ions1section.1

Probability, Measure and Martingales

November 17, 2022

1 Measurable sets and functions

given a set Ω,

- $\mathcal{P}(\Omega)$ is the power set of Ω
- $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ is an **algebra** on Ω if :
  (i) $\emptyset \in \mathcal{F}$
  (ii) if $E \in \mathcal{F}$ then $E^C = \Omega \backslash E \in \mathcal{F}$
  (iii) if $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$
- $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ is **σ-algebra/σ-field** on Ω if :
  (i) $\emptyset \in \mathcal{F}$
  (ii) if $E \in \mathcal{F}$ then $\Omega \backslash E \in \mathcal{F}$
  (iii) if $E_n \in \mathcal{F}$ for $n = 1, 2, ...$ then $\bigcup_{n=1}^{\infty} E_n \in \mathcal{F} \implies \bigcap_{n=1}^{\infty} E_n \in \mathcal{F}$

given $\mathcal{A} \in \mathcal{P}(\Omega)$, $\sigma(\mathcal{A})$ denotes the smallest σ-algebra containing all the sets in $\mathcal{A}$.
Note that the (countable) intersection of σ-algebras is also a σ-algebra [1.3].

Examples:

- trivial: $\{\emptyset, \Omega\}$
- simple: $\{\emptyset, \Omega, A, A^C\} = \sigma(\{A\}) = \sigma(A)$
- trace: given a set $E \subseteq \Omega$, a σ-algebra $\mathcal{F}$, then $\{E \cap A : A \in \mathcal{F}\}$ is a σ-algebra.
- Borel σ-algebra: the σ algebra generated by *open sets* on a topological space
- Borel σ-algebra on $\mathbb{R}$: $\sigma(\{\text{open sets in } \mathbb{R}\}) = \sigma(\{\text{open intervals in } \mathbb{R}\}) = \sigma(\{(-\infty, a] : a \in \mathbb{R}\})$ (you can write any open set as a countable union of open intervals, and any open interval in terms of half lines)

Measurable space: $(\Omega, \mathcal{F})$ a set and a σ-algebra on it

Product spaces

Given a collection of measurable spaces $(\Omega_i, \mathcal{F}_i)_{i\in I}$, their product space is $(\Omega, \mathcal{F})$, defined as follows:

- $\Omega = \prod_{i\in I} \Omega_i$ (a cartesian product)
- $\mathcal{F} = \sigma\left(\left\{A = \prod_{i\in I} A_i : \forall i \in I \ A_i \in \mathcal{F}_i, \forall \,\text{but infinite}\; i \ A_i = \Omega_i\right\}\right)$ (not a cartesian product - the A's are called cylinder sets)
  - note that if I is finite, $\mathcal{F}$ is just a cartesian product.

Note that we often write $\mathcal{F} = \times_{i\in I}\mathcal{F}_i$, but this is still not a Cartesian product in the countable case.

Product spaces on $\mathbb{R}$: $\mathscr{B}(\mathbb{R}^d) = \prod_1^d \mathscr{B}(\mathbb{R})$ clearly $\supseteq$, because $d < \infty$, and $\subseteq$ because any open set in $\mathbb{R}^d$ can be written as a product of open hypercubes.

π and λ systems

a collection of sets $\mathcal{A}$ is a

- **π-system** if it is stable under intersections - i.e. $A, B \in \mathcal{A} \implies A \cap B \in \mathcal{A}$
- **λ-system** if
  - $\Omega \in \mathcal{A}$
  - $A, B \in \mathcal{A}, A \subseteq B \implies B \backslash A \in \mathcal{A}$
  - $\{A_n\}_{n\geq 1} \subseteq \mathcal{A}$,for all $n \geq 1$ $A_n \subseteq A_{n+1}$ then $\bigcup_{n\geq 1} A_n \in \mathcal{A}$

$\mathcal{A}$ is a σ-algebra $\iff$ it is both a π-system and a λ-system. [proof: use $B_n = \bigcup_{k=1}^n A_k$ - note this is constructed using the π-system rule, as each B_n is finite]

$\pi - \lambda$ systems lemma: given $\mathcal{M}$ is a λ-system, $\mathcal{A}$ a π-system, then $\mathcal{A} \subseteq \mathcal{M} \implies \sigma(\mathcal{A}) \subseteq \mathcal{M}$

[proof: let $\lambda(\mathcal{A})$ be the $\cap$ of all λ-systems containing $\mathcal{A}$, clearly $\lambda(\mathcal{A}) \subseteq \mathcal{M}$, and prove $\lambda(\mathcal{A})$ is a π-system]

Random variables

given $(\Omega, \mathcal{F})$ and $(E, \mathcal{E})$ are measurable spaces, $f : \Omega \to E$ is a **measurable function/random variable** if $\forall A \in \mathcal{E}\ f^{-1}(A) = \{\omega \in \Omega : f(\omega) \in A\} \in \mathcal{F}$.

Random variables/measurable functions are closed under composition.

$A \subset \Omega$ is an event/measurable set $\iff$ $\mathbf{1}_A$ is measurable.

$\sigma(f_i : i \in I) :=$ the smallest σ-algebra on Ω st all $f_i : \Omega \to E_i$ are measurable wrt to it.

- Note that $\sigma(\{A_i : i \in I\}) = \sigma(\mathbf{1}_{A_i} : i \in I)$, as expected.
- for a single RV X from $(\Omega, \mathcal{F})$ to $(E, \mathcal{E})$ with $\mathcal{E} = \sigma(\mathcal{A})$, then $\sigma(X) = \{X^{-1}(A) : A \in \mathcal{E}\} = \sigma(X^{-1}(A) : A \in \mathcal{A})$.

in $\mathbb{R}$ or $[-\infty, \infty]$, thus f is measurable $\iff \forall t \in \mathbb{R}\ \{x : f(x) \leq t\} \in \mathcal{F}$

Measurable functions on $\mathbb{R}$ or $[-\infty, \infty]$ are closed under addition, multiplication, max, min, divison, composition, sup, inf, lim sup and lim inf.

f is a **simple function** if $f = \sum_{k=1}^{n} a_k \mathbf{1}_{E_k}$ for $n \geq 1, E_k \in \mathcal{F}, a_k \in \mathbb{R}$.

f is measurable iff it is a limit of simple functions.

if $X : \Omega \to E$ is an RV across $(\Omega, \mathcal{F})$ and $(E, \mathcal{E})$, g another RV on $(\Omega, \mathcal{F})$, g is $\sigma(X)$-measurable $\iff g = h \circ X$ for some RV h on $(E, \mathcal{E})$. [no proof]

Monotone Class theorem

Given $\mathcal{H}$ is a class of bounded functions $\Omega \to \mathbb{R}$ st:

1. $\mathcal{H}$ is a vector space over $\mathbb{R}$,
2. the constant function 1 is in $\mathcal{H}$
3. $\mathcal{H}$ is closed under upwards limits to bounded functions:

If $\mathcal{C} \subseteq \mathcal{H}$ is closed under pointwise multiplication, then $\mathcal{H}$ contains all bounded $\sigma(\mathcal{C})$-measurable functions

Correspondence to $\pi - \lambda$ systems lemma: let $\mathcal{C} = \{\mathbf{1}_A : A \in \mathcal{A}\}$ for $\mathcal{A}$ a π-system: then,

- $\mathbf{1}_A \times \mathbf{1}_B = \mathbf{1}_{\boldsymbol{A \cap B}}$, so $\mathcal{C} =$ a π-system
- $\mathbf{1} \in \mathcal{H}$ is just $\mathbf{1}_\Omega$
- vector space properties of $\mathcal{H}$ gives us complements/etc.
- upwards limits gives the union property of λ-systems

Proof is non-examinable.

