www.mymathscloud.com

Probability
3105

Based on lectures by
Dr N Sidorova
typed by John Sylvester

Jan 2012

Contents

www.mymathscloud.com
My Maths Cloud

1 Rigorous set up

1.1 Probability space, events and random variables

Definition 1.1 (σ-algebra of sets). *Let Ω be a set and Σ be a collection of sets. Then Σ is a σ - algebra if*

1. $\Omega,\ \phi \in \Sigma$
2. $A \in \Sigma$ *then* $\Omega / A \in \Sigma$
3. $A_1, A_2, \dots \in \Sigma$ *then* $\cup_1^\infty A_i \in \Sigma$

Definition 1.2 (Measure). $\mu : \Sigma \to [0, \infty]$ *is called a measure if*

1. $\mu(\phi) = 0$
2. $A_1, A_2, \dots$ *are disjoint then* $\mu(\cup_1^\infty A_i) = \sum_1^\infty \mu(A_i)$

Definition 1.3 (Probability Measure). *A measure μ is called a Probability Measure denoted by P if*

$$P(\Omega) = 1$$

Definition 1.4 (Probability Space). *A triple (Ω, Σ, P), where Ω is a set, Σ is a σ-algebra and P is a probability measure, is called a Probability Space.*

Definition 1.5 (Measurable Function). *a function X is called a Measurable function if*

$$\forall B \in \mathcal{B} \quad X^{-1}(B) = \{w \ : \ X(w) \in \mathcal{B}\} \in \Sigma$$

Definition 1.6 (Random Variable). *A random variable is called a measurable function*

$$X : \Omega \to \mathbb{R}$$
$$(\Omega, \Sigma) \to (\mathbb{R}, \underbrace{\mathcal{B}}_{borel\ \sigma-algebra})$$

The idea:
Ω - Random outcomes
Σ - All possible events
$P(E)$ - Probability of the event E

Example 1.1. *Bernulli = "tossing a coin" = 0 or 1 with probability $\frac{1}{2}$*

$$\Omega = \{H, T\}$$
$$\Sigma = \{\{H\}, \{T\}, \{H, T\}, \phi\} = 2^\Omega$$
$$P(\{H\}) = P(\{T\}) = \frac{1}{2}$$
$$P(\{H, T\}) = 1$$
$$P(\phi) = 0$$

www.mymathscloud.com
My Maths Cloud

$$X : H \to 1$$
$$T \to 0$$

$$\text{"Probability that } X = 1\text{"} = P(\omega : X(\omega) = 1) = P(\{H\}) = \frac{1}{2}$$

Example 1.2. *Roll a die, spell the number, take # of letters*

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$
$$\Sigma = 2^{\Omega} \quad (\sigma\text{-algebra of all subsets})$$
$$P(\{1\}) = \cdots = P(\{6\}) = \frac{1}{6}$$
$$P(\{1, 3, 5\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2} \ etc\ldots$$

$$\begin{aligned} X :\quad & 1 \to 3 \\ & 2 \to 3 \\ & 3 \to 5 \\ & 4 \to 4 \\ & 5 \to 4 \\ & 6 \to 3 \end{aligned}$$

Example 1.3. *Roll a die, spell the number, take # of letters but with the possibility of the dice rolling off the table and scoring* 0

$$\Omega = \{1, 2, 3, 4, 5, 6, 0\}$$
$$\Sigma = 2^{\Omega} \quad (\sigma\text{-algebra of all subsets})$$
$$P(\{1\}) = \cdots = P(\{6\}) = \frac{1}{6} \quad P(\{0\}) = 0$$

$$\begin{aligned} X :\quad & 0 \to 4 \\ & 1 \to 3 \\ & 2 \to 3 \\ & 3 \to 5 \\ & 4 \to 4 \\ & 5 \to 4 \\ & 6 \to 3 \end{aligned}$$

Example 1.4. *Tossing a fair coin infinitely many times*

$$\Omega = [0, 1]$$
$$\Sigma = ?$$

www.mymathscloud.com

we can then represnt each event as a real number in $[0,1]$ *for example all events where first three results are* HTH *rest unknown:*

$$\{\omega = 0.101 * **\} = \left[\frac{5}{8}, \frac{3}{4}\right]$$

All binary intervals must be in Σ. *The minimal* σ*-algebra with this property is* $\mathcal{B}$ *the borel* σ*-algebra.*

$$P(\omega : 0.101 * * * * * \dots) = leb\left[\frac{5}{8}, \frac{3}{4}\right] = \frac{1}{8}$$

$$\Rightarrow P \text{ is a lebesque measure}$$

some questions that could be asked:
if $\omega = \omega_1, \omega_2, \dots$

(a) *what is the number in the* n^{th} *position*
$[0,1] \to \mathbb{R}$
$0.\omega_1\omega_2 \dots \mapsto \omega_n$

(b) *how many* 1*'s out of the first* n *tosses?*
$[0,1] \to \mathbb{R}$
$0.\omega_1\omega_2 \dots \mapsto \omega_1 + \dots + \omega_n$

Definition 1.7. *An event is an element of* Σ

Suppose an event E occurs = suppose $\omega \in E$. An event occurs with probability $p = P(E)$. We are interested in: $P(X \in B) \equiv P(\omega \ : \ X(\omega) \subseteq B)$

Definition 1.8.

$$\mu_X(B) := P(X \in B), \quad B \in \mathcal{B}$$

This probability measure on $(\mathbb{R}, \mathcal{B})$ *is called the distribution of* X *or the law of* X.

$\mathcal{B}$ is generated by $\{(-\infty, t], \ t \in \mathbb{R}\}$

1.2 Distribution function

Definition 1.9.

$$F_X(t) := \mu_X((-\infty, t]) = P(X \leq t) \quad t \in \mathbb{R}$$

$F_X(t)$ *is called the distribution function of* X

Example 1.5. *Bernoullli:*

$$\mu_X(B) := P(X \in B) = \begin{cases} 1 & \text{if } 0,1 \in B \\ \frac{1}{2} & \text{if just one of } 0,1 \in B \\ 0 & \text{otherwise} \end{cases}$$

Definition 1.10. *Properties of* F_X*:*

My Maths Cloud
www.mymathscloud.com

1. *F_X is increasing*

2. *$F_X \to 1$ as $t \to \infty$ and $F_X \to 0$ as $t \to -\infty$*

3. *F_X is right continuous*

Proof.
(1):
$t_1 < t_2$

$$F_X(t_1) = P(X \leq t_1)$$
$$F_X(t_2) = P(X \leq t_2).$$

$$t_1 < t_2 \Rightarrow \{X \leq t_1\} \subset \{X \leq t_2\}$$
$$\therefore P(X \leq t_1) \leq P(X \leq t_2)$$

(2):
Let $t_n \to \infty$ (need countability with σ-algebras).

$$F_X(t_n) = P(X \leq t_n) \to P\left(\cup_{n \in \mathbb{N}} \{X \leq t_n\}\right) = P(\Omega) = 1$$

Let $t_n \searrow \infty$

$$F_X(t_n) = P(X \leq t_n) \to P\left(\cap_{n \in \mathbb{N}} \{X \leq t_n\}\right) = P(\phi) = 0$$

(3):
Let $t_n \searrow t$

$$F_X(t_n) = P(X \leq t_n) \to P\left(\cap_{n \in \mathbb{N}} \{X \leq t_n\}\right) = P(X \leq t) = F_X(t)$$

$\square$

Theorem 1.1 (Skorokhod Representation)**.** *If $F : \mathbb{R} \to [0,1]$ satisfies (1)-(3) from def 1.10 above then there is a randon variable X on the probability space $([0,1], \mathcal{B}, leb)$ such that*

$$F_X = F$$

Idea. *If F is invertable, then take $G = F^{-1}$ and $X(\omega) = G(\omega)$*

$$F_X(t) = leb(\omega : X(\omega) \leq t) = F(t)$$

Proof. Define $G : [0,1] \to \mathbb{R}$

$$G(\omega) = \inf\{t : F(t) > \omega\}$$

Define $X(\omega) = G(\omega)$
need to prove:

$$F_X(u) = leb\{\omega : G(\omega) \leq u\} \stackrel{?}{=} F(u)$$

$$i.e\ F_X(u) = leb\{\omega : \inf\{t : F(t) > \omega\} \leq u\} \stackrel{?}{=} F(u)$$

It suffices to show:

$$[0, F(u)) \subset \{\omega : \inf\{t : F(t) > \omega\} \leq u\} \subset [0, F(u)]$$

www.mymathscloud.com
My Maths Cloud

(a) Let $\omega \in [0, F(u))$

$$\begin{aligned}
&\Rightarrow \omega < F(u)\\
&\Rightarrow u \in \{t : F(t) > \omega\}\\
&\Rightarrow \inf\{t : F(t) > \omega\} \leq u\\
&\Rightarrow \omega \in \text{ "middle set"}
\end{aligned}$$

(b) Let ω be such that $\inf\{t : F(t) > \omega\} \leq u$
monotonicity of F:

$$F(\inf\{t : F(t) > \omega\}) \leq F(u)$$

right continuity:

$$\inf\{F(t) : F(t) > \omega\} < F(u)$$

So

$$\omega \leq \inf\{F(t) : F(t) > \omega\} < F(u)$$

$\square$

Example 1.6. *Uniform Distribution*

$$F(t) = \begin{cases} 0 & if\ t < 0 \\ t & if\ t \in [0,1] \\ 1 & if\ t > 1 \end{cases}$$

X ; $([0,1], \mathcal{B}, leb)$

$$X(\omega) = \omega \quad \textit{(jumps to } \infty, \ -\infty \textit{ outside of } [0,1])$$

$0.\omega_1\omega_2\omega_3\ldots$ - *uniform random variable on* $[0,1]$

Example 1.7. *Exponential random variable (with mean μ).*

$$F(t) = \begin{cases} 1 - e^{-t/\mu} & if\ t \geq 0 \\ 0 & otherwise \end{cases}$$

Example 1.8. *Normal,* $N(\underset{mean}{\mu}, \underset{variance}{\sigma^2})$

$$F(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{t} \exp(-\frac{(u-\mu)^2}{2\sigma^2}) du$$

Example 1.9. *Poisson Distribution*

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$$

www.mymathscloud.com
My Maths Cloud

$([0,1], \mathcal{B}, leb)$

$$
\begin{aligned}
&\Omega = \{0,1,2,\dots\} \\
&\Sigma = 2^{\Omega} \\
&P(\{k\}) = e^{-\lambda}\frac{\lambda^k}{k!} \qquad \leftarrow \textit{ probability measure complicated} \\
&X : 0 \to 0 \\
&\quad : 1 \to 1 \\
&\quad : 2 \to 2 \qquad\qquad \leftarrow \textit{ Simple} \\
&\quad : 3 \to 3
\end{aligned}
$$

By using Skorohod Representation, thm 1.1, we keep the probability measure simple and random variable function complicated.

Definition 1.11. *If one can write $F_X(t) = \int_{-\infty}^{t} f_X(u)du$ then the law/distribution is called continuous and f_X is called the density.*

$$\mu_X((-\infty,t]) = F_X(t)du = \int_{-\infty}^{t} f_X(u)du = \int_{(-\infty,t]} f_X(u)d\,leb(u)$$

Remark.

(1)

$$
\begin{aligned}
X \textit{ has a densty} &\Leftrightarrow \textit{the law is continuous} \\
&\Leftrightarrow X \textit{ is absolutely continuous w.r.t leb and:} \\
&\quad f_X = \frac{d\mu_X}{d\,leb} \quad \textit{(Radon - Nikodym Density)}
\end{aligned}
$$

(2) if F is differentiable then

$$f_x = F_x'$$

(3) Exponential:

$$F(t) = \begin{cases} 1 - e^{-t/\mu} & if\ t \geq 0 \\ 0 & otherwise \end{cases}$$

$$f_x(t) = \begin{cases} \frac{1}{\mu}e^{-t/\mu} & if\ t \geq 0 \\ 0 & otherwise \end{cases}$$

Normal:

$$f_X(t) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{(t-\mu)^2}{2\sigma^2})$$

www.mymathscloud.com
My Maths Cloud

(4)

$$\int_{-\infty}^{\infty} f_X(u)du = 1$$

$$\lim_{x\to\infty} \int_{-\infty}^{x} f_X(t)dt = \lim_{x\to\infty} F(X)$$

1.3 Expectation and varience

Remark (Reminder from measure theory).

1.

$$f = \sum_{i=1}^{n} a_i \mathbb{1}_{B_i} \quad \textit{(simple functions)}$$

$$\int f dP = \sum_{i=1}^{n} a_i P(B_i)$$

2. $f \geq 0$, *take simple functions* f_n *where* $f_n \nearrow f$. *Define*

$$\int f dP = \lim_{n\to\infty} \int f_n dP \quad \in [0,1]$$

3. *for arbitrary* f

$$f = f^+ + (-f^-)$$

Define

$$\int f dP = \int f^+ dP - \int f^- dP$$

if both are finite. We say the function is non-lebesque measurable otherwise.

MCT *suppose* $f_n : \Omega \to [0,\infty]$ *and* $f_n \nearrow f$ *a.s. Then:*

$$\int f_n dP \to \int f dP$$

DCT *suppose* $f_n \to f$ *a.s. and* $|f_n(\omega)| < g(\omega)$ *a.s. where* $\int g(\omega) < \infty$ *Then:*

$$\int f_n dP \to \int f dP$$

Definition 1.12. *Let* X *be a random variable on* (Ω, Σ, P). *If* X *is integrable then*

$$EX = \int X dP$$

this is called the expectation of X. *If* $X > 0$, *we allow the case* $EX = \infty$

www.mymathscloud.com
My Maths Cloud

Definition 1.13. *If X is square integrable (X^2 is integrable), then*

$$VarX = E(X - EX)^2$$

this is called the variance of X.

$$E(X - EX)^2 = E(X^2 - 2X \cdot EX + (EX)^2) = EX^2 - 2(EX)^2 + (EX)^2 = EX^2 - (EX)^2$$

Lemma 1.2. *If $EX^2 < \infty$ then $E|X| < \infty$ and so $EX < \infty$*

Proof.

$$E|X| = E|X| \cdot 1 \leq \underbrace{\sqrt{EX^2}}_{finite} \cdot \underbrace{\sqrt{E1^2}}_{1} < \infty$$

$\square$

Theorem 1.3 (Chebyshev inequality)**.**
Let X be a non negitive r.v. then for any $c > 0$:

$$P(X > c) \leq c^{-1} EX$$

Proof. Define

$$Y(\omega) = \begin{cases} c & if\ \omega\ is\ st\ X(\omega) > c \\ 0 & otherwise \end{cases}$$

$$Y \leq X\ a.s.\ \Rightarrow \underbrace{EY}_{=c \cdot P(X>c)} \leq EX$$

$\square$

Theorem 1.4. *Let X be a random variable on (Ω, Σ, P) and $h : \mathbb{R} \to \mathbb{R}$ integrable on $(\mathbb{R}, \mathcal{B}, leb)$. Then*

$$Eh(X) = \int h(u) d\mu_X(u)$$

$$\begin{tikzcd} \Omega \arrow[r, "X"] \arrow[dr] & \mathbb{R} \arrow[d, "h"] \\ & \mathbb{R} \end{tikzcd}$$

Proof.

