Stochastic Calculus

(Ref: [2015][O. Calin] An informal introduction to stochastic calculus w/ Applications).

2.9 Expectation.

Prop 2.9.2

X, Y: indep. r.v.

Then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

Def X, Y: r.v.

$\mathrm{Cov}(X,Y) := \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$, called the covariance of X and Y.

$\mathrm{Var}(X) := \mathrm{Cov}(X,X)$, called the variance of X.

Rmk:

① If X, Y: indep., then $\mathrm{Cov}(X,Y) = 0$.

② In general, X, Y: uncorrelated doesn't imply X, Y: indep. (see Exercise 2.9.6).

③ $\mathrm{Cov}(X,Y) = \mathbb{E}[(X-\mu_X)(Y-\mu_Y)]$ and hence $\mathrm{Var}(X) = \mathbb{E}[(X-\mu_X)^2]$.

Prop (Exercise 2.9.5).

μ: mean of X, σ: standard deviation of X.

Then $\mathbb{E}[X^2] = \mu^2 + \sigma^2$.

Def X, Y: r.v.

The correlation coefficient of X and Y is

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}.$$

Prop (Exercise 2.9.7).

X, Y: r.v. w/ $\rho(X,Y)$ well-defined.

Then ① $-1 \le \rho(X,Y) \le 1$ and

② $\rho(X,Y) = 1 \iff X = \lambda Y$, some $\lambda > 0$.

$\rho(X,Y) = -1 \iff X = \lambda Y$, some $\lambda < 0$.

<Pf>

Schwarz #

2.9.1 The best app. of a random variable.

Q: What is the best (nonrandom) number that app. a r.v. X in the "least square" sense?

A:

Let x be a nonrandom number.

$\mu := \mathbb{E}[X]$ and $\sigma^2 := \mathrm{Var}(X)$.

We want to find $\underset{x}{\operatorname{argmin}}\, \mathbb{E}[(X-x)^2]$.

$$\mathbb{E}[(X-x)^2] = \mathbb{E}[X^2] - 2x\,\mathbb{E}[X] + x^2$$

$$= \sigma^2 + \mu^2 - 2x\mu + x^2.$$

Thus, it is min. iff $2x - 2\mu = 0$ iff $x = \mu$.

Therefore, the mean μ of X is the best app. of X in the "least square" sense. #

2.10 Basic Dist.

Come back whenever necessary.

2.11 Sum of R.V.s.

Def

Given $f: [0,\infty) \to \mathbb{R}$.

The Laplace transform of f is

$$\mathcal{L}\{f\}(s) := \int_0^\infty e^{-sx} f(x)\,dx.$$

Rmk:

X: r.v. w/ values in $[0,\infty)$, and pdf $f(x)$.

Then $m_X(-s) = \mathbb{E}[e^{-sX}] = \int_0^\infty e^{-sx} f(x)\,dx$

$= \mathcal{L}\{f\}(s)$.

Prop (Exercise 2.11.2)

Given $f, g: [0,\infty) \to \mathbb{R}$.

Then $\mathcal{L}\{f*g\}(s) = \left(\mathcal{L}\{f\}(s)\right)\cdot\left(\mathcal{L}\{g\}(s)\right)$,

where $(f*g)(x) := \int_0^x f(x-z)g(z)\,dz$, is the convolution of f and g.

<Pf>

$$\mathcal{L}\{f*g\}(s) = \int_0^\infty \left[\int_0^x f(x-z)g(z)\,dz\right] e^{-sx}\,dx.$$

Use Fubini to change the order of integration. #

Thm 2.11.1

X, Y : nonnegative indep. r.v. w/ pdf f and g.

$h :=$ pdf of $X+Y$.

Then $h(t) = (f*g)(t), \ \forall t \geq 0$.

<Pf>

$$m_{X+Y}(-s) = \mathbb{E}[e^{-s(X+Y)}] \overset{\text{indep. of } X \text{ and } Y}{=} \mathbb{E}[e^{-sX}] \cdot \mathbb{E}[e^{-sY}]$$

$$= m_X(-s)\, m_Y(-s) = (\mathcal{L}\{f\}(s)) \cdot (\mathcal{L}\{g\}(s))$$

$$\overset{\text{by prop above}}{=} \mathcal{L}\{(f*g)\}(s). \ \#$$

2.12 Conditional Expectations.

Def

X : r.v. in $L^2(\Omega, \mathcal{F}, P)$.

$\mathcal{G}$: a σ-field in $\mathcal{F}$.

The conditional expectation of X given $\mathcal{G}$ is

$$\mathbb{E}[X \mid \mathcal{G}] := \underset{Z \in L^2(\Omega, \mathcal{G}, P)}{\operatorname{argmin}} \mathbb{E}[(X-Z)^2].$$

For $\mathbb{E}[X \mid \mathcal{G}]$ to be well-defined, we need to prove its existence and uniqueness.

Thm 6.13 (Projection) (in Hunter's "Applied Analysis")

M : a closed subset of a Hilbert space $\mathcal{H}$.

Then

(a) $\forall x \in \mathcal{H}, \exists!$ closest $y \in M$ s.t.

$$\|x-y\| \leq \|x-z\|, \ \forall z \in M.$$

(b) The y in (a) is the unique element of M w/ $x-y \perp M$. This y is called the orthogonal projection of x onto M.

Prop

$X \in L^2(\Omega, \mathcal{F}, P)$, $\mathcal{G} \subseteq \mathcal{F}$.

Then $\mathbb{E}[X \mid \mathcal{G}]$ is the orthogonal projection of X onto $L^2(\Omega, \mathcal{G}, P)$.

<Pf>

Note that $L^2(\Omega, \mathcal{G}, P)$ is a closed subset of the Hilbert space $L^2(\Omega, \mathcal{F}, P)$.

Thus, the existence and uniqueness of $\mathbb{E}[X \mid \mathcal{G}]$ follows from (a) in Thm 6.13. #

Q : What if X is simply integrable but not L^2?

Can we still define $\mathbb{E}[X \mid \mathcal{G}]$?

Prop

$X \in L^2(\Omega, \mathcal{F}, P)$, $\mathcal{G} \subseteq \mathcal{F}$.