2 Measures and measure spaces

Given $(\Omega, \mathcal{F})$ is a **measurable space**, a **measure** on $(\Omega, \mathcal{F})$ is a function $\mu : \mathcal{F} \to [0, \infty]$ st

(i) *set function*: $\mu(\emptyset) = 0$

(ii) *countably additive:* $\mu(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu(E_n)$ if the E_n are disjoint sets in $\mathcal{F}$

then $(\Omega, \mathcal{F}, \mu)$ is **measure space**.

useful properties of a measure space $(\Omega, \mathcal{F}, \mu)$ [2.3]:

1. additive: $A \cap B = \emptyset \implies \mu(A \cup B) = \mu(A) + \mu(B)$
2. increasing: if $A \subseteq B$ then $\mu(A) \leq \mu(B)$
3. $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$
4. continuous from below: if $A_n \uparrow A$ then $\mu(A_n) \uparrow \mu(A)$
5. continuous from above: if $B_n \downarrow B$ and $\mu(B_k) < \infty$ for some k, then $\mu(B_n) \downarrow \mu(B)$
6. $\mu(\bigcup_{n\geq 1} A_n) \leq \sum_{n\geq 1} \mu(A_n)$ (σ-subadditive)

[2.4] given $\mu : \mathcal{A} \to [0, \infty)$ is an additive set function (not a measure) on an algebra $\mathcal{A}$ taking only finite values, then μ is countably additive (i.e. a measure) $\iff$ for all sequences $(A_n) \subseteq \mathcal{A}$ with $A_n \downarrow \emptyset$ $\mu(A_n) \to 0$

a measure μ is **finite** if $\mu(\Omega) < \infty$, and a **probability measure** if $\mu(\Omega) = 1$, so $(\Omega, \mathcal{F}, \mu/\mathbb{P})$ is a probability space.

μ is **σ-finite** if $\exists (K_n)_{n\geq 1} \in \mathcal{F}$ st $\mu(K_n) < \infty$ for all n, and $\bigcup_{n\geq 1} K_n = \Omega$

$A \in \mathcal{F}$ is a **null set** if $\mu(A) = 0$, and a property holds **almost everywhere** if it it true for $\forall \omega \in \Omega \backslash A$ for A null. In probabilty measures, we normally call this **almost surely**

A measure ν is **absolutely continuous** wrt to a measure μ (and write $\nu \ll \mu$) if $\forall A$ $\mu(A) = 0 \implies \nu(A) = 0$. ν and μ are equivalent ($\mu \sim \nu$) if $\nu \ll \mu$ and $\mu \ll \nu$.

Uniqueness of extension: given μ_1, μ_2 are measures on $(\Omega, \mathcal{F})$ and $\mathcal{A} \subseteq \mathcal{F}$ with $\mathcal{F} = \sigma(\mathcal{A})$, if $\mu_1(\Omega) = \mu_2(\Omega) < \infty$ and $\mu_1 = \mu_2$ on $\mathcal{A}$, then $\mu_1 = \mu_2$ on $\mathcal{F}$.

Carathéodory Extension Theorem: given $(\Omega, \mathcal{F} = \sigma(\mathcal{A}))$, where $\mathcal{A}$ is an *algebra*, if $\mu_0 : \mathcal{A} \to [0, \infty]$ is a countably additive set function, then $\exists \mu : \mathcal{F} \to [0, \infty]$ a measure on $(\Omega, \mathcal{F})$ st $\mu = \mu_0$ on $\mathcal{A}$. [proof like part A outer measure]

- General outer measure: given the same setup as the C. Ext. Theorem, $\mu^*(B) = \inf \left\{ \sum\limits_{n=1}^{\infty} m(A_n) : A_n \in \mathcal{A}, B \subseteq \bigcup\limits_{n=1}^{\infty} A_n \right\}$
- measurable sets under the outer measure: $B \subseteq \Omega$ is measurable if $\forall E \in \Omega$, $\mu^*(E) = \mu^*(E \cap B) + \mu^*(E \cap B^C)$

If $(\Omega, \mathcal{F}, \mu)$ is a measure space and $\mathcal{G} \subseteq \mathcal{F}$ a σ-algebra, then $(\Omega, \mathcal{G}, \mu|_{\mathcal{G}})$ is the **restriction** of the measure space, and is itself a measure space.

The **sum** of a countable sequence of probability measures is a probability measure - if $(\Omega, \mathcal{F})$ is a measurable space, and $(\mu_n)_{n\geq 1}$ a seq of prob measures, $(a_n)_{n\geq 1}$ a sequence with $\sum_n a_n = 1$, then $\mu(A) := \sum_n a_n \mu_n(A)$ is a prob measure.

Basic conditional probability

$$\mu(A) = \mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \text{ for } A \in \mathcal{F} \tag{1}$$

is the **conditional probability** measure on $(\Omega, \mathcal{F}, \mathbb{P})$ given a set B with $\mathbb{P}(B) > 0$.

$$\mathbb{P}(A \mid \sigma(B))(\boldsymbol{\omega}) := \mathbb{P}(A \mid B)\mathbf{1}_B(\boldsymbol{\omega}) + \mathbb{P}(A \mid B^C)\mathbf{1}_{B^C}(\boldsymbol{\omega})$$

is the conditional probability given a σ-algebra - given a fixed $\boldsymbol{\omega}$, it is a probability measure over sets A, but for a fixed A, it is an RV.

Measures on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$

$F : \mathbb{R} \to \mathbb{R}$ is a **distribution function** if:

1. $F : \mathbb{R} \to [0, 1]$
2. F is increasing
3. $\lim_{x\to\infty} F(x) = 1$, $\lim_{x\to-\infty} F(x) = 0$
4. F is right-continuous

Given a measure μ on $\mathscr{B}(\mathbb{R})$, the **distribution function** of μ is $F_\mu(x) = \mu((-\infty, x])$, which satisfies the requirements above.

Lebesgue's theorem: for any distribution function F, there is a unique Borel measure μ_F on $\mathscr{B}(\mathbb{R})$ st $F = F_{\mu_F}$. i.e. there is a 1-1 correspondence.

Leb is the unique Borel measure on $\mathbb{R}$st $\forall a, b\ a \leq b\ \mathrm{Leb}((a, b]) = b - a$.