1. $h = \mathbb{1}_B,\ b \in \mathcal{B}$

$$Eh(X) = E\mathbb{1}_B(X) = 1 \cdot P(X \in B)$$

$$\int h(u) d\mu_X(u) = \int \mathbb{1}_B(u) d\mu_X(u) = 1 \cdot \mu_X(B) = P(X \in B)$$

www.mymathscloud.com
My Maths Cloud

2. $h = \sum_{i=1}^{m} a_i \mathbb{1}_{B_i}$

The formula holds by linearity of the integral

3. $h \geq 0$.

$$\underset{\substack{(simple,\\ positive)}}{h_n} \nearrow h \overset{MCT}{\Rightarrow} \int h_n d\mu_X \to \int h d\mu_X$$

$$\underset{\substack{(simple,\\ positive)}}{h_n(X)} \nearrow h(X) \overset{MCT}{\Rightarrow} Eh_n(X) \to Eh(X)$$

4. h arbitrary

$$h = \underbrace{h^+}_{\geq 0} - \underbrace{h^-}_{\geq 0} \Rightarrow Eh(X) = \int h(u) d\mu_X(u)$$

$\square$

Corollary 1.1. *if X has density $f(u)$ then*

$$EX = \int h(u) f(u) du$$

in particular if $h(X) = X$

$$EX = \int u f(u) du \quad \textit{(old formula)}$$

and for $h(X) = X^2$

$$EX^2 = \int u^2 f(u) du$$

If X has finitely, or countably, many values

$$Eh(X) = \sum_{i=1}^{n} h(a_i) \cdot P(X = a_i)$$

in particular if $h(X) = X$

$$EX = \sum_{i=1}^{n} a_i \cdot P(X = a_i) \quad \textit{(old formula)}$$

Example 1.10. *X- Bernulli*

$$EX = 1 \cdot P(X = 1) + 0 \cdot P(X = 0) = \frac{1}{2}$$

$$EX^2 = 1^2 \cdot P(X = 1) + 0 \cdot P(X = 0) = \frac{1}{2}$$

$$VarX = EX^2 - (EX)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

www.mymathscloud.com
My Maths Cloud

Example 1.11. *N(0,1)*

$$EX = \int_{-\infty}^{\infty} t \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = 0$$

$$EX^2 = \int_{-\infty}^{\infty} t^2 \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = 0$$

Example 1.12. *n people collecting their suitcases at random*

$$\begin{array}{cccc} 1, & 2, & \dots, & n \\ \downarrow & \downarrow & & \downarrow \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{array}$$

Pick a random permutaton σ uniformly with probability $\frac{1}{n!}$
Probability of everyone getting wrong suitcase?
Expected number of correct suitcases?

$$N = X_1 + X_2 + \cdots + X_n \quad X_i = \begin{cases} 1 & \text{if the ith passenger collected correct suitcase} \\ 0 & \text{otherwise} \end{cases}$$

$$EN = \sum_{i=1}^{n} EX_i = \sum_{i=1}^{n} 1 \cdot P(X_i = 1) = \sum_{i=1}^{n} \frac{(n-1)!}{n!} = \frac{1}{n} n = 1$$

2 Independence

Definition 2.1. *Let (Ω, Σ, P) be a probability space. Events $A, B \in \Sigma$ are independent if*

$$P(A \cap B) = P(A) \cdot P(B)$$

two σ-algebras $\Sigma_1, \Sigma_2 \subset \Sigma$ are independent if

$$P(A \cap B) = P(A) \cdot P(B) \quad \text{for any } A \in \Sigma_1 \ B \in \Sigma_2$$

finitely many σ-algebras $\Sigma_1, \Sigma_2, \dots, \Sigma_n \subset \Sigma$ are independent if

$$P\left(\cap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i) \text{ whenever } A_i \in \Sigma_i \ 1 \le i \le n$$

a sequence of σ-algebras $\Sigma_1, \Sigma_2, \dots \subset \Sigma$ are independent for any n
let X, Y be random variables, they are independent if

$$\sigma(X) \coprod \sigma(Y)$$

Recall:

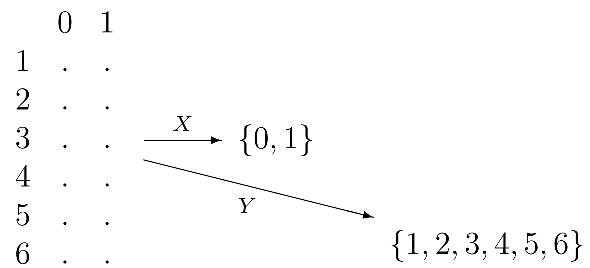
$$\sigma(X) = \{X^{-1}(B),\ B \in \mathcal{B}\} = \{X \in B,\ B \in \mathcal{B}\}$$

"information which we can get from X".

www.mymathscloud.com
My Maths Cloud

Example 2.1.
$X \sim$ coin toss: Bernulli
$Y \sim$ roll a die: $P(X = 1) = \cdots = P(X = 6) = \frac{1}{2}$

$$\begin{aligned} &\Omega = \{0, 1\} \times \{1, 2, 3, 4, 5, 6\} \\ &\Sigma = 2^{\Omega} \\ &P(\textit{each point}) = \frac{1}{12} \end{aligned}$$



$$\left.\begin{aligned} X(\omega_1, \omega_2) = \omega_1 \\ Y(\omega_1, \omega_2) = \omega_2 \end{aligned}\right\} \quad \textit{are they independent?}$$

$$\sigma(X) = \{\phi, \Sigma, \underbrace{\{(1, i),\ i = 1 \ldots 6\}}_{\{X=1\}}, \underbrace{\{(0, i),\ i = 1 \ldots 6\}}_{\{X=0\}}\}$$

$$\sigma(Y) = \{\phi, \Sigma, \{Y = 1\}, \ldots \{Y = 6\}, \{Y = 3 \textit{ or } 4 \textit{ or } 6\} \textit{ etc} \ldots\}$$

$$\underbrace{P\{X = 1,\ Y = 6\} = \frac{1}{12}, \quad P\{X = 1\} = \frac{1}{2}\ P\{Y = 6\} = \frac{1}{6}}_{\frac{1}{12} = \frac{1}{2} \cdot \frac{1}{6}}$$

similarly check this for all pairs of sets!

Definition 2.2. *Let $\mathcal{I}$ be a collection of sets, it is called a π-system if $\forall A, B \in \mathcal{I}$*

$$A \cap B \in \mathcal{I}$$

Example 2.2.

$$\left.\begin{aligned} &\left.\begin{aligned} &\{(-\infty, t),\ t \in \mathbb{R}\} \\ &\{(-\infty, t],\ t \in \mathbb{R}\} \\ &\{\{(a, b) : a < b\}, \phi\} \end{aligned}\right\} \textit{ generate } \mathcal{B} \\ &\{\{1\}, \{2\}, \{3\}, \phi\} \end{aligned}\right\} \quad \pi\textit{-system on } \mathcal{B}$$

Example 2.3.

$$\left.\begin{aligned} \{\{X < t\},\ t \in \mathbb{R}\} \\ \{\{X \leq t\},\ t \in \mathbb{R}\} \end{aligned}\right\} \quad \pi\textit{-system generating } \sigma(X)$$

because

$$\{X < t\} \cap \{X < s\} = \{X < \min(t, s)\} \textit{ etc}$$

Example 2.4. *If X, $\underbrace{\text{takes finitely or countably many values}}_{\text{discrete}}$ $a_1, a_2, \dots$*

$$\{\{X = a_1\}, \{X = a_2\}, \dots\ \phi\} - \pi\text{-system generating } \sigma(X)$$

Theorem 2.1. *Let (Ω, Σ) be a set with σ-algebra and mathcal$I$ be a π-system generating Σ. Let μ_1, μ_2 be measures such that*

1. $\mu_1(\Omega) = \mu_2(\Omega) < \infty$
2. $\mu_1(I) = \mu_2(I)$ *for any* $I \in \mathcal{I}$

then

$$\mu_1(A) = \mu_2(A) \quad \forall A \in \Sigma$$

Proof. Stated without proof □

Theorem 2.2. *Let X, Y be rv's and $\mathcal{I}, \mathcal{J}$ be π-systems generating $\sigma(X)$ and $\sigma(Y)$.*

$$\mathcal{I} \text{ and } \mathcal{J} \text{ are independent} \Rightarrow X \text{ and } Y \text{ are independent}$$

Proof. Stated without proof □

Corollary 2.1.

(a) To check the independence of X and Y it suffices to check

$$P(X < t, Y < s) = P(X < t)P(Y < s)$$

(b) if X and Y are discrete, taking values $a_1, a_2, \dots$ and $b_1, b_2, \dots$ then

$$P(X = a_i, Y = b_i) = P(X = a_i)P(Y = b_i) \quad \forall i, j$$

Proof. fix $I \in \mathcal{I}$

$$\mu_1(B) = P(I \cap B)$$
$$\mu_2(B) = P(I)P(B)$$

on $\sigma(Y)$, μ_1, μ_2 are measures such that $\mu_1(\Omega) = \mu_2(\Omega) = P(I) < \infty$ and they agree on $\mathcal{J}$

$$\overset{Thm}{\underset{2.1}{\Rightarrow}} \mu_1(B) = \mu_2(B) \quad \forall B \in \sigma(Y)$$
$$\Rightarrow P(X = a_i, Y = b_i) = P(X = a_i)P(Y = b_i) \quad \forall I \in \mathcal{I},\ B \in \sigma(Y)$$

fix $B \in \sigma(Y)$. Define

$$v_1(A) = P(A \cap B)$$
$$v_2(A) = P(A)P(B)$$

on $\sigma(X)$, v_1, v_2 are measures such that $v_1(\Omega) = v_2(\Omega) = P(B) < \infty$ and they agree on $\mathcal{I}$

$$\overset{Thm}{\underset{2.1}{\Rightarrow}} \text{they agree on } \sigma(X)$$
$$\Rightarrow P(A \cap B) = P(A)P(B) \quad \text{whenever } A \in \sigma(X),\ B \in \sigma(Y)$$

□

www.mymathscloud.com My Maths Cloud

www.mymathscloud.com
My Maths Cloud

Example 2.5. $X_1, X_2, \ldots$

- *independent*
- *each X_i has a prescribed distribution function F_i*

do they always exist?

Trick Model*:*
It suffices to construct $U_1, U_2, \ldots,$ which are independent and have uniform distribution, because $X_1 = G_1(U_1), X_2 = G_2(U_2), \ldots$ generalised as in the Skorokhod representation, 1.1.

$$G = F^{-1} \qquad G(\omega) = \inf\{t : F(t) > \omega\}$$

- *X_i has distribution F_i*
- *they are independent since $U_1, U_2, \ldots$ are independent*

How do we construct $U_1, U_2, \ldots$

$$([0,1], \mathcal{B}, leb) \quad U(\omega) = \omega \ (Uniform!) = .\omega_1\omega_2\ldots$$

$\omega_1, \omega_2, \ldots$ *are*

(a) bernulli

(b) independent

$$P(\omega_2 = 1) = \frac{1}{2} \quad \underbrace{P(\omega_1 = 0, \omega_2 = 1)}_{\frac{1}{4}} = \underbrace{P(\omega_1 = 0)}_{\frac{1}{2}}\underbrace{P(\omega_2 = 1)}_{\frac{1}{2}}$$

$\omega = .\omega_1\omega_2\ldots$

$$\begin{aligned} U_1(\omega) &= .\omega_1\omega_2\omega_{25}\omega_{100} \\ U_2(\omega) &= .\omega_2\omega_3\omega_{200} \\ U_3(\omega) &= .\omega_4\omega_{500}\omega_{1000} \end{aligned}$$

Uniform + Independent

2.1 Finite and infinite occurance of Events

Example 2.6.

$$\begin{array}{cccccc} X_1, & X_2 & X_3 & X_4 & \ldots & Bernulli \\ 1, & 0, & 1, & 1, & \ldots & \end{array}$$

"there will be infinitely many 1's in the sequence with probability 1."

$$E_n = \{X_n = 1\}$$

$$E = \{\text{ infinitely many 1's in the sequence}\} = \{\forall N \in \mathbb{N} \exists n \geq N \ X_n = 1\}$$

$$= \bigcap_{N=1}^{\infty} \bigcup_{n \geq N} \{X_n = 1\}$$

www.mymathscloud.com

Notation. *Let (E_n) be a sequence of events*

$$\{E_n \text{ i.o.}\} = \bigcap_{N=1}^{\infty} \bigcup_{n\geq N} E_n \quad \textit{infinitely many of } E_n \textit{ occur}$$

$$\{E_n \text{ i.o.}\}^c = \bigcup_{N=1}^{\infty} \bigcap_{n\geq N} E_n^c \quad \textit{finitely many events occur}$$

Note. *i.o. - infinitely often*

Theorem 2.3 (Strong Law of Large Numbers)**.** *Let (X_n) be independent identically distributed random variables such that $E|X_1| < \infty$, then*

$$\frac{X_1 + \cdots + X_n}{n} \underset{a.s.}{\to} EX_1 = EX_n \quad \forall n$$

Outline of proof:
$S_n = X_1 + \cdots + X_n$ denote $\mu = EX_1$ we need to prove:

$$P\left(\omega \in \Omega : \frac{S_n}{n} \to \mu\right) = 1$$

$$\begin{aligned} P\left(\omega \in \Omega : \frac{S_n}{n} \to \mu\right) &= P\left(\forall k > 0 \ \exists N \in \mathbb{N} \ \forall n > N \quad \left|\frac{S_n}{n} - \mu\right| < \frac{1}{k}\right) \\ &= P\left(\bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n\geq N} \left\{\left|\frac{S_n}{n} - \mu\right| < \frac{1}{k}\right\}\right) \end{aligned}$$

We need to show $\forall k \in \mathbb{N}$

$$P\left(\bigcup_{N=1}^{\infty} \bigcap_{n\geq N} \left\{\left|\frac{S_n}{n} - \mu\right| < \frac{1}{k}\right\}\right) = 1$$

that is equivalent to

$$P\left(\bigcup_{N=1}^{\infty} \bigcap_{n\geq N} \left\{\left|\frac{S_n}{n} - \mu\right| \geq \frac{1}{k}\right\}^c\right) = 1$$

ie

$$P\left(\left\{\left|\frac{S_n}{n} - \mu\right| \geq \frac{1}{k} \ \text{ i.o.}\right\}^c\right) = 1$$

$$P\left(\left\{\left|\frac{S_n}{n} - \mu\right| \geq \frac{1}{k} \ \text{ i.o.}\right\}\right) = 0$$

Proven later.

2.2 The Borel-Cantelli Lemmas

Lemma 2.4 (Borel-Cantelli Lemma 1 - BC1)**.** *Let (E_n) be events such that $\sum\limits_{n=1}^{\infty} P(E_n) < \infty$, then*

$$P(E_n\ i.o.) = 0$$

Proof.

$$P(\bigcup_{n\geq N} E_n) \leq \sum_{n\geq N} P(E_n) \quad \forall n \in \mathbb{N}$$

$$P(E_n\ i.o) = P(\bigcap_{N=1}^{\infty} \bigcup_{n\geq N} E_n) = \lim_{n\to\infty} P(\bigcup_{n\geq N} E_n) \leq \lim_{n\to\infty} \sum_{n\geq N} P(E_n) = 0$$

since $\sum\limits_{n=1}^{\infty} P(E_n) < \infty$. □

Lemma 2.5 (Borel-Cantelli Lemma 2 - BC2)**.** *Let (E_n) be events independent events such that $\sum\limits_{n=1}^{\infty} P(E_n) = \infty$, then*

$$P(E_n\ i.o.) = 1$$

Proof.

$$P(E_n\ i.o.) = 1 \Rightarrow P(\{E_n\ i.o.\}^c) = P(\bigcup_{N=1}^{\infty} \bigcap_{n\geq N} E_n^c) = 0$$

i.e

$$P(\bigcap_{n\geq N} E_n^c) = 0 \quad \forall n \in \mathbb{N}$$

$$P(\bigcap_{n\geq N} E_n^c) = \lim_{k\to\infty} P(\bigcap_{n=N}^{k} E_n^c) = \lim_{k\to\infty} \prod_{n=N}^{k} \underbrace{P(E_n^c)}_{1-P(E_n)}$$

$$\leq \lim_{k\to\infty} \prod_{n=N}^{k} e^{-P(E_n)} = \lim_{k\to\infty} e^{-\sum_{n=N}^{k} P(E_n)} = 0 \quad \forall N$$

□

Remark. *$E_n = E$ with some E such that $0 < P(E) < 1$, $\{E_n\ i.o.\} = E$ so $P(E_n\ i.o.) = p$ not $0, 1$. BC2, 2.5, doesn't hold for dependent events.*

To summarize if (E_n) are independent then

$$P(E_n\ i.o.) = \begin{cases} 1 & if\ \sum\limits_{i=1}^{\infty} P(E_n) = \infty \\ 0 & if\ \sum\limits_{i=1}^{\infty} P(E_n) < \infty \end{cases}$$

www.mymathscloud.com
My Maths Cloud

www.mymathscloud.com
My Maths Cloud

Applications. *X_n iid assume all X_n have exponential distribution. we wish to show:*

$$\lim_{n\to\infty} \sup \frac{X_n}{log(n)} = 1 \quad a.s.$$

$\limsup a_n = a$ *means a is the largest value such that there is a subsequence a_{n_k} converging to a*

Example 2.7.

