第 7 章　离散鞅引论

鞅 (martingale) 论目前已成为研究概率论以及应用概率论和其他随机过程的有力工具. 在统计，序贯决策，最优控制，随机微分方程等方面均得到了广泛应用．鞅论的发展与现今的竞争社会是分不开的．有奖彩票，保险，投资建设等均与鞅论有关.

7.1　定义与例子

定义 7.1.1　过程 $\{X_n, n \geqslant 0\}$ 是**鞅**, 如果 $\forall n \geqslant 0$, 有

(1) $E|X_n| < \infty$,

(2) $E(X_{n+1} \big| X_0, X_1, \cdots, X_n) = X_n \qquad$ a.s. (几乎处处).

鞅的背景来源于公平赌博．上式表明，如第 n 次赌博后资金为 X_n, 则第 $n+1$ 次赌博后的平均资金恰等于 X_n, 即每次赌博胜负机会均等.

有时 $\{X_n, n \geqslant 0\}$ 不能直接观察，而只能观察另一过程 $\{Y_n, n \geqslant 0\}$. 故做如下定义：

定义 7.1.2　设有两个过程 $\{X_n, n \geqslant 0\}$ 及 $\{Y_n, n \geqslant 0\}$, 称 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是鞅，如果

(1) $E|X_n| < \infty$,

(2) $E(X_{n+1} \big| Y_0, Y_1, \cdots, Y_n) = X_n \qquad$ a.s.(几乎处处).

解释　① 因为 $X_n = E(X_{n+1} \big| Y_0, \cdots, Y_n)$ 是 $(Y_0, Y_1, \cdots, Y_n)$ 的函数,故有 $E(X_n \big| Y_0, Y_1, \cdots, Y_n) = X_n$.

② $EX_{n+1}=E[E(X_{n+1}|Y_0,\cdots,Y_n)]=EX_n=EX_0$. 这说明鞅 $\{X_n,n\geqslant 0\}$ 在任何时刻的期望值均相等.

下面介绍一些鞅的典型例子.

例 1 独立同分布随机变量之和

设 $Y_0=0,\{Y_n,n\geqslant 1\}$ 独立同分布，$EY_n=0,E|Y_n|<\infty,X_0=0,X_n=\sum\limits_{i=1}^{n}Y_i$, 则 $\{X_n,n\geqslant 0\}$ 关于 $\{Y_n,n\geqslant 0\}$ 是鞅.

证明 因为

$$\begin{aligned}E|X_n|=&E\left|\sum_{i=1}^{n}Y_i\right|<\infty,\\E(X_{n+1}|Y_0,Y_1,\cdots,Y_n)=&E(X_n+Y_{n+1}|Y_0,Y_1,\cdots,Y_n)\\=&E(X_n|Y_0,\cdots,Y_n)+E(Y_{n+1}|Y_0,\cdots,Y_n)=X_n.\quad\square\end{aligned}$$

例 2 和的方差

设 $Y_0=0,\{Y_n,n\geqslant 1\}$ 独立同分布，$EY_n=0,EY_n^2=\sigma^2,X_0=0,X_n=\left(\sum\limits_{k=1}^{n}Y_k\right)^2-n\sigma^2$, 则 $\{X_n,n\geqslant 0\}$ 关于 $\{Y_n,n\geqslant 0\}$ 是鞅.

证明 因为

$$\begin{aligned}E|X_n|=&E\left|\left(\sum_{k=1}^{n}Y_k\right)^2-n\sigma^2\right|\leqslant E\left|\left(\sum_{k=1}^{n}Y_k\right)^2\right|+n\sigma^2\\=&E\left(\sum_{k=1}^{n}Y_k^2+\sum_{i\neq j}Y_iY_j\right)+n\sigma^2=2n\sigma^2<\infty;\end{aligned}$$

所以

$$
\begin{aligned}
&E(X_{n+1}\big|Y_0,\cdots,Y_n)\\
=&E\Big[\Big\{\Big(Y_{n+1}+\sum_{k=1}^{n}Y_k\Big)^2-(n+1)\sigma^2\Big\}\big|Y_0,\cdots,Y_n\Big]\\
=&E\Big[\Big\{Y_{n+1}^2+2Y_{n+1}\sum_{k=1}^{n}Y_k+\Big(\sum_{k=1}^{n}Y_k\Big)^2-(n+1)\sigma^2\Big\}\big|Y_0,\cdots,Y_n\Big]\\
=&E[Y_{n+1}^2\big|Y_0,\cdots,Y_n]+2E\Big(Y_{n+1}\sum_{k=1}^{n}Y_k\big|Y_0,\cdots,Y_n\Big)+E(X_n\big|Y_0,\cdots,Y_n)-\sigma^2\\
=&EY_{n+1}^2+2E(Y_{n+1}\big|Y_0,\cdots,Y_n)\Big(\sum_{k=1}^{n}Y_k\Big)+X_n-\sigma^2\\
=&\sigma^2+0+X_n-\sigma^2=X_n.
\end{aligned}
$$

□

由上两例知，由独立同分布随机变量的和或者和的方差所构成的序列都可以构造鞅，那么更一般的结论呢？

例 3　一般和

设 $\{Y_n, n\geqslant 0\}$ 为一随机序列, $Z_i=g_i(Y_0,\cdots,Y_i), g_i$ 为一般函数. 函数 f 满足 $E|f(Z_k)|<\infty$. $a_k(y_0,\cdots,y_{k-1})\,(k\geqslant 0)$ 为 k 元有界实函数，即

$$|a_k(y_0,\cdots,y_{k-1})|\leqslant A_k,\ \forall\, y_0,\cdots,y_{k-1}.$$

约定

$$a_0(Y_{-1})=a_0,\quad E[f(Z_0)\big|Y_{-1}]=E[f(Z_0)].$$

令

$$X_n = \sum_{k=0}^{n} \{f(Z_k) - E[f(Z_k)\big|Y_0, Y_1, \cdots, Y_{k-1}]\} \cdot a_k(Y_0, \cdots, Y_{k-1}),$$

可以验证，$\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是鞅.

证明 (1)

$$\begin{aligned} E|X_n| &\leqslant \sum_{k=0}^{n} E|\{f(Z_k) - E[f(Z_k)\big|Y_0, \cdots, Y_{k-1}]\}a_k(Y_0, \cdots, Y_{k-1})| \\ &\leqslant \sum_{k=0}^{n} A_k\{E[|f(Z_k)|] + E\{E[|f(Z_k)|\big|Y_0, \cdots, Y_{k-1}]\}\} \\ &\leqslant \sum_{k=0}^{n} 2A_k E|f(Z_k)| < \infty. \end{aligned}$$

(2) 记

$$B_k = \{f(Z_k) - E[f(Z_k)\big|Y_0, \cdots, Y_{k-1}]\}a_k(Y_0, \cdots, Y_{k-1}),$$

则

$$\begin{aligned} E(B_k\big|Y_0, \cdots, Y_{k-1}) =& a_k(Y_0, \cdots, Y_{k-1})\{E[f(Z_k)\big|Y_0, \cdots, Y_{k-1}] - \\ & E(E[f(Z_k)\big|Y_0, \cdots, Y_{k-1}]\Big|Y_0, \cdots, Y_{k-1})\} \\ =& a_k(Y_0, \cdots, Y_{k-1})\{E[f(Z_k)\big|Y_0, \cdots, Y_{k-1}] - \\ & E[f(Z_k)\big|Y_0, \cdots, Y_{k-1}]\} \\ =& 0, \end{aligned}$$

上式推导中用到了 $a_k(Y_0,\cdots,Y_{k-1})$ 及 $E[f(Z_k)\big|Y_0,Y_1,\cdots,Y_{k-1}]$ 都是 $Y_0,\cdots,Y_{k-1}$ 的函数的事实. 由于 $E(B_k\big|Y_0,\cdots,Y_{k-1})=0$, 且 X_n 是 $Y_0,\cdots,Y_n$ 的函数，故

$$E(X_{n+1}\big|Y_0,\cdots,Y_n)=E(X_n\big|Y_0,\cdots,Y_n)+E(B_{n+1}\big|Y_0,\cdots,Y_n)=X_n. \qquad \square$$

由此可知，由一个一般的随机序列也可以构造出鞅来. 这是一个有意义的结论. 下面介绍几个具有特殊实用价值的鞅.

例 4 由马尔可夫链导出的鞅

设 $\{Y_n,n\geqslant 0\}$ 是马尔可夫链 (其状态空间为 S), 具有转移概率矩阵 $\boldsymbol{P}=(p_{ij})$, f 是 $\boldsymbol{P}$ 的有界右正则序列 (调和函数), 即 $f(i)\geqslant 0$, 且

$$f(i)=\sum_{j\in S}p_{ij}f(j),\quad |f(i)|<M,\quad i\in S.$$

令 $X_n=f(Y_n)$, 则 $\{X_n,n\geqslant 0\}$ 关于 $\{Y_n,n\geqslant 0\}$ 是鞅.

证明 因为

$$E|X_n|<\infty,$$

又由于

$$\begin{aligned}E(X_{n+1}\big|Y_0,Y_1,\cdots,Y_n)&=E(f(Y_{n+1})\big|Y_0,Y_1,\cdots,Y_n)\\&=E(f(Y_{n+1})\big|Y_n)\quad (\text{由马氏性得到})\\&=\sum_{j\in S}f(j)P(Y_{n+1}=j\big|Y_n)\\&=\sum_{j\in S}f(j)p_{Y_nj}=f(Y_n)=X_n.\end{aligned}$$

因此 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是鞅. □

例 5 由转移概率特征向量导出的鞅

设 $\{Y_n, n \geqslant 0\}$ 是一马尔可夫链, $\boldsymbol{P} = (p_{ij})$. 向量 $\boldsymbol{f} = (f(0), f(1), \cdots, f(i), \cdots)$ 称为 $\boldsymbol{P}$ 的右特征向量, 如果对某个 λ(称 λ 为特征值) 有

$$\lambda f(i) = \sum_{j \in S} p_{ij} f(j), \quad \forall\, i \in S,$$

且 $E|f(Y_n)| < \infty$, $\forall\, n$. 令 $X_n = \lambda^{-n} f(Y_n)$, 则 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是鞅.

证明 因为

$$\begin{aligned} E|X_n| &= E|\lambda^{-n}\ f(Y_n)| = \lambda^{-n} E|f(Y_n)| < \infty, \\ E(X_{n+1}\big|Y_0, Y_1, \cdots, Y_n) &= E[\lambda^{-n-1} f(Y_{n+1})\big|Y_0, Y_1, \cdots, Y_n] \\ &= \lambda^{-n}\ \lambda^{-1}\ E[f(Y_{n+1})\big|Y_n] \end{aligned}$$

$$= \lambda^{-n}\lambda^{-1} \sum_{j \in S} [f(j)\ p_{Y_n j}] = \lambda^{-n} f(Y_n) = X_n.$$

更一般地, 设 $\{Y_n, n \geqslant 0\}$ 是一离散时间马尔可夫过程, 具有转移分布函数

$$F(y\big|z) = P\{Y_{n+1} \leqslant y\big|Y_n = z\}.$$

如果对所有 n, 有 $E|f(Y_n)| < \infty$, 且

$$\lambda f(y) = \int f(z)\,\mathrm{d}F(z|y),$$

则 $\{X_n = \lambda^{-n} f(Y_n), n \geqslant 0\}$ 是一个鞅. □

例 4 和例 5 将上一章讨论的马尔可夫链与本章的鞅这两个重要的随机过程有机地联系起来了，在实际中这样的应用非常广泛.

