西北工業大學
NORTHWESTERN POLYTECHNICAL UNIVERSITY

概率论与数理统计

下 回 停

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、常用的四种大数定律

一、问题的提出

在第一章有关概率的统计定义中讲到，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.

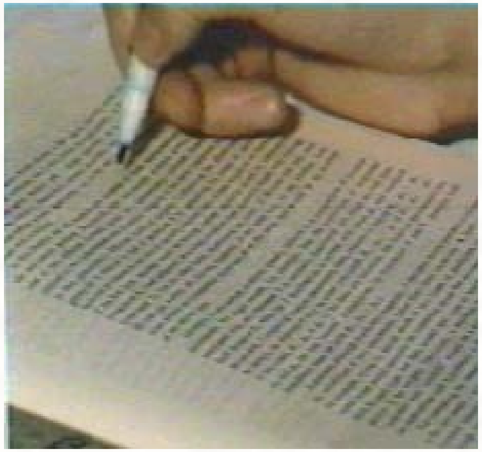
贝努里于1713年首先提出关于频率稳定性的定理，被称为贝努里大数定律.

大数定律的客观背景

在实践中，人们认识到大量测量值的算术平均值也具有稳定性. 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.

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

二、随机变量序列的收敛性

定义4.1 设随机变量 $Y_n(n=1,2,\cdots)$ 和随机变量Y的分布函数分别为$F_n(x)(n=1,2,\cdots)$和$F(x)$,若在$F(x)$ 的所有连续点 x 上都有

$$\lim_{n\to\infty} F_n(x)=F(x)$$

则称随机变量序列 $\{Y_n\}$ 依分布收敛与随机变量Y,简记为

$$Y_n \xrightarrow{L} Y$$

依分布收敛表示：当n充分大时，Y_n 的分布函数 $F_n(x)$ 收敛于Y的分布函数$F(x)$，它是概率论中较弱的一种收敛性.

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

$$Y_n \xrightarrow{P} Y$$

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， $Y_n \xrightarrow{P} C$ 且 (常数)，又函数$g(\cdot)$ 在点C处连续，则有

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

证 由 $g(\cdot)$ 在C处连续可知，对任意实数 $\varepsilon > 0$, 存在实数 $\delta > 0$, 使当 $|y - C| < \delta$ 时，总有

$|g(y) - g(C)| < \varepsilon$, 从而

$$\{|Y_n - C| < \delta\} \subset \{|g(Y_n) - g(C) < \varepsilon|\},$$

$$1 \geq P\{|g(Y_n) - g(C)| < \varepsilon\} \geq P\{|Y_n - C| < \delta\}$$

$$= 1 - P\{|Y_n - C| \geq \delta\} \to 1, n \to \infty$$

这就表明：

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

定义 4.3 设随机变量序列 $\{Y_n\}$和随机变量 Y 对 $r>0$ 时，有 $E|Y_n|^r<\infty(n=1,2\cdots)$ 和 $E|Y|^r<\infty$ ，若

$$\lim_{n\to\infty} E|Y_n-Y|^r=0$$

则称随机变量序列 $\{Y_n\}$ 阶收敛于随机变量Y，简记为

$$Y_n \xrightarrow{r} Y$$

特别的有

1-阶收敛又称为平均收敛，

2-阶收敛又称为均方收敛。

可以证明：均方收敛则平均收敛。

定义 4.4 设随机变量序列 $\{Y_n(\omega)\}$ 和随机变量 $Y(\omega)$，若

$$P\{\omega: \lim_{n\to\infty} Y_n(\omega) = Y(\omega)\} = 1$$

或简记为

$$P\{\lim_{n\to\infty} Y_n = Y\} = 1$$

则称随机变量序列 $\{Y_n\}$以概率1（或几乎处处）收敛于随机变量 Y

简记为

$$Y_n \xrightarrow{a.e} Y$$

目 上 下 回 停 续

下面定理揭示了三种收敛之间的关系。

定理 4.2 设随机变量序列$\{X_n\}$和随机变量 X

(1)若 $X_n \xrightarrow{a.e} X$，则 $X_n \xrightarrow{P} X$;

(2) 若 $X_n \xrightarrow{r} X$，则 $X_n \xrightarrow{P} X$;

(3) 若 $X_n \xrightarrow{P} X$，则 $X_n \xrightarrow{L} X$.

三、常用的四种大数定理

定义4.5 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

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

则称随机变量序列$\{X_n\}$服从大数定律.

