n} p_k\right| < \varepsilon\right\} = 1$$

即 $\frac{\mu_n}{n} \xrightarrow{p} \frac{1}{n}\sum_{k=1}^{n} p_k$

证 令

$$X_k = \begin{cases} 0, & \text{第}k\text{次试验中}A\text{不发生} \\ 1, & \text{第}k\text{次试验中}A\text{发生} \end{cases} \quad (k = 1, 2, \cdots, n)$$

X_k独立且服从$B(1, p_k)$ 由定理4.3可得结论.

定理4.5 辛钦大数定律

设随机变量 $X_1, X_2, \cdots, X_n$ 相互独立,服从同一分布,且具有数学期望 $E(X_i)=\mu \quad (i=1,2,\cdots,n)$, 则对任意的 $\varepsilon>0$, 都有

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

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

注1° 与切比谢夫大数定理相比, 不要求方差存在且有界.

2° 伯努利大数定理是辛钦大数定理的特例.

例2 设$X_1, X_2, \ldots, X_n$是独立同分布的随机变量序列, $E(X_i)=\mu$, $D(X_i)=\sigma^2$均存在, 证明

$$Y_n=\frac{2}{n(n+1)}\sum_{i=1}^{n} iX_i$$

依概率收敛到 μ.

解 大数定律

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

切比谢夫不等式

$$0 \le P\left\{\left|Y_n - E(Y_n)\right| \ge \varepsilon\right\} \le \frac{D(Y_n)}{\varepsilon^2}$$

其中 $$Y_n = \frac{2}{n(n+1)}\sum_{i=1}^{n} iX_i$$

因为 $$E(Y_n) = E\left[\frac{2}{n(n+1)}\sum_{i=1}^{n} iX_i\right]$$

$$= \frac{2}{n(n+1)}\sum_{i=1}^{n} iE(X_i) = \frac{2\mu}{n(n+1)}\sum_{i=1}^{n} i = \mu$$

$$D(Y_n)=\frac{4}{n^2(n+1)^2}\sum_{i=1}^{n} i^2 D(X_i)=\frac{4\sigma^2}{n^2(n+1)^2}\sum_{i=1}^{n} i^2$$

$$=\frac{4n(n+1)(2n+1)\sigma^2}{6n^2(n+1)^2}=\frac{2(2n+1)\sigma^2}{3n(n+1)}$$

从而对任意给定的$\varepsilon>0$, 由**切比谢夫不等式**得

$$0\le P\{|Y_n-\mu|\ge\varepsilon\}\le\frac{D(Y_n)}{\varepsilon^2}$$

$$=\frac{2(2n+1)\sigma^2}{3n(n+1)\varepsilon^2}\to 0\quad(n\to\infty)$$

因此 $Y_n \xrightarrow{P} \mu.$

例3 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 由**辛钦大数定律**知,

$$\frac{1}{n}\sum_{k=1}^{n} X_k^2 \xrightarrow{P} \frac{1}{n}\sum_{k=1}^{n} E(X_k^2)=E(X_k^2)$$

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**辛钦大数定律**知, 对于任意正数 ε, 有

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

即 $\frac{1}{n}\sum_{k=1}^{n} X_k^2 \xrightarrow{P} E(X_k^2)$

该结论为数理统计中的**矩估计法**的理论基础。

内容小结

四个大数定理

切比谢夫大数定律

不相关、方差有界：$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \frac{1}{n}\sum_{i=1}^{n} EX_i$

伯努利大数定律

独立同分布，$X \sim B(1,p)$ ：$\frac{\mu_n}{n} \xrightarrow{p} p$

泊松大数定律

独立不同分布，$X \sim B(1,p_k)$ ：$\frac{\mu_n}{n} \xrightarrow{p} \frac{1}{n}\sum_{k=1}^{n} p_k$

辛钦大数定律

独立同分布：$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} E(X_i)$

4-1 大数定律

Thank You !

