第四章　大数律和中心极限定理

4.1　大数律

4.1.1　随机变量的收敛性

定义 4.1.1. 如果对任何$\varepsilon > 0$, 都有

$$\lim_{n\to\infty} P(|\xi_n - \xi| \geq \varepsilon) = 0,$$

那么我们就称随机变量序列$\{\xi_n,\ n \in \mathbb{N}\}$依概率收敛到随机变量$\xi$, 记为$\xi_n \xrightarrow{p} \xi$.

引理 4.1.1 (Chebyshev Inequality)**.** 设$r.v.\ \xi$的r阶矩存在，则

$$P(|\xi| > x) \leq \frac{E|\xi|^r}{x^r}, \quad \forall\, x > 0$$

例4.1.1. 如果以ζ_n表示n重Bernoulli试验中的成功次数, 则有

$$\frac{\zeta_n}{n} \xrightarrow{p} p.$$

4.1.2　大数律

定义 4.1.2. 设$\{\xi_n\}$是一列随机变量，如果存在常数序列$\{a_n\}$使得

$$\frac{1}{n}\sum_{k=1}^{n} \xi_k - a_n \xrightarrow{p} 0$$

则称$\{\xi_n\}$服从大数定律。

定理 4.1.1. 设随机变量序列$\{X_n\}$相互独立且具有相同的分布*(独立同分布,记为$i.i.d$)*，具有数学期望$EX_k = \mu, k = 1, 2, \cdots$. 则有

$$\overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{p} \mu$$

即$\{X_n\}$服从大数定律。

4.2 中心极限定理

我们仅叙述独立同分布场合下的中心极限定理，更一般的情形参考其它专业的概率论教材。

定理 4.2.2. 设$\{X_n\}$为$i.i.d$的随机变量序列，具有数学期望$EX_k = \mu$ 和方差$\sigma^2 = D(X_k)$,$k = 1, 2, \cdots$. 则有

$$\lim_{n\to\infty} P\left(\frac{1}{\sqrt{n}\sigma}(X_1 + \cdots + X_n - n\mu) \le x\right) = \lim_{n\to\infty} P\left(\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma}\right) = \Phi(x) \qquad \forall x \in \mathbb{R}$$

其中$\overline{X}_n = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$.

此定理我们称为**独立同分布场合下的中心极限定理**.

Proof. 由于标准正态分布的特征函数为$f(t) = e^{-t^2/2}$，因此我们只需证明$\eta_n = \sum\limits_{i=1}^{n} \frac{X_i - \mu}{\sigma}$的特征函数的极限是$f(t)$就可以了。

记$\{X_i - \mu\}$的共同特征函数为$g(t)$，则

$$g\Big(\frac{t}{\sigma\sqrt{n}}\Big) = 1 - \frac{t^2}{2n} + o\Big(\frac{t^2}{n}\Big)$$

而η_n的特征函数为$g^n(\frac{t}{\sigma\sqrt{n}})$. 由于

$$\left|g^n\Big(\frac{t}{\sigma\sqrt{n}}\Big) - \Big(1 - \frac{t^2}{2n}\Big)^n\right| \le n\left|g\Big(\frac{t}{\sigma\sqrt{n}}\Big) - \Big(1 - \frac{t^2}{2n}\Big)\right| = no\Big(\frac{t^2}{n}\Big) \longrightarrow 0$$

所以

$$\lim_{n\to\infty} g^n\Big(\frac{t}{\sigma\sqrt{n}}\Big) = e^{-t^2/2}$$

即

$$\lim_{n\to\infty} P(\eta_n \le x) = \Phi(x)$$

□

例4.2.1. 求极限$\lim\limits_{n\to\infty}\sum\limits_{k=1}^{n}\frac{n^k}{k!}e^{-n}$.

定理 4.2.3. 设$X \sim B(n, p)$，则有

$$\lim_{n\to\infty} P(\frac{X - np}{\sqrt{npq}} \le x) = \Phi(x), \quad \forall\, x \in \mathbb{R}$$

即

$$\frac{X - np}{\sqrt{npq}} \overset{asy.}{\sim} N(0, 1).$$

Proof. 由二项分布随机变量和0-1分布随机变量之间的关系及中心极限定理易证。 □

在仅有独立性和二阶矩有限场合下，我们有

定理 4.2.4. 设$\{X_n\}$为独立的随机变量序列，而且具有数学期望$EX_k = \mu_k$ 和方差$D(X_k) = \sigma_k^2 < \infty$,$k = 1, 2, \cdots$. 记

$$B_n^2 = \sum_{k=1}^{n} \sigma_k^2$$

若存在正数δ，使得当$n \to \infty$时

$$\frac{1}{B_n^{2+\delta}} \sum_{k=1}^{n} E|X_k - EX_k|^{2+\delta} \to 0$$

则有

$$\lim_{n\to\infty} P\left(\sum_{k=1}^{n} \frac{X_k - \mu_k}{B_n} \leq x\right) = \Phi(x) \qquad \forall x \in \mathbb{R} \tag{4.2.1}$$

例题参考课本.

定义 4.2.1. 如果独立随机变量序列$\{X_n,\ n \in \mathbb{N}\}$同上述定理, 并且对任何$\tau > 0$, 都有

$$\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{k=1}^{n} E\left\{(X_k - a_k)^2 I(|X_k - a_k| \geq \tau B_n)\right\} = 0, \tag{4.2.2}$$

则称该随机变量序列满足*Linderberg*条件.

定理 4.2.5. 设随机变量序列$\{X_n\}$满足*Linderberg*条件*(4.2.2)*, 则$\{X_n\}$满足中心极限定理, 即*(4.2.1)*式成立.

参考文献

[1] 苏淳., 概率论, 北京: 科学出版社, 2004.