1. $1,\ -1,\ 1,\ -1, \ldots$
$$\limsup = 1$$
2. $1,\ 0,\ 3,\ 0,\ 5, \ldots$
$$\limsup = \infty$$

So if I want to prove $\limsup a_n = a$*:*

1. $\forall b > a, \quad a_n > b$ *occurs for finitely many n*
2. $\forall b < a, \quad a_n > b$ *occurs for infinitely many n*

So for

$$\lim_{n\to\infty} \sup \frac{X_n}{log(n)} = 1$$

1. $b > 1$
$$P\{\frac{X_n}{log(n)} > b \;\; i.o\} = 0$$
$$\sum_{n=1}^{\infty} P(\frac{X_n}{log(n)} > b) = \sum_{n=1}^{\infty} P(X_n > b\,log(n)) = \sum_{n=1}^{\infty} e^{-b\,log(n)} = \sum_{n=1}^{\infty} \frac{1}{n^b} < \infty$$
2. $b < 1$
$$P\{\frac{X_n}{log(n)} > b \;\; i.o\} = 1$$
$$\sum_{n=1}^{\infty} P(\frac{X_n}{log(n)} > b) = \sum_{n=1}^{\infty} \frac{1}{n^b} = \infty$$

Example 2.8. *For the following two iid sequences (X_n):*
If (X_n) is Exponentially distributed such that $F(x) = 1 - e^{-x},\ x > 0$

$$\lim_{n\to\infty} \sup \frac{X_n}{log(n)} = 1 \quad a.s.$$

If (X_n) is normally distributed $N(0,1)$

$$\lim_{n\to\infty} \sup \frac{X_n}{\sqrt{2log(n)}} = 1 \quad a.s.$$

www.mymathscloud.com
My Maths Cloud

(this is an exersize in homework.)
Are the any distributions such that (X_n) *grows along a straight line?*

$$\lim_{n\to\infty}\sup \frac{|X_n|}{n} = \alpha?$$

answer

$$\lim_{n\to\infty}\sup \frac{|X_n|}{n} = \begin{cases} o & if\ E|X_1| < \infty \leftarrow \textit{follows from SLLN} \\ \infty & if\ E|X_1| = \infty \end{cases}$$

Proof.

$$\frac{X_n}{n} = \underset{\substack{\downarrow \text{SLLN} \\ EX_1}}{\frac{X_1 + \cdots + X_n}{n}} - \underset{\substack{\downarrow \text{SLLN} \\ EX_1}}{\frac{X_1 + \cdots + X_{n-1}}{n-1}} \cdot \underset{\substack{\downarrow \\ 1}}{\frac{n-1}{n}} \to 0$$

We want to show the if $E|X_1| = \infty$

$$P\left(\lim_{n\to\infty}\sup \frac{|X_n|}{n} = \infty\right) = 1$$

i.e.

$$P\left(\forall m\ \frac{X_n}{n} > m\ i.o.\right) = P\left(\bigcap_{m\in\mathbb{N}} \frac{X_n}{n} > m\ i.o.\right) = 1$$

i.e

$$P\left(\frac{X_n}{n} > m\ i.o.\right) = 1 \quad \forall m$$

Use BC2, 2.5

$$\sum_{n=1}^{\infty} P\left(\frac{|X_1|}{n} > m\right) = \sum_{n=1}^{\infty} P\left(\frac{|X_1|}{n} > m\right) = \sum_{n=1}^{\infty} E\mathbb{1}_{\{\frac{|X_1|}{n} > m\}}$$

Idea $m = 1$

$$P(|X_1| > 1)$$
$$P(|X_1| > 2)$$
$$\vdots \ \vdots$$

$$\sum_{n=1}^{\infty} E\mathbb{1}_{\{\frac{|X_1|}{n} > m\}} \overset{MCT}{=} E\sum_{n=1}^{\infty} \mathbb{1}_{\{n < \frac{|X_1|}{m}\}} \geq E\left(\frac{|X_1|}{m} - 1\right) = \infty$$

□

Theorem 2.6 (Expectation and Variance for independent variables)**.** *Let X and Y be independent random variables*

(a) if $E|X| < \infty, E|Y| < \infty$ *then* $E|XY| < \infty$

$$E(XY) = E(X) \cdot E(Y)$$

www.mymathscloud.com
My Maths Cloud

(b) if $E(X^2), E(Y^2) < \infty$ *then*

$$Var(X+Y) = Var(X) + Var(Y)$$

Proof of (a).
(i) Lets check this for simple r.v's

$$X = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}, \quad Y = \sum_{j=1}^{m} b_j \mathbb{1}_{B_j} \quad \text{mutually different}$$

$$\begin{aligned} E(XY) &= E((\sum_{i=1}^{n} a_i \mathbb{1}_{A_i})(\sum_{j=1}^{m} b_j \mathbb{1}_{B_j})) = E(\sum_{j=1}^{n}\sum_{j=1}^{m} a_i b_j \mathbb{1}_{A_i \cap B_j}) \\ &= \sum_{j=1}^{n}\sum_{j=1}^{m} a_i b_j \underbrace{P(A_i \cap B_j)}_{=P(A_i)P(B_j)} = (\sum_{i=1}^{n} a_i P(A_i))(\sum_{i=1}^{m} b_j P(B_j)) \\ &= E(X)E(Y) \end{aligned}$$

(ii) $X \geq 0,\ Y \geq 0\ \leftarrow$ by approximation $X,\ Y$ by simple r.v's and using MCT.
(iii) X, Y arbitrary, take

$$X = X^+ - X^- \quad Y = Y^+ - Y^-$$

and use linearity. □

Proof of (b).

$$\begin{aligned} Var(X+Y) &= E(X+Y)^2 - (E(X+Y))^2 \\ &= E(X^2 + 2XY + Y^2) - ((EX)^2 + 2EXEY + (EY)^2) \\ &= EX^2 - (EX)^2 + EY^2 - (EY)^2 \\ &= VarX + VarY \end{aligned}$$

□

Example 2.9. $n \in \mathbb{N},\ X_1, \dots, X_n$ *− independent.*

$$P(X_i = 1) = p \quad P(X_i = 0) = 1 - p$$

$$Y = X_1 + \cdots + X_n \quad \textit{number of heads over } n\textit{-tosses}$$

Y has Binomial distribution

$$P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$EY = \sum_{k=1}^{n} kP(Y=k) \quad VarY = \sum_{k=1}^{n} k^2 P(Y=k)$$

$$\vdots \qquad \vdots$$

$$EY = E(X_1 + \cdots + X_n) = EX_1 + \cdots + EX_n = np$$

$$VarY = Var(X_1 + \cdots + X_n) = VarX_1 + \cdots + VarX_n = np(1-p)$$

www.mymathscloud.com
My Maths Cloud

2.3 Bernsteins inequality

Theorem 2.7 (Bernsteins inequality). *Let $X_1, \ldots, X_n$ be independent and such that*

$$P(X_i = 1) = \frac{1}{2}, \quad P(X_i = -1) = \frac{1}{2} \quad \forall 1 \leq i \leq n$$

and $a_1, \ldots, a_n \in \mathbb{R}$ then

$$P\left(\left|\sum_{i=1}^n a_i X_i\right| \geq t\right) \leq 2\exp\left(\frac{-t^2}{2\sum\limits_{i=1}^n a_i^2}\right) \quad \forall t > 0$$

Proof. Denote $c = \sum\limits_{i=1}^n a_i^2$ and let $\lambda > 0$

$$E\left(\exp\left(\lambda\sum_{i=1}^n a_i X_i\right)\right) = E\left(\prod_{i=1}^n e^{\lambda a_i X_i}\right) = \prod_{i=1}^n E\left(e^{\lambda a_i X_i}\right) = \prod_{i=1}^n \frac{e^{\lambda a_i} + e^{-\lambda a_i}}{2}$$
$$= \prod_{i=1}^n \cosh \lambda a_i \leq \prod_{i=1}^n \exp\left(\frac{\lambda^2 a_i^2}{2}\right) = \exp\left(\frac{\lambda^2 c}{2}\right)$$

now

$$P\left(\lambda\sum_{i=1}^n a_i X_i > \lambda t\right) = P\left(\exp\left(\lambda\sum_{i=1}^n a_i X_i\right) > \exp(\lambda t)\right)$$
$$\overset{Chebyshev}{\leq} E\left(\exp\left(\lambda\sum_{i=1}^n a_i X_i\right) \cdot \exp(-\lambda t)\right)$$
$$\leq \exp\left(\frac{\lambda^2 c}{2} - \lambda t\right)$$

Take λ in such a way that it minimises $\frac{\lambda^2 c}{2} - \lambda t$

$$P\left(\sum_{i=1}^n a_i X_i > t\right) \leq \exp\left(\frac{t^2}{c^2}c\frac{1}{2} - \frac{t^2}{c}\right) = \exp\frac{-t^2}{2c}$$

$$P\left(\sum_{i=1}^n \underbrace{a_i}_{b_i = -a_i} X_i < -t\right) = P\left(\sum_{i=1}^n b_i X_i > t\right) \leq \exp\frac{-t^2}{2c}$$

$$P\left(\left|\sum_{i=1}^n a_i X_i\right| > t\right) = P\left(\sum_{i=1}^n a_i X_i > t \cup \sum_{i=1}^n a_i X_i < -t\right)$$
$$\leq P\left(\sum_{i=1}^n a_i X_i > t\right) + P\left(\sum_{i=1}^n a_i X_i < -t\right)$$
$$\leq 2\exp\frac{-t^2}{2c}$$

$\square$

www.mymathscloud.com

Theorem 2.8 (SLLN for r.v's taking values ± 1)**.** *Let $X_1, \dots, X_n$ be a sequence of iid random variables such that*

$$P(X_i = 1) = \frac{1}{2}, \quad P(X_i = -1) = \frac{1}{2} \quad \forall i$$

then

$$\frac{X_1 + \cdots + X_n}{n} \to 0 \; (= EX_1) \; a.s$$

Proof. we want to prove

$$P\left(\forall k \, \exists N \, \forall n \geq N \; \left|\frac{X_1 + \cdots + X_n}{n}\right| < \frac{1}{k}\right) = 1$$

i.e

$$P\left(\bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n \geq N} \left\{\left|\frac{X_1 + \cdots + X_n}{n}\right| < \frac{1}{k}\right\}\right) = 1$$

i.e

$$P\left(\bigcup_{N=1}^{\infty} \bigcap_{n \geq N} \left\{\left|\frac{X_1 + \cdots + X_n}{n}\right| < \frac{1}{k}\right\}\right) = 1 \quad \forall k$$

$$P\left(\bigcap_{N=1}^{\infty} \bigcup_{n \geq N} \left\{\left|\frac{X_1 + \cdots + X_n}{n}\right| \geq \frac{1}{k}\right\}\right) = 0$$

this means

$$P\left(\left|\frac{X_1 + \cdots + X_n}{n}\right| \geq \frac{1}{k} \; i.o.\right) = 0 \quad \forall k$$

Use BC1:

$$\sum_{n=1}^{\infty} \left(\left|\frac{X_1 + \cdots + X_n}{n}\right| \geq \frac{1}{k}\right) \overset{(*)}{\leq} \sum_{n=1}^{\infty} 2 \exp \frac{-n}{2k^2} < \infty$$

Note (*)**.** *this is an application of bernsteins inequality with $a_1 = \cdots = a_n = 1$ and $t = \frac{n}{k}$ since*

$$\sum_{n=1}^{\infty} e^{-n} = \sum_{n=1}^{\infty} (\underbrace{e^{-1}}_{<1})^n$$

□

2.4 Joint Laws

Recall.

$$X \rightsquigarrow \mu_X \quad on \; (\mathbb{R}, \mathcal{B})$$

Let $X,\; Y$ be two r.v's

Definition 2.3 (Joint Law)**.** *The Joint Law of X and Y is a probability measure*

$$\mu_{X,Y} \quad on \; (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$$

Defined by

$$\mu_{X,Y} = P((x, y) \in \mathcal{B})$$

www.mymathscloud.com My Maths Cloud

Definition 2.4 (The Joint distribution function)**.**

$$F_{X,Y}(x,y) = P(X \leq x,\ Y \leq y) = \mu_{X,Y}((-\infty, x] \times (-\infty, y])$$

Theorem 2.9. *Let X and Y be independent*

(a)

$$F_{X,Y}(x,y) = F_X(x)F_Y(y)$$

(b) if X and Y have densities f and g then $\mu_{X,Y}(x,y)$ has density

$$f(x)g(y)$$

with respect to lebesgue measure on $\mathbb{R}^2$

(c) if X and Y have densities f and g, then $X+Y$ has density $(f*g)(t)$ where

$$(f*g)(t) = \int_{-\infty}^{\infty} f(x)g(t-x)dx$$

which is called the convolution of f and g.

Proof of (a).

$$F_{X,Y}(x,y) = P(X \leq x,\ Y \leq y) = P(X \leq x)P(Y \leq y) = F_X(x)F_Y(y)$$

□

Proof of (b). we want to show

$$\mu_{X,Y}(B) = \int_B f(x)g(y)dxdy \quad \forall B \subset \mathcal{B}(\mathbb{R}^2)$$

it suffices to look at $B = (-\infty, x] \times (-\infty, y],\ \forall x, y$ since they form a π-system generating $\mathcal{B}(\mathbb{R}^2)$.

$$\mu_{X,Y}((-\infty, x] \times (-\infty, y]) = F_{X,Y}(x,y) = F_X(x)F_Y(y) = \int\limits_{-\infty}^{x} f(u)du \int\limits_{-\infty}^{y} g(v)dv$$

$$= \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} f(u)g(v)dudv$$

□

www.mymathscloud.com
My Maths Cloud

Proof of (c).

$$F_{X+Y}(t) = P(X+Y \leq t) \overset{(b)}{=} \iint\limits_{U+V\leq t} f(u)g(v)dudv = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{t-u} f(u)g(v)dudv$$

$$= \int\limits_{-\infty}^{\infty} f(u) \underbrace{\left(\int\limits_{-\infty}^{t-u} g(v)dv\right)}_{v=z-u} du = \int\limits_{-\infty}^{\infty} f(u) \int\limits_{-\infty}^{t} g(z-u)dzdu$$

$$= \int\limits_{-\infty}^{t} \left(\int\limits_{-\infty}^{\infty} f(u)g(z-u)du\right) dz$$

$$F_{X+Y}(t) = \int\limits_{-\infty}^{\infty} f(u)g(t-u)du$$

$\square$

Example 2.10. *Density of $X \cdot Y$?*

$$F_{X\cdot Y}(t) = \int\limits_{uv\leq t} f(u)g(v)dudv$$

2.5 Tail events and Kolmogorov $0-1$ law

(X_n) - r.v's

$$\begin{array}{lcl} & & \text{tail events} \\ \{\lim X_n > 0\} & \leftarrow & \boxed{\text{doesn't depend on any finite no. r.v's}} \\ \{\sup X_n > 0\} & \leftarrow & \text{is not like that} \end{array}$$

Definition 2.5. *Let X_n be a sequence of r.v.'s*

$$\mathcal{T}_n = \sigma(X_{n+1} > X_{n+2} > \cdots) \quad n^{th} \text{ tail } \sigma\text{-algebra}$$

this is the information contained in $X_{n+1}, X_{n+2}, \ldots$

$$\mathcal{T} = \bigcap_{n=1}^{\infty} \mathcal{T}_n \quad \text{the tail } \sigma\text{-algebra}$$

each $A_i \in \mathcal{T}$ is called a Tail event.

Example 2.11.

www.mymathscloud.com

1.

$$\{X_n \to a\} \; - \; \textit{tail event}$$

since

$$\{X_n \to a\} = \{X_1, X_2, X_3, \cdots \to a\} = \{X_{m+1}, X_{m+2}, \cdots \to a\} \in \mathcal{T}_m$$
$$\Rightarrow \{X_n \to a\} \in \cap \mathcal{T}_m = \mathcal{T}$$

2.

$$\{\lim_{n\to\infty} X_n \; \textit{exists}\} \; - \; \textit{tail event}$$
$$\left\{\sum_{n=1}^{\infty} X_n < \infty\right\} \; - \; \textit{tail event}$$

3.

$$\{\sum_{n=1}^{\infty} X_n < 10\} \; - \; \textit{not a tail event}$$

4.

$$\left\{\frac{X_1 + \cdots + X_n}{n} \; \textit{converges}\right\} \; - \; \textit{tail event}$$

since

$$\left\{\frac{X_1 + \cdots + X_n}{n} \; \textit{converges}\right\} = \left\{\frac{X_1}{1}, \frac{X_1 + X_2}{2}, \frac{X_1 + X_2 + X_3}{3}, \ldots \; \textit{converges}\right\}$$

$$\frac{X_1 + \cdots + X_n}{n} = \underbrace{\frac{X_1 + \cdots + X_m}{n}}_{\to 0} + \frac{X_{m+1} + \cdots + X_n}{n}$$

$$\forall m \; \left\{\frac{X_{m+1} + \cdots + X_n}{n} \; \textit{converges}\right\} \in \mathcal{T}_n \Rightarrow \; \textit{Tail event}$$

5.

$$\{\sum_{n=1}^{\infty} X_n > 0\} \; - \; \textit{not a tail event}$$

6.

$$A = \{\sup_n X_n > 0\} \quad - \quad \textit{not a tail event}$$

Consider

$$X_i = \left\{1, -1 \; \textit{each with} \; P = \frac{1}{2}\right\} \quad X_i = 0, \; \forall i \geq 2$$

we want to show

$$A \notin \mathcal{T} = \bigcap_{n=1}^{\infty} \mathcal{T}_n, \quad \mathcal{T}_n = \sigma(X_{n+1}, X_{n+2}, dots)$$

www.mymathscloud.com
My Maths Cloud

$$
\begin{array}{rcl}
\sigma(X_2) & = & \{\phi, \Omega\} \\
\sigma(X_3) & = & \{\phi, \Omega\} \\
\vdots & & \vdots \\
\sigma(X_n) & = & \{\phi, \Omega\}
\end{array}
$$

$$\Rightarrow \mathcal{T}_n = \{\phi, \Omega\} \Rightarrow \mathcal{T} = \{\phi, \Omega\}$$

$$P(A) = \frac{1}{2} \Rightarrow A \neq \phi, \; A \neq \Omega$$

Theorem 2.10 (Kolmogorov $0-1$ law)**.** *If (X_n) is a sequence of independent random variables, then each tail event has probability 0 or 1.*

Proof.

$$\sigma_n = \sigma(X_1, \dots, X_n)$$

$$\mathcal{T}_n = \sigma(X_{n+1}, X_{n+2}, \dots)$$

since (X_n) are independent

$$\sigma_n \perp\!\!\!\perp \mathcal{T}_n, \quad \forall n$$

$$\mathcal{T} = \bigcap_{i=1}^{\infty} \mathcal{T}_n \Rightarrow \mathcal{T} \subset \mathcal{T}_n \Rightarrow \sigma_n \perp\!\!\!\perp \mathcal{T}, \quad \forall n$$

Denote

$$\sigma_\infty(X_1, X_2 \dots) \Rightarrow \textit{ each } \mathcal{T}_n \subset \sigma_\infty \Rightarrow \mathcal{T} \subset \sigma_\infty$$

On the other hand

$$\sigma_\infty \perp\!\!\!\perp \mathcal{T}$$

since σ_∞ is generated by the π-system $\bigcup_{n \in \mathbb{N}} \sigma_n$ which is independent of $\mathcal{T}$.

$$\Rightarrow \mathcal{T} \perp\!\!\!\perp \mathcal{T}$$

$$\forall A \in \mathcal{T} \quad P(A) = P(\underset{\in \mathcal{T}}{A} \cap \underset{\in \mathcal{T}}{A}) = P(A)P(A) = P(A)^2$$

$$\Rightarrow P(A) = 0 \quad or \quad P(A) = 1$$

□

Example 2.12.

