Last class

- Markov chain
- Chapman-Kolmogorov equations and classification of states
- Stationary distribution
- Transitions and gambler's ruin problem
- Branching processes
- Time-reversible Markov chains and MCMC
- Semi-Markov processes

References: Chapter 4, Markov Chains, 2nd edition, 1995, *by Sheldon M. Ross*

Stochastic Processes

Lecture 5: Martingales

Chao Qian （钱超）

Associate Professor, Nanjing University, China

Email: qianc@nju.edu.cn
Homepage: http://www.lamda.nju.edu.cn/qianc/

Martingales

A stochastic process $\{Z_n, n \geq 1\}$ is said to be a ***martingale*** process if $\forall n$: $E[|Z_n|] < \infty$ and

$$E[\, Z_{n+1} \mid Z_1, Z_2, \dots, Z_n \,] = Z_n$$

$$\Rightarrow \quad E[Z_{n+1}] = E[Z_n] = \cdots = E[Z_1]$$

Example 1: Let $X_1, X_2, \dots$ be independent random variables with mean 0; and let $Z_n = \sum_{i=1}^{n} X_i$. Then $\{Z_n, n \geq 1\}$ is a martingale.

Proof:

$$\begin{aligned}
& E[\, Z_{n+1} \mid Z_1, Z_2, \dots, Z_n \,] \\
&= E[\, Z_n + X_{n+1} \mid Z_1, Z_2, \dots, Z_n \,] \\
&= E[\, Z_n \mid Z_1, Z_2, \dots, Z_n \,] + E[\, X_{n+1} \mid Z_1, Z_2, \dots, Z_n \,] \\
&= Z_n + E[X_{n+1}] = Z_n
\end{aligned}$$

X_i is independent $\qquad E[X_i] = 0$

Martingales

Example 2: Let $X_1, X_2, \ldots$ be independent random variables with $E[X_i] = 1$; and let $Z_n = \prod_{i=1}^{n} X_i$. Then $\{Z_n, n \geq 1\}$ is a martingale.

Proof:

$$\begin{aligned}
&E[\,Z_{n+1} \mid Z_1, Z_2, \ldots, Z_n\,] \\
&= E[\,Z_n \cdot X_{n+1} \mid Z_1, Z_2, \ldots, Z_n\,] \\
&= Z_n \cdot E[\,X_{n+1} \mid Z_1, Z_2, \ldots, Z_n\,] \\
&= Z_n \cdot E[X_{n+1}] = Z_n
\end{aligned}$$

X_i is independent

Example 3: Consider a branching process, and let X_n denote the size of the nth generation. If m is the mean number of offspring per individual, then $\{Z_n, n \geq 1\}$ is a martingale when

$$Z_n = X_n / m^n$$

Leave as the exercise

Martingales

Another way to prove martingale

$$E[Z_{n+1} \mid Z_1, Z_2, \ldots, Z_n, \boldsymbol{Y}] = Z_n \quad \Rightarrow \quad \text{Martingale}$$

some other random vector

Why?

$$E[Z_{n+1} \mid Z_1, Z_2, \ldots, Z_n]$$

$$= E[E[Z_{n+1} \mid Z_1, Z_2, \ldots, Z_n, \boldsymbol{Y}] \mid Z_1, Z_2, \ldots, Z_n]$$

$$= E[Z_n \mid Z_1, Z_2, \ldots, Z_n]$$

$$= Z_n$$

Martingales

Example 4: Let $X, Y_1, Y_2, \ldots$ be arbitrary random variables such that $E[|X|] < \infty$; and let $Z_n = E[X \mid Y_1, \ldots, Y_n]$. Then $\{Z_n, n \geq 1\}$ is a martingale.

Proof:

$$E[\, Z_{n+1} \mid Z_1, \ldots, Z_n, Y_1, \ldots, Y_n \,]$$

$Z_1, \ldots, Z_n$ are determined by $Y_1, \ldots, Y_n$

$$= E[\, Z_{n+1} \mid Y_1, \ldots, Y_n \,]$$

$$= E[\, E[X \mid Y_1, \ldots, Y_n, Y_{n+1}] \mid Y_1, \ldots, Y_n \,]$$ Definition of Z_{n+1}

$$= E[\, X \mid Y_1, \ldots, Y_n \,] = Z_n$$ Conditional expectation

Martingales

Example 5: For any random variables $X_1, X_2, \dots,$ let

$$Z_n = \sum_{i=1}^{n} (X_i - E[X_i \mid X_1, \dots, X_{i-1}])$$

If $E[|Z_n| < \infty]$, then $\{Z_n, n \geq 1\}$ is a martingale.

Proof:

$$Z_{n+1} = Z_n + X_{n+1} - E[\, X_{n+1} \mid X_1, \dots, X_n \,]$$

$$E[\, Z_{n+1} \mid Z_1, \dots, Z_n, X_1, \dots, X_n \,]$$

$$= E[\, Z_{n+1} \mid X_1, \dots, X_n \,]$$

$Z_1, \dots, Z_n$ are determined by $X_1, \dots, X_n$

$$= Z_n + E[\, X_{n+1} \mid X_1, \dots, X_n \,] - E[\, X_{n+1} \mid X_1, \dots, X_n \,]$$

$$= Z_n$$

Stopping times

Random time: The positive integer-valued, possibly infinite, random variable N is said to be a ***random time*** for the process $\{Z_n, n \geq 1\}$ if the event $\{N = n\}$ is determined by the random variables $Z_1, \ldots, Z_n$.

Stopping time: If $P(N < \infty) = 1$, then the random time N is said to be a stopping time

Stopping time: An integer-valued random variable N is said to be a *stopping time* for the sequence of independent random variables $X_1, X_2, \ldots$, if the event $\{N = n\}$ is independent of $X_{n+1}, X_{n+2}, \ldots$, for all $n = 1, 2, \ldots$

From Lecture 3

Stopped process

Let N be a random time for the process $\{Z_n, n \geq 1\}$ and let

$$\bar{Z}_n = \begin{cases} Z_n & \text{if } n \leq N \\ Z_N & \text{if } n > N \end{cases}$$

$\{\bar{Z}_n, n \geq 1\}$ is called the stopped process

Proposition: If N is a random time for the martingale $\{Z_n, n \geq 1\}$, then the stopped process $\{\bar{Z}_n, n \geq 1\}$ is also a martingale

