概率论与数理统计

第五章　大数定律及中心极限定理

练习：设有一大批种子，其中良种占1/6. 试估计在任选的 6000 粒种子中, 良种所占比例与1/6 比较上下小于1%的概率.

解 设 X 表示 6000 粒种子中的良种数，

$$X \sim B(6000, 1/6)$$

$$E(X) = 1000,\ D(X) = \frac{5000}{6}$$

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

$$= P(|X - 1000| < 60) \geq 1 - \frac{\frac{5000}{6}}{60^2} = \frac{83}{108} = 0.7685$$

实际精确计算

$$P\left(\left|\frac{X}{6000}-\frac{1}{6}\right|<0.01\right)=P(940<X<1060)$$

$$=\sum_{k=941}^{1059} C_{6000}^{k}\left(\frac{1}{6}\right)^{k}\left(\frac{5}{6}\right)^{6000-k}=0.959036$$

用泊松分布近似计算

取$\lambda=1000$

$$P\left(\left|\frac{X}{6000}-\frac{1}{6}\right|<0.01\right)=P(940<X<1060)$$

$$=\sum_{k=941}^{1059}\frac{1000^{k}e^{-1000}}{k!}=0.937934$$

第五章　大数定律及中心极限定理

本章要解决的问题　　答复

1. 为何能以某事件发生的频率作为该事件的 概率的估计？
2. 为何能以样本均值作为总体期望的估计？

→ **大数定律**

3. 为何正态分布在概率论中占有极其重要的地位？
4. 大样本统计推断的理论基础是什么？

→ **中心极限定理**

§1　大数定律

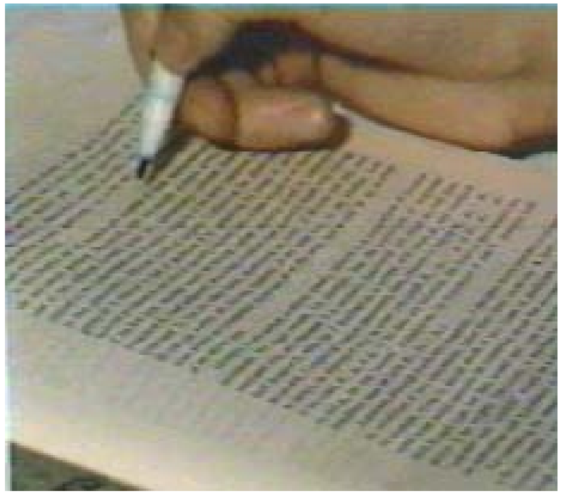
- ◆大数定律
- ◆以概率收敛定义及性质
- ◆小结

大数定律的客观背景

大量随机试验中 $\begin{cases}\text{事件发生的频率稳定于某一常数}\\\text{测量值的算术平均值具有稳定性}\end{cases}$

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

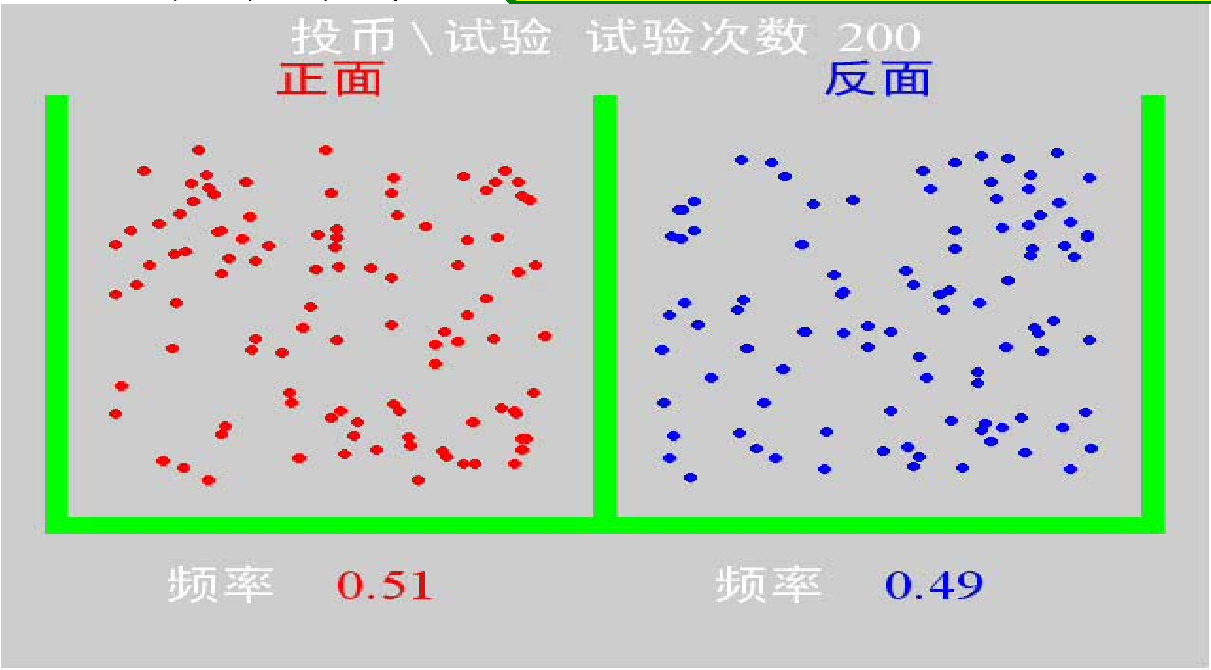
……

问题的引入

实例　频率的稳定性

随着试验次数的增加, 事件发生的频率逐渐稳定于某个常数.

单击图形播放/暂停　ESC键退出



启示:从实践中人们发现大量测量值的算术平均值有稳定性.

一、大数定律

定理1（切比雪夫定理的特殊情况）

设随机变量$X_1, X_2, \cdots, X_n, \cdots$相互独立，且具有相同的数学期望和方差：

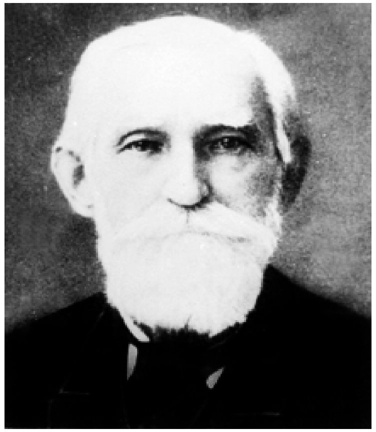
$$E(X_k)=\mu, D(X_k)=\sigma^2 (k=1,2,\cdots).$$

做前 n 个随机变量的算术平均 $\bar{X}=\frac{1}{n}\sum_{k=1}^{n}X_k$