1. *for all tail events in example 2.11 above*

$$P(\dots) = 0 \textit{ or } 1 \quad \textit{(if } (X_n) \textit{ are independent)}$$

www.mymathscloud.com
My Maths Cloud

2. $\frac{|X_n|}{n}$ *if* $E|X_1| < \infty$ *then by SLLN*
$$\frac{|X_n|}{n} \to 0$$
if $E|X_1| = \infty$
$$\frac{|X_n|}{n} \text{ diverges.}$$
$$P\left(\left\{\frac{|X_n|}{n} \quad \text{converges.}\right\}\right) = \begin{cases} 1 & \text{if } E|X_1| < \infty \\ 0 & \text{if } E|X_1| = \infty \end{cases}$$

3. *Percolation. Have a lattice and flip a coin with probability p, if heads keep edge, if tails remove edge.*
$$\{\text{there is no infinite cluster}\} \; - \; \text{tail event}$$
X_n = *edge n edges away from start point. Infinite cluster exists with either* $P = 1$ *or* 0 *depending on value of p.*

3 Weak convergence

X, Y independent bernulli random variables
$$X \neq Y \quad a.s.$$
but
$$X \underset{Law}{\overset{in}{=}} Y \quad \text{since} \quad \mu_X = \mu_Y$$
$$\frac{X_1 + \cdots X_n}{n} \overset{SLLN}{\to} \mu \quad a.s.$$
$$\frac{X_1 + \cdots X_n - n\mu}{n\sigma} \underset{Law}{\overset{in}{=}} N(1,0) \quad (CLT)$$
where here σ means varience of (X_i)

Definition 3.1. *we say that* $X_n \to X$ *in law, in distribution or weakly if*
$$F_{X_n}(t) \to F_X(t) \quad \forall t \in \mathbb{R}$$
where F_X *is continuous.*
$$\Leftrightarrow \mu_{X_n} \to \mu_X \quad \text{weakly}$$

Notation. *to say that* $X_n \to X$ *in law, in distribution or weakly:*
$$\underbrace{\overset{d}{\to}}_{\text{we will use this}} \qquad \underbrace{\overset{w}{\to}, \Rightarrow}_{\text{some people use these}}$$

www.mymathscloud.com
My Maths Cloud

Remark. *why do we exclude discontinuity points?*

$$X_n = \frac{1}{n} \text{ with } P = 1 \qquad X = 0 \text{ with } P = 1$$

Let μ_n be the law of X_n, so μ_n is 1 at $\frac{1}{n}$ and let μ be the law of X. Let h be a continuous function then

$$\mu_n(h) = h(\tfrac{1}{n}) \to h(0) = \mu(h)$$

however

$$F_n(0) = 0 \not\to F(0) = 1$$

not a distribution function since not right-continuous

Theorem 3.1 (Relation between a.s. and weak convergence)**.**

1. *if $X_n \to X$ a.s. then $X_n \xrightarrow{d} X$*
2. *if $\mu_X \to \mu$ weakly then there are random variables (X_n) and X such that*

$$\begin{aligned} &X_n \text{ has law } \mu_n \quad \forall n \\ &X \text{ has law } \mu \end{aligned}$$

and $X_n \to X$ a.s.

Theorem 3.2 (Useful definition of weak convergence)**.**

$$\mu_n \to \mu \text{ weakly } \Rightarrow \int_{\mathbb{R}} h d\mu_n \to \int_{\mathbb{R}} h d\mu$$

$$\forall h : \mathbb{R} \to \mathbb{R} \qquad \text{continuous and bounded}$$

Idea.

$$\begin{aligned} X_n \xrightarrow{d} X \;\Rightarrow\; EX_n &\to EX \qquad \int X d\mu_n \to \int X d\mu \\ EX_n^2 &\to EX^2 \qquad \int X^2 d\mu_n \to \int X^2 d\mu \end{aligned}$$

$\Leftarrow$ *mabie we can get this if we check for a sufficently large class of test functions.*

$$\boxed{\int_{\mathbb{R}} h d\mu_n \to \int_{\mathbb{R}} h d\mu} \quad \to \quad Eh(X_n) \to Eh(X)$$

Proof of theorems 3.1 and 3.2. Plan: (2), $\Rightarrow$, $\Leftarrow$, (1)

www.mymathscloud.com
My Maths Cloud

Proof of (2). we are given $\mu_n \to \mu$ weakly i.e.

$$F_n(t) \to F(t) \quad \forall t, \text{ where } F \text{ is continuous}$$

we use Skorokhod representation, 1.1, to construct all X_n and X.

$$\begin{aligned}([0,1], \mathcal{B}, Leb), \quad X(\omega) &= \inf\{u : F(u) > \omega\} \leftarrow \text{ has law } \mu \\ X_n(\omega) &= \inf\{u : F_n(u) > \omega\} \leftarrow \text{ has law } \mu_n\end{aligned}$$

$$B = \{\omega \in [0,1] : \exists x, y \in \mathbb{R} \; s.t. \; F(x) = F(y)\}$$

we need to

(a) prove that B is at most countable and $Leb(B) = 0$

(b) prove that $X_n(\omega) \to X(\omega) \quad \forall \omega \in [0,1] \backslash B$

proof of (a): each $\omega \in B$ generate an interval (x, y), all these intevals don't intersect. Each interval contains a rational number.

$$\begin{aligned}&\Rightarrow \text{ at most countably many intervals}\\ &\Rightarrow \text{ at most countably many } \omega\end{aligned}$$

proof of (b): lets prove that the set of discontinuity points of F is at most countable, (same arguement as for (a) but considering intervals fromed on the y axis). Now let $\omega \in [0,1] \backslash B$, $\epsilon > 0$. Choose $0 < \delta < \epsilon$ so that $X(\omega) \pm \delta$ are continuity points of F.

$$F(X(\omega) - \delta) < \omega < F(X(\omega) + \delta)$$

$$\left.\begin{aligned}F_n(X(\omega) - \delta) \to F(X(\omega) - \delta)\\ F_n(X(\omega) + \delta) \to F(X(\omega) + \delta)\end{aligned}\right\} \text{Since } X(\omega) \pm \delta \text{ are are continuity points of } F \text{ and weak convergence}$$

$$\Rightarrow \exists N \; \forall n \geq N \quad F_n(X(\omega) - \delta) < \omega < F_n(X(\omega) + \delta)$$

$$\Rightarrow X(\omega) - \delta < X_n(\omega) < X(\omega) + \delta$$

□

Proof of ($\Rightarrow$). Assume $\mu_n \to \mu$ weakly. BY (2) we can pick X_n, X with laws μ_n, μ such that

$$X_n \to X \; a.s.$$

$$\int h d\mu_n = Eh(X_n) \underset{*}{\overset{DCT}{\to}} Eh(X) = \int h d\mu$$

($*$) since $X_n \to X$ a.s $\Rightarrow h(X_n) \to h(X)$ continutiy of h and since $h(X)$ bounded. □

www.mymathscloud.com

Proof of ($\Leftarrow$).

Idea.

$$F_X(X) = \mu_n((-\infty, X]) = \int \mathbb{1}_{(-\infty,X]} d\mu_n \xrightarrow{*} \int \mathbb{1}_{(-\infty,X]} d\mu = F(X)$$

($*$) *Problem: h is not continuous so we need to make it continuous.*

Let X be a continuity point of F. Let $\delta > 0$ and h be a continuous function such that:

$$h(x) = \begin{cases} 1 & x \in [-\infty, X-\delta) \\ \left.\begin{array}{l}\text{smooth decreasing function s(x)} \\ s.t: \ s(X-\delta)=1 \quad s(X)=0\end{array}\right\} & x \in [X-\delta, X] \\ 0 & x \in (X, \infty] \end{cases}$$

$$F_n(X) = \int \mathbb{1}_{(-\infty,X]} d\mu_n \geq \int h d\mu_n$$

$$\lim_{n\to\infty} \inf F_n(X) \geq \lim_{n\to\infty} \int h d\mu_n = \int h d\mu \geq \int \mathbb{1}_{(-\infty,X-\delta]} d\mu = F(X-\delta)$$

Let $\delta \to 0$

$$\lim_{n\to\infty} \inf F_n(X) \geq \lim_{\delta\downarrow 0} F(X-\delta) \stackrel{*}{=} F(X)$$

($*$) as X is a continuity point of F. Take some $\delta > 0$ and take h to be :

$$h(x) = \begin{cases} 1 & x \in [-\infty, X) \\ \left.\begin{array}{l}\text{smooth decreasing function s(x)} \\ s.t: \ s(X)=1 \quad s(X+\delta)=0\end{array}\right\} & x \in [X, X+\delta] \\ 0 & x \in (X+\delta, \infty] \end{cases}$$

$$F_n(X) = \int \mathbb{1}_{(-\infty,X]} d\mu_n \leq \int h d\mu_n$$

$$\lim_{n\to\infty} \sup F_n(X) \leq \lim_{n\to\infty} \int h d\mu = \int h d\mu \leq \int \mathbb{1}_{(-\infty,X]} d\mu = F(X+\delta)$$

Let $\delta \to 0$

$$\lim_{n\to\infty} \sup F_n(X) \leq \lim_{\delta\downarrow 0} F(X+\delta) \stackrel{*}{=} F(X)$$

($*$) since any distribution fInction is right continuous.

$$F(X) \leq \lim_{n\to\infty} \inf F_n(X) \leq \lim_{n\to\infty} \sup F_n(X) \leq F(X)$$

$$\Rightarrow \lim_{n\to\infty} F_n(X) = F(X)$$

$\square$

www.mymathscloud.com
My Maths Cloud

Proof of (1). Suppose $X_n \to X$ a.s.

$$
\begin{aligned}
&\Rightarrow h(X_n) \to h(X) \quad \text{for } h \text{ continuous and bounded} \\
&\overset{DCT}{\Rightarrow} Eh(X_n) \to Eh(X) \\
&\Rightarrow \int h d\mu_n \to \int h d\mu \underset{3.2}{\overset{thm}{\Rightarrow}} \mu_n \to \mu \text{ weakly}
\end{aligned}
$$

and so

$$X_n \xrightarrow{d} X$$

□

□

Definition 3.2. *A function is $h : \mathbb{R} \to \mathbb{R}$ is called a C^2 test function if it is twice differentiable, h'' is continuous and $h = 0$ outside some bounded interval.*

Definition 3.3. *μ_n is called tight if for any $\epsilon > 0$ $\exists M > 0$ such that*

$$\mu_n([-M, M]) \geq 1 - \epsilon \quad \forall n$$

$$\mu_n(\mathbb{R} \backslash [-M, M]) < \epsilon \quad \forall n$$

Theorem 3.3 (Usefull definition of weak convergence 2)**.**

$$\mu_n \to \mu \textit{ weakly } \Leftrightarrow \int h d\mu_n \to \int h d\mu$$

for any C^2 test function.

Proof. ($\Rightarrow$) obvious since any test function is continuous and bounded.
($\Leftarrow$)

(a) Right hand side $\Rightarrow$ $\{\mu_n\}$ is tight. Let $\epsilon > 0$ choose M_1 such that

$$\mu[-M_1, M_1] \geq 1 - \tfrac{\epsilon}{2}$$

and define h to be a smooth function such that

$$
h(x) = \begin{cases}
0 & x \in [-\infty, -M_1 - 1) \\
\left.\begin{array}{l} \text{smooth increasing function s(x)} \\ s.t: \ s(-M_1 - 1) = 0 \ s(-M_1) = 1 \end{array}\right\} & x \in [M_1, M_1 + 1] \\
1 & x \in (-M_1, M_1) \\
\left.\begin{array}{l} \text{smooth decreasing function s(x)} \\ s.t: \ s(M_1) = 1 \ s(M_1 + 1) = 0 \end{array}\right\} & x \in [M_1, M_1 + 1] \\
0 & x \in (M_1 + 1, \infty]
\end{cases}
$$

www.mymathscloud.com My Maths Cloud

we need

$$\begin{aligned}\mu_n([-M_1-1, M_1+1]) &= \int \mathbb{1}_{[-M_1-1,M_1+1]} d\mu_n \geq \int h d\mu_n \to \int h d\mu \\ &\geq \int \mathbb{1}_{[-M_1,M_1]} d\mu = \mu[-M_1, M_1] \geq 1 - \tfrac{\epsilon}{2}\end{aligned}$$

$$\mu_n([-M_1-1, M_1+1]) \geq 1 - \tfrac{\epsilon}{2} \quad \forall n \geq \ some\ N$$

Choose M_2 such that

$$\mu[-M_2, M_2] \geq 1 - \tfrac{\epsilon}{2} \quad \forall n \leq N$$

Choose $M = \max\{M_1 + 1, M_2\}$

$$\Rightarrow \mu_n[-M, M] \geq 1 - \tfrac{\epsilon}{2} \quad \forall n$$

(b) Let x be a continuity point of F. Let $\epsilon > 0$, $\delta > 0$, choose M according to tightness property and now redefine h to be

$$h(x) = \begin{cases} 0 & x \in [-\infty, -M_1 - 1) \\ \left.\begin{array}{l}\text{smooth increasing function s(x)} \\ s.t: \ s(-M_1-1) = 0 \ s(-M_1) = 1\end{array}\right\} & x \in [M_1, M_1 + 1] \\ 1 & x \in (-M_1, X - \delta) \\ \left.\begin{array}{l}\text{smooth decreasing function s(x)} \\ s.t: \ s(X-\delta) = 1 \ s(X) = 0\end{array}\right\} & x \in [X - \delta, X] \\ 0 & x \in (X, \infty] \end{cases}$$

$$F_n(X) = \int \mathbb{1}_{(-\infty,X)} d\mu_n \geq \int h d\mu_n$$

$$\begin{aligned}\lim_{n\to\infty} \inf F_n(X) \geq \lim_{n\to\infty} \int h d\mu_n = \int h d\mu \geq \int \mathbb{1}_{(-\infty,X-\delta)} d\mu - \int \mathbb{1}_{(-\infty,-M)} d\mu \\ = F(x-\delta) - \mu(-\infty, -M) \geq F(X-\delta) - \epsilon\end{aligned}$$

$$\boxed{\lim_{n\to\infty} \inf F_n(X) \geq F(X) \quad as\ \epsilon, \delta \to 0}$$

We will now do lim sup in a simular way. Redefine h to be:

$$h(x) = \begin{cases} 0 & x \in [-\infty, -M_1 - 1) \\ \left.\begin{array}{l}\text{smooth increasing function s(x)} \\ s.t: \ s(-M_1-1) = 0 \ s(-M_1) = 1\end{array}\right\} & x \in [M_1, M_1 + 1] \\ 1 & x \in (-M_1, X) \\ \left.\begin{array}{l}\text{smooth decreasing function s(x)} \\ s.t: \ s(X) = 1 \ s(X+\delta) = 0\end{array}\right\} & x \in [X, X + \delta] \\ 0 & x \in (X + \delta, \infty] \end{cases}$$

$$F_n(X) = \int \mathbb{1}_{(-\infty,X]} d\mu_n = \mu(-\infty,-M] + \int \mathbb{1}_{(-M,X]} d\mu_n \leq \epsilon + \int h d\mu_n$$

$$\begin{aligned} \lim_{n\to\infty} \sup F_n(X) &\leq \epsilon + \lim_{n\to\infty} \int h d\mu_n = \epsilon + \int h d\mu \\ &\leq \epsilon + \int \mathbb{1}_{(-\infty,X+\delta]} d\mu = \epsilon + F(X+\delta) \end{aligned}$$

$$\boxed{\lim_{n\to\infty} \sup F_n(X) \leq F(X) \quad as\ \epsilon, \delta \to 0}$$

$$\Rightarrow \lim_{n\to\infty} F_n(X) = F(X)$$

$\square$

3.1 Characteristic functions and Fourier transforms

Definition 3.4. *for a probability measure μ on $(\mathbb{R}, \mathcal{B})$*

$$\left.\begin{aligned} &\tilde{\mu} : \mathbb{R} \to \mathbb{C} \\ &\tilde{\mu}(t) = \int_{\mathbb{R}} e^{itX} d\mu(x) \end{aligned}\right\} \quad \textit{fourier transform of } \mu$$

Definition 3.5. *for an integrable function $h : \mathbb{R} \to \mathbb{R}$*

$$\left.\begin{aligned} &\tilde{h} : \mathbb{R} \to \mathbb{C} \\ &\tilde{h}(t) = \int_{\mathbb{R}} e^{itX} h(x) dx \end{aligned}\right\} \quad \textit{fourier transform of } h$$

Definition 3.6. *if X is a random variable*

$$\varphi_X : \mathbb{R} \to \mathbb{C} \qquad \varphi_X(t) = Ee^{itX}$$

is the Characteristic function of X

Remark.