例 6 由分支过程构成的鞅

设 $\{Y_n, n \geqslant 0\}$ 表示一分支过程，并设生成后代分布的均值为 $m < \infty$, 则 $X_n = m^{-n} Y_n$ 关于 $\{Y_n, n \geqslant 0\}$ 是鞅.

证明 设 $Z^{(n)}(j)$ 为第 n 代的第 j 个个体产生的个体的数目，$Z^{(n)}(i)(i = 1, 2, \cdots)$ 独立同分布, $E\{Z^{(n)}(i)\} = m$, 并设 Y_n 为第 n 代的个体数. 若 $Y_n = 0$, 取 $Y_{n+1} = 0$; 若 $Y_n \neq 0$, 则

$$Y_{n+1} = Z^{(n)}(1) + Z^{(n)}(2) + \cdots + Z^{(n)}(Y_n).$$

显然

$$\begin{aligned} E(Y_{n+1}|Y_n) =& E\{Z^{(n)}(1) + \cdots + Z^{(n)}(Y_n)\big|Y_n\} \\ =& E\{Z^{(n)}(1) + \cdots + Z^{(n)}(Y_n)\} = Y_n E(Z^{(n)}(1)) = mY_n. \end{aligned}$$

所以,m 是函数 $f(y) = y$ 的特征值. 根据上例的结论, 容易导出 $X_n = m^{-n} Y_n$ 关于 $\{Y_n, n \geqslant 0\}$ 是鞅. □

例 7 Wald 鞅

设 $Y_0 = 0, Y_1, Y_2, \cdots, Y_n, \cdots$ 独立同分布, 且存在一有限矩生成函数 $\phi(\lambda) = E[\exp(\lambda Y_n)]$, 对 $\lambda \neq 0$, 令 $X_0 = 1, X_n = \phi^{-n}(\lambda) \exp[\lambda(Y_1 + \cdots + Y_n)]$, 那么，$\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是鞅.

证明 首先证明函数 $f(y) = \exp(\lambda y)$ 是部分和 $S_n = Y_1 + Y_2 + \cdots + Y_n$ 马尔可夫过程的特征函数，对应的特征值是 $\phi(\lambda)$.

事实上，若 G 是 Y_n 的分布函数，则有 $P\{S_{n+1} \leqslant y | S_n = x\} = G(y-x)$, 因而有

$$\int \exp(\lambda y)\,\mathrm{d}G(y-x) = \exp(\lambda x)\int \exp(\lambda\zeta)\,\mathrm{d}G(\zeta) = \exp(\lambda x)\phi(\lambda).$$

由例 5 推广的结论知，当

$$X_n = \phi^{-n}(\lambda)\ f(S_n) = \phi^{-n}(\lambda)\ \exp[\lambda S_n]$$

时，$\{X_n, n \geqslant 0\}$ 关于 $\{S_n, n \geqslant 0\}$ 是鞅，即 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是鞅. □

在上例中，假设 $Y_1, Y_2, \cdots, Y_n \cdots$ 独立同分布，且服从 $N(0, \sigma^2)$, 则

$$\begin{aligned}\phi(\lambda) =& E(\exp(\lambda Y_1)) = \exp\left(\frac{1}{2}\lambda^2\sigma^2\right),\\ X_n =& \exp\left\{\lambda(Y_1 + \cdots + Y_n) - \frac{n}{2}\lambda^2\sigma^2\right\}.\end{aligned}$$

若令 $\lambda = \dfrac{\mu}{\sigma^2}$, 得到

$$X_n = \exp\left\{\frac{\mu}{\sigma^2}(Y_1 + Y_2 + \cdots + Y_n) - \frac{n\mu^2}{2\sigma^2}\right\},$$

则 $\{X_n, n \geqslant 1\}$ 关于 $\{Y_n, n \geqslant 1\}$ 是鞅.

这是一个非常有用的结论，因为正态分布是经常遇到且长于研究的一种分布.

例 8 似然比构成的鞅

设 $Y_0, Y_1, \cdots, Y_n, \cdots$ 是独立同分布随机变量序列，f_0 和 f_1 是概率密度函数，令

$$X_n = \frac{f_1(Y_0)f_1(Y_1)\cdots f_1(Y_n)}{f_0(Y_0)f_0(Y_1)\cdots f_0(Y_n)}, \quad n \geqslant 0.$$

假设 $\forall y,\ f_0(y) > 0$. 当 Y_n 的概率密度函数为 f_0 时, 则 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是鞅.

证明 因为

$$E|X_n| = E\left[\frac{f_1(Y_0)\cdots f_1(Y_n)}{f_0(Y_0)\cdots f_0(Y_n)}\right] = 1 < \infty,$$

且

$$E(X_{n+1}\big|Y_0, \cdots, Y_n) = E\left[X_n \frac{f_1(Y_{n+1})}{f_0(Y_{n+1})}\big|Y_0, Y_1, \cdots, Y_n\right] = X_n E\left[\frac{f_1(Y_{n+1})}{f_0(Y_{n+1})}\right].$$

由于

$$E\left[\frac{f_1(Y_{n+1})}{f_0(Y_{n+1})}\right] = \int \frac{f_1(y)}{f_0(y)}\ f_0(y)\,\mathrm{d}y = \int f_1(y)\,\mathrm{d}y = 1,$$

因此 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是鞅. □

在上例中, 设 f_0 是正态概率密度, 均值为 0, 方差为 σ^2; f_1 也是一正态概率密度, 均值为 μ, 方差为 σ^2, 则

$$\begin{aligned}\frac{f_1(y)}{f_0(y)} &= \exp\left\{\frac{2\mu y - \mu^2}{2\sigma^2}\right\},\\ X_n &= \exp\left\{\frac{\mu}{\sigma^2}(Y_1 + Y_2 + \cdots + Y_n) - \frac{n\mu^2}{2\sigma^2}\right\}.\end{aligned}$$

由此可得出与例 7 结尾相同的结论.

例 9 Doob 鞅过程

设 $Y_0, Y_1, \cdots, Y_n, \cdots$ 是一随机序列, 有一随机变量 X, $E|X| < \infty$, 令

$$X_n = E(X\big|Y_0, \cdots, Y_n),$$

则 $\{X_n, n \geqslant 0\}$ 是关于 $\{Y_n, n \geqslant 0\}$ 的鞅，并称之为 Doob 过程.

证明 因为

$$E|X_n| = E\{|E(X\big|Y_0,\cdots,Y_n)|\} \leqslant E\{E(|X|\big|Y_0,Y_1,\cdots,Y_n)\} = E|X| < \infty,$$

$$\begin{aligned}E(X_{n+1}\big|Y_0,\cdots,Y_n) =& E\{E(X\big|Y_0,\cdots Y_{n+1})\big|Y_0,\cdots,Y_n\}\\ =& E(X\big|Y_0,\cdots,Y_n) = X_n.\end{aligned}$$

□

这个例子很“奇特”，以一系列任意随机变量为条件的条件数学期望构成鞅.

例 10 随机 Radon-Nikodym 导数构成的鞅

设 $Z \sim U[0,1]$, 即 Z 是 $[0,1]$ 上均匀分布. 设 f 是 $[0,1]$ 上的有限函数，令 $X_n = 2^n\{f(Y_n+2^{-n}) - f(Y_n)\}$, 而 $Y_n = \sum\limits_{k=0}^{2^n-1} \dfrac{k}{2^n} I_{(\frac{k}{2^n} \leqslant Z < \frac{k+1}{2^n})}$. 则 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是鞅.

证明 首先注意到，在 $Y_0, Y_1, \cdots, Y_n$ 条件下，Z 服从 $[Y_n, Y_n+2^{-n})$ 上的均匀分布 ($Y_n \leqslant Z < Y_n + 2^{-n}$), 且 Y_{n+1} 以相等的概率等于 Y_n 或等于 $Y_n + 2^{-(n+1)}$. 于是有

$$\begin{aligned}E(X_{n+1}\big|Y_0,\cdots,Y_n) &= 2^{n+1}[E(f(Y_{n+1}+2^{-(n+1)}) - f(Y_{n+1})\big|Y_0,Y_1,\cdots,Y_n)]\\ &= 2^{n+1}\left\{\frac{1}{2}[f(Y_n+2^{-(n+1)}) - f(Y_n)]\ +\right.\\ &\qquad \left.\frac{1}{2}[f(Y_n+2^{-n}) - f(Y_n+2^{-(n+1)})]\right\}\\ &= 2^n\{f(Y_n+2^{-n}) - f(Y_n)\} = X_n.\end{aligned}$$

□

注意，$X_n = \dfrac{f(Y_n + 2^{-n}) - f(Y_n)}{2^{-n}}$ 近似等于 f 对 Z 的随机导数.

以上列举了 10 个鞅的例子，许多时候用鞅可以解决原来不易解决的问题. 但关键是如何构造出鞅来.

7.2 上鞅 (下鞅) 及分解定理

定义 7.2.1 设 $\{X_n, n \geqslant 0\}$ 与 $\{Y_n, n \geqslant 0\}$ 是随机过程, 称 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是一个**上鞅**, 如果:

(1) $E(X_n^-) > -\infty$, 其中 $x^- = \min(x, 0)$;

(2) $E(X_{n+1} \big| Y_0, Y_1, \cdots, Y_n) \leqslant X_n$;

(3) X_n 是 $Y_0, Y_1, \cdots, Y_n$ 的函数.

定义 7.2.2 设 $\{X_n, n \geqslant 0\}$ 与 $\{Y_n, n \geqslant 0\}$ 是随机过程, 称 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是一个**下鞅**, 如果:

(1) $E(X_n^+) < \infty$, 其中 $x^+ = \max(x, 0)$;

(2) $E(X_{n+1} \big| Y_0, Y_1, \cdots, Y_n) \geqslant X_n$;

(3) X_n 是 $Y_0, Y_1, \cdots, Y_n$ 的函数.

注意，若 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是上鞅 $\Leftrightarrow$ $\{-X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是下鞅.

上 (下) 鞅可用不公平赌博来解释.

例 1 设 $\{Y_n, n \geqslant 0\}$ 是马尔可夫链，$\boldsymbol{P} = (p_{ij})$, f 是 $\boldsymbol{P}$ 的有界右超正则函数，即

$$\sum_{j \in S} p_{ij} f(j) \leqslant f(i) \text{ 且 } |f(i)| \leqslant M.$$

若令 $X_n = f(Y_n)$, 则 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是上鞅.

证明 按定义验证，显然 (1),(3) 成立，只需证明 (2) 成立即可.

