* Chapter 1 : Measure theory.

* 1.1) Probability spaces.

* Def:

A probability space is a triple (Ω, F, P)

where
- Ω : nonempty set = set of "outcomes"
- F : a σ-algebra of subsets of Ω = set of "events".
- $P: F \to [0,1]$ is a function that assigns probability to "events".

* F is a σ-algebra, iff F is a nonempty collection of subsets $A \subset \Omega$ s.t.

i) $A \in F$ then $A^c \in F$ $\quad \Omega \in F$

ii) If $\underbrace{A_1, A_2, \dots}_{\text{finite / countable}} \in F \Rightarrow \bigcup_{i=1}^{\infty} A_i \in F$ (closed under countable union) (also countable intersection).

* Properties of F
- $\phi, \Omega \in F$ since Ω is a subset of Ω and $\phi = \Omega^c \in F$
- If $A_1, A_2, \dots \in F$ then $\bigcap_{i=1}^{\infty} A_i \in F$ • $A, B \in F \Rightarrow B \setminus A \in F$.

* $A \in F$ then A is called F-measurable.

(Ω, F) is a measurable space, is a space that on which we can put a meas

* A measure is a nonnegative countably additive set function

$\mu : F \to R^+$ is a measure if

i) $\mu(A) \geq 0 \ \forall A \in F \quad \mu(\phi) = 0$

ii) μ is countable additive on F (include $+\infty = +\infty$

If $A_1, A_2, \dots$ are pairwise disjoint on F then $\mu(\underbrace{\bigcup_{i=1}^{\infty} A_i}_{\in F}) = \underbrace{\sum_{i=1}^{\infty} \mu(A_i)}_{\text{could be } +\infty \text{ (see example)}}$

- When $\mu(\Omega) = 1$ then μ is a probability measure.

Example

* Take Lebesgue measure μ on $(\mathbb{R}, F)$ Borel subset of $\mathbb{R} = \Omega$

Let $A_i = [i, i+1)$ then $\{A_i\}$ pairwise disjoint.

Then $\bigcup_{i=1}^{\infty} A_i = [1, +\infty)$ case $+\infty = +\infty$

And we have $+\infty = \mu([1, +\infty)) \ \mu(\bigcup_{i=1}^{\infty} A_i) \overset{\uparrow}{=} \sum_{i=1}^{\infty} \mu(A_i)$

Borel algebra on Ω i[s] the smallest σ-alge[bra] containing all open se[ts]

* Remark:

If $A, B \in \mathcal{F}$, then $\mu(A\cup B) = \underbrace{\mu(A) + \mu(B) - \mu(A\cap B)}_{\text{not well defined when } \mu(A\cap B) = +\infty}$.

However $\mu(A\cup B) + \mu(A\cap B) = \mu(A) + \mu(B)$ is well defined (includes $+\infty = +$

* Theorem 1.1.1.

Let μ is a measure on $(\Omega, \mathcal{F})$, then we have

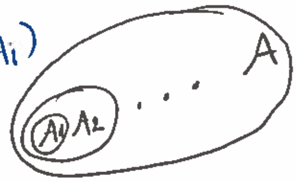
i) μ is monotone: If $A \subset B \Rightarrow \mu(A) \leq \mu(B)$ and $\mu(B\setminus A) = \mu(B) - \mu(A)$

ii) μ is subadditivity: $A \subset \bigcup_{i=1}^{\infty} A_i \Rightarrow \mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i)$ (not neccessary disjoint

$\left.\begin{array}{l} \text{If } \underbrace{A_1, A_2, \ldots \in \mathcal{F}}_{\text{not neccessary disjoint}} \\ A \subset \bigcup_{i=1}^{\infty} A_i \end{array}\right\}$ then $\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i)$

$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i)$

not neccessary disjoint.



iii) μ is continuous from below

$\left.\text{If}\begin{cases} A_1, A_2, \ldots \in \mathcal{F} \\ A_i \uparrow A \ (\text{i.e. } A_1 \subset A_2 \subset \cdots \subset A,\ \bigcup_{i=1}^{\infty} A_i = A) \end{cases}\right\} \Rightarrow \mu(A_i) \longrightarrow \mu(A)$

$\lim_{i\to\infty} \mu(A_i) = \mu(A)$

iv) μ is continuous from above

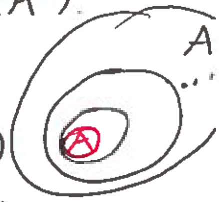
If $A_1, A_2, \ldots \in \mathcal{F}$

$A_i \downarrow A$ (ie $A_1 \supset A_2 \supset \cdots \supset A$, $\bigcap_{i=1}^{\infty} A_i = A$

$\mu(A_1) < +\infty$

$\Bigg\}$ $\mu(A_i) \xrightarrow{i\to\infty} \mu(A)$

$\lim_{i\to\infty} \mu(A_i) = \mu(A)$.



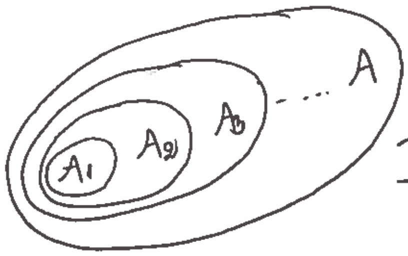
* Proof i)

We have $B = A \cup (B\setminus A) \xRightarrow{\text{disjoint}} \mu(B) = \mu(A) + \mu(B\setminus A)$

$\Rightarrow \mu(A) \leq \mu(B)$ and $\mu(B\setminus A) = \mu(B) - \mu(A)$.

• Proof ii)

* Proof iii) Proof that if $A_1, A_2, \ldots \in \mathcal{F}$, $A_1 \subset A_2 \subset \cdots \subset A$, $A = \bigcup_{i=1}^{\infty} A_i$ $\Big\} \Rightarrow$ then $\mu(A_i) \longrightarrow \mu(A)$.



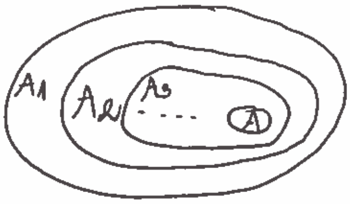
Put $B_1 = A_1$
$B_2 = A_2 \setminus A_2$
$B_3 = A_3 \setminus A_2$
$\vdots$
$B_i = A_i \setminus A_{i-1}$

then we have $\{B_i\}$ disjoint and $\bigcup_{i=1}^{\infty} B_i = A$
$\Rightarrow \mu(\bigcup_{i=1}^{\infty} B_i) = \mu(A)$.

which implies $\mu(A) = \mu(\bigcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} \mu(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(B_i) = \lim_{n\to\infty} A_n$

important step.

* Proof (iv) from (iii). Proof if $A_1, A_2, \ldots \in \mathcal{F}$, $A_i \downarrow A$ (i.e. $A_1 \supset A_2 \supset \cdots \supset A$, $\bigcap_{i=1}^{\infty} A_i = A$), $\mu(A_1) < +\infty$ $\Big\} \Rightarrow \mu(A_n) \to$



Idea: since we want to use iii) we want to construct an "increasing" sequence of sets.
the biggest set is A_1, $A_1 \setminus A_1 = \phi$
$A_1 \setminus A_2 \supset A_1 \setminus A_1$ ---

we have $(A_1 \setminus A_i) \uparrow (A_1 \setminus A)$
by (3) $\mu(A_1 \setminus A_i) \uparrow \mu(A_1 \setminus A)$ as $i \to +\infty$ $\Big\} \Rightarrow \mu(A_1) - \mu(A_i) \longrightarrow \mu(A_1) - \mu$

- If $\mu(A_1) < +\infty$ then $LHS = \mu(A_1) - \mu(A_i)$
  $RHS = \mu(A_1) - \mu(A)$

or

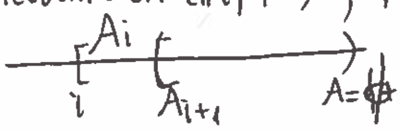
or $\lim_{i\to\infty} (\mu(A_1) - \mu(A_i)) = \mu(A_1) - \mu(A)$

$\Rightarrow \underbrace{\mu(A_1)}_{<+\infty} - \lim(\mu(A_i)) = \underbrace{\mu(A_1)}_{<+\infty} - \mu(A) \Rightarrow \lim_{i\to+\infty} \mu(A_i) = \mu(A)$ (iv) is done □

* Remark:
- A probability measure is continuous from above
- An example of a measure that is not continuous from above

Consider Lebesgue measure on $(\mathbb{R}, \mathcal{F})$, $\mathcal{F}$: Borel set.

Let $A_i = [i, +\infty)$



then $A_i \downarrow \phi$ $\quad \mu(A_i) \not\longrightarrow \mu(\phi)$
and $\mu(A_i) = +\infty$, $\mu(\phi) = 0$ $\Rightarrow$ not continuous from above

* Probability measure:

$\mathbb{P}: \mathcal{F} \longrightarrow [0,1]$ is a probability measure if

i) $\mathbb{P}(A) \geq 0, \forall A \in \Omega$

ii) $\mathbb{P}$ is countable additive on $\mathcal{F}$

If $A_1, A_2, \dots$ pairwise disjoint elements of $\mathcal{F}$, then $\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$

iii) $\mathbb{P}(\Omega) = 1$

* Example (Discrete probability space)

Let Ω = countable set (finite / countable infinite)

$= \mathbb{N}$ (for example)

$\mathcal{F}$ = set of all subsets of Ω

Ω	w_1	w_2	w_3	----
p	p_1	p_2	p_3	----

Then for all $w \in \Omega$, $\exists\ p(w)$ then $\forall A \in \mathcal{F}$, $\boxed{\mathbb{P}(A) = \sum_{w \in A} p(w)}$

This is the discrete probability space.

- A special case is when Ω is a finite set, $\Omega = \{1, 2, \dots, n\}$

$$p_i = \begin{cases} \frac{1}{n}, & i = \overline{1,n} \\ 0, & i > n \end{cases}$$

* Example: $\Omega = [0,1]$, $\mathcal{F}$ Borel subsets of $[0,1]$, $\mathbb{P}$: Lebesgue measure.

* Lemma: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space

Let $A, B \in \mathcal{F}$

1) If $A \subset B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$ $\qquad$ $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$

2) If $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ $\qquad$ $\mathbb{P}(A \cup B) \overset{\text{disjoint}}{=\!=} \mathbb{P}(A) + \mathbb{P}(B)$.

* Stieltjes measure function — Lebesgue measure

* The idea explain why there exists σ-algebra generated by A.

- If $\mathcal{F}_i$, $i \in I$ then $\bigcap_{i\in I} \mathcal{F}_i$ is a σ algebra. (Exercise 1.1.1 any intersection of σ algebra is a σ algebra)
  $I \neq \phi$ is an arbitrary index set

- Def: Borel σ-algebra

Let $\mathcal{A}$ be any collection of subsets of a set Ω then the σ-algebra generated by A, denoted σ(A) is the smallest σ algebra contain A

* Def (example 1.1.2) d: Euclidian space

Let $(\mathbb{R}, \mathcal{B})$ Borel measure

Define Stieltjes measure function: $F: (\mathbb{R}, \mathcal{B}) \to \mathbb{R}$ is a measure such that

i) F is nondecreasing $x \leq y \Rightarrow F(x) \leq F(y)$
   (which implies F has left limit $F(y^-) = \lim_{x \uparrow y} F(x)$)

ii) F is right continuous, i.e $\lim_{y \downarrow x} F(y) = F(x)$.

$F(\infty) = \lim_{x\uparrow\infty} F(x)$

$F(-\infty) = \lim_{x \downarrow -\infty} F(x)$

possibly $\pm\infty$

* Theorem 1.1.2. Let $F: \mathbb{R} \to \mathbb{R}$.

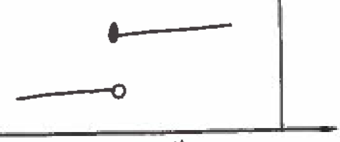
Associated with each Stieltjes measure, there exists a unique nonnegative measure μ on $(\mathbb{R}, \mathcal{B})$ such that:

for $a < b$, $a, b \in \mathbb{R}$, $\mu((a, b]) = F(b) - F(a)$

the choice of "closed on the right", $(a, b] \Rightarrow$ $b_n \downarrow b$ then $\bigcap (a, b_n] = (a, b]$

- When $F(x) = x$, Lebesgue measure $\mu([a, b]) = b - a$

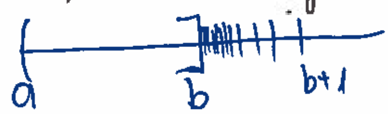
- Example of Stieltjes measure: $F(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$



- If $F(\infty) = 1$, $F(-\infty) = 0$ then μ is a probability measure

* Idea of theorem 1.1.2: Explain the reason of "closed on the right"

- F is increasing and right continuous on $\mathbb{R}$.



- Why do we require right continuity?

Consider $(a, b + \frac{1}{n}]$, $n = 1, 2, \dots$ Then $(a, b+\frac{1}{n}] \xrightarrow{\downarrow} (a, b]$ as $n \to \infty$

$\Rightarrow \mu(a, b+\frac{1}{n}] \longrightarrow \mu(a, b] \Rightarrow F(b+\frac{1}{n}) - F(a) \xrightarrow{n\to\infty} F(b) - F(a)$ $\Rightarrow F(b+\frac{1}{n}) \to$ right contin

* Some definitions to explain the choice of "open on the left" of theorem 1.1.2

Semialgebra

A nonempty collection $\mathcal{S}$ of subsets of Ω is a semialgebra if i) $\mathcal{S}$ is closed under finite intersection

If $A_1, \dots, A_n \in \mathcal{S}$ then $\bigcap_{i=1}^{n} A_i \in \mathcal{S}$

ii) $A \in \mathcal{S}$ then $A = \bigcap_{i=1}^{n} A_i$ for some $A_1, \dots, A_n \in \mathcal{S}$

= finite intersection of elements of $\mathcal{S}$

* Example 1.1.3 (an important example)

$\mathcal{S}$ = all intervals in $\mathbb{R}$ of the form $(a, b]$, where $-\infty \leq a < b \leq +\infty$ $\Rightarrow$ $\mathcal{S}$ is an semialgebra

Algebra

2) A nonempty collection S of subsets of Ω is an algebra if i) $\mathcal{S}$ is closed under finite union $A_1, \dots, A_n \in \mathcal{S} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{S}$

ii) $\mathcal{S}$ is closed under taking complements $A \in \mathcal{S} \Rightarrow A^C \in \mathcal{S}$.

$\Rightarrow$ ~~$\mathcal{S}$ is closed under finite intersection~~ $(A \cap B) = (A^c \cup B^c)^c \in \mathcal{S}$

* An algebra $\Longrightarrow$ an semialgebra

$\not\Longleftarrow$ the above example is a semialgebra but algebra

* The algebra generated by a collection of set S is

$\mathcal{A}(S) = \bigcap_I A$, I : collection of all algebra containing S.

Measure on ~~the~~ algebra

* Def (a measure on an algebra)

Let $\mathcal{A}$ is an algebra, $\mu : \mathcal{A} \longrightarrow [0, \infty]$ is a measure on the algebra $\mathcal{A}$ iff.

i) $\mu(A) \geq 0 = \mu(\phi)$, $\forall A \in \mathcal{A}$

ii) If $A_1, A_2, \dots \in \mathcal{A}$ disjoint, $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ (since we don't always have this with algebra) $\Big\} \Rightarrow \mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$

* Def (σ finite measure on an algebra).

A measure μ on an algebra $\mathcal{A}$ is σ-finite iff

$\exists A_1, A_2 \dots \in \mathcal{A}$ such that $A_n \uparrow_{n \to \infty} \Omega$ such that $\mu(A_n) < \infty, \forall n$.

$(\Leftrightarrow \Omega = \bigcup_{n=1}^{\infty} A_n$

- example ($\mathbb{R}$, Lebesgue measure) $A_n = [-n, n]$ $\Rightarrow$ $A_n \uparrow \mathbb{R}$ and $\mu(A_n) < +\infty$

$\Rightarrow$ Lebesgue measure is σ-finite.

* Lemma 1.1.3.

Let $\mathcal{S}$ be a semialgebra

$\overline{\mathcal{S}} = \{$ finite disjoint unions of sets in $\mathcal{S}\} = \{A = \bigcup_{i=1}^{n} A_i \,;\, A_i \in \mathcal{S} \,;\, A_i, i=\overline{1,n}$ pairwise disjoint $\}$

Then $\overline{\mathcal{S}}$ is called the algebra generated by $\mathcal{S}$, $\overline{\mathcal{S}} = a(\mathcal{S})$

$\overline{\mathcal{S}}$ is an algebra.

* Proof: we need to prove that $\overline{\mathcal{S}}$ is an algebra, NTP

NTP: $\begin{cases} \text{closed under taking complement } A \in \overline{\mathcal{S}} \Rightarrow A^c \in \overline{\mathcal{S}} & (1) \\ \text{closed under finite unions.} & (2) \end{cases}$

- Proof (1).

Let $A = \bigcup_{i=1}^{n} A_i \,;\, A_i \in \mathcal{S} \,;\, A_i, i=\overline{1,n}$ disjoint, NTP $A^c \in \overline{\mathcal{S}}$

We have $A^c = \Omega \setminus \left(\bigcup_{i=1}^{n} A_i\right) = \bigcap_{i=1}^{n} A_i \in \mathcal{S} \subset \overline{\mathcal{S}} \Rightarrow A^c \in \overline{\mathcal{S}}$

- Proof (2) NTP for $A_1, A_2, \dots, A_n \in \overline{\mathcal{S}}$ then $\bigcup_{i=1}^{n} A_i \in \overline{\mathcal{S}}$

Each $A_i = \bigcup_{j=1}^{n_i} B_{ij}$ where $B_{ij} \in \mathcal{S}$ and $B_{i1}, \dots, B_{i,n_i}$ are pairwise disjoint

then $\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} \bigcup_{j=1}^{n_j} B_{ij}$

$= \bigcup_{i=1}^{N} C_i$ where $C_i \in \mathcal{S}$ and $C_i, i=\overline{1,N}$ are pairwise disjoint $\square$.

* Def: Given μ: measure defined on semialgebra $\mathcal{S}$.

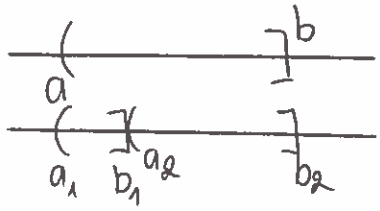
Define $\overline{\mu}$: defined on the algebra $\overline{\mathcal{S}}$ by

Since $A \in \overline{\mathcal{S}}$, $A = \bigcup_{i=1}^{n} A_i$ for $A_i \in \mathcal{S}$ (p. disjoint) $\Rightarrow \overline{\mu}(A) = \overline{\mu}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i)$ (p. disjoint)

- Problem: Is $\overline{\mu}$ well defined?

Let $A_1, \dots, A_n \in \mathcal{S}$ pairwise disjoint
$B_1, \dots, B_m \in \mathcal{S}$ pairwise disjoint
$\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{m} B_m$
$\overset{?}{\longrightarrow}$ Do we have $\overline{\mu}\left(\bigcup_{i=1}^{n} A_i\right) = \mu\left(\bigcup_{i=1}^{m} B_i\right)$

- In case we consider $(a,b]$ in $(\mathbb{R}, \mathcal{B})$



Let $(a,b] = (a_1, b_1] \cup (a_2, b_2]$

$\Rightarrow \mu(a,b] = \mu(a,b_1] + \mu(a_2, b_2]$

$\mu(a,b] = F(b) - F(a)$, $\quad \mu(a,b_1] + \mu(a_2,b_2] = F(\cancel{b_1}) - F(a) + F(b_2) - F(\cancel{a_2}) = F(b_2) - F(a) = F(b) - F(a)$.

* Lemma 1.1.5 :

Let $\mathcal{S}$: be a semialgebra — not always true for semialg.

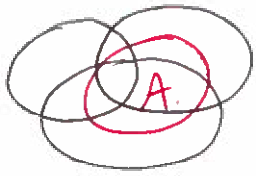
$\mu : \mathcal{S} \longrightarrow [0, +\infty]$, $\mu(\phi) = 0$,

μ is finite additive on $\mathcal{S}$ ($A_1, .., A_n$ pairwise disjoint, and $\bigcup_{i=1}^{n} A_i \in \mathcal{S}$) $\Rightarrow \mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n}$

a) If $A, \{A_i\}_{i=1}^{n} \in \bar{\mathcal{S}}$, $A = \sum_{i=1}^{n} A_i$ $\Big\} \Rightarrow \bar{\mu}(A) = \sum_{i=1}^{n} \bar{\mu}(A_i)$ $\bar{\mu}$ is finitely additive on $\bar{\mathcal{S}}$

b) If $A, \{A_i\}_{i=1}^{n} \in \bar{\mathcal{S}}$, $A \subset \bigcup_{i=1}^{n} A_i$ $\Big\} \Rightarrow \bar{\mu}(A) \leq \sum_{i=1}^{n} \bar{\mu}(A_i)$

not necessary disjoint



$A_1, A_2, A_3 \in \bar{\mathcal{S}}$
$A \subset \bigcup_{i=1}^{3} A_i$

Proof: see text : idea : construct disjoint $B_1, .., B_n$ which are pairwise disjoint

$A = \bigcup_{i=1}^{n} B_i$

* Idea: we have μ on $\mathcal{S}$ semialgebra $\Rightarrow$ want to go to algebra.
$\bar{\mu}$ on $\bar{\mathcal{S}} = \mathcal{A}(\mathcal{S})$ algebra

Theorem 1.1.4

Let $\mathcal{S}$ be a semialgebra,

$\mu: \mathcal{S} \longrightarrow [0, +\infty]$, $\mu(\phi) = 0$

Assume: 1) μ is finite additive on $\mathcal{S}$ (Let $A_1, \dots, A_n \in \mathcal{S}$, pairwise disjoint, $\bigcup_{i=1}^{n} A_i \in \mathcal{S}$ then $\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu$

2) If $A_1, A_2, \dots \in \mathcal{S}$ are pairwise disjoint and $\bigcup_{i=1}^{\infty} A_i \in \mathcal{S}$, then, $\mu(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$ — need to check $\leq$

then $\exists !$ extension $\bar{\mu}$ on $\mathcal{A}(\mathcal{S})$ which is a measure on $\mathcal{A}(\mathcal{S})$

- If $\bar{\mu}$ is σ-finite, then $\exists$ a unique extension ν which is a measure on $\sigma(\mathcal{S})$

$\nu(A) = \mu(A)$ for $A \in \mathcal{S}$

* Remind theorem 1.1.2: Let $F: \mathbb{R} \longrightarrow \mathbb{R}$ is a Stieltjes measure then

$\exists ! \mu$ on $(\mathbb{R}, \mathcal{B})$ so that $\mu(a, b] = F(b) - F(a)$

* Note that if $F(\infty) = 1$ and $F(-\infty) = 0$, μ is a probability measure.

since $\mu(\mathbb{R}) = \lim_{n \to \infty} \mu(-n, n]$ because $(-n, n] \uparrow \mathbb{R}$

$\mu(\Omega) = \lim_{n \to \infty} F(n) - F(-n) = F(\infty) - F(0) = 1 - 0 = 1$.

* Proof

Recall $\mathcal{S}$ = collection of $(a, b]$ from before

$\mathcal{S}$ is a semialgebra. Define μ on $\mathcal{S}$ by $\mu(a, b] = F(b) - F(a)$

* Check (1) Let $A_1, \dots, A_n \in \mathcal{S}$, pairwise disjoint and $\bigcup_{i=1}^{n} A_i \in \mathcal{S}$.

We want $\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i)$

- Must have $\bigcup_{i=1}^{n} A_i = (a, b]$ for $a < b$ and each A_i has the form $(\alpha_i, \beta_i]$ where α_i

Relabel $A_1, A_2, \dots, A_n$ in the form $(a_1, b_1], (a_2, b_2], \dots, (a_n, b_n]$ such that

$a_1 = a$, $b_n = b$ and $a_{i+1} = b_i$, $\forall i = \overline{1, <n}$

Now

• $\mu(\bigcup_{i=1}^{n} A_i) = \mu(a,b]$

• $\sum_{i=1}^{n} \mu(A_i) = \sum_{j=1}^{n} \mu(a_j,b_j) = \sum_{j=1}^{n} [F(b_j) - F(a_j)] =$

$= [F(b_1) - F(a_1)] + [F(b_2) - F(a_2)] + \cdots + [F(b_n) - F(a_n)]$

$= F(b) - F(a)$

+ Check (2): If $A_1, A_2, \ldots \in \mathcal{S}$, pairwise disjoint and $\bigcup_{i=1}^{\infty} A_i \in \mathcal{S}$ then $\mu(\underbrace{\bigcup_{i=1}^{\infty} A_i}_{\in \mathcal{S}}) \le \sum_{i=1}^{\infty} \mu(\underbrace{A_i}_{\in \mathcal{S}})$

• Suppose $(a,b] \subset \bigcup_{i=1}^{\infty} (a_i,b_i]$. Suppose $-\infty < a < b < +\infty$ (read text for the case when a, b are infini

Fix $\varepsilon > 0$. By right continuity, $\exists\ \delta > 0$, $F(a+\delta) - F(a) < \varepsilon$

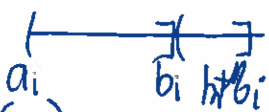
Also, by right continuity of F, for each i, $\exists\ \eta_i > 0$ such that $F(b_i + \eta_i) - F(b_i) < \frac{\varepsilon}{2^i}$

Thus $\mu(a, b+\delta] < \varepsilon$ and $\mu(b_i, b_i + \eta_i] < \frac{\varepsilon}{2^i}$

By part (1), $\mu(a,b] = \mu(a, a+\delta] + \mu(a+\delta, b] < \varepsilon + \mu(a+\delta, b]$ (I)

• Also, by $\mu(a_i, b_i + \eta_i] = \mu(a_i, b_i] + \mu(b_i, b_i + \eta_i)$

$\le \mu(a_i, b_i] + \frac{\varepsilon}{2^i}$ (II)



• Now consider $\overbrace{[a+\delta, b]}^{\subset (a,b]} \subset (a,b] \subset \bigcup_{i=1}^{\infty} (a_i, b_i] \subset \bigcup_{i=1}^{\infty} (a_i, b_i + \eta_i)$

we have an open cover of compact $[a+\delta, b]$, so there exist a finite subcover,

$[a+\delta, b] \subset \bigcup_{i \in I} (a_i, b_i + \eta_i)$, I finite

$\subset \bigcup_{i \in I} (a_i, b_i + \eta_i]$, I finite and so $\underbrace{(a+\delta, b]}_{\in \mathcal{S}} \subset \bigcup_{I} \underbrace{(a_i, b_i + \eta_i]}_{\in \mathcal{S}}$

Then by Lemma 1.1.5 b or exercise Set 1 #7, $\mu(a+\delta, b] \le \mu(a,b] \le \sum_{i \in I} \mu(a_i, b_i + \eta_i]$

$\le \sum_{i=1}^{\infty} \mu(a_i, b_i + \eta_i]$

$\underset{(II)}{\le} \sum_{i=1}^{\infty} \left(\mu(a_i, b_i] + \frac{\varepsilon}{2^i}\right) = \sum_{i=1}^{\infty} \mu(a_i, b_i] + \varepsilon \underbrace{\sum_{i=1}^{\infty} \frac{1}{2^i}}_{=1} = \sum_{i=1}^{\infty} \mu(a_i, b_i] + \varepsilon$

• So by I, $\mu(a,b] < \varepsilon + \mu(a+\delta, b] < \varepsilon + \sum_{i=1}^{\infty} \mu(a_i, b_i] + \varepsilon$

We have shown that $\mu(a,b] \le \sum_{i=1}^{\infty} \mu(a_i, b_i] + 2\varepsilon$, $\forall \varepsilon > 0$

$\Rightarrow \mu(a,b] \le \sum_{i=1}^{\infty} \mu(a_i, b_i]$ □ for (2) as long as $(a,b] \in \mathcal{S}$ (no restric on pairwi $\subset \bigcup_{i=1}^{\infty} (a_i, b_i]$

⊛ This shows that μ has an extension ν which is a measure on $(\mathbb{R}, \sigma(\mathcal{S}))$

We need to show that $\sigma(\mathcal{S}) = \mathcal{B} \overset{def}{=} \sigma(\text{all open subsets of } \mathbb{R}) = \sigma(\mathcal{O})$

* Proof of equality

1) Suppose $(a,b] \in \mathcal{S}$, $(a,b] = \bigcap_{n=1}^{\infty} (a, b + \frac{1}{n}) \in \sigma(\mathcal{O}) = \mathcal{B}$

Thus $\mathcal{S} \subset \sigma(\mathcal{O}) \Rightarrow \sigma(\mathcal{S}) \subset \sigma(\mathcal{O}) = \mathcal{B}$ ← need to prove this.

2) Prove $\sigma(\mathcal{O}) \subset \sigma(\mathcal{S})$

We have every open $\mathcal{U} \subset \mathbb{R}$ is a countable union of open intervals, $U = \bigcup_{i=1}^{\infty} (a_i, b_i)$

since $(a_i, b_i) = \underbrace{\bigcup_{n=1}^{\infty} \underbrace{(a, b - \tfrac{1}{n}]}_{\in \mathcal{S}}}_{\in \sigma(\mathcal{S})}$, every open interval $\in \sigma(\mathcal{S})$

Thus $\mathcal{U} \in \sigma(\mathcal{S})$ which implies $\mathcal{O} \subset \sigma(\mathcal{S}) \Rightarrow \sigma(\mathcal{O}) \subset \sigma(\mathcal{S})$

1.2) Probability distribution and real valued random variables

* Def (measurable function and random variable)

Let $(\Omega, \mathcal{F})$: measurable space. Let $(\mathbb{R}, \mathcal{B})$ Borel σ-algebra.

