Probability
3105

Based on lectures by
Dr N Sidorova

Jan 2012

Contents

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*

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

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 $\frac{1}{2}$*

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

$$
\begin{aligned}
X : & H \to 1 \\
& T \to 0
\end{aligned}
$$

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

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

$$
\begin{aligned}
&\Omega = \{1, 2, 3, 4, 5, 6\} \\
&\Sigma = 2^{\Omega} \quad \textit{(σ-algebra of all subsets)} \\
&P(\{1\}) = \cdots = P(\{6\}) = \frac{1}{6} \\
&P(\{1, 3, 5\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2} \; etc \ldots
\end{aligned}
$$

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

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

$$
\begin{aligned}
&\Omega = \{1, 2, 3, 4, 5, 6, 0\} \\
&\Sigma = 2^{\Omega} \quad \textit{(σ-algebra of all subsets)} \\
&P(\{1\}) = \cdots = P(\{6\}) = \frac{1}{6} \quad P(\{0\}) = 0
\end{aligned}
$$

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

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

$$
\begin{aligned}
&\Omega = [0, 1] \\
&\Sigma = ?
\end{aligned}
$$

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?*
$[0,1] \to \mathbb{R}$
$0.\omega_1\omega_2 \dots \mapsto \omega_1 + \cdots + \omega_n$

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*:*

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):
$t_1 < t_2$

$$F_X(t_1) = P(x \le t_1)$$
$$F_X(t_2) = P(x \le t_2).$$

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

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

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

Let $t_n \searrow \infty$

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

(3):
Let $t_n \searrow t$

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

□

Theorem 1.1 (Skorokhod Representation)**.** *If $F : \mathbb{R} \to [0,1]$ satisfies (1)-(3) from def* **??** *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 $X(\omega) = G(\omega)$*

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

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

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

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

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

It suffices to show:

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

(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:

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

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

$\square$

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 ; $([0,1], \lfloor, leb)$

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

$0.\omega_1\omega_2\omega_3\ldots$ - *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*

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

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

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

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 $F_x = \int_{-\infty}^{t} f_X(u)du$ 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:

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

(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.

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

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

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 $EX = \infty$

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*

$$Eh(x) = \int h(x) d\mu_X(x)$$

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

Proof.

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

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

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

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

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

$\square$

Corollary 1.1. *if X has density $h(x)$ then*

$$EX = \int h(x) f(x) dx$$

in particular if $h(x) = x$

$$EX = \int x f(x) dx \quad \textit{(old formula)}$$

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

$$EX^2 = \int x^2 f(x) dx$$

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*

$$EX = 1 \cdot P(x = 1) + 0 \cdot P(x = 0) = \frac{1}{2}$$
$$EX^2 = 1^2 \cdot P(x = 1) + 0 \cdot P(x = 0) = \frac{1}{2}$$
$$VarX = EX^2 - (EX)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

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?

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

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

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

$$P\left(\cap - i = 1^n A_i\right) = \prod_{i=1}^{n} P(A_i) \text{ whenever } A_i \in \Sigma_i \ 1 \geq i \leq n$$

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".

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

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

$$
\begin{array}{ccc}
 & 0 & 1 \\
1 & . & . \\
2 & . & . \\
3 & . & . \\
4 & . & . \\
5 & . & . \\
6 & . & .
\end{array}
\quad
\begin{array}{l}
\xrightarrow{X} \{0,1\} \\
\xrightarrow[Y]{} \{1,2,3,4,5,6\}
\end{array}
$$

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

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

$$
\sigma(y) = \{\phi, \Sigma, \{y=1\}, \dots \{y=6\}, \{y = 3 \,or\, 4 \,or\, 6\} \; etc \dots\}
$$

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

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.

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

Example 2.3.

$$
\left.\begin{array}{l}
\{\{x < t\},\ t \in \mathbb{R}\} \\
\{\{x \le t\},\ t \in \mathbb{R}\}
\end{array}\right\} \quad \pi\textit{-system generating } \sigma(x)
$$

because

$$
\{x < t\} \cap \{x < s\} = \{x < \min(t,s)\} \; etc
$$

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

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

Theorem 2.1. *Let* (Ω, Σ) *be a set with* σ*-algebra and mathcalI 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* $\sigma(y)$*.*

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

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 of (a). fix $I \in \mathcal{I}$

$$\begin{aligned} \mu_1(B) &= P(I \cap B) \quad on\ \sigma(Y) \\ \mu_2(B) &= P(I)P(B) \end{aligned}$$

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

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

□

Proof of (b). fix $B \in \sigma(Y)$. Define

$$\begin{aligned} v_1(A) &= P(A \cap B) \quad \textit{on } \sigma(X) \\ v_2(A) &= P(A)P(B) \end{aligned}$$

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

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

□

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

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

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

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

Uniform + Independent

2.1 Finite and infinite occurance of Events

Example 2.6.

$$\begin{array}{ccccc} X_1, & X_2 & X_3 & X_4 & \dots \quad Bernulli \\ 1, & 0, & 1, & 1, & \dots \end{array}$$

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

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

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

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

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

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*

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

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

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

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

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

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

Proven later.