Distribution of an RV

Given X is a random variable from $(\Omega, \mathcal{F})$ to $(E, \mathcal{E})$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space:

- $\Omega \xrightarrow{X} \mathbb{R}$
- $[0, 1] \xleftarrow{\mathbb{P}} \mathcal{F} \xleftarrow{X^{-1}} \mathcal{B}$, or indeed $[0, 1] \xleftarrow{\mathbb{P}} \sigma(X) \xleftarrow{X^{-1}} \mathcal{B}$

Note that X^{-1} is the **pre-image** function $X^{-1}(A) = \{\omega : X(\omega) \in A\}$, not an inverse!

Define the **law/distribution** $\mu_X : \mathcal{B} \to [0, 1]$ by $\mu_X := \mathbb{P} \circ X^{-1}$. This is a probability measure on $(\mathbb{R}, \mathcal{B})$, also known as the **image measure** of $\mathbb{P}$ via X, or the **pushforward measure**

The **distribution function** of X is $F_X(a) : \mathbb{R} \to \mathbb{R}$ defined by $F_X(a) := \mathbb{P}(X \leq a) = \mu_X((-\infty, a])$ (see above)

We write $X \sim Y$ to mean $\mu_X = \mu_Y$, and X, Y can even be defined on different probability spaces $(\Omega, \mathcal{F})$

Given $(\Omega, \mathcal{F}, \mathbb{P})$ is a prob space, $(E, \mathcal{E})$ and $G, \mathcal{G})$ are meas spaces, $X : \Omega \to E; Y : E \to G$ then the image measure of "μ_X via Y" is the image measure of "μ via $Y \circ X$"

Existence of RVs

If F is a function with properties $1-4$ above, we can construct an RV on $(\Omega, \mathcal{F}, \mathbb{P}) = ((0,1), \mathcal{B}(0,1), \text{Leb})$ with distribution function $F_X = F$.

Define the right continuous inverse of F, also known as the *quantile* function

$$F^{-1}(z) = \inf\{y : F(y) > z\}$$

F^{-1} is increasing, and so measurable.

Then, $\{\omega : \omega < F(x)\} \subseteq \{\omega : F^{-1}(\omega) \leq x\} \subseteq \{\omega : \omega \leq F(x)\}$, and both outer sets have the same Lebesgue measure, $F(x)$.

Thus,

$$F_X(x) = \mathbb{P}(X \leq x) = \text{Leb}(X \leq x) = \text{Leb}(F^{-1} \leq x) = \text{Leb}\left(\{\omega : F^{-1}(\omega) \leq x\}\right) = \text{Leb}\left(\{\omega : \omega < F(x)\}\right) = F($$

[note $\mathbb{P} = \text{Leb}$, $X = F^{-1}$, $\Omega = (0,1)$ so the final inequality is true]

Product measure

Given $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i)_{i=1..N}$ are probability measures, there is a unique measure $\mathbb{P}$ on $(\Omega, \mathcal{F}) = (\prod_{i=1}^N \Omega_i, \times_{i=1}^N \mathcal{F}_i)$ st $\mathbb{P}(E_1 \times \cdots \times E_n) = \mathbb{P}_1(E_1) \cdots \mathbb{P}_N(E_n)$ for $E_i \in \mathcal{F}_i$. $\mathbb{P}$ is the **product measure**, and is often written $\otimes_{i \leq N} \mathbb{P}_i$.

Thus, the product measure is specified by its marginals. This is not true for general measures on product spaces (where two different measures could have the same marginals)

3 Independence

A collection of σ-algebras $\mathcal{G}_i : i \leq n$ is **independent** iff $\forall A_i \in \mathcal{G}_i, i \leq n \; \mathbb{P}(\bigcap_{i \leq n} A_i) = \prod_{i \leq n} \mathbb{P}(A_i)$

For sequences $(\mathcal{G}_n)_{n \geq 1}$ of σ-algebras by $\forall A_n \in \mathcal{G}_n, n \geq 1 \; \mathbb{P}(\bigcap_{n \geq 1} A_i) = \prod_{n \geq 1} \mathbb{P}(A_i)$ by continuity of measure

If our collection of σ-algebras $(\mathcal{G}_i)_{i \in \mathcal{I}}$ can be written as $(\sigma(\mathcal{A}_i))_{i \in I}$ for π-systems $\mathcal{A}_i \in \mathcal{F}$, then $(\mathcal{G}_i)_{i \in \mathcal{I}}$ are independent $\iff$ $\mathbb{P}(\bigcap_{i \in J} A_i) = \prod_{i \in J} \mathbb{P}(A_i)$ for any $A_i \in \mathcal{A}_i, i \in J$ for any finite subset $J \subseteq \mathcal{I}$. [see theorem 3.5]

A collection of **events** (finite or countable) is independent $\iff$ their generated σ-algebras are $\iff$ the standard condition from Part A.

A finite or countable collection of **RVs** $(X_i)_{i\in\mathcal{I}}$, maps from $(\Omega, \mathcal{F}, \mathbb{P})$ to measurable spaces $(E_i, \mathcal{E}_i)_{i\in\mathcal{I}}$, are independent

- $\iff$ $(\sigma(X_i))_{i\in\mathcal{I}}$ are independent
- $\iff$ $\mathbb{P}(X_i \in A_i$ for all $i \in J) = \prod_{i\in J} \mathbb{P}(X_i \in A_i)$ for any $A_i \in \mathcal{E}_i, i \in J$ for any finite subset $J \subseteq \mathcal{I}$.
- (for real-valued) $\iff$ $\forall n \geq 1, x_1, .., x_n \in \mathbb{R}$ or $\overline{\mathbb{R}}$ $\mathbb{P}(X_1 \leq x_1, ..., X_n \leq x_n) = \mathbb{P}(X_1 \leq x_1) \cdots \mathbb{P}(X_n \leq x_n)$ (using the π-systems def)

For a finite family of RVs, they are independent $\iff$ their joint dist $\mu_{(X_1,...,X_n)}$ on the product space $(\prod_{i\leq n} E_i, \times_{i\leq n}\mathcal{E}_i)$ is the product measure of the marginal dists μ_{X_i}

if $(X_i)_{i\in\mathcal{I}}$ are independent, and $f_i : E_i \to \mathbb{R}$ are measurable, then $(f_i(X_i))_{i\in\mathcal{I}}$ are independent RVs.