A map $f : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}, \mathcal{B})$ is called a Borel measurable function / random vario

$\overset{\text{def}}{\iff} f^{-1}(B) \in \mathcal{F}, \forall B \in \mathcal{B}$ $\quad f^{-1}(B) = \{ w \in \Omega, f(w) \in B \}$

* We typically use $X : \Omega \longrightarrow \mathbb{R}$ or Y, Z for random variables.

* Property (see example 1.1.1)

When (Ω), is a discrete probability space, then any $X : \Omega \longrightarrow \mathbb{R}$ is a random varic

* Example Indicator$_{\{0,1\}}$ function. check this.

$\mathbb{1} : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}, \mathcal{B})$

$$A \longmapsto \mathbb{1}_A(w) = \begin{cases} 1, & w \in A \\ 0, & w \notin A \end{cases}$$ is a random variable

* Def

* Let $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability measure space

Let $(X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B})$ be a random variable

Then X induces a probability measure on $\mathbb{R}$, called its distribution by setting (or law of X)

$\mu(B) = \mathbb{P}(\underbrace{X^{-1}(B)}_{\in \mathcal{F}}) = \mathbb{P}(X \in B)$ and μ is a probability measure on $(\mathbb{R}, \mathcal{B})$

* Def

- The distribution function (d.f) of X is F, where

$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X^{-1}(-\infty, x])$

- Convince yourself that μ is a probability on $(\mathbb{R}, \mathcal{B})$

If $B_1, B_2, \dots$ are pairwise disjoint elements of $\mathcal{B}$

Then $\begin{cases} X^{-1}(B_1), X^{-1}(B_2), \dots \text{ are pairwise disjoint elements of } \mathcal{F} \\ X^{-1}(\bigcup_{i=1}^{\infty} B_i) = \bigcup_{i=1}^{\infty} (X^{-1}(B_i)) \end{cases}$

$\Rightarrow \mu(\bigcup_{i=1}^{\infty} B_i) = \mathbb{P}^{-1}(X^{-1}(\bigcup_{i=1}^{\infty} B_i)) = \mathbb{P}(\bigcup_{i=1}^{\infty} X^{-1}(B_i)) \overset{\mathbb{P} \text{ is a measure}}{=} \sum_{i=1}^{\infty} \mathbb{P}(X^{-1}(B_i))$

$= \sum_{i=1}^{\infty} \mu(X^{-1}(B_i)) \Rightarrow \mu$ is countable additive.

* **Theorem 1.2.1**: Hide $(\Omega, \mathcal{F}, \mathbb{P})$

Let X be a random variable with d.f F

Then a) If $a < b$ then $\mathbb{P}(a < X \leq b) = F(b) - F(a)$

b) F is increasing on $\mathbb{R}$

c) $F(-\infty) = 0$, $F(\infty) = 1$

d) F is right continuous

e) $\mathbb{P}(X < x) = F(x-) = \lim_{y \uparrow x} F(y)$

f) $\mathbb{P}(X = x) = F(x) - F(x-)$

* Remark:

• F is a Stieltjes measure function, by theorem 1.1.2, $\exists!\, \nu$ on $(\mathbb{R}, \mathcal{B})$ so that $\nu(a,b] = F(b) - F$

• μ

Proof: These properties come from definition of F and the fact that $\mathbb{P}$ is a prob. m

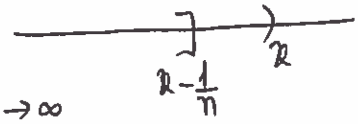
a) Since $(-\infty, b] = (-\infty, a] \cup (a, b]$

$$X^{-1}(-\infty, b] = X^{-1}(-\infty, a] \uplus X^{-1}(a, b].$$

$$\mathbb{P}(X^{-1}(-\infty, b]) = \mathbb{P}(X^{-1}(-\infty, a]) + \mathbb{P}(X^{-1}(a, b])$$

$$F(b) = F(a) + \mathbb{P}(a < X \leq b)$$

d) Prove that F is right continuous.



We have $(-\infty, x - \frac{1}{n}] \uparrow (-\infty, x)$ as $n \to \infty$

$$X^{-1}(-\infty, x - \tfrac{1}{n}] \uparrow X^{-1}(-\infty, x)$$

$$\mathbb{P}(X^{-1}(-\infty, x - \tfrac{1}{n}]) \uparrow \mathbb{P}(X^{-1}(-\infty, x)) \Leftrightarrow F(x - \tfrac{1}{n}) \uparrow F(x)$$

* Remark:

• F is a Stieltjes measure function, by theorem 1.1.2

$\exists!\, \nu$ a measure on $(\mathbb{R}, \mathcal{B})$ so that $\nu(a,b] = F(b) - F(a)$

• The dist μ of X denoted by $\mu(B) = \mathbb{P}(X^{-1}(B))$ is a probability measure on $(\mathbb{R}, \mathcal{B})$

$\mu(a,b] \overset{\text{def}}{=} \mathbb{P}(a < X \leq b) = F(b) - F(a)$ by a)

Thus $\mu = \nu$, the probability measure associated with F (by T1.1.2)

* Question:

Given a Stiejes measure function F with $F(-\infty)=0$, $F(\infty)=1$, is there a r.v with F as its d.f?

More precisely, is there a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with r.v X s.t $\mathbb{P}(X\leq x)=$

* Theorem 1.2.2.

If F is a Stiejes measure function then it is the distribution function of some random variable

$(\exists(\Omega,\mathcal{F},\mathbb{P})$ with a r.v. X such that $F(x)=\mathbb{P}(X\leq x)=\mathbb{P}(X^{-1}(-\infty,x])$

* Proof: Let F is fixed.

Let $\Omega=(0,1)$, $\mathcal{F}=$ Borel subset of $(0,1)$, $\mathbb{P}=$ Lebesgue measure on $(0,1)$

- Suppose F is continuous and stricly increasing.

Then F has a continuous, increasing inverse $F^{-1}(0,1)\longrightarrow\mathbb{R}$.

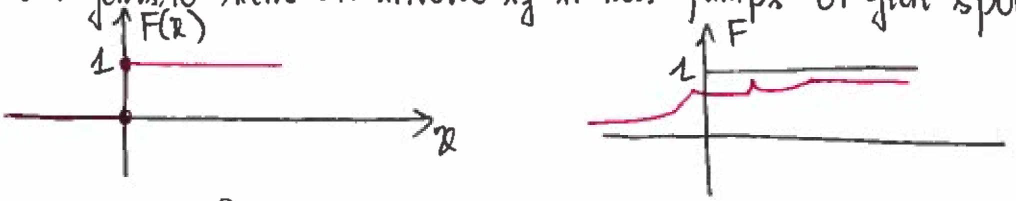
Define $\underline{|X(w)=F^{-1}(w)}$, $w\in(0,1)$. $F^{-1}(w)\leq x \Leftrightarrow w\leq F(x)$.

Then $\mathbb{P}(X\leq x)=\mathbb{P}\{w\in(0,1), F^{-1}(w)\leq x\}=\mathbb{P}(\{w\in(0,1), w\leq F(x)\})$

$=\mathbb{P}(0,F(x))=F(x)$.

i.e, X has d.f. F and X is a random variable.

- F fails to have an inverse if it has jumps or flat spots



- In general, define

$X(w)=\sup\{y, F(y)<w\}$, $w\in(0,1)$ — use this def to prove the claim.

claim: this set is always of the form $(-\infty,a)$ or $(-\infty,a]$

For all $x\in\mathbb{R}$, $\underbrace{\{w, X(w)\leq x\}}_{LHS}=\underbrace{\{w, w\leq F(x)\}}_{RHS}$

Given this, $\mathbb{P}(X\leq x)=\mathbb{P}\{w, 0<w\leq F(x)\}=F(x)$

Lebesgue measure.

- Prove the claim:

① Fix x, suppose $w\leq F(x)$. Then $x\notin\{y, F(y)<w\}$

Then $x \geq \sup\{y, F(y) < w\} \overset{\text{def}}{=} X(w)$

This show LHS $\subset$ RHS

⊕ Now want to show RHS $\subset$ LHS.

Suppose that $w > F(x)$. By right-continuity of F, $\exists \varepsilon$ s.t. $w > F(x+\varepsilon)$

If $x' = x + \varepsilon$, then $F(x') < w$

Thus $x' \in \{y, F(y) < w\}$, thus $x' \leq \sup\{y, F(y) < x\} =_{\text{def}} X(w)$

Thus $x' = x + \varepsilon \leq X(w)$

$\Rightarrow X(w) > x$

Thus $w \in RHS^c$ implies that $w \in LHC^c \Rightarrow RHS^c \subset LHS^c \Rightarrow$ ~~LHS ⊂ R~~ $\Rightarrow RHS \subset LHS$ cl

* Example.

For some $a < b$, define F by

$$F(x) = \begin{cases} 0 & x < a \\ \frac{1}{3} & a \leq x < b \\ 1 & x \geq b \end{cases}$$

This is a d.f. Convince yourself that $X(w) = \begin{cases} a & \text{if } w \in (0, \frac{1}{3}] \\ b & \text{if } w \in (\frac{1}{3}, 1] \end{cases}$

and that $\mathbb{P}(X = a) = \frac{1}{3}$, $\mathbb{P}(X = b) = \frac{2}{3}$

* **Def (density function)**

Suppose that X has d.f. F,

and $\exists$ measurable, non negative Borel function f defined on $\mathbb{R}$ such that $F(x) = \int_{-\infty}^{x} f(t)\, dt$

Then X (and F) is said to have density f

* Note

- $F(\infty) = 1 = \int_{-\infty}^{+\infty} f(t)\, dt$

1.3 Random variables (general)

Def

Let $(\Omega, \mathcal{F})$ and $(S, \mathcal{S})$ are measurable space ($\mathcal{S}$: σ-algebra)

A $\mathcal{F}$-measurable function is a function $X : (\Omega, \mathcal{F}) \longrightarrow (S, \mathcal{S})$ s.t. $X^{-1}(A) \in \mathcal{F}, \forall A \in$

- $(S, \mathcal{S}) \equiv (\mathbb{R}, \mathcal{B}) \to$ call X is a random variable
- $(S, \mathcal{S}) \equiv (\mathbb{R}^n, \mathcal{B}^n) \to$ " " " " " vector
- $(S, \mathcal{S})$ might be a metric space, Borel subsets.

Question: How can we tell if a given X is a $\mathcal{F}$-measurable function?

Theorem 1.3.1:

As in the def, suppose $\left.\begin{array}{l} \mathcal{a} \subset \mathcal{S} \text{ and } \sigma(\mathcal{a}) = \mathcal{S} \\ \text{If } X^{-1}(A) \in \mathcal{F}, \forall A \in \mathcal{a} \end{array}\right\} \Rightarrow X$ is a $\mathcal{F}$-measurable map

Example: Let $(S, \mathcal{S}) \equiv (\mathbb{R}, \mathcal{B})$ $\mathcal{a} = \{\text{all } (a,b]\}$

Proof the theorem: We have $X^{-1}(A) \in \mathcal{F}, \forall A \in \mathcal{a}$ want X is $\mathcal{F}$ measurable map

Let $\mathcal{A} = \{A \in \mathcal{S}, X^{-1}(A) \in \mathcal{F}\}$ we want $\mathcal{A} = \mathcal{S}$. need to prove $\mathcal{S} = \sigma(\mathcal{A})$ (1), $\sigma(\mathcal{A}) = \mathcal{A}$ (2)

(1) We know $\mathcal{a} \subset \mathcal{A} \Rightarrow \underset{= \mathcal{S}}{\sigma(\mathcal{a})} \subset \sigma(\mathcal{A}) \subset \mathcal{S} \Rightarrow \mathcal{S} = \sigma(\mathcal{A})$

(2) Now want to prove that $\sigma(\mathcal{A}) = \mathcal{A}$, it suffices to show that $\mathcal{A}$ is a σ-algebra.

⊕ Suppose $B \in \mathcal{A}$, NTP $B^c \in \mathcal{A} \Leftrightarrow B^c \in \mathcal{A}$ and $X^{-1}(B^c) \in \mathcal{F}$

$\Updownarrow$

$B \in \mathcal{A}$, $X^{-1}(B) \in \mathcal{F}$

we have $X^{-1}(B^c) = \{w, X(w) \in B^c\} = \{w, X(w) \in B\}^c = (X^{-1}(B))^c \in \mathcal{F}$ since $X^{-1}(B) \in \mathcal{F}$, σ-alg

⊕ Suppose $B_1, B_2, \ldots \in \mathcal{A}$, NTP $\bigcup_{i=1}^{\infty} B_i \in \mathcal{A}$

NTP $\underbrace{\bigcup B_i \in \mathcal{S}}_{\text{obvious since } \mathcal{S}\ \sigma\text{-algebra}}$ and $X^{-1}(\bigcup_{i=1}^{\infty} B_i) \in \mathcal{F}$

Consider $X^{-1}(\bigcup_{i=1}^{\infty} B_i) = \bigcup_{i=1}^{\infty} X^{-1}(B_i) \in \mathcal{F}$ because each $X^{-1}(B) \in \mathcal{F}$.

$(X^{-1}(\bigcup B_i) = \{w, X(w) \in \bigcup B_i\} = \{w, X(w) \in B_i, \underset{\text{some}}{\forall i}\} = \bigcup_{i=1}^{\infty} X^{-1}(B_i) \in \mathcal{F}$

σ-algebra generated by X, denoted σ(X) don't even need X is measurable

* Remark: We have shown in (2) that for any $X: (\Omega, F) \to (S, \mathcal{S})$, the collection of set $\{X^{-1}(B), B \in \mathcal{S}\}$ is a ~~sub~~ σ-algebra of F. (and don't even need the σ algebra F)

Example: $\Omega = \{1,2,3,4\}$ F all subsets $X: \Omega \to \mathbb{R}$ is

$X(w) = 0$ if w is odd } Then $\{X^{-1}(B), B \in \mathcal{B}\} = \{\phi, \Omega, \{1,3\}, \{2,4\}\}$

$X(w) = 1$ if w is even } which is a σ algebra and does not contain {1}

- Consider $X: (\Omega, F) \to (S, \mathcal{S})$

1) Suppose $A \in \{X^{-1}(B), B \in \mathcal{S}\}$

then $\exists B \in \mathcal{S}$, s.t $A = X^{-1}(B)$

then $A^c = (X^{-1}(B))^c = \{w, w \in B^c\} = X^{-1}(B^c)$

2) Check if $\{X^{-1}(B), B \in \mathcal{S}\}$ is closed under countable union

- when $X: (\Omega, F) \to (S, \mathcal{S})$

then the collection of $\{X^{-1}(B), B \in \mathcal{S}\}$ is a sub σ algebra of F.

* Example of applying theorem 1.3.1

1) $(\Omega, F), (S, \mathcal{S}) = (\mathbb{R}, \mathcal{B})$

we can take any of the form $\mathcal{A}$ so that $\sigma(\mathcal{A}) = \mathcal{B}$

all (a,b) $-\infty < a < b < +\infty$,

$(a,b]$ $-\infty < a < +\infty$

$[a,b]$ $-\infty < a < b < \infty$

$[a,b)$ $-\infty < a < b < +\infty$

- To check $X: (\Omega, F) \to (S, \mathcal{S})$ is measurable it is enough to check that $X^{-1}(a,b) \in F$, $\forall -\infty < a < b < +\infty$

- can even just need to check for $a, b \in \mathbb{Q}$

2) For $(S, \mathcal{S}) = (\mathbb{R}^n, \mathcal{B}^n)$, we can $\mathcal{A}$ to be all sets of the form

$(a_1, b_1) \times (a_2, b_2) \times \dots \times (a_n, b_n)$, all $-\infty < a_i < b_i < +\infty$ (rectangles).

* Extended real line $\mathbb{R}^* = [-\infty, +\infty]$

$\mathcal{B}^*$ is the σ algebra generated by the collection of sets $\{[-\infty, a), a \in \mathbb{R}$; (a,b) $a < b$ $a, b \in \mathbb{R}\}$.

Fact: $B \in \mathcal{B}^* \Leftrightarrow B \cap \mathbb{R} \in \mathcal{B}$

A map $X: (\Omega, F) \to (\mathbb{R}^*, \mathcal{B}^*)$ such that X is F-$\mathcal{B}^*$ measurable is called an extended random variable.

* **Theorem 1.3.3**

Let $X_1, X_2, \dots, X_n$ be r.v's on (Ω, F)

$f: (\mathbb{R}^n, \mathcal{B}^n) \longrightarrow (\mathbb{R}, \mathcal{B})$ be Borel measurable

Then $Z(\omega) = f(X_1(\omega), X_2(\omega), \dots, X_n(\omega)) : \Omega \longrightarrow \mathbb{R}$ is a random variable

$\omega \longmapsto Z(\omega)$

* **Theorem 1.3.2**

Let (Ω, F), $(S, \mathcal{S})$, $(G, \mathcal{G})$ be measurable space

$X: (\Omega, F) \longrightarrow (S, \mathcal{S})$ is measurable

$Y: (S, \mathcal{S}) \longrightarrow (G, \mathcal{G})$ is measurable

Then f, $f: \Omega \longrightarrow G$, $\omega \longmapsto f(\omega) = Y(X(\omega))$ is $(F, \mathcal{G})$ measurable

* Proof theorem 1.3.2

If $\Gamma \in \mathcal{G}$, then $f^{-1}(\Gamma) = (Y \circ X)^{-1}(\Gamma) = \underbrace{X^{-1}(\underbrace{Y^{-1}(\Gamma)}_{\in \mathcal{S}})}_{\in F}$ because X and Y are measurable

* Proof theorem 1.3.3

- Let $Y(\omega) = (X_1(\omega), X_2(\omega), \dots, X_n(\omega))$, so $Y: \Omega \longrightarrow \mathbb{R}^n$ and so $Z(\omega) = f(Y(\omega))$, Z will be measurable if Y is measurable.
- Now we prove that Y is measurable,

i.e, $Y^{-1}((a_1, b_1) \times (a_2, b_2) \times \dots \times (a_n, b_n)) \in F$ for all $-\infty < a_i < b_i < +\infty$

$$Y^{-1}((a_1,b_1)\times \dots \times (a_n,b_n)) = \{\omega, (X_1(\omega), \dots, X_n(\omega)) \in (a_1,b_1)\times \dots \times (a_n,b_n)\}$$
$$= \{\omega, X_1(\omega) \in (a_1,b_1), \dots, X_n(\omega) \in (a_n,b_n)\}$$
$$= \bigcap_{i=1}^{n} \underbrace{X_i^{-1}(a_i,b_i)}_{\in F \text{ since each } X_i \text{ is a random variable}} \in F$$

This shows Y is measurable and then Z is measurable.

* **Theorem 1.3.4**

If $X_1, X_2, \dots, X_n$ are r.v's then $X_1 + X_2 + \dots + X_n$ is a random variable

* Proof

- By theorem 1.3.3, we need to prove that $f: \mathbb{R}^n \longrightarrow \mathbb{R}$ is Borel measurable

$(x_1, \dots, x_n) \mapsto f(x_1, \dots, x_n) = x_1 + \dots + x_n$

It suffices to show that for $-\infty < a < +\infty$, $f^{-1}(-\infty, a) \in \mathcal{B}^n$

$f^{-1}(-\infty, a) = \{\underbrace{(x_1, \dots, x_n)}_{\in \mathbb{R}^n}, x_1 + \dots + x_n < a\}$ an open set in $\mathbb{R}^n$, hence in $\mathcal{B}^n$

* Theorem 1.3.5

Let $X_1, X_2, \ldots$ be random variables

Then the following are (possibly $\mathbb{R}^*$ valued) random variables

$$\inf_{n\geq 1} X_n \quad \sup_{n\geq 1} X_n \quad \liminf_{n\to\infty} X_n \quad \limsup_{n\to\infty} X_n \leftarrow \text{some of these may be } \infty$$

* Proof

① $\left\{\inf_{n\geq 1} X_n < a\right\} = \underbrace{\bigcup_{n=1}^{\infty} \underbrace{\{X_n < a\}}_{\in F}}_{F}$ ② $\left\{\sup_{n\geq 1} X_n > a\right\} = \underbrace{\bigcup_{n=1}^{\infty} \underbrace{\{X_n > a\}}_{\in F}}_{\in F}$

③ $\liminf_{n\to\infty} X_n(\omega) = \sup_{m\geq 1}\left(\inf_{n\geq m} X_n(\omega)\right)$ measurable by (1) and (2).

• Remark. Let $\Omega_0 = \{\omega, \lim_{n\to\infty} X_n(\omega) \text{ exists and } (\in \mathbb{R})\}$

Then Ω_0 is measurable ($\in F$) because $\Omega = \left\{\omega,\ \limsup X_n(\omega) \in \mathbb{R},\ \liminf X_n(\omega) \in \mathbb{R},\ \limsup X_n(\omega) = \liminf X_n(\omega)\right.$

each set is measurable $\leftarrow$ $= \{\omega, \limsup \in \mathbb{R}\} \cap \{\omega, \liminf \in \mathbb{R}\}$

because limsup and liminf are measurable $\cap \{\omega, \limsup = \liminf\}$.

1.4.7 Integration

Def $(X, \mathcal{F}, \mu)$ a measurable space, μ is σ-finite

A property Q holds μ-a.e if $\mu\{x \mid Q \text{ does not hold}\} = 0$

* Example: $f \geq g$ μ-a.e $\Leftrightarrow$ $\mu\{x \mid f(x) < g(x)\} = 0$

* Definition (simple function)

$\phi : (X, \mathcal{F}, \mu) \longrightarrow (\mathbb{R})$ is a simple function $\overset{\text{def}}{\Longleftrightarrow}$ $\begin{cases} \phi(x) = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}(x) \\ \alpha_i \in \mathbb{R} ; \\ A_i \in \mathcal{F}, \ \mu(A_i) < +\infty \end{cases}$

* Outline of instruction of the integral

* Step 1: For simple function $\phi = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$

then $\int \phi \, d\mu = \sum_{i=1}^{n} a_i \mu(A_i)$

• We must need to check that $\int \phi \, d\mu$ is well defined

• If $\phi = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i} = \sum_{j=1}^{m} b_j \mathbb{1}_{B_j}$. We need to check $\sum_{i=1}^{n} a_i \mu(A_i) = \sum_{j=1}^{m} b_j \mu(B_j)$

We will use the lemma:

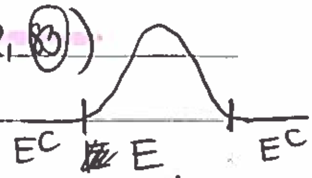
* Lemma 1.4.1 Let ϕ and ψ are simple functions, then

i) If $\phi \geq 0$ a.e then $\int \phi \, d\mu \geq 0$ iv) If $\phi \geq \psi$ a.e $\Rightarrow \int \phi \, d\mu \geq \int \psi \, d\mu$.

ii) $\forall a \in \mathbb{R}$, $\int a\phi \, d\mu = a \int \phi \, d\mu$

iii) $\int (\psi + \phi) \, d\mu = \int \psi \, d\mu + \int \phi \, d\mu$

* Step 2: Let $\begin{cases} f \text{ be a bounded (Borel) measurable function } f : (X, \mathcal{F}, \mu) \longrightarrow (\mathbb{R}, \mathcal{B}) \\ f \equiv 0 \text{ on } E^c \text{ where } E \in \mathcal{F}, \ \mu(E) \leq \infty \end{cases}$



(f is bounded Borel measurable and nonzero in a finite set)

$\int f \, d\mu = \sup\{ \int \phi \, d\mu, \ \phi \text{ is simple } (\phi = 0 \text{ on } E^c), \ \phi \leq f \text{ a.e} \}$ $E = \bigcup_{i=1}^{n} A_i$ in simple function

$= \inf\{ \int \psi \, d\mu, \ \psi \text{ is simple}, \ \psi = 0 \text{ on } E^c, \ \psi \geq f \text{ a.e} \}$.

$f = f^+ - f^-$ $\quad |f| = f^+ + |f^-|$ | If $a \in \mathbb{R}$, $a^+ = \max\{a, 0\}$.

$a^- = \cancel{\min\{a,0\}}$
$= \max\{-a, 0\}$.

$a = a^+ - a^-$

① Check if f is simple, we get the same value of $\int f d\mu$

② $\int f d\mu$ is finite

③ If $f = 0$ on D^c where $D \in \mathcal{F}$ and $\mu(D) < +\infty$, then we get the same value for $\int$.

The def must establish properties of $\int f d\mu$ (analogues of lemma 1.4.1)

* Step 3: Let $f: X \to [0, +\infty]$ be measurable (means f is nonnegative)

Define $\int f d\mu = \sup\{\int h d\mu$, h is measurable, $0 \le h \le f$, h is bounded, $\mu\{x,\ h(x) > 0\}$

This integral is well defined but may be $+\infty$

- Check this agrees with def of the integral in step 2.

* Step 4: For any measurable $f: X \to [-\infty, +\infty]$

if at least one of $\int f^+ d\mu$, $\int f^- d\mu$ is finite

Then let $\int f d\mu = \int f^+ d\mu - \int f^- d\mu$

* If $\int |f| d\mu < +\infty$ (f is measurable?) we say f is integrable

- Notation: if μ is Lebesgue measure on $(\mathbb{R}, \mathcal{B})$ then we write $\int f d\mu = \int f dx$
- For any $E \in \mathcal{F}$, define $\int_E f d\mu = \int (f \mathbb{1}_E) d\mu$.

* Theorem 1.4.7

Suppose f and g are integrable

i) $f \ge 0$ a.e then $\int f d\mu \ge 0$

ii) $\forall a \in \mathbb{R}$, $\int a f d\mu = a \int f d\mu$.

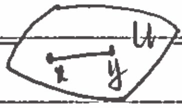
iii) $\int (f+g) d\mu = \int f d\mu + \int g d\mu$.

iv) $f \le g$ a.e $\Rightarrow \int f d\mu \le \int g d\mu$

v) $f = g$ a.e $\Rightarrow \int f d\mu = \int g d\mu$

vi) $\left|\int f d\mu\right| \le \int |f| d\mu$.

1.5) Properties of the intergrals



* Def:

- $U \subset \mathbb{R}^n$ is convex if $\forall x, y \in U$, $tx + (1-t)y \in U$ $\forall t \in [0,1]$
- $f : U \longrightarrow \mathbb{R}$ (is U need to be convex set)

then f is convex if $\forall x, y \in U$, $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$, $\lambda \in [0$

- Fact:

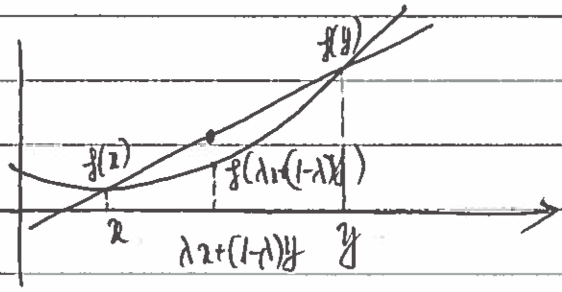
If $f : (a,b) \longrightarrow \mathbb{R}$, f'' exists on (a,b) } then f is convex $\Leftrightarrow$ $f'' \geq 0$ on (a,b).

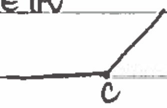
* Example of convex function

1) $f(x) = e^x$

2) $f(x) = |x|^p$ for any $p \geq 1$

3) $f(x) = (x-c)^+$ any $c \in \mathbb{R}$





* Theorem 1.5.1 (Jensen's inequality)

If μ is a probability measure on $(\mathbb{R}, \mathcal{B})$ $\quad \mu : (\mathbb{R}, \mathcal{B}) \longrightarrow (\mathbb{R}, \mathcal{B})$ $\mu(\mathbb{R}) = 1$

$\phi : \mathbb{R} \longrightarrow \mathbb{R}$ is convex

$f, \phi(f)$ are integrable function $\quad f : X \longrightarrow \mathbb{R}$, $\quad \varphi : X \longrightarrow \mathbb{R}$.

Then $\phi\left(\int f \, d\mu\right) \leq \int \phi(f) \, d\mu$.

* Def (norm)

For any measurable f, let $\|f\|_p = \left[\int |f|^p d\mu\right]^{1/p}$

$\|f\|_\infty = \inf\{K, \mu(|f| > K) = 0\}$

* Prop: If f is measurable, then $p \longrightarrow \|f\|_p$ is increasing in p

$\|f\|_p \leq \|f\|_q$ when $0 < p \leq q < \infty$

• In particular, If $0 < p < q$, $\int |f|^q < +\infty$ then $\int |f|^p d\mu < +\infty$

* Prop: If f is measurable, then $p \longmapsto \|f\|_p$ is increasing in p

$\|f\|_p \le \|f\|_q$ when $0 \le p < q \le +\infty$

* Proof Apply Jensen's inequality.

Let $0 < p < q < +\infty$, then $\phi(x) = x^{q/p}$ is (for $x \geq 0$) is convex function.

$\int |f|^q d\mu = \int (|f|^p)^{q/p} d\mu \xlongequal[\text{let } g = |f|^p]{} \int \phi(g) d\mu \geq \phi \int g \, d\mu = \phi \int |f|^p d\mu$.

$= \left(\int |f|^p d\mu\right)^{q/p}$? condition to apply Jensen inequality.

Take q root both sides to get $\int |f|^q d\mu \geq \int |f|^p d\mu$ when $q > p$

---

* Theorem 1.5.2 (Holder inequality)

If $p, q \in (1, \infty]$, $\frac{1}{p} + \frac{1}{q} = 1$, f, g measurable $\Big\}$ $\|fg\|_1 \le \|f\|_p \|g\|_q \xRightarrow{p=q=2} \|fg\|_1 \le \|f\|_2 \|g\|_2$

(Cauchy Schwarz inequality).

* Exercise 1.5.3 Minkowski inequality. $\|f+g\|_p \le \|f\|_p + \|g\|_p$ $\quad p \le +\infty$

1.5.1 Convergence Theorems

Always $(X,\mathcal{E},\mu)$ is a measurable function space, μ is finite measurable.