我们有

$$\begin{aligned} E(X_{n+1}\big|Y_0, Y_1, \cdots, Y_n) &= E(f(Y_{n+1})\big|Y_n) \\ &= \sum_{j\in S} f(j) p_{Y_n j} \leqslant f(Y_n) = X_n. \end{aligned}$$

□

为后面讲述方便，先介绍一下 Jensen 不等式.

设 $\phi(x)$ 为凸函数，即 $\forall\, x_1, x_2 \in \mathbb{R}\,, 0 < \alpha < 1$, 有

$$\alpha\phi(x_1) + (1-\alpha)\phi(x_2) \geqslant \phi(\alpha x_1 + (1-\alpha)x_2).$$

将其推广，设 $x_i\,(i = 1, 2, \cdots, m) \in \mathbb{R}\,, 0 \leqslant \alpha_i \leqslant 1, \sum_{i=1}^{m} \alpha_i = 1$, 则

$$\sum_{i=1}^{m} \alpha_i \phi(x_i) \geqslant \phi\Big(\sum_{i=1}^{m} \alpha_i x_i\Big),$$

因此

$$E(\phi(X)) \geqslant \phi(E(X)).$$

故，当 ϕ 是凸函数时，有

$$E[\phi(X)\big|Y_0, Y_1, \cdots, Y_n] \geqslant \phi(E(X\big|Y_0, Y_1, \cdots, Y_n)).$$

引理 7.2.1 如 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是鞅, $\phi(x)$ 是凸函数, 且 $\forall\, n$, $E(\phi(X_n)^+) < \infty$, 则 $\{\phi(X_n), n \geqslant 0\}$ 是关于 $\{Y_n, n \geqslant 0\}$ 的一个下鞅.

推论 1 如 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是鞅，$E(X_n^2) < \infty$, 则 $\{|X_n|, n \geqslant 0\}$ 与 $\{X_n^2, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是下鞅.

上 (下) 鞅有以下基本性质:

(1) 如 $\{X_n, n \geqslant 0\}$ 是关于 $\{Y_n, n \geqslant 0\}$ 的 (上) 鞅，则

$$E(X_{n+k}\big|Y_0, Y_1, \cdots, Y_n)(\leqslant) = X_n, \quad \forall k \geqslant 0. \tag{7.2.1}$$

证明 用数学归纳法证之. 仅证上鞅时的情形，当为鞅时，将 "$\leqslant$" 改成 "$=$" 即可.

当 $k = 1$ 时，由定义， (7.2.1) 式成立.

设 $E(X_{n+k}\big|Y_0, Y_1, \cdots, Y_n) \leqslant X_n$ 成立, 要证明 $E(X_{n+k+1}\big|Y_0, Y_1, \cdots, Y_n) \leqslant X_n$ 也成立. 因为

$$\begin{aligned} E(X_{n+k+1}\big|Y_0, Y_1, \cdots, Y_n) =& E\{E(X_{n+k+1}\big|Y_0, Y_1, \cdots, Y_{n+k})\big|Y_0, Y_1, \cdots, Y_n\} \\ \leqslant& E(X_{n+k}\big|Y_0, Y_1, \cdots, Y_n) \leqslant X_n, \end{aligned}$$

因此对所有 $k \geqslant 0$, (7.2.1) 式成立. □

(2) 若 $\{X_n, n \geqslant 0\}$ 是 (上) 鞅，则对 $0 \leqslant k \leqslant n$, 有

$$E(X_n)(\leqslant) = E(X_k)(\leqslant) = E(X_0).$$

证明 利用 (7.2.1) 式有 $E\{X_n\big|Y_0, \cdots, Y_k\}(\leqslant) = X_k$, 故

$$E(X_n) = E\{E[X_n\big|Y_0, \cdots, Y_k]\}(\leqslant) = E(X_k).$$

类似地可证 $E(X_k)(\leqslant) = E(X_0)$. □

(3) 如 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是 (上) 鞅, g 是关于 $Y_0, Y_1, \cdots, Y_n$ 的 (非负) 函数, 则

$$E\{g(Y_0, Y_1, \cdots, Y_n)X_{n+k}\big|Y_0, Y_1, \cdots, Y_n)\}(\leqslant) = g(Y_0, Y_1, \cdots, Y_n)X_n,\ \forall\, k \geqslant 0.$$

证明 因为 g 是关于 $Y_0, Y_1, \cdots, Y_n$ 的 (非负) 函数，因此

$$\begin{aligned} E(g(Y_0, Y_1, \cdots, Y_n)X_{n+k}\big|Y_0, Y_1, \cdots, Y_n) =& g(Y_0, Y_1, \cdots, Y_n)\times \\ & E(X_{n+k}\big|Y_0, Y_1, \cdots, Y_n) \\ (\leqslant) =& g(Y_0, Y_1, \cdots, Y_n)X_n. \end{aligned}$$

□

上一节讨论了许多鞅的例子，在实际中常常把上鞅和下鞅分解成鞅来处理.

对于上鞅和下鞅，有一分解定理，它是鞅论中的基本定理之一.

定理 7.2.1 对于任意一个 $\{X_n, n \geqslant 1\}$ 关于 $\{Y_n, n \geqslant 1\}$ 的下鞅，必存在过程 $\{M_n, n \geqslant 1\}$ 与 $\{Z_n, n \geqslant 1\}$, 使得

(1) $\{M_n, n \geqslant 1\}$ 关于 $\{Y_n, n \geqslant 1\}$ 是鞅;

(2) Z_n 是 $Y_1, \cdots, Y_{n-1}$ 的函数 $(n \geqslant 2)$, 且 $Z_1 = 0, Z_n \leqslant Z_{n+1}, EZ_n < +\infty$;

(3) $X_n = M_n + Z_n (n \geqslant 1)$.

且上述分解是唯一的.

证明 先证存在性. 令 $Z_1=0, M_0=X_0$, 及

$$
\begin{aligned}
M_n &= X_n - \sum_{k=1}^{n} E(X_k - X_{k-1} \big| Y_1, \cdots, Y_{k-1}), \quad n \geqslant 1,\\
Z_n &= X_n - M_n = \sum_{k=1}^{n} E(X_k - X_{k-1} \big| Y_1, \cdots, Y_{k-1}), \quad n \geqslant 2.
\end{aligned}
$$

因为 $\{X_n, n \geqslant 1\}$ 关于 $\{Y_n, n \geqslant 1\}$ 是下鞅，因此

$$
E(X_k \big| Y_1, \cdots, Y_{k-1}) \geqslant X_{k-1}, \quad E(X_{k-1} \big| Y_1, \cdots, Y_{k-1}) = X_{k-1},
$$

进而有

$$
E(X_k - X_{k-1} \big| Y_1, \cdots, Y_{k-1}) \geqslant 0.
$$

因此 Z_n 非负且单调非降，且 Z_n 是 $Y_1, \cdots, Y_{n-1}$ 的函数. 同时由 Z_n 的定义有

$$
E|Z_n| \leqslant E|X_n| + E|X_1| < +\infty.
$$

另外

$$
\begin{aligned}
&E(M_n \big| Y_1, \cdots, Y_{n-1}) \\
=&E\left\{\left[X_n - \sum_{k=1}^{n} E(X_k - X_{k-1} \big| Y_1, \cdots, Y_{k-1})\right] \Big| Y_1, \cdots, Y_{n-1}\right\} \\
=&E(X_n \big| Y_1, \cdots, Y_{n-1}) - \sum_{k=1}^{n} E\left(E(X_k - X_{k-1} \big| Y_1, \cdots, Y_{k-1}) \Big| Y_1, \cdots, Y_{n-1}\right) \\
=&E(X_n \big| Y_1, \cdots, Y_{n-1}) - \sum_{k=1}^{n} E(X_k - X_{k-1} \big| Y_1, \cdots, Y_{k-1}) \\
=&E(X_n \big| Y_1, \cdots, Y_{n-1}) - \sum_{k=1}^{n-1} E(X_k - X_{k-1} \big| Y_1, \cdots, Y_{k-1}) - \\
&\ E(X_n - X_{n-1} \big| Y_1, \cdots, Y_{n-1}) \\
=&X_{n-1} - \sum_{k=1}^{n-1} E(X_k - X_{k-1} \big| Y_1, \cdots, Y_{k-1}) = M_{n-1},
\end{aligned}
$$

又

$$E|M_n| = E|X_n - Z_n| \leqslant E|X_n| + E|Z_n| < \infty,$$

因此 $\{M_n, n \geqslant 1\}$ 关于 $\{Y_n, n \geqslant 1\}$ 是鞅，且 $X_n = M_n + Z_n$. 故存在性得证.

下面证唯一性. 设另一分解 M_n', Z_n' 满足上面定理要求，即

$$X_n = M_n' + Z_n',\ n \geqslant 1, \quad Z_1' = 0, \quad M_0' = X_0 = 0.$$

则

$$M_n + Z_n = M'_n + Z'_n = X_n.$$

令 $\Delta_n = M_n - M'_n = Z'_n - Z_n$. 因为 $\{M_n, n \geqslant 1\}$ 和 $\{M'_n, n \geqslant 1\}$ 均是关于 $\{Y_n, n \geqslant 1\}$ 的鞅，因此 $\{\Delta_n, n \geqslant 1\}$ 也是关于 $\{Y_n, n \geqslant 1\}$ 的鞅．所以有

$$E(\Delta_n \big| Y_1, \cdots, Y_{n-1}) = \Delta_{n-1}.$$

又因为 Z_n, Z'_n 是关于 $Y_1, \cdots, Y_{n-1}$ 的函数．因此 Δ_n 也是关于 $Y_1, \cdots, Y_{n-1}$ 的函数，于是

$$E(\Delta_n \big| Y_1, \cdots, Y_{n-1}) = \Delta_n,$$

进而

$$\Delta_n = \Delta_{n-1} = \Delta_{n-2} = \cdots = \Delta_1 = Z'_1 - Z_1 = 0,$$

故

$$Z_n = Z'_n, M_n = M'_n. \qquad \square$$

由本定理可知，一个下鞅总可分解为一个鞅与一增过程之和．

推论 若 $\{X_n, n \geqslant 1\}$ 关于 $\{Y_n, n \geqslant 1\}$ 是上鞅，则可分解为 $X_n = M_n - Z_n$ 使得

(1) $\{M_n, n \geqslant 1\}$ 关于 $\{y_n, n \geqslant 1\}$ 是鞅；

(2) Z_n 是 $Y_1, \cdots, Y_{n-1}$ 的函数 $(n \geqslant 2)$, $Z_1 = 0, Z_n \leqslant Z_{n+1}, EZ_n < +\infty$.

且上述分解是唯一的．

例 2 鞅在序贯决策模型中的应用

考虑一个可控的随机动态系统，状态空间有限，记为 $S=\{1,2,\cdots,p\}$. 行动集 $A=\{a,b,\cdots,l\}$ 有限，有时称 A 为**决策空间**. 设每经单位时间 (如每小时，每天，每月等) 观察即时的系统状态 i, 然后从 A 中选取一个行动 a, 有两件事情发生:

(1) 得到一个报酬 (或能量) $r(i,a)$;

(2) 在现时段状态为 i, 采取行动为 a 的条件下，系统下一时刻转移到状态 j 的概率为 $q(j|i,a)$.

