

- 随机现象规律的刻画:
 - (1) 随机现象具有可重复、可观察属性；任何有限次都只是真相的近似。
 - (2) 寻求极限是揭示随机现象规律的基本方法。

- 随机现象规律的刻画:
 - (1) 随机现象具有可重复、可观察属性；任何有限次都只是真相的近似。
 - (2) 寻求极限是揭示随机现象规律的基本方法。
- 1930-40年代以前
 - (1) 几何分析代数方程等数学学科相当成熟
 - (2) 概率还处于初级阶段：计算简单事件的概率，理解频率和概率的关系，确定测量误差的分布等
 - (3) 建立极限定理

- A.Kolomgorov

- Bernoulli 大数律

- Bernoulli 大数律
- De Moivre-Laplace 中心极限定理

- Bernoulli 大数律
- De Moivre-Laplace 中心极限定理
- Poisson 极限定理

以上三个定理都只涉及Bernoulli试验（Bernoulli随机变量）

- Bernoulli 及其著作

(1) Jacobi Bernoulli (1654—1705), 法国著名数学家;
其家族在200年间先后出现8位数学家

(2) “The Art of Guessing” 一书1713 出版, 标志者概率论学科发展的正式开始

- 大数律

给定 $0 < p < 1$, 假设 $S_n \sim B(n, p)$, 那么

$$\frac{S_n}{n} \rightarrow p, \quad n \rightarrow \infty$$

- 大数律

给定 $0 < p < 1$, 假设 $S_n \sim B(n, p)$, 那么

$$\frac{S_n}{n} \rightarrow p, \quad n \rightarrow \infty$$

- “收敛”的含义

(1) 固定 $0 < \varepsilon < \min\{p, 1 - p\}$. 无论 n 有多大, 总会发生

$$\left| \frac{S_n}{n} - p \right| > \varepsilon$$

(2) 即

$$P(\omega : \left| \frac{S_n(\omega)}{n} - p \right| > \varepsilon) > 0$$

(3) 事实上,

$$P(\omega : \left| \frac{S_n(\omega)}{n} - p \right| > \varepsilon) = \sum_{k: |\frac{k}{n} - p| > \varepsilon} C_n^k p^k (1-p)^{n-k}$$

(4) Bernoulli的贡献

$$\sum_{k: |\frac{k}{n} - p| > \varepsilon} C_n^k p^k (1-p)^{n-k} \rightarrow 0, \quad n \rightarrow \infty$$

- Bernoulli大数律: 给定 $0 < p < 1$, 假设 $S_n \sim B(n, p)$, 那么

$$P(\omega : \left| \frac{S_n(\omega)}{n} - p \right| > \varepsilon) \rightarrow 0, \quad n \rightarrow \infty$$

- Bernoulli大数律: 给定 $0 < p < 1$, 假设 $S_n \sim B(n, p)$, 那么

$$P(\omega : \left| \frac{S_n(\omega)}{n} - p \right| > \varepsilon) \rightarrow 0, \quad n \rightarrow \infty$$

- Bernoulli大数律的意义

(1) 给出了“频率接近概率真值”的数学解释;

(2) 引入了“依概率收敛”的概念

(Ω, Σ, P) 是一个概率空间, $X, X_n, n \geq 1$ 是一列随机变量, 如果对任意 $\varepsilon > 0$,

$$P(\omega : |X_n(\omega) - X(\omega)| > \varepsilon) \rightarrow 0, \quad n \rightarrow \infty$$

称 X_n 依概率收敛到 X , 记做 $X_n \xrightarrow{P} X$.

按此概念, Bernoulli 大数律可写成

$$\frac{S_n}{n} \xrightarrow{P} p$$

(3) 问题: 如何推广和应用Bernoulli大数律?

- De Moivre 和Laplace

- (1) De Moivre

(2) Laplace

- 实际问题

(1) 通知200名学生参加某讲座，但每名学生出现讲座的概率仅为0.5，问应选择何种规模的报告厅？

(2) 计算

$$P(S_{200} \leq 80), \quad P(80 \leq S_{200} \leq 120), \quad P(S_{200} \geq 200)$$

(3)

$$P(80 \leq S_{200} \leq 120) = \sum_{80 \leq k \leq 120} C_{200}^k 0.5^{120}$$

- De Moivre-Laplace 中心极限定理

假设 $S_n \sim B(n, p)$, 那么

$$P\left(\frac{S_n - np}{\sqrt{np(1-p)}} \leq x\right) \asymp \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

(1) 左边: 规范化随机变量的分布函数

(2) 右边: 正态分布函数

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

(3)

$$P(a \leq S_n \leq b) = P\left(\frac{a - np}{\sqrt{np(1-p)}} \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq \frac{b - np}{\sqrt{np(1-p)}}\right)$$

$$P(a \leq S_n \leq b) \asymp \Phi\left(\frac{b - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - np}{\sqrt{np(1-p)}}\right)$$

查正态分布表

- 证明

(1) $p = \frac{1}{2}$ — De Moivre

(2) $p \neq \frac{1}{2}$ — Laplace

(3) 基本出发点:

$$P\left(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right) = \sum_{k: a \leq \frac{kn - np}{\sqrt{np(1-p)}} \leq b} C_n^k p^k (1-p)^{n-k}$$