* def:

Let $f_1, f_2, \ldots$ be measurable functions. Def of $f_n \to f$ as $n \to \infty$

- $f_n \to f$ a.e $\iff \mu\{x : \lim_{n\to\infty} f_n(x) \neq f(x)\} = 0$
- $f_n \to f$ in measure $\iff \forall \epsilon > 0,\ \mu\{x, |f_n(x) - f(x)| > \epsilon\} \longrightarrow 0$ (in probabil
- $f_n \to f$ in L^p means $\|f_n - f\|_p \xrightarrow{n\to\infty} 0$

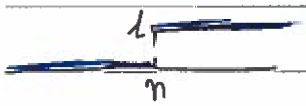
When can we say $\int f_n \, d\mu \to \int f \, d\mu$?

* In general $f_n \xrightarrow{a.e} f \not\Rightarrow f_n \xrightarrow{measure} f$

Example: Consider $(\mathbb{R}, \mathcal{B})$, μ = Lebesgue measure

$f_n = 1_{[n,+\infty)}$ $\quad f(x) = 0, \forall x \in \mathbb{R}$ $\quad$ then $f_n(x) \xrightarrow{n\to\infty} f(x), \forall x \in \mathbb{R} \Rightarrow f_n \xrightarrow{a.e} f$

- $\mu(\{x : |f_n(x) - f(x)| > \frac{1}{2}\}) = \mu[n,+\infty) = +\infty, \forall n$.



$\not\to 0$ as $n \to \infty$

which means $f_n \not\xrightarrow{measure} f$

* Fact:

If μ is a finite measure, and $f_n \xrightarrow{a.e} f$ μ a.e then $f_n \xrightarrow{measure} f$.

Probability

Proof: Let $E = \{x, \lim_{n\to\infty} f_n(x) = f(x)\}$ then $\mu(E^c) = 0$

Fix any $\epsilon > 0$,

Let $E_n' = (\{x \in E, |f_n(x) - f(x)| > \epsilon\}) = \{\cancel{x \in E}, \sup_{k \geq n}$

note $E_n' \subset E_n$

$$E_n = \{x \in E, \sup_{k \geq n} |f_k(x) - f(x)| > \epsilon\}$$

Then $\underbrace{\mu(E_n) \longrightarrow 0}$, and since $\mu(E^c) = 0$, $f_n \to f$ in measure

why $E_{n+1} \subset E_n$ so $E_n \downarrow \emptyset$ as $n \to \infty$ $\Big\}$ $\Rightarrow \mu(E_n) \longrightarrow 0$

since μ is a finite measure

- why not use E_n' and show that $\mu(E_n') \to 0$? Because in general $E_{n+1}' \not\subset E_n'$

$(X, \mathcal{F}, \mu) \longrightarrow (E, \mathcal{F}|_E, \mu|_E)$ Or can be taken in X where $\mu(X) < +\infty$

* Theorem 1.5.3 (Bounded convergence theorem)

Let E be measurable set, $\mu(E) < +\infty$ (Think of $E \subset X$)

If $f_1, f_2, \dots$ be measurable functions; $f_n = 0$ on E^c

$\exists\, M < +\infty$ such that $|f_n(x)| < M \quad \forall x \in E, \forall n = 1, 2, \dots$

$f_n \xrightarrow[n\to\infty]{\text{measure}} f$ or $f_n \xrightarrow[n\to\infty]{a.e} f$ on E μ

$\Big\}$ $\lim \int f_n \, d\mu = \int f \, d\mu$

Exercise: why does f integrable?

Recall $f(x) = \begin{cases} \frac{\sin x}{x} & x \neq 0 \\ 1 & x = 0 \end{cases} \notin L^1(\mathbb{R})$ means $\int_0^\infty \left|\frac{\sin x}{x}\right| dx = \infty$

f not integrable although $\int_0^\infty \frac{\sin x}{x} dx$ exist an improper Riemann integral.

*Proof: Fix $\varepsilon > 0$, note $|f_n| < M$ a.e since $\{x, |f(x)| > M + \delta\} \subset \{|f_n(x) - f(x)| \dots$

$\mu\{x, |f(x)| > M + \delta\} \leq \mu\{|f_n(x) - f(x)| \dots$

so $\mu\{x, |f(x)| > M + \delta\} = 0$ for $\delta \xrightarrow{n\to\infty} 0$

Let $\delta \downarrow 0$, $\{x, |f(x)| > M + \delta\} \uparrow \{x, |f(x)| > M\}$

$\Rightarrow \mu(\{x, |f(x)| \geq M\}) = 0$

$$\left|\int f_n \, d\mu - \int f \, d\mu\right| = \left|\int (f_n - f) \, d\mu\right| \leq \int |f_n - f| \, d\mu = \int_{|f_n - f| > \varepsilon} |f_n - f| \, d\mu + \int_{|f_n - f| \leq \varepsilon} |f_n - f| \, d\mu$$

$$\leq \int_{|f_n - f| > \varepsilon} 2M \, d\mu + \int_{|f_n - f| \leq \varepsilon} \varepsilon \leq 2M \, \mu\{|f_n - f| > \varepsilon\} + \varepsilon \, \mu\{|f_n - f| \leq \varepsilon\}$$

$$= 2M \, \mu\{|f_n - f| > \varepsilon\} + \varepsilon \, \mu(E)$$

$$\limsup_{n\to\infty} \left|\int f_n \, d\mu - \int f \, d\mu\right| \leq 0 + \varepsilon \underbrace{\mu(E)}_{< +\infty} \longrightarrow 0$$

$$\Rightarrow \lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu$$

All functions are measurable

★ Theorem 1.5.4 Fatou's lemma

If f_n measurable, $f_n \geq 0, \forall n$ $\int \left(\liminf_{n\to\infty} f_n \right) d\mu \leq \liminf_{n\to\infty} \left(\int f_n \, d\mu \right)$

• Note: ① This is often used when $\lim_{n\to\infty} f_n$ exists

② Strict inequality '<' can hold

③ Both sides equal to $+\infty$ is possible

④ One main use of this is if the RHS is finite then so LHS, $\liminf_{n\to\infty} f_n \in$

★ Theorem 1.5.5 (Monotone convergence theorem) (MCT)

If f_n measurable; $f_n \geq 0, \forall n$, $f_n \uparrow f$ as $n\to\infty$ then $\int f_n \, d\mu \uparrow \int f \, d\mu$

$\lim \int f_n \, d\mu = \int \lim f_n \, d\mu$

• Note: ① This includes $+\infty = +\infty$, but if one side is finite, so is the other

★ Theorem 1.5.6 (Dominated convergence theorem) (DCT)

If f, f_n measurable, $f_n \xrightarrow{a.e} f$

(*) $\exists g$ s.t. $|f_n| \leq g$ a.e $\forall n$

$g \in L^1(\mu)$, ~~then f~~

$\Bigg\}$ f is integrable

$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu$

• Note: (*) is not enough to replace (*) with $\|f_n\|_1 \leq \|g\|_1$, $\forall n$

we have to have pointwise domination (see text for counterexample).

1.6) Expected value (Change of terminologies)

* Let $(\Omega, \mathcal{F}, P)$ be probability space

  X be a random variable — an abstract integral — Mean / of X

Then we write $E(X) = \int X \, dP$ (as well as EX) $\leftarrow$ Expected value / Expectation

- We will write $E(X; A) = E(X \mathbb{1}_A) = \int_A X \, dP$, $A \in \mathcal{F}$

* Theorem 1.6.1

If X and Y are random variables.

Either both X and Y are integrable or both X and X are nonnegative. Then

a) $E(X+Y) = E(X) + E(Y)$ (possibly $+\infty = +\infty$)

b) $E(aX+b) = aE(X) + b$, $\forall a, b \in \mathbb{R}$, ($a \neq 0$ if $X \geq 0$, $E(X) = +\infty$)

c) If $X \geq Y$ then $E(X) \geq E(Y)$

- Note for integrable X & Y, see theorem 1.4.7 (end of construction of integral)

For X and $Y \geq 0$ (possibly $E(X)$ or $E(Y) = +\infty$) in step 3 of the construction of the integ

Lema 1.4.5. In this case, one allows $+\infty = +\infty$ in a)

- Otherwise, to prove $X, Y \geq 0$ case,

define $X_n = \underset{\text{minimum}}{X \wedge n}$, $Y_n = Y \wedge n$, these are bounded, hence integrable,

then $E(X_n + Y_n) = E(X_n) + E(Y_n)$, $\forall n$.

Note that $X_n \uparrow X$, $Y_n \uparrow X$ and $X_n + Y_n \uparrow X + Y$ as $n \to \infty$

Apply MCT 3 times, to get $E(X+Y) = E(X) + E(Y)$

* Markov inequality:

$\left.\begin{array}{l}\text{If } X \text{ is a random variable, } (X \geq 0) \\ a > 0\end{array}\right\} \to P(X \geq a) \leq \frac{E(X)}{a} \Rightarrow P(|X| \geq a) \leq \frac{E(|X|)}{a}$

* Proof: Let $A = \{X \geq a\}$

$E(X) \geq \underbrace{E(X \mathbb{1}_A) \geq E(a \mathbb{1}_A)}_{X\mathbb{1}_A \geq a\mathbb{1}_A \text{ since } A = \{X \geq a\}} = a E(\mathbb{1}_A) = a P(A) \Rightarrow P(A) \leq \frac{E(X)}{a}$ $\square$

* Def:

- Let $\ell = 1, 2, \dots$

The ℓ^{th} moment of an X is $E(X^\ell)$ provided $E(X^\ell)$ exists $\Big\langle \begin{array}{l} X^\ell \geq 0 \\ E(|X|^\ell) < +\infty \end{array}$

- The first moment of X is just $E(X)$

* If X is integrable ($E(|X|) < +\infty$) Let $\mu = E(X)$

Define $Var(X) = E((X - E(X))^2)$ variance of X is the 2^{nd} moment of $(X - E(X))$

always exists but may be ∞

* Fact $Var(X) = E((X - EX)^2) = EX^2 - (EX)^2$ (include $+\infty = +\infty$)

If $E(X^2) < +\infty$, then $Var(X) = E((X-EX)^2) = E(X^2 - 2X\mu + \mu^2) =$

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

* Fact

a) $Var(aX + b) = a^2 Var(X)$ if $Var(X) < +\infty$

b) $Var(X) \leq E(X^2)$

* Theorem (Chebyshev inequality) $\Rightarrow P(|X| \geq a) \leq \frac{E(|X|^2)}{a^2}$ better than Markov ineq for large

If X has finite mean μ, $a > 0$ Then $P(|X - \mu| \geq a) \leq \frac{Var(X)}{a^2}$

- Proof

$$P(|X - \mu| \geq a) = P(|X - \mu|^2 \geq a^2) \leq \frac{E((X-\mu)^2)}{a^2} = \frac{Var(X)}{a^2}$$

---

* Theorem 1.6.9 (Change of variables)

Suppose (Ω, F, P) is a ~~measurable~~ probability space

$(S, \mathcal{S})$ is a measurable space

$X : (\Omega, F, P) \longrightarrow (S, \mathcal{S})$ is a S random variable

F is a distribution of X $(F(B) = P(X \in B))$, $B \in \mathcal{S}$

$f : S \longrightarrow \mathbb{R}$ is a Borel measurable function such that $f(x)$ is integrable ($f \geq 0$, $E(|f|) < +\infty$)

Then $(\Omega, F, P) \xrightarrow{X} (S, \mathcal{S}, F) \xrightarrow{f} (\mathbb{R}, \mathcal{B})$

$(f \circ X)$ is a real random variable on (Ω, F)

second probability space (Think of $(S, \mathcal{S})$ as $(\mathbb{R}^n, \mathcal{B}^n)$)

Then $E(f(X)) = \int_S f \, dF$ $(*)$

$\qquad \| $

$\int_\Omega (f \circ X) \, dP$

* Example: $f(x) = x^k$ for some fixed k $(S,\mathcal{S}) = (\mathbb{R}, \mathcal{B})$ $f \circ X = X^k$ $E(X^k) = \int_{\mathbb{R}} x^k d\mu$ $[0,+\infty)$

If X has exponential, parameter 1 distribution, then $= \int_{\mathbb{R}} x^k \mu(dx)$

$$E(X^k) = \int_{[0,+\infty)} x^k e^{-x} dx \underset{\text{check}}{=} k!$$

means X has distribution μ, $\mu(B) = \int_{[0,+\infty)} 1\!\!1_B(x) e^{-x} dx$ or $\mu = e^{-x} 1\!\!1_{[0,+\infty)}(x)\, dx$

X has density function $h(x) = e^{-x} 1\!\!1_{[0,+\infty)}(x)$

* Proof: Successively verifying $(*)$ for f of the following types:

① Indicators:

If $f = 1\!\!1_B$ then $E(f(X)) = E(1\!\!1_B(X)) \overset{= \int_B x\,dP}{=} P(X \in B) \overset{\text{def}}{=} \mu(B) = \int_S 1\!\!1_B d\mu = \int_S f d\mu$

② Simple function

If $f = \sum_{i=1}^{n} c_i 1\!\!1_{B_i}$, $c_i \in \mathbb{R}$, $B_i \in \mathcal{S}$ (since finite sum → assume B_i pairwise disjoint

$$E(f(X)) = E\left(\sum_{i=1}^{n} c_i 1\!\!1_{B_i}(X)\right) \underset{\text{of the integral}}{\overset{\text{linearity}}{=}} \sum_{i=1}^{n} c_i E(1\!\!1_{B_i}(X)) = \sum_{i=1}^{n} c_i P(X \in B_i)$$

on

$$= \sum_{i=1}^{n} c_i \mu(B_i) = \sum_{i=1}^{n} c_i \int_S 1\!\!1_{B_i} d\mu \underset{\text{of the intgl}}{\overset{\text{linearity}}{=}} \int_S \left(\sum c_i 1\!\!1_{B_i}\right) d\mu = \int_S f d\mu.$$

③ Suppose $f \geq 0$ (nonnegative case), choose simple function f_n, $f_n \geq 0$, $f_n \uparrow f$ pointwise on S

Then by (2) $E(f_n(X)) = \int_S f_n d\mu$, $\forall n$

Applying monotone convergence theorem to both sides

For LHS, $f_n(X(\omega)) \geq 0$ and $f_n(X(\omega)) \uparrow f(X(\omega))$

So $E(f_n(X)) \to E(f(X))$

Also $\int_S f_n d\mu \to \int_S f d\mu$.

$\Rightarrow E(f(X)) = \int_S f d\mu$

④ General case $f = f^+ - f^-$ as usual where $f^+ \geq 0$, $f^- \geq 0$

Assume $E(|f(X)|) < +\infty$ which is the same as $E(f^+(X)) < +\infty$ $E(f^-(X)) < +\infty$

$\Rightarrow E(f(X)) = E(f^+(X)) - E(f^-(X)) = \int_S f^+ d\mu - \int_S f^- d\mu = \int_S f d\mu.$

* 1.7 Product measure

Assume $(\Omega_1, \mathcal{A}_1, \mu_1)$ and $(\Omega_2, \mathcal{A}_2, \mu_2)$ are σ-finite measure spaces

Let $\Omega = \Omega_1 \times \Omega_2 = \{(w_1, w_2), w_1 \in \Omega_1, w_2 \in \Omega_2\}$

We want a σ-algebra $\mathcal{F}$ on Ω and a σ-finite measure on Ω such that

$\forall A \in \mathcal{A}_1, A_2 \in \mathcal{A}_2$, $\mu(A_1 \times A_2) = \mu(A_1) * \mu(A_2)$

Let $\mathcal{S} = \{A_1 \times A_2, A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\}$ (collection of rectangles)

Check that $\mathcal{S}$ is a semialgebra

Let $\mathcal{F} = \sigma(\mathcal{S}) \overset{\text{notation}}{=\!=} \mathcal{A}_1 \times \mathcal{A}_2$ $\leftarrow$ bad notation since $\mathcal{F} \neq \{A_1 \times A_2, A_1 \in \mathcal{A}_1, \dots$

---

* Theorem 1.7.1 (helps Fubini theorem)

There is a unique σ-finite measure μ on $(\Omega, \mathcal{F})$ so that $\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$ $\forall A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2$

* Proof:

① Define $\mu(A_1 \times A_2) := \mu_1(A_1) \cdot \mu_2(A_2)$ for all $A_1 \times A_2 \in \mathcal{S}$

② Check conditions of theorem 1.1.4

See text for details ! Short and helps with Fubini theorem

---

* Theorem 1.7.2 (Fubini's theorem)

Let $(\Omega_1, \mathcal{A}_1, \mu_1)$ and $(\Omega_2, \mathcal{A}_2, \mu_2)$ be σ-finite measurable spaces

$\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F}$, μ as above

Let $f: \Omega = \Omega_1 \times \Omega_2 \longrightarrow \mathbb{R}$ be $\mathcal{F}$-measurable

Then if $f \geq 0$ or f is $\mu_1 \times \mu_2$ integrable

$$(*) \quad \underbrace{\int_{\Omega_1} \int_{\Omega_2} f(x,y)\, \mu_2(dy)\, \mu_1(dx)}_{A} = \int_{\Omega_1 \times \Omega_2} f \, d\mu = \underbrace{\int_{\Omega_2} \int_{\Omega_1} f(x,y)\, \mu_1(dx)\, \mu_2(dy)}_{B}$$

Includes $+\infty = +\infty$ when $f \geq 0$

* Remark:

① For RHS to make sense, we need $\begin{cases} \text{for fixed } y,\ f(x,\cdot) \text{ is } \mathcal{A}_1 \text{ measurable} \\ y \longmapsto \int f(x,y)\, d\mu_1(x) \text{ is } \mathcal{A}_2 \text{ measurable} \end{cases}$

② In $(*)$, if one side is finite, so is the other

* Exercise 1.7.1 (How do we check if f is $\mu_1 \times \mu_2$ integrable i.e. $\int_{\Omega_1 \times \Omega_2} |f| \, d\mu < +\infty$

Claim: f is $\mu_1 \times \mu_2$ integrable iff $\int_{\Omega_2} \left(\int_{\Omega_1} |f(x,y)| \, d\mu_1(x) \right) d\mu_2(y) < \infty$ (or when reverse the order of x and y)

- If f is $\mathcal{a}_1 \times \mathcal{a}_2$ measurable and either of using $|f|$ is finite, then $f \in L^1(\mu_1 \times \mu_2)$, so $(*)$ applies to f.

* Proof: apply Fubini twice.
First, apply to $|f| \geq 0$, so $(*)$ apply to $|f|$ (instead of f)
But if either A or B is finite (for $|f|$), then $\int_{\Omega_1 \times \Omega_2} |f| \, d(\mu_1 \times \mu_2) < +\infty$
So $f \in L^1(\mu_1 \times \mu_2)$
So now $(*)$ applies to f (conclusion of Fubinni) for f

- See text for example for $f \notin L^1(\mu_1 \times \mu_2)$, both Ⓐ and Ⓑ are finite but not equal.

* Chapter 2: Law of Large numbers

* 2.1) Independence $(\Omega, \mathcal{F}, \mathbb{P})$

* Def

- Even A and $B \in \mathcal{F}$ are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$
- Two random variables X and Y are independent if
  $\mathbb{P}(X \in C, Y \in D) = \mathbb{P}(X \in C)\,\mathbb{P}(Y \in D)$, $\forall$ Borel C, D
- Two algebra $\mathcal{A}, \mathcal{B} \in \mathcal{F}$ are independent iff (def) $\forall A \in \mathcal{A}$, $\forall B \in \mathcal{B}$
  then A and B are independent

* Exercise 2.1.1

i) If X and Y are independent rvs then $\sigma(X)$ and $\sigma(Y)$ are independent

ii) If $\left.\begin{array}{l} \mathcal{A} \text{ and } \mathcal{B} \text{ are independent } \sigma\text{-algebras} \\ X \text{ is } \mathcal{A}\text{-measurable} \quad (\sigma(X) \subset \mathcal{A}) \\ Y \text{ is } \mathcal{B}\text{-measurable} \quad (\sigma(Y) \subset \mathcal{B}) \end{array}\right\} \Rightarrow$ X and Y are independent

because $\sigma(X) \subset \mathcal{A}$, $\sigma(Y) \subset \mathcal{B}$ $\{X \in C\} \in \mathcal{A}$ $\{Y \in D\} \in \mathcal{B}$ $\Rightarrow$ they are independe

* Exercise 2.1.2

If events A and B are independent then so are $(A \text{ and } B^c)$
$(A^c \text{ and } B)$
$(A^c \text{ and } B^c)$

This implies events A and B are independent $\overset{\text{iff}}{\Longleftrightarrow}$ $\mathbb{1}_A$ and $\mathbb{1}_B$ are independent

Proof: Let A and B are independent. Prove that $\mathbb{P}(A \cap B^c) = \mathbb{P}(A)\,\mathbb{P}(B^c)$

~~$\mathbb{P}(A \cap B^c) =$~~ Have $\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c)$

$\Rightarrow \mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A \cap B) \underset{\substack{A, B \\ \text{independent}}}{=\!=} \mathbb{P}(A) - \mathbb{P}(A)\,\mathbb{P}(B) = \mathbb{P}(A)(1 - \mathbb{P}(B)) =$
$= \mathbb{P}(A)\,\mathbb{P}(B^c)$

~~$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$~~

- A, B independent $\Rightarrow$ Prove that $\mathbb{1}_A$ and $\mathbb{1}_B$ independent

$\mathbb{P}(\mathbb{1}_A \in C \text{ and } \mathbb{1}_B \in D) = \mathbb{P}\{\omega, \mathbb{1}_A(\omega) \in C \text{ and } \mathbb{1}_B(\omega) \in D\}$

check $C \cap \{0, 1\} = \{1\}$ and $D \cap \{0, 1\} = \{0\}$

since $\mathbb{P}(\mathbb{1}_A \in C) = \mathbb{P}\{\omega, \mathbb{1}_A(\omega) \in C\} = \begin{cases} \mathbb{P}(A) & \text{if } 1 \in C \text{ and } 0 \notin C \\ \mathbb{P}(A^c) & \text{if } 0 \in C \text{ and } 1 \notin C \end{cases}$

$\rightarrow \mathbb{P}(\mathbb{1}_A \in C, \mathbb{1}_B \in D) = \mathbb{P}(A B^c) \underset{A, B \text{ indep}}{=\!=} \mathbb{P}(A)\,\mathbb{P}(B^c) = \mathbb{P}(\mathbb{1}_A \in C)\,\mathbb{P}(\mathbb{1}_B \in D)$ ✓ ...

$\mathbb{1}_A = \begin{cases} 1 & \omega \in A \\ 0 & \omega \in A \end{cases}$

$\mathbb{1} : \Omega \rightarrow (S,$

$\omega \mapsto \mathbb{1}_A(\omega$

$\mathbb{1}_B(\omega$

C2 Law of large numbers
2.1 Independent

* Definition:

- Events $A_1, A_2, \dots, A_n$ are independent iff $\forall I \subset \{1, \dots, n\}$ $P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i)$
(the requirement is stronger than $P(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i)$)
- Rv's $X_1, \dots, X_n$ are independent iff $\forall$ Borel sets $B_1, \dots, B_n$,
$P(X_i \in B_i$, for all $i = 1, \dots, n) = \prod_{i=1}^{n} P(B_i)$

⊕ could take $B_i = \mathbb{R}$ for $i \neq 1, 3$
this say $P(X_1 \in B_1, X_3 \in B_3) = \prod_{i=1}^{n} P(X_i \in B_i) = P(X_1 \in B_1)\, P(\Omega)\, P(X_3 \in B_3)\, P(\Omega) \dots$
$= P(X_1 \in B_1) P(X_3 \in B_3)$.

- σ algebra $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$ are independent iff $\forall A_1 \in \mathcal{A}_1, \dots, A_n \in \mathcal{A}_n$ then $A_1, \dots, A_n$ are indep $\forall A_1, \dots, A_n$ $P(A_1 \cap \dots \cap A_n) = \prod_{i=1}^{n} P(A_i) \dots$
- We will say collection of events $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$ are independent ($\Leftrightarrow$ iff) (def)
$\forall I \subset \{1, \dots, n\}$ $P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i)$
If each $\mathcal{A}_i$ contains Ω, this equivalent to $P(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i)$, $\forall A_i \in \mathcal{A}_i$!

* Example: Suppose A and B are independent for all $B \in \mathcal{A}$ (a collection of set)
we can say that A is independent of $B \in \sigma(\mathcal{A})$
- Approach:
Let $\mathcal{L} = \{B \in \mathcal{F}, P(A \cap B) = P(A) P(B)\}$
Then $\mathcal{A} \in \mathcal{L}$. If we could show $\mathcal{L}$ is a σ algebra, then $\sigma(\mathcal{A}) \subset \mathcal{L}$
which could show $P(A \cap B) = P(A) P(B)$ $\forall B \in \sigma(\mathcal{A})$

⊕ This is hard to do, independence does not "fit well" with σ algebra requirement.
+ Check: if $B_1, B_2 \in \mathcal{L}$, does $B_1 \cap B_2 \in \mathcal{L}$
That is, is $P(A \cap \underbrace{(B_1 \cap B_2)}_{B}) \overset{?}{=} P(A)\, P(B_1 \cap B_2)$ given $\begin{cases} P(A \cap B_1) = P(A) P(B_1) \\ P(A \cap B_2) = P(A) P(B_2) \end{cases}$?

* Question: what conditions on collection of independent events $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$
will gurantee that $\sigma(\mathcal{A}_1), \sigma(\mathcal{A}_2), \dots, \sigma(\mathcal{A}_n)$ are independent?

2.1.1 Sufficient conditions for Independence : answer the question :

What conditions on collections of independent events $\mathcal{A}_1, \mathcal{A}_2, .., \mathcal{A}_n$ will guarantee that $\sigma(\mathcal{A}_1), .., \sigma(\mathcal{A}_n)$ are

* Dynkin's π-λ Theorem

* Definition:

- A collection of sets $\mathcal{A}$ is a π system if $\mathcal{A}$ is closed under finite intersections

  $A_1, A_2 \in \mathcal{A} \Rightarrow A_1 \cap A_2 \in \mathcal{A}$

- A collection of set $\mathcal{L}$ is called a λ-system if

  1) $\Omega \in \mathcal{L}$

  2) If $A, B \in \mathcal{L}$, $A \supset B$ then $A \setminus B = \{ w \in A, w \notin B \} \in \mathcal{L}$

  3) If $\left.\begin{array}{l} A_1, A_2, ... \in \mathcal{L} \\ A_n \uparrow A \quad n \to \infty \end{array}\right\} \Rightarrow A \in \mathcal{L}$ A indeed $A = \bigcup_{i=1}^{\infty} A_i$

- A σ algebra is a λ system

* Point : Sometimes, it is easier to verify a collection is a λ system instead of checking it is a σ algebra

* Theorem 2.1.2 (just need to use this to prove 2.1.3) (π-λ Dynkin's theorem).

$\left.\begin{array}{l} \text{If } \mathcal{A} \text{ is a } \pi \text{ system} \\ \quad \mathcal{L} \text{ is a } \lambda\text{-system that contains } \mathcal{A} \end{array}\right\} \Rightarrow \sigma(\mathcal{A}) \subset \mathcal{L}$

- Proof : in appendix :

* Theorem 2.1.3

Suppose $\mathcal{A}_1, ..., \mathcal{A}_n$ are independent collections, each $\mathcal{A}_i$ is a π system

$\Rightarrow \sigma(\mathcal{A}_1), .., \sigma(\mathcal{A}_n)$ are independent.

- Fact : Example :

  Suppose A is independent of all sets in $\mathcal{A}$, where $\mathcal{A}$ is a π system.

  Let $\mathcal{L} = \{ B \in \mathcal{F} \ (\sigma(\mathcal{A})), P(A \cap B) = P(A) P(B) \}$, $\mathcal{A} \subset \mathcal{L}$

  Then $\mathcal{L}$ is a λ system

  By Dynkin's theorem, $\mathcal{L}$ contains $\sigma(\mathcal{A})$

* Example (of theorem 2.1.2)

$\mathcal{A}$ is a collection of all intervals of the form $(-\infty, x]$, $x \in \mathbb{R}$

$\mathcal{L}$ is a λ system that contains $\mathcal{A}$

$\Rightarrow \mathcal{L}$ contains all Borel sets $(\underbrace{\sigma(\mathcal{A}) = \mathcal{B}) \subset \mathcal{L}}_{\text{apply DK theorem}})$

* Theorem 2.1.4 (corrolary of theorem 2.1.3)

If $\forall x_1, x_2, \dots, x_n \in (-\infty, +\infty]$
$P(X_1 \le x_1, X_2 \le x_2, \dots, X_n \le x_n) = \prod_{i=1}^{n} P(X_i \le x_i)$ $\Big\} \Rightarrow X_1, \dots, X_n$ are independent.

* Proof: $X_1, \dots, X_n : (\Omega, \mathcal{F}, P) \longrightarrow (\mathbb{R}, \mathcal{B})$.

Let $a_i \in \mathcal{F}$ be all sets of the form $a_i = \{X_i \le x_i\} = \{X_i^{-1}(-\infty, x_i]\}$.

Then $\begin{cases} \forall i = \overline{1,n}, \text{ each } a_i \text{ is a } \pi \text{ system} & (1) \\ \sigma(a_i) = \sigma(X_i) & (2) \\ \Omega \in a_i & (3) \end{cases}$

- By assumption, $P(X_1 \le x_1, \dots, X_n \le x_n) = \prod_{i=1}^{n} P(X_i \le x_i) \Leftrightarrow P(a_1, \dots, a_n) = \prod_{i=1}^{n} P(a_i)$

$\Omega \in a_i$

$\Rightarrow a_1, a_2, \dots, a_n$ are independent events (as a group) because each a_i is a π system $\Big\} \overset{DK}{\Rightarrow} \sigma(a_1), \dots, \sigma(a_n)$ are independent

* Theorem 2.1.5

Let $\mathcal{F}_{11}, \mathcal{F}_{12}, \dots, \mathcal{F}_{1m_1}$
$\mathcal{F}_{21}, \mathcal{F}_{22}, \dots, \mathcal{F}_{2m_2}$ are independent
$\mathcal{F}_{n1}, \mathcal{F}_{n2}, \dots, \mathcal{F}_{nm_n}$

$\mathcal{G}_1 = \sigma\left(\bigcup_{i=1}^{m_1} \mathcal{F}_{1i}\right)$ then $\mathcal{G}_1, \dots, \mathcal{G}_n$ are independent.