Then $\mathbb{E}[X \mid \mathcal{G}]$ is the unique r.v.

s.t. ① $\mathbb{E}[X \mid \mathcal{G}]$ is $\mathcal{G}$-mea.

② $\int_A \mathbb{E}[X \mid \mathcal{G}]\, dP = \int_A X \, dP, \ \forall A \in \mathcal{G}$.

<Pf>

① follows from def. $\#_1$

For ②, let $Z = \mathbb{E}[X \mid \mathcal{G}]$ for simplicity.

By (b) in Thm 6.13, $X - Z \perp 1_A, \ \forall A \in \mathcal{G}$.

i.e. $\int_\Omega (X-Z) \cdot 1_A \, dP = 0, \ \forall A \in \mathcal{G}$.

i.e. $\int_A X - Z \, dP = 0, \ \forall A \in \mathcal{G}$.

i.e. $\int_A X \, dP = \int_A Z \, dP = \int_A \mathbb{E}[X \mid \mathcal{G}]\, dP, \ \forall A \in \mathcal{G}$ $\#_2$

For uniqueness, let Z' satisfy ① and ②.

Then $\int_A Z - Z' \, dP = 0, \ \forall A \in \mathcal{G}$.

By Exercise 2.12.1, $Z = Z'$ $\mathcal{G}$-a.s. $\#_3$

see next page

Rmk :

By Prop above, the two properties characterize $\mathbb{E}[X \mid \mathcal{G}]$ for $X \in L^2(\Omega, \mathcal{F}, P)$.

We may use these two properties to define $\mathbb{E}[X \mid \mathcal{G}]$ for $X \in L^1(\Omega, \mathcal{F}, P)$.

Def X : r.v. in $L^1(\Omega, \mathcal{F}, P)$. $\mathcal{G} \subseteq \mathcal{F}$.

The conditional expectation of X given $\mathcal{G}$ is the $\mathcal{G}$-mea. r.v. satisfying the 2 properties in the Prop above.

Still, existence and uniqueness need to be proved.

This time, the existence comes from the Radon-Nikodym Thm. The uniqueness follows from Exercise 2.12.1.

Thm 2.12.2 (Radon-Nikodym).

$(\Omega, \mathcal{F}, P)$: prob. sp. $\mathcal{G} \subseteq \mathcal{F}$.

$X \in L^1(\Omega, \mathcal{F}, P)$.

Then $\exists\ Y \in L^1(\Omega, \mathcal{G}, P)$ s.t.

$$\int_A X\,dP = \int_A Y\,dP, \ \forall A \in \mathcal{G}.$$

Exercise 2.12.1

$(\Omega, \mathcal{F}, P)$: prob. sp. w/ $\mathcal{G} \subseteq \mathcal{F}$.

X : $\mathcal{G}$-mea. s.t. $\int_A X\,dP = 0, \ \forall A \in \mathcal{G}$.

Then $X = 0$ $\mathcal{G}$-a.s.

<Pf>

$A_+ := \{X > 0\}$, $A_0 := \{X = 0\}$, $A_- := \{X < 0\}$.

If $P(A_+) \neq 0$, since $A_+ = \bigcup_{n=1}^{\infty} \{X > 1/n\}$, there must exist n s.t. $\{X > 1/n\}$ has positive mea.

$\Rightarrow \int_{\{X > 1/n\}} X\,dP \geq \frac{1}{n} \cdot P(\{X > 1/n\}) > 0$. $\divideontimes$

Thus $P(A_+) = 0$. Similarly, $P(A_-) = 0$ and the result follows. #

Prop (Example 2.12.3~5)

① If $\underbrace{\mathcal{G} = \{\phi, \Omega\}}_{\boxed{\text{no information}}}$, then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.

② $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$. All cond. exp. has the same exp. i.e. $\mathbb{E}[X]$.

③ X : $\mathcal{F}$-mea.

Then $\mathbb{E}[X|\mathcal{F}] = X$. Whole information will recover X.

Prop 2.12.6

① (Linearity) $\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}]$

② (Factoring out mea. part).

If X : $\mathcal{G}$-mea., then $\mathbb{E}[XY|\mathcal{G}] = X \cdot \mathbb{E}[Y|\mathcal{G}]$

In particular, $\mathbb{E}[X|\mathcal{G}] = X$.

③ (Tower) ("least info. wins").

If $\mathcal{H} \subseteq \mathcal{G}$, then

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}]$$
$$= \mathbb{E}[X|\mathcal{H}].$$

④ (Positivity) If $X \geq 0$, then $\mathbb{E}[X|\mathcal{G}] \geq 0$

⑤ If c is a const., then $\mathbb{E}[c|\mathcal{G}] = c$.

⑥ (Indep. info).

If X : indep. of $\mathcal{G}$, then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.

2.13 Ineq. of R.V.

Thm 2.13.2 (Jensen's ineq.).

X : L^1 r.v.

$\varphi : \mathbb{R} \to \mathbb{R}$ convex s.t. $\varphi(X)$: L^1.

Then $\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)]$.

App. 2.13.3

$X : L^2 \Rightarrow X : L^1$.

<Pf>

$\varphi(x) = x^2$ is convex.

Thus by Jensen's, $\mathbb{E}[X]^2 \leq \mathbb{E}[X^2]$.

Hence, $X : L^2 \Leftrightarrow \mathbb{E}[X^2] < \infty \Rightarrow \mathbb{E}[X]^2 < \infty \Rightarrow X : L^1$. #

App. 2.13.4

X : r.v. w/ mean μ and mgf $m_X(t)$.

Then $m_X(t) \geq e^{t\mu}$, $\forall t$.

<Pf>

$\varphi(x) = e^x$ is convex.

By Jensen's, $e^{\mathbb{E}[tX]} \leq \mathbb{E}[e^{tX}] = m_X(t)$.

$\Rightarrow e^{t\mu} = e^{\mathbb{E}[tX]} \leq m_X(t)$, $\forall t$. #

Rmk :

X : r.v. It is meaningful to talk about variance and standard deviation only when $X \in L^2$.

Prop (Ex 2.13.6)

X : nonconst r.v. (i.e. $\nexists\, c$ s.t. $X = c$ a.s.)