(4) 将给出更现代的证明

- De Moivre 的证明思路:

$$P(S_n = \frac{n}{2} + k) = \frac{1}{2^n} \binom{n}{\frac{n}{2} + k}, \quad \text{假设 } \frac{n}{2} \text{ 是整数}$$

运用De Moivre 和他的朋友Stirling 1730年左右共同发展起来的公式:

$$n! = n^n e^{-n} \sqrt{2\pi n} (1 + o(1))$$

得到

$$\frac{1}{2^n} \binom{n}{\frac{n}{2} + k} \approx \frac{2}{\sqrt{2\pi n}} e^{-\frac{2k^2}{n}}, \quad \frac{1}{2^n} \binom{n}{\frac{n}{2}} \approx \frac{2}{\sqrt{2\pi n}}$$

$$\begin{aligned}P\left(a \leq S_n - \frac{n}{2} \leq b\right) &= \sum_{a \leq k \leq b} P\left(S_n = \frac{n}{2} + k\right) \\&= \sum_{a \leq k \leq b} \frac{2}{\sqrt{2\pi n}} e^{-\frac{2k^2}{n}} \\&= \frac{4}{\sqrt{2\pi}} \int_{a/\sqrt{n}}^{b/\sqrt{n}} e^{-2y^2} dy\end{aligned}$$

- 应用
例1.

- De Moivre - Laplace 中心极限定理的意义

(1) 给出近似计算公示

(2) 引入“依分布收敛”的概念

假设 (Ω, Σ, P) 是概率空间, $X, X_n, n \geq 1$ 是一列随机变量, $F, F_n, n \geq 1$ 是一列相应的分布函数, 如果对任意 x ,

$$F_n(x) \rightarrow F(x), \quad n \rightarrow \infty$$

称 F_n 依分布收敛 F , 记 $F_n \xrightarrow{d} F$ 或者 $X_n \xrightarrow{d} X$.

按上述概念，中心极限定理可写成

$$\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0, 1)$$

(3) 问题：如何推广和应用中心极限定理？

- Poisson 及其分布

- Poisson 极限定理

令 $0 < p_n < 1$, 假设 $S_n \sim B(n, p_n)$. 如果 $np_n \rightarrow \lambda$, 并且 $0 < \lambda < 1$ 那么对任何 $k = 0, 1, 2, \dots$

$$P(S_n = k) \rightarrow \frac{\lambda^k e^{-\lambda}}{k!}, \quad n \rightarrow \infty$$

- Poisson 极限定理

令 $0 < p_n < 1$, 假设 $S_n \sim B(n, p_n)$. 如果 $np_n \rightarrow \lambda$, 并且 $0 < \lambda < 1$ 那么对任何 $k = 0, 1, 2, \dots$

$$P(S_n = k) \rightarrow \frac{\lambda^k e^{-\lambda}}{k!}, \quad n \rightarrow \infty$$

- 证明

第二章已证, 略去

- 应用
例

实例. 保费的确定

Consider a group of n insurance policies. Let X_j be the amount of claim expense paid by the insurer to the j -th policy-holder during a particular future time period and let $S = X_1 + X_2 + \cdots + X_n$. Then S is the total payments made by the insurer on this group of policies during the period of insurance.

Suppose the X_j are independent and identically distributed and the insurer charges each policyholder the same premium at the beginning of the insured period. Let P be the premium charged per policy.

Clearly, the insurer would like the total premium collected to exceed (or at worst equal) the total payments made with high probability. That is the insurer would like $P(nP \geq S)$ to be close to 1.

Our goal is to analyze the behavior of P with respect to n under varying assumption about $P(nP \geq S)$

具体的问题:

Suppose that each $X_j \sim \exp(1/5)$ and the premium P is set in such a way that

$$P(nP \geq S) = 0.95$$

Determine P as a function of n .

What happens to the premium P as the size of the insured group becomes large?

$$ES = EX_1 + \cdots + EX_n = 5n$$

$$Var(S) = Var(X_1 + \cdots + X_n) = 25n$$

$$\begin{aligned} P(S \leq nP) &= P\left(\frac{S - 5n}{5\sqrt{n}} \leq \frac{nP - 5n}{5\sqrt{n}}\right) \\ &= ? \end{aligned}$$

● 假如

$$P\left(\frac{S - 5n}{5\sqrt{n}} \leq \frac{nP - 5n}{5\sqrt{n}}\right) \approx \Phi\left(\frac{nP - 5n}{5\sqrt{n}}\right)$$

