Random Models of Dynamical Systems
Introduction to SDE's (1/5)

4GM–AROM

François Le Gland
INRIA Rennes + IRMAR

http://www.irisa.fr/aspi/legland/insa-rennes/

November 9, 2018

Introduction

Stochastic processes

Brownian motion

Continuous martingales

objective: find (and study) a continuous–time analogue to discrete–time stochastic models, such as

$$X_k = f(X_{k-1}, W_k)$$

where W_k's are independent (non necessarily Gaussian) random variables

shall we succeed? yes and no
concept of a *stochastic differential equation* (SDE)

$$dX(t) = b(X(t))\, dt + \sigma(X(t))\, dB(t)$$

where $dB(t)$'s are independent random variables, precisely: *Brownian motion* increments $B(t_n) - B(t_{n-1}), \cdots, B(t_1) - B(t_0)$ are independent random variables for any finite subset $t_0 < t_1 < \cdots < t_n$, and for any $0 \leq s \leq t$ the distribution of the r.v. $B(t) - B(s)$ depends only on $(t - s)$

interpretation as random perturbation of (ordinary) differential equation

$$\dot{X}(t) = b(X(t))$$

loss of generality: increments should *necessarily* be Gaussian + noise–dependence is additive

yet some benefit: stochastic differential calculus, e.g. *Itô formula* (chain rule) yields SDE for $\phi(X(t))$

$$d\phi(X(t)) = L\,\phi(X(t))\,dt + \phi'(X(t))\,\sigma(X(t))\,dB(t)$$

this is in constrast with discrete–time counterpart: indeed, if

$$X_k = f(X_{k-1}) + W_k$$

holds with additive noise, this structure is not preserved under mapping, i.e.

$$\phi(X_k) = \phi(f(X_{k-1}) + W_k)$$

does not exhibit additive noise structure

Introduction

Stochastic processes

Brownian motion

Continuous martingales

Stochastic processes

Definition a stochastic process is a collection $X = (X(t)\,,\, 0 \leq t \leq T)$ or $X = (X(t)\,,\, t \geq 0)$ of r.v.'s (measurable maps defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in a space $(E, \mathcal{E})$ (typically $E = \mathbb{R}^d$ with its Borel σ–field $\mathcal{E}$) indexed by $I = [0, T]$ or $I = [0, \infty)$ respectively

Definition finite–dimensional distributions of the stochastic process X are joint probability distributions of r.v.s such as $(X(t_1), \cdots, X(t_n))$ for any finite subset $t_1 < \cdots < t_n$ of indices, i.e.

$$\mu_{t_1 \cdots t_n}(A_1 \times \cdots \times A_n) = \mathbb{P}[X(t_1) \in A_1, \cdots, X(t_n) \in A_n]$$

Theorem 1* [Kolmogorov extension theorem] given the collection of finite–dimensional distributions defined for all possible finite subsets of I, there exists a unique probability distribution μ^X (called the probability distribution of the process X) on the set E^I (of all mappings defined on I and taking values in E), whose restriction (marginals) to any finite subset of indices coincides with the prescribed finite–dimensional distribution

in other words: the distribution of a stochastic process is completely characterized by the collection of all its finite–dimensional distributions

Definition a process X has almost surely continuous sample paths iff the set

$$\{\omega \in \Omega : \text{the mapping } t \mapsto X(t,\omega) \text{ is continuous on } I\}$$

has probability 1

in other words: a process with almost surely continuous sample paths on $I = [0, T]$ can be seen as a r.v. on the functional space $C([0,T],E)$ of continuous mappings

Theorem 2* [Kolmogorov continuity criterion] if there exist positive constants $\alpha, \beta > 0$ and $C > 0$ such that for any $t, s \geq 0$

$$\mathbb{E}|X(t) - X(s)|^{\beta} \leq C\,|t-s|^{1+\alpha}$$

then almost surely the process X has continuous sample paths

Introduction

Stochastic processes

Brownian motion

Continuous martingales

Subdivisions

Definition for any $n \geq 1$, let $0 = t_0^n < t_1^n < \cdots < t_n^n = t$ be a subdivision of $[0, t]$ with $\Delta_n = \max_{i=1\cdots n} (t_i^n - t_{i-1}^n)$

- a convergent subdivision is such that $\Delta_n \to 0$ as $n \uparrow \infty$
- a *fast* subdivision is any subsequence such that $\sum_{k=1}^{\infty} \Delta_{n(k)} < \infty$

Remark clearly, $\Delta_n \geq t/n$ hence

$$\sum_{n=1}^{\infty} \Delta_n \geq t \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

i.e. the condition does not hold without taking a subsequence

Remark the dyadic subdivision, with $n(k) = 2^k$ and $t_i^{(k)} = t\, i\, 2^{-k}$ for $i = 0 \cdots 2^k$, is a *fast* subdivision: indeed $\Delta_{n(k)} = t\, 2^{-k}$ and

$$\sum_{k=1}^{\infty} \Delta_{n(k)} = t \sum_{n=1}^{\infty} 2^{-k} = t < \infty$$