Then $\mathrm{Var}(X) \neq 0$.

<Pf>

$\mathrm{Var}(X) = 0 \iff \mathbb{E}[(X - \mathbb{E}[X])^2] = 0 \iff X = \mathbb{E}[X]$ a.s. #

Prop (Ex 2.13.7) (Conditional Jensen's ineq.)

$\varphi : \mathbb{R} \to \mathbb{R}$ convex.

$\mathcal{G} \subseteq \mathcal{F}$. X : $\mathcal{F}$-mea.

Then $\varphi(\mathbb{E}[X \mid \mathcal{G}]) \leq \mathbb{E}[\varphi(X) \mid \mathcal{G}]$.

<Pf> Def of cond. exp. and stand Jensen's. #

Thm 2.13.9 (Markov's ineq.)

$\forall \lambda, p > 0$, we have

$$P(\omega \mid |X(\omega)| \geq \lambda) \leq \frac{1}{\lambda^p} \mathbb{E}[|X|^p].$$

<Pf> $\lambda^p \cdot P(|X| \geq \lambda) \leq \mathbb{E}[|X|^p]$. #

Thm 2.13.10 (Tchebychev's ineq.)

X : r.v. w/ mean μ and var. σ^2.

Then $P(|X - \mu| \geq \lambda) \leq \frac{\sigma^2}{\lambda^2}$, $\forall \lambda > 0$.

<Pf> Use Markov's. #

Thm 2.13.11 (Chernoff bounds).

X : r.v.

$\lambda \in \mathbb{R}$.

Then ① $P(X \geq \lambda) \leq \dfrac{\mathbb{E}[e^{tX}]}{e^{\lambda t}}$, $\forall t > 0$.

② $P(X \leq \lambda) \leq \dfrac{\mathbb{E}[e^{tX}]}{e^{\lambda t}}$, $\forall t < 0$.

<Pf>

under $t>0$.

① $X \geq \lambda \iff tX \geq t\lambda \iff e^{tX} \geq e^{t\lambda}$. Apply Markov's. #

② is similar. #

App. (Prop 2.13.12)

$X \sim N(\mu, \sigma^2)$.

Then, $\forall \lambda > \mu$, $P(X \geq \lambda) \leq e^{-\frac{(\lambda - \mu)^2}{2\sigma^2}}$.

<Pf>

Recall that $\mathbb{E}[e^{tX}] = m_X(t) = e^{\mu t + \frac{1}{2} t^2 \sigma^2}$. Apply Chernoff ①. #

Rmk:

Markov's, Tchebychev's and Chernoff's will be useful for computing lim. of r.v.

Prop 2.13.14

X : r.v.

$f, g : \mathbb{R} \to \mathbb{R}$, both $\uparrow$ or both $\downarrow$.

Then $\mathbb{E}[f(X) g(X)] \geq \mathbb{E}[f(X)]\, \mathbb{E}[g(X)]$.

<Pf>

We need the following lemma:

Lemma:

$X : \Omega \to \mathbb{R}$ r.v.

product sp.

Define $X_1, X_2 : \Omega^2 \to \mathbb{R}$ by $X_1(\omega_1, \omega_2) = X(\omega_1)$, $X_2(\omega_1, \omega_2) = X(\omega_2)$.

Then X_1 and X_2 are iid and

$\mathbb{E}[X_1] = \mathbb{E}[X] = \mathbb{E}[X_2]$.

<Pf of lemma>

For A, B Borel in $\mathbb{R}$, $\{X_1 \in A\} = A \times \Omega$ and $\{X_2 \in B\} = \Omega \times B$. Their intersection is $A \times B$.

Thus, $P(X_1 \in A, X_2 \in B) = P(A \times B) = P(A) \times P(B)$

$= P(X_1 \in A)\, P(X_2 \in B)$. $\Rightarrow X_1, X_2$: indep. #

$\mathbb{E}[X_1] = \int_{\Omega \times \Omega} X(\omega_1)\, d\omega_1\, d\omega_2 = \int_\Omega \left[\int_\Omega X(\omega_1)\, d\omega_1 \right] d\omega_2$

$= \int_\Omega \mathbb{E}[X]\, d\omega_2 = \mathbb{E}[X]$. Similarly, $\mathbb{E}[X_2] = \mathbb{E}[X]$. #

By abuse of notation, denote X_1 and X_2 in the lemma by X and Y.

By cond. of f and g, $(f(X) - f(Y))(g(X) - g(Y)) \geq 0$.

Applying $\mathbb{E}$ and using indep. of X and Y, we have

$\mathbb{E}[f(X) g(X)] + \mathbb{E}[f(Y) g(Y)] \geq \mathbb{E}[f(X)]\, \mathbb{E}[g(Y)] + \mathbb{E}[f(Y)]\, \mathbb{E}[g(X)]$.

By lemma, since X and Y are iid, the result follows. #

2.14 Limits of Seq. of R.V.

Def $\{X_n\}_{n=1}^\infty$, X : r.v.

① X_n converges almost surely to X if

$$P(\omega : \lim_{n \to \infty} X_n(\omega) = X(\omega)) = 1.$$

We write $X = \text{as-}\lim_{n \to \infty} X_n$.

② X_n converges to X in the mean square if

$$\lim_{n \to \infty} \mathbb{E}[(X - X_n)^2] = 0.$$

We write $\text{ms-}\lim_{n \to \infty} X_n = X$. ⊛ useful when defining Ito integral.

③ X_n converges to X in probability if

$$\forall \varepsilon>0, \lim_{n\to\infty} P(|X-X_n| \geq \varepsilon) = 0.$$

We write $\displaystyle p\text{-}\lim_{n\to\infty} X_n = X$.

④ X_n converges to X in distribution if

$\forall$ cont. bdd φ, we have

$$\lim_{n\to\infty} \mathbb{E}[\varphi(X_n)] = \mathbb{E}[\varphi(X)].$$

Prop 2.14.1

$\{X_n\}_{n=1}^{\infty}$: a seq. of r.v.

$\mathbb{E}[X_n] \to K$, some const. K.

$\mathrm{Var}[X_n] \to 0$.