Then the requirement $P(nP \geq S) = 0.95$ is approximately equivalent to the requirement

$$\Phi\left(\frac{nP - 5n}{5\sqrt{n}}\right) = 0.95$$

注意, $\Phi(1.64) = 0.95$, 那么

$$\frac{nP - 5n}{5\sqrt{n}} = 1.645$$

$$P = 5 + \frac{8.225}{\sqrt{n}}$$

$$P \rightarrow 5. \quad \text{Why?}$$

$$P = \lim_{n \rightarrow \infty} \frac{ES}{n}$$

因此，我们需要建立：
大数律和中心极限定理

- Chebyshev-Khinchine 大数律

(1) $S_n \sim B(n, p),$

$$S_n = \sum_{k=1}^n \xi_k, \quad \xi_k = 0, 1$$

(2)

$$S_n = \sum_{k=1}^n \xi_k,$$

ξ_k 一般随机变量

- Chebyshev-Khinchine 大数律
- Levy-Feller 中心极限定理

$$(1) S_n \sim B(n, p),$$

$$S_n = \sum_{k=1}^n \xi_k, \quad \xi_k = 0, 1$$

$$(2)$$

$$S_n = \sum_{k=1}^n \xi_k,$$

ξ_k 一般随机变量

- Chebyshev-Khinchine 大数律
- Levy-Feller 中心极限定理
- Lyapunov 中心极限定理

$$(1) S_n \sim B(n, p),$$

$$S_n = \sum_{k=1}^n \xi_k, \quad \xi_k = 0, 1$$

$$(2)$$

$$S_n = \sum_{k=1}^n \xi_k,$$

ξ_k 一般随机变量

- Chebyshev 及其不等式
- (1) Chebyshev

(2) Chebyshev 不等式

$$P(|X - EX| > \varepsilon) \leq \frac{\text{Var}(X)}{\varepsilon^2}$$

(3) 应用Chebyshev 不等式证明Bernoulli大数律:

$$P\left(\left|\frac{S_n}{n} - p\right| > \varepsilon\right) = P(|S_n - np| > n\varepsilon)$$

$$P(|S_n - np| > n\varepsilon) \leq \frac{\text{Var}(S_n - np)}{n^2\varepsilon^2}$$

$$= \frac{np(1-p)}{n^2\varepsilon^2} \rightarrow 0, \quad n \rightarrow \infty$$

- Chebyshev 大数律

假设 $\xi_k, k \geq 1$ 是一列随机变量, $E\xi_k = \mu$.

记 $S_n = \sum_{k=1}^n \xi_k$, 如果

$$\frac{\text{Var}(S_n)}{n^2} \rightarrow 0, \quad n \rightarrow \infty$$

那么

$$\frac{S_n}{n} \xrightarrow{P} \mu, \quad n \rightarrow \infty$$

更一般地

假设 $\xi_k, k \geq 1$ 是一列随机变量, $E\xi_k = \mu_k$.

记 $S_n = \sum_{k=1}^n \xi_k$, 如果

$$\frac{Var(S_n)}{n^2} \rightarrow 0, \quad n \rightarrow \infty$$

那么

$$\frac{S_n}{n} - \frac{\sum_{k=1}^n \mu_k}{n} \xrightarrow{P} 0, \quad n \rightarrow \infty$$

- Chebyshev 大数律的证明

对任意 $\varepsilon > 0$,

$$P\left(\left|\frac{S_n}{n} - \frac{\sum_{k=1}^n \mu_k}{n}\right| > \varepsilon\right) \leq \frac{Var(S_n)}{n^2 \varepsilon^2} \\ \rightarrow 0$$

- Chebyshev 大数律的意义
 - (1) 样本均值渐近总体均值
 - (2) 没有独立性要求

- Chebyshev 大数律的意义
 - (1) 样本均值渐近总体均值
 - (2) 没有独立性要求
- Chebyshev 大数律的不足之处
要求方差存在

- Khinchine

- Khinchine 大数律

假设 $\xi_k, k \geq 1$ 是一列独立同分布的随机变量, $E\xi_k = \mu$. 记 $S_n = \sum_{k=1}^n \xi_k$, 那么

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

- Khinchine 大数律的证明

(1) 由于没有方差条件，因此不能直接使用Chebyshev不等式

(2) 但可以先截尾，然后使用Chebyshev不等式，
(略)

(3) 后面将给出另一种证明

- 应用

例1. 令 ξ_1, ξ_2, \dots , 是一列独立同分布随机变量, 服从参数为 $\lambda > 0$ 的指数分布。

记 $S_n = \sum_{k=1}^n \xi_k$ 。证明:

$$\frac{S_n}{n} \xrightarrow{P} \frac{1}{\lambda}.$$

事实上, $E\xi_1 = \frac{1}{\lambda}$. 直接使用Khinchine 大数律即可。

- 例2. 令 ξ_1, ξ_2, \dots , 是一列独立随机变量, ξ_k 分布如下: $\xi_1 \equiv 0; k \geq 2$,

$$P(\xi_k = k) = P(\xi_k = -k) = \frac{1}{2k \log k}$$

$$P(\xi_k = 0) = 1 - \frac{1}{k \log k}$$

记 $S_n = \sum_{k=2}^n \xi_k$. 证明:

$$\frac{S_n}{n} \xrightarrow{P} 0$$

事实上, $E\xi_k = 0, \text{Var}(\xi_k) = \frac{k}{\log k}$

$$\text{Var}(S_n) = \sum_{k=2}^n \text{Var}(\xi_k) = \sum_{k=2}^n \frac{k}{\log k} \sim \frac{n^2}{\log n}$$