Brownian motion

Definition a Brownian motion B is a process

- with independent and stationary increments, i.e.
  for any finite subset $t_0 < t_1 < \cdots < t_n$ of indices the r.v.'s $B(t_n) - B(t_{n-1}), \cdots, B(t_1) - B(t_0)$ are independent, and for any $0 \leq s \leq t$ the distribution of the r.v. $B(t) - B(s)$ depends only on $(t - s)$
- with uniformly continuous sample paths on subdivisions of compact subsets of I, i.e. for any $0 \leq s \leq t$, for any convergent subdivision $s = t_0^n < t_1^n < \cdots < t_n^n = t$ of $[s, t]$ and for any $\delta > 0$

$$\mathbb{P}[\max_{i=1\cdots n} |B(t_i^n) - B(t_{i-1}^n)| > \delta] \rightarrow 0$$

as $n \uparrow \infty$

Remark * necessarily, such a process is Gaussian, and for any $0 \leq s \leq t$ the variance of the increment $B(t) - B(s)$ is proportional $(t - s)$

if X is a Gaussian r.v. with zero mean and variance σ^2, then $\mathbb{E}|X|^4 = 3\,\sigma^4$, hence

$$\mathbb{E}|B(t) - B(s)|^4 = C\,|t - s|^2$$

and it follows from the Kolmogorov criterion that a Brownian motion has almost surely continuous sample paths

Remark necessarily, these sample paths cannot be differentiable (even in a weak sense) since

$$\mathbb{E}|\frac{B(t+h) - B(t)}{h}|^2 = C\,\frac{1}{h}$$

does not have a finite limit as $h \downarrow 0$

this discussion justifies the following equivalent

Definition a Brownian motion B is a process

- with independent and Gaussian increments, i.e. for any finite subset $t_0 < t_1 < \cdots < t_n$ of indices the r.v.'s $B(t_n) - B(t_{n-1}), \cdots, B(t_1) - B(t_0)$ are independent, and for any $0 \leq s \leq t$ the distribution of the r.v. $B(t) - B(s)$ is $\mathcal{N}(0, (t-s)\,\sigma^2)$
- with almost surely continuous sample paths

without loss of generality, it is assumed that $B(0) = 0$, i.e. a Brownian motion starts at zero

if $\sigma^2 = 1$ in the definition, the Brownian motion is called a standard Brownian motion

Proposition 3 a process B is a Brownian motion iff B is a zero mean Gaussian process with correlation function

$$K(s,t) = \mathbb{E}[B(t)\,B(s)] = (s \wedge t)\,\sigma^2$$

and almost surely continuous sample paths

Proof 'only if' part: for any finite subset $t_0 < t_1 < \cdots < t_n$ of indices, the r.v. $(B(t_0), B(t_1), \cdots, B(t_n))$ is a linear transformation of the r.v. $(B(t_0) - B(0), B(t_1) - B(t_0), \cdots, B(t_n) - B(t_{n-1}))$ (a Gaussian r.v. since its components are Gaussian independent r.v.'s) hence it is Gaussian

clearly, if $0 \leq s \leq t$ then

$$\mathbb{E}[B(t)] = \mathbb{E}[B(t) - B(s)] + \mathbb{E}[B(s)] = \mathbb{E}[B(s)] = \mathbb{E}[B(0)] = 0$$

and

$$K(s,t) = \mathbb{E}[B(t)\,B(s)] = \mathbb{E}[(B(t) - B(s))\,B(s)] + \mathbb{E}|B(s)|^2 = s\,\sigma^2$$

'if' part: conversely, for any finite subset $t_0 < t_1 < \cdots < t_n$ of indices, the r.v. $(B(t_1) - B(t_0), \cdots, B(t_n) - B(t_{n-1}))$ is a linear transformation of the Gaussian r.v. $(B(t_0), B(t_1), \cdots, B(t_n))$ hence it is Gaussian

clearly, for any $i = 1 \cdots n$

$$\mathbb{E}[(B(t_i) - B(t_{i-1}))^2]$$

$$= K(t_i, t_i) - 2\,K(t_{i-1}, t_i) + K(t_{i-1}, t_{i-1})$$

$$= (t_i - 2\,t_{i-1} + t_{i-1})\,\sigma^2 = (t_i - t_{i-1})\,\sigma^2$$

and for any $i, j = 1 \cdots n$ with $i \neq j$, for instance $t_{j-1} < t_j \leq t_{i-1} < t_i$

$$\mathbb{E}[(B(t_j) - B(t_{j-1}))\,(B(t_i) - B(t_{i-1}))]$$

$$= K(t_j, t_i) - K(t_j, t_{i-1}) - K(t_{j-1}, t_i) + K(t_{j-1}, t_{i-1})$$

$$= (t_j - t_j + t_{j-1} - t_{j-1})\,\sigma^2 = 0$$

hence the Gaussian r.v.'s $B(t_n) - B(t_{n-1}), \cdots, B(t_1) - B(t_0)$ are independent

