

Föreläsning 1

probability experiment:

Ω sample space

$w \in \Omega$: elementary events (outcomes)

Ex 1: Tossing a coin

$$\Omega = \{H, T\}$$

Ex 2: Tossing two distinguishable coins

$$\Omega = \{HH, HT, TH, TT\}$$

Ex 3: Tossing two indistinguishable coins

$$\Omega = \{HH, HT, TT\}$$

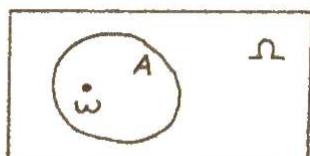
To conduct the experiment = choose one element of Ω
This is the outcome of exp

Consider the case of Ω discrete : either finite or countably infinite

$$\Omega = \{w_1, w_2, \dots\}$$

Def: any subset of (a discrete) Ω are called events.

An event E is set to occur in exp. with outcome $w \in \Omega$
if $w \in A$



Ex 2: $A = \{HH, TT\} \subseteq \Omega$

the coins show the same face

Def: \emptyset : impossible event ($\emptyset \subseteq \Omega$)

Ω : certain event ($\Omega \subseteq \Omega$)

Event A implies B if $A \subseteq B$

$$\bar{A} = A^c = \Omega \setminus A \quad \text{"not } A\text{"}$$

$$A \text{ and } B = A \cap B = AB$$

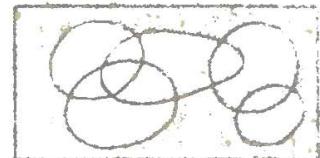
$$A \text{ or } B = A \cup B$$

$$\text{either } A \text{ or } B \Leftrightarrow A \cup B \setminus A \cap B = A \oplus B$$

De Morgan's law:

$$1) \overline{\bigcup_{\alpha \in I} A_\alpha} = \bigcap_{\alpha \in I} \bar{A}_\alpha$$

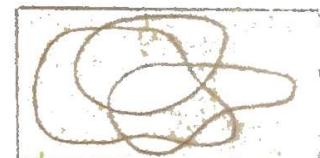
$$2) \overline{\bigcap_{\alpha \in I} A_\alpha} = \bigcup_{\alpha \in I} \bar{A}_\alpha$$



Proof 1)

$$\begin{aligned} \text{Assume } w \in \overline{\bigcup_{\alpha \in I} A_\alpha} &\Rightarrow w \notin A_\alpha \forall \alpha \in I \Rightarrow w \in \bar{A}_\alpha \forall \alpha \\ &\Rightarrow w \in \bigcap_{\alpha \in I} \bar{A}_\alpha \end{aligned}$$

$$\begin{aligned} \text{Assume } w \in \overline{\bigcap_{\alpha \in I} \bar{A}_\alpha} &\Rightarrow w \in \bar{A}_\alpha \forall \alpha \Rightarrow w \notin A_\alpha \forall \alpha \in I \\ &\Rightarrow w \in \bigcup_{\alpha \in I} A_\alpha \end{aligned}$$



Probability

Repeat experiment n times

A : Event $F_n(A)$: frequency of A happening

$$F_n(\emptyset) = 0$$

$$F_n(\Omega) = 1$$

$$F_n(\bar{A}) = 1 - F_n(A)$$

A & B are disjoint events ie they cannot happen together: $A \cap B = \emptyset$

$$F_n(A \cup B) = F_n(A) + F_n(B) \quad \text{Admittedly}$$

Anatoly Kolmogorov's axioms of probability

Prob. is a frequency P of events such that

$$1) 0 \leq P(A) \leq 1 \quad \forall \text{ events } A$$
$$P(\emptyset) = 0 \quad P(\Omega) = 1$$

$$2) P(A \cup B) = P(A) + P(B) \quad \forall \text{ events } A, B \text{ s.t. } AB = \emptyset$$

2*) A countable sequence of disjoint events A_1, A_2, \dots
s.t. $A_i \cap A_j = \emptyset \quad \forall i \neq j$

$$P(\cup A_i) = \sum P(A_i) \quad \text{Countable additivity}$$

Ex.