Proof: Let $I_n = \begin{cases} 1 & \text{if } N \geq n \text{ (i.e., not stopped after observing } Z_1, \ldots, Z_{n-1}) \\ 0 & \text{if } N < n \end{cases}$

$$\Rightarrow \bar{Z}_n = \bar{Z}_{n-1} + I_n(Z_n - Z_{n-1})$$

Stopped process

$$\bar{Z}_n = \bar{Z}_{n-1} + I_n(Z_n - Z_{n-1})$$

Verify the above equation:

$$\begin{cases} 1.\ N \geq n:\ \bar{Z}_n = Z_n,\ \bar{Z}_{n-1} = Z_{n-1},\ I_n = 1,\ \text{the equation holds} \\ 2.\ N < n:\ \bar{Z}_n = Z_N,\ \bar{Z}_{n-1} = Z_N,\ I_n = 0,\ \text{the equation holds} \end{cases}$$

$$E[\,\bar{Z}_n \mid Z_1, Z_2, \ldots, Z_{n-1}\,] = E[\,\bar{Z}_{n-1} + I_n(Z_n - Z_{n-1}) \mid Z_1, Z_2, \ldots, Z_{n-1}\,]$$

$\bar{Z}_{n-1}$ and I_n can be determined by $Z_1, \ldots, Z_{n-1}$

$$= \bar{Z}_{n-1} + I_n \cdot \underbrace{E[\,Z_n - Z_{n-1} \mid Z_1, Z_2, \ldots, Z_{n-1}\,]}$$

$$= \bar{Z}_{n-1}$$

$= 0$ because $\{Z_n, n \geq 1\}$ is a martingale

$$E[\,\bar{Z}_n \mid \bar{Z}_1, \ldots, \bar{Z}_{n-1}, Z_1, \ldots, Z_{n-1}\,] = E[\,\bar{Z}_n \mid Z_1, Z_2, \ldots, Z_{n-1}\,]$$

$\bar{Z}_1, \ldots, \bar{Z}_{n-1}$ are determined by $Z_1, \ldots, Z_{n-1}$

$$= \bar{Z}_{n-1}$$

Martingale stopping theorem

Martingale stopping theorem: If either

- ✓ $\bar{Z}_n$ are uniformly bounded, or
- ✓ N is bounded, or stopping time
- ✓ $E[N] < \infty$, and there is an $M < \infty$ such that

$$E[\,|Z_{n+1} - Z_n| \mid Z_1, \dots, Z_n\,] < M$$

then

$$E[Z_N] = E[Z_1]$$

Proof:

$$E[\bar{Z}_n] = E[\bar{Z}_1] = E[Z_1]$$

$\bar{Z}_n \to Z_N$ as $n \to \infty$ with probability 1

Martingale stopping theorem

Martingale stopping theorem: If either

- ✓ $\bar{Z}_n$ are uniformly bounded, or
- ✓ N is bounded, or stopping time
- ✓ $E[N] < \infty$, and there is an $M < \infty$ such that

$$E[\,|Z_{n+1} - Z_n| \mid Z_1, \ldots, Z_n\,] < M$$

then

$$E[Z_N] = E[Z_1]$$

Proof:

$$E[\bar{Z}_n] = E[\bar{Z}_1] = E[Z_1]$$

$$E[\bar{Z}_n] \to E[Z_N] \text{ as } n \to \infty$$

Martingale stopping theorem

Wald's Equation: If $X_1, X_2, \ldots$ are iid random variables having finite expectations, and if N is a stopping time for $X_1, X_2, \ldots$ such that $E[N] < \infty$, then

$$E\left[\sum_{n=1}^{N} X_n\right] = E[N]E[X]$$

Another Proof using martingale stopping theorem:

Let $E[X] = \mu$

$$Z_n = \sum_{i=1}^{n} (X_i - \mu) \Rightarrow \text{a martingale}$$

Martingale stopping theorem

$$Z_n = \sum_{i=1}^{n}(X_i - \mu) \quad \Rightarrow \quad \text{a martingale}$$

Verify the third condition of martingale stopping theorem:

$E[N] < \infty$

$$E[\,|Z_{n+1} - Z_n| \mid Z_1, \dots, Z_n\,] = E[\,|X_{n+1} - \mu| \mid Z_1, \dots, Z_n\,]$$
$$= E[|X_{n+1} - \mu|]$$
$$\le E[|X_{n+1}|] + |\mu| < \infty$$

Apply martingale stopping theorem:

$E[Z_N] = E[Z_1] = 0$

$$= E\left[\sum_{i=1}^{N}(X_i - \mu)\right] = E\left[\sum_{i=1}^{N} X_i - N\mu\right] = E\left[\sum_{i=1}^{N} X_i\right] - E[N]\mu$$

Martingale stopping theorem

Example: At a party, n people put their hats in the center of a room where the hats are mixed together. Each person then randomly selects one. Those choosing their own hats depart, while the others (those without a match) put their selected hats in the center of the room, mix them up, and then reselect. Let R denote the number of rounds until all people have a match. What is $E[R]$?

Solution:

Let X_i denote the number of matches on the ith round

Note that $X_i = 1$ for $i > R$

$$Z_k = \sum_{i=1}^{k} (X_i - E[X_i \mid X_1, \ldots, X_{i-1}]) \Rightarrow \text{a martingale}$$

Martingale stopping theorem

Note: $X_i = 1$ for $i > R$

Let X_i denote the number of matches on the ith round

a martingale $$Z_k = \sum_{i=1}^{k} (X_i - E[X_i \mid X_1, \dots, X_{i-1}]) = \sum_{i=1}^{k} (X_i - 1)$$

R is the stopping time of $\{Z_k, k \geq 1\}$

$$E[\,|Z_{k+1} - Z_k| \mid Z_1, \dots, Z_k\,] = E[\,|X_{k+1} - 1| \mid Z_1, \dots, Z_k\,] \leq 2$$

Applying martingale stopping theorem:

$$0 = E[Z_1] = E[Z_R] = E\left[\sum_{i=1}^{R} (X_i - 1)\right] = E\left[\sum_{i=1}^{R} X_i\right] - E[R] = n - E[R]$$

Martingale stopping theorem

Example: Suppose that a sequence of iid discrete random variables is observed sequentially, one at each day. What is the expected number N that must be observed until some given sequence appears?