□

multi–dimensional version

Definition a d–dimensional Brownian motion B with $d \times d$ covariance matrix Σ is a process

- with independent and Gaussian increments, i.e.
  for any finite subset $t_0 < t_1 < \cdots < t_n$ of indices the r.v.'s $B(t_n) - B(t_{n-1}), \cdots, B(t_1) - B(t_0)$ are independent, and for any $0 \leq s \leq t$ the distribution of the r.v. $B(t) - B(s)$ is $\mathcal{N}(0, (t-s)\,\Sigma)$
- with almost surely continuous sample paths

Proposition * a process B is a d–dimensional Brownian motion with $d \times d$ covariance matrix Σ iff B is a zero mean Gaussian process with matrix–valued correlation function

$$K(s,t) = \mathbb{E}[B(t)\,B^*(s)] = (s \wedge t)\,\Sigma$$

and almost surely continuous sample paths

Exercise if B is a standard Brownian motion, then the processes defined by: *rescaling*

$$X(t) = \lambda\, B(\frac{t}{\lambda^2})$$

time inversion

$$X(t) = \begin{cases} t\, B(\frac{1}{t}) & \text{if } t > 0 \\ 0 & \text{if } t = 0 \end{cases}$$

refreshing

$$X(t) = B(t + t_0) - B(t_0)$$

time reversal for $0 \leq t \leq T$

$$X(t) = B(T - t) - B(T)$$

are also standard Brownian motions, i.e. have the same distribution as B

Quadratic variation

Proposition 4 [quadratic variation] let B be a standard Brownian motion and let $0 = t_0^n < t_1^n < \cdots < t_n^n = t$ be a convergent subdivision of $[0, t]$, then

$$V_n(t) = \sum_{i=1}^{n} (B(t_i^n) - B(t_{i-1}^n))^2 \to t$$

in $\mathbb{L}^2$ as $n \uparrow \infty$, and the convergence holds almost surely along a *fast* subdivision

Remark necessarily, Brownian motion sample paths cannot have finite variation since

$$V_n(t) \leq \max_{i=1\cdots n} |B(t_i^n) - B(t_{i-1}^n)| \sum_{i=1}^{n} |B(t_i^n) - B(t_{i-1}^n)|$$

Proof interpretation as a sum of independent zero–mean r.v.'s

$$V_n(t) - t = \sum_{i=1}^{n} [(B(t_i^n) - B(t_{i-1}^n))^2 - (t_i^n - t_{i-1}^n)]$$

expansion

$$|V_n(t) - t|^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} [(B(t_i^n) - B(t_{i-1}^n))^2 - (t_i^n - t_{i-1}^n)] [(B(t_j^n) - B(t_{j-1}^n))^2 - (t_j^n - t_{j-1}^n)]$$

and expectation yield

$$\mathbb{E}|V_n(t) - t|^2 = \sum_{i=1}^{n} \mathbb{E}|(B(t_i^n) - B(t_{i-1}^n))^2 - (t_i^n - t_{i-1}^n)|^2$$

if X is a Gaussian r.v. with zero mean and variance σ^2, then

$$\mathbb{E}|X^2 - \sigma^2|^2 = \mathbb{E}|X|^4 - \sigma^4 = 2\,\sigma^4$$

in particular for $X = B(t_i^n) - B(t_{i-1}^n)$, a Gaussian r.v. with zero mean and variance $\sigma^2 = t_i^n - t_{i-1}^n$, it holds

$$\mathbb{E}|(B(t_i^n) - B(t_{i-1}^n))^2 - (t_i^n - t_{i-1}^n)|^2 = 2\,(t_i^n - t_{i-1}^n)^2$$

hence

$$\begin{aligned}
\mathbb{E}|V_n(t) - t|^2 \;&=\; 2\sum_{i=1}^{n}(t_i^n - t_{i-1}^n)^2 \\
&\leq\; 2\sup_{i=1\cdots n}(t_i^n - t_{i-1}^n)\;\sum_{i=1}^{n}(t_i^n - t_{i-1}^n) \\
&=\; 2\,t\,\Delta_n \to 0
\end{aligned}$$

as $n \uparrow \infty$, which shows the first part it follows from the Markov inequality that for any $\delta > 0$

$$\mathbb{P}[|V_{n(k)}(t) - t| > \delta] \leq \frac{1}{\delta^2}\,\mathbb{E}|V_{n(k)}(t) - t|^2 \leq \frac{2\,t}{\delta^2}\,\Delta_{n(k)}$$

just as in the Borel–Cantelli lemma, notice that the events

$$A_p = \bigcup_{k \geq p} \{|V_{n(k)}(t) - t| > \delta\}$$

form a non–increasing sequence, i.e. $A_p \subseteq A_{p-1}$, hence

$$\begin{aligned}
\mathbb{P}[\bigcap_{p\geq 1} \bigcup_{k\geq p} \{|V_{n(k)}(t) - t| > \delta\}] &\leq \lim_{p\uparrow\infty} \mathbb{P}[\bigcup_{k\geq p} \{|V_{n(k)}(t) - t| > \delta\}] \\
&\leq \lim_{p\uparrow\infty} \sum_{k\geq p} \mathbb{P}[|V_{n(k)}(t) - t| > \delta] \\
&\leq \frac{2\,t}{\delta^2} \lim_{p\uparrow\infty} \sum_{k\geq p} \Delta_{n(k)} = 0
\end{aligned}$$

as $q \uparrow \infty$, hence

$$\mathbb{P}[\bigcup_{p\geq 1} \bigcap_{k\geq p} \{|V_{n(k)}(t) - t| \leq \delta\}] = 1 \qquad \square$$