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

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} EX_i\right| \ge \varepsilon\right\} = 0$$

证 因为$\{X_n\}$两两不相关, 故

$$D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)\le\frac{C}{n}$$

再由切比谢夫不等式得到

$$0\le P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}EX_i\right|\ge\varepsilon\right\}\le\frac{D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)}{\varepsilon^2}\le\frac{C}{n\varepsilon^2}$$

于是,当$n\to\infty$时,有式(4.2)成立,因此定理(4.3)得证.

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.

2° 切比谢夫大数定理的另一种叙述

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

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

则序列$\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$依概率收敛于$\dfrac{1}{n}\sum_{i=1}^{n} E(X_i)$,

即

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

目 上 下 回 停 续

例1 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 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 $Y_n \xrightarrow{P} \mu.$

目 上 下 回 停 续

例2 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	$\dfrac{1}{2n^2}$	$1-\dfrac{1}{n^2}$	$\dfrac{1}{2n^2}$

问是否满足切比谢夫大数定理？

检验是否有数学期望

解 由题意可知独立性.

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

可见，每个随机变量的数学期望都存在.

因为

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

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

检验是否有有限方差

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

因此，随机变量 $X_n\ (n=1,2,\cdots)$ 有有限的方差，且有公共上界.

故满足**切比谢夫大数定理**的条件.

定理4.4

设$\{Y_n\}$为一随机变量序列,且$Y_n \xrightarrow{P} C$(常数),又函数$g(\cdot)$在点C处连续,则$g(Y_n) \xrightarrow{P} g(C)$.

定理4.5 贝努里大数定理

设μ_n是n次独立重复贝努里试验中事件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

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

目 上 下 回 停 续

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

目 上 下 回 停 续

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - p\right| \geq \varepsilon\right\} = 0$$

即

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

证毕.

注1° 贝努里大数定理表明事件发生的频率 $\dfrac{\mu_A}{n}$ 依概率收敛于事件的概率 p.

用严格的数学形式表达频率的稳定性！

当 n 很大时，事件发生的频率与概率有较大偏差的可能性很小.在实际应用中，当试验次数很大时，便可以用事件发生的频率来代替事件的概率.

定理4.6 泊松大数定理

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

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu_n}{n}-\frac{1}{n}\sum_{k=1}^{n} p_k\right| \geq \varepsilon\right\} = 0$$

证 令

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

由定理4.3可得结论.

定理4.7 辛钦大数定理

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 $E(X_k) = \mu \quad (k = 1, 2, \cdots, n)$, 则对任意的 $\varepsilon > 0$, 都有

$$\lim_{n \to \infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right\} = 0$$

注1° 与切比谢夫大数定理相比, 不要求方差存在且有界.

2° 贝努里大数定理是辛钦大数定理的特例.

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

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

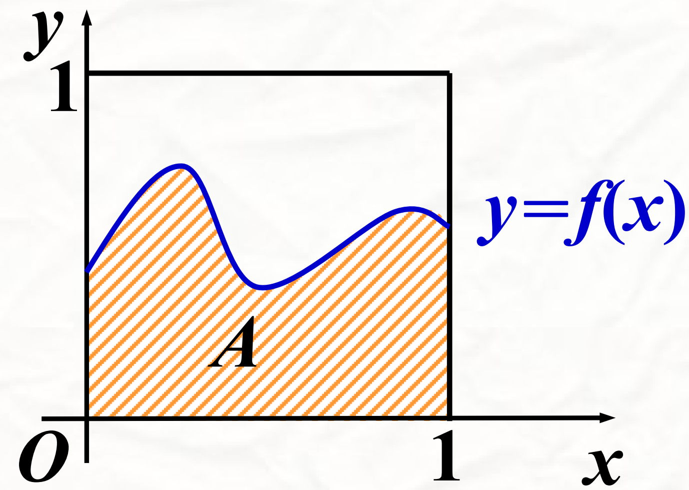
解 因为$X_1, X_2, \cdots, X_n, \cdots$是相互独立的，所以 $X_1^2, X_2^2, \cdots, X_n^2, \cdots$也是相互独立的. 由$E(X_k)=0$,

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

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

例4 设 $f(x)$ $(a\le x\le b)$ 是连续函数, 试用概率论方法近似计算积分 $\int_a^b f(x)\mathrm{d}x$.



解 设 $|f(x)|$ 的一个上界为 M $(M>0)$, $f(x)$ 的最小值为 h, 则 $0\le \dfrac{f(x)-h}{2M}\le 1$,

故不妨假定 $0\le f(x)\le 1$, 引进新变量 z: $x=(b-a)z+a$ 后, 可将 x 轴上的区间 $[a, b]$ 变为 z 轴上 $[0, 1]$, 故不妨设 $a=0$, $b=1$.