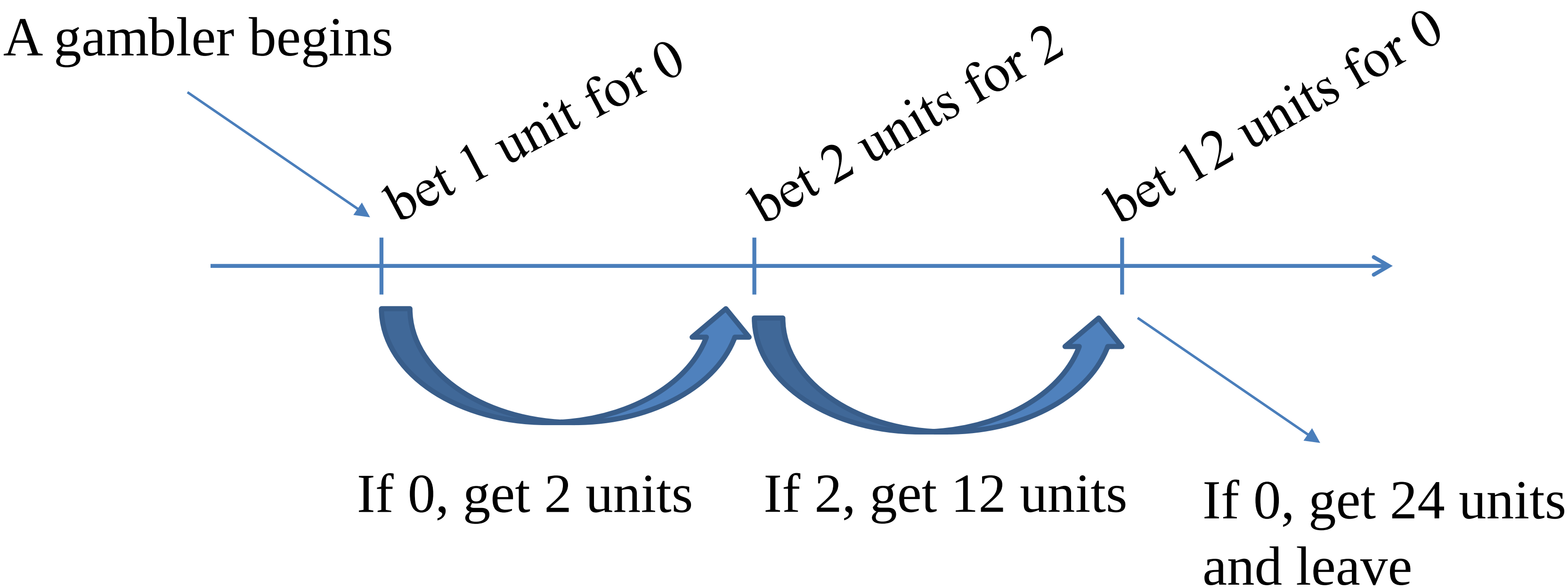
In Lecture 3, we have used delayed renewal process to compute it

Now, we will show how to use the martingale stopping theorem to compute it?

Example: More specifically, suppose that each outcome is either 0, 1, or 2 with respective probabilities $\frac{1}{2}$, $\frac{1}{3}$ and $\frac{1}{6}$, and we desire the expected time until the run 020 occurs

Martingale stopping theorem

Construct a fair gambling model for the pattern “020”



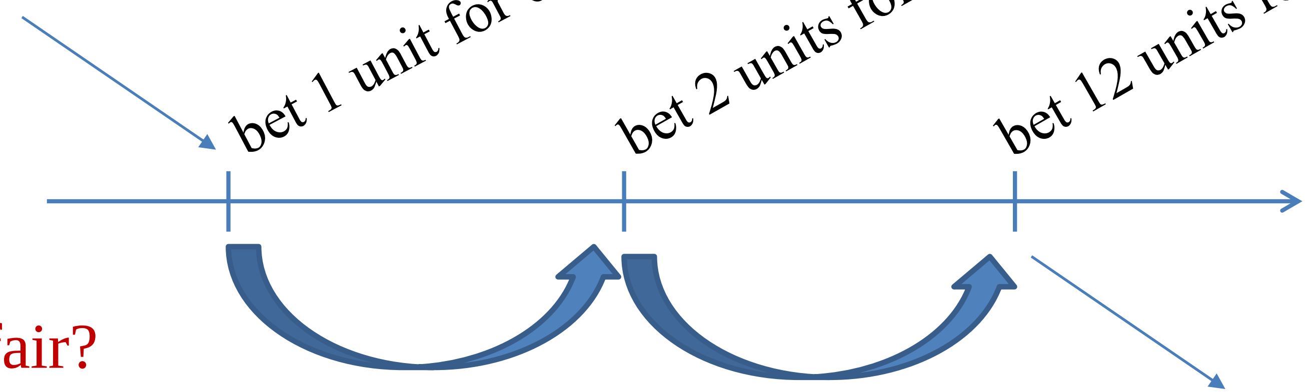
The gambler will lose 1 unit if any of her bets fails and will win 23 if all three of her bets succeed

Martingale stopping theorem

0 with probability 1/2

1 with probability 1/3

2 with probability 1/6



Why fair?

Suppose $P(Z=1)=p, P(Z=0)=1-p$

Bet a units for 1: $\begin{cases} \text{if 1 get } x \text{ units} \\ \text{if 0 lose } a \text{ units} \end{cases}$

Fair $\Leftrightarrow E[winning]=0$

$\Leftrightarrow p(x-a)+(1-p)(-a)=0 \Leftrightarrow x=\frac{a}{p}$

$\frac{1}{2}\times(2-1)+\frac{1}{2}\times(-1)=0$

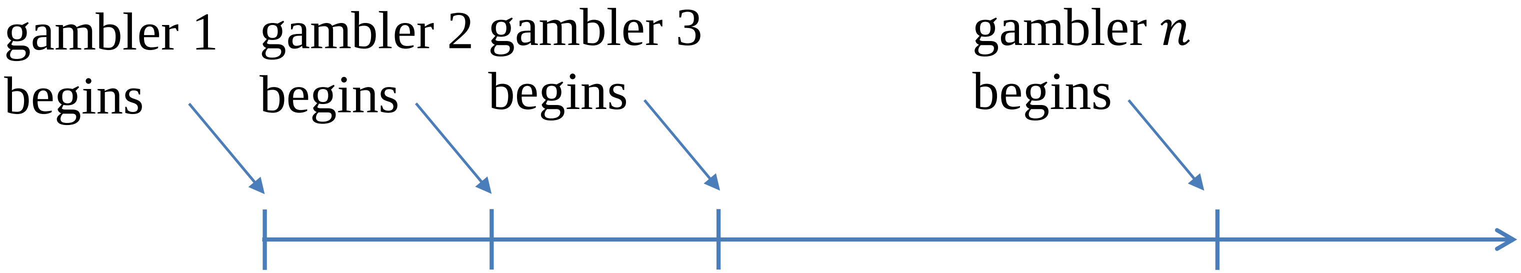
$\frac{1}{6}\times(12-2)+\frac{5}{6}\times(-2)=0$