Limits

The **tail σ-algebra** of a sequence of RVs $(X_n)_{n\geq 1}$ is $\mathcal{T} := \bigcap_{n=1}^{\infty} \mathcal{T}_n$, where $\mathcal{T}_n := \sigma(X_{n+1}, X_{n+2}, ...)$

Kolmogorov's $0-1$ **law**: The tail σ-algebra of an independent sequence of RVs contains only events of probability 0 or 1, so any $\mathcal{T}$-measurable RV is a.s. constant. e.g. $A = \{(X_n)_{n\geq 1}$ converges$\}$

For a sequence $(A_n)_{n\geq 1}$ of sets in $\mathcal{F}$:

$$\begin{aligned}\{A_n \text{ i.o.}\} &=: \limsup_{n\to\infty} A_n \\ &:= \bigcap_{n=1}^{\infty} \bigcup_{m\geq n} A_m \\ &= \{\omega \in \Omega : \omega \in A_m \text{ for infinitely many } m\}\end{aligned}$$

(Note i.o. = occurs infinitely often)

$$\begin{aligned}\liminf_{n\to\infty} A_n &:= \bigcup_{n=1}^{\infty} \bigcap_{m\geq n} A_m \\ &= \{\omega \in \Omega : \exists m_\omega \text{ st } \omega \in A_m \text{ for all } m \geq m_\omega\} \\ &= \{A_n \text{ eventually}\} = \{A_n^C \text{ i.o.}\}^C\end{aligned}$$

Note that $\mathbf{1}_{\limsup_{n\to\infty} A_n} = \limsup_{n\to\infty} \mathbf{1}_{A_n}$, and the same for $\liminf$

Fatou's lemma (and reverse): $\mathbb{P}(\liminf_{n\to\infty} A_n) \leq \liminf_{n\to\infty} \mathbb{P}(A_n)$, and $\mathbb{P}(A_n \text{ i.o.}) = \mathbb{P}(\limsup_{n\to\infty} A_n) \geq \limsup_{n\to\infty} \mathbb{P}(A_n)$

BC1: Borel-Cantelli Lemma #1: if $\sum_{n\geq 1} \mathbb{P}(A_n) < \infty$ then $\mathbb{P}(A_n \text{ i.o.}) = 0$.

BC2: for A_n independent, then if $\sum_{n\geq 1} \mathbb{P}(A_n) = \infty$ then $\mathbb{P}(A_n \text{ i.o.}) = 1$

4 Integration & Expectation FINISH

Notation:

$$\int f \, \mathrm{d}\mu \equiv \int_\Omega f \, \mathrm{d}\mu \equiv \int f(\boldsymbol{\omega}) \, \mu(\mathrm{d}\boldsymbol{\omega})$$

See part A for integration definitions, just using a more general measure. Includes the MCT, DCT, Fatou, Reverse Fatou etc.

Radon-Nikodym Theorem: given μ, ν are prob measures on $(\Omega, \mathcal{F})$, $\nu \ll \mu \iff \exists$ an RV $f \geq 0$ st $\nu(A) = \int_A f \, \mathrm{d}\mu$ for all $A \in \mathcal{F}$. f is the Radon-Nikodym derivative of ν wrt μ, and is often written $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$. $v \sim \mu \iff f \geq 0 \; \mu - a.s.$, and then $\frac{\mathrm{d}\mu}{\mathrm{d}\nu} = 1/f$

Scheffé: if $f_n, f \in \mathscr{L}^1(\Omega, \mathcal{F}, \mu)$ and $f_n \to f$ pointwise, then $\int |f_n - f| \, \mathrm{d}\mu \to 0 \iff \int |f_n| \, \mathrm{d}\mu \to \int |f| \, \mathrm{d}\mu$. [no proof]

If $X : \Omega \to E, g : E \to \mathbb{R}$ for a prob space $(\Omega, \mathcal{F}, \mathbb{P})$, measure space $(E, \mathcal{E})$ and μ_X is the law of X, then g is μ_X-integrable $\iff g \circ X$ is $\mathbb{P}$-integrable, and then $\int_E g(x) \, \mu_X(\mathrm{d}x) = \int_\Omega g(X(\boldsymbol{\omega})) \, \mathbb{P}(\mathrm{d}\mu)$

The **expectation** of X is $\mathbb{E}[X] := \int X \, \mathrm{d}\mathbb{P} = \int_\Omega X(\boldsymbol{\omega}) \, \mathbb{P}(\mathrm{d}\boldsymbol{\omega}) = \int_\mathbb{R} x \mu_X$

The **variance** of X is $\mathbb{E}[X - \mathbb{E}[X]^2]$, as expected. The n'th standardised moment...

Fubini/Tonelli: given $(\Omega, \mathcal{F}, \mathbb{P})$ is the product of $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i)$ for $i = 1, 2$, and $f(x, y)$ is measurable on $(\Omega, \mathcal{F})$:

- $x \mapsto \int_{\Omega_2} f(x, y) \, \mathbb{P}_2(\mathrm{d}y), y \mapsto \int_{\Omega_2} f(x, y) \, \mathbb{P}_1(\mathrm{d}x)$ are $\mathcal{F}_1, \mathcal{F}_2$-measurable resp.
- if f is $\mathbb{P}$-integrable on Ω, or $\boldsymbol{f \geq 0}$: $\int_\Omega f \, \mathrm{d}\mathbb{P}$ equals the repeated integrals (though if we are doing this because $f \geq 0$, it could be ∞)

X, Y on a prob space $(\Omega, \mathcal{F}, \mathbb{P})$ are independent $\iff \forall f, g$ measurable, $f, g \geq 0$ $\mathbb{E}[f(X)g(y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$

integration on product space - fubton

indep in terms of expectation of funcs

5 Complements & more integration

5.1 Modes of convergence

almost surely $X_n \to X$ a.s. $\iff \mathbb{P}[X_n \to X] = \mathbb{P}(\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1$

in probability $X_n \xrightarrow{\mathbb{P}} X \iff \forall \varepsilon > 0 \; \lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$

in $\mathcal{L}^p/L^p/$ pth moment $X_n \xrightarrow{L^p} X$ iff $X \in \mathcal{L}^p, \forall n\ X_n \in \mathcal{L}^p$ and $\lim_{n\to\infty} \mathbb{E}[|X_n - X|^p] = 0$

weakly in $\mathcal{L}^1$ if $X_n, X \in \mathcal{L}^1$ and $\forall$bounded RVs Y $\lim_{n\to\infty} \mathbb{E}[X_n Y] = \mathbb{E}[XY]$

weakly/in distribution if $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$ for every $x \in \mathbb{R}$ at which F_X is cont.