现在的问题是：在每一时刻如何选取行动，使前 N 时段的期望总报酬 (总能量) 达到最大？

令 Δ_k 表示 k 时段采取的行动;Y_k 表示 k 时段系统的状态;$h_{n-1}=\{i_0,a_0,i_1,a_1,\cdots,i_{n-1},a_{n-1}\}$ 表示 $n-1$ 时刻及以前系统的状态及采取行动的交互序列，称为 $n-1$ 之前的历史. 设 n 时刻采取的行动 (决策) a_n 依赖于 h_{n-1} 与 i_n, 记为

$$a_n=\pi_n(h_{n-1},i_n)=\pi_n(i_0,a_0,i_1,a_1,\cdots,i_{n-1},a_{n-1},i_n),$$

其中 π_n 称为 n 时刻的决策函数.

一个策略 $\pi=\{\pi_0,\pi_1,\cdots,\pi_{N-1}\}$ 是一个决策函数序列. 若给定一个策略 π 及初始状态 $Y_0=i$, 则直到 $N-1$ 时刻的期望总报酬为

$$V(\pi,i)=E_\pi\left\{\sum_{k=0}^{N-1}r(Y_k,\Delta_k)\big|Y_0=i\right\}.$$

式中 E_π 表示在 π 的条件下求期望. 我们的目的是选取一最优策略 π^*, 使对所有 $i\in S$, 有

$$V(\pi^*,i)=\max_\pi V(\pi,i). \tag{7.2.2}$$

为此，记 $V_N(i)=0,\ \forall\, i\in S$. $V_k(i) i\in S$, 满足

$$V_{k-1}(i)=\max_{a\in A}\left\{r(i,a)+\sum_{j\in S}q(j\big|i,a)V_k(j)\right\},\quad 1\leqslant k\leqslant N. \tag{7.2.3}$$

令

$$X_n=\sum_{k=1}^{n}\left\{V_k(Y_k)-E\left[V_k(Y_k)\big|Y_0,\varDelta_0,Y_1,\varDelta_1,\cdots,Y_{k-1},\varDelta_{k-1}\right]\right\}.$$

由 7.1 节例 3 可知，$\{X_n,n\geqslant 1\}$ 关于 $\{(Y_n,\varDelta_n),n\geqslant 0\}$ 是鞅，于是

$$EX_n=EX_1=0. \tag{7.2.4}$$

由 (7.2.3) 式有

$$V_{k-1}(i)\geqslant r(i,a)+\sum_{j\in S}q(j\big|i,a)V_k(j),\quad \forall\, i\in S,a\in A,$$

故

$$\begin{aligned}V_{k-1}(Y_{k-1})\geqslant& r(Y_{k-1},\varDelta_{k-1})+\sum_{j\in S}q(j\big|Y_{k-1},\varDelta_{k-1})V_k(j)\\ =&r(Y_{k-1},\varDelta_{k-1})+E\{V_k(Y_k)\big|Y_{k-1},\varDelta_{k-1}\}.\end{aligned}$$

因此由马尔可夫性得

$$V_{k-1}(Y_{k-1})\geqslant r(Y_{k-1},\varDelta_{k-1})+E\{V_k(Y_k)\big|Y_0,\varDelta_0,Y_1,\varDelta_1,\cdots,Y_{k-1},\varDelta_{k-1}\}. \tag{7.2.5}$$

于是有

$$\begin{aligned}
0 = EX_N =& E\left\{\sum_{k=1}^{N}\left[V_k(Y_k) - E\left(V_k(Y_k)\big|Y_0, \varDelta_0, \cdots, Y_{k-1}, \varDelta_{k-1}\right)\right]\right\}\\
\geqslant& E\left\{\sum_{k=1}^{N}[V_k(Y_k) + r(Y_{k-1}, \varDelta_{k-1}) - V_{k-1}(Y_{k-1})]\right\}\\
=& E\left\{\sum_{k=0}^{N-1} r(Y_k, \varDelta k) + V_N(Y_N) - V_0(Y_0)\right\}.
\end{aligned}$$

即

$$E(V_0(Y_0)) \geqslant E_\pi\left\{\sum_{k=0}^{N-1} r(Y_k, \varDelta_k)\big|Y_0\right\}.$$

若 $Y_0 = i$, 则上式说明，对任意 π, 有

$$V_0(i) \geqslant E_\pi\left\{\sum_{k=0}^{N-1} r(Y_k, \varDelta_k)\big|Y_0 = i\right\},$$

即 $V_0(i) \geqslant V(\pi, i)$ 对 $\forall\, i \in S$ 及所有 π 均成立.

选取 $a_{k-1}^* \in A$, 使 $a_{k-1}^*(i) = \pi_{k-1}^*(i_0, a_0^*, \cdots, i_{k-1})$ 满足 (7.2.3) 式，即

$$\begin{aligned}
V_{k-1}(i) =& r(i, a_{k-1}^*(i)) + \sum_{j\in S} q(j\big|i, a_{k-1}^*(i))V_k(j)\\
=& \max_a\left\{r(i, a) + \sum_j q(j\big|i, a)V_k(j)\right\}, \quad \forall\, i \in S.
\end{aligned}$$

则 $\pi^* = \{\pi_0^*, \pi_1^*, \cdots, \pi_{N-1}^*\}$ 是最优策略，即

$$V_0(i) = V(\pi^*, i), \quad \forall\, i \in S.$$

换言之，我们是通过使每一步均取最优来达到总体最优的. □

从上面的例子可以得到鞅应用的感性认识. 利用鞅的性质可以解决许多本来不易研究的问题，但关键是怎样构造出一个合适的鞅来. 这通常也不是件容易的事，需要经验，也需要技巧.

7.3 停时与停时定理

本节要研究当 T 是一随机变量时，EX_T 是否等于 EX_0. 为此引出停时的概念. 停时是一个不依赖于 “将来” 的随机时间. 先给出粗略的直观定义.

定义 7.3.1 设取值为非负整数 (包括 $+\infty$) 的随机变量 T, 及随机序列 $\{Y_n, n \geqslant 0\}$. 若对 $n \geqslant 0$, 事件 $\{T = n\}$ 的示性函数 $I_{\{T=n\}}$ 仅是 $Y_0, Y_1, \cdots, Y_n$ 的函数，则称 T 是关于

$\{Y_n, n \geqslant 0\}$ 的**停时**(stopping time)(或称马尔可夫时间).

为今后叙述方便，引入一些记号. 记 $\sigma(X)$ 为由随机变量 X 决定的事件及它们的有限和可列运算的全体构成的事件集，简称为由 X 生成的事件 σ 域，它表示由随机变量 X 可能提供的全部信息. 同样，记

$$\sigma(Y_0, Y_1, \cdots, Y_n) = \sigma(Y_k,$$

$0 \leqslant k \leqslant n), A3FV.N * SIKf; z1dA?PrAPY_0, Y_1, \cdots, Y_n$ 生成的事件 σ 域. 它表示由 $Y_0, Y_1, \cdots, Y_n$ 可能提供的全部信息. 例如，在样本空间 Ω 中定义事件 A 及 B 的示性函数 I_A 及 I_B, 那么

$$\sigma(I_A) = \{\varnothing, A, A^C, \Omega\}, \qquad \sigma(I_B) = \{\varnothing, B, B^C, \Omega\},$$

$$\begin{aligned}\sigma(I_A, I_B) = \{&\varnothing, A, B, A^C, B^C, A-B, B-A, AB, A\bigcup B, (A-B)^C, (B-A)^C,\\ &(AB)^C, (A\bigcup B)^C, (A-B)\bigcup(B-A), [(A-B)\bigcup(B-A)]^C, \Omega\}.\end{aligned}$$

通常简记

$$\mathcal{F}_n = \sigma(Y_k, 0 \leqslant k \leqslant n) = \sigma(Y_0, Y_1, \cdots, Y_n), \quad n \geqslant 0.$$

若 $\forall n \geqslant 0, \{T = n\} \in \mathcal{F}_n$, 则 $\{T = n\}$ 与 $\{T \neq n\}$ 完全取决于过程直到 n 时刻的信息 $(Y_0, Y_1, \cdots, Y_n)$, 而与过程的未来无关.

由定义知, 若 $\forall n \geqslant 0$, 事件 $\{T \leqslant n\}$, $\{T > n\}$, $\{T \geqslant n\}$, $\{T < n\}$ 均只由 $(Y_0, Y_1, \cdots, Y_n)$ 确定，则 T 是一停时.

因此，下面给出停时的确切定义.

定义 7.3.2 设有非负整数的随机变量 T 及随机序列 $\{Y_n, n \geqslant 0\}$, $\mathcal{F}_n = \sigma(Y_k, 0 \leqslant k \leqslant n)$. 若对 $\forall n \geqslant 0$, $\{T = n\} \in \mathcal{F}_n$, 则称 T 是 $\{Y_n, n \geqslant 0\}$ 的停时.

对随机过程 $\{Y_n, n \geqslant 0\}$, 令 $T = \min\{n{:}\, Y_n \in A\}$, 即 T 是首达 A 的时间，则 T 是关于 $\{Y_n, n \geqslant 0\}$ 的停时. 因为 $\{T > n\} = \{\omega{:}\, Y_k \notin A, \forall k \leqslant n\} \in \mathcal{F}_n$, 即事件 $\{T > n\}$ 发生与否完全由 $(Y_0, Y_1, \cdots, Y_n)$ 确定.

例 1 $\{N(t), t \geqslant 0\}$ 是参数为 λ 的时齐泊松过程，$S_0 = 0, S_n$ 为第 n 个事件发生时刻，则 $N(t)$ 关于 $\{S_n, n \geqslant 0\}$ 不是停时，但 $N(t)+1$ 关于 $\{S_n, n \geqslant 0\}$ 是停时.

显然 $T = k$(k 是一常数) 是一个停时.

停时有以下基本特性:

设 T, σ 是关于 $\{Y_n, n \geqslant 0\}$ 的两个停时，则 $T + \sigma, T \wedge \sigma = \min(T, \sigma), T \vee \sigma = \max(T, \sigma)$ 均是停时.

下面介绍有关停时定理 (optional stopping theorem 或 optional sampling theorem).

为此，先介绍以下几个引理.