$\frac{1}{2}\times(24-12)+\frac{1}{2}\times(-12)=0$

The expected winning of the gambler at each time is 0

Martingale stopping theorem

fair gambling casino



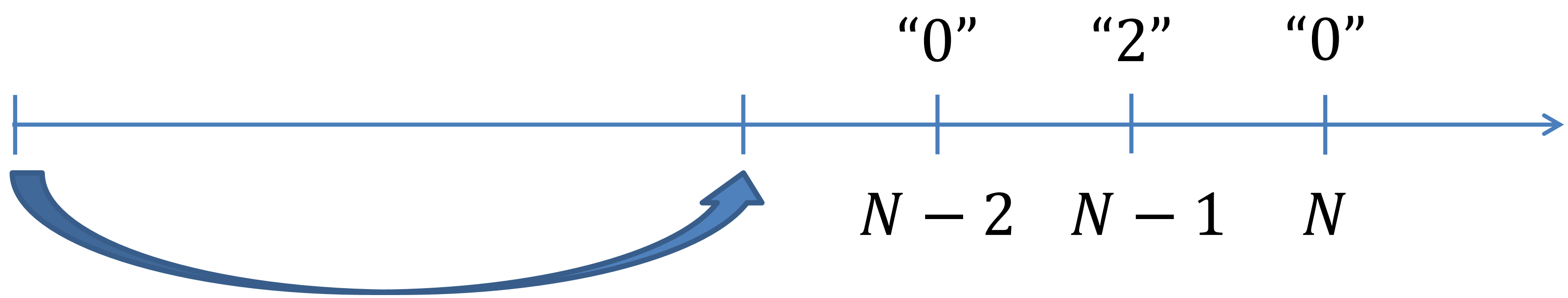
Let X_n denote the total winnings of the casino after the nth day

$\Downarrow$ all bets are fair

$\{X_n, n \geq 1\}$ is a martingale

Martingale stopping theorem

The required number N for "020" is a stopping time for $\{X_n, n \geq 1\}$



Each of the first $N-3$ gamblers loses 1 unit	Gambler $N-2$ wins 23 units	Gambler $N-1$ loses 1 unit	Gambler N wins 1 unit

$$X_N = (N-3) - 23 + 1 - 1 = N - 26$$

$$N = \min\{n: X_n = n - 26\}$$

Martingale stopping theorem

$\{X_n, n \geq 1\}$ is a martingale

The required number N for "020" is a stopping time for $\{X_n, n \geq 1\}$

$$|X_{n+1} - X_n| \leq 3 * 23$$

Applying martingale stopping theorem:

$$E[X_N] = E[X_1] = 0$$

$$X_N = N - 26$$

$$E[N] = 26$$

Martingale stopping theorem

Example: More specifically, suppose that each outcome is either H or T with respective probabilities p and $q = 1 - p$, and we desire the expected time until HHTTHH occurs

leave as the exercise

Azuma's inequality

Azuma's inequality: Let $Z_n, n \geq 1$ be a martingale with mean $\mu = E[Z_n]$. Let $Z_0 = \mu$ and suppose that for nonnegative constants $\alpha_i, \beta_i, i \geq 1$,

$$-\alpha_i \leq Z_i - Z_{i-1} \leq \beta_i$$

Then for any $n \geq 0, a > 0$

$$P(Z_n - \mu \geq a) \leq \exp\left\{-2a^2 / \sum_{i=1}^{n} (\alpha_i + \beta_i)^2\right\}$$

$$P(Z_n - \mu \leq -a) \leq \exp\left\{-2a^2 / \sum_{i=1}^{n} (\alpha_i + \beta_i)^2\right\}$$

Azuma's inequality

Lemma 1: Let X be such that $E[X] = 0$ and $P\{-\alpha \leq X \leq \beta\} = 1$. Then for any convex function f

$$E[f(X)] \leq \frac{\beta}{\alpha+\beta} f(-\alpha) + \frac{\alpha}{\alpha+\beta} f(\beta)$$

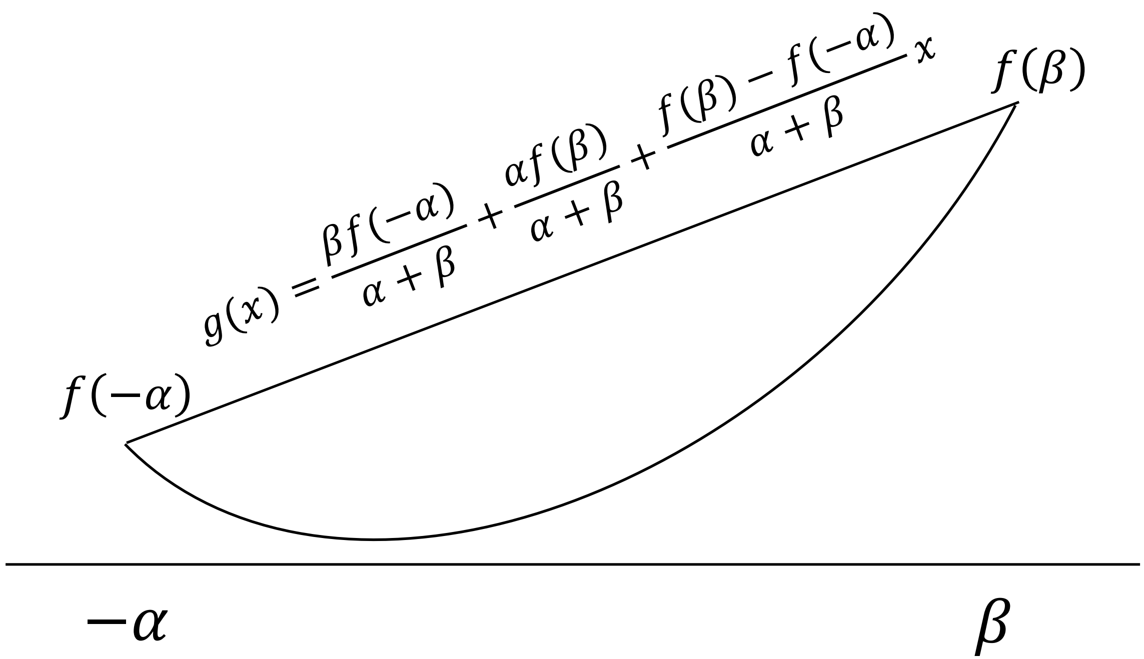
Proof:

$\forall -\alpha \leq x \leq \beta, f(x) \leq g(x)$

$$\Downarrow$$

$E[f(X)] \leq \frac{\beta f(-\alpha)}{\alpha+\beta} + \frac{\alpha f(\beta)}{\alpha+\beta}$