Ω	H	T
P	$\frac{1}{2}$	$\frac{1}{2}$

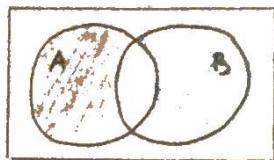
If Ω is countable, by 2*) it is suff. to def P on the elements of events

$$P\{\omega_i\} = p_i \quad 1 = P\{\Omega\} = P(\cup \{\omega_i\}) = \sum P(\{\omega_i\}) = \sum p_i = 1$$

Ω	H	T
P	0.49	0.51

non-symmetric coin

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



Probability is a measure s.t. total length is 1

Measures: length, mass, counts

Equiprobable space

Ω is finite, $|\Omega| = n$

All ω have the same probability $P(\omega) = \frac{1}{n}$

$$P(A) = \sum_{\omega \in A} P(\omega) = |A| \cdot \frac{1}{n} = \frac{|A|}{|\Omega|}$$

Föreläsning 2

Def: a system \mathcal{A} of subsets of Ω is called σ -field (or σ -algebra) if

- 1) $\emptyset, \Omega \in \mathcal{A}$
- 2) $A \in \mathcal{A}, \bar{A} \in \mathcal{A}$
- 3) $A_1, A_2, \dots \in \mathcal{A} \Rightarrow \bigcup A_i \in \mathcal{A}, \bigcap A_i \in \mathcal{A}$

Def: a measurable space is a pair $[\Omega, \mathcal{A}]$ where \mathcal{A} is a σ -field of subsets of Ω

a probability space is a triple (Ω, \mathcal{A}, P) where \mathcal{A} σ -field, P prob. det on \mathcal{A}

Kolmogorov's axioms:

... $[\Omega, \mathcal{A}]$ is a measurable space
elements of \mathcal{A} are called events ...

EX. Ω discrete

$\mathcal{F} = 2^\Omega$ set of all subset of Ω is a σ -field

EX. 1 $\Omega = \{H, T\}$

$2^\Omega = \{\emptyset, \{H\}, \{T\}, \{H, T\}\} : \sigma\text{-field}$

EX. 2. $\Omega = \{HH, HT, TH, TT\}$

$2^\Omega = \{\emptyset, \{HH\}, \dots, \{HH, HT\}, \dots, \{HH, HT, TH, TT\}\} \quad \sigma\text{-field}$

$\mathcal{F} = \{\emptyset, \Omega, \{HH, HT\}, \{TH, TT\}\}$ also σ -field

* 2^{Ω^2} coin is H: $\{HH, TH\} \notin \mathcal{F}$

Def: Elements of \mathcal{F} are called (\mathcal{F} -) measurable sets

* is not \mathcal{F} -measurable, but is 2^Ω -measurable

Ex. $\Omega = \mathbb{R}$

$[a, b] \cap (a, b) \cap [a, b] \cap (a, b)$ must be events

$R = \{[a, b], (a, b), [a, b], (a, b) \mid -\infty \leq a < b \leq +\infty\}$
not a σ -field

Def: Let A be some system of sets.

The minimal σ -field containing A is called
the σ -field generated by A

$$\sigma(A) = \bigcap_{F \ni A} F \quad \begin{matrix} \text{not intersection of sets,} \\ \text{intersection of set of sets} \end{matrix}$$

Exercise: Show that the intersection of σ -field is a σ -field

Let (X, \mathcal{O}) be a topological space

\mathcal{O} : system of 'open' subsets of X

Def: a Borel σ -field is $\sigma(\mathcal{O})$

in particular: the Borel σ -field in \mathbb{R} is the one generated
by $\{(a, b), a \leq b\}$

$$\sigma(\{(a, b) \mid a \leq b\})$$

$$[a, b] = \bigcap_k (a - \frac{1}{k}, b + \frac{1}{k}) \quad \text{Borel set}$$

$$(a, b) = \bigcup_k [a + \frac{1}{k}, b - \frac{1}{k}]$$

Let $F: \mathbb{R} \rightarrow [0, 1]$ be a fct s.t.

$$F(-\infty) = 0 \quad F(\infty) = 1 \quad F \nearrow$$

and right cont: $F(x+) = \lim_{\epsilon \rightarrow 0} F(x + \epsilon) = F(x) \quad \forall x$

$$\text{Define } P(a, b) = F(b) - F(a)$$

Check that P satisfies Kolmogorov's axioms

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

Similarly take $F(x) = x$, $P((a, b)) = F(b) - F(a) = b - a$
the length of the interval.