西北工業大學
NORTHWESTERN POLYTECHNICAL UNIVERSITY

概率论与数理统计

徐 爽

西北工业大学理学院概率统计教研室

第二节 中心极限定理

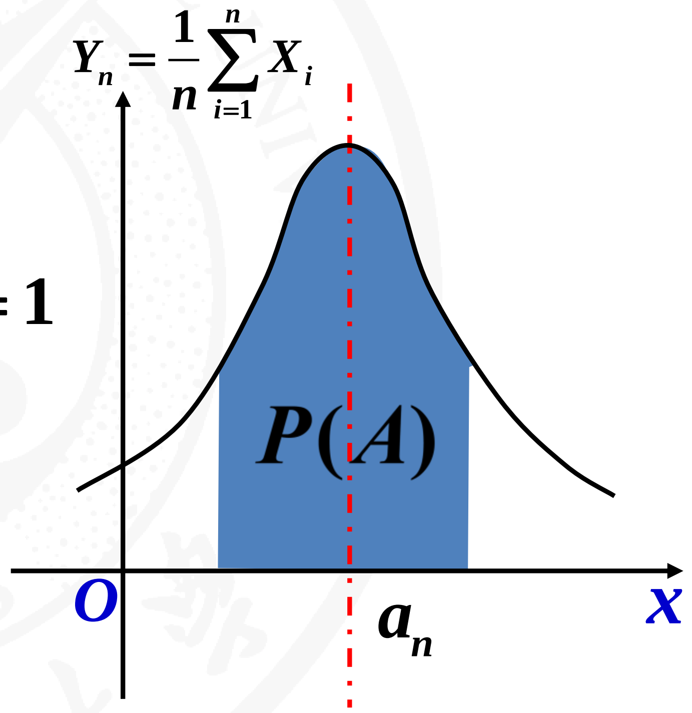
- 一、问题的提出
- 二、中心极限定理

一、问题的提出

大数定律： 满足一定条件的随机变量序列的**算数平均值依概率收敛。**

$$\lim_{n\to\infty} P\Bigg\{\underbrace{\left|\frac{1}{n}\sum_{i=1}^{n}X_i - \frac{1}{n}\sum_{i=1}^{n}E(X_i)\right| < \varepsilon}_{A}\Bigg\} = 1$$

$$\boxed{Y_n = \frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{p} E(Y_n)}$$



若$X_i \sim N(2,3)$, 则$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{p} 2$

X_1,	X_2,	X_3,	X_4,	X_5,	X_6,	X_7,	X_8,	X_9,	X_{10},	$\overline{X}$
4.5211	-4.4151	10.7240	0.9385	2.0687	3.5602	1.1187	-1.9960	-2.0851	1.4143	2.1200
-0.6641	-0.5188	4.4757	-0.4708	1.2140	1.9399	-0.5438	-4.9896	3.3651	1.3472	1.6740
2.3003	6.0638	6.1369	-2.7312	-3.2506	1.8957	-1.3604	-2.3473	-0.5461	1.0907	1.8494
0.3664	-1.2165	-1.1745	3.5239	1.1430	-0.3945	9.5780	3.0005	0.9953	2.0691	1.6199
2.9106	4.8829	0.5942	2.8460	-0.4941	5.0561	6.9665	3.1741	3.6584	2.1539	2.0553
0.1990	2.3721	1.1826	2.1004	-0.9376	1.6003	2.9226	3.3550	5.1173	4.4782	1.6360
3.4699	6.3101	5.2953	-2.0010	-1.4692	-0.1436	-1.7714	1.6091	-1.3529	6.5809	1.9810
4.2181	-3.8827	1.1664	5.3825	0.3993	6.0542	-0.5964	2.5511	5.7820	3.4007	1.8734
7.1357	1.4069	4.1046	3.0505	-4.0079	1.3257	1.4704	0.5715	3.9804	1.3709	2.1550
1.4176	-1.6235	-4.1554	1.1028	4.8927	0.2329	4.3742	4.5861	1.7964	3.8756	1.9405