引理 7.3.1 设 $\{X_n, n \geqslant 0\}$ 是一关于 $\{Y_n, n \geqslant 0\}$ 的 (上) 鞅，T 是一关于 $\{Y_n, n \geqslant 0\}$ 的停时，则 $\forall n \geqslant k$ 有

$$E(X_n\ I_{(T=k)})(\leqslant) = E(X_k\ I_{(T=k)}).$$

证明 注意到 T 是关于 $\{Y_n, n \geqslant 0\}$ 的停时, 所以 $I_{(T=k)}$ 是关于 $Y_0, Y_1, \cdots, Y_k$ 的函数. 因此

$$\begin{aligned} E(X_n\ I_{(T=k)}) =& E\left(E(X_n\ I_{(T=k)}\middle|Y_0, Y_1, \cdots, Y_k)\right) \\ =& E\left(I_{(T=k)}\ E(X_n\middle|Y_0, Y_1, \cdots, Y_k)\right)(\leqslant) = E(I_{(T=k)}\ X_k). \end{aligned}$$ □

引理 7.3.2 如 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是 (上) 鞅，T 关于 $\{Y_n, n \geqslant 0\}$ 是停时，则 $\forall n \geqslant 1$, 有

$$EX_0(\geqslant) = EX_{T\wedge n}(\geqslant) = EX_n.$$

证明 注意到 $I_{\{T<n\}}+I_{\{T\geqslant n\}}=1$, 并利用引理 7.3.1 得

$$\begin{aligned}EX_{T\wedge n}=&E\left\{X_{T\wedge n}\left(\sum_{k=0}^{n-1}I_{\{T=k\}}+I_{\{T\geqslant n\}}\right)\right\}\\=&E\left\{X_{T\wedge n}\ \sum_{k=0}^{n-1}I_{\{T=k\}}\right\}+E\{X_{T\wedge n}\ I_{\{T\geqslant n\}}\}\\=&\sum_{k=0}^{n-1}E\{X_{T\wedge n}\ I_{\{T=k\}}\}+E\{X_{T\wedge n}\ I_{\{T\geqslant n\}}\}\\=&\sum_{k=0}^{n-1}E\{X_k\ I_{\{T=k\}}\}+E\{X_n\ I_{\{T\geqslant n\}}\}\\(\geqslant)=&\sum_{k=0}^{n-1}E\{X_n\ I_{\{T=k\}}\}+E\{X_n\ I_{\{T\geqslant n\}}\}=EX_n.\end{aligned}$$

因此

$$EX_{T\wedge n}(\geqslant)=EX_n.$$

对于鞅，因为 $EX_n=EX_0$, 所以 $EX_{T\wedge n}=EX_0$.

对于上鞅，下面证明 $EX_0\geqslant EX_{T\wedge n}$. 设

$$\widetilde{X}_0=0,\qquad \widetilde{X}_n=\sum_{k=1}^{n}\big[X_k-E(X_k\big|Y_0,Y_1,\cdots,Y_{k-1})\big],\quad n\geqslant 1.$$

根据 7.1 节例 3 可知, $\{\widetilde{X}_n,n\geqslant 0\}$ 关于 $\{Y_n,n\geqslant 0\}$ 是鞅. 由鞅的性质可知

$$E\widetilde{X}_n=E\widetilde{X}_{T\wedge n}=E\widetilde{X}_0=0,$$

因此有

$$\begin{aligned}0 =&E\widetilde{X}_{T\wedge n} = E\left[\sum_{k=1}^{T\wedge n}\big(X_k - E(X_k\big|Y_0, Y_1, \cdots, Y_{k-1})\big)\right]\\ \geqslant&E\left[\sum_{k=1}^{T\wedge n}(X_k - X_{k-1})\right] \qquad\qquad (\text{由上鞅定义得})\\ =&E(X_{T\wedge n} - X_0) = E(X_{T\wedge n}) - E(X_0).\end{aligned}$$

因此

$$EX_{T\wedge n} \leqslant EX_0. \qquad \square$$

引理 7.3.3 设 X 是一随机变量，满足 $E|X| < \infty$. T 是关于 $\{Y_n, n \geqslant 0\}$ 的停时，且 $P(T < \infty) = 1$, 则

$$\lim_{n\to\infty} E(X\ I_{\{T>n\}}) = 0, \quad \lim_{n\to\infty} E(X\ I_{\{T\leqslant n\}}) = EX.$$

证明 因为

$$|X| = |X|I_{\{T\leqslant n\}} + |X|I_{\{T>n\}} \geqslant |X|I_{\{T\leqslant n\}},$$

并且

$$\lim_{n\to\infty} I_{\{T\leqslant n\}} = \lim_{n\to\infty}\sum_{k=1}^{n} I_{\{T=k\}} = \sum_{k=1}^{\infty} I_{\{T=k\}} = 1.$$

因此

$$E|X| \geqslant E(|X|I_{\{T\leqslant n\}}) \xrightarrow{n\to\infty} \sum_{k=0}^{\infty} E\{|X|I_{\{T=k\}}\} = E|X|.$$

于是有

$$\lim_{n\to\infty} E(|X|I_{\{T\leqslant n\}}) = E|X|, \quad \lim_{n\to\infty} E(|X|\ I_{\{T>n\}}) = 0.$$

由上式知

$$\lim_{n\to\infty} E(X\ I_{\{T>n\}}) = 0.$$

又因为

$$|E(X\ I_{\{T\leqslant n\}}) - EX| = |EX\ I_{\{T>n\}}| \leqslant E|X\ I_{\{T>n\}}| = E(|X|I_{\{T>n\}}) \overset{n\to\infty}{\longrightarrow} 0,$$

即

$$\lim_{n\to\infty} E(X\ I_{(T\leqslant n)}) = EX. \qquad \square$$

定理 7.3.1 设 $\{X_n, n\geqslant 0\}$ 是鞅，T 是停时，若 $P(T<\infty)=1$, 且

$$E\Big(\sup_{n\geqslant 0}|X_{T\wedge n}|\Big) < \infty,$$

则

$$EX_T = EX_0.$$

证明 记 $Z = \sup\limits_{n\geqslant 0}|X_{T\wedge n}|$. 因为

$$X_T = \sum_{k=0}^{\infty}(X_k\ I_{\{T=k\}}) = \sum_{k=0}^{\infty}(X_{T\wedge k}\ I_{\{T=k\}}),$$

因此

$$|X_T| = |\sum_{k=0}^{\infty}(X_{T\wedge k}\ I_{\{T=k\}})| \leqslant \sum_{k=0}^{\infty}(|X_{T\wedge k}|\ I_{\{T=k\}})$$
$$\leqslant \sup_{n\geqslant 0}|X_{T\wedge n}|\ \sum_{k=0}^{\infty} I_{\{T=k\}} = \sup_{n\geqslant 0}|X_{T\wedge n}| = Z.$$

所以有

$$E|X_T| \leqslant E(Z) < \infty,$$

即 EX_T 有意义. 又

$$\begin{aligned}|EX_{T\wedge n} - EX_T| =& |E[(X_{T\wedge n} - X_T)\ I_{\{T>n\}}] + E[(X_{T\wedge n} - X_T)\ I_{\{T\leqslant n\}}]| \\ =& |E(X_{T\wedge n} - X_T)\ I_{\{T>n\}}| \leqslant E|(X_{T\wedge n} - X_T)\ I_{\{T>n\}}| \\ \leqslant& E|(X_{T\wedge n}\ I_{\{T>n\}}| + E|X_T\ I_{(T>n)}| \leqslant 2E(Z\ I_{\{T>n\}}),\end{aligned}$$

由引理 7.3.3 知 $\lim\limits_{n\to\infty} E(Z\cdot I_{\{T>n\}}) = 0$, 因此

$$\lim_{n\to\infty} EX_{T\wedge n} = EX_T.$$

又由引理 7.3.2 得 $EX_{T\wedge n} = EX_0$, 所以

$$EX_T = \lim_{n\to\infty} EX_{T\wedge n} = \lim_{n\to\infty} EX_0 = EX_0. \qquad \square$$

推论 设 $\{X_n, n\geqslant 0\}$ 关于 $\{Y_n, n\geqslant 0\}$ 是鞅，T 是停时，且 $ET<\infty$. 若存在一常数 $b<\infty$, 满足对 $\forall n<T$, 有

$$E(|X_{n+1} - X_n|\big|Y_0, Y_1, \cdots, Y_n) \leqslant b,$$

则

$$EX_0 = EX_T.$$

证明 令

$$\begin{aligned} Z_0 =& |X_0|, \\ Z_n =& |X_n - X_{n-1}|, \quad n \geqslant 1, \\ W =& Z_0 + Z_1 + \cdots + Z_T. \end{aligned}$$

则

$$\begin{aligned} W =& |X_0| + |X_1 - X_0| + \cdots + |X_T - X_{T-1}|, \\ EW =& \sum_{n=0}^{\infty} \sum_{k=0}^{n} E(Z_k I_{\{T=n\}}) = \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} E(Z_k I_{\{T=n\}}) = \sum_{k=0}^{\infty} E(Z_k \ I_{\{T \geqslant k\}}). \end{aligned}$$

因为 $I_{\{T \geqslant k\}} = 1 - I_{\{T \leqslant k-1\}}$ 仅仅是 $Y_0, Y_1, \cdots, Y_{k-1}$ 的函数，又由已知条件知，对 $k \leqslant T$ 有 $E(Z_k \big| Y_0, \cdots, Y_{k-1}) \leqslant b$, 因此

$$\begin{aligned} \sum_{k=0}^{\infty} E(Z_k \ I_{\{T \geqslant k\}}) =& \sum_{k=0}^{\infty} E\{E(Z_k \ I_{\{T \geqslant k\}} \big| Y_0, Y_1, \cdots, Y_{k-1})\} \\ =& \sum_{k=0}^{\infty} E\{I_{\{T \geqslant k\}} \ E(Z_k \big| Y_0, Y_1, \cdots, Y_{k-1})\} \\ \leqslant& b \sum_{k=0}^{\infty} P(T \geqslant k) = b(1 + ET) \qquad \left(\text{利用} \sum_{k=1}^{\infty} P(T \geqslant k) = ET\right) \\ <& \infty, \end{aligned}$$

即 $EW < \infty$. 因为 $|X_T| \leqslant W$, 因此 $|X_{T\wedge n}| \leqslant W$, $\forall n \geqslant 0$, 即

$$\sup_{n\geqslant 0} |X_{T\wedge n}| \leqslant W,$$

所以有

$$E(\sup_{n\geqslant 0} |X_{T\wedge n}|) \leqslant EW < \infty.$$

又因为 $ET < \infty$, 因此有

$$P(T < \infty) = 1.$$

利用定理 7.3.1, 即得

$$EX_T = EX_0. \qquad \square$$

定理 7.3.2 (停时定理) 设 $\{X_n, n \geqslant 0\}$ 是鞅, T 是停时, 若:

(1) $P(T < \infty) = 1$;

(2) $E|X_T| < \infty$;

(3) $\lim\limits_{n\to\infty} E|X_n\ I_{\{T>n\}}| = 0$.