Corollary 5 let B be a standard Brownian motion, and let $0 = t_0^n < t_1^n < \cdots < t_n^n = t$ be a convergent subdivision of $[0, t]$, then

$$\sum_{i=1}^{n} \tfrac{1}{2}\left(B(t_i^n) + B(t_{i-1}^n)\right)\left(B(t_i^n) - B(t_{i-1}^n)\right) = \tfrac{1}{2}\, B^2(t)$$

and

$$\sum_{i=1}^{n} B(t_{i-1}^n)\left(B(t_i^n) - B(t_{i-1}^n)\right) \to \tfrac{1}{2}\left(B^2(t) - t\right)$$

in $\mathbb{L}^2$ as $n \uparrow \infty$, and the convergence holds almost surely along a *fast* subdivision

Proof interpretation as a telescopic sum yields

$$\sum_{i=1}^{n}(B(t_i^n)+B(t_{i-1}^n))\,(B(t_i^n)-B(t_{i-1}^n))$$

$$=\sum_{i=1}^{n}(B^2(t_i^n)-B^2(t_{i-1}^n))=B^2(t_n^n)-B^2(t_0^n)=B^2(t)$$

therefore, using the identity

$$x=\tfrac{1}{2}\,(x'+x)-\tfrac{1}{2}\,(x'-x)$$

yields

$$\sum_{i=1}^{n}B(t_{i-1}^n)\,(B(t_i^n)-B(t_{i-1}^n))$$

$$=\tfrac{1}{2}\sum_{i=1}^{n}(B(t_i^n)+B(t_{i-1}^n))\,(B(t_i^n)-B(t_{i-1}^n))$$

$$-\tfrac{1}{2}\sum_{i=1}^{n}(B(t_i^n)-B(t_{i-1}^n))^2 \qquad \square$$

multi–dimensional version

Proposition 6 [quadratic co–variation] let B be a d–dimensional Brownian motion with covariance matrix Σ, and let $0 = t_0^n < t_1^n < \cdots < t_n^n = t$ be a convergent subdivision of $[0, t]$, then

$$V_n(t) = \sum_{i=1}^{n} (B(t_i^n) - B(t_{i-1}^n))\,(B(t_i^n) - B(t_{i-1}^n))^* \to t\,\Sigma$$

in $\mathbb{L}^2$ as $n \uparrow \infty$, and the convergence holds almost surely along a *fast* subdivision

Proof for any $u \in \mathbb{R}^d$, the one–dimensional process $u^* B(t)$ is a Brownian motion with variance $\sigma^2 = u^* \Sigma u$, hence

$$
\begin{aligned}
u^* V_n(t)\, u &= \sum_{i=1}^{n} (u^* (B(t_i^n) - B(t_{i-1}^n)))^2 \\
&= \sum_{i=1}^{n} (\frac{u^* (B(t_i^n) - B(t_{i-1}^n))}{\sigma})^2 \, u^* \Sigma u \\
&\to t\, u^* \Sigma u
\end{aligned}
$$

and by polarization, for any $u, v \in \mathbb{R}^d$

$$
u^* V_n(t)\, v \to t\, u^* \Sigma v
$$

in $\mathbb{L}^2$ as $n \uparrow \infty$, and the convergence holds almost surely along a *fast* subdivision □

Introduction

Stochastic processes

Brownian motion

Continuous martingales

Filtrations

Definition a filtration is a non–decreasing collection $\mathcal{F} = (\mathcal{F}(t)\,,\, t \geq 0)$ of σ–algebras, and a stochastic process $X = (X(t)\,,\, t \geq 0)$ is said adapted w.r.t. $\mathcal{F}$ (or simply adapted) if for any $t \geq 0$ the r.v. $X(t)$ is measurable w.r.t. $\mathcal{F}(t)$

Definition an adapted standard Brownian motion B is a process

- with independent and Gaussian increments, i.e. for any $0 \leq s \leq t$ the r.v. $B(t) - B(s)$ is independent of $\mathcal{F}(s)$ and its distribution is $\mathcal{N}(0, (t - s))$

Martingales

Definition a stochastic process $M = (M(t)\,,\, t \geq 0)$ is a martingale (or a submartingale, or a supermartingale), iff