则对任意的$\varepsilon>0$，有

$$\lim_{n\to\infty}P\{|\bar{X}-\mu|<\varepsilon\}$$

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


切比雪夫，П. Л.

切比雪夫

证　由于

$$E\left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)=\frac{1}{n}\sum_{k=1}^{n}E(X_k)=\frac{1}{n}\cdot n\mu=\mu$$

$$D\left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)=\frac{1}{n^2}\sum_{k=1}^{n}D(X_k)=\frac{1}{n^2}\cdot n\sigma^2=\frac{\sigma^2}{n}$$

由切比雪夫不等式

$$P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu\right|<\varepsilon\right\}\geq 1-\frac{\sigma^2/n}{\varepsilon^2}$$

上式中令 $n\to\infty$ 得

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

说明

1、定理中$\{|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu|<\varepsilon\}$是指一个随机事件，当$n\to\infty$时，这个事件的概率趋于1.

2、定理以数学形式证明了随机变量$X_1,\cdots X_n$的算术平均$\overline{\mathrm{X}}=\frac{1}{n}\sum_{i=1}^{n}X_i$接近数学期望$\mathrm{E}(\mathrm{X}_k)=\mu$（$k=1,2\cdots,n$），这种接近说明其具有的稳定性.

这种稳定性的含义说明算术平均值是依概率收敛的意义下逼近某一常数.

二、依概率收敛定义及性质

定义　设$Y_1,Y_2,\cdots Y_n,\cdots$是一个随机变量序列，$a$是一个常数.若对于任意正数$\varepsilon$，有

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

则称序列$Y_1,Y_2,\cdots Y_n,\cdots$依概率收敛于$a$.记为

$$Y_n \xrightarrow{P} a.$$

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

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

证明略

请注意：

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

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.