- *μ is the law of X*

$$\varphi_X(t) = Ee^{itX} = \int e^{itX} d\mu(x) = \tilde{\mu}(t)$$

- *if μ has density h then*

$$\tilde{\mu} = \tilde{h}$$

Theorem 3.4 (Properties of the Characteristic Function).

www.mymathscloud.com
My Maths Cloud

www.mymathscloud.com
My Maths Cloud

(1) $\varphi(0) = 1$

(2) $\varphi_{\lambda X}(t) = Ee^{it\lambda X} = \varphi_X(\lambda t)$

(3) if X *and* Y *are independent then*

$$\varphi_{X+Y}(t) = Ee^{it(X+Y)} = Ee^{itX} \cdot Ee^{itY} = \varphi_X(t) \cdot \varphi_Y(t)$$

(4) φ *is continuous*
$t_n \to t$,

$$\varphi(t_n) = Ee^{it_nX} \underset{*}{\overset{DCT}{\Rightarrow}} Ee^{itX} = \varphi(t)$$

** since everything bounded by 1*

Example 3.1. *if* (X_i) *are iid with characteristic function* φ *then*

$$\varphi_{\frac{X_1+\cdots+X_n}{\sqrt{n}}}(t) = \varphi(\frac{t}{\sqrt{n}})^n$$

Example 3.2. *Uniform distribution on* $[-1, 1]$

$$\varphi(t) = \int_{-1}^{1} e^{itX} \cdot \frac{1}{2}dx = \frac{1}{2} \cdot \frac{e^{itX}}{it}\Big|_{-1}^{1} = \frac{1}{t} \cdot \frac{e^{it} - e^{-it}}{2i} = \frac{\sin t}{t}$$

Example 3.3 (Cauchy Distributiion: Density)**.**

$$f(X) = \frac{1}{\pi} \cdot \frac{1}{1 + X^2}$$

Cauchy distribution has ∞ *expectation since no integrable. for* $t > 0$

$$\varphi(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \cdot \frac{e^{itX}}{1 + X^2}dx = \frac{1}{\pi} \cdot 2\pi i \, Res(i)$$

repeat for $t < 0$ *to get*

$$\varphi(t) = e^{-|t|} \quad \textit{not differentiable since } E|X| \textit{ is } \infty$$

Example 3.4 (X, Y- cauchy independent. Distribution of $X + Y$?)**.**

$$f_{X+Y}(X) = \int_{-\infty}^{\infty} \frac{1}{\pi(1 + (t - X)^2)} \cdot \frac{1}{\pi(1 + X^2)}dX = \frown$$

$$\varphi_{X+Y}(t) = e^{-|t|}e^{-|t|} = e^{-2|t|}$$

$$\varphi_{\frac{X+Y}{2}}(t) = e^{\frac{-|t|}{2}}e^{\frac{-|t|}{2}} = e^{-|t|} \quad \Rightarrow \frac{X+Y}{2} \textit{ is cauchy.}$$

www.mymathscloud.com
My Maths Cloud

Example 3.5 (N(0,1)).

$$\varphi(t) = \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itX} e^{\frac{-X^2}{2}} dX = \underbrace{\left(\tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-(X-it)^2}{2}} dX \right)}_{=1} e^{\frac{-t^2}{2}} = e^{\frac{-t^2}{2}}$$

since by using contor integration form complex analysis:

$$\tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-(X-it)^2}{2}} dX = \tfrac{1}{\sqrt{2\pi}} \int_{-} e^{\frac{-z^2}{2}} dz = \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx = 1$$

Lemma 3.5. *Let h be a C^2 test function. Then*

$$|\tilde{h}(t)| \leq \tfrac{c}{t^2} \quad \textit{for some } c > 0$$

Proof.

$$|\tilde{h}(t)| = \left| \int e^{itX} h(X) dX \right| = \left| \int \overset{v}{h(X)} \Big(\overset{du}{\tfrac{1}{t} \tfrac{de^{itX}}{dX}} \Big) dX \right| = \left| \underbrace{\left[\tfrac{1}{t} h(X) e^{itX} \right]_{-\infty}^{\infty}}_{=0} - \tfrac{1}{t} \int_{-\infty}^{\infty} \overset{v}{h'(X)} \Big(\overset{du}{\tfrac{1}{t} \tfrac{de^{itX}}{dX}} \Big) dX \right|$$

$$= \left| - \underbrace{\left[\tfrac{1}{t^2} h''(X) e^{itX} \right]_{-\infty}^{\infty}}_{=0} + \tfrac{1}{t^2} \int_{-\infty}^{\infty} h''(X) e^{itX} dX \right|$$

$$= \tfrac{1}{t^2} \left| \int_{-\infty}^{\infty} e^{itX} h''(X) dX \right| \leq \tfrac{1}{t^2} \left(\int_{-\infty}^{\infty} e^{itX} \left| h''(X) \right| dX \right) = \frac{c}{t^2}$$

□

Remark.

(a) $\int\limits_{-\infty}^{\infty} \frac{\sin t}{t} dt = \frac{\pi}{2}$

(b) $\int\limits_{-\infty}^{X} \frac{\sin t}{t} dt$ *is a bounded function.*

3.2 Parseval-Plancherel

Theorem 3.6 (Parseval-Plancherel). *Let μ be a probability measure with a fourier transform φ then*

$$\int_{\mathbb{R}} h d\mu = \frac{1}{2\pi} \int_{-\infty}^{\infty} \overline{\hat{h}(t)} \varphi(t) dt$$

Proof.

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\overline{\hat{h}(t)}\varphi(t)dt = \lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\left(\int_{-\infty}^{\infty}e^{-itX}h(X)dX\right)\left(\int_{\mathbb{R}}e^{itY}d\mu(Y)\right)dt$$

$$= \lim_{T\to\infty}\frac{1}{2\pi}\int_{\mathbb{R}}\int_{-\infty}^{\infty}h(X)\int_{-T}^{T}e^{it(Y-X)}dt\,dx\,d\mu(Y)$$

Fubini: integrable and continuous with $Leb \times Leb \times \mu$ on $[T, T] \times \mathbb{R} \times \mathbb{R}$

$$|h(X)e^{it(Y-X)}| = |h(X)|$$

$$= \lim_{T\to\infty}\frac{1}{2\pi}\int_{\mathbb{R}}\int_{-\infty}^{\infty}h(X)\underbrace{\left.\frac{e^{it(Y-X)}}{i(Y-X)}\right|_{-T}^{T}}_{*}dx\,d\mu(Y)$$

$$(*) \qquad \left.\frac{e^{it(Y-X)}}{2i(Y-X)}\right|_{-T}^{T} = \frac{e^{iT(Y-X)} - e^{-iT(Y-X)}}{2i(Y-X)} = \frac{\sin(T(X-Y))}{X-Y}$$

$$= \lim_{T\to\infty}\frac{1}{\pi}\int_{\mathbb{R}}\int_{-\infty}^{\infty}h(X)\frac{\sin(T(X-Y))}{X-Y}dx\,d\mu(Y) \stackrel{?}{=} \int_{\mathbb{R}}h(Y)d\mu(Y)$$

To prove the theorem it suffices to show

$$\frac{1}{\pi}\int_{-\infty}^{\infty}h(X)\frac{\sin(T(X-Y))}{X-Y}dx \to h(Y) \quad as\ T\to\infty$$

$$\frac{1}{\pi}\int_{-\infty}^{\infty}h(X)\frac{\sin(T(X-Y))}{X-Y}dx = \frac{1}{\pi}\int_{-\infty}^{Y}h(X)\frac{\sin(T(X-Y))}{X-Y}dx + \frac{1}{\pi}\int_{Y}^{\infty}h(X)\frac{\sin(T(X-Y))}{X-Y}dx$$

www.mymathscloud.com

www.mymathscloud.com

$$\frac{1}{\pi}\int_Y^\infty h(X)\frac{\sin(T(X-Y))}{X-Y}dx = \frac{1}{\pi}\left[\underbrace{h(X)\int_Y^X \frac{\sin(T(U-Y))}{U-Y}dU}_{=0}\right]_Y^\infty$$

$$-\frac{1}{\pi}\int_Y^\infty h'(X)\int_Y^\infty \frac{\sin(T(X-Y))}{X-Y}dU dx \quad \leftarrow V = T(U-V)$$

$$= -\frac{1}{\pi}\int_Y^\infty h'(X)\int_0^{T(X-Y)} \frac{\sin(V)}{V}dV dx \quad (we\ want\ \to \tfrac{h(X)}{2}\ dominatedly)$$

$$\int_0^{T(X-Y)} \frac{\sin(V)}{V}dV \to \frac{\pi}{2} \quad \text{bounded by some } c$$

$$h'(X)\int_0^{T(X-Y)} \frac{\sin(V)}{V}dV \to \frac{\pi}{2}h'(X) \quad \text{bounded by } c\cdot|h'(X)| \text{ lebesque integrable}$$

Since

$$\left|-\frac{1}{\pi}\int_Y^\infty h'(X)\underbrace{\int_0^{T(X-Y)} \frac{\sin(V)}{V}dV}_{\le c} dx\right| \le \underbrace{\frac{c}{\pi}\int_{-\infty}^\infty |h'(X)|dx}_{constant\ i.e\ int.\ wrt\ \mu}$$

By dominated convergence theorem

$$-\frac{1}{\pi}\int_Y^\infty h'(X)\int_0^{T(X-Y)} \frac{\sin(V)}{V}dV dx \to -\frac{1}{\pi}\int_Y^\infty \frac{\pi}{2}h'(X) = \frac{-1}{2}h(X)\Big|_Y^\infty = \frac{h(Y)}{2}$$

□

Theorem 3.7 (Weak Convergence ≡ Convergence of Characteristic Functions).

$$X_n \to X \Leftrightarrow \varphi_{X_n}(t) \to \varphi_X(t) \quad \forall t \in \mathbb{R}$$

Proof.
($\Rightarrow$) suppose $X_n \xrightarrow{d} X$

$$\varphi_{X_n} = Ee^{itX_n} = \int \underbrace{e^{itX}d\mu_n(X)}_{*} \to \int e^{itX}d\mu = Ee^{itX} = \varphi_X(t)$$

* $h(x)$ continuous and bounded by 1, also $e^{itX_n} = \cos(tX_n) + i\sin(tX_n)$ same true for real and complex parts.
($\Leftarrow$) suppose $\varphi_{X_n} \to \varphi_X(t) \quad \forall t \in \mathbb{R}$. Take any C^2 test function h

$$\hat{h}(t)\varphi_{X_n}(t) \to \hat{h}(t)\varphi_X(t) \quad \forall t \in \mathbb{R}$$

www.mymathscloud.com
My Maths Cloud

Dominating:

$$|\hat{h}(t)\varphi_{X_n}(t)| = |\hat{h}(t)| \overset{*}{\leq} \min\{M, \tfrac{c}{t^2}\} \qquad |\varphi_{X_n}(t)| = |Ee^{itX_n}| \leq 1$$

* by lemma 3.5 on decay of the FT, $\min\{M, \frac{c}{t^2}\}$ is integrable on $\mathbb{R}$ wrt Leb.

$$\overset{Dom}{\Rightarrow} \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{\hat{h}}(t)\varphi_{X_n}(t)dt \to \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{\hat{h}}(t)\varphi_X(t)dt$$

$$\overset{PP}{\Rightarrow} \int h d\mu_n \to \int h d\mu$$

$$\underset{3.3}{\overset{Thm}{\Rightarrow}} X_n \xrightarrow{d} X$$

$\square$

We now have the following:

$$\begin{tikzcd}
\boxed{X_n \xrightarrow{d} X} \arrow[r, Leftrightarrow, "Thm", "3.2"'] \arrow[d, Leftrightarrow, "Thm"', "3.3"] & \boxed{\underset{h \text{ const+bounded}}{\int h d\mu_n \to \int h d\mu}} \arrow[r, Rightarrow, "Thm", "3.7"'] & \boxed{\underset{\forall t}{\varphi_n(t) \to \varphi(t)}} \arrow[dl, Rightarrow, "Thm", "3.7"'] \\
\boxed{\underset{h \text{ is a } C^2 \text{ test function}}{\int h d\mu_n \to \int h d\mu}} \arrow[r, Leftrightarrow, "Thm", "3.6"'] & \boxed{\underset{h \text{ is a } C^2 \text{ test function}}{\int \tilde{h}\varphi_n dt \to \int \tilde{h}\varphi dt}} &
\end{tikzcd}$$

Theorem 3.8. *If X and Y have the same Characteristic function φ then they have the same distribution.*

Proof.

$$\left.\begin{matrix} X, X, \dots, Y \\ \varphi(t), \varphi(t), \dots, \varphi(t) \end{matrix}\right\} \underset{3.6}{\overset{Thm}{\Rightarrow}} X, X, \dots \xrightarrow{d} Y$$

$$F_Y(t) = \lim_{n\to\infty} F_X(t) = F_X(t) \quad \text{for any } t \text{ where } F_Y \text{ is continuous}$$

$$F_Y(t) = \lim_{n\to\infty} F_Y(t_n) \quad \text{where } t_n \text{ are continuity points such that } t_n \downarrow t$$

$$= \lim_{n\to\infty} F_X(t_n) = F_X(t) \quad \forall t$$

$\square$

Example 3.6. (X_n) *independent cauchy random variables.*

$$\frac{X_1 + \cdots + X_n}{n} \quad - \quad \textit{distribution?}$$

$$\varphi_{\frac{X_1+\cdots+X_n}{n}}(t) = \varphi_{\frac{X_1}{n}}(t)\cdots\varphi_{\frac{X_n}{n}}(t) = (\varphi_{\frac{X_1}{n}}(t))^n = (\varphi_{X_1}(\tfrac{t}{n}))^n = (e^{-|\frac{t}{n}|})^n = e^{-|t|}$$

$$\frac{X_1 + \cdots + X_n}{n} \quad \textit{has cauchy distribution}$$

www.mymathscloud.com
My Maths Cloud

3.3 Central Limit Theorem

Theorem 3.9. *Let X be a square integrable random variable then*

(i) φ *is twice differentiable and* φ'' *is continuous at 0*

(ii) $\varphi'(0) = iEX$

(iii) $\varphi''(0) = -EX^2$

Proof.

$$\frac{\varphi(t+h) - \varphi(t)}{h} = \frac{Ee^{i(t+h)X} - Ee^{itX}}{h} = E\left(e^{itX}\frac{e^{iXh}-1}{h}\right) \overset{*DCT}{\to} E[e^{itX}iX]$$

* For the DCT:

$$\frac{e^{iXh}-1}{h} = iXe^{ihX}\big|_{h=0} = iX \quad (\textit{dominatedly by } |X|)$$

$$\left|\frac{e^{iXh}-1}{h}\right| = \left|\frac{1}{h}\int_0^{hX} e^{is}ds\right| \le \left|\frac{1}{h}\cdot hX\right|$$

$$\Rightarrow \varphi(t) \text{ is differentiable at any } t \text{ and } \varphi'(0) = iEX$$

$$\frac{\varphi'(t+h) - \varphi'(t)}{h} = \frac{E(e^{i(t+h)X}iX) - E(e^{itX}iX)}{h} = E\left(iXe^{itX}\frac{e^{ith}-1}{h}\right) \overset{DCT}{\to} -E[X^2e^{itX}]$$

$$\Rightarrow \varphi(t) \text{ is twice differentiable at any } t \text{ and } \varphi''(0) = -EX^2$$

Let $t_n \to 0$ and show $\varphi''(t_n) \to \varphi''(0)$

$$\varphi''(t_n) = -E[X^2e^{it_nX}] \overset{DCT}{\to} -EX^2 = \varphi''(0)$$

Where X^2 is dominated by X^2 ☐

Lemma 3.10. *if* $|z| < \frac{1}{2}$ *then*

$$|\ln(1+z) - z| \le |z|^2$$

Proof.

$$|\ln(1+z) - z| = \left|\int_\Gamma \left(\frac{1}{1+z} - 1\right) dz\right| \le \int_\Gamma \left|\frac{1}{1+z} - 1\right| dz \le 2\int_\Gamma |z|\, dz = 2\int_0^{|z|} x dx = |z|^2$$

☐

Theorem 3.11 (Central Limit Theorem - CLT)**.** *Let (X_n) be iid random variables with expectation μ and Varience σ^2 denote $S_n = X_1 + \cdots + X_n$ then*

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \to N(0,1)$$

www.mymathscloud.com
My Maths Cloud

Meaning.