$\mathcal{G}_2 = \sigma\left(\bigcup_{i=1}^{m_2} \mathcal{F}_{2i}\right)$

$\vdots$

$\mathcal{G}_n = \sigma\left(\bigcup_{i=1}^{m_n} \mathcal{F}_{ni}\right)$

Proof Let $a_i = \bigcap_{j=1}^{m_i} A_{ij}$ where $A_{ij} \in \mathcal{F}_{ij}$
then a_i are π system, $\forall i = \overline{1,n}$
a_i contains Ω, contains $\bigcup_{j=1}^{m_i} \mathcal{F}_{ij}$) $\Big\} \xrightarrow[\text{then}]{DK} \mathcal{G}_i = \sigma(a_i), i = \overline{1,n}$ are independent

* 2.1.2) Independence, Distribution and Expectation

* Theorem 2.1.6 (Cox) Given n probability distribution $\mu_1, \dots, \mu_n$ then there exist independent $X_1, \dots, X_n$ in (Ω, F) that has distribution μ_i.

If $\mu_1, \dots, \mu_n$ are probability measure on $(\mathbb{R}_i, \mathcal{B})$.
Then there exist random variable $X_1, \dots, X_n$ such that $\begin{cases} X_i \text{ has dist } \mu_i \\ X_1, \dots, X_n \text{ are independent} \end{cases}$

* Remind $X_i : (\Omega, \mathcal{F}, P) \longrightarrow (\mathbb{R}, \mathcal{B}) \mu_i)$
μ_i is a measure on $(\mathbb{R}, \mathcal{B})$ with $\mu_i(B) = P(X_i \in B) = P(X_i^{-1}(B))$

* Proof:
- Put $\Omega = \mathbb{R}^n$, $\mathcal{B}^n := \mathcal{B}^n$, $\mu = \mu_1 \times \mu_2 \times \dots \times \mu_n$ on $(\mathbb{R}^n, \mathcal{B}^n)$
$\underset{\sim}{w} = (w_1, w_2, \dots, w_n)$ and $X_i(\underset{\sim}{w}) := X_i(w_i)$.
- We need to prove that $\begin{cases} \mu(B_i) = P(X_i \in B_i) \\ P(X_1 \in B_1, \dots, X_n \in B_n) = \prod_{i=1}^{n} P(X_i \in B_i), \forall B_i \in \end{cases}$

* $P(X_1 \in B_1, \dots, X_n \in B_n) = P\{\underset{\sim}{w} = (w_1, \dots, w_n), w_1 \in B_1, \dots, w_n \in B_n)$.
$= P(\underset{\sim}{w} = (w_1, \dots, w_n), \underset{\sim}{w} \in B_1 \times B_2 \times \dots \times B_n\}$
$= (\mu_1 \times \mu_2 \times \dots \times \mu_n)(B_1 \times B_2 \times \dots \times B_n)$
$= \prod \mu_i(B_i)$
So we have $P(X_1 \in B_1, \dots, X_n \in B_n) = \prod_{i=1}^{n} \mu_i(B_i)$ (1)
- Let fix i, and let $B_j = \mathbb{R}$ for $j \neq i$ $\Rightarrow$ X_i has law μ_i (2)
(1)+(2) $\Rightarrow$ $X_1, \dots, X_n$ are independent □

* Remind + introduction for the next theorem
Random variables $X_1, X_2, \dots$ are independent $\overset{(def)}{\Longleftrightarrow}$ If any finite collection of the r.v's are independent
- Question: given $\mu_1, \mu_2, \dots$ of probabilities measure on $(\mathbb{R}, \mathcal{B})$
Is there a probability space $(\Omega, \mathcal{F}, P)$ with r.v's $X_1, X_2, \dots$ such that
$\begin{cases} X_i \text{ has law } \mu_i \\ X_1, X_2, \dots \text{ are independent} \end{cases}$

we do need $X_1,..,X_n$ are independent, to have the result

* Theorem 2.1.7 Cox

Suppose $X_1, X_2, ..., X_n$ are independent r.v's in probability space $(\Omega, \mathcal{F}, \mathbb{P})$.
For $i = \overline{1,n}$, let μ_i be the law of X_i
Let ν be the law of random vector $(X_1, .., X_n)$ in a probability space $(\mathbb{R}^n, \mathcal{B}^n)$,
then $\nu = \mu_1 \times \mu_2 \times \cdots \times \mu_n$.

* Proof:

- Let A be a rectangle, we want to prove that $\nu(A) = \mu_1 \times \cdots \times \mu_n(A)$

$\nu(A) = \mathbb{P}((X_1, .., X_n) \in A) = \mathbb{P}(X_1 \in A_1, X_2 \in A_2, \cdots, X_n \in A_n)$, $A = A_1 \times \cdots \times A_n$

$\underset{\text{independence}}{=} \prod_{i=1}^{n} \mathbb{P}(A_i) = \prod_{i=1}^{n} \mu_i(A_i) = (\mu_1 \times \cdots \times \mu_n)(A_1 \times \cdots \times A_n) = (\mu_1 \times \cdots \times \mu_n)(A)$

- So now we want to prove that $\nu(A) = (\mu_1 \times \cdots \times \mu_n)(A)$, $\forall A \in \mathcal{B}^n$

Let $\mathcal{A} = \{A \in \mathcal{B}^n, \nu(A) = (\mu_1 \times \cdots \times \mu_n)(A)\}$, it suffices to show that $\mathcal{B}^n \subseteq \mathcal{A}$

We will show that $\mathcal{B}^n \subseteq \mathcal{A}$ by showing that $\mathcal{A}$ is a λ-system containing π-system of rectangles A, then by a theorem $\mathcal{B}^n = \sigma(A) \subset \mathcal{A}$

- Now we will show that $\mathcal{A}$ is a λ-system

① Check $\Omega \in \mathcal{A}$; check $\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}^n \in \mathcal{A}$

② Check $A, B \in \mathcal{A}$, $A \subset B$ then $(B \setminus A) \in \mathcal{A}$

since $(\mu_1, .., \mu_n)$ is a measure

$\nu(B \setminus A) = \nu(B) - \nu(A) \underset{A,B \in \mathcal{A}}{=} (\mu_1 \times \cdots \times \mu_n)(B) - (\mu_1 \times \cdots \times \mu_n)(A) = (\mu_1 \times \cdots \times \mu_n)(B \setminus A)$

③ Check $A_1, A_2 \ldots \in \mathcal{A}$, $A_i \uparrow A$ then $\nu(A_i) \uparrow \nu(A)$

$\nu(A_i) = (\mu_1 \times \cdots \times \mu_n)(A_i) \uparrow (\mu_1 \times \cdots \times \mu_n)(A_i) \neq \nu(A)$.

Then by π-λ theorem, $\mathcal{B}^n = \sigma(A) \subset \mathcal{A}$ $\square$

$$X:(\Omega,\mathcal{F}) \longrightarrow (\mathbb{R},\mathcal{B}) \xrightarrow{h} (\mathbb{R},\mathcal{B})$$
$$Y(\Omega,\mathcal{F}) \longrightarrow \mathbb{R},\mathcal{B})$$

*Theorem 2.18 (connection of independence and expectation)

Suppose X and Y are independent r.v's, X has law μ, Y has law ν

a) Let $h: \mathbb{R}^2 \to \mathbb{R}$ measurable and either — $h \geq 0$ — $h(X,Y) \in L^1(P)$ ie $E(|h(X,Y)|)<+\infty$

Then $E(h(X,Y)) = \int_{\mathbb{R}}\int_{\mathbb{R}} h(x,y)\, \mu(dx)\, \nu(dy)$

b) If $h(x,y) = f(x)g(y)$ where $f,g: \mathbb{R} \to \mathbb{R}$ and either — $f \geq 0$, $g \geq 0$ — $E(|f(X)|)<+\infty$, $E(|g(Y)|)<+\infty$

then $E(h(X,Y)) = E(f(X))\, E(g(Y))$

- If $E(|f(X)|)<+\infty$ and $E(|g(Y)|)<+\infty \Rightarrow E(|f(X)g(Y)|)<+\infty$

* Proof

$$\begin{matrix} X \ (\Omega,\mathcal{F},P) \\ Y \ (\Omega,\mathcal{F},P) \end{matrix} \longrightarrow (\mathbb{R}^2, \mathcal{B}^2, \lambda = \mu\times\nu) \xrightarrow{h} (\mathbb{R},\mathcal{B})$$

Let λ be the law of (X,Y), λ = probability measure on $(\mathbb{R}^2,\mathcal{B}^2)$, $\lambda = \mu\times\nu$

- When $h \geq 0$

Then $E(h(X,Y)) = \int_\Omega h(X,Y)\, dP \xrightarrow{\text{change variable}} \int_{\mathbb{R}^2} h\, d\lambda = \int_{\mathbb{R}^2} h\, d(\mu\times\nu) =$

$\overset{\text{Fubini}}{=} \int_{\mathbb{R}}\int_{\mathbb{R}} h(x,y)\, \mu(dx)\, \nu(dy)$

- Case $E(|h(X,Y)|)<+\infty$,

b) Suppose $f,g \geq 0$, $h(x,y) = f(x)g(y) \Rightarrow h \geq 0$

by a), $E(h(X,Y)) = \iint h(x,y)\, \mu(dx)\, \nu(dy) = \iint f(x)g(y)\, \mu(dx)\, \nu(dy) =$

$= \int g(y) \underbrace{\int f(x)\, \mu(dx)}_{E(f(X))} \nu(dy) = E(f(X)) \int g(y)\, \nu(dy)$

- The case when $E(|f(X)|)<+\infty$ and $E(|g(Y)|)<+\infty$, then apply as above

$E(|h(X,Y)|) = E(|f(X)|)\, E(|g(Y)|) < +\infty \Rightarrow h(X,Y) \in L^1$

⊕ We can repeat the argument using the second part of a)

By change of variable

$E(h(X,Y)) = \int_{\mathbb{R}^2} h\, d(\mu\times\nu) \xrightarrow[\text{because } h(X,Y)\in L^1]{\text{Fubini}} \iint h(x,y)\, \mu(dx)\, \nu(dy) = \cdots$ $\square$

* Theorem 2.1.9

- If $X_1, \dots, X_n$ are independent random variables, and either (i) $X_i \geq 0, \forall i = \overline{1,n}$ (ii) $E(|X_i|) < \infty, \forall i = \overline{1,n}$
  Then $E(\prod_{i=1}^{n} X_i) = \prod_{i=1}^{n} E(X_i)$. special case $E(X_1 X_2) = E(X_1) E(X_2)$
- We also have if $X_1, \dots, X_n \in L^1 \Rightarrow$ then $\prod_{i=1}^{n} X_i$ is $\in L^1$.

* Prove: Suppose $n = 2$, If $X, Y \geq 0$ and $f(t) = g(t) = \max\{t, 0\}$
  Theorem 2.1.8b applies for $f(X) = X$, $g(Y) = Y$ so $E(XY) = E(X) E(Y)$)
- In case $E(|X|) < +\infty$ and $E(|Y|) < +\infty$ and $f(t) = g(t) = \max\{t, 0\}$
  Theorem 2.1.8 b applies for $f(X) = X$, $g(Y) = Y$ so $E(XY) = E(X) E(Y)$

* Suppose now $n \geq 2$, we prove by induction, using $X = X_1 \dots X_n$, $Y = X_{n+1} \Rightarrow \square$

* Note: from theorem 2.1.9
- $E(X_1 X_2) = E(X_1) E(X_2) \rightarrow$ if $\left.\begin{matrix} E(X_1) < +\infty & X_1 \in L^1 \\ E(|X|_2) < +\infty & X_2 \in L^1 \end{matrix}\right\} \Rightarrow X_1 X_2 \in L^1$ (stronger)
- Cauchy Swatch $E(|XY|) \leq \|X\|_2 \|Y\|_2 \Rightarrow \left.\begin{matrix} X_1 \in L^2 \\ X_2 \in L^2 \end{matrix}\right\} \rightarrow X_1 X_2 \in L^1$ (weaker).

* Notation: If a probability measure on $(\mathbb{R}, \mathcal{B})$ has distribution function F
  $F(x) = \mu(-\infty, x]$
  it is common to write $dF(y)$ for $\nu(dy)$ $\qquad \mu(dy) = dF(y) = f(y) dy$
  $$\int_{\mathbb{R}} h(y) \mu = \int_{\mathbb{R}} h(y) \mu(dy) = \int_{\mathbb{R}} h(y) dF(y).$$

$X : F, \mu$

* Theorem 2.1.10 $\quad Y : G, \nu$
  If X, Y are independent with distribution function $F, G = \int_{\mathbb{R}} \int_{-\infty}^{z-y} f(x-y) dG(y)$
  Then $\forall z \in \mathbb{R}$, $P(X + Y \leq z) = \int_{\mathbb{R}} F(z-y) dG(y) = \int P(X \leq z-y) dG(y)$
  $= \int_{\mathbb{R}} F(z-y) \mu(dy)$

* Proof: Fix $z \in \mathbb{R}$ Define $h : \mathbb{R}^2 \longrightarrow \mathbb{R}$
  $(x,y) \longmapsto h(x,y) = \begin{cases} 1 & x + y \leq z \\ 0 & x + y > z \end{cases}$

Apply theorem 2.1.8
$$P(X + Y \leq z) = E(h(X,Y)) = \int_{\mathbb{R}} \int_{\mathbb{R}} h(x,y) \mu(dx)) \nu(dy) =$$

* Theorem 2.19 $X_1, \dots, X_n$ independent $\Rightarrow E(X_1 X_2 \dots X_n) = \prod_{i=1}^{n} E(X_i)$

- Consider $\int h(x,y)\,\mu(dx)$

For fixed y, $x \longmapsto h(x,y)$ is the function $\mathbb{1}_{(-\infty, z-y]}(x)$ because $h(x,y) = 1 \Leftrightarrow x+y \leq z \Leftrightarrow x \leq z-y$

$\Rightarrow$

$$\int h(x,y)\,\mu(dx) = \int_{\mathbb{R}} \mathbb{1}_{(-\infty, z-y]}(x)\,\mu(dx) = \int_{-\infty}^{z-y} \mu(dx) = F(z-y)$$

plug in $\Rightarrow \square$

* Recall: A probability measure μ on $(\mathbb{R}, \mathcal{B})$ has $f: \mathbb{R} \to \mathbb{R}$ as a density if $\forall B \in \mathcal{B}$, $\mu(B) = \int_B f(x)\,dx$ (Lebesgue integer)

If X has law μ, this is $P(X \in B) = \int_B f(x)\,dx = \int_{\mathbb{R}} \mathbb{1}_B(x) f(x)\,dx$

- Fact If X has density f, then $\forall$ Borel measurable $h: \mathbb{R} \to \mathbb{R}$ such that

i) $h \geq 0$

ii) $E(|h(X)|) < +\infty$

Then $E(h(X)) = \int h(x) f(x)\,dx$

$\qquad\qquad \| $

$\int_{\mathbb{R}} h\,d\mu \quad \mu$: law of X

- Informally, "$\mu(dx) = f(x)\,dx$" $= dF(x)$

* The prove for the above inequality
  - i) true for $h(X) = \mathbb{1}_B$
  - ii) true for simple function h
  - iii) true for nonnegative h
  - iv) true for general h

* To prove that X has density f, it is suffices to prove that $P(X \leq z) = \mu(-\infty, z] = \int \mathbb{1}_{(-\infty, z]}) = \int \mathbb{1}_{(-\infty, z]} f(x)\,dx = \int_{-\infty}^{z} f(x)\,dx$ $\forall z \in \mathbb{R}$

* Theorem 2.1.11

Suppose X and Y are independent with distribution functions F and G
X has density f

Then $(X+Y)$ has density $\phi(z) \overset{(a)}{=} \int f(z-y)\,dG(y) \underset{g}{\overset{Y \text{ has density}}{=}} \int f(z-y)\,g(y)\,dy \overset{def}{=} \underbrace{(f*g)}_{\text{convolution}}$ (b)

Define ϕ as in (a)

* Prove a) We want to prove that $(X+Y)$ has density ϕ $\Rightarrow$ it suffices to prove that

$$P(X+Y \le z) = \int_{-\infty}^{z} \phi(x')\,d(x') = \int_{-\infty}^{z}\left[\int_{R} f(x'-y)\,dG(y)\right] d(x')$$

• By previous theorem,

$$(*)\ P(X+Y \le z) = \int_R F(z-y)\,dG(y) = \int_R \int_{-\infty}^{z-y} f(x)\,dx\,dG(y) \overset{\Theta}{=} \int_R \underbrace{\left(\int_R 1\!\!1_{(-\infty,z-y)}(x) f(x)\,dx\right)}_{\text{Lebesgue measurable}} dG(y)$$

Θ Lebesgue measurable is translation invariant, we can change variables in the integral to compute $\int_R 1\!\!1_{(-\infty,z-y)}(x) f(x)\,dx$

take y fixed, put $x' = x+y \Rightarrow x = x'-y$ and $dx' = dx$

$$1\!\!1_{(-\infty,z-y]}(x) = 1\!\!1_{(-\infty,z-y)}(x'-y) = 1\!\!1_{(-\infty,z)}(x')$$

then $\int_R 1\!\!1_{(-\infty,z-y)}(x) f(x)\,dx = \int_R 1\!\!1_{(-\infty,z)}(x') f(x'-y)\,dx'$

• Put it back to $(*)$, we have

$$P(X+Y \le z) = \int_R \left[\int_R 1\!\!1_{(-\infty,z)}(x')\,(f(x'-y)\,d(x')\right] dG(y)$$

$$\underset{\text{Fubini}}{=} \int_R \left(1\!\!1_{(-\infty,z)}\left(\int_R f(x'-y)\,dG(y)\right)\right) d(x')$$

$$= \int_{-\infty}^{z}\left(\int_R f(x'-y)\,dG(y)\right) d(x') = \int_{-\infty}^{z} \phi(x')\,d(x') \quad \square\ a)$$

* Sum of a discrete and a continuous random variable.

* Suppose X discrete
Y continuous

$$P(X+Y<z) = \sum_{k} P(Y<z-X) = \sum_{k} P(Y<z-x \mid X=x) P(X=x).$$

$$= \sum_{X=x} F_Y(z-x) \, P(X=x)$$

$\Rightarrow$ density of z $\quad f_Z(z) = \sum_{\ell \in K} f_Y(z-\ell)\, p_\ell$

* Definition

Random variables X and Y are uncorrelated (def) iff $\begin{cases} E(|X|), E(|Y|), E(|XY|) \text{ finite.} \\ E(XY) = E(X)\,E(Y) \end{cases}$

same as $\underbrace{E((X-EX)(Y-E(Y)))}_{\text{covariance}(X,Y)} = 0$

• (Usually, we assume $E(|X|^2)$, $E(|Y|^2)$ are finite, which implies $E(|X|)$, $E(|Y|)$, $E(|XY|)$ are finite since by Holder inequality $E(|XY|) \le \|X\|_2 \|Y\|_2$

* Prove that $E(XY) = E(X)E(Y) \Rightarrow E((X-EX)(Y-E(Y))) = 0$

$$E((X-EX)(Y-EY)) = E(XY - YE(X) - XE(Y) + EX\,EY)$$

$$= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y)$$

$$= E(XY) - E(X)E(Y) = 0.$$

$\Rightarrow$ Independent random variables are uncorrelated.

$$E\left(\left|\frac{B_n^c}{n}-b_c\right|\right) \leq E\left(\left|\frac{B_n^c}{n}\right|\right)+b_c=\frac{1}{n}E(|B_n^c|)+|b_c| \leq \frac{1}{n}\sum_{i=1}^{n}\underbrace{(E\,Y_i^c)}_{\varepsilon_c}+|b_c|$$

$$=\varepsilon_c+|b_c| \leq 2\varepsilon_c$$

$$\limsup E\left(\left|\frac{B_n^c}{n}-b_c\right|\right) \leq 2\lim_{c\to\infty}\varepsilon_c=0$$

- To sum up

$$S_n \leq \underbrace{E\left(\left|\frac{A_n^c}{n}-a_c\right|\right)}_{\longrightarrow 0}+\underbrace{E\left(\left|\frac{B_n^c}{n}-b_c\right|\right)}_{\longrightarrow 0} \Rightarrow S_n \longrightarrow 0$$

* Question

If $E(|X_1|)=+\infty$, can we generalize to: are there constants $\mu_n \in \mathbb{R}$, s.t. $\frac{S_n}{n}-\mu_n \xrightarrow{\text{prob}}$

---

* Theorem:

Let $X_1, X_2, \dots$ be iid r.v's such that $\lim_{t\to\infty}|t|\,P(|X|>t)=0$

Then if $\mu_n=E\{X_1\mathbb{1}_{\{|X_1|\leq n\}}\}$ then $\frac{S_n}{n}-\mu_n \longrightarrow 0$ in probability.

---

* Lemma 2.2.8

For any r.v Y, for $p>0$, $E(|Y|^p)=p\int_0^\infty x^{p-1}P(|Y|\geq x)\,dx$

* Note: • we can replace "$>$" with "$\geq$"

• Common case is $p=1$, $E(|Y|)=\int_0^\infty P(|Y|\geq x)\,dx=\int_0^1(1-F_{|Y|}(x))\,dx$

So we have: if $\int_a^\infty P(|Y|>x)\,dx=\int_a^1(1-F_{|Y|}(x))\,dx \iff E(|Y|)<\infty$

* Proof

$$p\int_0^\infty x^{p-1}P(|Y|>x)\,dx = p\int_0^\infty x^{p-1}\underbrace{E(\mathbb{1}_{(|Y|>x)})}_{=\int_\Omega \mathbb{1}_{|Y|>x}\,dP}\,dx \overset{\text{Fubini}}{=} E\left(\int_0^\infty p x^{p-1}\mathbb{1}_{\{|Y|>x\}}\,dx\right.$$

$$=E\left[\int_0^{|Y(\omega)|}px^{p-1}\,dx\right]=E\left[x^p\Big|_0^{|Y(\omega)|}\right]=E(|Y|^p) \quad \square$$

* Theorem 2.2.7 (Followed from Theorem 2.2.6 ← Triangular array version of WLOLN)

Let $X_1, X_2, \dots$ be iid r.v's, s.t. $\lim_{x\to\infty} x P(|X|>x) = 0$ (*)

If $\mu_n = E(X_1 \mathbb{1}_{\{|X_1|\le n\}})$

Then $\frac{S_n}{n} - \mu_n \xrightarrow{P} 0$

* Notes:

① (*) holds if $E(|X_1|) < +\infty$ by DCT

$E(|X_1| \mathbb{1}_{\{|X_1|>x\}}) \xrightarrow[DCT]{x\to\infty} 0$ $\qquad E(|X_1| \mathbb{1}_{\{|X_1|>x\}}) \geq E(x \mathbb{1}_{\{|X_1|>x\}}) = x P(|X_1|$

② (*) does not implies $E(|X_1|) < +\infty$

* Proof:

Assume (*), take $0<p<1$, then choose $a>0$ such that $x P(|X|>x) \le 1$, $\forall x \geq$

$$E(|X_1|^p) = p\int_0^\infty x^{p-1} P(|X_1|>x)\,dx \le \underbrace{p\int_0^a x^{p-1} P(|X_1|>x)\,dx}_{\text{finite}} + p\int_a^\infty x^{p-1}\frac{1}{x}\,dx$$

$$\int_a^\infty x^{p-1}\frac{1}{x}\,dx = \int_a^\infty x^{p-2}\,dx = \int_a^\infty \frac{1}{x^{2-p}}\,dx < \infty \quad \text{where } 2-p>1 \iff 1>p$$

So (*) implies $E(|X|^p) < +\infty$ only when $p<1$

• Let X has d.f. $F(x) = \begin{cases} 1 - \frac{e}{x\log x}, & x \geq e \\ 0 & x < e \end{cases}$

Then $P(X_1 > x) = \frac{e}{x\log x}$ for $x > e$

$$E(X_1) = \int_0^\infty P(|X_1|>x)\,dx = \int_0^e P(X_1>x)\,dx + \int_e^\infty P(X_1>x)\,dx = e + \int_e^\infty \frac{e}{x\log x}\,dx$$

$$= e\left(1+\int_e^\infty \frac{dx}{x\log x}\right)$$

put $u = \log x$, $du = \frac{1}{x}dx$

then $\int_e^\infty \frac{dx}{x\log x} = \int \frac{du}{u} = \log u = \log(\log x)$

$$= e\left(1 + \log\log x\Big|_0^\infty\right) = \infty.$$

which means (*) does not mean $E(X_1) < +\infty$

$\frac{S_n}{n} \xrightarrow{P} \mu$ (WLOLN) $\quad \frac{S_n}{n} \xrightarrow{a.s} \mu$ (Strog law of large number)

2.3 Borel Cantelli Lemmas

$B_m \downarrow B = \bigcap B_m \Rightarrow P(B_m) \downarrow P(B)$

* Def (limsup, liminf)

Let $A_1, A_2, A_3, \ldots$ Define $\limsup A_n = \bigcap_{m=1}^{\infty} \underbrace{\boxed{\bigcup_{n=m}^{\infty} A_n}}_{B_m \downarrow} = \lim_{m\to\infty} \left(\bigcup_{n=m}^{\infty} A_n\right) \overset{\text{notal}}{=\!=} \{A_n \text{ i}\}$ infinitely of

$\limsup A_n = \{w,$ w belongs to infinitely many of the $A_n\}$

If $w \in$ infinitely many A_n then $w \in \limsup A_n$

If $w \in \limsup A_n$, then $w \in \bigcup_{n=m}^{\infty} A_n$, $\forall m$

choose n_1 s.t. $w \in A_{n_1}$, consider $w \in \bigcup_{n=n_1+1}^{\infty} A_n$, choose $n_2 \geq n_1 + 1$ such that $w \in$

continue, produce $n_1 < n_2 < n_3 < \cdots$ s.t $w \in A_{n_l}$, for $l = 1, 2, \ldots \infty$

so $w \in$ infinitely many A_n

- Note $1\!\!1_{\limsup A_n}(w) = \limsup_{n\to\infty} 1\!\!1_{A_n}(w)$ for all w

If $\limsup 1\!\!1_{A_n}(w) = 1$, then $\{w \in \{A_n \text{ i.o}\}\}$, so $1\!\!1_{\limsup A_n}(w) = 1$

If $\limsup 1\!\!1_{A_n}(w) = 0$, then there exists $N = N(w)$ s.t $1\!\!1_{A_n}(w) = 0$, $\forall n \geq N$

which means $w \notin \{A_n \text{ i.o}\}$ so $1\!\!1_{\limsup A_n} = 0$

- If $X = \sum_{n=1}^{\infty} 1\!\!1_{A_n}$ then $\begin{cases} \{X = +\infty\} = \limsup A_n \\ \{X < +\infty\} = (\limsup A_n)^c \end{cases}$

- $P\left(\bigcup_{n=m}^{\infty} A_n\right) \downarrow P(A_n \text{ i.o.})$ as $m \to \infty$

* $\liminf A_n = \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n$

* Theorem 2.3.1 (1st Borel Cantelli lemma)

Let $A_1, A_2, \dots$ be a sequence of events.

If $\sum_{n=1}^{\infty} P(A_n) < +\infty$ then $\implies$ $\mathbb{P}(A_n \text{ io}) = 0$

* Proof

Let $X(w) = \sum_{n=1}^{\infty} \mathbb{1}_{A_n}(w)$, takes values in $[0, +\infty]$

By Fubini, to $\Omega \times \mathbb{N}$, P continuity measure

$E(X) = E(\sum_{n=1}^{\infty} \mathbb{1}_{A_n}) = \sum_{n=1}^{\infty} E(\mathbb{1}_{A_n}) = \sum_{n=1}^{\infty} P(A_n) < +\infty$

$E(X) < +\infty$ which implies X is finite a.s, i.e, $P(X < +\infty) = 1$ or $P(X = +\infty) = ($

$\{A_n \text{ i.o}\} = \{X = +\infty\}$

* Theorem 2.3.5 (Strong law of large number)

Suppose r.v's $X_1, X_2, \dots$ are iid $E(|X_1|^4) < +\infty$ } Then $\frac{S_n}{n} \xrightarrow{a.s} \mu$

If $\mu = E(X_1)$

* Proof

1) $\{\lim \frac{S_n}{n} = \mu\} = \bigcap_{j=1}^{\infty} \{\limsup |\frac{S_n}{n} - \mu| \leq \frac{1}{j}\}$

check $\subset$, $\supset$,

① Let $w \in$ RHS $\Rightarrow \limsup |\frac{S_n(w)}{n} - \mu| \leq \frac{1}{j}$, $\forall j = 1, 2, 2, \dots$

$\Rightarrow \limsup |\frac{S_n(w)}{n} - \mu| = 0$ or $\lim_{n \to \infty} \frac{S_n(w)}{n} = \mu$.