- 例3. 令 ξ_1, ξ_2, \dots , 是一列具有相同分布的随机变量, $E\xi_k = \mu$, $Var(\xi_k) = \sigma^2 < \infty$ 。假设 ξ_k 和 ξ_{k+1} 相关, $k \geq 1$ 。但是, 当 $|k - l| \geq 2$ 时, ξ_k 与 ξ_l 相互独立。记 $S_n = \sum_{k=1}^n \xi_k$ 。证明:

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

$$\text{计算 } Var(S_n) = ?, \quad \frac{Var(S_n)}{n^2} \xrightarrow{?} 0$$

$$Var(S_n) = \sum_{k=1}^n Var(\xi_k) + 2 \sum_{k < l} Cov(\xi_k, \xi_l)$$

$$|Cov(\xi_k, \xi_{k+1})| \leq (Var(\xi_k) \cdot Var(\xi_{k+1}))^{1/2} \leq \sigma^2$$

- Levy 和Feller
 - (1) Paul Levy
 - (2) William Feller

- Levy-Feller 中心极限定理

假设 $\xi_k, k \geq 1$ 是一列独立同分布随机变量,

$E\xi_k = \mu, Var(\xi_k) = \sigma^2$. 记 $S_n = \sum_{k=1}^n \xi_k$,

那么对任意 x ,

$$P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) \rightarrow \Phi(x)$$

即

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0, 1)$$

- Levy-Feller 中心极限定理的意义
 - (1) 应用于一般随机变量，推广了De Moivre-Laplace 中心极限定理
 - (2) 说明测量误差可用正态分布描述，即正态分布无处不在

- Levy-Feller 中心极限定理的证明
 - (1) 组合计算的方法不再适用
 - (2) 后面将用特征函数方法证明

- Levy-Feller 中心极限定理的应用

例1. 假设 ξ_1, ξ_2, \dots , 是一列独立同分布的随机变量, 服从的分布如下:

$$P(\xi_k = 0) = P(\xi_k = -1) = P(\xi_k = 1) = \frac{1}{3}.$$

记 $S_n = \sum_{k=1}^n \xi_k$. 求 $P(S_{10000} > 100) = ?$

$$E\xi_k = 0, \quad Var(\xi_k) = \frac{2}{3}.$$

$$\begin{aligned} & P(S_{10000} > 100) \\ &= P\left(\frac{S_{10000}}{\sqrt{10000 \times \frac{2}{3}}} > \frac{100}{\sqrt{10000 \times \frac{2}{3}}}\right) \\ &= 1 - \Phi\left(\frac{100}{81.65}\right) \approx 0.11 \end{aligned}$$

- 例2. 某车间有200台车床，工作时每台车床60%的时间在开动，每台开动时耗电1千瓦。问应供给这个车间多少电力才能有0.999把握保证正常生产？

$$P(S_{200} < x) \approx 0.999$$

求 $x = ?$

$$\begin{aligned} P(S_{200} < x) &= P\left(\frac{S_{200} - 200 \times 0.6}{\sqrt{200 \times 0.6 \times 0.4}} < \frac{x - 200 \times 0.6}{\sqrt{200 \times 0.6 \times 0.4}}\right) \\ &\approx \Phi\left(\frac{x - 120}{\sqrt{48}}\right) = 0.999 \end{aligned}$$

- Lyapunov

- Lyapunov 中心极限定理

假设 $\xi_k, k \geq 1$ 是一列独立随机变量(不一定同分布), $E\xi_k = \mu_k, \text{Var}(\xi_k) = \sigma_k^2$.

记 $S_n = \sum_{k=1}^n \xi_k, B_n = \sum_{k=1}^n \sigma_k^2$. 如果

(1) $B_n \rightarrow \infty$

(2) $E|X_k|^3 < \infty$, 且

$$\frac{\sum_{k=1}^n E|X_k|^3}{B_n^{3/2}} \rightarrow 0, \quad n \rightarrow \infty$$

那么对任意 x

$$P\left(\frac{\sum_{k=1}^n (\xi_k - \mu_k)}{\sqrt{B_n}} \leq x\right) \rightarrow \Phi(x)$$

即

$$\frac{\sum_{k=1}^n (\xi_k - \mu_k)}{\sqrt{B_n}} \xrightarrow{d} N(0, 1)$$

例. 假设 $\xi_k, k \geq 1$ 是一列独立随机变量,

$P(\xi_k = 1) = p_k, P(\xi_k = 0) = 1 - p_k$, 其中 $0 < p_k < 1$.

如果

$$\sum_{k=1}^n p_k(1 - p_k) \rightarrow \infty$$

那么

$$\frac{\sum_{k=1}^n (\xi_k - p_k)}{\sqrt{\sum_{k=1}^n p_k(1 - p_k)}} \xrightarrow{d} N(0, 1)$$

- 依概率收敛的概念

我们将给出Khinchine大数律和Levy-Feller中心极限定理的证明

- 依概率收敛的概念
- 依概率收敛的基本性质

我们将给出Khinchine大数律和Levy-Feller中心极限定理的证明

- 依概率收敛的概念
- 依概率收敛的基本性质
- 依分布收敛的概念

我们将给出Khinchine大数律和Levy-Feller中心极限定理的证明

- 依概率收敛的概念
- 依概率收敛的基本性质
- 依分布收敛的概念
- 依分布收敛的基本性质

我们将给出Khinchine大数律和Levy-Feller中心极限定理的证明

- 依概率收敛的概念

假设 (Ω, Σ, P) 是一个概率空间, $X, X_n, n \geq 1$ 是一列随机变量, 如果对任意 $\varepsilon > 0$

$$P(\omega : |X_n(\omega) - X(\omega)| > \varepsilon) \rightarrow 0, \quad n \rightarrow \infty$$

我们称 X_n 依概率收敛到 X , 并记做 $X_n \xrightarrow{P} X$.