Relationships:

$$\begin{array}{ccccc} \text{a.s.} & \Longrightarrow & \text{in prob} & \Longrightarrow & \text{in dist} \\ & & \Uparrow & & \\ & & \text{in } L^p & \Longrightarrow & \text{weakly in } L^p \end{array}$$

5.2 Useful inequalities

Markov's inequality: if $X \geq 0$ is an RV, $\forall \lambda > 0$ $\mathbb{P}[X \geq \lambda] \leq \lambda^{-1}\mathbb{E}[X]$

Chebyshev's inequality (general): for an RV X with $\mathrm{Im}(X) \subseteq A \subseteq \mathbb{R}$ for measurable A, and $\phi : A \to [0,\infty]$ is increasing and measurable, then $\forall \lambda \in A$ with $\phi(\lambda) < \infty$

$$\mathbb{P}[X \geq \lambda] \leq \frac{\mathbb{E}[\phi(X)]}{\phi(\lambda)}$$

Useful ϕ's: x^2 on $|X - \mathbb{E}[X]|$ gives the standard form, and $e^{\theta x}$ gives $\mathbb{P}[X \geq \lambda] \leq e^{-\lambda\theta}\mathbb{E}[e^{\theta X}]$

WLLN: if $(X_n)_{n\geq 1}$ is a sequence of *iid* random variables with mean μ, variance $\sigma^2 < \infty$, then $\frac{1}{n}\sum_{i=1}^n X_i \to \mu$ in probability

Jensen's inequality: for an rv X taking values in an interval I, and $f : I \to \mathbb{R}$ a convex function, then $\mathbb{E}[f(X)] \geq f(\mathbb{E}[X])$

L^p spaces: see FA1

A collection $\mathcal{C}$ of random variables is **UI/uniformly integrable** $\iff$

$$\lim_{K\to\infty} \sup_{X\in\mathcal{C}} \mathbb{E}[|X|\mathbf{1}_{\{|X|>K\}}] = 0$$

$\iff$ both of the following hold:

$$\sup_{X\in\mathcal{C}} \mathbb{E}[|X|] < \infty$$
$$\sup_{A\in\mathcal{F}:\mathbb{P}(A)\leq\delta} \sup_{X\in\mathcal{C}} \mathbb{E}[|X|\mathbf{1}_A] \to 0 \text{ as } \delta \to 0$$

Useful points:

- this definition works a.s.
- $\{X\}$ is UI $\iff$ X is integrable
- if $\forall X \in \mathcal{C}$ $|X| \le Y$ for $Y \in \mathcal{L}^1$, then $\mathcal{C}$ is UI.
- we can replace $|X|\mathbf{1}_{X>K}$ with $(|X|-K)^+$ (or similar), as $0 \le (|X|-K)^+ \le |X|\mathbf{1}_{X>2K} \le 2(|X|-K)^+$

5.23: if $X_n \to X$ in probability, and are all bounded by $K \in \mathbb{R}$, then $X_n \to X$ in L^1

Vitali's Convergence theorem: if $X_n \to X$ in prob, $X_n \in \mathcal{L}^1$, tfae:

- $\{X_n : n \ge 1\}$ is UI
- $X \in \mathcal{L}^1$ and $\mathbb{E}[|X_n - X|] \to 0$
- $X \in \mathcal{L}^1$ and $\mathbb{E}[|X_n|] \to \mathbb{E}[|X|] < \infty$

Thus, $X_n \to X$ in L^1 $\iff$ $X_n \to X$ in probability and $\{X_n : n \ge 1\}$ is UI.

6 Conditional expectation

[Entire chapter is on $(\Omega, \mathcal{F}, \mathbb{P})$]

If X is an integrable RV, and $\mathcal{G} \subseteq \mathcal{F}$ a σ-algebra. Then $\mathbb{E}[X|\mathcal{G}] := Y$ where Y is integrable, $\mathcal{G}$-measurable &

$\forall G \in \mathcal{G},\ \mathbb{E}[Y\mathbf{1}_G] = \mathbb{E}[X\mathbf{1}_G] \iff \int_G \mathbb{E}[X|\mathcal{G}]\, d\mathbb{P} = \int_G X\, d\mathbb{P}$ (the **defining relation**)

By theorem 6.3, Y exists, and is unique almost surely (i.e. if Y, Z both follow the conditions above, then $Y = Z$ a.s.)

If we verify the defining relation for X and a candidate Y for $\mathbb{E}[X|\mathcal{G}]$ for $G = \Omega$ and $G \in \mathcal{A}$, where $\mathcal{A}$ is a π-system generating $\mathcal{G}$, we have verified it for all $G \in \mathcal{G}$ [because the DCT applied to the defining relation shows that the set of G satifying it is a λ-system, so then apply $\pi - \lambda$)

For various cases: (where X is an integrable RV)

- $\mathcal{G} = \sigma(B)$, $X = \mathbf{1}_A$ for events A, B: $\mathbb{E}[\mathbf{1}_A|\sigma(B)](\boldsymbol{\omega}) = \mathbb{P}(A|\sigma(B))(\boldsymbol{\omega}) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}\mathbf{1}_B(\boldsymbol{\omega}) + \frac{\mathbb{P}(A\cap B^C)}{\mathbb{P}(B^C)}\mathbf{1}_{B^C}(\boldsymbol{\omega})$
- $\mathcal{G} = \sigma(B)$, for an event B: $\mathbb{E}[X|\sigma(B)](\boldsymbol{\omega}) = \frac{\mathbb{E}[X\mathbf{1}_B]}{\mathbb{P}(B)}\mathbf{1}_B(\boldsymbol{\omega}) + \frac{\mathbb{E}[X\mathbf{1}_{B^C}]}{\mathbb{P}(B^C)}\mathbf{1}_{B^C}(\boldsymbol{\omega})$
- $\mathcal{G} = \sigma(B_n : n \ge 1)$, for a sequence of events B_n: $\mathbb{E}[X|\sigma(B_n : n \ge 1)](\boldsymbol{\omega}) = \sum\limits_{n\ge 1} \frac{\mathbb{E}[X\mathbf{1}_{B_n}]}{\mathbb{P}(B_n)}\mathbf{1}_{B_n}(\boldsymbol{\omega})$
- $\mathcal{G} = \sigma(Z)$ for an RV Z: $\mathbb{E}[X|Z] := \mathbb{E}[X|\sigma(Z)]$.

- this is then defined as the RV Y, which is $\sigma(Z)$ measurable, so a function of Z.
- instead of checking the defining relation, we can check: $\forall A \in \mathscr{B}(\mathbb{R})$, $\mathbb{E}[Y\mathbf{1}_{\{Z\in A\}}] = \mathbb{E}[X\mathbf{1}_{\{Z\in A\}}]$, as $G \in \sigma(Z) = Z(A)$ for A Borel.

