随机过程及应用

Stochastic Processes with its Applications

电子科技大学 数学科学学院

武德安

第三章 二阶矩过程的均方微积分

§3.1 收敛性与极限定理

一、分布函数弱收敛

定义3.1.1 对于分布函数列$\{F_n(x)\}$,如果存在单调不降函数$F(x)$,使

$\lim\limits_{n\to\infty} F_n(x)=F(x)$,在$F(x)$的每一连续点成立,称$F_n(x)$弱收敛于$F(x)$.

记为$F_n(x)\xrightarrow{W}F(x)$.

注: 分布函数列的极限函数$F(x)$是有界非降函数,但不一定是分布函数.

定理3.1.1 连续性定理(列维—克拉美) 正极限定理设分布函数列$\{F_n(x)\}$弱收敛于一分布函数$F(x)$,则相应的特征函数列收敛于特征函数,且在t的任一有限区间内收敛是一致的.

$$F_n(x)\xrightarrow{W}F(x)\Rightarrow\{\varphi_n(t)\}\to\varphi(t)\text{ 一致成立}.$$

逆极限定理 设特征函数列$\{\varphi_n(t)\}$收敛于某一函数$\varphi(t)$，且$\varphi(t)$在$t=0$连续，则相应的分布函数列$\{F_n(x)\}$弱收敛于某一分布函数$F(x)$，而且是$F(x)$的特征函数。

$$\{\varphi_n(t)\} \to \varphi(t) \xrightarrow{\text{在 } t=0 \text{ 处连续}} F_n(x) \xrightarrow{W} F(x)$$

连续性定理可用来确定随机变量序列的极限分布.

Eg.1： 设随机变量序列$X_1, X_2, \ldots$相互独立,且$X_k \sim P(\lambda)(k=1,2,\ldots)$.

1)求$Y_n = \sum_{k=1}^{n} X_k$的概率分布;

2)证明:当$n \to \infty$时,$Y_n^* = \dfrac{Y_n - E(Y_n)}{\sqrt{D(Y_n)}}$趋于$N(0,1)$.

解1) $\varphi_k(t) = e^{\lambda(e^{it}-1)} \Rightarrow \varphi_{Y_n}(t) = \prod_{k=1}^{n} \varphi_k(t) = e^{n\lambda(e^{it}-1)}$

即$Y_n \sim P(n\lambda)$,且$E(Y_n) = D(Y_n) = n\lambda$.

2) Y_n^*的特征函数为

$$\varphi_{Y_n^*}(t) = e^{-i\sqrt{n\lambda}t}\varphi_{Y_n}\left(\frac{t}{\sqrt{n\lambda}}\right) = e^{-i\sqrt{n\lambda}t}e^{-n\lambda}e^{n\lambda\exp\left(i\frac{t}{\sqrt{n\lambda}}\right)}$$

$$= e^{-n\lambda}e^{-i\sqrt{n\lambda}t}\exp\left[n\lambda\left(1+\frac{it}{\sqrt{n\lambda}} - \frac{t^2}{2n\lambda} + \cdots\right)\right] = e^{n\lambda\left[-\frac{t^2}{2n\lambda} + \cdots\right)]}$$

有$\lim\limits_{n\to\infty}\varphi_{Y_k^*}(t) = e^{-\frac{t^2}{2}}, \quad t \in R$.由连续性定理的逆定理知当$n \to \infty$时,$Y_n^*$趋于正态分布.

二、随机变量的收敛性

定义3.1.2 设随机变量序列$\{X_n\}$的分布函数列$\{F_n(x)\}$弱收敛于随机变量X的分布函数$F(x)$,称$\{X_n\}$**依分布收敛**(**convergence in distribution**)于X,记为

$$X_n \xrightarrow{d} X, \quad \text{as} \quad n \to \infty$$

注: 直观上而言，依分布收敛只在乎随机变量的分布，而不在乎他们之间的相互关系。

例: 倘若已知$X_n \xrightarrow{d} X$，$X \sim N(0,1)$，假设$Y = -X$。对于任意一个发生的事件，Y与X的取值正好差了一个负号。

但这并不影响X与Y有相同的累积函数，即$F_X(z) = F_Y(z)$。如此一来，$X_n \xrightarrow{d} Y$。

因为依分布收敛仅仅在乎分布，而不在乎随机变量值相互之间的关系。

定义3.1.3 设$\{X_n\}, n=1,2,\ldots$是定义在(Ω, F, P)上的随机变量序列,若对

$$\forall \varepsilon>0, \lim_{n\to\infty} P\{|X_n - X| \geq \varepsilon\}=0, \text{或} \quad \lim_{n\to\infty} P\{|X_n - X| < \varepsilon\}=1.$$

称随机变量序列$\{X_n\}$**依概率收敛**于X,记为

$$\lim_{n\to\infty} X_n = X \quad (p) \quad \text{或} \quad X_n \xrightarrow{p} X.$$