X 可以是常数

- 极限唯一性

假设 $X_n \xrightarrow{P} X$, $X_n \xrightarrow{P} Y$, 那么

$$P(X = Y) = 1$$

证明：只需证明 $P(X \neq Y) = 0$

注意到，

$$P(X \neq Y) = P(|X - Y| > 0) = P(\cup_{m=1}^{\infty} \{|X - Y| > \frac{1}{m}\})$$

因此, 需要证明对任意 $\varepsilon > 0$,

$$P(|X - Y| > \varepsilon) = 0$$

给定 $\varepsilon > 0$, 对任意 $n \geq 1$

$$P(|X - Y| > \varepsilon) = P(|X_n - X - (X_n - Y)| > \varepsilon)$$

$$\begin{aligned} &\leq P(|X_n - X| + |(X_n - Y)| > \varepsilon) \\ &\leq P(|X_n - X| > \frac{\varepsilon}{2}) + P(|X_n - Y| > \frac{\varepsilon}{2}) \end{aligned}$$

令 $n \rightarrow \infty$,

$$P(|X_n - X| > \frac{\varepsilon}{2}) \rightarrow 0, \quad P(|X_n - Y| > \frac{\varepsilon}{2}) \rightarrow 0$$

$$P(|X - Y| > \varepsilon) = 0$$

- 依概率收敛的判别法则

如果

$$E|X_n - X|^r \rightarrow 0$$

那么

$$X_n \xrightarrow{P} X$$

证明：应用Markov不等式即可

- 依概率收敛的基本性质

(1) 运算性质

如果 $X_n \xrightarrow{P} X, Y_n \xrightarrow{P} Y$, 那么

(i) $X_n \pm Y_n \xrightarrow{P} X \pm Y$

(ii) $X_n \cdot Y_n \xrightarrow{P} X \cdot Y$

(iii) 如果 $P(Y \neq 0) = 1$, 那么 $\frac{X_n}{Y_n} \xrightarrow{P} \frac{X}{Y}$

以下仅证明(i)中加法成立, 其余类似

给定 $\varepsilon > 0$, 对任意 $n \geq 1$

$$P(|X_n + Y_n - (X + Y)| > \varepsilon) \leq P(|X_n - X| + |Y_n - Y| > \varepsilon)$$

$$\leq P(|X_n - X| > \frac{\varepsilon}{2}) + P(|Y_n - Y| > \frac{\varepsilon}{2})$$

$$P(|X_n - X| > \frac{\varepsilon}{2}) \rightarrow 0, \quad P(|Y_n - Y| > \frac{\varepsilon}{2}) \rightarrow 0$$

所以

$$P(|X_n + Y_n - (X + Y)| > \varepsilon) \rightarrow 0$$

- 连续映射

假设 $f: R \mapsto R$ 是连续映射, 如果 $X_n \xrightarrow{P} X$, 那么

$$f(X_n) \xrightarrow{P} f(X)$$

证明：对任意小的 ε, η ，需要证明存在 $N = N(\varepsilon, \eta)$ ，使得当 $n \geq N$

$$P(|f(X_n) - f(X)| > \varepsilon) \leq \eta$$

注意到，存在 $M > 0$ ，

$$P(|X| > M) \leq \frac{\eta}{8}$$

容易证明，存在 N_1 足够大，使得 $n \geq N_1$

$$P(|X_n| > M + 1) \leq \frac{\eta}{4}$$

所以

$$P(|X| > M + 1) \leq \frac{\eta}{4}, P(|X_n| > M + 1) \leq \frac{\eta}{4}$$

f 是连续映射, 因此 f 在 $[-M-1, M+1]$ 上一致连续
对任意 $\varepsilon > 0$, 存在 $\delta > 0$, 如果

$$|x - y| < \delta, \quad x, y \in [-M-1, M+1]$$

那么

$$|f(x) - f(y)| < \varepsilon$$

$$\begin{aligned} P(|f(X_n) - f(X)| > \varepsilon) &\leq P(|X_n - X| > \delta) \\ &\quad + P(|X_n| > M+1) + P(|X| > M+1) \end{aligned}$$

对于上述 $\delta > 0$, 存在 N_2 使得 $n \geq N_2$

$$P(|X_n - X| > \delta) \leq \frac{\eta}{2}$$

选取 $N = \max\{N_1, N_2\}$, 那么

$$P(|f(X_n) - f(X)| > \varepsilon) \leq \frac{\eta}{2} + \frac{\eta}{4} + \frac{\eta}{4} = \eta$$