(by $E[X] = 0$)



Azuma's inequality

Lemma 1: Let X be such that $E[X]=0$ and $P\{-\alpha \le X \le \beta\}=1$. Then for any convex function f

$$E[f(X)] \le \frac{\beta}{\alpha+\beta} f(-\alpha) + \frac{\alpha}{\alpha+\beta} f(\beta)$$

Lemma 2: For $0 \le \theta \le 1$,

$$\theta e^{(1-\theta)x} + (1-\theta) e^{-\theta x} \le e^{x^2/8}$$

Azuma's inequality

Proof of Azuma's inequality: Suppose first that $\mu = E[Z_n] = 0$

For any $c > 0$

Markov inequality

$$P(Z_n \geq a) = P(e^{cZ_n} \geq e^{ca}) \leq E[e^{cZ_n}] \cdot e^{-ca}$$

Let $W_n = e^{cZ_n}$, then $W_0 = 1$ and for $n > 0$

$$W_n = e^{cZ_{n-1}} \cdot e^{c(Z_n - Z_{n-1})}$$

$$E[W_n \mid Z_{n-1}] = e^{cZ_{n-1}} \cdot E[e^{c(Z_n - Z_{n-1})} \mid Z_{n-1}]$$

Azuma's inequality

$$E[W_n \mid Z_{n-1}] = e^{cZ_{n-1}} \cdot E[e^{c(Z_n - Z_{n-1})} \mid Z_{n-1}]$$

by Lemma 1

$$\leq W_{n-1} \cdot \left(\beta_n e^{-c\alpha_n} + \alpha_n e^{c\beta_n}\right) / (\alpha_n + \beta_n)$$

check conditions of Lemma 1:

- $f(x) = e^{cx}$ is convex
- $-\alpha_n \leq Z_n - Z_{n-1} \leq \beta_n$
- $E[Z_n - Z_{n-1} | Z_{n-1}] = E[Z_n | Z_{n-1}] - Z_{n-1} = 0$

Lemma 1: Let X be such that $E[X] = 0$ and $P\{-\alpha \leq X \leq \beta\} = 1$. Then for any convex function f

α_n β_n

$$E[f(X)] \leq \frac{\beta}{\alpha + \beta} f(-\alpha) + \frac{\alpha}{\alpha + \beta} f(\beta)$$

Azuma's inequality

$$E[W_n \mid Z_{n-1}] = e^{cZ_{n-1}} \cdot E[e^{c(Z_n - Z_{n-1})} \mid Z_{n-1}]$$

$$\leq W_{n-1} \cdot \left(\beta_n e^{-c\alpha_n} + \alpha_n e^{c\beta_n}\right)/(\alpha_n + \beta_n)$$

$$E[W_n] \leq E[W_{n-1}] \cdot \left(\beta_n e^{-c\alpha_n} + \alpha_n e^{c\beta_n}\right)/(\alpha_n + \beta_n)$$

$$E[W_n] \leq \prod_{i=1}^{n} \left(\beta_i e^{-c\alpha_i} + \alpha_i e^{c\beta_i}\right)/(\alpha_i + \beta_i)$$

Azuma's inequality

$$E[W_n] \leq \prod_{i=1}^{n} \left(\beta_i e^{-c\alpha_i} + \alpha_i e^{c\beta_i}\right)/(\alpha_i + \beta_i)$$

$$\leq \prod_{i=1}^{n} e^{c^2(\alpha_i+\beta_i)^2/8}$$

$$\Uparrow \quad \theta = \alpha_i/(\alpha_i + \beta_i)$$

$$x = c(\alpha_i + \beta_i)$$

Lemma 2: For $0 \leq \theta \leq 1$,

$$\theta e^{(1-\theta)x} + (1-\theta)e^{-\theta x} \leq e^{x^2/8}$$

Azuma's inequality

Proof of Azuma's inequality: Suppose first that $\mu = E[Z_n] = 0$

For any $c > 0$

$$P(Z_n \geq a) = P(e^{cZ_n} \geq e^{ca}) \leq E[e^{cZ_n}] \cdot e^{-ca}$$

$$\leq e^{-ca} \prod_{i=1}^{n} e^{c^2(\alpha_i+\beta_i)^2/8} = e^{-ca+c^2\sum_{i=1}^{n}(\alpha_i+\beta_i)^2/8}$$

$$\Downarrow \quad c = 4a/\sum_{i=1}^{n}(\alpha_i+\beta_i)^2$$

$$\leq e^{-2a^2/\sum_{i=1}^{n}(\alpha_i+\beta_i)^2}$$

Azuma's inequality

Suppose that $\mu = E[Z_n] = 0$, then $P(Z_n \geq a) \leq e^{-2a^2/\sum_{i=1}^{n}(\alpha_i+\beta_i)^2}$

$\Downarrow$ zero-mean martingale $\{Z_n - \mu\}$

Azuma's inequality: Let $Z_n, n \geq 1$ be a martingale with mean $\mu = E[Z_n]$. Let $Z_0 = \mu$ and suppose that for nonnegative constants $\alpha_i, \beta_i, i \geq 1$,

$$-\alpha_i \leq Z_i - Z_{i-1} \leq \beta_i$$

Then for any $n \geq 0, a > 0$

$$P(Z_n - \mu \geq a) \leq \exp\left\{-2a^2 / \sum_{i=1}^{n} (\alpha_i + \beta_i)^2\right\}$$

Azuma's inequality

Suppose that $\mu = E[Z_n] = 0$, then $P(Z_n \geq a) \leq e^{-2a^2/\sum_{i=1}^{n}(\alpha_i+\beta_i)^2}$

$\Downarrow$ zero-mean martingale $\{\mu - Z_n\}$

Azuma's inequality: Let $Z_n, n \geq 1$ be a martingale with mean $\mu = E[Z_n]$. Let $Z_0 = \mu$ and suppose that for nonnegative constants $\alpha_i, \beta_i, i \geq 1$,

$$-\alpha_i \leq Z_i - Z_{i-1} \leq \beta_i$$

Then for any $n \geq 0, a > 0$

$$P(Z_n - \mu \leq -a) \leq \exp\left\{-2a^2 / \sum_{i=1}^{n} (\alpha_i + \beta_i)^2\right\}$$

Azuma's inequality

Example: Let $X_1, \dots, X_n$ be random variables such that $E[X_1] = 0$ and $E[X_i \mid X_1, \dots, X_{i-1}] = 0, i > 1$. If $-\alpha_i \le X_i \le \beta_i$,

$$P\left(\sum_{i=1}^{n} X_i \ge a\right) \le \exp\left\{-2a^2 / \sum_{i=1}^{n} (\alpha_i + \beta_i)^2\right\}$$

Solution:

Azuma's inequality

$=\sum_{i=1}^{j}(X_i - E[X_i \mid X_1, \dots, X_{i-1}])$

$\sum_{i=1}^{j} X_i$: a zero-mean martingale

$$-\alpha_j \le \sum_{i=1}^{j} X_i - \sum_{i=1}^{j-1} X_i = X_j \le \beta_j$$

Azuma's inequality

Example: Suppose that n balls are put in m urns in such a manner that each ball, independently, is equally likely to go into any of the urns.