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \sim \underset{N(0,1)}{Z}$$

$$\frac{X_1 + \cdots + X_n}{n} \sim \mu + \frac{\sigma Z}{\sqrt{n}}$$

Proof. It suffices to prove CLT for $\mu = 0,\ \sigma = 1$ If (Y_n) is iid with μ, σ^2 then

$$X_n = \frac{Y_n - \mu}{\sigma} \quad \text{are iid with } \mu = 0,\ \sigma = 1$$

$$\frac{Y_1 + \cdots + Y_n - n\mu}{\sigma\sqrt{n}} = \frac{(\mu + \sigma X_1) + \cdots + (\mu + \sigma X_n) - n\mu}{\sigma\sqrt{n}} = \frac{X_1 + \cdots + X_n}{\sqrt{n}} \to N(0,1)$$

assume $\mu = 0,\ \sigma = 1$. We want to prove:

$$\frac{X_1 + \cdots + X_n}{\sqrt{n}} \to N(0,1)$$

$$\varphi_{\frac{X_1+\cdots+X_n}{\sqrt{n}}}(t) \to e^{\frac{-t^2}{2}} \quad \text{Characteristic function of } N(0,1)$$

Denote by φ the Characteristic function of all X_n.

$$\varphi_{\frac{X_1+\cdots+X_n}{\sqrt{n}}}(t) = \left(\varphi(\frac{t}{\sqrt{n}})\right)^n \overset{?}{\to} e^{\frac{-t^2}{2}} \quad \forall t$$

We need to understand φ in the vicinity of 0:

$$\varphi(t) = \underbrace{\varphi(0)}_{1} + \underbrace{\varphi'(0)t}_{iEX=0} + \frac{\varphi''(\zeta)t^2}{2}$$

where ζt is between 0 and t.

$$\varphi''(\zeta t) = \underbrace{\varphi''(0)}_{-EX^2=-1} + \epsilon(t)$$

Where $\epsilon(t) \to 0$ as $t \to 0$

$$\varphi(t) = 1 - \frac{t^2}{2} + \frac{\epsilon(t)t^2}{2}$$

$$(\varphi(\frac{t}{\sqrt{n}}))^n = \left(1 - \frac{t^2}{2n} + \frac{\epsilon(t)t^2}{2n}\right)^n =$$

$$\boxed{\left(1 - \frac{t^2}{2n}\right)^n = e^{\frac{-t^2}{2}}}$$

It suffices to show:

$$n \log \varphi(\tfrac{t}{\sqrt{n}}) \to \frac{-t^2}{2}$$

www.mymathscloud.com
My Maths Cloud

$$\theta_n = \frac{-t^2}{2n} + \frac{\epsilon(t)t^2}{2n}$$

$$\begin{aligned}\left|n\log(1+\theta_n(t)) + \frac{t^2}{2}\right| &\leq \left|n\theta_n(t) + \frac{t^2}{2}\right| + |n\log(1+\theta_n(t)) - n\theta_n(t)| \\ &\leq \left|\frac{\epsilon(t)t^2}{2}\right| + \underset{*}{n|\theta_n(t)|^2}\end{aligned}$$

* By lemma 3.10 since $|\theta_n(t)| < \frac{1}{2}$ for all large n.

$$= \underset{\to 0}{\left|\frac{\epsilon(t)t^2}{2}\right|} + \underset{\to 0}{\left|\frac{-t^2}{2\sqrt{n}} + \frac{\epsilon(t)t^2}{2\sqrt{n}}\right|^2} \to 0$$

□

4 Martingales

Example 4.1. $\Omega = \{1,2,3,4\},\ \Sigma = 2^{\Omega},\ P(\{1\}) = \cdots = P(\{4\}) = \frac{1}{4}$

$$X : \Omega \to \Omega \qquad X(i) = i$$

Lets look at some smaller σ*-algebra*

$$\mathcal{F} = \{\phi, \Omega, \{1,2\}, \{3,4\}\}$$

X *is not measurable with respect to* $\mathcal{F}$ *since*

$$X^{-1}(\{1\}) = \{1\} \notin \mathcal{F}$$

we want to construct Y *such that it is* $\mathcal{F}$ *measurable and resembles* X*. ie*

$$\int_{\{1,2\}} XdP = \int_{\{1,2\}} YdP \qquad \int_{\Omega} XdP = \int_{\Omega} YdP$$
$$\int_{\{3,4\}} XdP = \int_{\{3,4\}} YdP \qquad \int_{\phi} XdP = \int_{\phi} YdP$$

Theorem 4.1 (Existance & Uniqueness of Conditional Expectation)**.**

Definition 4.1. *Let* (Ω, Σ, P) *be a probability space and* $X : \Omega \to \mathbb{R}$ *be a random variable such that* $E|X| < \infty$ *and let* $\mathcal{F} \subset \Sigma$ *be a sub* σ*-algebra. Then* $\exists$ *a random variable* Y *such that*

(a) Y *is* $\mathcal{F}$ *measurable*

(b) $E|Y| < \infty$

www.mymathscloud.com
My Maths Cloud

(c) $\int\limits_A Y dP = \int\limits_A X dP \quad A \in \mathcal{F}$

If Y *and* $\tilde{Y}$ *are two such random variables then*

$$Y = \tilde{Y} \ a.s.$$

Each such random variable is called a Conditional Expectation of X *wrt* $\mathcal{F}$ *and is denoted* $E[X|\mathcal{F}]$

Proof. $X = X^+ - X^-$. Define:

$$\begin{cases} V^+(A) = \int\limits_A X^+ dP \\ V^-(A) = \int\limits_A X^- dP \end{cases} \qquad A \in \mathcal{F}$$

V^+, V^- are finite measures on $\mathcal{F}$

$$V^+ \ll p \qquad (\text{if } P(A) = 0 \text{ then } V^\pm = \int X^\pm dP = 0)$$

$$V^+ \gg p$$

RadonNikodym theorem: $\exists$ a density Y^+ *of* V^+ *wrt* P
$\exists$ a density Y^- *of* V^- *wrt* P

which means Y^+, Y^- are $\mathcal{F}$-measurable, P-integrable and

$$V^+(A) = \int\limits_A Y^+ dP \qquad V^-(A) = \int\limits_A Y^- dP \quad \forall A \in \mathcal{F}$$

Define $Y = Y^+ - Y^-$, $\mathcal{F}$-measurable and P-integrable

$$\int\limits_A Y dP = \int\limits_A (Y^+ - Y^-) dP = V^+(A) - V^-(A) = \int\limits_A X^+ dP - \int\limits_A X^- dP = \int\limits_A X dP \quad \forall A \in \mathcal{F}$$

Let Y and $\tilde{Y}$ be two random variables. Suppose

$$P(|Y - \tilde{Y}| \neq 0) > 0$$

$$P\left(\bigcup_{k=1}^{\infty} \{Y - \tilde{Y} > \tfrac{1}{k}\} \bigcup \bigcup_{k=1}^{\infty} \{Y - \tilde{Y} < -\tfrac{1}{k}\}\right) > 0$$

$\Rightarrow$ at least one of the sets has positive probability. Suppose

$$P(Y - \tilde{Y} > \tfrac{1}{k}) = c > 0 \quad \text{Denote this event by } A \in \mathcal{F}$$

$$0 = \int X dP - \int X dP = \int\limits_A Y dP - \int\limits_A \tilde{Y} dP = \int (Y - \tilde{Y}) dP > \tfrac{1}{k} P(A) = \tfrac{c}{k} > 0 \quad \textit{Contradiction!}$$

$\square$

www.mymathscloud.com
My Maths Cloud

Remark. $\mathcal{F} = \{\phi, \Omega\}$, $EX = c$ *(since $\mathcal{F}$ - measurable). What is c?*

$$\int_\Omega \underbrace{E[X|\mathcal{F}]}_{c} dP = \underbrace{\int_\Omega X dP}_{EX}$$

So $E[X|\mathcal{F}] = EX$

Theorem 4.2 (Properties of Conditional Convergence)**.** *Suppose all random variables below are integrable*

(a) $E\left[E[X|\mathcal{F}]\right] = EX$

(b) if X is $\mathcal{F}$-measurable then $E[X|\mathcal{F}] = X$

(c) (Linearity)

$$E[aX + bY \,|\, \mathcal{F}] = aE[X|\mathcal{F}] + bE[Y|\mathcal{F}]$$

(d) (Positivity) if $X \geq 0$ a.s. then

$$E[X|\mathcal{F}] a.s.$$

(e) (Conditional Mon) if $0 \leq X_n \uparrow X$ then

$$0 \leq E[X_n|\mathcal{F}] \uparrow E[X|\mathcal{F}]$$

(f) (Taking Out What Is Know) if X is $\mathcal{F}$-measurable then

$$E[XY|\mathcal{F}] = X \cdot E[Y|\mathcal{F}]$$

(g) (independence) if X is independent of $\mathcal{F}$ then

$$E[X|\mathcal{F}] = EX$$

(h) (Tower property) if $\mathcal{G} \subset \mathcal{F}$ are two sub-σ-algebras

$$E\left[E[X|\mathcal{F}] \,|\, \mathcal{G}\right] = E[X|\mathcal{G}]$$

(i) (Jenson's Inequality) Let φ be a convex function such that

$$E|\varphi(X)| < \infty$$

then

$$E[\varphi(X)|\mathcal{F}] \geq \varphi\left(E[X|\mathcal{F}]\right)$$

www.mymathscloud.com
My Maths Cloud

Proof.

(a):

$$E\left[E[X|\mathcal{F}]\right] = \int_\Omega E[X|\mathcal{F}]dP = \int_\Omega XdP = EX$$

(b): Obvious

(c): Obvious

(d): Denote $Y = E[X|\mathcal{F}]$ and assume $P(Y<0)>0$

$$P\left(\bigcup_{k=1}^{\infty}\{Y<-\tfrac{1}{k}\}\right)>0 \quad So \quad P\left(\underbrace{Y<-\tfrac{1}{k}}_{denote:\, A\in\mathcal{F}}\right)>0 \quad for\ some\ k$$

$$0=\int_A XdP - \int_A YdP = \int_A \underbrace{(\underset{\geq 0}{X} - \underset{<-\frac{1}{k}}{Y})}_{\geq \frac{1}{k}}dP$$

(e): denote $Y_n = E[X_n|\mathcal{F}]$

$$\Rightarrow Y_n \geq 0$$

$$X_{n+1}-X_n\geq 0 \overset{(c+d)}{\Rightarrow} Y_{n+1}-Y_n\geq 0 \Rightarrow Y_n\uparrow$$

Define

$$Y=\lim_{n\to\infty} Y_n$$

Why is $Y=E[X|\mathcal{F}]$? Y is $\mathcal{F}$-measurable

$$\int_A YdP \overset{MCT}{=} \lim_{n\to\infty}\int_A Y_n dP = \lim_{n\to\infty}\int_A X_n dP \overset{Mon}{=} \int_A XdP \quad A\in\mathcal{F}$$

(f): Let $\mathbb{1}_B,\ B\in\mathcal{F},\ \forall A\in\mathcal{F}$

$$\int_A \underset{\mathbb{1}_B}{Z} E[X|\mathcal{F}]dP = \int_{A\cap B\in\mathcal{F}} E[X|\mathcal{F}]dP = \int_{A\cap B} XdP = \int_A \mathbb{1}_B X dp = \int_A ZXdP$$

extend to whole of Z in standard way.

(g): Let $Z = \mathbb{1}_B,\ B \perp\!\!\!\perp \mathcal{F},\ A \in \mathcal{F}$

$$\int\limits_A \underbrace{Z}_{\mathbb{1}_B} dP = P(A \cap B) = P(A)P(B) = \int\limits_A \underbrace{EZ}_{=P(B)} dP$$

this can be extended to all Z in the the standard plane.

(h): Let $A \in \mathcal{G}$

$$\underbrace{\int\limits_A E[X|\mathcal{F}]dP}_{=\int\limits_A XdP} = \underbrace{\int\limits_A E[X|\mathcal{G}]dP}_{=\int\limits_A XdP}$$

(i): not proven here.

$\square$

Example 4.2. (X_n) *independent bernulli (p)*

$$\sigma(X_1, \ldots, X_m) \le X_1 + \cdots + X_n \quad n > m$$

$$E[S_n \,|\, \underbrace{\sigma(X_1, \ldots, X_m)}_{\mathcal{F}}] = X_1 + \cdots + X_m + p(n-m)$$

$$\begin{aligned} E[S_n \,|\, \underbrace{\sigma(X_1, \ldots, X_m)}_{\mathcal{F}}] &= E[X_1 + \cdots + X_n \,|\, \sigma(X_1, \ldots, X_m)] \\ &= \sum_{i=1}^{n} E[X_i \,|\, \sigma(X_1, \ldots, X_m)] \\ &= \sum_{i=1}^{m} EX_i + = \sum_{i=m+1}^{n} EX_i \\ &= \sum_{i=1}^{m} EX_i + p(n-m) \end{aligned}$$

Example 4.3. $0 < p_1 \neq p_2 < 1$ *Flip a coin and choose* p_1 *or* p_2 *with* $p = \frac{1}{2}$

$$X = \begin{cases} p_1 & with\, probability\ \frac{1}{2} \\ p_2 & with\, probability\ \frac{1}{2} \end{cases}$$

Now flip an unfair coin with this random probability X. Denote the result by Y

$$E[Y|\sigma(X)] = 1 \cdot X + 0 \cdot (1 - X) = X$$

$$\begin{aligned} &Y_1 = Bernulli\,(p_1) \qquad \textit{Both Independent \& Independent of } X \\ &Y_2 = Bernulli\,(p_2) \end{aligned}$$

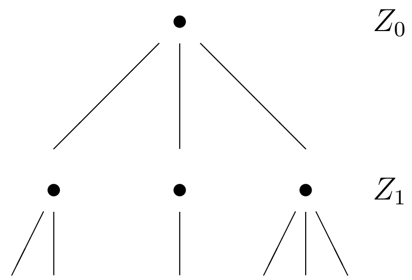
www.mymathscloud.com

www.mymathscloud.com
My Maths Cloud

$$Y = \mathbb{1}_{\{X=p_1\}} Y_1 + \mathbb{1}_{\{X=p_2\}} Y_2$$

$$\begin{aligned} E[Y|\sigma(X)] &= E[\mathbb{1}_{\{X=p_1\}} Y_1 + \mathbb{1}_{\{X=p_2\}} Y_2 \,|\, \sigma(X)] \\ &= \mathbb{1}_{\{X=p_1\}} E[Y_1 \,|\, \sigma(X)] + \mathbb{1}_{\{X=p_2\}} E[Y_2 \,|\, \sigma(X)] \\ &= \mathbb{1}_{\{X=p_1\}} \underset{p_1}{EY_1} + \mathbb{1}_{\{X=p_2\}} \underset{p_2}{EY_2} \\ &= X \end{aligned}$$

Example 4.4 (Galton Watson Tree). *Lets fix a random variable with values in* $\{1, 2, \dots\}$ *which chooses the number of edges decending from each vertex.*



$$Z_n = \text{ number of veticies in generation } n$$

$$a = \text{ the mean of offspring distribution}$$

$$E[Z_{n+1}|\sigma(Z_n)] = a \cdot Z_n$$

$$\begin{aligned} E[Z_{n+1}|\sigma(Z_n)] &= E\left[\sum_{i=1}^{\infty} \mathbb{1}_{\{Z_n=i\}} \underbrace{Z_{n+1}}_{N_1^i+\cdots+N_i^i} \,|\, \sigma(Z_n)\right] \\ &= \sum_{i=1}^{\infty} \mathbb{1}_{\{Z_n=i\}} \underset{ai}{E(N_1^i + \cdots + N_i^i)} \\ &= Z_n \cdot a \end{aligned}$$

Definition 4.2 (Filtration). *Let* (Ω, Σ, P) *be a probability space. A growing sequence of* σ*-algebras*

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$$

is called a Filtration.

Definition 4.3 (Adapted). *A sequence* (X_n) *of random variables is called adapted to* $\mathcal{F}$ *if* X_n *is* $\mathcal{F}_n$ *measurable* $\forall n$*.*

Definition 4.4 (Natural Filtration). *Let* (X_n) *be a sequence of random variables*

$$\sigma(X_1) \subset \sigma(X_1, X_2) \subset \cdots$$

www.mymathscloud.com
My Maths Cloud

is called the natural filtration for (X_n)*. It is natural since each* X_i *is measurable with respect to* $\sigma(X_1, \ldots, X_n)$*.* (X_n) *is always adapted to its natural filtration.*

$$S_n = X_1 + \cdots X_n \quad is \quad \sigma(X_1 + \cdots X_n)$$

S_n *is also adapted to the natural filtration of* (X_i)*.*

Example 4.5 (Fair Game)**.** *We will illstrate the concept of a fair game though an example.* (X_n) *bernulli* $(\frac{1}{2})$ *with values* ± 1

$$S_n = \sum_{i=1}^{n} Y_i X_i$$

Where Y_i *is the ammount which I bet at day* i*, this may be random but can also depend on* $X_1, \ldots, X_{n-1}$*.*

$$E[S_n|\sigma(X_1, \ldots, X_{n-1})] \stackrel{*}{=} S_{n-1}$$

** I want this equality to have a fair game.*

$$\begin{aligned} E[S_n|\sigma(X_1, \ldots, X_{n-1})] &= E[S_{n-1} + Y_n X_n|\sigma(X_1, \ldots, X_{n-1})] \\ &= S_{n-1} + Y_n \underbrace{EX_n}_{0} \\ &= S_{n-1} \end{aligned}$$

Definition 4.5 (Martingale)**.** *Let* (Ω, Σ, P) *be a probability space and let* $(\mathcal{F}_n)_{n\geq 0}$ *be a filtration. A sequence* (X_n) *of random variables is a martingale if*

(1) (X_n) *is adapted to* $(\mathcal{F}_n)$

(2) $E|X_n| < \infty$

(3) $E[X_n|\mathcal{F}_{n-1}] = X_{n-1} \quad \forall n \geq 1$

It is called a sub-martingale if we have conditions (1) + (2) +

(3)* $E[X_n|\mathcal{F}_{n-1}] \geq X_{n-1} \quad \forall n \geq 1$

Remark. *if* (X_n) *is a martingale*

(1) $E[X_n|\mathcal{F}_m] = X_m \quad \forall n \geq m$

$$E[X_n|\mathcal{F}_m] \stackrel{tower}{=} E[E[X_n|\mathcal{F}_{n-1}]|\mathcal{F}_m] = \cdots = E[X_{m+1}|\mathcal{F}_m] = X_m$$

(2) $EX_n = EX_0$

$$EX_n = E(\underbrace{E[X_n|\mathcal{F}_0]}_{X_0})$$

www.mymathscloud.com
My Maths Cloud

Example 4.6. *Let (X_n) be a sequence of integrable random variables with mean 0. $(\mathcal{F}_n)$ natural filtration. Then S_n is a martingale where*

$$S_n = X_1 + \cdots + X_n$$

need to check condition (3):

$$E[S_n|\mathcal{F}_{n-1}] = E[S_{n-1} + X_n|\mathcal{F}_{n-1}] = S_{n-1} + \underbrace{EX_n}_{0} = S_{n-1}$$

Example 4.7 (Galton Watson Tree)**.** (Z_n)

$$E[Z_n|\mathcal{F}_{n-1}] = a \cdot Z_{n-1}$$

if $a = 1$ then (Z_n) is usually a martingale.
if $a \neq 1$

$$W_n = \frac{Z_n}{a^n}$$

$$E\left[\frac{Z_n}{a^n}|\mathcal{F}_{n-1}\right] = \frac{1}{a^n} \cdot a \cdot Z_{n-1} = \frac{Z_{n-1}}{a^{n-1}} = W_{n-1}$$

