概率论与数理统计

第五章·大数定律及中心极限定理

2021 年 2 月 24 日

■ 统计与数学学院　■ 王官杰

主要内容

极限定理研究大量随机变量的规律性。

主要内容

极限定理研究大量随机变量的规律性。

1. 大数定律(Law of Large Numbers)：
   大量随机变量的平均结果的稳定性。

主要内容

极限定理研究大量随机变量的规律性。

1. 大数定律(Law of Large Numbers)：
   大量随机变量的平均结果的稳定性。
2. 中心极限定理(Central Limit Theorem)：
   大量随机变量的和的稳定性。

第一节 有理分式的积分

定义 设$Y_1, \ldots, Y_n$若对于任给的$\varepsilon > 0$，有

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

则称随机变量序列$\{Y_n\}$ 依概率收敛于Y，记作

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

切比雪夫不等式

定理 设随机变量X有期望和方差，则对于任给的$\varepsilon > 0$，有

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

即

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

定理 1 (切比雪夫大数定律) 设随机变量$X_1, \ldots, X_n, \cdots$相互独立，且有相同的期望μ和方差σ^2，定义Y_n为前n个随机变量的算术平均，即

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

则对任意正数ε，有 $\lim\limits_{n\to\infty} P\{|Y_n - \mu| < \varepsilon\} = 1.$

注记 切比雪夫定理的实际意义：多次试验求平均值能够有效地控制误差。

定理 2 (伯努利大数定律)　在独立重复试验中，记事件A的概率为p。以$f_n(A)$表示前n次试验中事件A发生的次数，则对任意正数ε，

$$\lim_{n\to\infty} P\left\{\left|\frac{f_n(A)}{n} - p\right| < \varepsilon\right\} = 1,$$

$$\lim_{n\to\infty} P\left\{\left|\frac{f_n(A)}{n} - p\right| \neq \varepsilon\right\} = 0.$$

伯努利大数定律

定理 2 (伯努利大数定律)　在独立重复试验中，记事件A的概率为p。以$f_n(A)$表示前n次试验中事件A发生的次数，则对任意正数ε,

$$\lim_{n\to\infty} P\left\{\left|\frac{f_n(A)}{n}-p\right|<\varepsilon\right\}=1,$$

$$\lim_{n\to\infty} P\left\{\left|\frac{f_n(A)}{n}-p\right|\neq\varepsilon\right\}=0.$$

注记　伯努利定理的实际意义：当重复试验次数充分大时，某事件发生的频率与该事件发生的概率有一定偏差的可能性很小。

定理 3 (辛钦大数定律) 设随机变量$X_1, \ldots, X_n, \cdots$相互独立，且服从同一分布，具有数学期望$E(X_k) = \mu \ (k = 1, 2, \cdots)$, 则对任意正数$\varepsilon$，有

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

注记 辛钦大数定律不要求方差存在，但要求随机变量服从同一分布。

第一节 有理分式的积分

中心极限定理的主要思想：如果

1. 一个随机现象由众多的随机因素所引起，
2. 且每一因素在总的变化里所起的作用不显著，

则描述这个随机现象的随机变量近似地服从正态分布。

这就是为什么实际中遇到的随机变量很多都服从正态分布的原因，也正因如此，正态分布在概率论和数理统计中占有极其重要的地位。

林德伯格—莱维定理

定理 4 (林德伯格—莱维定理) 若$\{X_n\}_{n\geq 1}$独立且同分布，$E(X_i)=\mu$，$\mathrm{Var}(X_i)=\sigma^2>0$，令

$$Z_n=\frac{\sum\limits_{i=1}^{n} X_i-n\mu}{\sqrt{n}\sigma}$$

则其分布函数$F_n(x)$收敛到$\Phi(x)$，即对任何实数x，

$$\lim_{n\to\infty} F_n(x)=\lim_{n\to\infty} P\{Z_n\leq x\}=\Phi(x)$$

其中$\Phi(x)$是标准正态分布$N(0,1)$的分布函数，

棣莫弗一拉普拉斯定理

设在某试验中事件A发生的概率为p，将该试验独立地进行n次。记η_n为n次试验中事件A发生的总次数，X_i为第i 次试验中事件A发生的次数，则

$$\eta_n \sim B(n, p),$$

$$X_i \sim B(1, p), i = 1, 2, \cdots, n,$$

且

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

棣莫弗—拉普拉斯定理

定理 5 (棣莫弗—拉普拉斯定理) 设随机变量η_n服从参数为n, p的二项分布，即$\eta_n \sim B(n, p)$，则对任意x有

$$\lim_{n\to\infty} P\left\{\frac{\eta_n - np}{\sqrt{np(1-p)}} \leq x\right\} = \Phi(x), \qquad \forall x \in \mathbb{R}.$$

注记 棣莫弗—拉普拉斯定理表明，当n充分大时，η_n近似服从正态分布，即可以近似认为

$$\eta_n \sim N(np, np(1-p)).$$

或者

$$\frac{\eta_n - np}{\sqrt{np(1-p)}} \sim N(0, 1).$$