定理1的另外一种叙述形式

设随机变量X_1，$X_2,\cdots,X_n,\cdots$相互独立，且具有相同的数学期望和方差：$E(X_k)=\mu, D(X_k)=\sigma^2$ $(k=1,2,\cdots)$，则序列$\overline{X}=\frac{1}{n}\sum_{k=1}^{n}X_k$依概率收敛于$\mu$，即

$$\overline{X}\xrightarrow{P}\mu.$$

定理的意义

具有相同数学期望和方差的独立随机变量序列的算术平均值依概率收敛于数学期望.

当 n 足够大时, 算术平均值几乎是一常数.

数学期望 可被 算术均值 近似代替

问题：

设n_A是n重伯努利(伯努利)试验中事件A发生的次数，p是事件A发生的概率，

雅各布第一·伯努利

伯努利

$\dfrac{n_A}{n}$ 是事件A发生的频率.

事件发生的频率能否代替事件的概率，频率是否具有稳定性呢？

定理2（伯努利大数定律）

设 n_A 是n次独立重复试验中事件A发生的次数，p是事件A在每次试验中发生的概率，则对于任意正数$\varepsilon>0$，有

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

伯努利

或
$$\lim_{n\to\infty} P\{|\frac{n_A}{n}-p|\geq\varepsilon\}=0$$

证明　　因为$n_A\sim b(n,p)$,由此可表示为

$$n_A=X_1+X_2+\cdots+X_n$$

其中,X_1，$X_2,\cdots,X_n$,相互独立，且都服从以p以为参数的（$0-1$）分布.因而$E(X_k)=p$，$D(X_k)=p(1-p)$,

由定理1即得

$$\lim_{n\to\infty} P\{|\frac{n_A}{n} - p| < \varepsilon\}$$

$$= \lim_{n\to\infty} P\{|\frac{1}{n}(X_1 + X_2 + \cdots + X_n) - p| < \varepsilon\} = 1$$

或

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

证毕

说明

伯努利定理表明事件发生的频率 n_A/n 依概率收敛于事件的概率 p，它以严格的数学形式表达了频率的稳定性.

因而当 $\boldsymbol{n}$ 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.

下面给出的独立同分布下的大数定律，不要求随机变量的方差存在.

辛钦

定理3（辛钦大数定律）

设随机变量序列$X_1, X_2, \ldots X_n, \ldots$ 相互独立，服从同一分布，具有数学期 $E(X_i)=\mu$, i=1,2,…， 则对于任意正数ε，有

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

注 1、辛钦大数定律为寻找随机变量的期望值提供了一条实际可行的途径.

2、伯努利大数定律是辛钦定理的特殊情况.

3、辛钦定理具有广泛的适用性.

要估计某地区的平均亩产量，要收割某些有代表性块，例如n 块地. 计算其平均亩产量，则当n 较大时，可用它作为整个地区平均亩产量的一个估计.

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

问是否满足切比雪夫定理?

解　独立性依题意可知, 检验是否具有数学期望?

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

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

三、小结

大数定律

伯努利大数定律 $\lim\limits_{n\to\infty} P\{|\frac{n_A}{n} - p| < \varepsilon\} = 1$ $\quad n_A \sim b(n, p)$

切比雪夫大数定律 $\lim\limits_{n\to\infty} P\{|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu| < \varepsilon\} = 1$ $\quad \begin{cases} E(X_k) = \mu \\ D(X_k) = \sigma^2 \end{cases}$

辛钦大数定律 $\lim\limits_{n\to\infty} P\{|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu| < \varepsilon\} = 1$ $\quad E(X_k) = \mu$

大数定律以严格的数学形式表达了随机现象最根本的性质之一：

平均结果的稳定性

练习：

设供电网有10000盏电灯，夜晚每盏电灯开灯的概率均为0.7，并且彼此开闭与否相互独立，试用切比雪夫不等式估算夜晚同时开灯数在6800到7200之间的概率.