- 依分布收敛的概念

假设 $X, X_n, n \geq 1$ 是一列随机变量, 其分布函数分别为 $F, F_n, n \geq 1$ 。如果对每个 F 的连续性点 x ,

$$F_n(x) \rightarrow F(x), \quad n \rightarrow \infty$$

那么称 F_n 依分布收敛到 F , 记做 $F_n \xrightarrow{d} F$

也称 X_n 依分布收敛到 X , 记做 $X_n \xrightarrow{d} X$

● 注解

- (1) 如果 F 是在 R 上连续, 那么 F_n 处处收敛到 F
- (2) 一般地, F 不是连续函数。
- (3) 既然 F 是单调有界函数, F 的不连续性点集最多可数个。
- (4) F 的连续性点集在 R 上稠密。

例1. 考虑 $X_n = \frac{1}{n}$, $X \equiv 0$.

$$F_n(x) = \begin{cases} 1 & x > \frac{1}{n} \\ 0 & x \leq \frac{1}{n} \end{cases}$$

$$F(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$

可以看出

- (1) F_n 不是处处收敛到 F
- (2) 除0点外, F_n 处处收敛到 F 。

$$X_n \xrightarrow{d} X$$

例2. 考虑 $X_n = n$

$$F_n(x) = \begin{cases} 1 & x > n \\ 0 & x \leq n \end{cases}$$

对每个 x ,

$$F_n(x) \rightarrow 0, \quad n \rightarrow \infty$$

但 $F(x) \equiv 0$ 不是分布函数

- 依概率收敛意味着依分布收敛

假设 $X_n \xrightarrow{P} X$, 那么 $X_n \xrightarrow{d} X$

证明: 对任意 $x \in R$ 和 $\varepsilon > 0$

$$\begin{aligned} P(X_n \leq x) &= P(X_n \leq x, X_n - X \geq -\varepsilon) \\ &\quad + P(X_n \leq x, X_n - X < -\varepsilon) \end{aligned}$$

$$\begin{aligned} P(X_n \leq x, X_n - X < -\varepsilon) &\leq P(X_n - X < -\varepsilon) \\ &\rightarrow 0 \end{aligned} \tag{1}$$

$$\limsup_{n \rightarrow \infty} P(X_n \leq x) \leq P(X \leq x + \varepsilon) \quad (2)$$

类似地,

$$\liminf_{n \rightarrow \infty} P(X_n \leq x) \geq P(X \leq x - \varepsilon) \quad (3)$$

如果 x 是 F 的连续性点, 那么令 $\varepsilon \rightarrow 0$ 得

$$\lim_{n \rightarrow \infty} P(X_n \leq x) = P(X \leq x)$$

结论成立

- 依分布收敛并不意味着依概率收敛

例. 假设 Y 是非退化对称随机变量,

令 $X_n = Y, n \geq 1, X = -Y$.

那么 $X, X_n, n \geq 1$ 的分布都相同, 因此 $X_n \xrightarrow{d} X$

但

$$P(|X_n - X| > \varepsilon) = P(2|Y| > \varepsilon)$$

所以 X_n 不依概率收敛到 X

● 如果 $X_n \xrightarrow{d} c$, 那么 $X_n \xrightarrow{P} c$

$$P(X_n \leq c + \varepsilon) \rightarrow 1, \quad n \rightarrow \infty$$

$$P(X_n > c + \varepsilon) \rightarrow 0, \quad n \rightarrow \infty$$

类似地,

$$P(X_n < c - \varepsilon) \rightarrow 0, \quad n \rightarrow \infty$$

所以

$$X_n \xrightarrow{P} c$$

- 依分布收敛的判别法则

假设 $X, X_n, n \geq 1$ 是一列随机变量，具有特征函数 $\phi, \phi_n, n \geq 1$. 那么

$$X_n \xrightarrow{d} X$$

当且仅当

$$\phi_n(t) \rightarrow \phi(t), \quad t \in R$$

- 依分布收敛的判别法则
- Levy连续性定理

假设 $X, X_n, n \geq 1$ 是一列随机变量，具有特征函数 $\phi, \phi_n, n \geq 1$. 那么

$$X_n \xrightarrow{d} X$$

当且仅当

$$\phi_n(t) \rightarrow \phi(t), \quad t \in \mathbb{R}$$

- Levy连续性定理的另一种形式

假设 $X_n, n \geq 1$ 是一列随机变量, 具有特征函数 $\phi_n, n \geq 1$.
如果

$$\phi_n(t) \rightarrow \phi(t), \quad t \in R$$

并且 ϕ 在 0 处连续, 那么 ϕ 一定是特征函数. 记与 ϕ 相应的随机变量为 X , 那么

$$X_n \xrightarrow{d} X$$

● Khinchine大数律的证明

回忆

$\xi_k, k \geq 1$ 独立同分布, $E\xi_k = \mu$, 那么

$$\frac{\sum_{k=1}^n \xi_k}{n} \xrightarrow{P} \mu$$

证明: 只需证明

$$X_n \doteq \frac{\sum_{k=1}^n \xi_k}{n} \xrightarrow{d} \mu$$

等价于, 证明

$$\phi_n(t) \doteq Ee^{itX_n} \rightarrow e^{it\mu}$$

$$\phi_n(t) = Ee^{itX_n} = [Ee^{i\frac{t}{n}\xi_1}]^n$$

在0处进行Taylor展开,

$$Ee^{i\frac{t}{n}\xi_1} = 1 + \frac{it\mu}{n} + o\left(\frac{1}{n}\right)$$

所以, 对每个 t

$$\phi_n(t) = \left[1 + \frac{it\mu}{n} + o\left(\frac{1}{n}\right)\right]^n \rightarrow e^{it\mu}$$