Then $\displaystyle \text{ms-}\lim_{n\to\infty} X_n = K$.

<Pf>

$\mathbb{E}[|X_n-K|^2] = \mathbb{E}[X_n^2] - 2K\,\mathbb{E}[X_n] + K^2$

$= \mathbb{E}[X_n^2] - \mathbb{E}[X_n]^2 + (\mathbb{E}[X_n]-K)^2$

$= \mathrm{Var}[X_n] + (\mathbb{E}[X_n]-K)^2 \to 0$ as $n\to\infty$. #

Rmk:

For con. properties of $\mathbb{E}$ and Var under $X = \text{ms-}\lim_{n\to\infty} X_n$, see Exercise 2.14.2 and Ex 2.14.3.

Prop (Ex 2.14.4)

$X = \text{ms-}\lim_{n\to\infty} X_n$. $\mathcal{H} \subseteq \mathcal{F}$.

Then $\mathbb{E}[X|\mathcal{H}] = \text{ms-}\lim_{n\to\infty} \mathbb{E}[X_n|\mathcal{H}]$.

<Pf>

$$(\mathbb{E}[X|\mathcal{H}] - \mathbb{E}[X_n|\mathcal{H}])^2 = (\mathbb{E}[X-X_n|\mathcal{H}])^2 \overset{\text{Jensen's}}{\leq} \mathbb{E}[(X-X_n)^2|\mathcal{H}]$$

Thus $\displaystyle \int_\Omega |\mathbb{E}[X|\mathcal{H}] - \mathbb{E}[X_n|\mathcal{H}]|^2 \leq \int_\Omega \mathbb{E}[(X-X_n)^2|\mathcal{H}]$

$\displaystyle = \int_\Omega (X-X_n)^2 \to 0$ as $n\to\infty$.

Hence, $\mathbb{E}[X_n|\mathcal{H}] \to \mathbb{E}[X|\mathcal{H}]$ in mean square. #

Prop 2.14.5 (ms-lim ⇒ p-lim).

Convergence in mean square implies convergence in probability. (i.e. L^2-con. ⇒ con. in prob.)

<Pf>

By Markov's, $0 \leq P(|Y_n - Y| \geq \varepsilon) \leq \frac{1}{\varepsilon^2}\mathbb{E}[|Y_n-Y|^2] \to 0$

if $Y = \text{ms-}\lim_{n\to\infty} Y_n$. #

Prop (Ex 2.14.6)

$X_n \to X$ in L^1.

⇒ $X_n \to X$ in prob.

<Pf>

$$0 \leq P(|X_n - X| \geq \varepsilon) \leq \frac{\mathbb{E}[|X_n-X|]}{\varepsilon} \to 0 \text{ as } n\to\infty. \ \#$$

Rmk:

W/ the same argument, we have:

$X_n \to X$ in L^p ⇒ $X_n \to X$ in prob.

Def

X: r.v. w/ pdf $p(x)$.

$\hat{p}_X(t) := \int_{-\infty}^{\infty} e^{itx} p(x)\,dx$ is called the characteristic function of X.

Note: This is the Fourier transform of $p(x)$.

Prop

① Convergence in prob. ⇒ convergence in dist.

② $X_n \to X$ in dist.

⇒ $\hat{p}_{X_n} \to \hat{p}_X$. (characteristic fun.)

⇒ $p_{X_n} \to p_X$ (pdf).

③ $X_n \to X$ in dist.

⇔ $\displaystyle \lim_{n\to\infty} F_{X_n}(x) = F_X(x)$, $\forall$ cont. pt x of F_X.

Rmk:

See Ex 2.14.8 for an example of con. in dist. but not con. in prob.

2.15 Properties of Mean-Square Limit.

(*) Will be useful in applications of Ito integral.

Lemma 2.15.1.

$\displaystyle \text{ms-}\lim_{n\to\infty} X_n = 0$ and $\displaystyle \text{ms-}\lim_{n\to\infty} Y_n = 0$.

Then $\displaystyle \text{ms-}\lim_{n\to\infty} (X_n + Y_n) = 0$.

<Pf>

$(x+y)^2 \leq 2x^2 + 2y^2$. #

Prop 2.15.2

$\displaystyle \text{ms-}\lim_{n\to\infty} X_n = X$ and $\displaystyle \text{ms-}\lim_{n\to\infty} Y_n = Y$.

Then ① $\displaystyle \text{ms-}\lim_{n\to\infty} (X_n + Y_n) = X + Y$ and

② $\displaystyle \text{ms-}\lim_{n\to\infty} (cX_n) = c \cdot X$, $\forall c \in \mathbb{R}$.

<Pf> Apply L2.15.1 #

Rmk:

In general, $\underset{n\to\infty}{\text{ms-lim}} (X_n Y_n) \neq (\underset{n\to\infty}{\text{ms-lim}} X_n) \cdot (\underset{n\to\infty}{\text{ms-lim}} Y_n)$.

See Ex 2.15.5 for an example.

2.16 Stochastic Processes.

Def

A stochastic process on $(\Omega, \mathcal{F}, P)$ is a family of r.v. X_t parameterized by $t \in T \subseteq \mathbb{R}$.

$(X_t)_{t\in T}$ is said to be in continuous time if T is an interval and in discrete time if $T = \{1, 2, 3, \cdots\}$.

For $\omega \in \Omega$, the function $t \mapsto X_t(\omega)$ is called a path or realization of X_t.

Rmk:

① We can define convergence concepts as $t \to \infty$ on stochastic processes in a similar way as before.

② "The study of stochastic processes via computer simulation" is based on retrieving info. of X_t given a large number of its realizations.

Def

① $\{\mathcal{F}_t\}_{t\in T}$, a family of σ-fields in $\mathcal{F}$, is a filtration if $\mathcal{F}_s \subseteq \mathcal{F}_t$, $\forall\ s \le t$ in T.

② $(X_t)_{t\in T}$ is adapted to the filtration $\{\mathcal{F}_t\}_{t\in T}$ if X_t is $\mathcal{F}_t$-mea., $\forall\, t \in T$.

Ex: (2.16.2)

X : r.v. $\{\mathcal{F}_t\}_{t\in T}$: a filtration.