② The set $\{\limsup |\frac{S_n}{n} - \mu| \leq \frac{1}{j}\}$ is $\downarrow$ in j, so

so $\underbrace{P\left(\bigcap_{j=1}^{\infty} (\limsup |\frac{S_n}{n} - \mu| < \frac{1}{j}\right)}_{P\left(\lim |\frac{S_n}{n} - \mu| = 0\right)} = \lim_{j \to \infty} P\left(\limsup |\frac{S_n}{n} - \mu| < \frac{1}{j}\right)$

To prove $\frac{S_n}{n} \xrightarrow{a.s} \mu$, it suffices to prove that $\forall \varepsilon > 0$, $P\left(\limsup_{n \to \infty} |\frac{S_n}{n} - \mu| \leq \varepsilon\right) =$

By Borel Cantelli lemma, it suffices to prove

$\sum_{n=1}^{\infty} P(|\frac{S_n}{n} - \mu| > \varepsilon) < \infty$ (~~because this would show~~ $P\left(\{|\frac{S_n}{n} - \mu| > \varepsilon \text{ io}\}^c\right) = 1$

and if $w \notin \{|\frac{S_n(w)}{n} - \mu| > \varepsilon\}$ then $\exists N = N(w)$ s.t

$|\frac{S_n}{n} - \mu| \leq \varepsilon$, $\forall n \geq N \Rightarrow \limsup |\frac{S_n(w)}{n} - \mu|$

$\frac{n!(n-2)!}{2!}$

This shows $\left\{\left|\frac{S_n}{n}-\mu\right|>\varepsilon \text{ i.o.}\right\}^c \subset \left\{\limsup\left|\frac{S_n}{n}-\mu\right|\le\varepsilon\right\}$

$\Rightarrow P\left\{\left|\frac{S_n}{n}-\mu\right|>\varepsilon \text{ i.o.}\right\}^c = 1 \Rightarrow P\left(\limsup\left|\frac{S_n}{n}-\mu\right|\le\varepsilon\right)=1$.

- Now we will prove that $\sum_{n=1}^{\infty} P\left(\left|\frac{S_n}{n}-\mu\right|>\varepsilon\right)<+\infty$

Idea: Try to estimate $P\left(\left|\frac{S_n}{n}-\mu\right|>\varepsilon\right) \underset{\text{Chebyshev}}{\le} \frac{Var\left(\frac{S_n}{n}\right)}{\varepsilon^2} = \frac{\frac{1}{n^2}Var(S_n)}{\varepsilon^2} =$

$= \frac{n\,Var(X_1)}{n^2\varepsilon^2} = \frac{E(|X_1-\mu|^2)}{n\varepsilon^2}$.

(we have $\sum \frac{Var(X_1)}{n^2\varepsilon^2} = +\infty$ not good).

2nd try: $P\left(\left|\frac{S_n}{n}-\mu\right|^4>\varepsilon^4\right) \le P\left(\left|\frac{S_n}{n}-\mu\right|^4 > \varepsilon^4\right) \le \frac{E\left(\left(\frac{S_n}{n}-\mu\right)^4\right)}{\varepsilon^4}$

$\underset{\text{will show}}{\le} \frac{C}{\varepsilon^4 n^2}$

- Let $Y_i = X_i - \mu$ then $S_n - n\mu = \sum_{i=1}^{n} Y_i$, the $\{Y_i\}$ are iid $E(Y_1)=0$, $E(Y_i^4)<\infty$

$P\left(\left|\frac{S_n}{n}-\mu\right|>\varepsilon\right) = P\left(|S_n-n\mu|>n\varepsilon\right) = P\left(|\sum Y_i|^4>(n\varepsilon)^4\right) \underset{\text{Markov}}{\le} \frac{E(|\sum Y_i|^4)}{n^4\varepsilon^4}$

- where $E(|\sum Y_i|^4) = E\left(\sum_{i=1}^{n} Y_i \sum_{j=1}^{n} Y_j \sum_{k=1}^{n} Y_k \sum_{l=1}^{n} Y_l\right) = \sum_{i,j,k,l=1}^{n} E(Y_iY_jY_kY_l)$

There are four kinds of terms in this case sum

1. $i=j=k=l$, n of these, $E(Y_iY_jY_kY_l) = E(Y_i^4) = E(Y_1^4)$
2. $i\ne j\ne k\ne l$ $E(Y_iY_jY_kY_l) = E(Y_i)E(Y_j)E(Y_k)E(Y_l) = 0$
3. there are two distinct indices, one occurs 3 times, the other occurs 1 time

   $E(Y_iY_jY_kY_l) = 0$

4. there two distinct indices, one occurs 2 times, the other occurs 2 times

   $E(Y_iY_iY_jY_j) = E(Y_i^2)E(Y_j^2) \ge [E(Y_i^2)]^2 = (\|Y_i\|_2)^4 \le (\|Y_i\|_4)^4 =$

   $= E(Y_1^4)$

How many terms are there

1) choose i', j' distinct $i', j' \in \{1,\dots,n\}$ $\frac{(n)(n-1)}{2} = \binom{n}{2}$ ways.

2) choose 2 position in 4 ~~4 3 2 1~~ $\binom{4}{2} = 6$ ways.

$\Big\}$ $6\binom{n}{2}$ ways $\frac{6(n)(n-1)}{2}$ ways.

1 2 __ __ __ __

* The contribution from these terms

$E(|\sum Y_i|^4) \leq n E(Y_1^4) + 3(n)(n-1) E(Y_1^4) = (3n^2 - 2n) E(Y_1^4) \leq Cn^2$

* And so,

$P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \leq \frac{Cn^2}{n^4\varepsilon^4} = \frac{C}{n^2\varepsilon^4}$ $\cancel{< +\infty}$ $\square$

$\Rightarrow \sum_{n=1}^{\infty} P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \leq \frac{C}{\varepsilon^4} \sum_{n=1}^{\infty} \frac{1}{n^2} < +\infty$

* Theorem 2.3.6 (2nd Borel Cantelli lemma)

If $A_1, A_2, \ldots$ be independent events

If $\sum_{n=1}^{\infty} P(A_n) = +\infty \xrightarrow{\text{then}} P(A_n \text{ i.o}) = 1$

* Proof

1) If events $E_n \uparrow \Gamma$ (or $\downarrow \Gamma$) then $P(\Gamma_n) \uparrow P(\Gamma)$ (or $P(\Gamma_n) \downarrow P(\Gamma)$)

2) By assumption $\sum_{n=m}^{\infty} P(A_n) = +\infty$, $\forall m = 1, 2, \ldots$

3) $1 - u \leq e^{-u}$ for all $u \geq 0$

put $f(u) = e^{-u} + u - 1$

$f'(u) = -e^{-u} + 1 \geq 0$ $\quad f'(u) = 0$ when $e^{-u} = 1 \rightarrow u = 0$

$\Rightarrow f(u) \geq f(0) = 1 + 0 - 1 = 0$

- Recall $\{A_n \text{ i.o}\} = \bigcap_{m=1}^{\infty} \underbrace{\bigcup_{n=m}^{\infty} A_n}$ and $\bigcup_{n=m}^{\infty} A_n \downarrow \{A_n \text{ i.o}\}$ as $m \rightarrow \infty$

- It suffices to show that $P(\bigcup_{n=m}^{\infty} A_n) = 1$ for all m

$\oplus$ $P(\bigcup_{n=m}^{\infty} A_n) = 1 - P\left(\left(\bigcup_{n=m}^{\infty} A_n\right)^c\right) = 1 - \left(P\left(\bigcap_{n=m}^{\infty} A_n^c\right)\right)$

$\Rightarrow$ It suffices to show that $P(\bigcap_{n=m}^{\infty} A_n^c) = 0$, $\forall m$

- Now we show $P(\bigcap_{n=m}^{\infty} A_n^c) = 0$, it suffices to show that $P\left(\bigcap_{n=m}^{N} A_n^c\right) \downarrow 0$ by (1)

By independence, $P\left(\bigcap_{n=m}^{N} A_n^c\right) = \prod_{n=m}^{N} P(A_n^c) = \prod_{n=m}^{N} (1 - P(A_n)) \underset{\text{by (3)}}{\leq} \prod_{n=m}^{N} e^{-P(A_n)} =$

we need independence

$= e^{-\sum_{n=m}^{N} P(A_n)} \downarrow 0$ as $N \longrightarrow \infty$

Can $\frac{S_n}{n} \xrightarrow{n\to\infty}$ something $\in \mathbb{R}$ possibly random if $E(|X_1|)<+\infty$ Theorem $\Rightarrow$ no

* Theorem 2.3.7

If $X_1, X_2, \dots$ are iid with $E(|X_1|) = +\infty$ then $P(|X_n| \geq n \text{ io}) = 1$.

If $S_n = X_1 + \dots + X_n$ Then $P(\omega, \lim_{n\to\infty} \frac{S_n(\omega)}{n} \text{ exist in } \mathbb{R}) = 0$

* Proof

① $E(|X_1|) = \int_0^\infty P(|X_1|>t)\,dt = \sum_{n=0}^{\infty} \int_n^{n+1} P(|X_1|>t)\,dt \leq$

$\leq \sum_{n=0}^{\infty} \int_n^{n+1} P(|X_1|>n)\,dt = \sum_{n=0}^{\infty} P(|X_1|>n)$ $\Rightarrow$ ~~$P(|X_n|>n) = 1$~~ $\sum_{n=0}^{\infty} P(|X_1|>n) = \infty$

Since $E(|X_1|) = \infty$

② We have $\sum_{n=0}^{\infty} P(|X_1|>n) = \sum_{n=0}^{\infty} P(|X_n|>n) = +\infty$ $\xrightarrow{2^{nd}BC}$ $P(|X_n| \geq n \text{ io}) = 1$.

event

Because $\{(|X_n|>n)\}$ events are independent

* Prove $P(\omega, \lim \frac{S_n(\omega)}{n} \text{ exists in } \mathbb{R}) = 0$

Let $\Omega_0 = \{\omega, \lim \frac{S_n(\omega)}{n} \text{ exists in } \mathbb{R}\}$

$\Omega_1 := \{\omega, |X_n| \geq n \text{ io}\}$

We know $P(\Omega_1) = 1$ and for all $\omega \in \Omega_1$, $\limsup_{n\to\infty} \frac{|X_n(\omega)|}{n} \geq 1$

- Now we have

$$\frac{S_n(\omega)}{n} - \frac{S_{n+1}(\omega)}{(n+1)} = \frac{S_n(\omega)}{n} - \frac{S_n(\omega)+X_{n+1}(\omega)}{n+1} = S_n(\omega)\left(\frac{1}{n} - \frac{1}{n+1}\right) - \frac{X_{n+1}(\omega)}{(n+1)}$$

$$= \frac{S_n(\omega)}{n(n+1)} - \frac{X_{n+1}(\omega)}{(n+1)}$$

- If $\omega \in \Omega_0 \cap \Omega_1$, then $\lim \frac{S_n(\omega)}{n(n+1)} = \lim \frac{S_n(\omega)}{n} \frac{1}{(n+1)} = 0$

Then $\limsup \left|\frac{S_n(\omega)}{n} - \frac{S_{n+1}(\omega)}{n+1}\right| = \limsup \left|\frac{S_n(\omega)}{n(n+1)} - \frac{X_{n+1}(\omega)}{(n+1)}\right| = \limsup \left|\frac{X_{n+1}(\omega)}{n+1}\right| =$

since $\omega \in \Omega_1$

contradiction

$\Rightarrow \Omega_0 \cap \Omega_1 = \phi \Rightarrow \Omega_0 \subset \Omega_1^c$ But $P(\Omega_1^c) = 0$ since $P(\Omega_1) = 1$ $\square$

since $\omega \in \Omega_0$ then $\limsup \left|\frac{S_n(\omega)}{n} - \frac{S_{n+1}(\omega)}{n+1}\right|$

* In section 2.3, there are a lot of interesting application of BC lemmas.

* Example 2.3.3

Let $X_1, X_2, \ldots$ be iid $P(X_1=0) = P(X_1=1) = \frac{1}{2}$ - $E(X) = np(1-p) =$

(Fair coin is tossed)

What can we say about the length of the longest consecutive of 1 in the first $X_1, \ldots$

Answer: For any n,

Let l_n: length of consecutive of 1s ending at n

$$l_n = \max\{m \le n, X_n = X_{n-1} = \cdots = X_{n-m+1} = 1\}$$

1 2 3 4 5 6 7 8 9

0 0 1 1 0 1 1 1 0 Let $L_n = \max l_k$, then $L_9 = 3$

l_i $l_1 = 0$ 0 1 2 0 1 2 3 0

* Prop $\lim_{n\to\infty} \frac{L_n}{\log_2 n} = 1$ $\log_2 1000 \simeq 10$ $\log_2 1000000 \simeq 20$

• Note $l_n \ge k = \{1 = X_n = X_{n-1} = \cdots = X_{n-k+1}\}$

$l_n = k = \{1 = X_n = X \quad = X_{n-k+1} = 0\}$

$P(l_n \ge k) = 2^{-k}$ $P(l_n = k) = \frac{1}{2^{k+1}}$

We will prove that, for any $\varepsilon > 0$

I) $\sum_{n=1}^{\infty} P(l_n \ge (1+\varepsilon)\log_2 n) < +\infty$ $\xrightarrow[\text{lemma}]{1^{st}\text{ BC}}$ $\Rightarrow P(l_n \ge (1+\varepsilon)\log_2 n \text{ io})$ $\Rightarrow P(L_n \ge (1+\varepsilon)\log_2 n \text{ io}) = ($; $\limsup \frac{L_n}{\log_2 n} < 1+\varepsilon$ a.s

II) $\sum_{n=1}^{\infty} P(L_n < (1-\varepsilon)\log_2 n) < +\infty$ $\longrightarrow$ $\liminf \frac{L_n}{\log_2 n} > 1-\varepsilon$ a.s ; $\Rightarrow P(L_n < (1-\varepsilon)\log_2 n \text{ io}) = 0$

why does this imply $\lim \frac{L_n}{\log_2 n} = 1$ a.s ?

$\to \forall$ any ε, $\Omega_\varepsilon = \{w, \limsup \frac{L_n}{\log_2 n} < 1+\varepsilon\} = 1$ $P(\Omega_\varepsilon) = 1$

$\to$ $\{w, \liminf \frac{L_n}{\log_2 n} > 1-\varepsilon\} = 1.$

For $\Omega_{i,j}$ then $P(\Omega_{i,j}) = 1 \to P(\bigcap_{j=1}^{\infty} \Omega_{ij}) = 1$

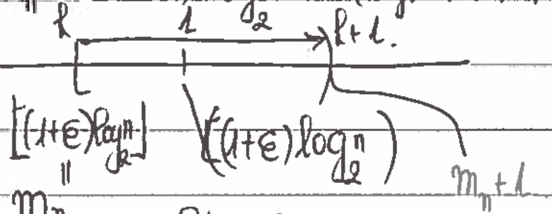
If $w \in \bigcap_{j=1}^{\infty} \Omega_{ij}$ then $\lim \frac{L_n}{\log_2 n} = 1.$

* Now we prove (I) $\sum_{n=1}^{\infty} P(L_n \geq (1+\epsilon)\log_2 n) \leq +\infty$

Choose n_0 such that $\frac{2\log_2(2n)}{n} \leq \frac{1}{2}$ for all $n \geq n_0$

Define $m_n = \lfloor (1+\epsilon)\log_2 n \rfloor$ (Greatest integer $\leq (1+\epsilon)\log_2 n$)

Note



we have the inequality $(1+\epsilon)\log_2 n - 1 \leq m_n \leq (1+\epsilon)\log_2 n < m_n + 1$

- For $n \geq n_0$, $m_n = \lfloor (1+\epsilon)\log_2 n \rfloor \leq (1+\epsilon)\log_2 n \leq 2\log_2 n \leq n$

Now $P(\ell_n \geq (1+\epsilon)\log_2 n) \leq P(\ell_n \geq m_n) = \frac{1}{2^{m_n}} \leq 2^{-((1+\epsilon)\log_2 n - 1)} =$

$= 2 * 2^{-(1+\epsilon)\log_2 n} = 2 * n^{-(1+\epsilon)} = \frac{2}{n^{(1+\epsilon)}}$

This implies $\sum_{i=1}^{\infty} P(\ell_n \geq (1+\epsilon)\log_2 n) < +\infty$

By BC lemma $P(\ell_n \geq (1+\epsilon)\log_2 n \text{ i.o}) = 0$. } we need to prove that

- To prove that $P(L_n \geq (1+\epsilon)\log_2 n \text{ i.o}) = 0$

we need to prove $\{w : L_n(w) \geq (1+\epsilon)\log_2 n \text{ i.o}\} \subset \{w, \ell_n(w) \geq (1+\epsilon)\log_2 n \text{ i.o}\}$

increasing function of n

To prove this, suppose $w \notin \{w, \ell_n(w) > (1+\epsilon)\log_2 n \text{ i.o}\}$, we want to prove that $\{w, L_n(w) \geq (1+\epsilon)\log_2 n \text{ i.o}\}$

means, suppose $w \in \{w, \ell_n(w) > (1+\epsilon)\log_2 n \text{ i.o}\}^c \subset \{w, L_n(w) \geq (1+\epsilon)\log_2 n \text{ i.o}$

w, then $\exists N_1(w) < +\infty$, $\forall k \geq N_1(w)$, $\ell_k(w) < (1+\epsilon)\log_2 n$

Let $N_2(w) = 2^{N_1(w)}$, so that $(1+\epsilon)\log_2 N_2(w) \geq \log_2 N_2(w) = N_1(w)$
$N_2 > N_1$

consider, $L_n(w) = \max\{\ell_1(w), \dots, \ell_n(w)\}$

$= \max\{\ell_1(w), \dots, \ell_{N_1}(w)\} \vee \max\{\ell_{N_1+1}(w), \dots, \ell_n(w)\}$

$\leq N_1(w) \vee \max\{\ell_k(w), N_1(w)+1 \leq k \leq n\}$

$\quad < (1+\epsilon)\log_2 k$ (by def of N_1

$\leq$ ~~$N_1(w) \vee (1+\epsilon)\log_2 n$~~ ... because $(1+\epsilon)\log_2 n > N_1$ if $n \geq N_2$.

$< (1+\epsilon)\log_2 n \vee \max\{\ell_k(w), N_1(w)+1 \leq k \leq n\}$

$< (1+\epsilon)\log_2 k \vee (1+\epsilon)\log_2 n = (1+\epsilon)\log_2 n$

That is, for all $n \ge N_2(w)$, $L_n(w) < (1+\epsilon)\log_2 n \Rightarrow w \in \{L_n(w) \ge (1+\epsilon)\log_2 n$ i

$*$ Now we want to prove II, $P(L_n < (1-\epsilon)\log_2 n \text{ i.o}) = 0$

For $i \ge$ let $m_n = \lceil (1-\epsilon)\log_2 n \rceil$

$m_n - 1 < (1-\epsilon)\log_2 n \le m_n \le (1-\epsilon)\log_2 n + 1$

$m_n - 1$ $\quad (1-\epsilon)\log_2 n \quad$ $\lceil (1-\epsilon)\log_2 n \rceil = m_n$

Define "intervals" I_ℓ as follow

$I_1 = \{1, 2, \dots, m_n\}$, $I_2 = \{m_n + 1, \dots, 2m_n\}$, $I_\ell = \{(\ell-1)m_n + 1, \dots, \ell m_n\}$

$0 \quad m_n \quad 2m_n \quad 3m_n \quad (k_n-1)m_n \quad n \quad k_n m_n$ $\quad = \{j \in \mathbb{N}, (\ell-1)m_n < j \le \ell m_n\}$

Define $k_n \in \mathbb{N}$, to satisfy $n \in I_{k_n}$

$L_n \ge m_n$ if $x_j = 1$ for all $j \in I_\ell$ for any $\ell \le k_n - 1$

If let $A_\ell = \{x_j = 1$, for all $j \in I_\ell\}$ then $A_1, A_2, \dots$ are independent and $P(A_\ell) = \left(\frac{1}{2}\right)^{m_n}$

Thus, $L_n \ge m_n$ on $\bigcup_{j=1}^{k_n-1} A_j$ $\quad (L_n(w) \ge m_n$ if $w \in \bigcup A_j)$

$$P(L_n < m_n) \le P\left(\left(\bigcup_{j=1}^{k_n-1} A_j\right)^c\right) = P\left(\bigcap_{j=1}^{k_n-1} A_j^c\right) = \prod_{j=1}^{k_n-1} P(A_j^c) = \prod_{j=1}^{k_n-1}\left(1 - \left(\frac{1}{2}\right)^{m_n}\right)$$

$$= \left(1 - \left(\frac{1}{2}\right)^{m_n}\right)^{k_n-1} \le \exp\left(-\left(\frac{1}{2}\right)^{m_n}\right)^{k_n-1} = \exp\left\{-(k_n-1)2^{-m_n}\right\}$$

$$\le \exp\left\{-(k_n-1)\, n^{-(1-\epsilon)}\right\} \quad \left(\text{because } 2^{-m_n} \ge 2^{-(1-\epsilon)\log_2 n} = 2^{\log_2(n^{-(1-\epsilon)})} = \right.$$

$$\le \exp\left\{-\frac{n-1}{m_n}\, n^{-(1-\epsilon)}\right\} \le \exp\left\{-\frac{(n-1)}{(1-\epsilon)\log_2 n + 1}\, n^{-1} n^{\epsilon}\right\}$$

(since $\frac{n}{2} \le n-1$ for $n \ge 2$)

$$\le \exp\left\{-\frac{n}{2}\,\frac{1}{(1-\epsilon)\log_2 n}\, n^{-1} n^{\epsilon}\right\}$$

$$= \exp\left\{-\frac{n^{\epsilon}}{2(1-\epsilon)\log_2 n}\right\}$$

Then $P(L_n < m_n) \le e^{-\frac{n^{\epsilon}}{2(1-\epsilon)\log_2 n}}$

$P(L_n < (1-\epsilon)\log_2 n) \le e^{-\frac{n^{\epsilon}}{2(1-\epsilon)\log_2 n}}$

$\Rightarrow \sum_{n=1}^{\infty} P(L_n < (1+\epsilon)\log_2 n) \le \sum_{n=1}^{\infty} e^{-\frac{n^{\epsilon}}{2(1-\epsilon)\log_2 n}} = \sum_{n=1}^{\infty} e^{-n^{\delta}}$ $\quad$ δ can be big small

For fixed $\delta > 0$, $\exists N$ s.t $n \ge N$, $n^{\delta} \ge 2\log n$

then $\sum_{n=N}^{\infty} e^{-n^{\delta}} < \sum_{n=N}^{\infty} e^{-2\log n} = \sum_{n=N}^{\infty} e^{\log(n^{-2})} = \sum_{n=N}^{\infty} \frac{1}{n^2} < +\infty$.

2.4* Strong Law of Large numbers

If $X_1, X_2, \dots$ are iid, $E(|X_1|) < +\infty$, $\mu = E(X_1)$, $S_n = X_1 + X_2 + \dots + X_n$, $\Big\}$ $\lim_{n\to\infty} \frac{S_n}{n} = \mu$ a.s

* Definition (Truncation)

Define $Y_k = X_k \mathbb{1}_{\{|X_k| \leq k\}}$, truncated r.v's, truncation level is not constant with k.

Define $T_n = Y_1 + Y_2 + \dots + Y_n$

* Lemma 2.4.2

If $\lim_{n\to\infty} \frac{T_n}{n} = \mu$ a.s then $\lim_{n\to\infty} \frac{S_n}{n} = \mu$ a.s.

* Proof of Lemma 2.4.2 Use B-C

- Since $E(|X_1|) = \int_0^\infty P(|X_1| > t)\,dt > \sum_{k=1}^{\infty} P(|X_1| > t)$ and $E(|X_1|) < +\infty$

$$\sum_{k=1}^{\infty} P(X_k \neq Y_k) = \sum_{k=1}^{\infty} P(|X_1| > k) < +\infty$$

By BC, $P(X_k \neq Y_k \text{ i.o}) = 0$

For some Ω_0, $P(\Omega_0) = 1$, $\omega \in \Omega_0$, $\exists N_1(\omega) < +\infty$ s.t $X_k(\omega) = Y_k(\omega)$, $\forall k \geq N_1$

- Now compare $\frac{S_n(\omega)}{n} - \frac{T_n(\omega)}{n} = \sum_{k=1}^{N_1(\omega)} \frac{X_k(\omega) - Y_k(\omega)}{n} + \sum_{k=N_1}^{n} \frac{X_k(\omega) - Y_k(\omega)}{n}$ $\to 0$

i.e, $$\frac{S_n(\omega)}{n} - \frac{T_n(\omega)}{n} = \frac{1}{n}\underbrace{\left[\sum_{k=1}^{N_1(\omega)} X_k(\omega) - Y_k(\omega)\right]}_{\text{finite sum of finite numbers}} \xrightarrow{n\to\infty} 0$$

Thus, for all $\omega \in \Omega_0$, $\lim_{n\to\infty}\left(\frac{S_n(\omega)}{n} - \frac{T_n(\omega)}{n}\right) = 0$

Thus, $\frac{T_n}{n} \xrightarrow{a.s} \mu$ then $\frac{S_n(\omega)}{n} \xrightarrow{a.s} \mu$

sequence of number (not random)

* Lemma

$$\boxed{\text{If } \lim_{k\to\infty}\left(\frac{T_k}{k}-\frac{E(T_k)}{k}\right)=0 \text{ a.s}, \text{ then } \lim_{k\to\infty}\frac{T_k}{k}=\mu \text{ a.s}}$$

* Proof of the Lemma: we want to prove that $\frac{E(T_k)}{k}\longrightarrow\mu$

• $\frac{1}{k}E(T_k)=\frac{1}{k}\sum_{j=1}^{k}E(Y_j)=\frac{1}{k}\sum_{j=1}^{k}E\left(X_j \mathbb{1}_{|X_j|\le j}\right)=\frac{1}{k}\sum_{j=1}^{k}E\left(X_1 \mathbb{1}_{|X_1|\le j}\right)$

$\xrightarrow{\text{as } k\to\infty}\mu$ if $E\left(X_1 \mathbb{1}_{|X_1|\le j}\right)\xrightarrow[j\to\infty]{a.s}\mu$ (Lemma $\lim a_n=a$ then $\frac{1}{n}\sum_{j=1}^{n}a_j\to a$ but converse fails)

• So, now we want to prove $E\left(X_1 \mathbb{1}_{|X_1|\le j}\right)\xrightarrow{a.s}\mu$

We have $E\left(X_1\mathbb{1}_{|X_1|\le j}\right)\longrightarrow E(X_1)=\mu$ as $j\to\infty$ by DCT

$\Big($ $X_1\mathbb{1}_{|X_1|\le j}(w)\longrightarrow X_1(w)$ as $j\to\infty$

and $\left|X_1\mathbb{1}_{|X_1|<j}\right|\le|X_1|$ and $E(|X_1|)<+\infty$ $\Big)$

* Lemma 2.4.4

$$\boxed{\text{For any } y>0,\quad y\sum_{k>y}\frac{1}{k^2}\le 2 \text{ (any constant } C \text{ would be fine)}}$$

(If $y>0$, this is $\sum_{k>y}\frac{1}{k^2}\le\frac{C}{y}$, approximately, $\int_y^\infty\frac{1}{x^2}=\frac{1}{y}$)

* Lemma 2.4.3 If $E(|X_i|^2)<+\infty$ then considering $Var(T_n)=\sum_{i=1}^{n}Var(Y_i)\to\infty$

$\sum_{k=1}^{\infty}\frac{Var(Y_k)}{k^2}<+\infty$

In fact $\frac{Var(T_n)}{n^2}\longrightarrow 0$

since? $n\,Var(T_n)\le n^2\varepsilon_n$, ε_n-

is the consequence of $E(|X_1|)<+\infty$

• Sketch the prove

$0\le Var(Y_k)=E(Y_k^2)-(E(Y_k))^2\le E(Y_k^2)=\int_0^\infty 2x\,P(|Y_k|\ge x)\,dx$

• Let $(k(n))$ be a subsequence of (n)

We will show $\frac{T_{k(n)}}{k(n)}\longrightarrow\mu$ a.s.