Let X the number of empty urns, then $X = \sum_{i=1}^{m} I(\text{urn } i \text{ is empty})$

$$\mu = E[X] = mP(\text{urn } i \text{ is empty}) = m\left(1 - \frac{1}{m}\right)^n$$

$$P(X - \mu \geq a) \leq ? \qquad P(X - \mu \leq -a) \leq ?$$

Solution: Let X_j denote the urn in which the jth ball is placed

$$Z_0 = E[X]$$

$$Z_n = E[X \mid X_1, \dots, X_n] = X$$

$Z_i = E[X \mid X_1, \dots, X_i]$: a martingale

Azuma's inequality

To analyze $$\boxed{-\alpha_i \leq Z_i - Z_{i-1} \leq \beta_i}$$

$Z_i = E[X \mid X_1, \cdots, X_i\,],\ Z_{i-1} = E[X \mid X_1, \cdots, X_{i-1}\,], Z_1 - Z_0 = E[X \mid X_1] - E[X] = 0$

When $i \geq 2$, let D denote the number of different values taken by $X_1, \dots, X_{i-1}$, i.e., the number of non-empty urns

$$E[X \mid X_1, \cdots, X_{i-1}\,] = (m - D)\left(1 - \frac{1}{m}\right)^{n-i+1}$$

$$E[X \mid X_1, \cdots, X_i\,] = \begin{cases} \text{if } X_i \in \{X_1, \dots, X_{i-1}\}, (m - D)\left(1 - \frac{1}{m}\right)^{n-i} \\ \text{if } X_i \notin \{X_1, \dots, X_{i-1}\}, (m - D - 1)\left(1 - \frac{1}{m}\right)^{n-i} \end{cases}$$

$$Z_i - Z_{i-1} = \frac{m-D}{m}\left(1 - \frac{1}{m}\right)^{n-i} \text{ or } -\frac{D}{m}\left(1 - \frac{1}{m}\right)^{n-i}$$

By $1 \leq D \leq \min\{i - 1, m\}$, we get $-\min\left\{\frac{i-1}{m}, 1\right\}\left(1 - \frac{1}{m}\right)^{n-i} \leq Z_i - Z_{i-1} \leq \left(1 - \frac{1}{m}\right)^{n-i+1}$

Azuma's inequality

Example: Suppose that n balls are put in m urns in such a manner that each ball, independently, is equally likely to go into any of the urns.

Let X the number of empty urns, then $X = \sum_{i=1}^{m} I(\text{urn } i \text{ is empty})$

$$\mu = E[X] = mP(\text{urn } i \text{ is empty}) = m\left(1 - \frac{1}{m}\right)^n$$

Apply Azuma's inequality:

$$P(X - \mu \geq a) \leq \exp\left\{-2a^2 / \sum_{i=2}^{n} (\alpha_i + \beta_i)^2\right\}$$

$$\sum_{i=2}^{n} (\alpha_i + \beta_i)^2 = \sum_{i=2}^{m} \left(\frac{m+i-2}{m}\right)^2 \left(1 - \frac{1}{m}\right)^{2(n-i)} + \sum_{i=m+1}^{n} \left(2 - \frac{1}{m}\right)^2 \left(1 - \frac{1}{m}\right)^{2(n-i)}$$

Azuma's inequality

Corollary: Let h be a function such that if the vectors $\boldsymbol{x} = (x_1, \dots, x_n)$ and $\boldsymbol{y} = (y_1, \dots, y_n)$ differ in at most one coordinate (i.e., for some k, $x_i = y_i$ for all $i \neq k$) then $|h(\boldsymbol{x}) - h(\boldsymbol{y})| \leq 1$. Let $X_1, \dots, X_n$ be independent random variables. Then, with $\boldsymbol{X} = (X_1, \dots, X_n)$, we have for $a > 0$ that

$$P(h(\boldsymbol{X}) - E[h(\boldsymbol{X})] \geq a) \leq \exp\{-a^2/(2n)\}$$

$$P(h(\boldsymbol{X}) - E[h(\boldsymbol{X})] \leq -a) \leq \exp\{-a^2/(2n)\}$$

Proof: Let $Z_i = E[h(\boldsymbol{X}) \mid X_1, \dots, X_i]$, then $\{Z_i, i \geq 1\}$ is a martingale. Now,

$$\begin{aligned}
&|E[h(\boldsymbol{X}) \mid X_1 = x_1, \dots, X_i = x_i] - E[h(\mathbf{X}) \mid X_1 = x_1, \dots, X_{i-1} = x_{i-1}]| \\
&= |E[h(x_1, \dots, x_i, X_{i+1}, \dots, X_n)] - E[h(x_1, \dots, x_{i-1}, X_i, \dots, X_n)]| \\
&= |E[h(x_1, \dots, x_i, X_{i+1}, \dots, X_n) - h(x_1, \dots, x_{i-1}, X_i, \dots, X_n)]| \\
&\leq E[|h(x_1, \dots, x_i, X_{i+1}, \dots, X_n) - h(x_1, \dots, x_{i-1}, X_i, \dots, X_n)|] \leq 1
\end{aligned}$$

Thus, $|Z_i - Z_{i-1}| \leq 1$ and the result holds by Azuma's inequality with $\alpha_i = \beta_i = 1$

Azuma's inequality

Example: Suppose that n balls are to be placed in m urns, with each ball independently going into urn j with probability $p_j, j = 1, \ldots, m$. Let Y_k denote the number of urns with exactly k balls, $0 \le k < n$, and use the preceding corollary to obtain a bound on its tail probabilities.