$X_t := \mathbb{E}[X \mid \mathcal{F}_t]$. Then X_t is adapted to $\{\mathcal{F}_t\}_{t\in T}$

Def

$X_t, t \in T$, is called a martingale w.r.t. $\{\mathcal{F}_t\}_{t\in T}$ if ① $X_t \in L^1$, $\forall\, t \in T$

② X_t : adapted to $\mathcal{F}_t$

③ $X_s = \mathbb{E}[X_t \mid \mathcal{F}_s]$, $\forall\ s < t$.

Rmk:

Condit'n ③ above asserts that the best forecast of $\underbrace{\text{the future}}_{X_t}$ is the $\underbrace{\text{last observation}}_{X_s}$ of X_t.

Note: $X_s = \mathbb{E}[X_t \mid \mathcal{F}_s] \iff \mathbb{E}[X_t - X_s \mid \mathcal{F}_s] = 0$.

Thus, some people interpret a martingale as a fair game.

Prop (Ex 2.16.7)

X : L^1 r.v. $(\mathcal{F}_t)_{t\in T}$: filtration

$X_t := \mathbb{E}[X \mid \mathcal{F}_t]$.

Then X_t is a martingale. w.r.t. $\mathcal{F}_t$.

<Pf>

Use properties of cond. exp. #

Prop (Ex 2.16.8)

X_t, Y_t : martingales w.r.t. $\mathcal{F}_t$.

$a, b, c \in \mathbb{R}$.

Then $Z_t := aX_t + bY_t + c$ is a martingale w.r.t. $\mathcal{F}_t$.

Def

X_t, Y_t are called conditionally uncorrelated given $\mathcal{F}_t$ if $\mathbb{E}[(X_t - X_s)(Y_t - Y_s) \mid \mathcal{F}_s] = 0$, $\forall\ 0 \le s < t < \infty$.

Prop (Ex 2.16.10).

X_t, Y_t : martingale w/ $Z_t := X_t Y_t$ L^1.

Then Z_t : martingale $\iff$ X_t, Y_t : conditionally uncorrelated given $\mathcal{F}_t$.

<Pf>

Easy check!! #

Rmk:

In the following, w/o further indication, given a process X_t, $\mathcal{F}_t := \sigma(X_s : s \le t)$.

Prop (Ex 2.16.11).

X_n, $n \ge 0$, : L^1 indep. r.v.

$S_0 := X_0$

$S_n := X_0 + X_1 + \cdots + X_n$.

Then

① $S_n - \mathbb{E}[S_n]$ is a martingale.

② if $\mathbb{E}[X_n] = 0$ and $\mathbb{E}[X_n^2] < \infty$, $\forall\ n \ge 0$, then $S_n^2 - \mathrm{Var}(S_n)$ is a martingale.

<Pf>

Easy check!! #

Prop (Ex 2.16.12)
$X_n, n\geq 0$: L^1 indep. r.v. w/ $\mathbb{E}[X_n]=1, \forall n\geq 0$.

Then $P_n := X_0 \cdot X_1 \cdot \cdots \cdot X_n$ is an $\mathcal{F}_n$-martingale.

<Pf>
Easy check!! #

Prop (Ex 2.16.13)
$X, (X_i)_{i\geq 0}$: iid normal w/ mean $\mu \neq 0$ and var. σ^2.

$\theta := -2\mu/\sigma^2$

Then ① $\mathbb{E}[e^{\theta X}]=1$.

② $S_n := \sum_{j=0}^{n} X_j$ and $Z_n := e^{\theta S_n}$.

$\Rightarrow Z_n$ is a martingale.

<Pf>
Easy check!! #

Chap 3 Useful Stochastic Processes.

In this chap, Brownian motion, Poisson processes and their "derivatives" are introduced.
Not in math sense.

3.1 The Brownian Motion.

Def
A Brownian motion process is a stochastic process B_t, $t\geq 0$ s.t.

① $B_0 = 0$

② B_t has indep. increments.

③ B_t is cont. in t. (i.e. cont. realizations).

④ $B_t - B_s$ is normally distributed w/ mean 0 and variance $|t-s|$. i.e. $B_t - B_s \sim N(0, |t-s|)$.

Def
A process X_t is said to have stationary increments if the dist. of $X_t - X_s$ dep. only on $t-s$.

Rmk:
① A Brownian motion has stationary increments.

② (Not proved) Even B_t is cont. in t, it is nowhere diff.

Prop
① $B_t \sim N(0,t)$, $\forall t>0$. In particular, $\mathbb{E}[B_t^2]=t$.

② $\mathbb{E}[B_s B_t] = s$, $\forall 0\leq s<t$. Hence, B_s, B_t: not indep.

<Pf>
① $B_t = B_t - B_0$. By ④ in the def. #

② $\mathbb{E}[B_s B_t] = \mathbb{E}[(B_s - B_0)(B_t - B_s) + B_s^2]$

$= \mathbb{E}[B_s - B_0]\mathbb{E}[B_t - B_s] + \mathbb{E}[B_s^2]$

$= 0\cdot 0 + s = s$. #

Def
A process w/ stationary and indep increments is called a Lévy process.

Rmk:
A Brownian motion is a Lévy process.

Prop 3.1.2
A Brownian motion process B_t is a martingale w.r.t. $\mathcal{F}_t := \sigma(B_s \mid s\leq t)$.

<Pf>
$\because B_t : L^2 \quad \therefore B_t : L^1$.
B_t : obviously adapted to $\mathcal{F}_t$.
For $0\leq s<t$,

$\mathbb{E}[B_t \mid \mathcal{F}_s] = \mathbb{E}[\underline{B_t - B_s} + \underline{B_s} \mid \mathcal{F}_s]$

$= \mathbb{E}[B_t - B_s] + B_s = 0 + B_s = B_s$. #

Def
A Weiner process W_t is a process adapted to $\mathcal{F}_t$ s.t.

① $W_0 = 0$.

② W_t : L^2 martingale w/ $\mathbb{E}[(W_t - W_s)^2] = t-s$, $s\leq t$.

③ W_t : cont. in t.

Cor (of Thm 10.2.1 (Lévy))
X_t is a Weiner process
$\Longleftrightarrow X_t$ is a Brownian motion process.