但我们无法得知其收敛的速度，即：

给定n, $P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-2\right|<\varepsilon\right\}=?$

$$Y_n=\frac{1}{n}\sum_{i=1}^{n}X_i \sim F(x)?$$

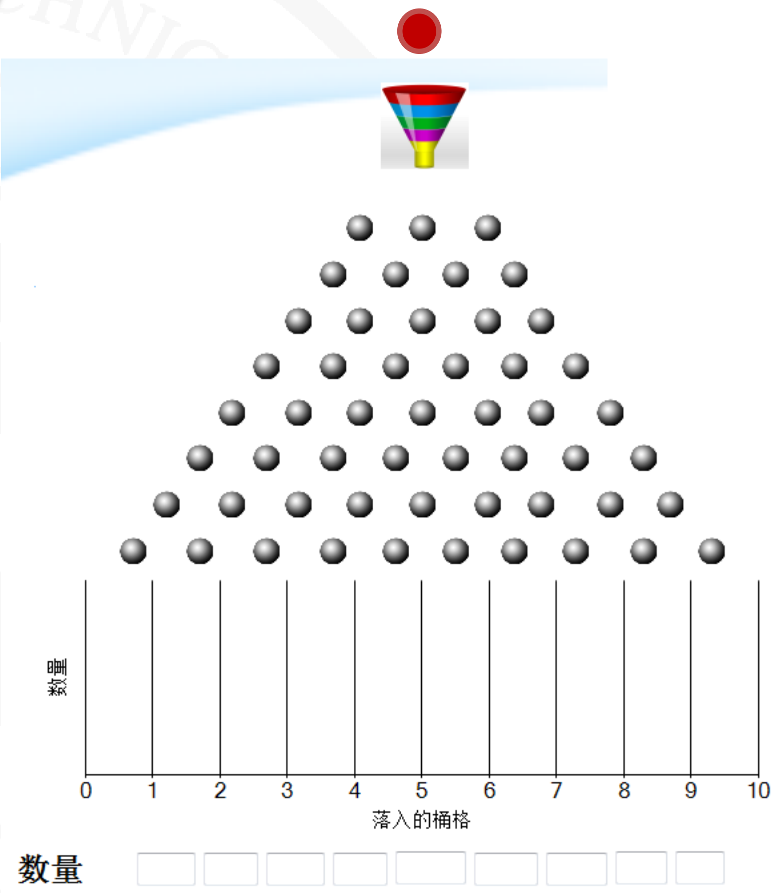
引例 高尔顿(Galton)板实验

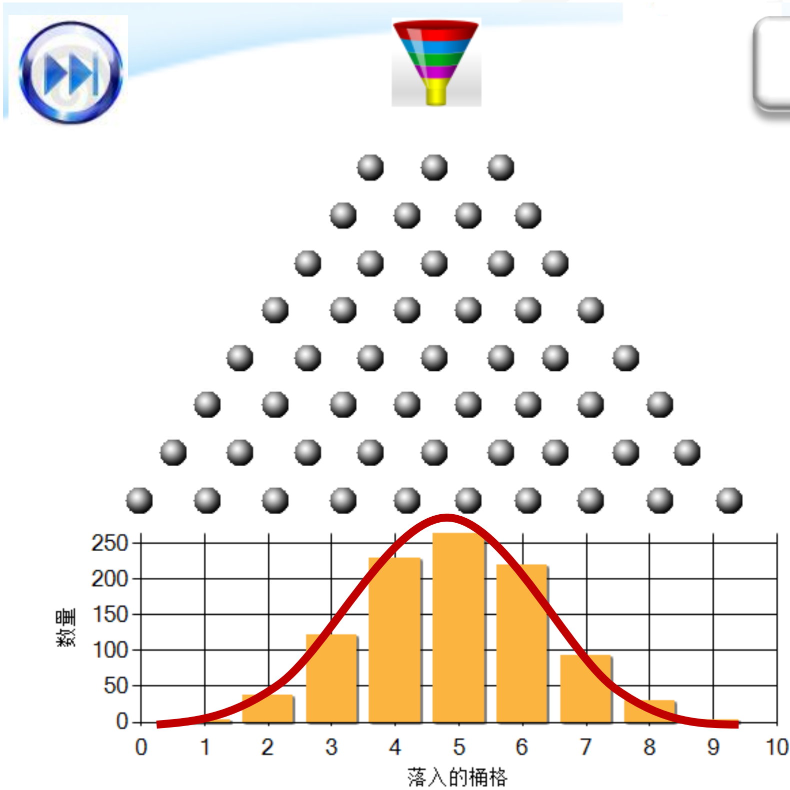
n层钉子

$P_{左}=P_{右}=1/2$

每个格子中小球的个数

统计特性 ?





数量	1	38	122	230	263	219	94	30	3

分析

令 $X_k=\begin{cases}1 & \text{第}k\text{层向左}\\ -1 & \text{第}k\text{层向右}\end{cases}$

则 $X_k\sim B(1,\frac{1}{2}),\ \ k=1,2,\cdots n,$

且 $X_1,X_2,\cdots,X_n$ 相互独立。

小球下落到最底层的水平位置

$$Y_n=\sum_{k=1}^{n}X_k\sim B(n,\frac{1}{2})$$

$$\sim N(\mu,\sigma^2)$$

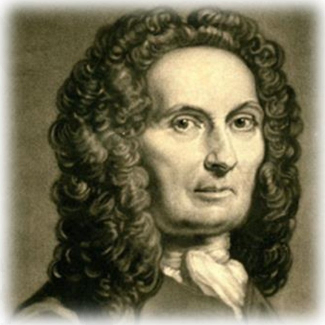
二、中心极限定理

定理4.7　棣莫佛-拉普拉斯($De\ Movire - Laplace$)中心极限定理

设随机变量Y_n服从二项分布$B(n, p)$,则对任意x，

有

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



A. De Moivre (1667-1754)

表 明　若$Y_n \sim B(n,p)$,

$$E(Y_n) = np$$
$$D(Y_n) = np(1-p)$$

则有$Y_n^* = \dfrac{Y_n - np}{\sqrt{np(1-p)}} \sim AN(0,1)$

P.-S. Laplace (1749-1827)