- Levy-Feller中心极限定理的证明

回忆: 如果 $\xi_k, k \geq 1$ 独立同分布, $E\xi_k = \mu, \text{Var}(\xi_k) = \sigma^2$, 那么

$$X_n \doteq \frac{\sum_{k=1}^n \xi_k - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0, 1)$$

证明: 只要证明

$$\phi_n(t) \doteq Ee^{itX_n} \rightarrow e^{-\frac{t^2}{2}}$$

注意到

$$Ee^{itX_n} = [Ee^{i\frac{t(\xi_k - \mu)}{\sigma\sqrt{n}}}]^n$$

在0处进行Taylor展开,

$$Ee^{i\frac{t(\xi_k - \mu)}{\sigma\sqrt{n}}} = 1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)$$

所以, 对任意 t ,

$$Ee^{itX_n} = \left[1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right]^n \rightarrow e^{-\frac{t^2}{2}}$$

- 依分布收敛的运算性质

(1) 线性运算

(i) 假设 $X_n \xrightarrow{d} X$, $b_n \rightarrow b$ 。那么

$$X_n + b_n \xrightarrow{d} X + b$$

证明：由于 $X_n \xrightarrow{d} X$ ，所以

$$Ee^{itX_n} \rightarrow Ee^{itX}$$

这样，

$$\begin{aligned} Ee^{it(X_n+b_n)} &= Ee^{itX_n} e^{itb_n} \\ &\rightarrow Ee^{itX} e^{itb} \\ &= Ee^{it(X+b)} \end{aligned}$$

(ii) 假设 $X_n \xrightarrow{d} X$, $a_n \rightarrow a$ 。那么

$$a_n X_n \xrightarrow{d} aX$$

证明：首先假设 $a > 0$ 。那么对任意 $\varepsilon > 0$ ，当 n 充分大

$$\frac{a}{1+\varepsilon} < a_n < \frac{a}{1-\varepsilon}$$

$$\begin{aligned} P(a_n X_n \leq x) &\leq P(X_n \leq \frac{(1+\varepsilon)x}{a}) \\ &\rightarrow P(X \leq \frac{(1+\varepsilon)x}{a}) \\ &= P(aX \leq (1+\varepsilon)x) \end{aligned} \quad (4)$$

类似地，

$$P(a_n X_n \leq x) \geq P(X_n \leq \frac{(1-\varepsilon)x}{a})$$

综合起来, 当 x 是 aX 的连续性点时,

$$P(a_n X_n \leq x) \rightarrow P(aX \leq x), \quad n \rightarrow \infty$$

假设 $a < 0$, 证明完全类似

假设 $a = 0$. 对任意 $\varepsilon, \eta > 0$, 当 n 充分大时, $|a_n| < \frac{\varepsilon}{\eta}$.

取 η 使得 $\pm \frac{1}{\eta}$ 是 X 的分布函数的连续性点,

$$\begin{aligned} P(|a_n X_n| > \varepsilon) &= P(|X_n| > \frac{1}{\eta}) \\ &\rightarrow P(|X| > \frac{1}{\eta}) \quad n \rightarrow \infty \end{aligned}$$

令 $\eta \rightarrow 0$,

$$P(|a_n X_n| > \varepsilon) \rightarrow 0 \quad (5)$$

综合起来

(iii) 假设 $X_n \xrightarrow{d} X$, $a_n \rightarrow a$, $b_n \rightarrow b$ 。那么

$$a_n X_n + b_n \xrightarrow{d} aX + b$$

(2)

假设 $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{P} c$ 。那么

$$Y_n \cdot X_n \xrightarrow{d} cX$$

证明：类似于(ii)

(3) 连续映射

假设 $f: R \mapsto R$ 是连续映射, $X_n \xrightarrow{d} X$ 。那么

$$f(X_n) \xrightarrow{d} f(X)$$

Helly引理

假设 $F, F_n, n \geq 1$ 是一列分布函数, 并且 $F_n \xrightarrow{d} F$, 那么对任意有界连续函数 g

$$\int g(x) dF_n(x) \rightarrow \int g(x) dF(x)$$

对任意 t ,

$$\begin{aligned} Ee^{itf(X_n)} &= \int e^{itf(x)} dF_n(x) \\ &\rightarrow \int e^{itf(x)} dF(x) \\ &= Ee^{itf(X)} \end{aligned} \tag{6}$$