Example 4.8 (Polya Urn)**.** *In each round $n = i$ a ball is added by a random variable*

$$\begin{array}{lll} n = 0 & \bullet \circ & X_0 = \frac{1}{2} \textit{ probability of black ball} \\ n = 1 & \bullet \circ \bullet & X_1 = \frac{2}{3} \textit{ probability of black ball} \\ n = 2 & \bullet \circ \bullet \bullet & X_2 = \frac{3}{4} \textit{ probability of black ball} \\ n = 3 & \bullet \circ \bullet \bullet \circ & X_3 = \frac{3}{5} \textit{ probability of black ball} \\ \vdots & \vdots & \vdots \qquad\qquad \vdots \end{array}$$

$$X_i = \textit{ Probability you add a black ball next time}$$

Looks like it is not a martingale, but it is! see Hw 7 something along the lines of $EX_n = \frac{1}{2}$

4.1 Stopping Times

Definition 4.6 (Stopping Time)**.** *Let (Ω, Σ, P) be a probability space and let $(\mathcal{F}_n)$ be a filtration. A random variable*

$$\tau : \Omega \to \{0, 1, 2, \dots\} \cup \{\infty\}$$

is called a stopping time if

$$\{\tau = n\} \in \mathcal{F}_n \quad \forall n \in \mathbb{N}$$

Example 4.9.

$$\tau = \textit{First rainy day in March}$$

www.mymathscloud.com
My Maths Cloud

$$\left.\begin{array}{ll} \tau = 0? & \text{"I know if it rains at day 0"} \\ \tau = 1? & \text{"I know if it rains at day 1"} \\ \vdots & \quad\vdots \\ \tau = n? & \text{"I know if it rains at day n"} \end{array}\right\} \quad \textit{Stopping Time}$$

$$\tau = \textit{Last sunny day in March}$$

$$\left.\begin{array}{ll} \tau = 0? & \text{"I dont know it at day 0"} \\ \tau = 1? & \text{"I dont know it at day 1"} \\ \vdots & \quad\vdots \\ \tau = n? & \text{"I dont know it at day n"} \end{array}\right\} \quad \begin{array}{c}\textit{NOT a} \\ \textit{Stopping Time}\end{array}$$

Example 4.10 (Random walk)**.** *let (X_n) be ± 1 with $p = \frac{1}{2}$ and independent.*

$$S_n = X_1 + \cdots + X_n$$

Let $\mathcal{F}_n$ be the natural filtration of (X_n), (to which S_n is always adapted).

$$\tau = \min\{n : S_n = 2\} \quad - \quad \textit{Stopping Time}$$

since

$$\{\tau = n\} = \{X_1 = 2, X_2 \neq 2, \ldots, X_{n-1} \neq 2, X_n = 2\}$$
$$\tau = \textit{ the first time I visit } 2 \in \mathcal{F}_n$$

More generally

$$\tau = \min\{n : S_n \in \{a, b\}\}$$

τ_i is a stopping time.

$$\{\tau = n\} = \{S_1 \notin \{a, b\}, \ldots, S_{n-1} \notin \{a, b\}, S_n \in \{a, b\}\} \in \mathcal{F}_n$$

Winnings $= bp_1 - a(1 - p_1) \overset{?}{\geq} 0$ *Can we beat a fair game?*

Definition 4.7. *Suppose $\tau < \infty$ a.s, let (X_n) be a random process, $(X_{n \wedge \tau}) \geq 0$ is another random process refered to as the origonal process stopped at a time τ*

$$X_\tau - \textit{ random variable (=value of the process } (X_n) \textit{ stopped at a time } \tau)$$

Lemma 4.3. *τ is a stopping time iff $\{\tau \leq n\} \in \mathcal{F}_n$, $\forall n = 0, 1, 2, \ldots$*

Proof.
$(\Rightarrow)$

$$\begin{array}{ccccccc} \{\tau \leq n\} = \{\tau \underset{\in}{=} 0\} & \cup & \{\tau \underset{\in}{=} 1\} & \cup \cdots & \cup & \{\tau \underset{\in}{=} n\} \\ \mathcal{F}_0 & \subset & \mathcal{F}_1 & \subset \cdots & \subset & \mathcal{F}_n \end{array}$$

$(\Leftarrow)$

$$\begin{array}{ccc} \{\tau = n\} = \{\tau \underset{\in}{\leq} n\} & \setminus & \{\tau \leq \underset{\in}{n} - 1\} \\ \mathcal{F}_n & \supset & \mathcal{F}_{n-1} \end{array}$$

$\square$

www.mymathscloud.com

Theorem 4.4. *Let (X_n) be a martingale / submartingale, then $(X_{n\wedge\tau})$ is a martingale / submartingale.*

Proof.

$$X_{n\wedge\tau} = \sum_{i=0}^{n-1} \mathbb{1}_{\tau=i} X_i + \underset{*}{\mathbb{1}_{\{\tau>n-1\}}} X_n$$

* - $\{\tau > n-1\} = \{\tau \le n-1\}^c \in \mathcal{F}_{n-1}$

- integrable
- $X_{n\wedge\tau}$ is $\mathcal{F}_n$-measurable
-

$$\begin{aligned}
E[X_{n\wedge\tau}|\mathcal{F}_{n-1}] &= E[\sum_{i=0}^{n-1} \mathbb{1}_{\tau=i} X_i + \mathbb{1}_{\tau>n-1} X_n | \mathcal{F}_n] \\
&= \sum_{i=0}^{n-1} \mathbb{1}_{\tau=i} X_i + \mathbb{1}_{\tau>n-1} \underbrace{E[X_n|\mathcal{F}_{n-1}]}_{=X_{n-1}} \\
&= \sum_{i=0}^{n-2} \mathbb{1}_{\tau=i} X_i + \mathbb{1}_{\tau>n-2} X_{n-1} \\
&= X_{\{n-1\}\wedge\tau}
\end{aligned}$$

□

Remark. $X_\tau = ?$
$n \wedge \tau \to \tau$ *a.s*
$X_{n\wedge\tau} \to X_\tau$ *a.s*
Can we say that

$$EX_0 = EX_{n\wedge\tau} \stackrel{?}{=} EX_\tau$$

Counter example: random walk S_n, $\tau =$ stopping time of 1, $S_\tau = 1$

$$ES_\tau = 1 \neq ES_0 = 0$$

This fails since $E\tau = \infty$

4.2 The Optional Stopping Theorem

Theorem 4.5 (Optimal Stopping Theorem)**.** *Suppose (X_n) is a martingale or submartingale and τ is a stopping time such that $\tau < \infty$ a.s then*

$$EX_\tau \underset{\geq}{=} EX_0$$

in each of the following cases:

www.mymathscloud.com
My Maths Cloud

(1) τ is bounded a.s

(2) $E\tau < \infty$ and for some $c > 0$:

$$|X_n - X_{n-1}| \leq c \quad \forall n$$

(3) X_n is bounded

Proof.
(1) assume $\tau \leq N$ a.s

$$X_\tau = X_{\tau\wedge N}$$
$$EX_\tau = EX_{\tau\wedge N} \underset{\geq}{=} EX_0$$

(2) $X_{\tau\wedge N} \to X_\tau$ a.s

$$X_{\tau\wedge N} = \underbrace{X_{\tau\wedge N} - X_{\tau\wedge N}} + \underbrace{X_{\tau\wedge N} - \cdots} + \cdots + \underbrace{\cdots - X_1} + \underbrace{X_1 - X_0} + X_0$$

$$|X_{\tau\wedge N}| = \underbrace{c\tau + |X_0|}_{\textit{integrable}}$$

$$\overset{DCT}{\Rightarrow} EX_\tau = \lim_{n\to 0} EX_{n\wedge\tau} \underset{\geq}{=} \lim_{n\to 0} EX_0 = 0$$

(3) $|X_{n\wedge\tau}| \leq c$

$$X_{n\wedge\tau} \to X_\tau \overset{DCT}{\Rightarrow} \underset{=EX_0}{EX_{n\wedge\tau}} \to EX_\tau$$

$\square$

Remark.

(1) it is enough to check the conditions for $(X_{n\wedge\tau})$ instead of (X_n)

(2) sometimes its easier to prove $EX_{n\wedge\tau} \to EX_\tau$ directly rather than using OST.

Applications.

(1) (S_n)-random walk. What are the probabilities that S_n hits b before $-a$?

$$\tau = \inf\{n : S_n = b \, or \, -a\}$$

S_n is a martingale. Lets stop S at time τ, $ES_\tau = ES_0 = 0$

$$ES_\tau = -aP(S_\tau = -a) + bP(S_\tau = b)$$

$$P(S_\tau = -a) + P(S_\tau = b) = 1$$
$$-a(1 - P(S_\tau = b)) + bP(S_\tau = b) = 0$$

www.mymathscloud.com
My Maths Cloud

$$P(S_\tau = b)(a+b) = a$$
$$P(S_\tau = b) = \tfrac{a}{a+b}$$
$$P(S_\tau = -a) = \tfrac{b}{a+b}$$

$\tau < \infty$ *a.s*

$$P(\tau = \infty) \leq P(-a < S_n < b) = P(\tfrac{-a}{\sqrt{n}} < \tfrac{S_n}{\sqrt{n}} < \tfrac{b}{\sqrt{n}}) \xrightarrow{CT} 0$$
$$\frac{S_n}{\sqrt{n}} \xrightarrow{d} N(0,1)$$

$S_{n\wedge\tau}$ *is bounded, so we can use OST(3)*

(2) *What is* $E\tau =$? *(from 1)*

$$X_n = S_n^2 - n$$

$$E[X_n|\mathcal{F}_{n-1}] = E[\underbrace{S_n^2}_{(S_{n-1}+X_n)^2} - n|\mathcal{F}_{n-1}] = E[S_{n-1}^2 + 2S_{n-1}X_n + X_n^2 - n|\mathcal{F}_{n-1}]$$
$$= S_{n-1}^2 + 2S_{n-1}\underset{=0}{EX_n} + \underset{E[X_n^2|\mathcal{F}_{n-1}]=1}{EX_n^2} - n$$
$$= S_{n-1}^2 - (n-1) = X_{n-1}$$

Stop X_n *at a time* τ: $EX_\tau = EX_0 = 0$

$$\boxed{P(S_\tau = b) = \tfrac{a}{a+b} \qquad P(S_\tau = -a) = \tfrac{b}{a+b}}$$

$$0 = EX_\tau = E(S_\tau^2 - \tau) = ES_\tau^2 - E\tau = a^2P(S_\tau = -a) + b^2P(S_\tau = b) - E\tau$$
$$E\tau = \frac{a^2b + b^2a}{a+b} = ab$$

Why can we use the OST?

$$ES_{n\wedge\tau}^2 - E(n\wedge\tau) \qquad \text{Why can I pass to the limit?}$$

$$\left.\begin{aligned} n\wedge\tau \to \tau &\overset{MCT}{\Rightarrow} E(n\wedge\tau) \to E(\tau) \\ |S_{n\wedge\tau}^2| \leq \max\{a^2+b^2\} &\overset{DCT}{\Rightarrow} E(S_{n\wedge\tau}^2) \to E(S_\tau^2) \end{aligned}\right\} \quad ES_\tau^2 - E\tau = 0 \quad E\tau = ES_\tau^2$$

(3) *Alice:* 666
Bob: 456
A dice is rolled each day and each time a player bets all the money they have. Each player starts with one pound and leaves the game when thre money is lost. First set up the game for Alice:
The Player wins 6 from a 1 bet if they roll a 6 any time while the game is still live, lose otherwise. game terminates when sequence 666 is acheived.

$$W_n \ - \ \text{total win up to day } n$$

www.mymathscloud.com
My Maths Cloud

$$I_n = n$$

$$\tau = \text{ the first time 666 occurs}$$

$$X_n = W_n - I_n \qquad \text{is a martingale}$$

$$0 = EX_0 = EX_\tau = EW_\tau - EI_\tau = EW_\tau - \tau \quad \Rightarrow E\tau = EW_\tau = 6^3 + 6^2 + 6$$

For Bob:
The Player wins 6 from a 1 bet if they roll a 4 first time then wins 6 per 1 if a 5 is hit, then wins 6 per 1 if a 6 is hit. Game terminates when sequence 456 is acheived.

$$E\tau = EW_\tau = 6^3$$

To show this is legit:

$$W_n = \sum_{i=1}^{n-2} 6^3 \mathbb{1}_{\{Y_i=6, Y_{i+1}=6, Y_{i+2}=6\}} + 6^2 \mathbb{1}_{\{Y_{n-1}=6, Y_n=6\}} + 6\mathbb{1}_{\{Y_n=6\}}$$

$$X_n = \sum_{i=1}^{n-2} 6^3 \mathbb{1}_{\{Y_i=6, Y_{i+1}=6, Y_{i+2}=6\}} + 6^2 \mathbb{1}_{\{Y_{n-1}=6, Y_n=6\}} + 6\mathbb{1}_{\{Y_n=6\}} - n$$

$$\begin{aligned}
E[X_n|\mathcal{F}_{n-1}] &= E\left[\sum_{i=1}^{n-2} 6^3 \mathbb{1}_{\{Y_i=6, Y_{i+1}=6, Y_{i+2}=6\}} + 6^2 \mathbb{1}_{\{Y_{n-1}=6, Y_n=6\}} + 6\mathbb{1}_{\{Y_n=6\}} - n|\mathcal{F}_{n-1}\right] \\
&= \sum_{i=1}^{n-3} 6^3 \mathbb{1}_{\{Y_i=6, Y_{i+1}=6, Y_{i+2}=6\}} + 6^3 \mathbb{1}_{\{Y_{n-2}=6, Y_{n-1}=6\}} \underbrace{E\mathbb{1}_{\{Y_n=6\}}}_{=\frac{1}{6}} \\
&\quad + 6^2 \mathbb{1}_{\{Y_{n-1}=6\}} \underbrace{E\mathbb{1}_{\{Y_n=6\}}}_{=\frac{1}{6}} + 6\underbrace{E\mathbb{1}_{\{Y_n=6\}}}_{=\frac{1}{6}} - n \\
&= \sum_{i=1}^{n-3} 6^3 \mathbb{1}_{\{Y_i=6, Y_{i+1}=6, Y_{i+2}=6\}} + 6^2 \mathbb{1}_{\{Y_{n-2}=6, Y_{n-1}=6\}} + 6\mathbb{1}_{\{Y_n=6\}} - (n-1) \\
&= W_{n-1} - (n-1) \\
&= X_{n-1}
\end{aligned}$$

Show that: $P(\tau = \infty) = 0$

$$P(\tau = \infty) \leq (1 - \tfrac{1}{6^3})^n \to 0$$

$\tau < \infty$ *a.s.*

$$E\tau < \infty \qquad (do - able)$$

$$0 = EX_{n\wedge\tau} = EW_{n\wedge\tau} - E(n \wedge \tau)$$

$$(n \wedge \tau) \uparrow \tau \overset{MCT}{\Rightarrow} E(n \wedge \tau) \to E\tau$$

$$|W_{n\wedge\tau}| \leq 6^3 + 6^2 + 6 \overset{DCT}{\Rightarrow} EW_{n\wedge\tau} \to EW_\tau$$

$$E\tau = EW_\tau = 6^3 + 6^2 + 6$$

www.mymathscloud.com
My Maths Cloud

4.3 Doob's and Kolmogorov's Submartingale Inequalities

Theorem 4.6 (Doobs Submartingale Inequality). *Let (X_n) be a submartingale, (non-negitive). Then for any $\epsilon > 0$*

$$cP(\max_{1\leq n\leq N}(X_i) > c) \leq EX_N$$

Proof.

$$\tau = \min\{n \leq N : X_n > c\} \wedge N$$

$\tau \leq N$ so lets prove $EX_\tau \leq EX_N$

$$EX_\tau = E\sum_{i=1}^N \mathbb{1}_{\tau=i}X_i = \sum_{i=1}^N \int_{\tau=1} X_i dP \leq \sum_{i=1}^N \int_{\tau=1} X_N dP$$

$$\boxed{E[X_N|\mathcal{F}_i] \geq X_i}$$

$$= \sum_{i=1}^N E(\mathbb{1}_{\tau=1}X_N) = EX_N$$

Denote $A = \{\max_{1\leq n\leq N} X_n > c\}$

$$EX_\tau \leq EX_N$$

$$(1 = \mathbb{1}_A + \mathbb{1}_{\Omega\backslash A})$$

So

$$EX_\tau\mathbb{1}_A + EX_\tau\mathbb{1}_{\Omega\backslash A} \leq EX_N\mathbb{1}_A + EX_N\mathbb{1}_{\Omega\backslash A}$$

On $\Omega\backslash A$, $\tau = N$. So the two expressions are equal

$$\Rightarrow EX_\tau\mathbb{1}_A \leq EX_N\mathbb{1}_A$$

$$EX_N \geq EX_N\mathbb{1}_A \geq EX_\tau\mathbb{1}_A \overset{*}{\geq} c\cdot P(A)$$

* Chebyshev inequality, theorem 1.3 □

Theorem 4.7 (Kolmogorov's Inequality). *Let (X_n) be a sequence of independent random variables, square integrable, with means 0. Then $\forall c > 0$*

$$c^2P\left(\max_{1\leq n\leq N}|\sum_{i=1}^N X_i| > c\right) \leq \sum_{i=1}^N EX_i^2$$

Proof. $S_n = \sum_{i=1}^n X_i$ is a sub-martingale

$$E[S_n|\mathcal{F}_{n-1}] = X_1 + \cdots + X_{n-1} + \underset{=0}{EX_n} = S_{n-1}$$

$$\Rightarrow S_n^2 \text{ is a sub-martingale}$$

$$E[S_n^2|\mathcal{F}_{n-1}] \overset{*}{\geq} (E[S_n|\mathcal{F}_{n-1}])^2 = S_{n-1}^2$$

www.mymathscloud.com

My Maths Cloud

* Conditional Jenson's inequality,Part (i) of theorem 4.2, with $\varphi(x) = x^2$