近似服从标准正态分布

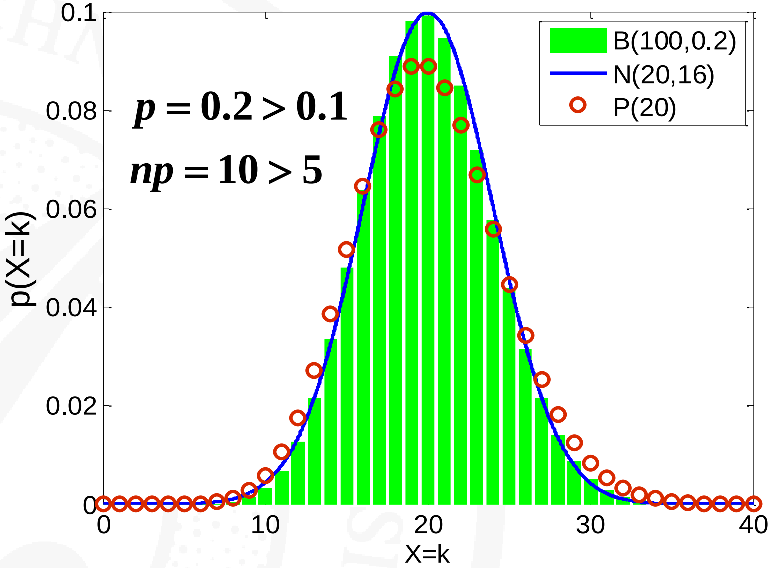
正态分布是二项分布的极限分布

$np \geq 5 \ \& \ n(1-p) \geq 5$

二项分布 $B(n,p)$ $\xrightarrow{n\to\infty}$ 正态分布 $N(\mu,\sigma^2)$

$p \leq 0.1, np \to \lambda < 5$

二项分布 $B(n,p)$ $\xrightarrow{n\to\infty}$ 泊松分布 $P(\lambda)$



$$X \sim B(100, 0.2)$$

$$P(X \leq b) = \sum_{k=0}^{b} C_{100}^{k} 0.2^{k} 0.8^{100-k}$$

$$\approx \Phi\left(\frac{b-20}{\sqrt{16}}\right)$$

中心极限定理

De Moivre-Laplace中心极限定理

$$Y_n \sim B(n,p) \Longrightarrow \frac{Y_n - np}{\sqrt{np(1-p)}} \sim AN(0,1)$$

$$\downarrow \qquad\qquad \downarrow$$

$$Y_n = \sum_{i=1}^{n} X_i \qquad\qquad \boxed{\frac{\sum_{i=1}^{n} X_i - np}{\sqrt{np(1-p)}} \sim AN(0,1)}$$

其中$\boxed{X_i \sim B(1,p)}$

特殊 $X_i \sim B(1,p)$ $\Rightarrow$ 一般$X_i \sim F(x)$?

定理4.6 林德伯格-列维(*Lindeberg-Levy*)中心极限定理

设随机变量$X_1, X_2, \ldots, X_n$相互独立, 服从同一分布, 且存在

$$E(X_i)=\mu,\quad D(X_i)=\sigma^2\neq 0\quad (i=1, 2, \ldots, n)$$

则对任意x,有

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

（证明略）

表明

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

独立同分布

$$\Rightarrow E\left(\sum_{i=1}^{n} X_i\right)=n\mu, D\left(\sum_{i=1}^{n} X_i\right)=n\sigma^2$$

大量相互<u>独立同分布</u>的随机变量的和

标准化后近似服从标

X_i之间没有相互依赖关系

对总和的影响在概率意义下均衡

$$Y_n^*=\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \sim AN(0,1)$$

$$\sum_{i=1}^{n} X_i \sim AN\left(n\mu, n\sigma^2\right)$$

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

例2 一船舶在某海区航行, 已知每遭受一次海浪的冲击, 纵摇角大于 3° 的概率为1/3, 若船舶遭受了90000次波浪冲击, 问其中有29500～30500次纵摇角大于 3° 的概率是多少?

解 设A:纵摇角大于 3°，则P(*A*)=1/3

假设海浪冲击相互独立。

X： 90000次波浪冲击中纵摇角大于 3° 的次数

则$X \sim B(90000, 1/3)$. 分布律为

$$P\{X=k\}=\binom{90000}{k}\left(\frac{1}{3}\right)^{k}\left(\frac{2}{3}\right)^{90000-k},k=1,\cdots,90000.$$

所求概率为 $P\{29500 < X \le 30500\}$

$$= \sum_{k=29501}^{30500} \binom{90000}{k} \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{90000-k}$$

直接计算很麻烦，利用**切比雪夫不等式**

$$\because n = 90000, \quad p = \frac{1}{3},$$

$P\{29500 < X \le 30500\}$

$$= P\{29500 - np < X - np \le 30500 - np\}$$

$$= P\left\{\left|X - 30000\right| \le 500\right\} \ge 1 - \frac{npq}{500^2} \approx 0.92$$