2.2 The Borel-Cantelli Lemmas

Lemma 2.4 (Borel-Cantelli Lemma 1 - BC1)**.** *Let* (E_n) *be events such that* $\sum_{i=1}^{\infty} P(E_n)<\infty$, *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 $\sum_{i=1}^{\infty}P(E_n)<\infty$. □

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

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

$$P(\bigcap_{n\geq N}E_n^c)=\lim_{k\to\infty}P(\bigcap_{n=N}^{k}E_n^c)=\lim_{k\to\infty}\prod_{n=N}^{k}\underbrace{P(E_n^c)}_{1-P(E_n)}$$
$$\leq\lim_{k\to\infty}\prod_{n=N}^{k}e^{-P(E_n)}=\lim_{k\to\infty}e^{-\sum_{n=N}^{k}P(E_n)}=0\ \forall N$$

□

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

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

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

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

$\limsup a_n = a$ 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

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

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

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

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

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

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

(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.

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

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 $E|XY| < \infty$

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

(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. *$n \in \mathbb{N},\ X_1, \ldots, X_n -$ independent.*

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

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

Y has Binomial distribution

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

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

$$\vdots \qquad \vdots$$

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

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

2.3 Bernsteins inequality

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

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

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

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

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

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

now

$$\begin{aligned}P\left(\lambda\sum_{i=1}^{n}a_iX_i>\lambda t\right)&=P\left(\exp\left(\lambda\sum_{i=1}^{n}a_iX_i\right)>\exp(\lambda t)\right)\\&\overset{Chebyshev}{\le}E\left(\exp\left(\lambda\sum_{i=1}^{n}a_iX_i\right)\cdot\exp(-\lambda t)\right)\\&\le\exp\left(\frac{\lambda^2c}{2}-\lambda t\right)\end{aligned}$$

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

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

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

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

□

Theorem 2.8 (SLLN for r.v's taking values ± 1)**.** *Let $X_1, \ldots, X_n$ 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:

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

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

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

□

2.4 Joint Laws

Recall.

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

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

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

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

Defined by

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

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

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

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

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

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

□

Proof of (c).

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

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

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

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

□

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

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

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*

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

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

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

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

Example 2.11.

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.

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

Consider

$$X_1 = \left\{1, -1 \textit{ each with } P = \frac{1}{2}\right\} \quad X_i = 0, \; \forall i \geq 2$$

we want to show

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

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

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

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

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 $\bigcup_{i=1}^n \sigma_i$ 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*

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

2. $\frac{|X_n|}{n}$ *if* $E|X_1| < \infty$ *then by SLLN*

$$\frac{|X_n|}{n} \to 0$$

if $E|X_1| = \infty$

$$\frac{|X_n|}{n} \text{ diverges.}$$

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

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

$$X \overset{in}{\underset{Law}{=}} Y \quad since \quad \mu_X = \mu_Y$$

$$\frac{X_1 + \cdots X_n}{n} \overset{SLLN}{\to} \mu \quad a.s.$$

$$\frac{X_1 + \cdots X_n - n\mu}{n\sigma} \overset{in}{\underset{Law}{=}} N(1, 0) \quad (CLT)$$

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*

$$F_{X_n}(t) \Rightarrow F_X(t) \quad \forall t \in \mathbb{R}$$

where F_X *is continuous.*

$$\Leftrightarrow \mu_{X_n} \to \mu_X \quad weakly$$

Notation. *to say that $X_n \to X$ in law, in distribution or weakly:*

$$\underbrace{\xrightarrow{d}}_{\text{we will use this}} \qquad \underbrace{\xrightarrow{w}, \Rightarrow}_{\text{some people use these}}$$

Remark. *why do we exclude discontinuity points?*

$$X_n = \frac{1}{n} \textit{ with } P = 1$$

Need to check someone elses notes cant read mine!

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*

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

and $X_n \to X$ a.s.

Theorem 3.2. *[Useful definition of weak convergence]*

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

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

Idea.

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

$\Leftarrow$ *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)

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.

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

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

$$\begin{aligned}&\Rightarrow \exists N\ \forall n \geq N \quad F_n(X(\omega) - \delta) < \omega < F_n(X(\omega) + \delta)\\ &\Rightarrow X(\omega) - \delta < X_n(\omega) < X(\omega) + \delta\end{aligned}$$

□

Proof of (⇒). we are given □

Proof of (⇐). we are given □

Proof of (1). we are given □

□