则

$$EX_T = EX_0.$$

证明 由

$$X_T = X_T I_{\{T\leqslant n\}} + X_T I_{\{T>n\}},$$

及

$$\begin{aligned}X_T I_{\{T\leqslant n\}} =& X_{T\wedge n} I_{\{T\leqslant n\}} = X_{T\wedge n}(1 - I_{\{T>n\}})\\ =& X_{T\wedge n} - X_{T\wedge n} I_{\{T>n\}} = X_{T\wedge n} - X_n I_{\{T>n\}},\end{aligned}$$

得

$$X_T = X_{T\wedge n} - X_n\ I_{\{T>n\}} + X_T\ I_{(T>n)}.$$

因此

$$EX_T = EX_{T\wedge n} - E(X_n\ I_{\{T>n\}}) + E(X_T\ I_{\{T>n\}}).$$

由已知 $\lim\limits_{n\to\infty} E|X_n\ I_{\{T>n\}}| = 0$, 则

$$\lim_{n\to\infty} E(X_n\ I_{\{T>n\}}) = 0.$$

由引理 7.3.3 得

$$\lim_{n\to\infty} E(X_T\ I_{\{T>n\}}) = 0,$$

因此

$$\lim_{n\to\infty} EX_{T\wedge n} = EX_T.$$

而由引理 7.3.2 知 $EX_{T\wedge n} = EX_0$, 故

$$EX_0 = EX_T. \qquad \square$$

这个基本定理有以下简单推论.

推论 1 设 $\{X_n, n \geqslant 0\}$ 是鞅, T 是停时, 若:

(1) $P(T < \infty) = 1$;

(2) 对某个 $k < \infty, \forall n \geqslant 0, E(X_{T\wedge n}^2) \leqslant k$.

则

$$EX_0 = EX_T.$$

证明 显然 $X_{T\wedge n}^2 \geqslant 0$. 由 (2) 知

$$E(X_{T\wedge n}^2 I_{\{T\leqslant n\}}) \leqslant E(X_{T\wedge n}^2) \leqslant k.$$

而

$$\begin{aligned} E(X_{T\wedge n}^2 I_{\{T\leqslant n\}}) &= \sum_{k=0}^{n} E[X_T^2 \big| T = k]\ P(T = k) \\ &\xrightarrow{n\to\infty} \sum_{k=0}^{\infty} E(X_T^2 \big| T = k)\ P(T = k) = EX_T^2, \end{aligned}$$

因此

$$EX_T^2 \leqslant k < \infty.$$

由 Schwartz 不等式可得

$$E|X_T| = E|1 \cdot X_T| \leqslant [E(X_T^2)]^{\frac{1}{2}} < \infty,$$

及

$$(E(X_n I_{\{T>n\}}))^2 = [E(X_{T\wedge n}\ I_{\{T>n\}})]^2 \leqslant E(X_{T\wedge n}^2)\ E(I_{\{T>n\}}^2),$$

即

$$(E(X_n\ I_{\{T>n\}}))^2 \leqslant k\ P\{T>n\} \xrightarrow{n\to\infty} 0.$$

因此

$$\lim_{n\to\infty} E[X_n\ I_{\{T>n\}}] = 0.$$

利用定理 7.3.2 得

$$EX_0 = EX_T. \qquad \square$$

推论 2 设 $Y_0=0, \{Y_k, k\geqslant 1\}$ 独立同分布，$EY_k=\mu, DY_k=\sigma^2<\infty, S_0=0, S_n=\sum\limits_{k=1}^{n} Y_k, X_n=S_n-n\mu$. 若 T 为停时，$ET<\infty$, 则 $E|X_T|<\infty$, 且

$$EX_T = ES_T - \mu ET = 0.$$

证明 因为 $ET=\sum\limits_{k=0}^{\infty} kP(T=k)<+\infty$, 从而余项

$$\sum_{k=n}^{\infty} kP(T=k)\to 0 \quad (n\to\infty),$$

又

$$\sum_{k=n}^{\infty} kP(T=k) \geqslant \sum_{k=n}^{\infty} nP(T=k) = nP(T\geqslant n)\to 0,$$

故 $nP(T \geqslant n) \to 0\ (n \to \infty)$. 因此

$$P(T \geqslant n) \to 0 \quad (n \to \infty).$$

从而

$$P(T < \infty) = 1.$$

$$E|X_T| = E|S_T - T\mu| \leqslant E\Big(\sum_{k=1}^{T} |Y_k - \mu|\Big).$$

因为 $\{Y_k\}$ 独立同分布，所以 $\{|Y_k - \mu|\}$ 独立同分布，于是

$$E\Big(\sum_{k=1}^{T} |Y_k - \mu|\Big) = ET\ E|Y_k - \mu| < \infty,$$

所以

$$E|X_T| < +\infty.$$

由 Schwartz 不等式得

$$[E(X_n I_{\{T>n\}})]^2 \leqslant EX_n^2\ E(I_{\{T>n\}}) \leqslant n\sigma^2 P(T \geqslant n) = \sigma^2 (n\ P(T \geqslant n)).$$

由前面的证明过程的中间结论 $nP(T \geqslant n) \to 0\ (n \to \infty)$ 知

$$\lim_{n\to\infty} E(X_n\ I_{\{T>n\}}) = 0.$$

利用定理 7.3.2 得

$$0 = EX_0 = EX_T = ES_T - \mu ET. \qquad \square$$

如何在实际中应用停时定理与推论呢？ 这里介绍一个应用停时定理的例子.

例 2 随机游动

令 $Y_0 = 0, \{Y_k, k \geqslant 1\}$ 独立同分布，$P(Y_k = 1) = p \geqslant 0, P(Y_k = -1) = q = 1 - p \geqslant 0$. 令 $X_0 = 0, X_n = \sum_{k=1}^{n} Y_k$, 记

$$\begin{aligned}
T_{0j} &= \min\{n\colon X_0 = 0, X_n = j\},\\
T &= \min\{n\colon X_n = a \text{ 或} X_n = b\}, b > 0 \text{ 为正整数}, a < 0 \text{ 为负整数},\\
{}_bT_a &= \min\{n\colon X_0 = 0, X_l \neq b, 1 \leqslant l \leqslant n-1, X_n = a\},\\
V_a &= P({}_bT_a < \infty \big| X_0 = 0),\ V_a\text{表示从 0 出发先到达 } a \text{ 的概率}.
\end{aligned}$$

则 $P({}_aT_b < \infty \big| X_0 = 0) = 1 - V_a$.

若以赌博 (或投资) 为背景．设甲乙两人赌博，$|a|$ 表示甲原有的资金，Y_n 表示甲第 n 次得到的钱，b 表示乙原有的资金，那么 V_a 表示甲先输光的概率.

分两种情况讨论

(1) 当 $p = q = 1/2$ 时：易证 $P(T_{01} < \infty \big| X_0 = 0) = 1$(见第 3 章练习题)．由 $X_{T_{01}}$ 定义知，$X_{T_{01}} = 1$, 故 $EX_{T_{01}} = 1$．而 $EX_0 = 0$, 所以 $EX_{T_{01}} \neq EX_0$．易知 $\{X_n, n \geqslant 1\}$ 关于 $\{Y_n, n \geqslant 1\}$ 是鞅，T_{01} 关于 $\{Y_n, n \geqslant 1\}$ 是停时．由定理 7.3.1 的推论知，$ET_{01} < \infty$ 不成立

(因为若 $ET_{01} < \infty$, 则 $EX_T = EX_0 = 0$), 故

$$ET_{01} = +\infty.$$

易证 $P(T < \infty \big| X_0 = 0) = 1$, 而 $|X_{T\wedge n}| \leqslant \max(|a|, b),\ \forall n \geqslant 0$, 因此有

$$E\left(\sup_{n\geqslant 0} |X_{T\wedge n}|\right) < \infty.$$

由定理 7.3.1 得

$$EX_T = EX_0 = 0.$$

但

$$EX_T = V_a a + (1 - V_a) b = 0,$$

解之，得

$$V_a = \frac{b}{|a| + b},$$

这就是甲先输光的概率. 同理

$$V_b = 1 - V_a = \frac{|a|}{|a| + b}.$$

如何求 $E(T|X_0 = 0)$? 即任何一方输光的平均时间是多少呢？这需要构造一个合适的鞅来解决.

设 $Z_n = X_n^2 - n$, 先验证它关于 $\{Y_n, n \geqslant 0\}$ 是鞅. 因为

$$
\begin{aligned}
E|Z_n| =& E|X_n^2 - n| \leqslant EX_n^2 + n < \infty, \\
E(Z_{n+1} \big| Y_0, Y_1, \cdots, Y_n) =& E(X_{n+1}^2 - n - 1 \big| Y_0, Y_1, \cdots, Y_n) \\
=& E((X_n + Y_{n+1})^2 \big| Y_0, Y_1, \cdots, Y_n) - (n+1) \\
=& E(X_n^2 \big| Y_0, \cdots, Y_n) + 2E(X_n Y_{n+1} \big| Y_0, \cdots, Y_n) + \\
& E(Y_{n+1}^2 \big| Y_0, \cdots, Y_n) - (n+1) \\
=& X_n^2 + 2X_n \cdot 0 + 1 - (n+1) = X_n^2 - n = Z_n,
\end{aligned}
$$

所以 $\{Z_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是鞅.

由于 T 为停时，且 $ET < \infty$(由第 3 章练习题), 另外

$$
\begin{aligned}
E(|Z_{n+1} - Z_n| \big| Y_0, \cdots, Y_n) =& E(|X_{n+1}^2 - (n+1) - X_n^2 + n| \big| Y_0, \cdots, Y_n) \\
=& E(|2X_n \ Y_{n+1} + Y_{n+1}^2 - 1| \big| Y_0, \cdots, Y_n) \\
\leqslant& 2E(|X_n \ Y_{n+1}| \big| Y_0, \cdots, Y_n) + E(Y_{n+1}^2 \big| Y_0, \cdots, Y_n) + 1 \\
=& 2|X_n| \ E|Y_{n+1}| + 1 + 1 \leqslant 2(\max(|a|, b)) + 2 < \infty,
\end{aligned}
$$

由定理 7.3.1 的推论知

$$EZ_T = EZ_0 = 0.$$

而由于

$$EZ_T = E(X_T^2 - T) = EX_T^2 - ET,$$

因此

$$ET = EX_T^2 = a^2V_a + b^2(1-V_a) = a^2\ \frac{b}{|a|+b} + b^2\ \frac{|a|}{|a|+b} = |a|b.$$

(2) 当 $p-q=\mu>0$ 时：这是不公平赌博，甲方赢的概率大，这时的 V_a 是多少呢？为此仍需要构造鞅. 注意到此时 $EY_n = p-q=\mu$. 令 $U_n = X_n - n\mu$.

因为

$$\begin{aligned}
E|U_n| = E|X_n - n\mu| \leqslant & E|X_n| + n\mu \leqslant n + n\mu < +\infty, \\
E(U_{n+1}\big|Y_0, Y_1, \cdots, Y_n) = & E(X_{n+1} - (n+1)\mu\big|Y_0, Y_1, \cdots, Y_n) \\
= & E(X_n + Y_{n+1}\big|Y_0, Y_1, \cdots, Y_n) - (n+1)\mu \\
= & X_n + \mu - (n+1)\mu = X_n - n\mu = U_n,
\end{aligned}$$

因此 $\{U_n, n\geqslant 0\}$ 关于 $\{Y_n, n\geqslant 0\}$ 是鞅.