$k(1)$ $k(2)$ $k(3)$

(Not $\sum\frac{1}{n}=\infty$ but $\sum_{1}^{\infty}\frac{1}{k(n)}<+\infty$ if $k(n)\overset{\text{say}}{=}2^n$

1 2 3 4 5 6 [illegible] 8

• Fix $\varepsilon > 0$, given some $k(n)$

$$\sum_{n=1}^{\infty} P\left(|T_{k(n)} - ET_{k(n)}| \geq \varepsilon k(n)\right) \leq \sum_{n=1}^{\infty} \frac{Var(T_{k(n)})}{\varepsilon^2 k^2(n)} = \sum_{n=1}^{\infty} \frac{1}{\varepsilon^2 k^2(n)} \sum_{j=1}^{k(n)} Var($$

$$= \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} \frac{1}{\varepsilon^2 k^2(n)} Var(Y_j) 1\!\!1_{\{j \leq k(n)\}} \overset{\text{Fubini}}{=} \frac{1}{\varepsilon^2} \sum_{j=1}^{\infty} Var(Y_j) \sum_{n,\, k(n) \geq j} \frac{1}{k^2(n)}$$

Chose $k(n)$, fix α, $1 < \alpha < +\infty$,

put $k(n) = \lfloor \alpha^n \rfloor \geq \frac{\alpha^n}{2}$ (If $x \geq 1$, then $\lfloor x \rfloor \geq \frac{x}{2}$ check)

~~Note that $k(n+1) - k(n) = \lfloor \alpha^{n+1} \rfloor - \lfloor \alpha^n \rfloor \geq \frac{\alpha^{n+1}}{2}$~~

note that $\lim_{n\to\infty} \frac{k(n+1)}{k(n)} = \lim_{n\to\infty} \frac{\lfloor \alpha^{n+1} \rfloor}{\lfloor \alpha^n \rfloor} = \alpha$ (check this) $\quad \lfloor x \rfloor \quad x \quad \lfloor x \rfloor + 1$

• Define $n_0 = n_0(\alpha, j)$ to be the smallest interger such that $\lfloor \alpha^n \rfloor \geq j$

then with $k(n) = \lfloor \alpha^n \rfloor$

$$\sum_{n=1}^{\infty} P\left(\left|T_{k(n)} - E(T_{k(n)})\right| \geq \varepsilon k(n)\right) \leq \frac{1}{\varepsilon^2} \sum_{j=1}^{\infty} Var(Y_j) \left(\sum_{n=n_0(\alpha,j)}^{\infty} \frac{1}{k^2(n)}\right)$$

$$\sum_{n=n_0}^{\infty} \frac{1}{k^2(n)} = \sum_{n=n_0}^{\infty} \frac{1}{(\lfloor \alpha^n \rfloor)^2} \leq \sum_{n=n_0}^{\infty} \frac{2}{(\alpha^n)^2} = 2 \sum_{n=n_0}^{\infty} \frac{1}{(\alpha^2)^{n-n_0+n_0}} = \frac{2}{(\alpha^2)^{n_0}} \sum_{k=0}^{\infty} \left(\frac{1}{\alpha^2}\right)^k \quad k = n$$

$$= \frac{2}{\alpha^{2n_0}} \left(\frac{1}{1 - (1/\alpha^2)}\right) \qquad (n_0 : \text{integer depended on } j)$$

Thus, $\displaystyle\sum_{n=1}^{\infty} P\left(|T_{k(n)} - ET_{k(n)}| \geq \varepsilon k(n)\right) \leq \frac{1}{\varepsilon^2} \sum_{j=1}^{\infty} Var(Y_j) \frac{1}{(1-\alpha^{-2})} \frac{2}{(\alpha^2)^{n_0(j)}}$

? sthing about $\alpha^{n_0} \geq j \iff \frac{1}{\alpha^{n_0}} \leq \frac{1}{j}$ $\qquad \displaystyle\leq \frac{2}{\varepsilon^2} \left(\frac{1}{1 - \frac{1}{\alpha^2}}\right) \sum_{j=1}^{\infty} Var(Y_j) \frac{1}{j^2} < \infty$

by lemma 2.4.3

By B-C, this shows $P\left(\left\{\frac{T_{k(n)} - E(T_{k(n)})}{k(n)} > \varepsilon \text{ io}\right\}\right) = 0$

This show $\frac{T_{k(n)} - E(T_{k(n)})}{k(n)} \xrightarrow[n\to\infty]{a.s} 0$

1, Durrett 1.1.1 (Do not turn in)

i, Prove that if F_i, $i \in I$ are σ-algebras then $\bigcap_{i \in I} F_i$ is a σ-algebra

ii, Use the result in i, to show that if we are given a set Ω and $\mathcal{A}$ = a collection of subsets of Ω then there is a smallest σ-algebra containing $\mathcal{A}$

(we call this the σ-algebra generated by $\mathcal{A}$, denoted by $\sigma(\mathcal{A})$)

2) Durret 1.1.2 (Do not turn in)

Let $\Omega = \mathbb{R}$,

$\mathcal{F}$ = all subsets so that A or A^c is countable,

$\mathbb{P}(A) = 0$ in the first case and $= 1$ in the second

Show that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space.

* Durrett 1.2.1 (Random variable)

Suppose X and Y are random variables on $(\Omega, \mathcal{F}, P)$

$A \in \mathcal{F}$

Show that if $Z(\omega) = \begin{cases} X(\omega), & \omega \in A \\ Y(\omega) & \omega \in A^c \end{cases}$

then Z is a random variable.

(Be careful with this problem since $A \in \mathcal{F}$ and we consider $B \in \mathcal{B}$).

We want to prove that Z is a random variable

$\Rightarrow$ NTP that $Z^{-1}(B) \in \mathcal{F}$, $\forall B \in \mathcal{B} \Leftarrow$ Borel σ-algebra

$$Z^{-1}(B) = \{\omega \in \Omega, Z(\omega) \in B\} = \{\omega \in A \cup A^c, Z(\omega) \in B\}.$$

$$= \{\omega \in A, \overset{X}{Z}(\omega) \in B\} \cup \{\omega \in A^c, \overset{Y}{Z}(\omega) \in B\}$$

$$= \{\overset{X}{Z}{}^{-1}(B) \cap A\} \cup \{\overset{Y}{Z}{}^{-1}(B) \cap A^c\}$$

(note that $A \in \mathcal{F}$, $B \in \mathcal{B}$, the above $\neq \cancel{Z^{-1}(B \cap A)}$) wrong.

$$= \{X^{-1}(B) \cap A\} \cup \{Y^{-1}(B) \cap A^c\}$$

$$= \underbrace{\underbrace{X^{-1}(B)}_{\substack{\in \mathcal{F} \\ \text{since } X \text{ is a r.v}}} \cup \underbrace{Y^{-1}(B)}_{\substack{\in \mathcal{F} \\ \text{since } Y \text{ is a r.v}}}}_{\in \mathcal{F} \text{ since } \sigma\text{-algebra is closed under countable union.}}$$

* Durrett 1.2.2

Let $Z \sim N(0,1)$

Use theorem 1.2.3 to get upper and lower bound on $P(Z \geq 4)$

(Theorem 1.2.3: For $x > 0$, $(x^{-1} - x^{-3}) e^{-\frac{x^2}{2}} \leq \int_x^{+\infty} e^{-\frac{y^2}{2}} dy \leq x^{-1} e^{-\frac{x^2}{2}}$

We have $P(Z \geq 4) = \int_4^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$,

by theorem 1.2.3,

$$\frac{1}{\sqrt{2\pi}} (x^{-1} - x^{-3}) e^{-\frac{x^2}{2}} \leq \int_4^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \leq \frac{1}{\sqrt{2\pi}} x^{-1} e^{-\frac{x^2}{2}} \quad (\text{where } x = 4).$$

$$\frac{1}{\sqrt{2\pi}} \left(\frac{1}{4} - \frac{1}{4^3}\right) e^{-\frac{4^2}{2}} \leq P(Z \geq 4) \leq \frac{1}{\sqrt{2\pi}} \left(\frac{1}{4}\right) e^{-\frac{4^2}{2}}$$

$$\Rightarrow \quad 0.0000314 \leq P(Z \geq 4) \leq 0.0000335.$$

$F : (\mathbb{R}, \mathcal{B}) \to \mathbb{R}$
$x \longmapsto F(x) = P(X \le x) = P(X^{-1}(-\infty, x])$

* Durret 1.2.3

(cdf $F(x) = P(X \le x) = P(X^{-1}(-\infty, x])$)

Show that a distribution function has at most countably many discontinuities.

* Let $D = \{x \in \mathbb{R},\ F$ is discontinuous at $x\}$

* We want to prove that D is at most countably

We want to prove that $D = \bigcup_n D_n$ for some D_n finite

countable union

- Claim: $x \in D \Rightarrow P(X = x) > 0$

Prove claim: $x \in D \Rightarrow F$ is not continuous at x

$\Rightarrow P(X = x) = P(X \le x) - P(X < x) =$

$= F(x) - F(x^-) > 0$

- Since $x \in D \Rightarrow P(X = x) > 0$

Then let $D_n = \{x \in D,\ P(X = x) = \frac{1}{n} > 0\}$

Because $P(\Omega) = 1 < +\infty$, there can be at much n elements in D_n. (2)

- Claim $D = \bigcup_{n=1}^{\infty} D_n$ (1)

⊕ Prove $\bigcup_{n=1}^{\infty} D_n \subset D$: this is because $D_n \subset D,\ \forall n$.

⊕ Prove $D \subseteq \bigcup_{n=1}^{\infty} D_n$

Let $x \in D \Rightarrow P(X = x) > 0$

we have $\forall \varepsilon > 0, \exists n_0 \in \mathbb{N},\ P(X = x) = \frac{1}{n_0} = \varepsilon > 0 \Rightarrow x \in D_{n_0} \subseteq \bigcup_{n=1}^{\infty} D_n$

From (1) and (2) we have that D is at most countable

$\Rightarrow F$ has at most countably discontinuity

* $\mathcal{H}$ is closed under pointwise limit since

Durrett 1.3.7 Let $(f_n) \subset \mathcal{H}$, $f_n(w) \to f(w)$, $\forall w \Rightarrow f(w) = \lim_{n\to\infty} f_n(w) \Rightarrow f$ is measurable $\to f \in \mathcal{H}$.

A function $\varphi : (\Omega, \mathcal{F}) \longrightarrow \mathbb{R}$ is said to be simple if

$$w \longmapsto \varphi(w) = \sum_{m=1}^{n} c_m \mathbb{1}_{A_m}(w) \text{ where } c_m \in \mathbb{R}, A_m \in \mathcal{F},$$

Show that the class of $\mathcal{F}$ measurable functions is the smallest class containing the simple functions and closed under pointwise limit.

Let $\mathcal{H} = \{f, f \text{ is } \mathcal{F} \text{ measurable function}\}$

we need to prove that i) Simple functions contained in $\mathcal{H}$.

ii) every $f \in \mathcal{H}$ is a pointwise limit of simple functions

iii) $\mathcal{H}$ is the smallest class that has above properties

* Prove i) Let $\varphi(w) = \sum_{m=1}^{n} c_m \mathbb{1}_{A_m}(w)$, we NTP that $\varphi \in \mathcal{H}$

Note that since φ only receive finite values in $\mathbb{R}$, we can assume that these $\{A_m\}_{m=1}^n$ are pairwise disjoint.

then let $x \in \mathbb{R}$, $\varphi^{-1}(-\infty, x] = \{w, \sum_{m=1}^{n} c_m \mathbb{1}_{A_m}(w) \le x\} =$

$$\overset{\text{disjoint}}{=} \{w, c_m \le x, \forall x = \overline{1,n}\} = \underbrace{\bigcup_{i=1}^{k} \underbrace{A_i}_{\in \mathcal{F}}}_{\in \mathcal{F}}$$

where $\{A_i\}_{i=1}^k$ are pairwise disjoint set where $c_i \le x$.

$\Rightarrow \varphi$ is in $\mathcal{H}$. done for i)

* For every $f \in \mathcal{H}$, prove that $\exists \{f_n\}$ simple function, $f_n \to f$

Define $f_n : \Omega \to \mathbb{R}$

$$w \longmapsto f_n(w) = \begin{cases} n & , f(w) \ge n \\ m\frac{1}{2^n}, & \frac{m}{2^n} \le f(w) < \frac{(m+1)}{2^n}, \quad \frac{-n}{2^n} \le m < \frac{n}{2^n} \\ -n & , f(w) \le n \end{cases}$$

Then $|f(w) - f_n(w)| \overset{n \text{ large enough}}{=} f(w) - m2^{-n} < \frac{(m+1)}{2^n} - \frac{m}{2^n} = 2^{-n} \xrightarrow{n\to\infty} 0$

this means $f_n \to f$ pointwise

* Prove iii) Let $\mathcal{H}^*$ be another class of measurable functions that contains all simple functions and closed under pointwise limit.

Let $\varphi \in \mathcal{H}$, $\Rightarrow \exists \varphi_n$ simple such that $\varphi_n \to \varphi$

But since φ_n simple $\Rightarrow \varphi_n \in \mathcal{H}^*$

$\Rightarrow \varphi \in \mathcal{H}^*$ since $\mathcal{H}^*$ is closed under pointwise limit.

$\Rightarrow \mathcal{H} \subseteq \mathcal{H}^*$.

Tran Le

$\frac{9}{10}$

Mat 721 Problem Set 1, due 1/31

Refer to theorems in Durrett by number, as in Theorem 2.1.14.

1. Durrett 1.1.1 – Do not turn in.

2. Durrett 1.1.2: Do not turn in.

3. Durrett 1.1.5: Let $\mathcal{F}_1, \mathcal{F}_2, \ldots$ be σ-algebras of subsets of Ω such that $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$, and let $\mathcal{F} = \cup_i \mathcal{F}_i$. (i) Prove that $\mathcal{F}$ is an algebra. (ii) Give an example to show that $\mathcal{F}$ need not be a σ-algebra.

4. Durrett 1.3.1 Let $(\Omega, \mathcal{F})$ and $(S, \mathcal{S})$ be measurable spaces, and let $X : \Omega \to S$. Suppose $\mathcal{A}$ is a collection of subsets of S such that $\sigma(\mathcal{A}) = \mathcal{S}$. Prove that $X^{-1}(\mathcal{A}) = \{\{X \in A\} : A \in \mathcal{A}\}$ generates $\sigma(X) = \{\underbrace{\{X \in B\}}_{X^{-1}(B)} : B \in \mathcal{S}\}$.

5. Durrett 1.3.9: Let $(\Omega, \mathcal{F})$ be a measurable space. A function $\phi : \Omega \to \mathbb{R}$ is called simple if
$$\phi(\omega) = \sum_{m=1}^{\infty} c_m 1_{A_m}(\omega)$$
where each $c_m \in \mathbb{R}$ and each $A_m \in \mathcal{F}$. Show that every $\mathcal{F}$-measurable function f is the pointwise limit of simple functions. *Hint.* Define
$$\phi_n = \sum_{k=-n2^n}^{n2^n} \frac{k}{2^n} 1_{\{k2^{-n} \le f < (k+1)2^{-n}\}}$$

6. Let Ω be a nonempty set and let $\mathcal{A}$ be an algebra of subsets of Ω. Let $\mu : \mathcal{A} \to [0, 1]$ be a finitely additive measure on $\mathcal{A}$. Prove that μ is a countably additive measure on $\mathcal{A}$ if and only if: for every sequence (A_n) of elements of $\mathcal{A}$ such that $A_n \downarrow \emptyset$ as $n \to \infty$, $\mu(A_n) \downarrow 0$ as $n \to \infty$.

7. Let $\mathcal{S}$ be a semialgebra of sets. (a) Prove directly that if $A_1, \ldots, A_n \in \mathcal{S}$ then there are finitely many $B_1, \ldots, B_m \in \mathcal{S}$, pairwise disjoint, such that $\cup_{i=1}^n A_i = \sum_{j=1}^m B_j$. (b) Suppose that $\mu : \mathcal{S} \to [0, \infty]$ such that $\mu(\emptyset) = 0$ and μ is finitely additive on $\mathcal{S}$. Prove that μ has the monotonicity property: if $A, A_1, \ldots, A_n \in \mathcal{S}$ and $A \subset \cup_{i=1}^n A_i$ then $\mu(A) \le \sum_{i=1}^n \mu(A_i)$. *Hint:* $\cup_{i=1}^n A_i = A \cup \left(\cup_{i=1}^n (A_i \cap A^c)\right)$. Use part (a).

3) Durret 1.1.5

Let $\mathcal{F}_1, \mathcal{F}_2, \ldots$ be σ-algebras of subsets of Ω

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

Let $\mathcal{F} = \bigcup_i \mathcal{F}_i$

i) Prove that $\mathcal{F}$ is an algebra

ii) Give an example, to show that $\mathcal{F}$ need not to be a σ-algebra

i) Prove that $\mathcal{F}$ is an algebra $\Rightarrow$ we need to prove $\begin{cases} ① \ \Omega \in \mathcal{F} \\ ② \ \mathcal{F} \text{ is closed under taking complem} \\ ③ \ \mathcal{F} \text{ is closed under finite unio} \end{cases}$

* Prove (1)

Since $\mathcal{F}_1, \mathcal{F}_2 \ldots$ σ algebra $\Rightarrow \Omega \in \mathcal{F}_i, \forall i \Rightarrow \Omega \in \bigcup_i \mathcal{F}_i = \mathcal{F}$

* Prove (2) Let $A \in \mathcal{F}$, we need to prove $A^c \in \mathcal{F}$

$$\left.\begin{array}{l} \text{Let } A \in \mathcal{F} = \bigcup_i \mathcal{F}_i \Rightarrow \exists i_0 \text{ so that } A \in \mathcal{F}_{i_0} \\ \text{we also have } \mathcal{F}_{i_0} \text{ is a } \sigma\text{-algebra} \end{array}\right\} \Rightarrow A^c \in \mathcal{F}_{i_0} \Rightarrow A^c \in \bigcup_i \mathcal{F}_i$$

* Prove (3), Let $A_1, A_2, \ldots, A_n \in \mathcal{F}$. We need to prove $\bigcup_{i=1}^n A_i \in \mathcal{F}$

$A_1 \in \mathcal{F} \Rightarrow \exists i_1$ so that $A_1 \in \mathcal{F}_{i_1} \xrightarrow{\mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots} A_1 \in \mathcal{F}_j, \forall j \geq i_1$

$A_2 \in \mathcal{F} \Rightarrow \exists i_2$ so that $A_2 \in \mathcal{F}_{i_2} \Longrightarrow A_2 \in \mathcal{F}_j, \forall j \geq i_2$

$\vdots$

$A_n \in \mathcal{F} \Rightarrow \exists i_n$ so that $A_n \in \mathcal{F}_{i_n} \xrightarrow{\mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots} A_n \in \mathcal{F}_j, \forall j \geq i_n$

So chose $i = \max\{i_1, i_2, \ldots, i_n\}$, we have $A_1, A_2, \ldots, A_n \in \mathcal{F}_j, \forall j \geq i$

$\Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{F}_j, \forall j \geq i$

$\Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{F} = \bigcup_i \mathcal{F}_i$

ii) Give an example to show that $\mathcal{F}$ need not to be a σ-algebra

We let $X = \mathbb{N}$,

Define $\mathcal{F}_n = \sigma$ algebra that contain all subset of $\{1, \ldots, n\}$.

Now let $x_i = \{2i\}$ then $x_i \in \bigcup \mathcal{F}_i, \forall i$

however $\bigcup\{x_i\} \notin \mathcal{F}_i \Rightarrow \bigcup\{x_i\} \notin \bigcup \mathcal{F}_i \Rightarrow$ $\mathcal{F} = \bigcup \mathcal{F}_i$ is not a σ alg

4) Durrett 1.3.1

Let $(\Omega, \mathcal{F})$ and $(S, \mathcal{S})$ be measurable spaces.

Let $X : \Omega \longrightarrow S$

Suppose $\mathcal{A}$ is a collection of subsets of S such that $\sigma(\mathcal{A}) = \mathcal{S}$.

Prove that $X^{-1}(\mathcal{A}) = \{\{X \in A\} : A \in \mathcal{A}\}$ generates $\sigma(X) = \{\{X \in B\}, B \in \mathcal{S}\}$

* We have

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

So we need to prove $\sigma(X^{-1}(\mathcal{A})) = \sigma(X) = X^{-1}(\mathcal{S})$.

* Prove $\sigma(X^{-1}(\mathcal{A})) \subset X^{-1}(\mathcal{S})$

$$\mathcal{A} \subset \mathcal{S} \Rightarrow X^{-1}(\mathcal{A}) \subset X^{-1}(\mathcal{S}) \Rightarrow \sigma(X^{-1}(A)) \subset \sigma(X^{-1}(\mathcal{S})) \overset{X^{-1}(\mathcal{S}) \text{ is a } \sigma\text{-algebra}}{=\!=} X^{-1}(\mathcal{S})$$

* Prove $X^{-1}(\mathcal{S}) \subset \sigma(X^{-1}(\mathcal{A}))$

$$X^{-1}(\mathcal{S}) \overset{\mathcal{S} = \sigma(\mathcal{A})}{=\!=} X^{-1}(\sigma(\mathcal{A})) = X^{-1}\Big(\bigcap_{I} a\Big) = \bigcap_{I} \underbrace{X^{-1}(a)}_{\sigma \text{ algebra that contain } X^{-1}(\mathcal{A})} = \sigma(X^{-1}(a))$$

I = the σ-algebras containing $\mathcal{A}$

57 Durrett 1.3.9.

Let $(\Omega, \mathcal{F})$ be a measurable space

A function $\phi : \Omega \longrightarrow \mathbb{R}$ is called a simple function (def) if $\phi(w) = \sum_{m=1}^{\infty} c_m \mathbb{1}_{A_m}(w)$

for each $c_m \in \mathbb{R}$ and $A_m \in \mathcal{F}$

Show that every $\mathcal{F}$ measurable function is a pointwise limit of simple functions

* Define $\phi_n : \Omega \longrightarrow \mathbb{R}$ by

$$w \longmapsto \phi_n(w) = \begin{cases} n, & \phi(w) \geq n \\ \frac{l}{2^n}, & \frac{l}{2^n} \leq \phi(w) < \frac{l+1}{2^n}, \quad -n2^n \leq l < n2^n \\ -n, & \phi(w) \leq n \end{cases}$$

* Then now we will prove that $\phi_n \xrightarrow[\text{pointwise}]{n\to\infty} \cancel{\phi}$ F

we have $|\phi_n(w) - \phi(w)| \underset{\text{large enough}}{\overset{\text{when } n}{=}} \left|\phi(w) - \frac{k}{2^n}\right| < \frac{l+1}{2^n} - \frac{l}{2^n} = \frac{1}{2^n} \xrightarrow[n\to\infty]{} 0$

which means $\phi_n \xrightarrow[\text{pointwise}]{n\to\infty} \phi$ $\square$

$n \geq n_0$ where $|F(w)| < n_0$

6) Let Ω be a nonempty set

$\mathcal{A}$ be an algebra of subset of Ω

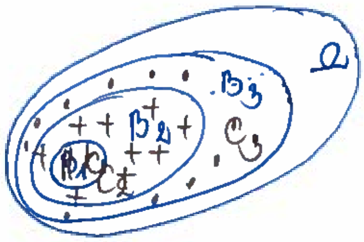
Let $\mu: \mathcal{A} \longrightarrow [0,1]$ be a finite additive measure on $\mathcal{A}$

Prove that μ is a countablly additive measure on $\mathcal{A}$ iff for every sequence $(A_n) \in \mathcal{A}$ such that $A_n \downarrow \phi$ as $n \to \infty$, then $\mu(A_n) \downarrow 0$ as $n \to \infty$

$(\Longrightarrow)$: Have: μ is countablly additive measure on $\mathcal{A}$

Need to prove: $\forall \{A_n\}$ in $\mathcal{A}$, $A_n \downarrow \phi$ as $n \to \infty$, then $\mu(A_n) \downarrow 0$ as $n \to \infty$

- First we will prove that for $\forall \{B_n\}$ in $\mathcal{A}$, $B_n \uparrow \Omega$ as $n \to \infty$, then $\mu(B_n) \uparrow 1$ as $n \to \infty$



Let $C_1 = B_1$

$C_2 = B_2 \setminus B_1$

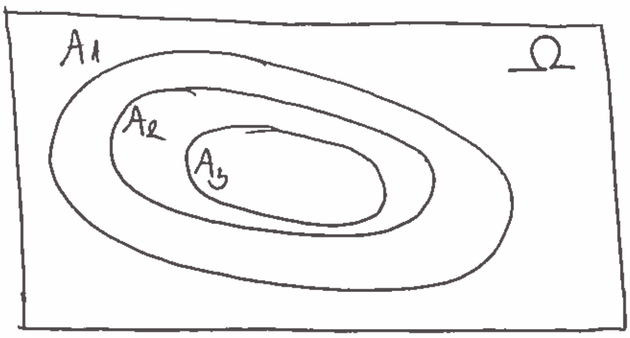
$C_3 = B_3 \setminus B_2$

$\vdots$

then we have $\{C_n\}_{n=1}^{\infty}$ are disjoint sets in $\mathcal{A}$ and $\bigcup_{n=1}^{\infty} C_n = \Omega$.

$$\Rightarrow 1 = \mu(\Omega) = \mu\left(\bigcup_{n=1}^{\infty} C_n\right) \underset{\mu \text{ is countablly additive}}{=\!=\!=} \sum_{n=1}^{\infty} \mu(C_n) = \lim_{m \to \infty} \sum_{n=1}^{m} C_n = \lim_{m \to \infty} B_m$$

- Second, we will use the result that we have just proved to prove that $\forall \{A_n\}_{n=1}^{\infty}$ in $\mathcal{A}$, $A_n \downarrow \phi$ as $n \to \infty$ then $\mu(A_n) \downarrow 0$ as $n \to \infty$



- Note that since $\mu: \mathcal{A} \longrightarrow [0,1]$

$\Rightarrow \mu(\Omega) = 1 < +\infty$

- We have $(\Omega \setminus A_n) \uparrow (\Omega \setminus \phi)$ as $n \to \infty$

$(\Omega \setminus A_n) \uparrow \Omega$ as $n \to \infty$

From above $\Rightarrow \mu(\Omega \setminus A_n) \uparrow 1$ as $n \to \infty$

or $\lim_{n\to\infty} \mu(\Omega \setminus A_n) = 1 \Rightarrow \lim_{n\to\infty} [\mu(\Omega) - \mu(A_n)] = 1$

$$\underset{\mu(\Omega) < \infty}{\Longrightarrow} \mu(\Omega) - \lim_{n\to\infty} \mu(A_n) = 1$$

$$\lim_{n\to\infty} A_n = 1 - \mu(\Omega) = 0$$

$\Leftarrow$ Have: For every sequence $(A_n) \in \mathcal{A}$, $A_n \downarrow \phi$ as $n \to \infty$ then $\mu(A_n) \downarrow 0$ as $n \to \infty$
Need to prove μ is countable additive measure on $\mathcal{A}$.

* Let $A = \bigcup\limits_{i=1}^{\infty} A_i$ (disjoint), we want to prove that $\mu(A) = \sum\limits_{i=1}^{\infty} \mu(A_i)$

- We first observe that $\mu(A_i) \downarrow 0$ as $n \to \infty$ since if $\mu(A_i) \not\downarrow 0$ then $\mu(A) \uparrow \infty$

- So now we rearrange $\{A_i\}$ so that $A = \bigcup\limits_{\substack{\text{pairwise} \\ \text{disjoint}}} A_i$ where $\mu(A_i) \downarrow 0$ as $n \to \infty$

- Let $B_i = A \setminus \left(\bigcup\limits_{k=1}^{i} A_k\right)$ then we have $B_i \downarrow 0$

⊕ From the assumption, since $B_i \downarrow 0$ as $n \to \infty$, we have $\mu(B_i) \downarrow 0$ as $n \to \infty$

⊕ Now consider $A \setminus B_i = \bigcup\limits_{k=1}^{i} A_k$

$$\Rightarrow \underset{\parallel}{\mu(A \setminus B_i)} = \mu\Big(\underbrace{\bigcup_{k=1}^{i} A_k}_{\text{p. disjoint}}\Big) \overset{\mu \text{ is finite}}{\underset{\text{additive}}{=}} \sum_{k=1}^{i} \mu(A_k)$$

$$\mu(A) - \mu(B_i)$$

$$\Rightarrow \mu(A) - \mu(B_i) = \sum_{k=1}^{i} \mu(A_k)$$

$$\lim_{i \to \infty} \mu(A) = \lim_{i \to \infty} \sum_{k=1}^{i} \mu(A_k) + \underbrace{\lim_{i \to \infty} \mu(B_i)}_{=0}$$

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{k=1}^{\infty} \mu(A_k). \quad \square$$

$\boxed{\begin{array}{l} \text{* Durret 1.5.1?} \\ \text{Let } \|f\|_\infty = \inf\{M,\ \mu\{x, |f(x)| > M\} = 0\}. \\ \text{Prove that } \int |fg|\,d\mu \leq \|f\|_1 \|g\|_\infty \end{array}}$

* $\|g\|_\infty = \inf\{M,\ \mu\{x, |g(x)| > M\} = 0\}$.