注: 也就是说，当n很大的时候，对任意发生的事件，X_n的值和X的值差不多，即$|(X_n - X)(\omega)|$很小。直观上而言，依概率收敛考虑的是随机变量的值。

以此思路来考察一下**依分布收敛与依概率收敛的关系**:

如果$X_n \xrightarrow{P} X$，但对于任何一个与X分布一样的Y，但$P(X=Y)<1$，$X_n \xrightarrow{P} Y$一定不成立，因为X与Y只是分布相同，而值不同。

反之，如果$X_n \xrightarrow{P} X$，即它们的值都差不多了，那么它们的分布一定也差不多，即$X_n \xrightarrow{d} X$。因此，依概率收敛比依分布收敛要强，即$X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$。

特殊情况: 但在某种情况下，取值就可以确定分布。即X取某个常数的情况下。此时X的取值和X的分布唯一确定。即此时会有依分布收敛和依概率收敛等价，即$X_n \xrightarrow{d} c \Leftrightarrow X_n \xrightarrow{P} c$。

定义3.1.4 设$\{X_n\}, n=1,2,\ldots$是定义在(Ω, F, P)上的随机变量序列,若$E(|X_n|^2)<\infty$,且$\lim\limits_{n\to\infty} E[|X_n - X|^2]=0$,称随机变量序列$\{X_n\}$**均方收敛**于$X$,记为

$$\underset{n\to\infty}{\mathrm{l.i.m}}\, X_n = X.$$

注: 直观上而言,均方收敛也关注是随机变量的值,但其要求比依概率收敛更加严格。之所以更加严格,是因为概率测度被均方测度所限制,其思想可以近似由Chebyshev不等式看到。

$$P(|X-\mu|\geq\varepsilon)\leq\frac{E(X-\mu)^2}{\varepsilon^2}$$ 。因此 $X_n \xrightarrow{\mathrm{l.i.m}} X \Rightarrow X_n \xrightarrow{P} X.$

定义3.1.5： 设$\{X_n\}, n=1,2,\dots$是定义在(Ω, F, P)上的随机变量序列,若存在一个随机变量X(可以是常数),使$P\left\{\lim\limits_{n\to\infty} X_n = X\right\}=1$称随机变量序列$\{X_n\}$以**概率为1收敛**于$X$,或称**几乎处处收敛**于$X$,记为

$$X_n \xrightarrow{\text{a.s.}} X. \text{ 或 } \quad \lim_{n\to\infty} X_n = X \quad (\text{a.s.})$$

注： 直观上而言，几乎处处收敛在乎的也是随机变量的值，但其要求也比依概率收敛更加严格。如果没有接触过实变函数的知识，几乎处处收敛对于连续型随机变量可能比较难以理解。