- it is adapted and integrable, i.e. for any $t \geq 0$ the r.v. $M(t)$ is measurable w.r.t. $\mathcal{F}(t)$ and $\mathbb{E}|M(t)| < \infty$
- for any $0 \leq s \leq t$
$$\mathbb{E}[M(t) \mid \mathcal{F}(s)] = M(s)$$
(or
$$\mathbb{E}[M(t) \mid \mathcal{F}(s)] \geq M(s) \quad \text{or} \quad \mathbb{E}[M(t) \mid \mathcal{F}(s)] \leq M(s)$$
respectively)

Proposition 7 let M be martingale and ϕ be a convex function if the process N defined by $N(t) = \phi(M(t))$ is integrable, then it is a submartingale

Proof for any $0 \leq s \leq t$, the Jensen inequality yields

$$\begin{aligned} \mathbb{E}[N(t) \mid \mathcal{F}(s)] &= \mathbb{E}[\phi(M(t)) \mid \mathcal{F}(s)] \\ &\geq \phi(\mathbb{E}[M(t) \mid \mathcal{F}(s)]) = \phi(M(s)) = N(s) \qquad \square \end{aligned}$$

Example let B be a Brownian motion, then B and the processes M and Z defined by

$$M(t) = B^2(t) - t \qquad \text{and} \qquad Z(t) = \exp\{\lambda\, B(t) - \tfrac{1}{2}\, \lambda^2\, t\}$$

are martingales

Proof for any $0 \le s \le t$, the r.v. $B(t) - B(s)$ is zero mean and is independent of $\mathcal{F}(s)$, hence

$$\mathbb{E}[B(t) \mid \mathcal{F}(s)] - B(s) = \mathbb{E}[B(t) - B(s) \mid \mathcal{F}(s)] = 0$$

for any $0 \le s \le t$

$$M(t) - M(s) = (B(t) - B(s))^2 - (t - s) + 2\,B(s)\,(B(t) - B(s))$$

and the r.v. $B(t) - B(s)$ is zero mean with variance $(t - s)$ and is independent of $\mathcal{F}(s)$, hence

$$\mathbb{E}[M(t) \mid \mathcal{F}(s)] - M(s) = \mathbb{E}[M(t) - M(s) \mid \mathcal{F}(s)]$$

$$= \mathbb{E}[(B(t) - B(s))^2 \mid \mathcal{F}(s)] - (t - s)$$

$$+ 2\,B(s)\,\mathbb{E}[B(t) - B(s) \mid \mathcal{F}(s)]$$

$$= 0$$

for any $0 \leq s \leq t$

$$Z(t) = \exp\{\lambda\,(B(t) - B(s))\}\ \exp\{-\tfrac{1}{2}\,\lambda^2\,(t-s)\}\ Z(s)$$

and the r.v. $B(t) - B(s)$ is Gaussian, with zero mean and variance $(t-s)$ and is independent of $\mathcal{F}(s)$, hence the Laplace transform

$$\mathbb{E}[\exp\{\lambda\,(B(t) - B(s))\} \mid \mathcal{F}(s)] = \exp\{\tfrac{1}{2}\,\lambda^2\,(t-s)\}$$

and

$$\mathbb{E}[Z(t) \mid \mathcal{F}(s)] \ = \ \mathbb{E}[\exp\{\lambda\,(B(t) - B(s))\} \mid \mathcal{F}(s)]$$

$$\exp\{-\tfrac{1}{2}\,\lambda^2\,(t-s)\}\ Z(s)$$

$$= \ Z(s) \qquad \square$$

Doob inequality

Theorem 8 [Doob maximal inequality] let M be a continuous martingale with finite p–th moments (i.e. $\mathbb{E}|M(t)|^p < \infty$ for any $t \geq 0$) for some $p > 1$, then for any $\lambda > 0$

$$\mathbb{P}[\max_{0\leq s\leq t} |M(s)| \geq \lambda] \leq \frac{1}{\lambda^p} \mathbb{E}|M(t)|^p$$

Remark the maximum is controlled in term of the final value, i.e. uniform control holds in term of the final value

Remark this inequality generalizes the Markov inequality valid in the static case for a single square integrable r.v.

Doob maximal inequality is a consequence of the following

Proposition 9 let X be a continuous non–negative submartingale, then for any $\lambda > 0$

$$\mathbb{P}[\max_{0\leq s\leq t} X(s) \geq \lambda] \leq \frac{1}{\lambda} \mathbb{E}[X(t)\, 1_{\{\max_{0\leq s\leq t} X(s) \geq \lambda\}}] \leq \frac{1}{\lambda} \mathbb{E}[X(t)]$$

Proof of Doob maximal inequality (as a consequence of the Proposition) if M is a continuous martingale with finite p–th moments, then $|M|^p$ is a continuous non–negative submartingale, and applying the Proposition yields

$$\mathbb{P}[\max_{0 \leq s \leq t} |M(s)| \geq \lambda] = \mathbb{P}[\max_{0 \leq s \leq t} |M(s)|^p \geq \lambda^p] \leq \frac{1}{\lambda^p} \mathbb{E}|M(t)|^p$$