Then $\exists$ E such that $|g| < \|g\|_\infty$, $\forall x \in E$.

and E^c has $\mu(E^c) = 0$ and $|g| > \|g\|_\infty$, $\forall x \in E^c$

$$\int |fg|\,d\mu = \int |f||g| \left(\mathbb{1}_E(x) + \mathbb{1}_{E^c}(x)\right) d\mu$$

$$= \underbrace{\int \mathbb{1}_E |f||g|\,d\mu}_{< \int \mathbb{1}_E |f| \|g\|_\infty d\mu} + \underbrace{\int \mathbb{1}_{E^c} |f||g|\,d\mu}_{= 0 \text{ since } \mu(E) = 0}$$

since $\|g\|_\infty > |g|$, $\forall x \in E$.

$$= \|g\|_\infty \int \mathbb{1}_E |f|\,d\mu = \|g\|_\infty \int_\Omega |f|\,d\mu = \|g\|_\infty \|f\|_1.$$

* We know if $\|g\| < c$ a.e on μ then $\int |fg|\,d\mu \leq \int |f| c\,d\mu = c \int |f|\,d\mu$

If we take $c \downarrow \|g\|_\infty$, we get $\|fg\|_1 \leq \|f\|_1 \|g\|_\infty$

* Durret 1.5.4

a> If f is integrable, $\{E_n\}$ are disjoint sets with $\bigcup_{n=1}^{\infty} E_n = E$. Then $\sum_{n=0}^{\infty} \int_{E_n} f \, d\mu = \int_E f \, d\mu$.

b> If $f \geq 0$ then $\nu(E) = \int_E f \, d\mu$ defines a measure

a> Put $f_n = f \mathbb{1}_{E_n}$

Put $g_m = \sum_{n=1}^{m} f_n$ then since E_n are pairwise disjoint, we have that $g_m \longrightarrow f$ point wise (1)

• We also have $|g_m| < |f|$, a L^1 function (2)

(1)+(2), by LDC theorem $\int_E f \, d\mu = \lim_{m\to\infty} \int_E g_m \, d\mu = \lim_{m\to\infty} \int_E \sum_{n=1}^{m} f_n \, d\mu$

$\underset{\text{linearity}}{=} \lim_{m\to\infty} \sum_{n=1}^{m} \int_E f \mathbb{1}_{E_n} \, d\mu = \lim_{m\to\infty} \sum_{n=1}^{m} \int_{E_n} f \, d\mu = \sum_{n=1}^{\infty} \int_{E_n} f \, d\mu$ □a>

b> Prove that if $f \geq 0$, then $\nu(E) = \int_E f \, d\mu$ defines a measure

we need to prove
i> $\nu(E) \geq 0$, $\forall E$
ii> $E_1, E_2, \ldots$ pairwise disjoint, then $\nu(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \nu(E_i)$

i> is obvious since $f \geq 0 \Rightarrow \int_E f \, d\mu \geq 0$

ii> comes from a>, since

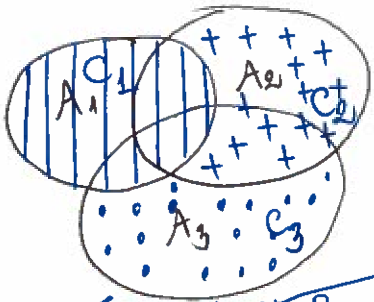
$\nu(\bigcup_{i=1}^{\infty} E_i) = \nu(E) = \int_E f \, d\mu$

$\sum_{i=1}^{\infty} \nu(E_i) = \sum_{i=1}^{\infty} \int_{E_i} f \, d\mu$

7) Let $\mathcal{S}$ be a semialgebra.

a) Prove directly that if $A_1, A_2, \dots, A_n \in \mathcal{S}$, then $\exists \begin{cases} B_1, \dots, B_m \in \mathcal{S} \\ \text{pairwise disjoint} \end{cases}$ such that $\bigcup_{i=1}^{n} A_i = \sum_{j=1}^{m} B_j$

b) Suppose $\left.\begin{array}{l} \mu : \mathcal{S} \longrightarrow [0, +\infty] \\ \mu(\phi) = 0 \\ \mu \text{ is finite additive on } \mathcal{S} \end{array}\right\} \Rightarrow$ Prove that μ has monotonicity property, that is If $A_1, \dots, A_n \in \mathcal{S}$, $A \subset \bigcup_{i=1}^{n} A_i$ then $\mu(A) \leq \sum_{i=1}^{n} \mu(A_i)$



Comment: we don't know if $C_i \in \mathcal{S}$ or not but we know C_i is finite pairwise disjoint union of sets in $\mathcal{S}$

Put $C_1 = A_1$

$C_2 = A_2 \setminus A_1 = \bigcup_{i=1}^{k_2} B_{2i}$ where B_{2i}, $i = 1, k_2$ are pairwise disjoint by the properties of semialgebra

$C_3 = A_3 \setminus (A_1 \cup A_2) =$

$= (A_3 \setminus A_1) \cap (A_3 \setminus A_2) = \underbrace{\left(\bigcup_{i=1}^{k_1} D_i\right)}_{\text{pairwise disjoint}} \cap \underbrace{\left(\bigcup_{j=1}^{k_j} D_j\right)}_{\text{pairwise disjoint}}$

$= \bigcup_{i=1}^{k_3} B_{3i}$

$\vdots$

Similarly $C_n = A_i \setminus \left(\bigcup_{j=1}^{n-1} A_j\right) = \bigcup_{i=1}^{k_n} B_{ni}$

$*$ First, we will prove that $\bigcup_{i=1}^{n} C_i = \bigcup_{j=1}^{n} A_j$

- Prove $\bigcup_{i=1}^{n} C_i \subset \bigcup_{j=1}^{n} A_j$

  Let $x \in \bigcup_{i=1}^{n} C_i$ then $\exists i_0$ such that $x \in C_{i_0} = A_{i_0} \setminus \left(\bigcup_{j=1}^{i_0-1} A_j\right) \subset A_{i_0}$

  $\Rightarrow x \in A_{i_0} \subset \bigcup_{j=1}^{n} A_j$

  $\Rightarrow \bigcup_{i=1}^{n} C_i \subset \bigcup_{j=1}^{n} A_j$

- Prove $\bigcup_{j=1}^{n} A_j \subset \bigcup_{i=1}^{n} C_i$

  Let $x \in \bigcup_{j=1}^{n} A_j$ then there exist at least one index j in $\{1, \dots, n\}$ so that $x \in A_i$

  Choose i_0 is the <u>minimum</u> index in $\{1, \dots, n\}$ so that $x \in A_{i_0}$, because i_0 is minimum, this means $x \notin \bigcup_{i=1}^{i_0-1} A_i$

  $\Rightarrow x \in A_{i_0} \setminus \left(\bigcup_{i=1}^{i_0-1} A_i\right) = C_{i_0} \subset \bigcup_{i=1}^{n} C_i$.

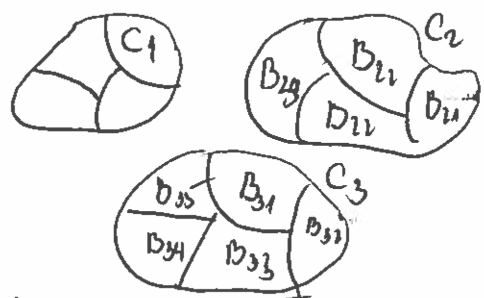
  So $\bigcup_{j=1}^{n} A_j \subset \bigcup_{i=1}^{n} C_i$

* By the way we set up $\{C_i\}_{i=1}^{n}$, we have that these C_i are pairwise disj

Futhermore, each $C_i = \bigcup_{j=1}^{k_j} B_{i,j}$ ✓

pairwise disjoint (by Semialgebra properties)

$$\Rightarrow \bigcup_{j=1}^{m} B_{j} := \underbrace{\bigcup_{\substack{i=\overline{1,n}\\ j=\overline{1,k_i}}} B_{i,j}}_{\text{pairwise disjoint}} = \bigcup_{i=1}^{n} C_i \underset{\text{from above}}{=\!=} \bigcup_{i=1}^{n} A_i$$



So a) has been proved.

b) Suppose $\mu : \mathcal{S} \longrightarrow [0, +\infty]$, $\mu(\emptyset) = 0$, μ is finite additive on $\mathcal{S}$. Prove that μ is monotonicity property, that is If $A_1, \ldots, A_n \in \mathcal{S}$, $A \subset \bigcup_{i=1}^{n} A_i$ then $\mu(A) \leq \sum_{i=1}^{n}$

We have that $\bigcup_{i=1}^{n} A_i = A \cup \left(\left(\bigcup_{i=1}^{n} A_i\right) \setminus A\right) = \underbrace{A \cup \left(\bigcup_{i=1}^{n} (A_i \cap A^c)\right)}_{\text{disjoint}}$ (1)

(2) since μ is finitely additive on $\mathcal{S}$

$$(1)+(2) \Rightarrow \mu(A) + \mu\left(\bigcup_{i=1}^{n}(A_i \cap A^c)\right) = \mu\left(\bigcup_{i=1}^{n} A_i\right) \underset{\text{from a)}}{=\!=} \mu\Big(\underbrace{\bigcup_{j=1}^{m} B_j}_{\text{pairwise disjoint}}\Big) =$$

$$\underset{\text{from the way we set up above}}{=} \mu\Big(\underbrace{\bigcup_{\substack{i=\overline{1,n}\\ j=\overline{1,k_i}}} B_{i,j}}_{\text{pairwise disjoint}}\Big)$$

$$\underset{\mu \text{ is finitely additive}}{=\!=} \sum_{i=1}^{m} \sum_{j=1}^{k_i} B_{ij} \leq \sum_{i=1}^{n} \mu(A_i)$$

something's unclear here.

Mat 721 Problem Set 2, due 2/21

Refer to theorems in Durrett by number, as in Theorem 2.1.14.

1. (For use later in the semester, do not turn in.) Fix $\gamma > 0$, let $x_1 = 1$ and let $x_2, x_3, \ldots$ be positive numbers which satisfy

$$x_n \leq \frac{\gamma}{\sqrt{n}} + \frac{x_{n-1}}{2}, \qquad n \geq 2.$$

   Prove that there is a finite constant C, not depending on γ, such that

$$\limsup_{n \to \infty} \sqrt{n}\, c_n \leq C\gamma$$

2. The statement of the Monotone Convergence Theorem in the text is really: if (a) $f_n(x) \geq 0$ for all $x \in X$, (b) $f_1(x) \leq f_2(x) \leq \ldots$ for all $x \in X$, and (c) $\lim_{n\to\infty} f_n(x) = f(x)$ for all $x \in X$, then $\lim_{n\to\infty} \int f_n d\mu = \int f d\mu$. State and prove from this theorem a version where the conditions only hold μ-a.e.

3. Durrett 1.5.6 and 1.5.10

4. Durrett 1.6.8

5. Durrett 1.7.5

6. Durrett 2.1.13 and 2.1.14

2) The statement of Monotone Convergence theorem in the text:

a) $f_n(x) \geq 0, \forall x \in X$

b) $f_1(x) \leq f_2(x) \leq \cdots \forall x \in X$

c) $\lim_{n\to\infty} f_n(x) = f(x), \forall x \in X$

$\Bigg\}$ $\lim_{n\to\infty}\int f_n d\mu = \int f d\mu$

State and prove from this theorem a version where the conditions only hold μ a.e

* Theorem (version where the conditions only hold μ almost everywhere)

a) $f_n(x) \geq 0$ a.e $[\mu]$

b) $f_1(x) \leq f_2(x) \leq \cdots$ $[\mu]$ a.e on X

c) $f_n(x) \xrightarrow[\text{pointwise}]{n\to\infty} f(x)$ $[\mu]$ a.e on X

$\Bigg\} \Rightarrow$ then $\lim_{n\to\infty}\int f_n d\mu = \int f d\mu$

We have a) $f_n(x) \geq 0$ $[\mu]$ a.e means $\exists F$, $f_n(x) \geq 0, \forall x \in F$ where $\mu(X\setminus F) = 0$

b) $f_1 \leq f_2 \leq \cdots$ $[\mu]$ a.e on X means $\exists E$, $f_1(x) \leq f_2(x) \leq \cdots, \forall x \in E$ where $\mu(X\setminus E)=0$

c) $\lim_{n\to\infty} f_n(x) = f(x)$ $[\mu]$ a.e on X means $\exists D$, $f_n(x) \xrightarrow{n\to\infty} f(x), \forall x \in D$, where $\mu(X\setminus D) = 0$

Then let $C = F\cap E\cap D$, means $\mu(X\setminus C) = \mu(X\setminus(F\cap E\cap D)) = \mu((X\setminus F)\cup(X\setminus E)\cup(X\setminus D))$

$\leq \mu(X\setminus F) + \mu(X\setminus E) + \mu(X\setminus D) = 0$.

Now let $g_n(x) = \begin{cases} f_n(x) & , \forall x \in C \\ 0 & , \forall x \in X\setminus C \end{cases}$ $\quad g(x) = \begin{cases} f(x) & , x \in C \\ 0 & , x \in X\setminus C \end{cases}$

- Then we have $\begin{cases} g_n(x) \xrightarrow[n\to\infty]{} g(x), \forall x \in C \text{ since } C \subset D \\ g_n(x) \longrightarrow g(x), \forall x \in X\setminus C \text{ since } g_n(x) = g(x) = 0 \end{cases}$

  this means $g_n(x) \xrightarrow[n\to\infty]{} g(x), \forall x \in X$

  we also have $g_n(x) \geq 0, \forall x \in X$ by the way we define g

  $g_1 \leq g_2 \leq \cdots \forall x \in X$ by the way we define g

  Then by MCT, $\int g_n d\mu \xrightarrow{n\to\infty} \int g d\mu$ (1)

- We also have $\int g_n d\mu = \int f_n d\mu$ since C is dense in X (2)

  $\int g d\mu = \int f d\mu$ since C is dense in X (3)

(1)+(2)+(3) $\Longrightarrow$ The MCT with a.e assumptions has been proved □

3) Durret 1.5.6 and 1.5.10

1.5.6) Prove if $g_m \geq 0$ then $\int \sum_{m=0}^{\infty} g_m \, d\mu = \sum_{m=0}^{\infty} \int g_m \, d\mu$

1.5.10 Prove that if $\sum_{n=1}^{\infty} \int |f_n| d\mu < +\infty$ then $\sum_{n=1}^{\infty} \int f_n \, d\mu = \int \sum_{n=1}^{\infty} f_n \, d\mu$.

1.5.6)

Put $f_0 = g_0 \geq 0$

$f_1 = g_0 + g_1 \geq f_0$ (since $g_i \geq 0$)

$\vdots$

$f_n = \sum_{i=1}^{n} g_i$

Then we have $0 \leq f_0 \leq f_1 \leq \cdots \leq f_n \leq \sum_{m=0}^{\infty} g_m$ (since $g_i \geq 0$)

and we also have $f_n \xrightarrow{n \to \infty} f = \sum_{m=0}^{\infty} g_m$

Then by monotone convergence theorem $\int f \, d\mu = \lim \int f_n \, d\mu$ — linearity of integration

$\int \sum_{m=0}^{\infty} g_m \, d\mu \;\Big|\; \lim_{n \to \infty} \int \sum_{m=1}^{n} g_m \, d\mu = \lim_{n \to \infty} \sum_{m=1}^{n} \int g_m$

$= \sum_{m=1}^{\infty} \int g_m \, d\mu$

they are equal

1.5.10) Put $g_n = \sum_{i=1}^{n} f_i$, then we have

① $|g_n| = \left|\sum_{i=1}^{n} f_i\right| \leq \sum_{i=1}^{n} |f_i| \leq \sum_{i=1}^{\infty} |f_i| = \varphi$

Use the result we get from 1.5.6, we have $|f_i| \geq 0 \Rightarrow \underbrace{\int \sum_{i=1}^{\infty} |f_i| \, d\mu}_{\varphi} = \sum_{i=1}^{\infty} \int |f_i| d\mu <$

which means we have $|g_n| \leq \varphi$ a.e where φ is a $L^1(\mu)$ function

② $g_n = \sum_{i=1}^{n} f_i \longrightarrow g = \sum_{i=1}^{\infty} f_i$

③ g_n is measurable since it is a finite sum of measurable functions

①+②+③ $\Rightarrow$ By Dominated Convergence Theorem $g = \sum_{i=1}^{\infty} f_i$ is integrable and

$\int g \, d\mu = \lim_{n \to \infty} \int g_n \, d\mu \iff \int \sum_{i=1}^{\infty} f_i \, d\mu = \lim_{n \to \infty} \int \sum_{i=1}^{n} f_i \, d\mu = \lim_{n \to \infty} \sum_{i=1}^{n} \int f_i \, d\mu$

$= \sum_{i=1}^{\infty} \int f_i \, d\mu$.

* Durrett 1.6.6 A useful lower bound

Let $Y \geq 0$ with $E(Y^2) < +\infty$

Apply the Cauchy Schwarz inequality to $Y \mathbb{1}_{Y>0}$ and conclude $E(Y>0) \geq \frac{(E(Y))^2}{E(Y^2)}$

$E(Y) = E(Y \mathbb{1}_{Y>0})$ since $Y \geq 0$

$\Rightarrow E(Y) = E(Y \mathbb{1}_{Y>0}) \underset{CS}{\leq} \|Y\|_2 \, \|\mathbb{1}_{Y>0}\|_2 = \left(\int Y^2 d\mu\right)^{1/2} \left(\int \mathbb{1}_{Y>0} \, d\mu\right)^{1/2}$

$\Rightarrow [E(Y)]^2 \leq \int Y^2 d\mu \int \mathbb{1}_{Y>0} \, d\mu = E(Y^2) \, P(Y>0)$

$\Rightarrow P(Y>0) \geq \frac{(EY)^2}{E(Y^2)}$

we need this condition to apply MCT

cdf $\Rightarrow f \geq 0$

47. Durret 1.6.8

Suppose that the probability measure μ has $\mu(A) = \int_A f(x)\,dx \quad \forall A \in \mathcal{B}$

Use the proof technique of theorem 1.6.9 to show that

for any g, $g \geq 0$ or $\int |g(x)|\,\mu(dx) < +\infty$, we have

$$\int g(x)\,\mu(dx) = \int g(x) f(x)\,dx.$$

$$\begin{tikzcd}(\Omega, \mathcal{F}, P) \arrow[r, "X"] \arrow[rr, bend right, "g\circ X"'] & (\mathbb{R}, \mathcal{B}, \mu) \arrow[r, "g"] & (\mathbb{R}, \mathcal{B})\end{tikzcd}$$

① When $g = \mathbb{1}_B$

$$\int g\,d\mu = \int \mathbb{1}_B\,d\mu = \int_B d\mu = \mu(B) = \int_B f(x)\,dx = \int \mathbb{1}_B f(x)\,dx = \int g(x) f(x)\,dx$$

② When g is a simple function $g = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{B_i}$ (since this is a finite sum, without lost of generality, we can assume that $\{B_i\}_{i=1}^{n}$ are pairwise disjoint.

$$\int g\,d\mu = \int \sum_{i=1}^{n} \alpha_i \mathbb{1}_{B_i}(x)\,\mu(dx) = \sum_{i=1}^{n} \alpha_i \int \mathbb{1}_{B_i}(x)\,\mu(dx) \overset{①}{=} \sum_{i=1}^{n} \alpha_i \int \mathbb{1}_{B_i}(x) f(x)\,dx =$$

$$= \sum_{i=1}^{n} \int \alpha_i \mathbb{1}_{B_i}(x) f(x)\,dx = \int \sum_{i=1}^{n} \alpha_i \mathbb{1}_{B_i}(x) f(x)\,dx = \int g(x) f(x)\,dx$$

③ In case $g \geq 0$, then $\exists (g_n) \geq 0$, g_n simple function, $g_n \uparrow g$

we need this condition to apply MCT

$$\int g\,d\mu = \int \lim g_n\,d\mu \overset{MCT}{=} \lim_{n\to\infty} \int g_n\,d\mu \overset{②}{=} \lim_{n\to\infty} \int g_n(x) f(x)\,dx \underset{g_n f \uparrow g f}{\overset{MCT}{=}} \int \lim_{n\to\infty} g_n(x) f(x)\,dx$$

$$= \int g(x) f(x)\,dx$$

since $f \geq 0$, $g_n \geq 0$

④ In general case when g is integrable, Let $g = g^+ - g^-$ $\quad g^+ \geq 0$, $g^- \geq 0$

$$\int g\,d\mu = \int g^+\,d\mu - \int g^-\,d\mu \overset{③}{=} \int g^+ f(x)\,dx - \int g^-(x) f(x)\,dx =$$

$$= \int \left(g^+ - g^-(x)\right) f(x)\,dx = \int g(x) f(x)\,dx.$$

5) Durret 17.5

Show that $e^{-xy}\sin x$ is integrable in the strip $0 < x < a$, $0 < y$.
Perform the double integral in the two orders to get

$$\int_0^a \frac{\sin x}{x}\,dx = \arctan(a) - (\cos a)\int_0^\infty \frac{e^{-ay}}{(1+y^2)}\,dy - (\sin a)\int_0^\infty \frac{y e^{-ay}}{1+y^2}\,dy \quad (*)$$

and replace $(1+y^2)$ by 1 to conclude that $\left|\int_0^a \frac{\sin x}{x}\,dx - \arctan(a)\right| \leq \frac{2}{a}$ for $a \geq 1$

* Show that $e^{-xy}\sin x$ is integrable on the strip $0 < x < a$, $0 < y$:

We have

$$\int_0^a \int_0^\infty \left|e^{-xy}\sin x\right| dy\,dx = \int_0^a |\sin x| \left(\int_0^\infty \left|e^{-xy}\right| dy\right) dx$$

$$\left(\int_0^\infty \left|e^{-xy}\right| dy = \frac{1}{x}\int_0^\infty e^{-xy}\,d(xy) \underset{u=xy}{=} \frac{1}{x}\int_0^\infty e^{-u}\,du = \frac{1}{x}\left[-e^{-u}\Big|_0^\infty\right] = \frac{1}{x}\right)$$

Then $\displaystyle\int_0^a \int_0^\infty \left|e^{-xy}\sin x\right| dy\,dx = \int_0^a \frac{|\sin x|}{|x|}\,dx \leq \int_0^a \frac{1}{x}\,dx < +\infty$

*) Now we prove $(*)$

$$\int_0^a \frac{\sin x}{x}\,dx = \int_0^a \sin x\left(\int_0^\infty e^{-xy}\,dy\right) dx \overset{\text{Fubini}}{=} \int_0^\infty \underbrace{\left(\int_0^a e^{-xy}\sin x\,dx\right)}_{(**)} dy$$

• Now we want to compute $(**)$

$$\int_0^a e^{-xy}\sin x\,dx \;\underset{\Rightarrow \begin{cases} du = -y e^{-x} \\ dv = -\cos x \end{cases}}{\overset{\text{put } \begin{cases} u = e^{-xy} \\ dv = \sin x\,dx \end{cases}}{=}}\; -\cos x\, e^{-xy}\Big|_0^a + y\int_0^a \cos x\, e^{-xy}\,dx.$$

$$= 1 - \cos a\, e^{-ay} + y\underbrace{\int_0^a \cos x\, e^{-xy}\,dx}_{(***)}$$

• Compute $(***)$:

$$\int_0^a \cos x\, e^{-xy}\,dx \;\underset{\Rightarrow \begin{cases} du = -y e^{-x} \\ v = \sin x \end{cases}}{\overset{\text{put } \begin{cases} u = e^{-xy} \\ dv = \cos x\,dx \end{cases}}{=}}\; \sin x\, e^{-xy}\Big|_{x=0}^{x=a} - y\int_0^a \sin x\, e^{-xy}\,dx$$

Input $(***)$ into the formula of $(**)$, we have

$$\int_0^a e^{-xy}\sin x\,dx = 1-\cos a\, e^{-ay} + y\left[\sin a\, e^{-ay} - y\int_0^a e^{-xy}\sin x\,dx\right]$$

$$= 1-\cos a\, e^{-ay} + y\sin a\, e^{-ay} - y^2\int_0^a e^{-xy}\sin x\,dx.$$

$$\Rightarrow (1+y^2)\int_0^a e^{-xy}\sin x\,dx = 1-\cos a\, e^{-ay} + y\sin a\, e^{-ya}$$

$$(**) = \int_0^a e^{-xy}\sin x\,dx = \frac{1-\cos a\, e^{-ay} + y\sin a\, e^{-ya}}{1+y^2}$$

- Input $(**)$ to the above, we have

$$\int_0^a \frac{\sin x}{x}dx = \int_0^\infty\left(\int_0^a e^{-xy}\sin x\,dx\right)dy = \int_0^\infty \frac{1-\cos a\, e^{-ay} + y\sin a\, e^{-ya}}{1+y^2}dy$$

$$= \int_0^\infty \frac{1}{1+y^2}dy - \cos a\int_0^\infty \frac{e^{-ay}}{1+y^2} + \sin a\int_0^\infty \frac{y e^{-ya}}{1+y^2}dy$$

I don't understand why the question write arctan (a) here

$$= \underbrace{\arctan(\infty)}_{=\frac{\pi}{2}} - (\cos a)\int_0^\infty \frac{e^{-ya}}{1+y^2}dy + \sin a\int_0^\infty \frac{y e^{-ya}}{1+y^2}dy.$$

* Replace $(1+y^2)$ by 1, to conclude that $\left|\int_0^a \frac{\sin x}{x}dx - \arctan(a)\right| \le \frac{2}{a}$ for $a \geq$

$$\Rightarrow \left|\int_0^a \frac{\sin x}{x} - \arctan(\infty)\right| \le \left|\cos a\int_0^\infty \frac{e^{-ya}}{1+y^2}dy\right| + \left|\int_0^\infty \sin a\frac{y e^{-ya}}{1+y^2}dy\right|$$

$$\le \left|\int_0^\infty \frac{e^{-ya}}{1}dy\right| + \left|\int_0^\infty \frac{y e^{-ya}}{1}dy\right|$$

$$= \left|\frac{1}{a}\left(e^{-ya}\Big|_{y=0}^{y=\infty}\right)\right| + \left|\left(-\frac{1}{a}e^{-ya}\Big|_{y=0}^{y=+\infty} - \frac{1}{a}\int_0^\infty e^{-ya}dy\right)\right|$$

$$= \left|\frac{1}{a}\right| + \left|\left(-\frac{1}{a} - \frac{1}{a}\cdot\frac{1}{a}\right)\right| \le \frac{2}{a} \quad \text{for } a \geq 1.$$

makes sense when $a \to \infty$, then $\int_0^a \frac{\sin x}{x}dx \longrightarrow \frac{\pi}{2}$ (what I have known)

$X:(\Omega,\mathcal{F})\to(\mathbb{R},\mathcal{B})$ ⎫ independent (they are from the same
*Durrett 2.1.1/37 $Y:(\Omega,\mathcal{F})\to(\mathbb{R},\mathcal{B})$ ⎭ space to another same)

i) Show that if X and Y are independent then $\sigma(X)$ and $\sigma(Y)$ are

ii) If $\mathcal{F}$ and $\mathcal{G}$ are independent σ algebras
X is $\mathcal{F}$ measurable
Y is $\mathcal{G}$ measurable
$\Rightarrow$ X and Y are independent.

i) Show that if X and Y are independent then $\sigma(X)$ and $\sigma(Y)$ are independent.

Have: X, Y independent

$\Leftrightarrow \forall C, D\in\mathcal{B}$ then $P(X\in C \text{ and } Y\in D) = P(X\in C)P(Y\in D)$

$\Leftrightarrow$ events $\{\omega, X(\omega)\in C\}$ and $\{\omega, Y(\omega)\in D\}$ independent

NTP $\sigma(X)$ and $\sigma(Y)$ independent

NTP any events $A\in\sigma(X)$ and $B\in\sigma(Y)$ then A & B independent

* Take $A\in\sigma(X)=\{X^{-1}(C), C\in\mathcal{B}\}$ then $\exists C\in\mathcal{B}$, $A=\{X^{-1}(C)\}$

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

because X and Y are independent then

$P(A\cap B)=P\{\cancel{\omega}\, X^{-1}(C)\cap X^{-1}(D)\} = P(X\in C \text{ and } Y\in D) \underset{\text{independent}}{\overset{X,Y}{=}} P(X\in C)P(Y\in D)$

$= P(A)P(B)$.

ii) Have: X is $\mathcal{F}$ measurable $\Leftrightarrow \forall B\in\mathcal{B}, X^{-1}(B)\in\mathcal{F}$

Y is $\mathcal{G}$ measurable $\Leftrightarrow \forall C\in\mathcal{B}, X^{-1}(C)\in\mathcal{G}$

$\mathcal{F},\mathcal{G}$ independent $\Rightarrow X^{-1}(B)$ and $X^{-1}(C)$ independent

NTP: X, Y are independent

$\Leftrightarrow$ events $\{X\in C\}$ and $\{Y\in D\}$ are independent.

* Consider $\forall C\in\mathcal{B}$ and $D\in\mathcal{B}$, then

$P(X\in C \text{ and } Y\in D\} = P\{X^{-1}(C)\cap Y^{-1}(D)\} \underset{\text{independent}}{\overset{\mathcal{F},\mathcal{G}}{=}} P(X^{-1}(C))\,P(Y^{-1}(D)) =$

$= P(X\in C)\,P(Y\in D)$

* Observe:

X and Y are independent $\Leftrightarrow$ $\sigma(X)$ and $\sigma(Y)$ are independent.

Mat 721 Problem Set 1, Problem #4

Durrett 1.3.1: Let $(\Omega, \mathcal{F})$ and $(S, \mathcal{S})$ be measurable spaces, and let $X : \Omega \to S$. Suppose $\mathcal{A}$ is a collection of subsets of S such that $\sigma(\mathcal{A}) = \mathcal{S}$. Prove that $X^{-1}(\mathcal{A}) = \{\{X \in A\} : A \in \mathcal{A}\}$ generates $\sigma(X) \equiv \{\{X \in B\} : B \in \mathcal{S}\}$.

Using $X^{-1}(\mathcal{S})$ for $\sigma(X)$, the goal is to prove

$$\sigma(X^{-1}(\mathcal{A})) = X^{-1}(\mathcal{S}) \tag{0}$$

1. Recall (comment after Thm 1.3.1) or check that $X^{-1}(\mathcal{S})$ is a σ-algebra.
2. $\mathcal{A} \subset \mathcal{S}$ implies $X^{-1}(\mathcal{A}) \subset X^{-1}(\mathcal{S})$, which implies $\sigma(X^{-1}(\mathcal{A})) \subset \sigma(X^{-1}(\mathcal{S}))$. Since $X^{-1}(\mathcal{S})$ is a σ-algebra, this shows
$$\sigma(X^{-1}(\mathcal{A})) \subset X^{-1}(\mathcal{S}) \tag{1}$$
3. Define $\mathcal{S}' = \{B \in \mathcal{S} : X^{-1}(B) \in \sigma(X^{-1}(\mathcal{A}))\}$. Then $\mathcal{S}' \subset \mathcal{S}$ and $\mathcal{A} \subset \mathcal{S}'$.
4. Check that $\mathcal{S}'$ is a σ-algebra:
   - If $C \in \mathcal{S}'$ then $X^{-1}(C) \in \sigma(X^{-1}(\mathcal{A}))$, and since $\sigma(X^{-1}(\mathcal{A}))$ is closed under complements,
   $$(X^{-1}(C))^c \in \sigma(X^{-1}(\mathcal{A})).$$
   Since $X^{-1}(C^c) = (X^{-1}(C))^c$, this shows $C^c \in \mathcal{S}'$.
   - If $C_1, C_2, \cdots \in \mathcal{S}'$ then $X^{-1}(C_i) \in \sigma(X^{-1}(\mathcal{A}))$ for each i, and since $\sigma(X^{-1}(\mathcal{A}))$ is closed under countable unions,
   $$\cup_i X^{-1}(C_i) \in \sigma(X^{-1}(\mathcal{A})).$$
   Since $X^{-1}(\cup_i C_i) = \cup_i X^{-1}(C_i)$, this shows $\cup_i C_i \in \mathcal{S}'$.

   Thus $\mathcal{S}'$ is a σ-algebra.
5. $\mathcal{A} \subset \mathcal{S}'$ implies $\sigma(\mathcal{A}) \subset \sigma(\mathcal{S}')$, and therefore that $\mathcal{S} \subset \mathcal{S}'$. By definition of $\mathcal{S}'$, this shows that if $B \in \mathcal{S}$ then $X^{-1}(B) \in \sigma(X^{-1}(\mathcal{A}))$, proving
$$X^{-1}(\mathcal{S}) \subset \sigma(X^{-1}(\mathcal{A})). \tag{2}$$
6. (1) and (2) imply (0)

*Durret 2.1.5 / page 40

Suppose $X_1, .., X_n$ are random variables that take values in countable sets $S_1, S_2, ..., S_n$
Then in order for $X_1, ..., X_n$ to be independent, it is sufficient that whenever $x_i \in S_i$
$P(X_1 = x_1, ..., X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i)$, $\forall x_i \in S_i$, $i = \overline{1,n}$

6) Durret 2.1.13 and 2.1.14.