只能估计范围，利用**棣莫佛-拉普拉斯定理**

$$
\begin{aligned}
&P\{29500 < X \le 30500\} \\
&= P\left\{\frac{29500-np}{\sqrt{np(1-p)}} < \frac{X-np}{\sqrt{np(1-p)}} \le \frac{30500-np}{\sqrt{np(1-p)}}\right\} \\
&= \Phi\left(\frac{30500-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{29500-np}{\sqrt{np(1-p)}}\right) \\
&\approx \Phi\left(\frac{5\sqrt{2}}{2}\right) - \Phi\left(-\frac{5\sqrt{2}}{2}\right) \\
&= 2\Phi(3.53) - 1 \ = 0.9995.
\end{aligned}
$$

例3 设随机变量$X_1, X_2,\ldots, X_n$相互独立，且 X_i 在区间[–1, 1] 上服从均匀分布(i=1, 2,…, n)，试证当 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 记 $Y_i = X_i^2, \quad (i = 1,2,\cdots,n)$

因为$X_1, X_2,\ldots, X_n$相互独立，所以$Y_1, Y_2,\ldots,Y_n$相互独立，根据**林德伯格–列维中心极限定理**

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 = \frac{1}{n}\sum_{i=1}^{n} Y_i \quad \sim AN\left(E(Y_i), \frac{D(Y_i)}{n}\right)$$

$X_i \sim U[-1, 1]$ $\quad E(X_i) = 0, D(X_i) = \dfrac{1}{3}$

$$E(Y_i) = E(X_i^2) = D(X_i) + E^2(X_i) = \frac{1}{3}$$

$$D(Y_i) = E(Y_i^2) - [E(Y_i)]^2 = E(X_i^4) - [E(Y_i)]^2$$

因为 $E(X_i^4) = \int_{-1}^{1} x_i^4 \cdot \dfrac{1}{2} \mathrm{d}x_i = \dfrac{1}{5},$

所以 $D(Y_i) = \dfrac{1}{5} - \left(\dfrac{1}{3}\right)^2 = \dfrac{4}{45},$

$$\therefore Z_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \sim N\left(\frac{1}{3}, \frac{4}{45n}\right).$$

例5 设总体X服从参数为2的指数分布，$X_1, X_1, \cdots X_n$为来自总体X的简单随机样本，则当$n\to\infty$时，随机变量序列$Y_n=\frac{1}{n}\sum_{i=1}^{n}X_i^2$依概率收敛于 $\underline{\quad \frac{1}{2}. \quad}$。

例6 设随机变量$X_1, X_1, \cdots X_n$相互独立，均服从$B(1, 0.2)$，则由中心极限定理知，当$n\to\infty$时，随机变量$U=\sum_{i=1}^{n}X_i$近似服从 $\underline{N(0.2n, 0.16n)}$ 分布.（写出分布参数）.

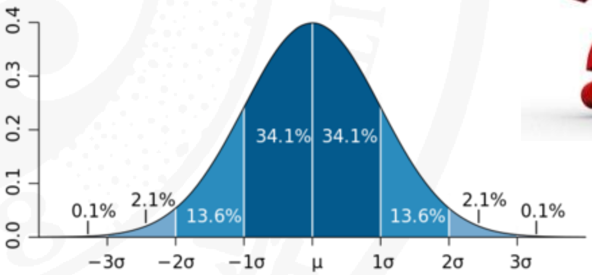
课后思考

中心极限定理

独立同分布

独立不同分布

$$\sum_{i=1}^{n} X_i \overset{\text{近似}}{\sim} \text{正态分布}$$



内容小结

$E(X_i)=\mu \quad D(X_i)=\sigma^2$

中心极限定理

独立同分布情形

林德贝格－列维中心极限定理

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \sim AN(0,1)$$

独立不同分布情形

李雅普诺夫定理

$$Y_n^* = \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mu_i}{B_n} \sim AN(0,1)$$

二项分布的正态近似

棣莫佛－拉普拉斯定理

$$Y_n = \sum_{k=1}^{n} X_k \sim B(n,p)$$

$$\frac{Y_n - np}{\sqrt{np(1-p)}} \sim AN(0,1)$$

4-2 中心极限定理

Thank You !