Proof of the Proposition the estimate if first proved for the maximum over any finite subdivision $0 = t_0^n < t_1^n < \cdots < t_n^n = t$ of $[0, t]$

the submartingale property yields

$$\mathbb{E}[X(t) \mid \mathcal{F}(t_i^n)] \geq X(t_i^n)$$

let $K = \min\{i = 0 \cdots n : X(t_i^n) \geq \lambda\}$ or $K = +\infty$ if such an index does not exist, clearly $\{K = i\} \in \mathcal{F}(t_i^n)$ and

$$\mathbb{E}[1_{\{K = i\}} X(t_i^n)] \geq \lambda \, \mathbb{P}[K = i]$$

$$
\begin{aligned}
\mathbb{P}[\max_{i=0\cdots n} X(t_i^n) \geq \lambda] = \mathbb{P}[K \leq n] &= \sum_{i=0}^{n} \mathbb{P}[K = i] \\
&\leq \frac{1}{\lambda} \sum_{i=0}^{n} \mathbb{E}[1_{\{K = i\}} X(t_i^n)] \\
&\leq \frac{1}{\lambda} \sum_{i=0}^{n} \mathbb{E}[1_{\{K = i\}} \mathbb{E}[X(t) \mid \mathcal{F}(t_i^n)]] \\
&= \frac{1}{\lambda} \sum_{i=0}^{n} \mathbb{E}[1_{\{K = i\}} X(t)] \\
&= \frac{1}{\lambda} \mathbb{E}[X(t)\, 1_{\{K \leq n\}}] \\
&\leq \frac{1}{\lambda} \mathbb{E}[X(t)\, 1_{\{\max_{0 \leq s \leq t} X(s) \geq \lambda\}}]
\end{aligned}
$$

notice that the dyadic subdivision at level k is a refined subdivision of the dyadic subdivision at coarser level $(k-1)$, since

$$\{t\, i\, 2^{-(k-1)}\,,\; i = 0 \cdots 2^{k-1}\} \;=\; \{t\, i\, 2^{-k}\,,\; i = 0 \cdots 2^{k}\,,\ \text{for even } i\}$$

$$\subset\; \{t\, i\, 2^{-k}\,,\; i = 0 \cdots 2^{k}\}$$

hence the events

$$A_k = \{\max_{i=0\cdots 2^k} X(t_i^{(k)}) \geq \lambda\}$$

form a non–decreasing sequence, i.e. $A_k \supseteq A_{k-1}$

furthermore, continuity of sample paths yields

$$\mathbb{P}[\max_{0 \leq s \leq t} X(s) \geq \lambda] \;=\; \mathbb{P}[\bigcup_{k \geq 1} \{\max_{i=0\cdots 2^k} X(t_i^{(k)}) \geq \lambda\}\,]$$

$$= \lim_{k \uparrow \infty} \mathbb{P}[\max_{i=0\cdots 2^k} X(t_i^{(k)}) \geq \lambda]$$

$$\leq \frac{1}{\lambda}\, \mathbb{E}[X(t)\, 1_{\{\max_{0 \leq s \leq t} X(s) \geq \lambda\}}] \leq \frac{1}{\lambda}\, \mathbb{E}[X(t)] \qquad \square$$

Corollary 10 [Doob inequality] let M be a continuous martingale with finite p–th moments (i.e. $\mathbb{E}|M(t)|^p < \infty$ for any $t \geq 0$) for some $p > 1$, then

$$\{\mathbb{E}(\max_{0 \leq s \leq t} |M(s)|)^p\}^{1/p} \leq \frac{p}{p-1} \{\mathbb{E}|M(t)|^p\}^{1/p}$$

Lemma let Y and Z be two non–negative r.v.'s such that for any $\lambda > 0$

$$\mathbb{P}[Y \geq \lambda] \leq \frac{1}{\lambda} \mathbb{E}[Z \, 1_{\{Y \geq \lambda\}}]$$

let F be a continuous non–decreasing function defined on $[0, \infty)$ (hence F has finite variation) and null at 0, then

$$\mathbb{E}[F(Y)] \leq \mathbb{E}[Z \int_0^Y \frac{1}{\lambda} F(d\lambda)]$$

in particular, if Z has finite p–th moments, then

$$\{\mathbb{E}[Y^p]\}^{1/p} \leq \frac{p}{p-1} \{\mathbb{E}[Z^p]\}^{1/p}$$

Proof by definition

$$
\begin{aligned}
\mathbb{E}[F(Y)] &= \mathbb{E}[\int_0^Y F(d\lambda)] \\
&= \mathbb{E}[\int_0^\infty 1_{\{0 \le \lambda \le Y\}} F(d\lambda)] \\
&= \int_0^\infty \mathbb{P}[Y \ge \lambda] \, F(d\lambda) \\
&\le \int_0^\infty \frac{1}{\lambda} \mathbb{E}[Z \, 1_{\{Y \ge \lambda\}}] \, F(d\lambda) \\
&= \mathbb{E}[Z \int_0^\infty \frac{1}{\lambda} 1_{\{Y \ge \lambda\}} F(d\lambda)] \\
&= \mathbb{E}[Z \int_0^Y \frac{1}{\lambda} F(d\lambda)]
\end{aligned}
$$