另外，由于

$$\begin{aligned}
E(|U_{n+1} - U_n|\big|Y_0, \cdots, Y_n) = & E(|Y_{n+1} - \mu|\big|Y_0, \cdots, Y_n) \\
= & E|Y_{n+1} - \mu| = 4pq < \infty,
\end{aligned}$$

由第 3 章练习题 3.19 第 3 小题知，有 $ET<\infty$, 因此由定理 7.3.2 的推论知，在 $ET<\infty$ 的条件下，$EU_T = EU_0$, 故

$$EX_T = \mu ET.$$

令 $V_n = \left(\dfrac{q}{p}\right)^{X_n}$, 由于 $p>q$, 故 $0<\dfrac{q}{p}<1$. 下面验证 $\{V_n, n\geqslant 0\}$ 是鞅.

因为

$$\begin{aligned}E|V_n|=E\Big|\Big(\frac{q}{p}\Big)^{X_n}\Big| =&E\Big\{\Big(\frac{q}{p}\Big)^{X_n}\Big\}=\prod_{k=1}^{n}E\Big\{\Big(\frac{q}{p}\Big)^{Y_k}\Big\}=1<\infty,\\ E(V_{n+1}\big|Y_0,\cdots,Y_n)=&E\Big\{\Big(\frac{q}{p}\Big)^{X_{n+1}}\big|Y_0,\cdots,Y_n\Big\}\\ =&E\Big\{\Big(\frac{q}{p}\Big)^{X_n}\ \Big(\frac{q}{p}\Big)^{Y_{n+1}}\big|Y_0,\cdots,Y_n\Big\}\\ =&\Big(\frac{q}{p}\Big)^{X_n}\ E\Big\{\Big(\frac{q}{p}\Big)^{Y_{n+1}}\Big\}\\ =&\Big(\frac{q}{p}\Big)^{X_n}\ \Big[\Big(\frac{q}{p}\Big)\ p+\Big(\frac{q}{p}\Big)^{-1}\ q\Big]\\ =&\Big(\frac{q}{p}\Big)^{X_n}=V_n,\end{aligned}$$

因此 $\{V_n,n\geqslant 0\}$ 关于 $\{Y_n,n\geqslant 0\}$ 是鞅.

由 $a\leqslant X_n\leqslant b,0<\dfrac{q}{p}<1$ 得

$$\Big(\frac{q}{p}\Big)^{b}\leqslant V_n\leqslant\Big(\frac{q}{p}\Big)^{a}.$$

因此

$$E|V_T|=E(V_T)\leqslant\Big(\frac{q}{p}\Big)^{a}<+\infty.$$

故

$$0\leqslant\lim_{n\to\infty}E(V_nI_{\{T>n\}})\leqslant\lim_{n\to\infty}\Big(\frac{q}{p}\Big)^{a}E\{I_{\{T>n\}}\}=\Big(\frac{q}{p}\Big)^{a}\lim_{n\to\infty}P\{T>n\}=0.$$

此处 $\lim\limits_{n\to\infty} P\{T>n\}=0$ 是因为当 $p-q=\mu>0$ 时，$[a,b]$ 为马尔可夫链 $\{X_n, n\geqslant 0\}$ 的非常返状态集．$(X_n=i)\in[a,b]$ 至多有限次，故有 $P(T<\infty)=1$，即 $\lim\limits_{n\to\infty} P\{T>n\}=0$. 由此得

$$\lim_{n\to\infty} E(V_n\ I_{\{T>n\}})=0.$$

于是由定理 7.3.2 得

$$EV_0=EV_T.$$

另外，由于 $EV_0=E\Big(\dfrac{q}{p}\Big)^{X_0}=1$, 因此

$$EV_T=1.$$

同样地记 $V_a=P({}_bT_a<\infty\big|X_0=0)$, 有

$$EV_T=V_a\Big(\frac{q}{p}\Big)^a+(1-V_a)\Big(\frac{q}{p}\Big)^b=1.$$

解之，得

$$V_a=\frac{1-\Big(\dfrac{q}{p}\Big)^b}{\Big(\dfrac{q}{p}\Big)^a-\Big(\dfrac{q}{p}\Big)^b}<\frac{b}{|a|+b}.$$

故从 0 状态出发，当 $p>q$ 时先到达 a 的概率小于 $p=q$ 时先到达 a 的概率.

再由 $EX_T=\mu ET$ 得

$$aV_a+b(1-V_a)=(p-q)ET,$$

故

$$ET = \frac{b}{p-q} - \frac{(b-a)\left[1-\left(\dfrac{q}{p}\right)^b\right]}{(p-q)\left[\left(\dfrac{q}{p}\right)^a - \left(\dfrac{q}{p}\right)^b\right]}. \qquad \square$$

7.4 鞅收敛定理

设 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是 (上、下) 鞅. 研究在各种意义下，$\lim\limits_{n\to\infty} X_n$ 是否存在的问题，即鞅收敛的问题. 为此，介绍鞅收敛定理，该定理是利用鞅解决许多不同领域问题的重要工具之一.

在叙述鞅收敛定理之前, 先介绍一个重要引理, 即上穿不等式. 为此, 先考虑实数列 $\{a_n\}$ 的收敛问题. $\{a_n\}$ 没有极限时的充要条件是 $\varliminf\limits_{n\to\infty} a_n < \varlimsup\limits_{n\to\infty} a_n$, 即必存在两个实数 a, b, 使

$$\varliminf_{n\to\infty} a_n \leqslant a < b \leqslant \varlimsup_{n\to\infty} a_n.$$

用图形来直观解释 (如图 7.1), 就是 $\{a_n\}$ 无穷多次到达 a 之下，b 之上，即 $\{a_n\}$ 上穿 (a, b) 无穷多次. 于是研究 $\{X_n\}$ 的收敛性就要研究穿越 (a, b) 的次数问题. 设 $\{X_n\}$ 是随机序列，令 $V^{(n)}(a, b)(\omega)$ 是 $(X_0(\omega), X_1(\omega), \cdots, X_n(\omega))$ 向上穿越 (a, b) 的次数. 令 $\alpha_0 = 0$, 记 α_1 为 $\{X_n\}$ 首次到达 $(-\infty, a]$ 的时间，α_2 是 α_1 之后首次到达 $[b, +\infty)$ 的时间，即

$$\begin{aligned}
\alpha_1 &= \min\{n\colon n \geqslant 0, X_n \leqslant a\}, \\
\alpha_2 &= \min\{n\colon n > \alpha_1, X_n \geqslant b\},
\end{aligned}$$

依次类推，归纳定义为

$$\alpha_{2k-1} = \min\{n\colon n \geqslant \alpha_{2k-2}, X_n \leqslant a\},$$
$$\alpha_{2k} = \min\{n\colon n > \alpha_{2k-1}, X_n \geqslant b\},$$

则

$$V^{(n)}(a,b) = \max\{k\colon k \geqslant 1, \alpha_{2k} \leqslant n\} = \sum_{k=1}^{[\frac{n}{2}]} I_{(\alpha_{2k} \leqslant n)}. \tag{7.4.1}$$

记号 $[x]$ 表示不超过 x 的整数，显然 $V^{(n)}(a,b)$ 的可能取值为 $0,1,2,\cdots,\left[\dfrac{n}{2}\right]$. 以下有名的上穿不等式是由 J.Doob 给出的.

引理 7.4.1 (上穿不等式) 设 $\{X_n, n \geqslant 0\}$ 关于 $\{Y_n, n \geqslant 0\}$ 是下鞅，$V^{(n)}(a,b)$ 表示 $\{X_k, 0 \leqslant k \leqslant n\}$ 上穿区间 (a,b) 的次数，$a < b$. 则

$$E(V^{(n)}(a,b)) \leqslant \frac{E(X_n - a)^+ - E(X_0 - a)^+}{b-a} \leqslant \frac{EX_n^+ + |a|}{b-a}, \tag{7.4.2}$$

这里记 $a^+ = \max(a,0) = a \vee 0$.

证明 因为 $\{X_n\}$ 关于 $\{Y_n\}$ 是下鞅，因此 $\{(X_n - a)^+\}$ 关于 $\{Y_n\}$ 也是下鞅. 这是由于

$$\begin{aligned}
E|(X_n - a)^+| &\leqslant E|X_n| + |a| < \infty,\\
E\{(X_{n+1} - a)^+ \big| Y_0, Y_1, \cdots, Y_n\} &= E(X_{n+1} \vee a - a \big| Y_0, Y_1, \cdots, Y_n)\\
&= E(X_{n+1} \vee a \big| Y_0, Y_1, \cdots, Y_n) - a\\
&\geqslant E(X_{n+1} \big| Y_0, Y_1, \cdots, Y_n) \vee a - a\\
&\geqslant X_n \vee a - a = (X_n - a)^+.
\end{aligned}$$

另外 $\{(X_k-a)^+, 0\leqslant k\leqslant n\}$ 上穿过区间 $(0,b-a)$ 的次数也为 $V^{(n)}(a,b)$, 故以下只要证明 $\{\widetilde{X}_k=(X_k-a)^+, 0\leqslant k\leqslant n\}$ 上穿区间 $(0,b-a)$ 的次数 $V^{(n)}(a,b)$ 满足

$$(b-a)E\{V^{(n)}(a,b)\}\leqslant E\widetilde{X}_n-E\widetilde{X}_0.$$

为此，引进随机序列 $\{\varepsilon_i, i\geqslant 1\}$, 满足

$$\{\varepsilon_i=1\}=\bigcup_{k=1}^{\infty}(\alpha_{2k-1}<i\leqslant\alpha_{2k}),\quad \{\varepsilon_i=0\}=\bigcup_{k=1}^{\infty}(\alpha_{2k}<i\leqslant\alpha_{2k+1}).$$

于是

$$\{\varepsilon_i=1\}=\bigcup_{k=1}^{\infty}(\alpha_{2k-1}<i\leqslant\alpha_{2k})=\bigcup_{k=1}^{\infty}\{(\alpha_{2k-1}<i)-(\alpha_{2k}<i)\}=\bigcup_{k=1}^{\infty}\{(\alpha_{2k-1}\leqslant i-1)-(\alpha_{2k}\leqslant i-1)\}\in\sigma(Y_l, 0\leqslant l\leqslant i-1),$$

即 ε_i 是 $Y_0,Y_1,\cdots,Y_{i-1}$ 的函数. $\{\varepsilon_i=1\}$ 发生，则 i 之前的最大的 α_{2k-1} 是一个下穿，即 $\widetilde{X}_{\alpha_{2k-1}}\leqslant 0$; 而 i 以后最小的 α_{2k} 是一个上穿，即 $\widetilde{X}_{\alpha_{2k}}\geqslant(b-a)$. 故

$$(b-a)V^{(n)}(a,b)\leqslant\sum_{k=1}^{V^{(n)}(a,b)}(\widetilde{X}_{\alpha_{2k}}-\widetilde{X}_{\alpha_{2k-1}})=\sum_{i=1}^{n}(\widetilde{X}_i-\widetilde{X}_{i-1})\,\varepsilon_i,$$