Rmk:
By the above Cor, Brownian motion and Weiner process are equivalent. We will use B_t and W_t interchangeably, meaning either of them.

Prop 3.1.5
$Y_t := W_t^2 - t$ is a martingale.

<Pf>
$\mathbb{E}[|Y_t|] \leq \mathbb{E}[W_t^2 + t] = 2t < \infty. \Rightarrow Y_t \in L^1$.

For $0 \le s < t$,

$\mathbb{E}[W_t^2 | \mathcal{F}_s] = \mathbb{E}[(W_t - W_s + W_s)^2 | \mathcal{F}_s]$

$= \mathbb{E}[(W_t - W_s)^2 | \mathcal{F}_s] + \mathbb{E}[2W_s(W_t - W_s) | \mathcal{F}_s]$

$+ \mathbb{E}[W_s^2 | \mathcal{F}_s]$

$= (t-s) + 2W_s \cdot 0 + W_s^2 = t + W_s^2 - s.$

Thus, $\mathbb{E}[W_t^2 - t | \mathcal{F}_s] = W_s^2 - s.$

i.e. $\mathbb{E}[Y_t | \mathcal{F}_s] = Y_s$. #

Prop 3.1.6 (Memoryless Property).

The cond. dist. of W_{t+s}, given W_t: the present and $W_u, u<t$: the past, dep. only on the present.

<Pf>

We need to prove: $\forall c \in \mathbb{R}$,

$P(W_{t+s} \le c | W_t = x, W_u, 0 \le u < t)$ ⊛

$\stackrel{?}{=} P(W_{t+s} \le c | W_t = x)$.

⊛ $= P(W_{t+s} - W_t \le c - x | W_t = x, W_u, 0 \le u < t)$

$= P(W_{t+s} - W_t \le c - x)$

$= P(W_{t+s} - W_t \le c - x | W_t = x)$

$= P(W_{t+s} \le c | W_t = x)$. #

Prop 3.1.7

Let $0 \le s \le t$. Then

① $Cov(W_s, W_t) = s$. $= \min\{s,t\}$.

② $Corr(W_s, W_t) = \sqrt{\frac{s}{t}}$. $= \sqrt{\frac{\min\{s,t\}}{\max\{s,t\}}}$.

<Pf>

① $Cov(W_s, W_t) = \mathbb{E}[(W_s - \mathbb{E}[W_s])(W_t - \mathbb{E}[W_t])]$

$= \mathbb{E}[W_s W_t] = s$. #

② $Corr(W_s, W_t) = \frac{Cov(W_s, W_t)}{\sqrt{Var(W_s)}\sqrt{Var(W_t)}} = \frac{s}{\sqrt{s}\sqrt{t}} = \sqrt{\frac{s}{t}}$. #

Prop (① translation invariant / ② scaling invariant).

① Let $t_0 \ge 0$. $X_t := W_{t+t_0} - W_{t_0}$, $t \ge 0$.

Then X_t: Brownian motion.

② Let $\lambda > 0$. $X_t := \frac{1}{\sqrt{\lambda}} W_{\lambda t}$.

Then X_t: Brownian motion.

3.2 Geometric Brownian Motion.

Def

The **geometric Brownian motion** w/ **drift** μ and **volatility** σ is the process

$$X_t = e^{\sigma W_t + (\mu - \sigma^2/2)t}, \quad t \ge 0.$$

Prop (Ex 3.2.4)

The driftless process is always a martingale.

Lemma 3.2.1

$\mathbb{E}[e^{\alpha W_t}] = e^{\alpha^2 t/2}$, $\forall \alpha \ge 0$.

<Pf>

Use $W_t \sim N(0,t)$. #

Prop 3.2.2

The **exponential Brownian motion** $X_t = e^{W_t}$ is log-normally distributed w/ mean $e^{t/2}$ and var. $e^{2t} - e^t$.

3.3 Integrated Brownian Motion.

Def

$Z_t := \int_0^t W_s \, ds$, $t \ge 0$ is called the **integrated Brownian motion.**

Lemma (Thm 3.3.1)

X_j: indep. r.v. normally distributed w/ mean μ_j and var. σ_j^2.

Then $X_1 + \cdots + X_n$ is normally distributed w/ mean $\mu_1 + \cdots + \mu_n$ and var. $\sigma_1^2 + \cdots + \sigma_n^2$.

Prop 3.3.2

Z_t has normal dist. w/ mean 0 and var. $t^3/3$.

<Pf>

Given a partition $0 = s_0 < s_1 < \cdots < s_n = t$, where $s_k = kt/n$.

$\because W_s$: cont. w.r.t. s

$\therefore Z_t = \lim_{n \to \infty} \sum_{k=1}^{n} W_{s_k} \cdot \frac{t}{n} = t \cdot \lim_{n \to \infty} \frac{W_{s_1} + \cdots + W_{s_n}}{n}$.

$W_{s_1} + \cdots + W_{s_n} = \underline{n(W_{s_1} - W_0)} + \underline{(n-1)(W_{s_2} - W_{s_1})} + \cdots + \underline{(W_{s_n} - W_{s_{n-1}})}$

$=: X_1 + \cdots + X_n$.

X_k is normally dist. w/ mean 0 and var. $(n-k+1)^2 \cdot \frac{t}{n}$.

Thus, by Lemma,

$$\frac{t(W_{s_1}+\cdots+W_{s_n})}{n} \sim N\left(0, \left(\frac{t}{n}\right)^2 \cdot (n^2+(n-1)^2+\cdots+1^2) \cdot \frac{t}{n}\right)$$

$$= N\left(0, t^3 \cdot \frac{(n+1)(2n+1)}{6n^2}\right)$$

LHS $\to Z_t$ a.s.

RHS $\to N(0, t^3/3)$ in dist.

Thus, $Z_t \sim N(0, t^3/3)$. #

Prop (Ex 3.3.9). $\mathcal{F}_t := \sigma(W_s : s \le t)$. Then

$\mathbb{E}[Z_T | \mathcal{F}_t] = Z_t + W_t \cdot (T-t)$, $\forall\, T > t$.

Thus $M_t := Z_t - tW_t$ is an $\mathcal{F}_t$-martingale.