in particular for $F(\lambda) = \lambda^p$, it holds

$$\mathbb{E}[Y^p] \leq p\,\mathbb{E}[Z \int_0^Y \frac{1}{\lambda}\,\lambda^{p-1}\,d\lambda] = \frac{p}{p-1}\,\mathbb{E}[Z\,Y^{p-1}]$$

the Hölder inequality with conjugate exponents p, p' yields

$$\mathbb{E}[Z\,Y^{p-1}] \leq \{\mathbb{E}[Z^p]\}^{1/p}\,\{\mathbb{E}[Y^{(p-1)\,p'}]\}^{1/p'} = \{\mathbb{E}[Z^p]\}^{1/p}\,\{\mathbb{E}[Y^p]\}^{1/p'}$$

since $(p-1)\,p' = p$, and finally

$$\mathbb{E}[Y^p] \leq \frac{p}{p-1}\,\mathbb{E}[Z\,Y^{p-1}] \leq \frac{p}{p-1}\,\{\mathbb{E}[Z^p]\}^{1/p}\,\{\mathbb{E}[Y^p]\}^{1/p'}$$

or equivalently

$$\{\mathbb{E}[Y^p]\}^{1/p} \leq \frac{p}{p-1}\,\{\mathbb{E}[Z^p]\}^{1/p} \qquad \square$$

Proof of Doob inequality if M is a continuous martingale (with finite p–th moments), then $|M|$ is a continuous non–negative submartingale (also with finite p–th moments), hence

$$\mathbb{P}[\max_{0\leq s\leq t} |M(s)| \geq \lambda] \leq \frac{1}{\lambda}\,\mathbb{E}[\,|M(t)|\,1_{\{\max_{0\leq s\leq t}|M(s)|\geq\lambda\}}]$$

and the result follows from applying the Lemma with

$$Y = \max_{0\leq s\leq t} |M(s)| \qquad \text{and} \qquad Z = |M(t)| \qquad \square$$

Stopping times

Definition a stopping time τ is a r.v. with values in $[0, +\infty) \cup \{+\infty\}$ such that for all $t \geq 0$

$$\{\tau \leq t\} \in \mathcal{F}(t)$$

i.e. whether $\tau \leq t$ or not, can be decided given events up to time t

Example let X be a continuous process with values in $\mathbb{R}^d$ and let $F \subseteq \mathbb{R}^d$ be a closed subset, then the hitting time

$$\tau_F = \begin{cases} \inf\{t \geq 0 \,:\, X(t) \in F\} & \text{if such a time exists} \\ +\infty & \text{otherwise} \end{cases}$$

is a stopping time

Definition the σ–algebra of *events determined prior to the stopping time* τ is defined by: $A \in \mathcal{F}(\tau)$ iff for any $t \geq 0$

$$A \cap \{\tau \leq t\} \in \mathcal{F}(t)$$

Theorem 11 [optional sampling] let M be a martingale (or a submartingale), and let $0 \leq \sigma \leq \tau \leq \text{cst} < \infty$ be two bounded stopping times, then

$$\mathbb{E}[M(\tau) \mid \mathcal{F}(\sigma)] = M(\sigma)$$

(or

$$\mathbb{E}[M(\tau) \mid \mathcal{F}(\sigma)] \geq M(\sigma)$$

respectively)

Proof assume $0 \leq \sigma \leq \tau \leq T < \infty$, and let $0 = t_0^n < t_1^n < \cdots < t_n^n = T$ be a convergent subdivision of $[0, T]$, so that the sequence defined by $M_k^n = M(t_k^n)$ is a discrete–time martingale for the filtration $\mathcal{F}_k^n = \mathcal{F}(t_k^n)$

clearly, the r.v. defined by

$$\tau_n = \begin{cases} t_k^n & \text{if } t_{k-1}^n < \tau \leq t_k^n \\ t_1^n & \text{if } \tau \leq t_1^n \end{cases}$$

is a stopping time, since $\tau_n \geq \tau$ (hence $\{\tau_n \leq t\} \subseteq \{\tau \leq t\} \in \mathcal{F}(t)$), and moreover $\tau_n \to \tau$ almost surely as $n \uparrow \infty$, since $0 \leq \tau_n - \tau \leq \Delta_n$, so that $M(\tau_n) \to M(\tau)$ almost surely as $n \uparrow \infty$ by continuity of the sample paths

note also that the r.v. defined by

$$K = \begin{cases} k & \text{if } t_{k-1}^n < \tau \leq t_k^n \\ 1 & \text{if } \tau \leq t_1^n \end{cases}$$

is a stopping time (for the discrete–time filtration), and $M(\tau_n) = M_K^n$

similarly, let

$$\sigma_n = \begin{cases} t_k^n & \text{if } t_{k-1}^n < \sigma \leq t_k^n \\ t_1^n & \text{if } \sigma \leq t_1^n \end{cases}$$