Solution:

$$E[Y_k] = E\left[\sum_{i=1}^{m} I(\text{urn } i \text{ has exactly } k \text{ balls})\right] = \sum_{i=1}^{m} \binom{n}{k} p_i^k (1 - p_i)^{n-k}$$

Let X_j denote the urn in which the jth ball is placed

$$Y_k = h_k(X_1, \ldots, X_n)$$

Azuma's inequality

For $k = 0$

If the vectors $\boldsymbol{x} = (x_1, \dots, x_n)$ and $\boldsymbol{y} = (y_1, \dots, y_n)$ differ in at most one coordinate, then

$$\boxed{|h_0(\boldsymbol{x}) - h_0(\boldsymbol{y})| \le 1}$$

$$\Downarrow \text{ Apply Corollary}$$

$$P\left(Y_0 - \sum_{i=1}^{m} (1 - p_i)^n \ge a\right) \le \exp\{-a^2/(2n)\}$$

$$P\left(Y_0 - \sum_{i=1}^{m} (1 - p_i)^n \le -a\right) \le \exp\{-a^2/(2n)\}$$

Azuma's inequality

For $0 < k < n$

If the vectors $\boldsymbol{x} = (x_1, \dots, x_n)$ and $\boldsymbol{y} = (y_1, \dots, y_n)$ differ in at most one coordinate, then

$$\boxed{|h_k(\boldsymbol{x}) - h_k(\boldsymbol{y})| \leq 2}$$

Azuma's inequality

For $0 < k < n$

If the vectors $\boldsymbol{x} = (x_1, \dots, x_n)$ and $\boldsymbol{y} = (y_1, \dots, y_n)$ differ in at most one coordinate, then

$$\boxed{\left|\frac{h_k(\boldsymbol{x})}{2} - \frac{h_k(\boldsymbol{y})}{2}\right| \le 1}$$

$\Downarrow$ Apply Corollary

$$P\left(Y_k - \sum_{i=1}^{m} \binom{n}{k} p_i^k (1-p_i)^{n-k} \ge 2a\right) \le \exp\{-a^2/(2n)\}$$

$$P\left(Y_k - \sum_{i=1}^{m} \binom{n}{k} p_i^k (1-p_i)^{n-k} \le -2a\right) \le \exp\{-a^2/(2n)\}$$

Azuma's inequality

Example: Consider a set of n components that are to be used in performing certain experiments. Let X_i equal 1 if component i is in functioning condition and let it equal 0 otherwise, and suppose that the X_i are independent with $E[X_i] = p_i$.

Suppose that in order to perform experiment $j, j = 1, \ldots, m$, all of the components in the set A_j must be functioning.

If any component is needed in at most three experiments, show that

for $a > 0$

$$P\left(X - \sum_{j=1}^{m} \prod_{i \in A_j} p_i \geq 3a\right) \leq \exp\{-a^2/(2n)\}$$

#experiments that can be performed

leave as the exercise

$$P\left(X - \sum_{j=1}^{m} \prod_{i \in A_j} p_i \leq -3a\right) \leq \exp\{-a^2/(2n)\}$$

Submartingales and supermartingales

A stochastic process $\{Z_n, n \geq 1\}$ is said to be a ***submartingale*** process if $\forall n$: $E[|Z_n|] < \infty$ and

$$E[\, Z_{n+1} \mid Z_1, Z_2, \dots, Z_n \,] \geq Z_n$$

$$\Rightarrow \quad E[Z_{n+1}] \geq E[Z_n] \geq \cdots \geq E[Z_1]$$

A stochastic process $\{Z_n, n \geq 1\}$ is said to be a ***supermartingale*** process if $\forall n$: $E[|Z_n|] < \infty$ and

$$E[\, Z_{n+1} \mid Z_1, Z_2, \dots, Z_n \,] \leq Z_n$$

$$\Rightarrow \quad E[Z_{n+1}] \leq E[Z_n] \leq \cdots \leq E[Z_1]$$

Martingale stopping theorem

Martingale stopping theorem: If either

- ✓ $\bar{Z}_n$ are uniformly bounded, or
- ✓ N is bounded, or stopping time
- ✓ $E[N] < \infty$, and there is an $M < \infty$ such that

$$E[\,|Z_{n+1} - Z_n| \mid Z_1, \dots, Z_n\,] < M$$

then

$$E[Z_N] \geq E[Z_1] \qquad \text{for a submartingale}$$

$$E[Z_N] \leq E[Z_1] \qquad \text{for a supermartingale}$$

Martingale convergence theorem

Martingale convergence theorem: If $\{Z_n, n \geq 1\}$ is a martingale such that for some $M < \infty$

$$E[|Z_n|] \leq M, \text{ for all } n$$

then, with probability 1, $\lim_{n \to \infty} Z_n$ exists and is finite

Martingale convergence theorem

Lemma: If $\{Z_n, n \geq 1\}$ is a martingale and f is a convex function, then $\{f(Z_n), n \geq 1\}$ is a submartingale

Proof:

By Jansen's inequality, we have

$$E[f(Z_{n+1})|Z_1, \ldots, Z_n] \geq f(E[Z_{n+1} \mid Z_1, \ldots, Z_n]) = f(Z_n)$$

Then,

$$\begin{aligned}
&E[f(Z_{n+1})|f(Z_1), \ldots, f(Z_n)] \\
&= E[E[f(Z_{n+1}) \mid f(Z_1), \ldots, f(Z_n), Z_1, \ldots, Z_n] \mid f(Z_1), \ldots, f(Z_n)] \\
&= E[E[f(Z_{n+1}) \mid Z_1, \ldots, Z_n] \mid f(Z_1), \ldots, f(Z_n)] \\
&\geq E[f(Z_n) \mid f(Z_1), \ldots, f(Z_n)] \\
&= f(Z_n)
\end{aligned}$$

Martingale convergence theorem

Kolmogorov's Inequality for Submartingales: If $\{Z_n, n \geq 1\}$ is a nonnegative submartingale, then for $a > 0$

$$P(\max\{Z_1, \dots, Z_n\} > a) \leq \frac{E[Z_n]}{a}$$

Proof: Let $N = \begin{cases} \min\{i: Z_i > a, i \leq n\} & \text{if } \{i: Z_i > a, i \leq n\} \neq \emptyset \\ n & \text{otherwise} \end{cases}$

Then, $\max\{Z_1, \dots, Z_n\} > a \Leftrightarrow Z_N > a$

Thus, $P(\max\{Z_1, \dots, Z_n\} > a) = P(Z_N > a) \leq E[Z_N]/a$

Note that $N \leq n$ and N is a stopping time,

$$E[Z_n|Z_1, \dots, Z_N, N = k] = E[Z_n|Z_1, \dots, Z_k, N = k] = E[Z_n|Z_1, \dots, Z_k]$$
$$= E[E[Z_n \mid Z_1, \dots, Z_k, Z_{k+1}, \dots, Z_{n-1}] \mid Z_1, \dots, Z_k] \geq E[Z_{n-1} \mid Z_1, \dots, Z_k]$$
$$\geq \cdots \geq E[Z_{k+1} \mid Z_1, \dots, Z_k] \geq Z_k = Z_N$$