3.4 Exponential Integrated Brownian Motion.

Def

$V_t := e^{Z_t}$ is called the exponential integrated Brownian motion.

Prop (Ex 3.4.1)

$\mathbb{E}[V_T | \mathcal{F}_t] = V_t e^{(T-t)W_t + \frac{(T-t)^3}{3}}$, $\forall\, t < T$.

3.5 Brownian Bridge.

Def

$X_t := W_t - tW_1$ is called the Brownian bridge fixed at 0 and 1, $0 \le t \le 1$.

Rmk:

① $X_0 = X_1 = 0$. (X_t : Brownian bridge at 0 and 1)

② $X_t = W_t - tW_1 = W_t - tW_t + tW_t - tW_1$

$= (1-t)(W_t - W_0) - t(W_1 - W_t)$.

Prop

$X_t \sim N(0, t(1-t))$.

⟨Pf⟩

$(1-t)(W_t - W_0) \sim N(0, (1-t)^2 t)$ and

$-t(W_1 - W_t) \sim N(0, t^2(1-t))$ are indep.

$\Rightarrow X_t \sim N(0, (1-t)^2 t + t^2(1-t)) = N(0, t(1-t))$. #

3.6 Brownian Motion w/ Drift.

Def

$Y_t := \mu t + W_t$, $t \ge 0$, μ : const. in $\mathbb{R}$, is called a Brownian motion w/ drift. μ : drift rate.

Prop

$Y_t \sim N(\mu t, t)$.

⟨Pf⟩

For fixed t, Y_t is simply W_t plus a const. μt.

$\because W_t \sim N(0,t)$ $\therefore Y_t \sim N(\mu t, t)$. #

3.7 Bessel Process.

Def

$W_1(t), \cdots, W_n(t)$: indep. Brownian motions.

$W(t) := (W_1(t), \cdots, W_n(t))$.

$R_t := \text{dist}(0, W(t))$ (dist = distance)

$= \sqrt{W_1(t)^2 + \cdots + W_n(t)^2}$, is called the n-dimensional Bessel process.

Prop 3.7.1

The pdf of R_t, $t > 0$, is given by

$$P_t(\rho) = \begin{cases} \dfrac{2}{(2t)^{n/2} \cdot \Gamma(n/2)} \cdot \rho^{n-1} e^{-\frac{\rho^2}{2t}}, & \rho \ge 0 \\ 0, & \rho < 0 \end{cases}, \text{ where}$$

$$\Gamma\left(\frac{n}{2}\right) = \begin{cases} \left(\frac{n}{2}-1\right)!, & n : \text{even} \\ \left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)\cdots\left(\frac{3}{2}\right)\left(\frac{1}{2}\right)\sqrt{\pi}, & n : \text{odd} \end{cases}$$

⟨Pf⟩

Standard calculation using spherical coordinates. #

3.8 The Poisson Process.

Main Idea:

A Poisson process describes the number of occurrences of a certain event before time t.

Def

A Poisson process with rate $\lambda (>0)$ is a stochastic process N_t, $t \ge 0$, s.t.

① $N_0 = 0$

② N_t has indep. increments.

③ N_t is right cont. in t, w/ left hand limits.

④ $N_t - N_s$, w/ $0 < s < t$, has a Poisson dist. w/ parameter $\lambda(t-s)$. i.e. $P(N_t - N_s = k) = \frac{\lambda^k (t-s)^k}{k!} e^{-\lambda(t-s)}$.

Rmk:

④ states that N_t has stationary increments.

Prop

① $\mathbb{E}[N_t] = \lambda t$, $Var[N_t] = \lambda t$.

② $\mathbb{E}[N_s N_t]$

$= \lambda^2 st + \lambda s$, $\forall\ 0 < s < t$.

⟨Pf⟩

① $N_t - N_0 \sim Poisson(\lambda t)$.

$\because N_0 = 0 \quad \therefore N_t \sim Poisson(\lambda t)$.

$\Rightarrow \mathbb{E}[N_t] = \lambda t = Var[N_t]$. #

② $\mathbb{E}[N_s N_t] = \mathbb{E}[(N_s - N_0)(N_t - N_s + N_s)]$

$= \mathbb{E}[N_s - N_0]\mathbb{E}[N_t - N_s] + \mathbb{E}[N_s^2]$

$= \lambda s \cdot \lambda(t-s) + (\mathbb{E}[N_s]^2 + Var[N_s])$

$= \lambda^2 s(t-s) + (\lambda s)^2 + \lambda s = \lambda^2 st + \lambda s$. #

Prop 3.8.4

Let $0 \le s \le t$. Then

① $Cov(N_s, N_t) = \lambda s$.

② $Corr(N_s, N_t) = \sqrt{s/t}$.

⟨Pf⟩

Apply prop above. #

Prop 3.8.5

Let N_t be $\mathcal{F}_t$-adapted.

Then $M_t := N_t - \lambda t$ is an $\mathcal{F}_t$-martingale.

⟨Pf⟩

$\because N_t$ is L^1 $\therefore M_t$ is L^1.

$\because N_t : \mathcal{F}_t$-mea. $\therefore M_t : \mathcal{F}_t$-mea.

$\mathbb{E}[N_t | \mathcal{F}_s] = \mathbb{E}[N_s + (N_t - N_s) | \mathcal{F}_s]$

$= N_s + \mathbb{E}[N_t - N_s] = N_s + \lambda(t-s)$.

$\Rightarrow \mathbb{E}[N_t - \lambda t | \mathcal{F}_s] = N_s - \lambda s$. i.e. $\mathbb{E}[M_t | \mathcal{F}_s] = M_s$.

Thus, M_t is an $\mathcal{F}_t$-martingale. #

Def

$M_t := N_t - \lambda t$ is called the compensated Poisson process.

Rmk By proof above, N_t itself is NOT a martingale.

3.9 Interarrival Times.

Rmk:

For each $\omega \in \Omega$, the path $(t \mapsto N_t(\omega))$ is a step function exhibiting unit jumps.

Def

T_1 := the r.v. describing the time of the 1st jump.

For $n \ge 2$,

T_n := the r.v. " " " elapsed between the $(n-1)$th and nth jump.