so that $M(\sigma_n) \to M(\sigma)$ almost surely as $n \uparrow \infty$, and let

$$J = \begin{cases} k & \text{if } t_{k-1}^n < \sigma \leq t_k^n \\ 1 & \text{if } \sigma \leq t_1^n \end{cases}$$

so that $M(\sigma_n) = M_J^n$

clearly $0 \leq \sigma_n \leq \tau_n \leq T$ and $1 \leq J \leq K \leq n$, and the optional sampling theorem for discrete–time martingales yields

$$\mathbb{E}[M(\tau_n)] = \mathbb{E}[M_K^n] = \mathbb{E}[M_J^n] = \mathbb{E}[M(\sigma_n)]$$

the optional sampling theorem for discrete–time martingales yields also

$$\mathbb{E}[M(T) \mid \mathcal{F}(\tau_n)] = \mathbb{E}[M_n^n \mid \mathcal{F}_K^n] = M_K^n = M(\tau_n)$$

hence the sequence $M(\tau_n)$ is uniformly integrable, and similarly the sequence $M(\sigma_n)$ is uniformly integrable, therefore $\mathbb{E}[M(\tau_n)] \to \mathbb{E}[M(\tau)]$ and similarly $\mathbb{E}[M(\sigma_n)] \to \mathbb{E}[M(\sigma)]$ as $n \uparrow \infty$, and uniqueness of the limit yields

$$\mathbb{E}[M(\tau)] = \mathbb{E}[M(\sigma)]$$

this identity holds for any stopping times σ and τ such that $0 \leq \sigma \leq \tau \leq T < \infty$ holds
notice that for any $B \in \mathcal{F}(\sigma)$, the r.v.'s

$$\sigma_B = \sigma \, 1_B + T \, 1_{B^c} \qquad \text{and} \qquad \tau_B = \tau \, 1_B + T \, 1_{B^c}$$

are stopping times, since $\sigma_B \geq \sigma$ and $\tau_B \geq \tau$, and the condition $0 \leq \sigma_B \leq \tau_B \leq T < \infty$ holds, hence

$$\mathbb{E}[M(\tau) \, 1_B] + \mathbb{E}[M(T) \, 1_{B^c}] = \mathbb{E}[M(\sigma) \, 1_B] + \mathbb{E}[M(T) \, 1_{B^c}]$$

or equivalently

$$\mathbb{E}[M(\tau) \mid \mathcal{F}(\sigma)] = M(\sigma) \qquad \square$$

Corollary 12 let M be a martingale (or a submartingale), and let $0 \leq s \leq \tau \leq \text{cst} < \infty$ for a bounded stopping time τ, then

$$\mathbb{E}[M(\tau) \mid \mathcal{F}(s)] = M(s)$$

(or

$$\mathbb{E}[M(\tau) \mid \mathcal{F}(s)] \geq M(s)$$

respectively)

Theorem 13 [stopped martingale] let M be a martingale (or a submartingale) and let τ be a (not necessarily finite) stopping time, then the stopped process

$$X(t) = M(t \wedge \tau) = \begin{cases} M(t) & \text{if } \tau \geq t \\ M(\tau) & \text{if } \tau \leq t \end{cases}$$

is a martingale (or a submartingale, respectively)

Proof let $t \geq s$ and notice that $\{\tau \leq s\} \in \mathcal{F}(s)$, hence

$$
\begin{aligned}
1_{\{\tau \leq s\}} \mathbb{E}[M(t \wedge \tau) \mid \mathcal{F}(s)] &= \mathbb{E}[1_{\{\tau \leq s\}} M(t \wedge \tau) \mid \mathcal{F}(s)] \\
&= \mathbb{E}[1_{\{\tau \leq s\}} M(s \wedge \tau) \mid \mathcal{F}(s)] \\
&= 1_{\{\tau \leq s\}} M(s \wedge \tau)
\end{aligned}
$$

on the other hand, the optional sampling theorem for the bounded stopping time $s \leq (t \wedge \tau) \vee s \leq t$ yields

$$
\begin{aligned}
1_{\{\tau > s\}} \mathbb{E}[M(t \wedge \tau) \mid \mathcal{F}(s)] &= \mathbb{E}[1_{\{\tau > s\}} M(t \wedge \tau) \mid \mathcal{F}(s)] \\
&= \mathbb{E}[1_{\{\tau > s\}} M((t \wedge \tau) \vee s) \mid \mathcal{F}(s)] \\
&= 1_{\{\tau > s\}} M(s) \\
&= 1_{\{\tau > s\}} M(s \wedge \tau) \qquad \square
\end{aligned}
$$

Definition a stochastic process X is a local martingale if there exists a non–decreasing sequence of stopping times τ^n such that

- the sequence $\tau^n \to \infty$ almost surely as $n \uparrow \infty$
- for any $n \geq 1$, the stopped process X^n defined by $X^n(t) = X(t \wedge \tau^n)$ is a uniformly integrable martingale