结论成立

例

- 几乎处处收敛的概念

(1) 几乎处处收敛比依概率收敛、依分布收敛强

(2) Borel 大数律比Bernoulli 大数律强

(3) Kolmogorov 大数律比Khinchine 大数律强

- 几乎处处收敛的概念
- Borel-Cantelli 引理

- (1) 几乎处处收敛比依概率收敛、依分布收敛强
- (2) Borel 大数律比Bernoulli 大数律强
- (3) Kolmogorov 大数律比Khinchine 大数律强

- 几乎处处收敛的概念
- Borel-Cantelli 引理
- Borel 大数律

- (1) 几乎处处收敛比依概率收敛、依分布收敛强
- (2) Borel 大数律比Bernoulli 大数律强
- (3) Kolmogorov 大数律比Khinchine 大数律强

- 几乎处处收敛的概念
- Borel-Cantelli 引理
- Borel 大数律
- Kolmogorov 强大数律

(1) 几乎处处收敛比依概率收敛、依分布收敛强

(2) Borel 大数律比Bernoulli 大数律强

(3) Kolmogorov 大数律比Khinchine 大数律强

- 几乎处处收敛的概念

- (1) 处处收敛

假设 (Ω, Σ, P) 是一个概率空间, $X, X_n, n \geq 1$ 是一列随机变量, 如果对每个 $\omega \in \Omega$,

$$X_n(\omega) \rightarrow X(\omega), \quad n \rightarrow \infty$$

那么称 X_n 处处收敛于 X

(2) 几乎处处收敛

假设 (Ω, Σ, P) 是一个概率空间, $X, X_n, n \geq 1$ 是一列随机变量, 如果存在 Ω_0 使得

(i) $P(\Omega_0) = 0$

(ii) 对每个 $\omega \in \Omega \setminus \Omega_0$,

$$X_n(\omega) \rightarrow X(\omega), \quad n \rightarrow \infty$$

那么称 X_n 几乎处处收敛于 X , 记做 $X_n \rightarrow X, a.s.$

即除一个零概率事件外, X_n 处处收敛于 X

- 几乎处处收敛的判别法则

$$X_n \rightarrow X, \quad a.s.$$

当且仅当

对任意 $\varepsilon > 0$,

$$P\left(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} \{|X_n(\omega) - X(\omega)| > \varepsilon\}\right) = 0$$

或者说

$$P(\{|X_n(\omega) - X(\omega)| > \varepsilon\}, i.o.) = 0$$

等价地

对任意 $\varepsilon > 0$,

$$\lim_{N \rightarrow \infty} P\left(\bigcup_{n=N}^{\infty} \{|X_n(\omega) - X(\omega)| > \varepsilon\}\right) = 0$$

由此, 也可看出几乎处处收敛比依概率收敛强。

- Borel-Cantelli 引理

(1) 假设 $A_n, n \geq 1$ 是一列事件, 如果

$$\sum_{n=1}^{\infty} P(A_n) < \infty$$

那么

$$P(A_n, i.o.) = 0$$

证明:

$$P(A_n, i.o.) = \lim_{N \rightarrow \infty} P\left(\bigcup_{n=N}^{\infty} A_n\right)$$

$$P\left(\bigcup_{n=N}^{\infty} A_n\right) \leq \sum_{n=N}^{\infty} P(A_n) \rightarrow 0$$

(2) 假设 $A_n, n \geq 1$ 是一列独立事件, 如果

$$\sum_{n=1}^{\infty} P(A_n) = \infty$$

那么

$$P(A_n, i.o.) = 1$$

(1) 证明略

(2) Borel-Cantelli 引理也被称为Borel-Cantelli 0-1律

- Borel

- Borel 大数律

假设 (Ω, Σ, P) 是一个概率空间, $\xi_k, k \geq 1$ 是一列独立同分布随机变量,

$$P(\xi_k = 1) = p, \quad P(\xi_k = 0) = 1 - p$$

记 $S_n = \sum_{k=1}^n \xi_k$, 那么

$$\frac{S_n}{n} \rightarrow p, \quad a.s.$$

- Borel 大数律的证明

需要证明, 对任意 $\varepsilon > 0$

$$P(|\frac{S_n}{n} - p| > \varepsilon, i.o.) = 0$$

由Borel-Cantelli引理, 只需验证

$$\sum_{n=1}^{\infty} P(|\frac{S_n}{n} - p| > \varepsilon) < \infty$$

由Markov不等式,

$$P(|\frac{S_n}{n} - p| > \varepsilon) \leq \frac{E|S_n - np|^4}{n^4 \varepsilon^4}$$

容易计算得,

$$E|S_n - np|^4 = np(1-p)[p^3 + (1-p)^3] + n(n-1)p^2(1-p)^2$$

$$\sum_{n=1}^{\infty} P(|\frac{S_n}{n} - p| > \varepsilon) \leq K(\varepsilon, p) \frac{1}{n^2} < \infty$$

- Kolmogorov

- Kolmogorov 强大数律

假设 (Ω, Σ, P) 是一个概率空间, $\xi_k, k \geq 1$ 是一列独立同分布随机变量. 如果 $E\xi_k = \mu$, 那么

$$\frac{S_n}{n} \rightarrow \mu, \quad a.s.$$

(1) Kolmogorov 强大数律推广了Borel 强大数律

(2) Kolmogorov 强大数律推广了Khinchine 大数律