- if $\mathcal{G} = \sigma(Z)$, where Z is a discrete RV (taking values $z_1, ...$):
  - then we don't need to check all $A \in \mathscr{B}(\mathbb{R})$, but simply the sets $\{Z = z_i\}$ (as the A's will just be countable unions of such sets)
  - we can now define $Y = \mathbb{E}[X|Z]$ explicitly as:
    * $Y(\omega) = \mathbb{E}[X|Z](\omega) := \mathbb{E}[X|\sigma(\{Z = z_i\})](\omega) = \sum_{n\geq 1} \frac{\mathbb{E}[X\mathbf{1}_{\{Z=z_i\}}]}{\mathbb{P}(Z=z_i)} \mathbf{1}_{Z=z_i}(\omega)$ where z_i st $Z(\omega) = z_i$ (note the last equality is because z_i is chosen st. $\omega \in \{Z = z_i\}$)
  - checking the defining relation:
    * $\forall z_i \in \mathbb{R}$, $\mathbb{E}[Y\mathbf{1}_{\{Z=z_i\}}] = \mathbb{E}\left[\frac{\mathbb{E}[X\mathbf{1}_{\{Z=z_i\}}]}{\mathbb{P}(Z=z_i)} \mathbf{1}^2_{Z=z_i}\right] = \mathbb{E}[X\mathbf{1}_{\{Z=z_i\}}]\mathbb{E}[\mathbf{1}^2_{Z=z_i}] = \mathbb{E}[X\mathbf{1}_{\{Z=z_i\}}]$

Useful and important properties: (6.5)

- $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$ (take $G = \Omega \in \mathcal{G}$)
- linear (a.s.)
- $\mathbb{E}[X|\mathcal{G}] = X$ a.s. if X is $\mathcal{G}$-measurable (satisfies def rel)
- $\mathbb{E}[c|\mathcal{G}] = c$ a.s.
- $\mathbb{E}[X|\{\emptyset, \Omega\}] = \mathbb{E}[X]$
- $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$ a.s. if $\sigma(X), \mathcal{G}$ are independent
- $X \leq Y$ a.s. $\implies$ $\mathbb{E}[X|\mathcal{G}] \leq \mathbb{E}[Y|\mathcal{G}]$ a.s.
  $\implies$ $|\mathbb{E}[X|\mathcal{G}]| \leq \mathbb{E}[|X|\,|\mathcal{G}]$ a.s.

Convergence theorems:

- **cMCT:** $X_n \geq 0, X_n \uparrow X \implies \mathbb{E}[X_n|\mathcal{G}] \uparrow \mathbb{E}[X|\mathcal{G}]$ a.s.
- **cFatou**: $X_n \geq 0 \implies \mathbb{E}[\liminf_{n\to\infty} X_n|\mathcal{G}] \leq \liminf_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}]$ a.s.
- **cDCT:** Y integrable, $|X_n| \leq Y, X_n \to X$ a.s $\implies \mathbb{E}[X_n|\mathcal{G}] \to \mathbb{E}[X|\mathcal{G}]$ a.s.

taking out what's known: X, Y rvs with X, Y, XY integrable, Y $\mathcal{G}$-measurable. Then $\mathbb{E}[XY|\mathcal{G}] = Y \cdot \mathbb{E}[X|\mathcal{G}]$ a.s.

tower property: $X \in \mathcal{L}^1, \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}$ then $\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{F}_2\right]|\mathcal{F}_1\right] = \mathbb{E}\left[X|\mathcal{F}_1\right]$ a.s.

cJensen's: $X \in \mathcal{L}^1, \mathrm{Im}(X) \subseteq I$, an interval, and $f : I \to \mathbb{R}$ is convex, $\mathbb{E}[|f(X)|] < \infty$ then $\mathbb{E}[f(X)|\mathcal{G}] \geq f(\mathbb{E}[X|\mathcal{G}])$ a.s.

For an integrable RV X, a family of σ-algebras $\{\mathcal{F}_\alpha : \alpha \in I\}$ where $\forall\alpha\ \mathcal{F}_\alpha \subseteq \mathcal{F}$ then $\{X_\alpha := \mathbb{E}[X|\mathcal{F}_\alpha] : \alpha \in I\}$ is UI.

Law of total expectation (for calc. use only): $\mathbb{E}[X] = \sum_i \mathbb{E}[X\mathbf{1}_{A_i}] = \sum_i \int X\ \mathbb{P}(\mathrm{d}\boldsymbol{\omega} \mid A_i) \cdot \mathbb{P}(A_i) \approx \sum_i \mathbb{E}[X \mid A_i]\mathbb{P}[A_i]$ given $\{A_i\}_{i\geq 1}$ forms a finite/countable partition of Ω, where $\mathbb{P}(\cdot \mid A_i)$ is the conditional probability measure defined at (1), and $\mathbb{E}[X \mid A_i]$ is the old conditional expectation.

Orthogonal projection (just for proof of existence of cond exp)

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

X, Y are **uncorrelated** if $\mathrm{Cov}(X,Y) = 0$ (indep $\implies$ uncorrelated)

$\langle X, Y\rangle := \mathbb{E}[XY]$ is a scalar product for $X, Y \in \mathscr{L}^2$, and X, Y are **orthogonal** if $\langle X, Y\rangle = 0$.

Pythagoras's theorem: if $X, Y \in \mathscr{L}^2$ are othorgonal, then $\|X+Y\|_2^2 = \|X\|_2^2 + \|Y\|_2^2$

If $\mathcal{H}$ is a complete vector subspace of $\mathscr{L}^2$,for any $X \in \mathscr{L}^2$the infinimum $\inf_{Z\in\mathcal{H}} \|X-Z\|_2$ is achieved by some $Y \in \mathcal{H}$, and $(X-Y)$ is othorgonal to all $Z \in \mathcal{H}$.

Thus, for $\mathbb{E}[X \mid \mathcal{G}]$, $\mathcal{H} := \mathscr{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ is a complete vector space, so we can project X onto $\mathcal{H}$ using the theorem above to get $Y \in \mathcal{H}$ (i,e, Y is $\mathcal{G}$-measurable), and then Y (by $\mathbf{1}_G$ for $G \in \mathcal{G}$ and cMCT) is a version of $\mathbb{E}[X \mid \mathcal{G}]$.

7 Filtration & stopping times

A **filtration** on $(\Omega, \mathcal{F}, \mathbb{P})$ is a sequence $(\mathcal{F}_n)_{n\geq 0}$ of σ-algebras $\mathcal{F}_n \subseteq \mathcal{F}$ st $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ for all n.

$\mathcal{F}_\infty := \sigma\left(\bigcup_{n\geq 0} \mathcal{F}_n\right)$ is the σ-algebra **generated by the filtration**.

$(X_n)_{n\geq 0}$ is **adapted** to $(\mathcal{F}_n)_{n\geq 0}$ if $\forall n$ X_n is $\mathcal{F}_n$-measurable. It is **integrable** if each X_n is integrable.

The natural filtration of $(X_n)_{n\geq 0}$ is $\mathcal{F}_n^X = \sigma(X_0, X_1, ..., X_n)$

An RV $\tau : \Omega \to \mathbb{N} \cup \{\infty\}$ is a **stopping time** wrt $(\mathcal{F}_n)_{n\geq 0}$ if $\forall n$ $\{\tau = n\} \in \mathcal{F}_n$, or equivalently $\{\tau \leq / \geq / < / > n\} \in \mathcal{F}_n$.

A constant is a stopping time, as are the max & min of 2 stopping times.

The hitting time $h_B := \inf\{n \geq 0 : X_n \in B\}$ of an adapted process $(X_n)_{n\geq 0}$ of a Borel set B is a stopping time.

The σ-algebra at time τ (a stopping time) is $\mathcal{F}_\tau := \{A \in \mathcal{F}_\infty : \forall n \geq 0 \ A \cap \{\tau = n\} \in \mathcal{F}_n\}$.