由上可得

$$
\begin{aligned}
(b-a)E(V^{(n)}(a,b)) \leqslant & E\Big[\sum_{i=1}^{n}(\widetilde{X}_i-\widetilde{X}_{i-1})\ \varepsilon_i\Big] \\
= & \sum_{i=1}^{n}E[(\widetilde{X}_i-\widetilde{X}_{i-1})\ \varepsilon_i] \\
= & \sum_{i=1}^{n}E\{E[(\widetilde{X}_i-\widetilde{X}_{i-1})\varepsilon_i|Y_0,Y_1,\cdots,Y_{i-1}]\} \\
= & \sum_{i=1}^{n}E\{\varepsilon_i[E(\widetilde{X}_i|Y_0,Y_1,\cdots,Y_{i-1})-\widetilde{X}_{i-1}]\} \\
& \qquad\qquad (\text{由} E(\widetilde{X}_i|Y_0,Y_1,\cdots,Y_{i-1})-\widetilde{X}_{i-1}\geqslant 0) \\
\leqslant & \sum_{i=1}^{n}E[E(\widetilde{X}_i|Y_0,Y_1,\cdots,Y_{i-1})-\widetilde{X}_{i-1}] \\
= & \sum_{i=1}^{n}(E\widetilde{X}_i-E\widetilde{X}_{i-1}) \\
= & E\widetilde{X}_n-E\widetilde{X}_0,
\end{aligned}
$$

得

$$
(b-a)E\{V^{(n)}(a,b)\}\leqslant E\widetilde{X}_n-E\widetilde{X}_0=E(X_n-a)^+-E(X_0-a)^+,
$$

于是有

$$
E\{V^{(n)}(a,b)\}\leqslant \frac{E(\widetilde{X}_n-\widetilde{X}_0)}{b-a}=\frac{E(X_n-a)^+-E(X_0-a)^+}{b-a}\leqslant \frac{EX_n^+ + |a|}{b-a}. \quad \square
$$

推论 $\{X_n, n \geqslant 0\}$ 是关于 $\{Y_n, n \geqslant 0\}$ 的上鞅, $\overline{V}^{(n)}(a,b)$ 是 X_n 下穿 (a,b) 的次数, 则

$$E(\overline{V}^{(n)}(a,b)) \leqslant \frac{1}{b-a}[E(b \wedge X_0) - E(X_n \wedge b)]. \tag{7.4.3}$$

如果 $X_n \geqslant 0, b > a \geqslant 0$, 则

$$E((\overline{V}^{(n)}(a,b)) \leqslant \frac{b}{b-a}.$$

下面证明在什么情况下, 一个鞅 $\{X_n\}$ 在 $n \to \infty$ 时将趋向一个期望有限的随机变量.

定理 7.4.1 设 $\{X_n\}$ 是一个下鞅. $\sup\limits_n E|X_n| < \infty$, 则存在一随机变量 X_∞, 使 $\{X_n\}$ 以概率 1 收敛于 X_∞, 即

$$P\Big(\lim_{n\to\infty} X_n = X_\infty\Big) = 1, \tag{7.4.4}$$

且 $E|X_\infty| < \infty$.

证明 首先, 由于

$$EX_n^+ \leqslant E|X_n| \leqslant 2EX_n^+ - EX_n,$$

故

$$\sup_n E|X_n| < \infty \quad \text{当且仅当} \quad \sup_n EX_n^+ < \infty.$$

另一方面, 当 $n \to \infty$ 时, $V^{(n)}(a,b) \to V(a,b) = \{X_n \text{ 上穿}(a,b) \text{ 的次数}\}$. 故

$$\begin{aligned} E(V(a,b)) =& E\big\{\lim_n V^{(n)}(a,b)\big\} = \lim_{n\to\infty} E(V^{(n)}(a,b)) \\ \leqslant& \lim_n \frac{EX_n^+ + |a|}{b-a} \leqslant \frac{\sup\limits_n EX_n^+ + |a|}{b-a} < \infty. \end{aligned}$$

因此

$$P(V(a,b)<\infty)=1.$$

从而

$$P(V(a,b)=+\infty)=0.$$

于是

$$\begin{aligned}
&P\{\omega\colon n\to\infty \text{ 时} X_n(\omega) \text{ 无极限}\}\\
=&P\left(\bigcup_{a<b \text{ 且为有理数}}\left\{\omega\colon \varliminf_{n\to\infty} X_n(\omega)\leqslant a<b<\varlimsup_{n\to\infty} X_n(\omega)\right\}\right)\\
=&P\left(\bigcup_{a<b \text{ 且为有理数}}\{\omega\colon V(a,b)=+\infty\}\right)=0.
\end{aligned}$$

因此

$$P(\omega\colon n\to\infty \text{ 时} X_n(\omega) \text{ 有极限})=1.$$

设 $\lim\limits_{n\to\infty} X_n(\omega)=X_\infty(\omega)$, 则

$$P\left\{\omega\colon \lim_{n\to\infty} X_n(\omega)=X_\infty(\omega)\right\}=1.$$

另外，由 Fatou 引理

$$E|X_\infty|=E(\lim_{n\to\infty}|X_n|)\leqslant \varliminf_{n\to\infty} E|X_n|\leqslant \sup_n E|X_n|<\infty,$$

即

$$E|X_\infty| < \infty. \qquad \square$$

有关鞅的收敛定理内容极其丰富. 下面再介绍一个有名的最大值不等式与另一个收敛定理.

最大值不等式 (maximal inequality).

当 $Y_0, Y_1, \cdots, Y_n, \cdots$ 独立同分布，且当 $EY_i = 0, EY_i^2 = \sigma^2 (i \geqslant 0), X_0 = 0, X_n = \sum_{i=1}^{n} Y_i$ 时，根据切比雪夫不等式，对任意 $\varepsilon > 0$, 有

$$\varepsilon^2 P(|X_n| > \varepsilon) \leqslant n\sigma^2.$$

根据柯尔莫哥洛夫 (Kolmogorov) 不等式, 有

$$\varepsilon^2 P\Big(\max_{0 \leqslant k \leqslant n} |X_k| > \varepsilon\Big) \leqslant n\sigma^2. \tag{7.4.5}$$

这些不等式应用到鞅中将非常简单而有效.

引理 7.4.2 设 $\{X_n\}$ 是下鞅，且 $\forall n \geqslant 0$ 有 $X_n \geqslant 0$, 则对任何 $\lambda > 0$, 有

$$\lambda P\Big(\max_{0 \leqslant k \leqslant n} X_k > \lambda\Big) \leqslant EX_n. \tag{7.4.6}$$

证明 令 $T = \min\{k\colon k \geqslant 0, X_k > \lambda\}$, 则 $X_T > \lambda$. 利用引理 7.3.2 得

$$EX_n \geqslant EX_{T\wedge n} \geqslant E\Big(X_{T\wedge n}\, I_{(\max\limits_{0 \leqslant k \leqslant n} X_k > \lambda)}\Big).$$

注意到当 $\left\{\max\limits_{0\leqslant k\leqslant n} X_k > \lambda\right\}$ 事件发生时，$T\leqslant n$, 于是有 $T\wedge n=T$, 故

$$\begin{aligned}EX_n \geqslant & E\Big(X_T\ I_{(\max\limits_{0\leqslant k\leqslant n} X_k>\lambda)}\Big)\\ \geqslant & \lambda E\Big(I_{(\max\limits_{0\leqslant k\leqslant n} X_k>\lambda)}\Big)\\ = & \lambda P\Big(\max_{0\leqslant k\leqslant n} X_k>\lambda\Big).\end{aligned}$$

□

推论　设 $\{X_n\}$ 是鞅，则对任意 $\lambda>0$, 有

$$\lambda P\Big(\max_{0\leqslant k\leqslant n}|X_k|>\lambda\Big)\leqslant E|X_n|.$$

证明　只要证明 $|X_n|\geqslant 0$ 是个下鞅，利用引理 7.4.2 即可证明．因为 $\{X_n\}$ 是鞅，因此

$$E\|X_n\|=E|X_n|<\infty,$$

$$E(|X_{n+1}|\big|Y_0,\cdots,Y_n)\geqslant|E(X_{n+1}\big|Y_0,\cdots,Y_n)|=|X_n|.$$

故 $\{|X_n|\}$ 是下鞅．由引理 7.4.2 得

$$\lambda P\Big(\max_{0\leqslant k\leqslant n}|X_k|>\lambda\Big)\leqslant E|X_n|.$$

□

定理 7.4.2　设 $\{X_n\}$ 关于 $\{Y_n\}$ 是鞅，且存在一常数 k, 使 $\forall n,\ EX_n^2\leqslant k<\infty$, 则存在一有限随机变量 X_∞, 使得

$$\begin{aligned}P\Big(\lim_{n\to\infty}X_n=X_\infty\Big)=&1,\\ \lim_{n\to\infty}E|X_n-X_\infty|^2=&0.\end{aligned}$$

更一般地

$$EX_0 = EX_n = EX_\infty, \ \ \forall n.$$

证明 本定理证明较长，在此略去．有兴趣的读者可参看 [2]. □

7.5 连续参数鞅

定义 7.5.1 $\{X(t), t \geqslant 0\}$ 是一随机过程，记 $\mathcal{F}_t = \sigma(X(s), 0 \leqslant s \leqslant t)$. 过程 $\{X(t), t \geqslant 0\}$ 是**鞅**, 如果：

(1) $\forall\, t \geqslant 0$, 有 $E|X(t)| < \infty$;

(2) $\forall\, s, t \geqslant 0$, 有 $E(X(t+s)\big|\mathcal{F}_t) = X(t)$ a.s. (几乎处处) 。

(3) $\forall\, t \geqslant 0$, $X(t)$ 关于 $\mathcal{F}$ 是可测的.

将这一定义离散化，则可写为：

定义 7.5.2 随机过程 $\{X(t), t \geqslant 0\}$ 是**鞅**, 如果：

(1) $\forall\, t \geqslant 0$, 有 $E|X(t)| < \infty$;

(2) $\forall\, 0 \leqslant t_0 < t_1 < \cdots < t_n < t_{n+1}$, 有 $E(X(t_{n+1})\big|X(t_1), \cdots, X(t_n)) = X(t_n)$ a.s. (几乎处处).

同样可以对连续参数定义上鞅和下鞅.

与离散的情形类似，连续参数鞅的停时定义为：

定义 7.5.3 设有非负随机变量 T 及随机序列 $\{X(t)t \geqslant 0\}$, $\mathcal{F}_t = \sigma(X(s), 0 \leqslant s \leqslant t)$. 若对 $\forall\, t \geqslant 0$, $\{T \leqslant t\} \in \mathcal{F}_n$, 称则 T 是 $\{X(t), t \geqslant 0\}$ 的停时.

定理 7.5.1(停时定理) 设 $\{X(t), t \geqslant 0\}$ 是鞅，T 是停时，若 $P(T < \infty) = 1$, 且 $E\Big(\sup\limits_{t \geqslant 0}|X_{T\wedge t}|\Big) < \infty$, 则 $EX(T) = EX(0)$. 定理的证明与离散鞅停时定理的证明类似，留

给读者作为练习.