Take expectations on both side, we have $E[Z_n] \geq E[Z_N]$

Martingale convergence theorem

Kolmogorov's Inequality for Submartingales: If $\{Z_n, n \geq 1\}$ is a nonnegative submartingale, then for $a > 0$

$$P(\max\{Z_1, \dots, Z_n\} > a) \leq \frac{E[Z_n]}{a}$$

$|x|$ and x^2 are convex $\Downarrow$ $\{|Z_n|, n \geq 1\}$ and $\{Z_n^2, n \geq 1\}$ nonnegative submartingale

Corollary: Let $\{Z_n, n \geq 1\}$ be a martingale, then for $a > 0$

$$P(\max\{|Z_1|, \dots, |Z_n|\} > a) \leq \frac{E[|Z_n|]}{a}$$

$$P(\max\{|Z_1|, \dots, |Z_n|\} > a) \leq \frac{E[Z_n^2]}{a^2}$$

Martingale convergence theorem

Martingale convergence theorem: If $\{Z_n, n \geq 1\}$ is a martingale such that for some $M < \infty$

$$E[|Z_n|] \leq M, \text{ for all } n$$

then, with probability 1, $\lim_{n\to\infty} Z_n$ exists and is finite

Proof: Under the stronger assumption that $E[Z_n^2]$ is bounded

To show that $\{Z_n, n \geq 1\}$ is, with probability 1, a Cauchy sequence, i.e., with probability 1, for any $k \geq 1$

$$|Z_{m+k} - Z_m| \to 0, \text{ as } m \to \infty$$

Note that $\{Z_{m+k} - Z_m, k \geq 1\}$ is a martingale

Martingale convergence theorem

$$P\left(\max_{1\le k\le n}|Z_{m+k}-Z_m|>\epsilon\right)\le\frac{E[(Z_{m+n}-Z_m)^2]}{\epsilon^2}$$

$$=E\left[Z_{m+n}^2-2Z_mZ_{m+n}+Z_m^2\right]/\epsilon^2\ =E\left[Z_{m+n}^2-Z_m^2\right]/\epsilon^2,$$

where the last equality is by

$$E[Z_mZ_{m+n}]=E\big[E[Z_mZ_{m+n}\mid Z_m]\big]=E\big[Z_mE[Z_{m+n}\mid Z_m]\big]=E[Z_m^2]$$

Note that $\{Z_n, n\ge 1\}$ is a martingale, $f(x)=x^2$ is convex, thus $\{Z_n^2, n\ge 1\}$ is a submartingale, implying $E\left[Z_{n+1}^2\right]\ge E[Z_n^2]$. By the assumption that $E[Z_n^2]$ is bounded, we get $\lim_{n\to\infty}E[Z_n^2]=\mu<\infty$

Let $n\to\infty$, we have $P\left(\max_{k\ge 1}|Z_{m+k}-Z_m|>\epsilon\right)\le(\mu-E[Z_m^2])/\epsilon^2$. Thus,

$$P\left(\max_{k\ge 1}|Z_{m+k}-Z_m|>\epsilon\right)\to 0\text{ as }m\to\infty$$

Martingale convergence theorem

Martingale convergence theorem: If $\{Z_n, n \geq 1\}$ is a martingale such that for some $M < \infty$

$$E[|Z_n|] \leq M, \text{ for all } n$$

then, with probability 1, $\lim_{n\to\infty} Z_n$ exists and is finite

$$\Downarrow \quad E[|Z_n|] = E[Z_n] = E[Z_1]$$

Corollary: If $\{Z_n, n \geq 1\}$ is a nonnegative martingale, then, with probability 1, $\lim_{n\to\infty} Z_n$ exists and is finite

Martingale convergence theorem

Strong Law of Large Numbers: If $X_1, X_2, \ldots$ are independent and identically distributed with mean μ, then

$$P\left(\lim_{n\to\infty}(X_1+\cdots+X_n)/n=\mu\right)=1$$

Proof: Let $S_n = X_1 + \cdots + X_n$

To show that for a given $\epsilon > 0$, $P\left(\lim_{n\to\infty}\frac{S_n}{n} \geq \mu+\epsilon\right)=0$

$$\psi(t)=E[e^{tX}] \qquad g(t)=e^{t(\mu+\epsilon)}/\psi(t)$$

Then, $g(0)=1$ and

$$g'(0)=\frac{(\mu+\epsilon)e^{t(\mu+\epsilon)}\psi(t)-\psi'(t)e^{t(\mu+\epsilon)}}{\psi^2(t)}\Big|_{t=0}=\mu+\epsilon-\mu>0$$

Martingale convergence theorem

there exists $t_0 > 0$ such that $g(t_0) > 1$

$$\boxed{\frac{S_n}{n} \geq \mu + \epsilon \quad \Rightarrow \quad \frac{e^{t_0 S_n}}{\psi^n(t_0)} \geq \left(\frac{e^{t_0(\mu+\epsilon)}}{\psi(t_0)}\right)^n = g^n(t_0)}$$

$$\frac{\prod_{i=1}^n e^{t_0 X_i}}{\psi^n(t_0)} = \prod_{i=1}^n \frac{e^{t_0 X_i}}{\psi(t_0)}$$

By Example 2 in page 4, $\prod_{i=1}^n \frac{e^{t_0 X_i}}{\psi(t_0)}$ is a martingale

By martingale convergence theorem:

With prob. 1, $\lim_{n\to\infty} \frac{e^{t_0 S_n}}{\psi^n(t_0)}$ exists and is finite

$$\lim_{n\to\infty} g^n(t_0) \to \infty$$

$$\left.\begin{matrix} \\ \\ \end{matrix}\right\} P\left(\lim_{n\to\infty} \frac{S_n}{n} \geq \mu + \epsilon\right) = 0$$

Martingale convergence theorem

To show that for a given $\epsilon > 0$, $P\left(\lim_{n\to\infty}\frac{S_n}{n} \le \mu - \epsilon\right) = 0$

Leave as the exercise

$$\forall \epsilon > 0: P\left(\mu - \epsilon \le \lim_{n\to\infty}\frac{S_n}{n} \le \mu + \epsilon\right) = 1$$

$$\Downarrow$$

$$P\left(\lim_{n\to\infty}\frac{S_n}{n} = \mu\right) = 1$$

Summary

- Martingales
- Martingale stopping theorem
- Azuma's inequality for martingales
- Submartingales, supermartingales and martingale convergence theorem

References: Chapter 6, Martingales, 2nd edition, 1995, *by Sheldon M. Ross*