- (Note that $\tau = n$ can have any inequality operator).
- $\tau \leq \rho \implies \mathcal{F}_\tau \subseteq \mathcal{F}_\rho$

X_τ is an RV if $\tau < \infty$, defined as $\omega \mapsto (X_{\tau(\omega)})(\omega)$, and is $\mathcal{F}_\infty$ and $\mathcal{F}_\tau$-measurable.

The **stopped process** of $(X_n)_{n\geq 0}$ and τ is $X^\tau = (X_{\tau\wedge n})_{n\geq 0}$, which is adapted to the filtrations $(\mathcal{F}_{\tau\wedge n})_{n\geq 0}$ and $(\mathcal{F}_n)_{n\geq 0}$

8 Martingales in discrete time

An integrable, $\mathcal{F}_n$-adapted stochastic process $(X_n)_{n\geq 0}$ is a

martingale if $\forall n \geq 0\ \mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$ a.s.

submartingale if $\forall n \geq 0\ \mathbb{E}[X_{n+1}|\mathcal{F}_n] \geq X_n$ a.s. **sub = bigger**

supermartingale if $\forall n \geq 0\ \mathbb{E}[X_{n+1}|\mathcal{F}_n] \leq X_n$ a.s. **super = smaller**

Basics:

- $(X_n)_{n\geq 0}$ is a submartingale wrt to $(\mathcal{F}_n)_{n\geq 0} \iff (-X_n)_{n\geq 0}$ is a supermartingale wrt to $(\mathcal{F}_n)_{n\geq 0}$
- $(X_n)_{n\geq 0}$ wrt to $(\mathcal{F}_n)_{n\geq 0}$ is a martingale $\iff$ it is a sub & super martingale.
- $(X_n)_{n\geq 0}$ wrt to $(\mathcal{F}_n)_{n\geq 0}$ is a submartingale $\implies (X_n)_{n\geq 0}$ is a submartingale wrt any smaller filtration (incl. $\mathcal{F}^X$)
- $\mathbb{E}[X_n|\mathcal{F}_m] = X_m$ a.s. if $n \geq m$ where X is a (-/sub/super)martingale [$(= / \geq / \leq)$]
- $\mathbb{E}[X_n] = \mathbb{E}[X_m] = \mathbb{E}[X_0]$ a.s. if $n \geq m$ where X is a (-/sub/super)martingale [$(= / \geq / \leq)$ for both $=$ signs]

An integrable sequence $(Y_n)_{n\geq 1}$is a **martingale difference sequence** wrt $(\mathcal{F}_n)$ if $\forall n \geq 0\ \ \mathbb{E}[Y_{n+1} \mid \mathcal{F}_n] = 0$ a.s.

if f is a convex function on $\mathbb{R}$, and $(X_n)_{n\geq 0}$ is a martingale wrt to $(\mathcal{F}_n)_{n\geq 0}$, $(f(X_n))_{n\geq 0}$ is a submartingale wrt to $(\mathcal{F}_n)_{n\geq 0}$. E.g.: $|X_n|, X_n^2, e^{X_n}e^{-X_n}, \max\{X_n, K\}$

Given a filtration $(\mathcal{F}_n)_{n\geq 0}$, a sequence $(V_n)_{n\geq \mathbf{1}}$ is **predictable** if $\forall n \geq 1$, V_n is $\mathcal{F}_{n-1}$-measurable

Given a predictable sequence $(V_n)_{n\geq \mathbf{1}}$ and a (-/super/sub)martingale $(X_n)_{n\geq 1}$ on a filtration $(\mathcal{F}_n)_{n\geq 0}$ their **martingale transform** is, which is a martingale wrt $(\mathcal{F}_n)_{n\geq 0}$

$$((V \circ X)_n)_{n\geq 0} := \left(\sum_{k=1}^{n} V_k(X_k - Y_{k-1})\right)_{n\geq 0}$$

(note the 0th term is either 0 or X_0, depending)

Doob's Decomposition theorem: given an integrable adapted process $\boldsymbol{X} = (X_n)_{n\geq 0}$ on $(\mathcal{F}_n)_{n\geq 0}$,

- $\boldsymbol{X}$ has a Doob decomposition $X_n = X_0 + M_n + A_n$, where M_n is a martingale and A_n predictable, both on $(\mathcal{F}_n)_{n\geq 0}$, and $M_0 = A_0 = 0$.
- This decomposition is unique in probability - i.e. $\mathbb{P}(M_n = M_n', A_n = A_n'$ for all $n \geq 0) = 1$

- $\boldsymbol{X}$ is a submartingale $\iff$ $(A_n)_{n\geq 0}$ is an increasing process a.s., and a supermartingale $\iff$ $(A_n)_{n\geq 0}$ is an decreasing process a.s.. Thus, $\boldsymbol{X}$ is a martingale $\iff$ $A_{n+1} = A_n$ a.s.

Given a L^2-martingale $\boldsymbol{M}$, i.e. $\mathbb{E}[M_n^2] < \infty$, we can consider the Doob decomposition $M_n^2 = M_0^2 + N_n + A_n$, where A_n is increasing, as x^2 is a convex function. we call $(\langle M \rangle_n)_{n\geq 0} = (A_n)_{n\geq 0}$

Given a martingale $\boldsymbol{X}$ and a finite stopping time τ, $\boldsymbol{X}^\tau$ is the **stopped process** of $\boldsymbol{X}$ and τ, and is a martingale wrt $(\mathcal{F}_{\tau\wedge n})_{n\geq 0}$ and $(\mathcal{F}_n)_{n\geq 0}$

Doob's Optional Sampling/Stopping Theorem: given a martingale $\boldsymbol{X}$ wrt $(\mathcal{F}_n)_{n\geq 0}$, and bounded stopping times $\tau \leq \rho$:

- $\mathbb{E}[X_\rho] = \mathbb{E}[X_\tau] = \mathbb{E}[X_0]$
- $\mathbb{E}[X_\rho | \mathcal{F}_\tau] = X_\tau$ a.s.
- (same for sub/super)

Variants of the above: for a martingale $\boldsymbol{X}$ wrt $(\mathcal{F}_n)_{n\geq 0}$, and τ is $a.s.$ finite (i.e. $\tau < \infty$ except on a null set), then $\mathbb{E}[X_\tau] = \mathbb{E}[X_\tau \mathbf{1}_{\tau<\infty}] = \mathbb{E}[X_0]$ if either:

1. $\{X_n : n \geq 0\}$ is UI $\impliedby$ SOMETHING CHECK LAST SHEET
2. $\mathbb{E}[\tau] < \infty$ and $\exists L \in \mathbb{R}$ st $\forall n \, \mathbb{E}[|M_{n+1} - M_n| \, |\mathcal{F}_n] \leq L$ a.s.

