Theory of Stochastic Processes
10. Martingales

Tomonari Sei
sei@mist.i.u-tokyo.ac.jp

Department of Mathematical Informatics, University of Tokyo

June 22, 2017

http://www.stat.t.u-tokyo.ac.jp/~sei/lec.html

Handouts & Announcements

Handouts:

- Slides (this one)
- Copy of §12.1 and §12.2 of PRP

Outline today

1. Review of last week's material
2. Martingales
   - Examples
   - Doob martingale
   - Hoeffding's inequality
3. Recommended problems

Review of last week's material

The reasons why we learned about stationary processes.

- Why thinking stationarity?
  1. Output of MCMC is stationary.
  2. Statistical methods often assume independence of data (or error term). But this assumption is sometimes too strong. A tractable class of dependent data is the set of stationary processes.
- Why using spectrum?
  1. The spectral density (= power spectrum) is convenient for visualizing the stationary processes.
  2. Statistical inference reduces to a simpler form (e.g. Whittle likelihood).

Details are beyond the scope of this lecture.

Martingales

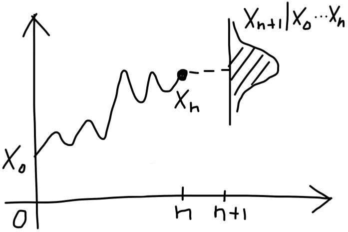
Today's topic

Definition

Let $\{X_n\}_{n\geq 0}$ be a process. A process $\{Y_n\}_{n\geq 0}$ is called a martingale with respect to $\{X_n\}$ if

$$E[Y_{n+1} \mid X_0, \ldots, X_n] = Y_n.$$

Typically $\{Y_n\} = \{X_n\}$.



Martingales are useful for describing fluctuation of random variables.

Review of conditional expectation

Important properties

Let X, Y, Z be random variables. Then

- Linearlity: $E[\alpha X + \beta Y | Z] = \alpha E[X|Z] + \beta E[Y|Z]$ for $\alpha, \beta \in \mathbb{R}$.
- If $X = f(Z)$ is a function of Z, then $E[XY|Z] = XE[Y|Z]$.
- If X and Z are independent, then $E[X|Z] = E[X]$.
- Tower property: $E[E[X|Y,Z]|Z] = E[X|Z]$.

Example (symmetric random walk is a martingale)

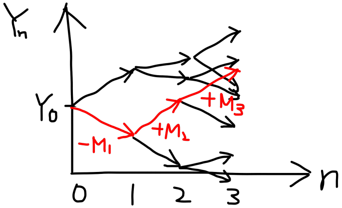
Let $Y_n = X_1 + \cdots + X_n$, where X_i are independent and $E[X_i] = 0$. Then

$$E[Y_n | X_1, \ldots, X_i] = Y_i.$$

$\longrightarrow$ blackboard

Example (betting game)

- Your initial capital is $Y_0 > 0$.
- For each $n \geq 1$,
  - You bet a money $M_n = M_n(X_1, \ldots, X_{n-1}) \geq 0$.
  - Let $X_n \in \{-1, 1\}$ be a Bernoulli trial.
  - Then $Y_n = Y_{n-1} + M_n X_n$.
- Then the capital process $\{Y_n\}$ is a martingale.



An advanced topic:

- Shafer and Vovk (2001). *Probability and Finance, It's only a Game!*, Wiley.

Remark

Two remarks.

1. Martingales are not necessarily Markov, and vice versa.
   For example, letting $\{X_i\}$ be independent and $E[X_i] = 0$,

	Markov	not Markov
martingale	$Y_n = Y_{n-1} + X_n$	$Y_n = Y_{n-1} + X_{n-1}X_n$
not martingale	$Y_n = Y_{n-1} + X_n + 1$	$Y_n = Y_{n-1} + X_{n-1}X_n + 1$

2. Y is called a submartingale if $Y_n \leq E[Y_{n+1}|X_0, \ldots, X_n]$ for all n.
   Y is called a supermartingale if $Y_n \geq E[Y_{n+1}|X_0, \ldots, X_n]$ for all n.

   Example: In the betting game, a fee may be charged in each round.

Seemingly artificial example: Doob martingale

For any function $S = f(X_1, \ldots, X_n)$ of a sequence $\{X_i\}_{i=0}^n$, the process

$$Y_i = E[S|X_1, \ldots, X_i], \quad 0 \leq i \leq n,$$

defines a martingale with respect to $\{X_i\}$. Here $Y_0 = E[S]$.

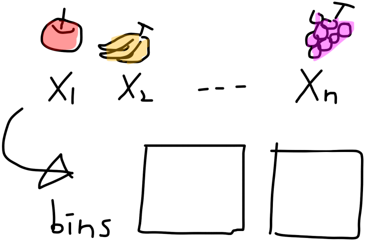
Exercise

Prove that $\{Y_i\}_{i=0}^n$ is indeed a martingale.

Definition

This martingale is called a Doob martingale.