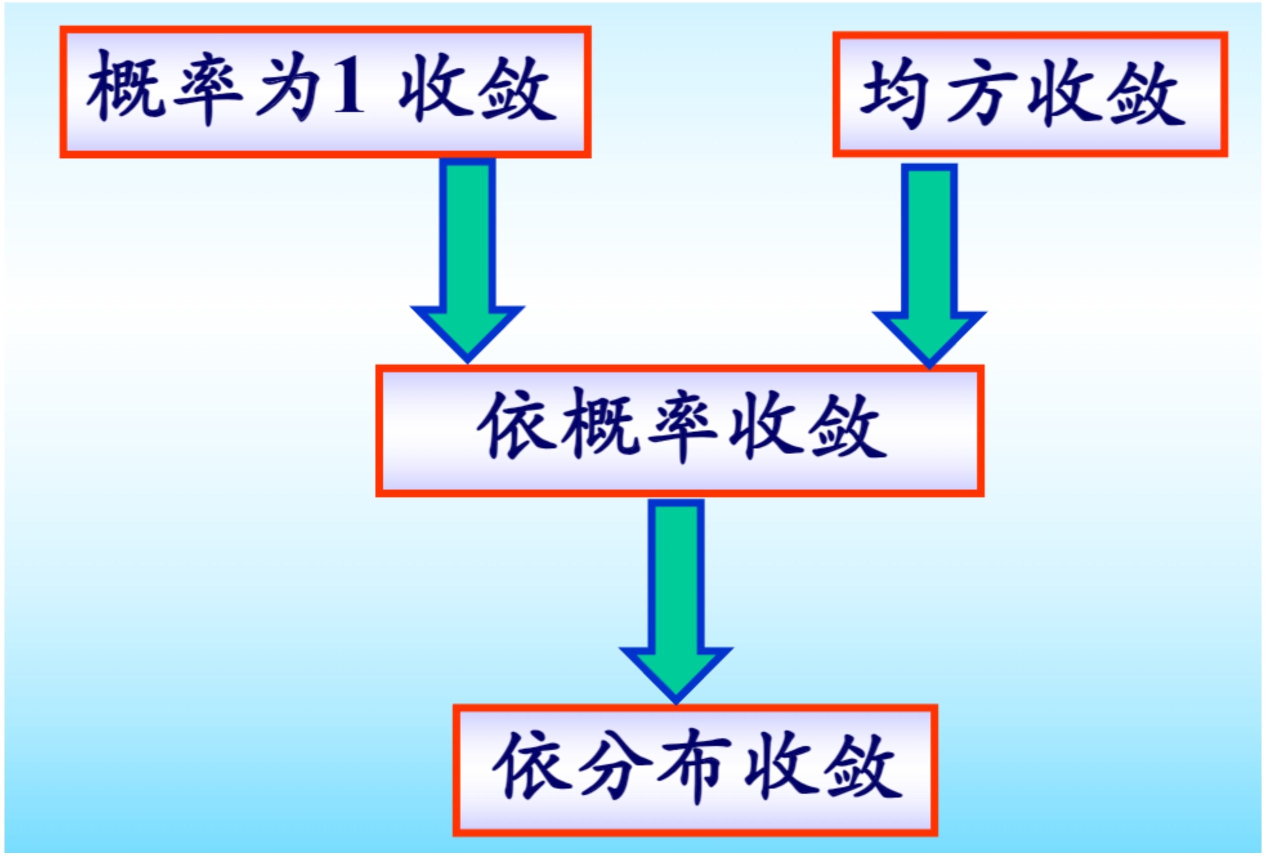
例： 用离散型随机变量进行直观解释，以避免0测度下的一些问题。

对于$X_n \sim \mathrm{Ber}(p_n)$，即以$P\{X_n=1\}=p_n$，$P\{X_n=0\}=1-p_n$的随机变量。

$X_n \xrightarrow{p} 1$：其依概率收敛到1意味着，X_n和1的值都差不多，而且随着n越来越大，不相等的概率越来越小。换而言之，出现0的概率越来越小，极限为0。

但几乎处处收敛至1要求，存在N，$n>N$时，$X_n=1$，即X_n和1的值都在n很大时必须相等，即X_n取0的概率在某个N后必须为0。前者限制其尾部概率收敛至0，但后者限制尾部概率为0。

三、几种收敛性的关系



证明随机变量序列$\{X_n\}$均方收敛于X,则一定依概率收敛于X.

证: 由马尔科夫不等式,对$\forall\varepsilon>0$,有$P\{X_n-X|\geq\varepsilon\}\leq\dfrac{E[|X_n-X|^2]}{\varepsilon^2}$

从而$\lim\limits_{n\to\infty}P\{X_n-X|\geq\varepsilon\}=0$.

四、极限定理

定理4.1.2.切比雪夫(Chebyshev)大数定律 设$X_k,k=1,2\ldots$是相互独立的随机变量序列,其数学期望和方差都存在,且存在常数C,*使得*$D(X_n)<C$,$k=1,2,\ldots$则对于任意的$\varepsilon>0$,有

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

定理4.1.3 独立同分布大数定律 设$X_k,k=1,2\dots$是相互独立且同分布的随机变量序列,且$E(X_k)=\mu,D(X_k)=\sigma^2,k=1,2,\dots$则$\{X_k\},k=1,2\dots$服从大数定律,即对任意的$\varepsilon>0$,有

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

定理4.1.4辛钦大数定律 设$X_k,k=1,2\dots$是相互独立且同分布的随机变量序列,且$E(X_k)=\mu,k=1,2,\dots$则$\{X_k\},k=1,2\dots$服从大数定律,即对任意的$\varepsilon>0$,有

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

定理4.1.5贝努里(Bernulli)大数定律 设$\frac{m}{n}$是n次重复独立试验中事件A发生的频率，p是事件A在每次试验中发生的概率。则对任意的$\forall \varepsilon > 0$，有

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

定理4.1.6独立同分布中心极限定理 设$\{X_k\}, k = 1, 2 \ldots$为相互独立,具有相同分布的随机变量序列,且$E(X_k) = \mu, D(X_k) = \sigma^2$,则$\{X_k\}$满足中心极限定理,即有

$$\lim_{n \to \infty} P\left\{ \frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\,\sigma} \le x \right\} = \Phi(x)$$

其中$\Phi(x)$是标准正态分布$N(0, 1)$的分布函数。

Appendix

Cauchy-Schwarz不等式

Let us introduce two *real indeterminants* [dummy variables] ξ and η and consider the identity:

$$E\{(X\xi + Y\eta)^2\} = E(X^2)\xi^2 + 2E(XY)\xi\eta + E(Y^2)\eta^2.$$

The right member is a quadratic form in (ξ, η) and it is never negative because the left member is the expectation of a random variable which

does not take negative values. A well-known result in college algebra says that the coefficients of such a quadratic form $a\xi^2 + 2b\xi\eta + c\eta^2$ must satisfy the inequality $b^2 \leq ac$. Hence in the present case

$$E(XY)^2 \leq E(X^2)E(Y^2).$$

This is called the *Cauchy–Schwarz inequality*.