解：设X表示夜晚同时开灯的盏数

则$X \sim B(10000, 0.7)$ $E(X)=7000$, $D(X)=2100$,

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

故 $\boldsymbol{P}\left\{6800<\boldsymbol{X}<7200\right\}=P\{|X-7000|<200\}$

$$\geq 1-\frac{2100}{200^2}=0.9475$$

§2　中心极限定理

- 中心极限定理

中心极限定理的客观背景

在实际问题中许多随机变量是由相互独立随机因素的综合影响所形成的.

例如：炮弹射击的落点与目标的偏差，就受着许多随机因素（如瞄准，空气阻力，炮弹或炮身结构等）综合影响的. 每个随机因素的对弹着点（随机变量和）所起的作用都是很小的. 那么弹着点服从怎样分布哪 ？

自从高斯指出测量误差服从正态分布之后，人们发现，正态分布在自然界中极为常见.

高斯

如果一个随机变量是由大量相互独立的随机因素的综合影响所造成，而每一个别因素对这种综合影响中所起的作用不大. 则这种随机变量一般都服从或近似服从正态分布.

现在我们就来研究独立随机变量之和所特有的规律性问题.

当n无限增大时，这个和的极限分布是什么呢？

由于无穷个随机变量之和可能趋于∞，故我们不研究n个随机变量之和本身而考虑它的标准化的随机变量. 即考虑随机变量$X_k(k=1,\cdots n)$的和$\sum_{k=1}^{n} X_k$的标准化变量

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

讨论Y_n的极限分布是否为标准正态分布

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.

一、中心极限定理

定理4（独立同分布下的中心极限定理）

（林德伯格–列维中心极限定理- **Lindberg-levi**）

设随机变量$X_1, X_2, \cdots X_n, \cdots$相互独立，服从同一分布，且具有数学期望和方差：$E(X_k)=\mu, D(X_k)=\sigma^2$ $(k=1,2,\cdots)$，则随机变量之和$\sum_{k=1}^{n} X_k$的标准化变量

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

的分布函数$F_n(x)$对于任意x满足

$$\lim_{n\to\infty} F_n(x)=\lim_{n\to\infty} P\left\{\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \le x\right\}$$

$$=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2}\mathrm{dt} \ =\Phi(x)$$

注 1、定理表明，独立同分布的随机变量之和$\sum_{k=1}^{n} X_k$，当n充分大时，随机变量之和与其标准化变量分别有

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

2、独立同分布中心极限定理的另一种形式可写为

$$\overline{X} \overset{\text{近似地}}{\sim} N(\mu, \sigma^2/n) \quad 或 \quad \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \overset{\text{近似地}}{\sim} N(0,1)$$

其中$\overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$

3、虽然在一般情况下，我们很难求出$\sum_{k=1}^{n} X_k$的分布的确切形式，但当n很大时，可以求出近似分布.

作业：课后习题 2、3

练习：

1、 根据以往经验，某种电器元件的寿命服从均值为100小时的指数分布. 现随机地取16只，设它们的寿命是相互独立的. 求这16只元件的寿命的总和大于1920小时的概率.

解： 设第i只元件的寿命为X_i, i=1,2, …,16

由题给条件知，诸X_i独立,且$E(X_i)$=100， $D(X_i)$=10000

16只元件的寿命的总和为

$$Y = \sum_{k=1}^{16} X_k$$

依题意，所求为$P(Y>1920)$

$E(Y)$=1600， $D(Y)$=160000

由中心极限定理, $\dfrac{Y-1600}{400}$ 近似$N(0,1)$

$$P(Y>1920)=1-P(Y\le 1920)\approx 1-\Phi\left(\frac{1920-1600}{400}\right)$$

$$=1-\Phi(0.8)=1-0.7881=0.2119$$

2、 设$X_1, X_2, \cdots, X_n$相互独立,同服从参数为1/2的指数分布,则当$n \to \infty$时, $Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ 依概率收敛于____.

分析：由辛钦大数定理知,独立同分布且期望存在的随机变量序列的平均值依概率收敛于它的期望. 因此只须求Y_n的期望.

解 因$X_1, X_2, \cdots, X_n$独立同分布,故$X_1^2, X_2^2, \cdots, X_n^2$独立同分布.

$$E(X_i) = \frac{1}{2},\ \ D(X_i) = \frac{1}{2^2} = \frac{1}{4}.\ E(X_i^2) = D(X_i) + \left[E(X_i)\right]^2 = \frac{1}{2},$$

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