目 上 下 回 停 续

考虑几何型随机试验E:向矩形 $0\le x\le 1,\ 0\le y\le 1$ 中均匀分布地掷点, 将E独立地重复做下去, 以 A表示此矩形中曲线 $y=f(x)$下的区域, 即

$$A=\{(x,y):0\le y\le f(x);\ x\in[0,1]\}$$

并定义随机变量序列

$$X_k=\begin{cases}1, & \text{第}k\text{次掷的点落于}A\text{中}\\ 0, & \text{反之}\end{cases}$$

则$\{X_k: k\ge 1\}$独立同分布, 而且

$$E(X_k)=P(X_k=1)=|A|=\int_0^1 f(x)\mathrm{d}x$$

$|A|$表示A的面积, 由**贝努里大数定理**知

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} E(X_k) = \int_0^1 f(x)\,\mathrm{d}x$$

这表示当n充分大时, 前n次试验中落于A中的点数 $\sum_{k=1}^{n} X_k$ 除以n后, 以任意接近于1的概率与 $\int_0^1 f(x)\,\mathrm{d}x$ 近似.

这种近似计算法叫**蒙特卡洛**(Monte-Carlo)方法.

内容小结

四个大数定理

- 切比谢夫大数定理
- 贝努里大数定理
- 泊松大数定理
- 辛钦大数定理

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


再见

备用题

例3-1 设$\{X_n\}$为独立同分布的随机变量序列，其共同分布为

$$p(X_n = \frac{2^k}{k^2}) = \frac{1}{2^k}, k = 1,2\cdots.$$

试问$\{X_n\}$是否服从大数定律?

解

$$E(X_n) = \sum_{k=1}^{+\infty} \frac{2^k}{k^2} \cdot \frac{1}{2^k} = \sum_{k=1}^{+\infty} \frac{1}{k^2} = \frac{\pi^2}{6} < +\infty$$

即$E(X_n)$存在，由**辛钦大数定律**知服从大数定律.

贝努里(Jacob Bernoulli)

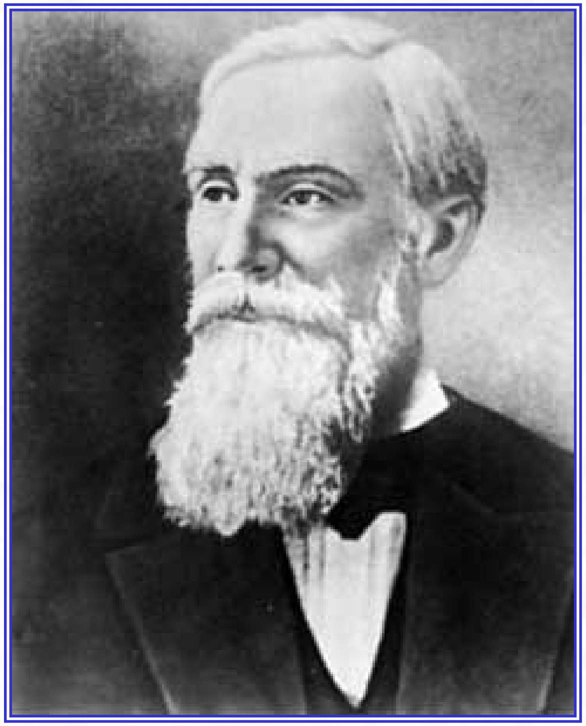
1654-1705

瑞士人, 贝努里家族的三大杰出的数学家之一.

首先发展无穷小分析, 1960年提出悬连线问题,首创积分 “integral”这一术语.

提出贝努里大数定理, 建立了贝努里概型. 在无穷级数理论、变分法和概率论等反面都有贡献.

切比谢夫(Pafnuty Chebyshev)



1821-1894

俄国数学家、机械学家. 对数论、积分理论、概率论和力学都有很大贡献.

证明了贝尔特兰公式, 关于自然数列中素数分布的定理, 大数定律的一般公式以及中心极限定理. 创立了切比谢夫多项式.

辛钦(Aleksandr Yakovlevich Khinchin)

1894-1959

苏联数学家, 现代概率论的奠基人之一.

他最早的概率论成果是贝努里试验序列的重对数律.

辛钦在函数的度量理论、数论、概率论、数学分析、信息论等方面都有重要的研究成果.