§6 大数定律和中心极限定理

一、问题的提出

事件发生的频率为什么能作为事件概率的估计?

样本均值为什么可以作为总体期望的估计?

在概率统计中，正态分布为什么极其重要?

统计推断的理论基础是什么?

定义： 设$\{\xi_n\}$（即$\xi_1,\xi_2,\cdots,\xi_n,\cdots$ ）为随机变量序列，

若对于$\forall\,\varepsilon>0$， 成立 $\lim\limits_{n\to\infty}P(|\xi_n|<\varepsilon)=1$

则称随机变量序列$\{\xi_n\}$依概率收敛于 0，

记为 $\xi_n\to 0$.

设 ξ 是随机变量，若 $\xi_n-\xi\xrightarrow{P}0$ ，

则称 $\{\xi_n\}$ 依概率收敛于 ξ，记为 $\xi_n\xrightarrow{P}\xi$.

定理（依概率收敛性质）：设$\xi_n\xrightarrow{P}A$，$\eta_n\xrightarrow{P}B$，

且二元函数 f 在点 (A,B)处连续，

则 $f(\xi_n,\eta_n)\xrightarrow{P}f(A,B)$.

二、大数定律

1、切比雪夫（*Chebyshev*）不等式：

若随机变量ξ具有数学期望 $E\xi$ 和方差 $D\xi$，若对于 $\forall\,\varepsilon>0$，成立

$$P(|\xi-E\xi|\geq\varepsilon)\leq\frac{D\xi}{\varepsilon^2} \text{ 和 } P(|\xi-E\xi|<\varepsilon)\geq 1-\frac{D\xi}{\varepsilon^2}.$$

切比雪夫不等式的意义：

反映的是随机变量 X 的取值落在其数学期望的 ε 邻域内的概率不小于 $1-\sigma^2/\varepsilon^2$. 它的意义在于当随机变量的数学期望和方差已知时，可以估计随机变量 X 落在以数学期望为中心的某一区间内的概率的一个下限。

证明：仅对离散型随机变量证明，

设 ξ 是一个离散型随机变量，其分布律为

$$P(\xi = x_k) = p_k \ (k = 1, 2, \cdots)$$

则 $$P(|\xi - E\xi| \geq \varepsilon) = \sum_{|x_k - E\xi| \geq \varepsilon} P(\xi = x_k) = \sum_{|x_k - E\xi| \geq \varepsilon} p_k$$

$$\leq \sum_{|x_k - E\xi| \geq \varepsilon} \frac{(x_k - E\xi)^2}{\varepsilon^2} p_k$$

$$\leq \sum_{k \geq 1} \frac{(x_k - E\xi)^2}{\varepsilon^2} p_k$$

$$= \frac{1}{\varepsilon^2} \sum_{k \geq 1} (x_k - E\xi)^2 p_k = \frac{1}{\varepsilon^2} D\xi$$

且 $$P(|\xi - E\xi| < \varepsilon) = 1 - P(|\xi - E\xi| \geq \varepsilon) \geq 1 - \frac{D\xi}{\varepsilon^2}.$$

例1、在每次试验中，事件 A 发生的概率为0.75，利用切比雪夫不等式，求 n 需要多大时才能使得在 n 次重复独立试验中事件 A 出现的频率在 $0.74\sim0.76$ 之间的概率至少为0.90 .

2、切比雪夫（Chebyshev）大数定律：

设 $\xi_1,\xi_2,\cdots,\xi_n,\cdots$ 是相互独立的随机变量序列（即对每个正整数 $n\geq 2$，$\xi_1,\xi_2,\cdots,\xi_n$ 是相互独立的）相应的数学期望依次为 $E\xi_1,E\xi_2,\cdots,E\xi_n,\cdots$，方差依次为 $D\xi_1,D\xi_2,\cdots,D\xi_n,\cdots$，若 $D\xi_i\leq L\,(i=1,2,\cdots)$，这里 L 是与 i 无关的常数，则对于 $\forall\,\varepsilon>0$，成立

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

证明：$\because \xi_1,\xi_2,\cdots,\xi_n$ 相互独立，所以

$$D\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D\xi_i\leq\frac{1}{n^2}nL=\frac{L}{n}$$

$$\because E\left(\sum_{i=1}^{n}\xi_i\right)=\frac{1}{n}\sum_{i=1}^{n}E\xi_i$$

由切比雪夫不等式得，对于 $\forall\varepsilon>0$，成立

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

令 $n\to\infty$，由极限的夹逼性得

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

推论：设 $\{\xi_n\}$ 是相互独立、具有相同分布的随机变量序列，且 $E\xi_i=\mu$，$D\xi_i=\sigma^2\ (i=1,2,\cdots)$，则对于 $\forall\varepsilon>0$，成立

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

3、辛钦大数定律：

设$\{\xi_n\}$是相互独立、具有相同分布的随机变量序列，且 $E\xi_i = \mu \ (i = 1, 2, \cdots)$，

则对于 $\forall\, \varepsilon > 0$，成立 $\lim\limits_{n\to\infty} P\left(\left|\dfrac{1}{n}\sum\limits_{i=1}^{n}\xi_i - \mu\right| < \varepsilon\right) = 1$.