$T_1, T_2, T_3, \cdots$ are called interarrival times.

Prop 3.9.1

The r.v. T_n are iid exponentially distributed w/ mean $\mathbb{E}[T_n] = 1/\lambda$.

⟨Pf⟩

Use defining properties of Poisson process.

For independence, use property ②.

For exponentially distributed, use ④. #

3.10 Waiting Times.

Def

$S_n := T_1 + \cdots + T_n$ is called the waiting time until the nth jump.

Prop

The pdf of S_n is

$$f_{S_n}(t) = \frac{t^{n-1} e^{-\lambda t}}{(1/\lambda)^n \Gamma(n)}.$$

i.e. S_n has gamma distribution w/ parameters $\alpha = n$ and $\beta = 1/\lambda$.

Thus $\mathbb{E}[S_n] = n/\lambda$, $Var[S_n] = n/\lambda^2$.

⟨Pf⟩

Use $\{S_n \le t\} = \{N_t \ge n\}$ to find cdf of S_n and differentiate to find pdf. #

3.11 The Integrated Poisson Process.

Def

$U_t := \int_0^t N_u \, du$ is called the integrated Poisson process.

Rmk For $\omega \in \Omega$, $u \mapsto N_u(\omega)$ is Riemann integrable.

Prop 3.11.1

$$U_t = t N_t - \sum_{k=1}^{N_t} S_k.$$

⟨Pf⟩

See Fig 3.6. #

Thm 3.11.2

Given $N_t = n$.

Then the waiting times $S_1, \cdots, S_n$ have joint pdf $f(s_1, \cdots, s_n) = \frac{n!}{t^n}$, $0 < S_1 \le S_2 \le \cdots \le S_n < t$.

3.12 Submartingales.

Def

A stochastic process X_t is called a submartingale w.r.t. the filtration $\mathcal{F}_t$ if

① $X_t : L^1, \forall t$

② X_t is $\mathcal{F}_t$-mea., $\forall t$.

③ $\mathbb{E}[X_t | \mathcal{F}_s] \geq X_s, \ \forall \ 0 \le s < t$.

Prop (Ex 3.12.1).

$X_t := \mu t + \sigma W_t$, w/ $\mu > 0$, is a submartingale.

⟨Pf⟩

$\because \mu t, \sigma$ are const. $\therefore X_t$ is L^1.

Clearly, X_t is $\mathcal{F}_t$-mea. (w/ $\mathcal{F}_t = \sigma(W_s : s \le t)$).

$\mathbb{E}[X_t | \mathcal{F}_s]$

$= \mathbb{E}[\mu t + \sigma W_t | \mathcal{F}_s] = \mu t + \sigma \mathbb{E}[W_t | \mathcal{F}_s]$

$= \mu t + \sigma W_s \geq \mu s + \sigma W_s = X_s$. #

Prop (Ex 3.12.2)

W_t^2 is a submartingale.

⟨Pf⟩

Recall that $W_t^2 - t$ is a martingale.

Thus $\mathbb{E}[W_t^2 | \mathcal{F}_s] = \mathbb{E}[W_t^2 - t + t | \mathcal{F}_s]$

$= \mathbb{E}[W_t^2 - t | \mathcal{F}_s] + t = W_s^2 - s + t \geq W_s^2$.

$\because W_t \sim N(0,t) \therefore \mathbb{E}[W_t^2] = Var[W_t] = t < \infty$.

$\Rightarrow W_t^2 : L^1$. It's clear W_t^2 is adapted to $\mathcal{F}_t = \sigma(W_s : s \le t)$. #

Prop 3.12.3 (Ways to construct submartingales).

(a) X_t : martingale. ϕ : convex w/ $\phi(X_t) : L^1$.

Then $Y_t := \phi(X_t)$ is a submartingale.

(b) X_t : submartingale. ϕ : increasing convex w/ $\phi(X_t) : L^1$.

Then $Y_t := \phi(X_t)$ is a submartingale.

(c) X_t : martingale. $f(t)$: increasing finite

can be generalized to submartingale.

Then $Y_t := X_t + f(t)$ is a submartingale.

⟨Pf⟩

(a)

$\mathbb{E}[Y_t | \mathcal{F}_s] = \mathbb{E}[\phi(X_t) | \mathcal{F}_s] \overset{\text{Jensen}}{\geq} \phi(\mathbb{E}[X_t | \mathcal{F}_s]) = \phi(X_s) = Y_s$ #

(b)

$\mathbb{E}[Y_t | \mathcal{F}_s] \underset{\text{Jensen}}{\geq} \phi(\mathbb{E}[X_t | \mathcal{F}_s]) \underset{\text{increasing}}{\geq} \phi(X_s) = Y_s$. #

(c)

Obvious. #

Cor 3.12.4

(a) X_t : martingale.

Then $X_t^2, |X_t|, e^{X_t}$ are submartingales, if they are L^1.

(b) Let $\mu > 0$. Then $e^{\mu t + \sigma W_t}$ is a submartingale.

Prop 3.12.5 (Doob's Submartingale Ineq.)

(a) X_t : nonnegative submartingale.

Then $P(\sup_{s \le t} X_s \geq x) \le \frac{\mathbb{E}[X_t]}{x}, \ \forall x > 0$.

(b) X_t : right cont. submartingale.

Then $P(\sup_{s \le t} X_s \geq x) \le \frac{\mathbb{E}[X_t^+]}{x}, \ \forall x > 0$,

where $X_t^+ = \max\{X_t, 0\}$.

Rmk:

Doob's ineq. implies the Markov ineq.

(i.e. consider the stochastic process $X_t := X, \forall t$).

Thm 3.12.11 (Doob's inequality)

X_t : cont. martingale. $T > 0$.

Then $\mathbb{E}[\sup_{0 \le t \le T} X_t^2] \le 4 \cdot \mathbb{E}[X_T^2]$

In general, if X_t is cont. L^p martingale w/ $p > 1$ and $T > 0$, then

$$\mathbb{E}[\sup_{0 \le t \le T} |X_t|^p] \le \left(\frac{p}{p-1}\right)^p \mathbb{E}[|X_T|^p].$$