Example: the bin packing problem (p.477 of PRP)



- Let $X_1, \ldots, X_n$ be the size of objects, assumed to be independent and distributed on $[0, 1]$.
- Let S be the minimum number of bins (of size 1) to pack them.
- The Doob martingale $Y_i = E[S|X_1, \ldots, X_i]$ is helpful as seen later.

Example: a martingale you may be familiar with

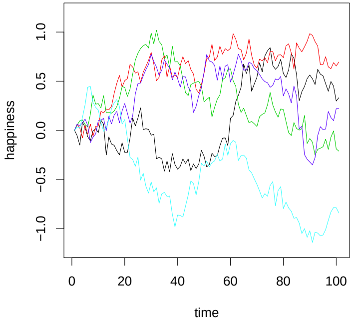
`http://www.tokyometro.jp/`

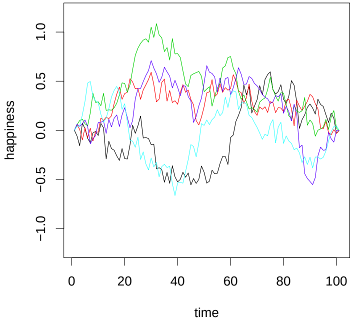
Search word	Number of hits	$E[S\|X_1, \ldots, X_i]$
(all)	142	$1/142$
K	21	$1/21$
Ki	6	$1/6$
Kit	3	$1/3$
Kita	3	$1/3$
Kita-	3	$1/3$
Kita-s	2	$1/2$
Kita-se	1	$1/1$
...		
Kita-senju	1	$1/1$

Let $S = 1$ if the search word $X = \{X_i\}$ is "Kita-senju" and 0 otherwise. Suppose that the distribution of X is uniform over all the stations.

Unnecessary addition (蛇足)

Which do you like better?





"Life is a martingale."

$$X_n$$

"Life is not a martingale."

$$X_n - \frac{n}{N} X_N$$

Hoeffding's inequality (p.476 of PRP)

An amazing result.

Theorem (Azuma-Hoeffding inequality)

Let Y_n be a martingale such that $|Y_n - Y_{n-1}| \leq 1$ for each n. Then

$$P(|Y_n - Y_0| \geq x) \leq 2\exp\left(-\frac{x^2}{2n}\right)$$

for any $x > 0$ and n.

- It means that Y_n is concentrated around Y_0.
- A sketch of proof will be given on the blackboard.

Application: large deviation

If $X_1, \ldots, X_n$ are i.i.d. with $|X_i| \leq 1$ and $E[X_i] = \mu$, then

$$P(|\bar{X} - \mu| \geq \varepsilon) \leq 2\exp\left(-\frac{n\varepsilon^2}{2}\right) \to 0 \quad (n \to \infty).$$

Application

McDiarmid's inequality

Let $X = (X_1, \dots, X_n)$ be a sequence of independent random variables. If

$$|f(x_1, \dots, x_i, \dots, x_n) - f(x_1, \dots, \tilde{x}_i, \dots, x_n)| \le 1, \quad \forall i, x_i, \tilde{x}_i,$$

then

$$P\left(|f(X) - E[f(X)]| \ge t\right) \le 2\exp\left(-\frac{t^2}{2n}\right).$$

Example (the bin packing problem; cont.)

For any fixed $\varepsilon > 0$, we have

$$P(|S - E[S]| \ge n\varepsilon) \le 2\exp\left(-\frac{n\varepsilon^2}{2}\right) \to 0 \quad (n \to \infty).$$

Chromatic number of random graphs $\rightarrow$ see §12.2, Problem 2.

Other topics we do not discuss in detail

- Let $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ be the "whole information" of $X_1, \ldots, X_n$. Mathematically, this is the smallest σ-field such that $X_1, \ldots, X_n$ are measurable. The sequence $\boldsymbol{\mathcal{F}} = \{\mathcal{F}_n\}_{n\geq 0}$ is called a filtration.
- A random variable T taking values in $\{0, 1, \cdots\}$ is called a stopping time if the event $\{T \leq n\}$ is $\mathcal{F}_n$-measurable for all n.
  - the time when Hayao started producing movies. $\to$ stopping time
  - the time when Hayao stops producing movies. $\to$ not stopping time

Optional sampling theorem

If Y is a martingale and T is a stopping time, then the "stopped" process $\{Y_{T\wedge n}\}$ is also a martingale. In particular, $E[Y_{T\wedge n}|Y_0] = Y_0$.

Example (absorbing barrier)

Let S_n be the symmetric simple random walk with $0 < S_0 < b$. Put $T = \inf\{n \mid S_n = 0 \text{ or } S_n = b\}$. Then $E[S_{T\wedge n}] = S_0$.

- Further topics: maximal inequality, convergence theorem etc.

Recommended problems

Recommended problems:

- §12.1, Problems 1, 2, 3, 4, 5, 6, 7*, 8, 9*.
- §12.2, Problems 1, 2.

The asterisk (*) shows difficulty.