4、Bernoulli大数定律：

设 ξ 为 n 重 Bernoulli 试验中事件 A 发生的次数，则当 n 无限增大时，事件 A 发生的频率 $\dfrac{\xi}{n}$ 依依概率收敛于 A 发生的概率 $p = P(A)$，

即对于 $\forall\, \varepsilon > 0$，成立 $\lim\limits_{n\to\infty} P\left(\left|\dfrac{\xi}{n} - p\right| < \varepsilon\right) = 1$.

大数定律的意义：

大数定律深刻地揭示了随机事件概率与频率间的关系，大数定律从大量独立重复试验中测量值的平均值出发，以严格的数学形式证明了平均值和频率的稳定性，同时表达了这种稳定性的含义。

具有相同数学期望和方差的独立随机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。

算术均值 可近似代替 **数学期望**

例2、设随机变量 $X_1, X_2, \cdots, X_n, \cdots$相互独立，服从相同的分布，$E(X_i)=0$， $D(X_i)=\sigma_2$，且 $E(X_i^4)\ (i=1,2,\cdots)$ 存在，

试证明：对$\forall\, \varepsilon>0$，有

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i^2-\sigma^2\right|<\varepsilon\right)=1.$$

三、中心极限定律

1、独立同分布的中心极限定理：

设$\{\xi_n\}$是相互独立、服从同分布的随机变量序列，且数学期望 $E\xi_i=\mu$，方差 $D\xi_i=\sigma^2$ $(i=1,2,\cdots)$，则

$$\lim_{n\to\infty}P\left(\frac{1}{\sqrt{n}\sigma}\sum_{i=1}^{n}(\xi_i-\mu)\le x\right)=\Phi_0(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt\,.$$

这一定理可以得到如下结论：

随机变量序列$\{X_n\}$满足独立同分布、方差存在的条件，不论 X_n服从什么分布，只要 n 充分大，就有 $Y_n=\sum\limits_{i=1}^{n}X_i \overset{\text{近似}}{\sim} N(n\mu,n\sigma^2)$

$$\frac{1}{\sqrt{n}\sigma}\sum_{i=1}^{n}(X_i-\mu)\overset{\text{近似}}{\sim}N(0,1)\,,\quad \bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i\overset{\text{近似}}{\sim}N(\mu,\frac{\sigma^2}{n})$$

例3、计算机在进行加法时，对每个加数取整（取为最接近于它的整数），设所有的取整误差是相互独立的，且它们都在 (-0.5, 0.5) 上服从均匀分布。

1）若取1500个数相加，问误差总和的绝对值超过15的概率是多少？

2）可将几个数加在一起使得误差总和的绝对值小于10的概率为0.90？

2、棣莫弗-Laplace积分极限定理：

设$\xi \sim B(n, p)$，a, b $(a<b)$ 为常数，则当n充分大时，

成立 $P(a<\xi \le b) \approx \Phi_0\left(\dfrac{b-np}{\sqrt{npq}}\right) - \Phi_0\left(\dfrac{a-np}{\sqrt{npq}}\right)$.

其中 $q = 1- p$.

3、棣莫弗-Laplace局部极限定理：

设$\xi \sim B(n, p)$，则当n充分大时，成立

$$P(\xi = k) = p_k \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}} = \frac{1}{\sqrt{npq}} \Phi_0\left(\frac{k-np}{\sqrt{npq}}\right).$$

其中 $q = 1- p$.

中心极限定理的意义：

中心极限定理说明，即使有些原本并不服从正态分布的一些独立的随机变量，它们的总和的分布渐近地服从正态分布。一般来说，这些随机变量受到大量独立的因素中每项因素的影响是均匀的，微小的，没有一项因素起特别突出的影响。那么这些随机变量的和的分布近似于正态分布。

例4、从一大批发芽率为0.9 的种子中随意抽取1000粒，试估计这1000粒种子发芽率不低于0.88的概率。

例5、假定100 00个同龄人参加人寿保险，在一年里这些人死亡率为0.1%，参加保险的人在一年的第一天交付保险费80元，死亡时，家属可从保险公司领取20 000元的抚恤金。求

1）保险公司一年中获利不少于60 000元的概率；

2）保险公司亏本的概率是多少？

大数定律与中心极限定理有什么异同和联系：

大数定律与中心极限定理的相同点

大数定律与中心极限定理的不同点

大数定律与中心极限定理的联系