Doob: $\forall c > 0$

$$c^2 P(\max_{1 \le n \le N} (\sum_{i=1}^{n} X_i)^2 > c^2) \le ES_n^2$$

$$\Rightarrow c^2 P(\max_{1 \le n \le N} \left|\sum_{i=1}^{n} X_i\right| > c) \le E(X_1 + \cdots + X_n)^2 = \sum_{i=1}^{N} EX_i^2 + \underbrace{\sum_{i \ne j} EX_i EX_j}_{=0}$$

$\square$

Theorem 4.8 (Kolmogorov's Theorem)**.** *Let (X_n) be a sequence of independent square integrable random variables with means 0. If*

$$\sum_{n=1}^{\infty} EX_n^2 \le \infty$$

Then

$$\sum_{n=1}^{\infty} X_n \le \infty \qquad a.s.$$

Proof. We want

$$P\left(\sum_{n=1}^{\infty} X_n \le \infty\right) = 1$$

i.e.

$$P\left(\left(\sum_{i=1}^{n} X_i\right)_{n=1}^{\infty} \text{ is a cauchy sequence}\right) = 1$$

$$\boxed{\begin{array}{l} a_n \text{ is a cauchy sequence if } \forall \epsilon > 0\ \exists N \text{ such that } \forall n, m \ge N\ |a_n - a_m| < \epsilon \\ \qquad\qquad \Leftrightarrow \forall \epsilon > 0\ \exists N \textit{ such that } \forall n \ge N\ |a_n - a_N| < \epsilon \end{array}}$$

i.e.

$$P\left(\bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n \ge N} \left\{\left|\sum_{i=1}^{n} X_i - \sum_{i=1}^{N} X_i\right| < \frac{1}{k}\right\}\right) = 1$$

i.e.

$$P\left(\bigcup_{N=1}^{\infty} \bigcap_{n \ge N} \left\{\left|\sum_{i=N+1}^{n} X_i\right| < \frac{1}{k}\right\}\right) = 1 \qquad \forall k$$

i.e.

$$P\left(\bigcap_{N=1}^{\infty} \bigcup_{n \ge N} \left\{\left|\sum_{i=N+1}^{n} X_i\right| \ge \frac{1}{k}\right\}\right) = 0 \qquad \forall k$$

$$P\left(\bigcap_{N=1}^{\infty}\bigcup_{n\geq N}\left\{\left|\sum_{i=N+1}^{n}X_i\right|\geq\frac{1}{k}\right\}\right)=\lim_{N\to\infty}P\left(\bigcup_{n\geq N}\left\{\left|\sum_{i=N+1}^{n}X_i\right|\geq\frac{1}{k}\right\}\right)$$

For any $M\geq N$

$$P\left(\bigcup_{n\geq N}^{M}\left\{\left|\sum_{i=N+1}^{n}X_i\right|\geq\frac{1}{k}\right\}\right)=P\left(\max_{N\leq n\leq M}\left|\sum_{i=N+1}^{n}X_i\right|\geq\frac{1}{k}\right)\overset{theorem}{\underset{4.7}{\leq}}k^2\sum_{i=N+1}^{M}EX_i^2$$

Let $M\to\infty$

$$P\left(\bigcup_{n\geq N}\left\{\left|\sum_{i=N+1}^{n}X_i\right|\geq\frac{1}{k}\right\}\right)\leq k^2\sum_{i=N+1}^{\infty}EX_i^2\to 0\qquad as\ N\to\infty$$

□

Example 4.11.

$$\begin{aligned}\tfrac{1}{1}+\tfrac{1}{2}+\tfrac{1}{3}+\tfrac{1}{4}+\ldots&=\infty\\ \tfrac{1}{1}-\tfrac{1}{2}+\tfrac{1}{3}-\tfrac{1}{4}+\ldots&<\infty\end{aligned}$$

What happens if we put ±1 randomly? i.e:

$$\tfrac{1}{1}+\tfrac{1}{2}-\tfrac{1}{3}-\tfrac{1}{4}+\ldots=?$$

i.e.

$$\sum_{n=1}^{\infty}\frac{X_n}{n}\quad where\quad P(X_n=\pm1)=\frac{1}{2}$$

$$\sum_{n=1}^{\infty}E\left(\frac{X_n}{n}\right)^2=\sum_{n=1}^{\infty}\frac{1}{n^2}<\infty$$

So

$$\sum_{n=1}^{\infty}\frac{X_n}{n}<\infty$$

We know

$$\sum_{n=1}^{\infty}\frac{1}{n^{\alpha}}<\infty\qquad\alpha\geq1+\epsilon$$

So

$$\sum_{n=1}^{\infty}\frac{X_n}{n^{\frac{1}{2}+\epsilon}}<\infty\quad a.s.$$

Lemma 4.9 (Cesaro)**.**

$$a_n\to a\Rightarrow\frac{a_1+\cdots+a_n}{n}\to a$$

Proof. analysis 1 □

www.mymathscloud.com
My Maths Cloud

www.mymathscloud.com

Lemma 4.10 (Kronekker). *If* $\sum_{n=1}^{\infty} \frac{a_n}{n} < \infty$ *then* $\frac{a_1+\cdots+a_n}{n} \to 0$

Proof.

$$S_n = \sum_{i=1}^{n} \frac{a_i}{i} \quad n \in \mathbb{N} \text{ and } S_0 = 0$$

$$S = \lim_{n\to\infty} S_n$$

$$S_n = S_{n-1} + \frac{a_n}{n} \Rightarrow a_n = n(S_n - S_{n-1})$$

$$\begin{aligned} \frac{a_1+\cdots+a_n}{n} &= \frac{(S_1-S_0)+2(S_2-S_1)+\cdots+n(S_n-S_{n-1})}{n} \\ &= \frac{nS_n - S_0 - S_1 - S_2 - \cdots - S_{n-1}}{n} \\ &= \underbrace{S_n}_{\to S} - \underbrace{\frac{+S_0+S_1+\cdots+S_{n-1}}{n}}_{\to S \text{ by Cesaro}} \to 0 \end{aligned}$$

□

4.4 Strong Law of Large Numbers

Theorem 4.11 (Strong Law of Large Numbers for Square-Integrable r.v's). *Let* (X_n) *be iid square-integrable random variables with mean 0. Then*

$$\frac{X_1+\cdots+X_n}{n} \to 0 \qquad a.s.$$

Proof.

$$\sum_{n=1}^{\infty} E\left(\frac{X_n}{n}\right)^2 = EX_1^2 \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty \overset{Thm\,4.8}{\Rightarrow} \sum_{n=1}^{\infty} \frac{X_n}{n} < \infty \ a.s.$$

$$\overset{Thm\,4.8}{\Rightarrow} \frac{X_1+\cdots+X_n}{n} \to 0 \qquad a.s.$$

□

Lemma 4.12 (Truncation Lemma). *Let* (X_n) *be iid random variables each with* $E(|X_n|) < \infty$ *and* $EX_n = \mu$ *and define*

$$Y_n = X_n \mathbb{1}_{\{|X_n| \le n\}}$$

then

(1) $X_n = Y_n$ *eventually for all* n *a.s. i.e.*

$$\exists N(\omega) \text{ such that } \forall n \ge N(\omega)\ X_n = Y_n$$

(2) $EY_n \to \mu$

(3) $\sum_{n=1}^{\infty} \frac{var(Y_n)}{n^2} < \infty$

www.mymathscloud.com

Proof. (1) We want:

$$P(X_n \neq Y_n \ i.o.) = 0$$

By BC1, 2.4:

$$\sum_{n=1}^{\infty} P(X_n \neq Y_n \ i.o.) = \sum_{n=1}^{\infty} P(|X_n| > n) = \sum_{n=1}^{\infty} P(|X_1| > n) \overset{MCT}{=} E\sum_{n=1}^{\infty} \mathbb{1}_{\{|X_n|>n\}} \leq E|X_1| < \infty$$

Result follows.
(2)

$$EY_n = E(X_n \mathbb{1}_{\{|X_n|\leq n\}}) = E(X_1 \underbrace{\mathbb{1}_{\{|X_n|\leq n\}}}_{\to 1 \ a.s.}) \xrightarrow{*} EX_1 = \mu$$

* dominatedly by $|X_1|$ which is integrable
(3)

$$\sum_{n=1}^{\infty} \frac{EY_n^2}{n^2} = \sum_{n=1}^{\infty} \frac{E(X_n^2 \mathbb{1}_{\{|X_n|\leq n\}})}{n^2} = \sum_{n=1}^{\infty} \frac{E(X_1^2 \mathbb{1}_{\{|X_n|\leq n\}})}{n^2}$$

$$= E\left(X_1^2 \sum_{n=1}^{\infty} \frac{\mathbb{1}_{\{|X_1|\leq n\}}}{n^2}\right) = E\left(X_1^2 \sum_{n\geq |X_1|} \frac{1}{n^2}\right)$$

$$\boxed{\sum_{n\geq m} \frac{1}{n^2} \leq \sum_{n\geq m} \frac{2}{n(n+1)} = 2(\frac{1}{m} - \frac{1}{m+1}) + 2(\frac{1}{m+1} - \frac{1}{m+2}) + \cdots = \frac{2}{m}}$$

$$\leq E\left(X_1^2 \cdot \frac{2}{|X_1|}\right) = 2E|X_1| < \infty$$

$$\sum_{n=1}^{\infty} \frac{(EY_n)^2}{n^2} \leq \sum_{n=1}^{\infty} \frac{c^2}{n^2} < \infty$$

$$\sum_{n=1}^{\infty} \frac{var(Y_n)}{n^2} = \sum_{n=1}^{\infty} \frac{EY_n^2}{n^2} - \sum_{n=1}^{\infty} \frac{(EY_n)^2}{n^2} < \infty$$

□

Theorem 4.13 (Strong Law of Large Numbers, SLLN)**.** *Let (X_n) be iid r.v's with $E|X_k| < \infty \ \forall k$ and mean μ then*

$$\frac{X_1 + \cdots + X_n}{n} \to \mu \qquad a.s.$$

Proof. Let $Y_n = X_n \mathbb{1}_{\{|X_n|\leq n\}}$ by (1) of the Truncation Lemma, 4.12

$$\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = \lim_{n\to\infty} \frac{Y_1 + \cdots + Y_n}{n}$$

It suffices to prove that

$$\frac{Y_1 + \cdots + Y_n}{n} \to \mu$$

www.mymathscloud.com
My Maths Cloud

$$\frac{Y_1+\cdots+Y_n}{n}=\frac{(Y_1-EY_1)+\cdots+(Y_n-EY_n)}{n}+\underbrace{\frac{EY_1+\cdots+EY_n}{n}}_{\to\mu \text{ by } TL +;Cesaro}$$

It suffices to prove:

$$\frac{(Y_1-EY_1)+\cdots+(Y_n-EY_n)}{n}\to 0$$

$$\sum_{n=1}^{\infty}E\left(\frac{Y_n-EY_n}{n}\right)^2=\sum_{n=1}^{\infty}\frac{var(Y_n)}{n^2}\overset{TL(3)}{<}\infty$$

$$\overset{4.8}{\Rightarrow}\sum_{n=1}^{\infty}E\frac{Y_n-EY_n}{n}<\infty \qquad a.s.$$

$$\overset{4.10}{\Rightarrow}\frac{(Y_1-EY_1)+\cdots+(Y_n-EY_n)}{n}\to 0$$

□

4.5 Convergence of Martingales

Definition 4.8. *A sequence (C_n) of r.v's is call previsable if*

$$C_n\in\mathcal{F}_{n-1}\forall n\geq 1$$

Let (X_n) be a martingale sequence and (C_n) be a previsable sequence

$$\begin{cases}M_n=\sum_{i=1}^n C_i(X_i-X_{i-1}) & n\geq 1\\ M_0=0\end{cases}$$

Is called the martingale transform of (X_n) via (C_n)

Theorem 4.14. *If (C_n) are uniformly bounded then (M_n) is a martingale*

Proof. Integrability and M_n is $\mathcal{F}_n$-measurable (obvious)

$$E[M_n|\mathcal{F}_{n-1}]=E[\sum_{i=1}^n C_i(X_i-X_{i-1})|\mathcal{F}_{n-1}]=M_{n-1}+C_nE[X_n-X_{n-1}|\mathcal{F}_{n-1}]=M_{n-1}$$

□

Definition 4.9 (Upcrossing)**.** *Let (X_n) be a martingale and $a<b$. The number of upcrossings up to time N, $U_N[a,b]$, is the largest number such that there are*

$$0\geq s_1<t_1<\cdots<s_k<t_k\leq N$$

Such that

$$X_{s_1},\ldots,X_{s_k}<a$$

$$X_{t_1},\ldots,X_{t_k}>b$$

www.mymathscloud.com
My Maths Cloud

Lemma 4.15 (Doob's Upcrossing Lemma)**.**

$$(b-a)EU_N[a,b] \leq E[(X_N-a)^-]$$

Proof.

$$\begin{array}{lll} S_1 = \min\{n : X_n \leq a\} & \dots & S_n = \min\{n \geq T_{n-1} : X_n \leq a\} \\ T_1 = \min\{n > S_1 : X_n \geq b\} & \dots & T_n = \min\{n \geq S_n : X_n \leq b\} \end{array}$$

$C_n = 0$ on $[0, S_1] \cap \bar{\mathbb{N}}$ and all $[T_k + 1, S_{k+1}] \cap \bar{\mathbb{N}}$. $C_n = 1$ on $[S_k + 1, T_k] \cap \bar{\mathbb{N}}$.

$$\begin{cases} Y_n = \sum_{i=1}^n C_i(X_i - X_{i-1}) & n \geq 1 \\ Y_0 = 0 \end{cases}$$

(Y_n) is a martingale.

$$0 = EY_0 = EY_n \leq (b-a)U_N[a,b](+0) - E(X_n - a)^-$$

$$\Rightarrow (b-a)U_N[a,b] \leq E(X_n-a)^-$$

□

Theorem 4.16 (Martingale Convergence Theorem)**.** *If (X_n) is a c-bounded martingale ie if*

$$\exists c > 0 \textit{ such that } E|X_n| \leq c \; \forall n$$

then it converges a.s.

Proof. Assume

$$\{\liminf_{n\to} X_n < \limsup X_n\} = \{\exists a,b \in \mathbb{Q} \; s.t. \; U_\infty[a,b] = \infty\} = \bigcup_{\substack{a<b \\ a,b,\in\mathbb{Q}}} \{U_\infty[a,b] = \infty\}$$

$$(b-a)U_\infty[a,b] \overset{MCT}{=} (b-a) \lim_{n\to\infty} U_n[a,b] \overset{4.15}{\leq} \lim_{n\to\infty} \sup E[(X_N-a)^-] < \infty$$

since $E[(X_N-a)^-] < E|X_n| + a \leq c + a$

$$EU_\infty[a,b] < \infty \Rightarrow U_\infty[a,b] < \infty \quad a.s.$$

$$\Rightarrow P(U_\infty[a,b] = \infty) = 0$$

$$\Rightarrow P(\bigcup_{\substack{a<b \\ a,b,\in\mathbb{Q}}} \{U_\infty[a,b] = \infty\}) = 0$$

$$\Rightarrow \liminf X_n = \limsup X_n \quad a.s.$$

$$X = \liminf X_n$$

$$EX = E \liminf |X_n| \overset{Fatou}{\leq} \liminf \underbrace{E|X_n|}_{=c} < \infty$$

$$\Rightarrow X < \infty \quad a.s.$$

□

www.mymathscloud.com
My Maths Cloud

Example 4.12 (Polya Urn).

$$X_n \in [0,1] \Rightarrow \ Converges$$

Example 4.13. *Any non-negitive martingale converges a.s.!*

$$E|X_n| = EX_n = EX_0 \quad is\ in\ L^1\ -\ holds$$

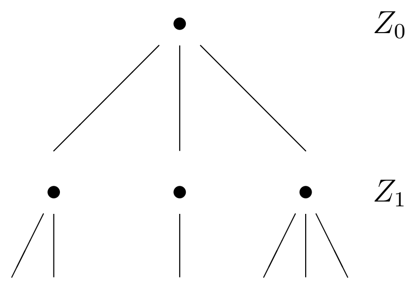
Example 4.14 ((S_n)- Random Walk). *$S_n \not\to$ any finite number. Why $S_n \not\to \pm\infty$?*

$$P(hit\ a\ before\ -b) = \frac{b}{a+b}$$

as $-b$ goes $-\infty$

$$P(hit\ a\ eventually) = 1$$

Example 4.15 (Galton Watson Tree). *Lets fix a random variable with values in $\{1, 2, \dots\}$ which chooses the number of edges decending from each vertex.*



$$Z_n = \ number\ of\ veticies\ in\ generation\ n$$

If

$$Q = EZ_1 = 1$$

Then Z_n is a non-negitive martingale

$$\Rightarrow Z_n\ converges\ to\ zero$$

$$Z_n \to k$$

$$\Rightarrow Z_n = k \quad eventually\ for\ some\ k$$

$$P\left(\exists N\ \forall n > N\ Z_n = k\right) = P\left(\cup_N \{\forall n > N\ Z_n = k\}\right) = 0$$

$$\{\forall n > N\ Z_n = k\} = \lim_{m\to\infty} P(Z_N = k)P(k)^m = 0 \quad P(k) < 1 \quad \forall k \neq 0$$

$$\Rightarrow Z_n \Rightarrow 0\ a.s.$$