Doob's maximal inequality: if $(X_n)_{n\geq 0}$is a submartingale, $\forall \lambda > 0$, $Y_n^\lambda := (X_n - \lambda)\mathbf{1}_{\{\max_{k\leq n} X_k \geq \lambda\}}$ is a submartingale, and in particular $\lambda \mathbb{P}[\max_{k\leq n} X_n \geq \lambda] \leq \mathbb{E}[X_n \mathbf{1}_{\{\max_{k\leq n} X_k \geq \lambda\}}] \leq \mathbb{E}[|X_n|]$

a corollary: for $p \geq 1, (M_n)_{n\geq 0}$ a martingale with $M_n \in \mathscr{L}^p$,then $\forall N \geq 0, \lambda > 0$ $\mathbb{P}\left[\max_{n\leq N} |M_n| \geq \lambda\right] \leq \dfrac{\mathbb{E}\left[|M_N|^p\right]}{\lambda^p}$

Doob's $\boldsymbol{L}^p$ inequality: for $p > 1, (X_n)_{n\geq 0}$a non-negative submartingale, $X_n \in \mathscr{L}^p$, $\overline{X}_n := \max_{k\leq n} X_K$ is in $\mathscr{L}^p$, and $\mathbb{E}[X_n^p] \leq \mathbb{E}[\max_{k\leq N} X_k^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[X_n^p]$

if $(X_n)_{n\geq 0}$ is a supermartingale, $\forall \lambda, n \geq 0$ then $\lambda \mathbb{P}(\max_{k\leq n} |X_k| \geq \lambda) \leq \mathbb{E}[X_0] + 2\mathbb{E}[X_0] + 2\mathbb{E}[X_n^-]$

Upcrossings

For a sequence $\boldsymbol{x} = (x_n)_{n\geq 0}$ of real numbers, fixed $a < b$, define sequences $(\rho_k)_{k\geq \mathbf{1}}, (\tau_k)_{k\geq \mathbf{0}}$ by:

$$\tau_0 = 0$$
$$\rho_k = \inf\{n \geq \tau_{k-1} : x_n \leq a\}$$
$$\tau_k = \inf\{n \geq \rho_k : x_n \geq b\}$$

$U_n([a,b],\boldsymbol{x}) := \max\{k \geq 0 : \tau_k \leq n\}$ is the **number of upcrossings** of $[a,b]$ by x **by time** n, and $U([a,b],\boldsymbol{x}) := \sup_n U_n([a,b],\boldsymbol{x}) = \sup\{k \geq 0 : \tau_k < \infty\}$ is the **total number of upcrossings**.

Doob's upcrossings lemma: $\boldsymbol{X} = (X_n)_{n\geq 0}$ is a supermartingale, $a < b$ fixed, $\forall n \geq 0$: $\mathbb{E}[(U_n([a,b],\boldsymbol{X})] \leq \dfrac{\mathbb{E}[(X_n - a)^-]}{b-a}$

a real sequences $\boldsymbol{x}$ converges iff $\forall a, b \in \mathbb{Q}, a < b\ U([a,b],\boldsymbol{x}) < \infty$

$(X_n)_{n\geq 1}$ is **bounded in** L^p if $\sup_n \mathbb{E}[|X_n|^p] < \infty$

Doob's Forward convergence theorem: if $\boldsymbol{X}$ is a sub-/super-martingale, and bounded in L^1, then it converges a.s. to a limit X_∞ (precisely, $\mathbb{P}[X_n \to X_\infty] = 1$), which is integrable.

Thus, if $(X_n)_{n\geq 0}$ is a non-negative supermartingale, then $X_\infty = \lim_{n\to\infty} X_n$ exists a.s. (note no constraints on L^1, as $\mathbb{E}[|X_n|] = \mathbb{E}[X_n] \leq \mathbb{E}[X_0]$)

UI again:

TFAE, for a martingale $(M_n)_{n\geq 0}$:

- $\boldsymbol{M}$ is UI
- $\exists M_\infty$, which is $\mathcal{F}_\infty$ measurable st $M_n \to M_\infty$ a.s. and in L^1
- $\exists M_\infty$, which is $\mathcal{F}_\infty$ measurable st $\forall n : M_n = \mathbb{E}[M_\infty \mid \mathcal{F}_n]$ a.s.

and if $M_\infty \in \mathscr{L}^p$ for some $p > 1$, then $M_n \to M_\infty$ in $\mathscr{L}^p$.

If $\boldsymbol{M}$ is a UI martingale, for all (potentially unbounded) stopping times $\tau \leq \rho$ $\mathbb{E}[M_\rho \mid \mathcal{F}_\tau] = M_\tau$ a.s., and $\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$

If $\boldsymbol{M}$ is a UI martingale, let $M_\infty^* := \max_{n\geq 0} |M_n|$ then $\lambda\mathbb{P}[M_\infty^* \geq \lambda] \leq \mathbb{E}|M_\infty|\mathbf{1}_{\{M_\infty^* \geq \lambda\}}]$ for $\lambda \geq 0$.

Further, if $M_\infty \in \mathscr{L}^p$ for $p > 1$, let q st $1/p + 1/q = 1$ then $\|M_\infty\|_p \leq \|M_\infty^*\|_p \leq q\|M_\infty\|_p$, and $M_n \to M_\infty$ in $\mathscr{L}^p$.

9 Some applications

Backwards martingales: time is indexed by $I = \{t \in \mathbb{Z} : t \leq 0\}$, and a backwards martingaleis written $(M_{-n})_{n\geq 0}$, and ends at 0.

Given a sequence of σ-algebras $(\mathcal{F}_{-n})_{n\geq 0}$, with $\mathcal{F}_{-n} \subseteq \mathcal{F}$, and $\mathcal{F}_{-k} \subseteq \mathcal{F}_{-k+1}$ for all $k \leq -1$, M_{-n} is a **backwards martingale** if $\forall n : M_{-n}$ is integrable and $\mathcal{F}_{-n}$ measurable, and $\mathbb{E}[M_{-n+1} \mid \mathcal{F}_{-n}] = M_{-n}$ a.s.

Thus, $M_{-n} = \mathbb{E}[M_0 \mid \mathcal{F}_{-n}]$ a.s., and so $(M_{-n})_{n\geq 0}$ is UI.

Doob's Upcrossing lemma automatically holds, as it is actually a result about about finite martingales

And, as $n \to \infty, M_{-n}$ converges a.s. to a random limit $M_{-\infty}$.

$\mathcal{F}_{-\infty} = \bigcap_{k=0}^{\infty} \mathcal{F}_{-k}$ - note that the σ-algrebras get smaller as k increases

$M_{-\infty}$ is $\mathcal{F}_{-\infty}$ and $\mathcal{F}_{-k}$ integrable

convergence to $M_{-\infty} = \mathbb{E}[M_0 \mid \mathcal{F}_{-\infty}]$ is both a.s. and in L^1.

Kolmogorov's Strong LLN: For a sequence $(X_n)_{n\geq 1}$ of IID RVs, each integrable with mean m , set $S_n = \sum_{k=1}^{n} X_k$, and then $S_n/n \to m$ a.s. and in L^1 as $n \to \infty$.