2.1.13) Show that if X and Y are independent, integer-valued r.v.s, then

$P(X+Y=n) = \sum_m P(X=m)\,P(Y=n-m)$

2.1.14) In example 1.6.4)

Poisson (λ) : $P(Z=k) = \frac{e^{-\lambda}\lambda^k}{k!}$, $k=0,1,2\ldots$

Use 2.1.13 to show that $\left.\begin{array}{l} X \sim \text{Poisson}(\lambda) \\ Y \sim \text{Poisson}(\mu) \end{array}\right\}$ independent $\Rightarrow X+Y \sim \text{Poisson}(\lambda+\mu)$

2.1.13) Consider $n \in \mathbb{Z}^+$

Then $P(X+Y=n) = \sum_{\substack{m\in\mathbb{Z}^+ \\ m\le n}} P(X=m,\ Y=n-m) \underset{\text{independent}}{\overset{X,Y}{=}} \sum_{m\in\mathbb{Z}^+} P(X=m)\,P(Y=n-m)$

2.1.14)

We want to prove that $P(X+Y=n) = \frac{e^{-(\lambda+\mu)}(\lambda+\mu)^n}{n!}$

From 2.1.13) we have

$$P(X+Y=n) = \sum_{\substack{m\in\mathbb{Z}^+ \\ m\le n}} P(X=m)\,P(Y=n-m) =$$

$$= \sum_{\substack{m\in\mathbb{Z}^+ \\ m\le n}} \frac{e^{-\lambda}\lambda^m}{m!}\,\frac{e^{-\mu}\mu^{n-m}}{(n-m)!} = e^{-(\lambda+\mu)}\underbrace{\frac{1}{n!}\sum_{\substack{m\in\mathbb{Z}^+ \\ m\le n}} \underbrace{\frac{n!}{(m)!(n-m)!}}_{=\binom{n}{m}}\lambda^m\mu^{n-m}}_{=(\lambda+\mu)^n}$$

$$= \frac{e^{-(\lambda+\mu)}(\lambda+\mu)^n}{n!} \quad \square.$$

A

MAT 721, Spring 2019
Exam #1

NAME: Tran Le

1. Suppose X and Y are random variables on $(\Omega, \mathcal{F}, P)$ and $A \in \mathcal{F}$. Define

$$Z(\omega) = \begin{cases} X(\omega) & \text{for } \omega \in A \\ Y(\omega) & \text{for } \omega \notin A \end{cases}$$

Prove that Z is a random variable.

Have: • $X : (\Omega, F, P) \longrightarrow (\mathbb{R}, \mathcal{B})$ random variable

$\Leftrightarrow X^{-1}(B) \in F, \forall B \in \mathcal{B}$.

• $Y : (\Omega, F, P) \longrightarrow (\mathbb{R}, \mathcal{B})$ random variable

$\Leftrightarrow X^{-1}(B) \in F, \forall B \in \mathcal{B}$

* We need to prove that $Z : (\Omega, F, P) \longrightarrow (\mathbb{R}, \mathcal{B})$ is a random variable

$\Leftrightarrow$ Need to prove that $\forall C \in \mathcal{B}$, then $Z^{-1}(C) \in F$.

Let $C \in \mathcal{B}$

• Now consider $Z^{-1}(C) = \{w \in \Omega, Z(w) \in C\}$.

$= \{w \in A, Z(w) \in C\} \cup \{w \in A^c, Z(w) \in C\}$,

$= \{w \in A, X(w) \in C\} \cup \{w \in A^c, Y(w) \in C\}$.

$$= \underbrace{\underbrace{X^{-1}(C)}_{\in F} \cup \underbrace{Y^{-1}(C)}_{\in F}}_{\in F}$$

(since X and Y are random variables)

$\in F$ since F is a σ-algebra $\Rightarrow$ closed under countab[le] and finite union.

So we have shown $Z^{-1}(C) \in F, \forall C \in \mathcal{B}$ which means Z is a random variable $\square$

> Prove that H contains all bounded $\sigma(\mathcal{A})$ measurable function.

$\left.\begin{array}{l}\mathcal{A}: \pi \text{ system.} \\ \mathcal{G}: \lambda \text{ system containing } \mathcal{A}.\end{array}\right\} \xrightarrow[\text{theorem}]{\pi-\lambda} \sigma(\mathcal{A}) \subseteq \mathcal{G}.$ ✓

$= \{A \in F, 1\!\!1_A \in H\}.$

. Consider $f: (\Omega, \overset{\sigma(\mathcal{A})}{\cancel{\mathcal{F}}}, P) \longrightarrow (S, \mathcal{S})$ (bounded)

f is $\sigma(\mathcal{A})$ measurable $\cancel{\Leftrightarrow f^{-1}(B) \in \sigma(\mathcal{A}) \overset{\in F}{\subset} \mathcal{G}, \forall B \in \mathcal{S}}$.

~~we need to prove that $f \in H$.~~

$\Rightarrow f$ is $\mathcal{G}$-measurable, then by def $\mathcal{G}$, $\{f$ is in H.

and $\mathcal{G}$ is a π system$\}$.

Incomplete

2. Let $(\Omega, \mathcal{F}, P)$ be a probability space, let $\mathcal{A}$ be a π-system that contains Ω, and let $\mathcal{H}$ be a collection of real-valued functions satisfying:

(a) If $A \in \mathcal{A}$ then $1_A \in \mathcal{H}$.

(b) If $f, g \in \mathcal{H}$ then $f + g$ and $cf \in \mathcal{H}$ for all real numbers c. $\Rightarrow -g \in \mathcal{H}$ (applied $c = -1$) $\Rightarrow f - g \in \mathcal{H}$

(c) If $f_1, f_2, \cdots \in \mathcal{F}$, $0 \le f_1 \le f_2 \dots$, and for a bounded function f, $f_n \uparrow f$ as $n \to \infty$, then $f \in \mathcal{H}$.

Let $\mathcal{G} = \{A \in \mathcal{F} : 1_A \in \mathcal{H}\}$.

I. Prove that $\mathcal{G}$ is a λ-system.

II. Prove that $\mathcal{H}$ contains all bounded $\sigma(\mathcal{A})$-measurable functions.

* $\mathcal{G} = \{A \in \mathcal{F}, 1\!\!1_A \in \mathcal{H}\}$. I> Prove $\mathcal{G}$ is a λ system

① Check $\Omega \in \mathcal{G}$.

since $\Omega \in \mathcal{A}$ (by assumption) and by (a) $\Omega \in \mathcal{A} \Rightarrow 1\!\!1_\Omega \in \mathcal{H}$ $\Big\} \Rightarrow \Omega \in \mathcal{G}$.

we also have $\Omega \in \mathcal{F}$ since $\mathcal{F}$ σ algebra

② Check if $A, B \in \mathcal{G}$, $A \subset B$ then $B \setminus A \in \mathcal{G}$.

$A, B \in \mathcal{G} \Leftrightarrow A \in \mathcal{F}, 1\!\!1_A \in \mathcal{H}$
$\qquad B \in \mathcal{F}, 1\!\!1_B \in \mathcal{H}$

• $A \subset B$, $A \in \mathcal{F}$, $B \in \mathcal{F}$, $\mathcal{F}$ σ algebra $\Rightarrow B \setminus A \in \mathcal{F}$.

• $1\!\!1_A \in \mathcal{H}$, $1\!\!1_B \in \mathcal{H} \Rightarrow$ by (b) $1\!\!1_{B \setminus A} = 1\!\!1_B - 1\!\!1_A \in \mathcal{H}$ since b>

③ Check if $A_1 \subset A_2 \subset \cdots$, $A_n \uparrow A$ then $A \in \mathcal{G}$.

• we have since $A_1 \subset A_2 \subset \cdots$, $A_n \uparrow A$ $\Big\} \Rightarrow A = \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ (since $\mathcal{F}$ σ algebra (1) $\Rightarrow$ closed under countable uni

• We have $1\!\!1_{A_i} \in \mathcal{F}$ $\forall A_i$ (since $1\!\!1 : A_i \longrightarrow \{0, 1\}$ discrete $\Rightarrow$ a random variable.

$0 \le 1\!\!1_{A_1} \le 1\!\!1_{A_2} \le \cdots$ since $A_1 \subset A_2 \subset \cdots$

and $1\!\!1_{A_i} \nearrow 1\!\!1_A$ as $n \to \infty$ by

$\to 1\!\!1_A \in \mathcal{H}$ (2)

(1)+(2) $\Rightarrow A \in \mathcal{G}$

①+②+③ $\Rightarrow$ $\mathcal{G}$ is a λ system.

$X \geq 0$, not identically 0.

$X^2 \geq 0$ not identically 0 $\Rightarrow E(X^2) > 0$.

3. Let X be a nonnegative random variable not indentically 0. Prove that

$$P(X > 0) \geq \frac{(E(X))^2}{E(X^2)}.$$

We want to prove $(E(X))^2 \leq E(X^2)\, \mathbb{P}(X>0)$. (Holder's inequality should be applied)

$$\cancel{(E(X))^2 = E(X)\,E(X) \leq E(|X|)\,E(|X|)}$$

$$\cancel{E(|X|) = \|X\|_1 \leq \|X\|_2 = [E(X^2)]^{1/2}}$$

$$\cancel{E(X) = \int x\, d\mu \overset{x \geq 0}{=} \int X \mathbb{1}_{(X \geq 0)}\, d\mu = \mathbb{P}(X \geq 0)}$$

$$[E(X)]^2 = E(X)\,E(X) \leq E(X^2) \underbrace{\int X \, d(\mu x)}$$

$$= \int X \mathbb{1}_{X \geq 0}\, \mu(dx) \quad \text{since } X \geq 0.$$

?

$$= \mathbb{P}(X > 0)$$

$$E(X) < E(X^2)$$

since $\|X\|_1 = \|X \mathbb{1}\|_1 \leq \|X\|_2 \, \|\mathbb{1}\|_2 = \|X\|_2$.

$$\|X\|_1 = \|X \mathbb{1}_{\{X>0\}}\|_1 \leq \|X\|_2 \cdot \|\mathbb{1}_{\{X>0\}}\|_2$$

Now square + rearrange

I remeber that I had learned a theorem that may be similar with this problem but I'm not sure how to do this

4. Let X be a random variable such that $P(X > 0) = 1$. Prove that

$$\lim_{y \downarrow 0} yE\left(\frac{1}{X}1_{\{X>y\}}\right) = 0$$

Do NOT assume $E(1/X) < \infty$.

$$* \quad yE\left(\frac{1}{X}1_{X>y}\right) = E\left(\frac{1}{X}1_{(X>y)}y\right) = \int_\Omega \frac{1}{x}1_{(x>y)}y\,\mu(dx)$$

$$\lim_{y\downarrow 0} yE\left(\frac{1}{X}1_{X>y}\right) = \lim_{y\to 0}\int_\Omega \frac{1}{x}1_{\{x>y\}}y\,\mu(dx).$$

$$= \lim_{n\to\infty}\int_\Omega \underbrace{\frac{1}{x}1_{\{x>\frac{1}{n}\}}(x)\left(\frac{1}{n}\right)}_{\text{put}=g(x)}\mu(dx). \quad (*)$$

Won't work.

* If we can apply Monotone or Dominated convergence theorem, then $(*) = \int \underbrace{\lim_{n\to\infty}\frac{1}{x}1_{\{x>\frac{1}{n}\}}(x)}_{\longrightarrow 0 \; n\uparrow\infty}\frac{1}{n}\mu(dx) = \int 0 = 0.$

* Put $g_n(x) = \frac{1}{x}1_{\{x>\frac{1}{n}\}}(x)\left(\frac{1}{n}\right)$

X

Then $|g_n(x)| \le \left|\frac{1}{x}\right|$ a L^1 function.

$g_n(x) \uparrow g(x)$

$\xrightarrow{DCT} \lim\int = \int\lim.$

- $\frac{Y}{X}1_{\{X>Y\}} = \frac{Y}{X}1_{\{\frac{Y}{X}\le 1\}} \le 1$
- $\frac{Y}{X}1_{\{X>Y\}} \to 0$ as $Y \to \infty$

Apply Bounded Conv. Thm

Tran le
Homework 3

Mat 721 Problem Set 3, Thur March 7

Refer to theorems in Durrett by number, as in Theorem 2.1.14.

1. Let X, Y be independent random variables. By Theorem 2.19 in the text, if either $X, Y \geq 0$ or both $E|X|$ and $E|Y|$ are finite, then $E(XY) = E(X)E(Y)$. Given a second proof of this fact by verifying it first for indicator r.v.'s, then simple r.v.'s, then nonnegative r.v.'s, then L^1 r.v.'s.
2. Durrett 2.2.4 Use Thm 2.2.7
3. Durrett 2.2.6 For (b), there should be no square roots in your answer.
4. Durrett 2.3.8, 2.3.9 – Just read these.
5. Durrett 2.3.14
6. Durrett 2.3.18

1) Let X and Y be independent random variables.

By theorem 2.1.9 in the text, if $\begin{cases} X, Y \geq 0 \\ E(|X|) < +\infty, E(|Y|) < +\infty \end{cases}$ then $E(XY) = E(X)E(Y)$

Given a second prove of this fact by veryfying it for indicators rv's ⇒ simple rv's ⇒ non negative rv's ⇒ L^1 rv's.

① Consider when $X = 1\!\!1_A$, $Y = 1\!\!1_B$ and X and Y are independent (⟺ A and B are independen

$$E(1\!\!1_A 1\!\!1_B) = E(1\!\!1_{A\cap B}) = P(A\cap B) \overset{\text{independent}}{=\!=} P(A)P(B)$$
$$E(1\!\!1_A)E(1\!\!1_B) = P(A)P(B)$$
$\Rightarrow E(1\!\!1_A 1\!\!1_B) = E(1\!\!1_A)E(1\!\!1_B)$

$E(XY) = E(X)E(Y)$.

② Consider when $X = \sum_{i=1}^n \alpha_i 1\!\!1_{A_i}$, $Y = \sum_{j=1}^m \beta_j 1\!\!1_{B_j}$, $\{A_i\}_{i=1}^n$ pairwise disjoint, $\{B_j\}_{j=1}^m$ pai disjon

where all $\{A_i\}_{i=1}^n$ are independent of all $\{B_j\}_{j=1}^m$

$$E(XY) = E\left(\left(\sum_{i=1}^n \alpha_i 1\!\!1_{A_i}\right)\left(\sum_{j=1}^m \beta_j 1\!\!1_{B_j}\right)\right) = \sum_{\substack{i=1,n\\ j=1,m}} \alpha_i\beta_j P(A_i\cap B_j)$$

$$E(X)E(Y) = E\left(\sum_{i=1}^n \alpha_i 1\!\!1_{A_i}\right)\left(E\left(\sum_{j=1}^m \beta_j 1\!\!1_{B_j}\right)\right) = \left(\sum_{i=1}^n \alpha_i P(A_i)\right)\left(\sum_{j=1}^m \beta_j P(B_j)\right) =$$

$$= \sum_{\substack{i=1,n\\ j=1,m}} \alpha_i\beta_j P(A_i)P(B_j) = \sum_{\substack{i=1,n\\ j=1,m}} \alpha_i\beta_j P(A_i\cap B_j).$$

③ When X and Y are non negative and independent random variables

We have $\exists X_n \geq 0$, simple, $X_n \uparrow X$

$\exists Y_n \geq 0$, simple, $Y_n \uparrow Y$

since X, Y are independent ⇒ X_n and Y_n have to be independent

Applying Monotone convergence for $g_n = X_nY_n$, we have $X_nY_n \uparrow XY$

and so we have $E(X_nY_n) \uparrow E(XY)$

We also have since $E(X_nY_n) = E(X_n)E(Y_n) \uparrow E(X)E(Y)$ ⇒ $E(XY) = E(X)$

④ When X and Y are independent and $E(|X|) < +\infty$ $E(|Y|) < +\infty$

Then $X = X^+ - X^-$ and $Y = Y^+ - Y^-$ and since X, Y independent ⇒ $X^{+,-}$ and $Y^{+,-}$ are in

$$E(XY) = E((X^+ - X^-)(Y^+ - Y^-)) = E(X^+Y^+ - X^-Y^+ - X^+Y^- + X^-Y^-) =$$

$$= E(X^+Y^+) - E(X^-Y^-) - E(X^+Y^-) + E(X^-Y^-).$$

$$\overset{\text{above}}{=} E(X^+)E(Y^+) - E(X^-)E(Y^-) - E(X^+)E(Y^-) + E(X^-)E(Y^-)$$

$$= E(X^+)[E(Y^+) - E(Y^-)] - E(X^-)[E(Y^+) - E(Y^-)]$$

$$= [E(X^+) - E(X^-)][E(Y^+) - E(Y^-)] = E(X)E(Y)$$

2) Durrett 2.2.4 Use Theorem 2.2.7

Let $X_1, X_2, \ldots \overset{iid}{\sim}$ $P(X_i = (-1)^k k) = \frac{c}{k^2}(\log k)^{-1}$, $k \geq 2$

c is chosen to make the $\sum P(X_i = \ldots) = 1$

Show that $E(|X_i|) = \infty$, but there is a finite constant μ so that $\frac{S_n}{n} \xrightarrow{P} \mu$

NOTE: I DON'T KNOW HOW TO WORK with the f when (lo is the numerato So I changes it to make (log k in the denomin

* Theorem 2.2.7 (Weak law of large numbers)

Let $X_1, X_2, \ldots \overset{iid}{\sim}$ $x P(|X_i| > x) \xrightarrow{x \to \infty} 0$

Let $S_n = X_1 + \cdots + X_n$

$\mu_n = E(X_1 1_{(|X_1| < n)})$

Then $\frac{S_n}{n} \xrightarrow{P} \mu_n$

* Show that $E(|X_i|) = +\infty$

$$E(|X_i|) = \sum_{x_i \geq 2} x_i P(X = x_i) = \sum_{k \geq 2} |(-1)^k k| \frac{c}{k^2}(\log k)^{-1} = \sum_{k \geq 2} \frac{c(\log k)^{-1}}{k} = +\infty$$

(since the series $\sum \frac{1}{k(\log k)^p}$ converges if $p > 1$ and diverse if $p \leq 1$)

* Now we prove the condition of theorem 2.2.7

$$x P(|X_i| > x) = x \sum_{|X_i| = \lceil x \rceil} \frac{c}{k^2 \log k} < x \sum_{k = \lceil x \rceil}^{\infty} \frac{c}{k^2 \log k} = x \int_x^{\infty} \frac{c}{k^2 \log k} dk$$

$$= \frac{xc}{\log x} \int_x^{\infty} \frac{1}{k^2} dk = \frac{xc}{\log x} \left(\frac{-1}{k}\right)\Big|_{k=x}^{k=\infty} \xrightarrow{x \to \infty} 0 \quad (1)$$

• $$\mu_n = E(X 1_{|X| \leq n}) = \sum_{2}^{\lceil n \rceil} (-1)^k k \frac{c}{k^2 \log k} = \sum_{2}^{\lceil n \rceil} (-1)^k \frac{c}{k \log k}.$$

this is an alternative series with $\frac{1}{k \log k}$ decrease to 0 $\Rightarrow \mu_n \xrightarrow{n \to \infty} \mu$. (2)

(1) + (2) then by theorem 2.2.7 we have $\frac{S_n}{n} \xrightarrow{P} \mu$. □

3) Durret 2.2.6 For b) there should be no square roots in your answer
i) Show that if $X \geq 0$ is integer valued $\Rightarrow E(X) = \sum_{n \geq 1} P(X \geq n)$
ii) Find a similar expression for $E(X^2)$

a) We have lemma 2.2.8

$$EX \overset{X \geq 0}{=} E(|X|) \overset{\text{lemma 2.2.8}}{=} \int_0^\infty P(|X|>x)\,dx = \lim_{m\to\infty}\int_0^m P(|X|>x)\,dx =$$

$$= \lim_{m\to\infty}\sum_{n=0}^{m}\int_n^{n+1} P(|X|>x)\,dx = \sum_{n=0}^{\infty}\int_n^{n+1} P(|X|>x)\,dx =$$

$$= \sum_{n=0}^{\infty}\int_n^{n+1} P(X>n)\,dx = \sum_{n=0}^{\infty} P(X>n)\int_n^{n+1} dx = \sum_{n=0}^{\infty} P(X>n) = \sum_{n\geq 1} P(X\geq n)$$

b) $$E(X^2) \overset{\text{lemma 2.2.8}}{=} \int_0^\infty 2x^{2-1}P(X>x)\,dx = \int_0^\infty 2x\,P(X>x)\,dx$$

$$= \sum_{n=0}^{\infty}\int_n^{n+1} 2x\,P(X>x)\,dx = \sum_{n=0}^{\infty} P(X>n)\int_n^{n+1} 2x\,dx =$$

$$= \sum_{n=0}^{\infty} \cancel{2}P(X>n)\left(\frac{(n+1)^2}{\cancel{2}} - \frac{n^2}{\cancel{2}}\right) = \sum_{n=0}^{\infty} P(X>n)\,(2n+1) =$$

$$= \sum_{n\geq 1}(2n+1)\,P(X>n) \quad \square$$

5) Durrett 2.3.14.

Let $X_1, X_2, \ldots$ independent $\begin{cases} P(X_n = 1) = p_n \\ P(X_n = 0) = 1 - p_n \end{cases}$

i) Show that $X_n \xrightarrow{P} 0 \iff p_n \to 0$

ii) $X_n \xrightarrow{a.s} 0 \iff \sum p_n < +\infty$

i) We have

$X_n \xrightarrow{P} 0 \overset{def}{\iff} P\{w, |X_n - 0| < \varepsilon\} \xrightarrow{n \to \infty} 1$ for $0 < \varepsilon < 1$

$\underset{X_n \text{ only receive } 0 \text{ or } 1}{\iff} P\{w, X_n = 0\} \xrightarrow{n\to\infty} 1$

$\iff 1 - p_n \xrightarrow{n \to \infty} 1$

$\iff p_n \xrightarrow{n\to\infty} 0$

ii) $(\Rightarrow)$: Have: $X_n \xrightarrow{a.s} 0$ need to prove $\sum p_n < +\infty$

We will prove $\sum p_n = +\infty$ then $X_n \overset{a.s}{\not\to} 0$

We have $\sum p_n = \sum P(X_n = 1) = +\infty$

note that from the assumption, we have that $\{X_i\}$ are independent

so by 2nd B-C lemma, $P(X_n = 1 \text{ i.o}) = 1$, which means $X_n \overset{a.s}{\not\to} 0$

$(\Leftarrow)$: $\sum p_n < +\infty$, Need to prove $X_n \xrightarrow{a.s} 0$.

$\sum p_n < +\infty \iff \sum_{n=1}^{\infty} P(X_n = 1) < +\infty$

by 1st B-C lemma, $P(X_n = 1 \text{ i.o.}) = 0$, which means $X_n \xrightarrow{a.s} 0$

since X_n only receive 0 or 1 value $\square$

6) Durrett 2.3.18

Let $X_1, X_2, \ldots \overset{iid}{\sim} P(X_i > x) = e^{-x}$

Let $M_n = \max\limits_{1 \le m \le n} X_m$

i) Show that $\limsup\limits_{n\to\infty} \frac{X_n}{\log n} = 1$ a.s   ii) Show that $\frac{M_n}{\log n} \xrightarrow{a.s} 1$

i) We need to prove $\limsup \frac{X_n}{\log n} = 1$ a.s $\underset{\text{Need to prove}}{\Longleftrightarrow}$ $\begin{cases} P\left(\limsup \frac{X_n}{\log n} > 1-\epsilon\right) = 1 \\ P\left(\limsup \frac{X_n}{\log n} > 1+\epsilon\right) = 0 \end{cases}$

* Now we prove ①: $P\left(\limsup \frac{X_n}{\log n} > 1-\epsilon\right) = 1$

By Borel Cantelli lemma, it suffices to prove that $\sum\limits_{i=1}^{\infty} P\left(\frac{X_n}{\log n} > (1-\epsilon)\right) = +\infty$

$$\sum_{i=1}^{\infty} P\left(\frac{X_n}{\log n} > (1-\epsilon)\right) = \sum_{i=1}^{n} P\left(X_n > (1-\epsilon)\log n\right) \overset{\text{assumption}}{=} \sum_{i=1}^{n} e^{-(1-\epsilon)\log n} =$$

$$= \sum_{i=1}^{n} n^{-(1-\epsilon)} = \sum_{i=1}^{n} \frac{1}{n^{1-\epsilon}} \geq \sum_{i=1}^{n} \frac{1}{n} = +\infty$$

$\Rightarrow$ By BC $\Rightarrow$ ①

* Now we want to prove ② $P\left(\limsup \frac{X_n}{\log n} > (1+\epsilon)\right) = 0$

By Borel Cantelli lemma, it suffices to prove that $\sum\limits_{i=1}^{\infty} P\left(\frac{X_n}{\log n} > (1+\epsilon)\right) < \infty$

$$\sum_{i=1}^{\infty} P\left(\frac{X_n}{\log n} > (1+\epsilon)\right) = \sum_{i=1}^{\infty} P\left(X_n > (1+\epsilon)\log n\right) = \sum_{i=1}^{n} e^{-(1+\epsilon)\log n} =$$

$$= \sum_{n=1}^{\infty} \frac{1}{n^{1+\epsilon}} < \infty \text{ (converges)}$$

* Then from ① + ② $\Rightarrow \limsup \frac{X_n}{\log n} = 1$ a.s

ii) Need to prove $\frac{M_n}{\log n} \xrightarrow{a.s} 1 \xLeftrightarrow[\text{to prove}]{\text{Need}} \begin{cases} P\left(\frac{M_n}{\log n} < (1-\varepsilon)\ i.o\right) = 0 \\ P\left(\frac{M_n}{\log n} > (1+\varepsilon)\ i.o\right) = 0 \end{cases}$

(when $M_n = \max_{1\le i\le n} X_i$)

$\xLeftrightarrow[\text{to prove}]{\text{Need}} \begin{cases} P(M_n < (1-\varepsilon)\log n \ i.o) = 0 & (3) \\ P(M_n > (1+\varepsilon)\log n \ i.o) = 0 & (4) \end{cases}$

* Now we prove (3), prove $P(M_n < (1-\varepsilon)\log n\ i.o) = 0$

We need to prove that $\sum P(M_n < (1-\varepsilon)\log n) < +\infty$

$$P\left(M_n < (1-\varepsilon)\log n\right) = P\left(X_i < (1-\varepsilon)\log n,\ \forall i = \overline{1,n}\right)$$

$$\underset{\{X_i\}\ iid}{=} \prod_{i=1}^{n} P(X_i < (1-\varepsilon)\log n) = \prod_{i=1}^{n}\left(1 - e^{-(1-\varepsilon)\log n}\right)$$

$$= \prod_{i=1}^{n}\left(1 - n^{-(1-\varepsilon)}\right) \quad \text{since } 1-u \le e^{-u} \text{ for all } u \geq$$

$$\le \prod_{i=1}^{n} e^{-n^{-(1-\varepsilon)}}$$

$$\sum_{i=1}^{n} e^{-n^{(\varepsilon+1)}} = \sum_{i=1}^{n} e^{-n^{\delta}} \quad \text{where } \delta = \varepsilon + 1.$$

for fixed $\delta > 0$, we have $\exists N_\delta$ such that $\forall n \geq N$, $n^\delta \geq 2\log n$

then $\sum_{n=N}^{\infty} e^{-n^\delta} < \sum_{n=N}^{\infty} e^{-2\log n} = \sum_{n=N}^{\infty} e^{\log(n^{-2})} = \sum_{n=N}^{\infty} \frac{1}{n^2} < +\infty$

and so we have $\sum_{i=1}^{\infty} P(M_n < (1-\varepsilon)\log n) = \sum_{i=1}^{N-1} \text{mm} + \sum_{n=N}^{\infty} \text{mm} < +\infty$

By 1st B-C lemma $\Rightarrow P(M_n < (1-\varepsilon)\log n\ i.o) = 0$

* Now we prove ④, prove that $P(M_n > (1+\varepsilon)\log n \text{ i.o.}) = 0$.

In part ②, we have proved that $P(X_n > (1+\varepsilon)\log n \text{ i.o}) = 0$,

so it suffices to prove that

$$(M_n > (1+\varepsilon)\log n \text{ i.o}) \subset (X_n > (1+\varepsilon)\log n \text{ i.o.})$$

→ We will prove that $\underbrace{\{w, X_n > (1+\varepsilon)\log n \text{ i.o.}\}^c}_{A} \subset \underbrace{\{w, M_n > (1+\varepsilon)\log n \text{ i.o.}\}^c}_{B}$.

Let $w \in A$, then $\exists N_1(w)$, $\forall n \geq N_1$, $X_n < (1+\varepsilon)\log n$

We also know that when n is large enough, choose $n \geq N_2 > N_1$, we have

$$X_n < (1+\varepsilon)\log n$$

So we have

$M_n(w) = \max\{X_1, X_2, \dots, X_{N_1}, X_{N_1+1}, \dots, X_n\}$

$= \max\{X_1, \dots, X_{N_1}\} \vee \max\{X_{N_1+1}, \dots, X_n\}$

$\leq (1+\varepsilon)\log n \vee (1+\varepsilon)\log n = (1+\varepsilon)\log n$.

then $w \in B$

so we have $A \subset B$, which means we have proved ④.

* ③ + ④ $\Rightarrow \dfrac{M_n}{\log n} \xrightarrow{a.s} 1$. □