Carathéodory's extension thm

a system R of sets is called a semiring if $R_1, R_2, \dots \in R \Rightarrow$

$$1) A, B \in R \Rightarrow AB \in R$$

$$2) A \setminus B = \bigcup_{i=1}^n C_i \quad C_i \in R \text{ disjoint}$$

Thm

a countably additive fct defined on a semiring can be extended to σ -field generated by the semiring

So, to a measure

measure seen as prob but without requirement $\mu(\Omega) = 1$

$$1) \mu(\emptyset) = 0, \mu(A) \geq 0 \quad \forall A \in P$$

$$2) \mu(\bigcup A_i) = \sum \mu(A_i) \quad \forall A_i \text{ disjoint}$$

Lettre μ measure: extension to Borel σ -field of the length λ on \mathbb{R}

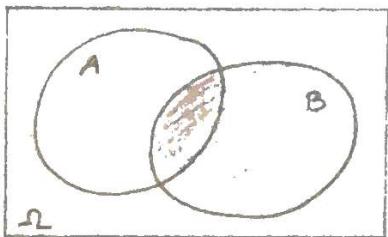
Conditional probability

(Ω, \mathcal{F}, P) prob. space

Let $A, B \in \mathcal{F}, P(B) > 0$

Def : cond. prob. of 'A given B' is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$



Ex. $Q = \{\text{Queen is drawn}\}$ $S = \{\text{Spade is drawn}\}$

$$P(Q \cap S) = 1/52 \quad P(Q) = 4/52$$

$$P(Q \cap S | Q) = \frac{1/52}{4/52} = \frac{1}{4}$$

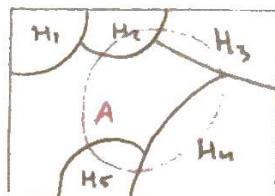
Joint prob. formula

If $P(B) > 0$ then $P(AB) = P(A|B) P(B)$

Def A system of sets H_1, H_2, \dots is called a partition of Ω if

1) $H_i \cap H_j = \emptyset \quad \forall i \neq j$

2) $\bigcup_i H_i = \Omega$



Full prob. formula

If $\{H_1, H_2, \dots\}$ partition s.t $P(H_i) > 0 \quad \forall i$ then

$$P(A) = \sum_i P(A|H_i) P(H_i) \quad \text{a prob.}$$

$$P(A) = P(\bigcup_i A \cap H_i) = \sum_i P(A \cap H_i) = \sum_i P(A|H_i) P(H_i)$$

Bayes rule

$$P(H_i | A) = \frac{P(A|H_i)}{P(A)} = \frac{P(A|H_i) P(H_i)}{\sum_i P(A|H_i) P(H_i)} \quad \text{a posteriori}$$

Föreläsning 3

(Ω, \mathcal{F}, P) : probability space

sample space
events

EX. $\Omega = \{HH, HT, TH, TT\}$: fair coin

$A = \text{"1st coin Head"} = \{HH, HT\}$

$B = \text{"2nd coin Tail"} = \{HT, TT\}$

log. of: there are 4 outcomes

check: $P(A|B) = \frac{P(HT)}{P(HT, TT)} = \frac{1/4}{1/4 + 1/4} = \frac{2}{4} = \frac{1}{2}$ ✓

Def: Two events $A \wedge B$ are called independent if (pairwise indep.)

$$P(AB) = P(A)P(B)$$

Def: A_1, A_2, A_3, \dots are called mutually independent if

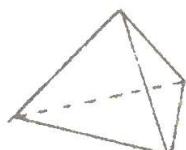
$\forall n \geq 2$ and any (i_1, \dots, i_n)

$$P(\bigcap_{j=1}^n A_{i_j}) = \prod_{j=1}^n P(A_{i_j})$$

EX. Pairwise indep $\not\Rightarrow$ mutual indep

Consider a right tetrahedron

Each of 4 faces is colored r, g, b, rgb



Event: R = 'face touching ground contains Red'

G = ' — II — Green'

B = ' — II — Blue'

Claim: R, G, B are pairwise indep but not mutually

$$P(R) = P(G) = P(B) = 1/2$$

$$P(RG) = P(\text{rgb-face}) = \frac{1}{4} = P(R)P(G)$$

Same for RB, GB

$$P(RGB) = P(\text{rgb-face}) = \frac{1}{4} \neq P(R)P(G)P(B) = \frac{1}{8}$$

Notation: $A \perp\!\!\!\perp B$ pairwise indep

PRODUCT SPACES

Two indep experiments : $(\Omega_1, \mathcal{F}_1, P_1)$ $(\Omega_2, \mathcal{F}_2, P_2)$

Result : (ω_1, ω_2)

Conducting two exp. corresponds to $\Omega = \Omega_1 \times \Omega_2$

$$A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2$$

$$A_1 \times A_2 = \{(\omega_1, \omega_2) \mid \omega_1 \in A_1, \omega_2 \in A_2\} \quad \text{Should be events}$$

↑
System of rectangles $A_1 \times A_2$ does not constitute a σ -field

Why?

Complement is not a rectangle

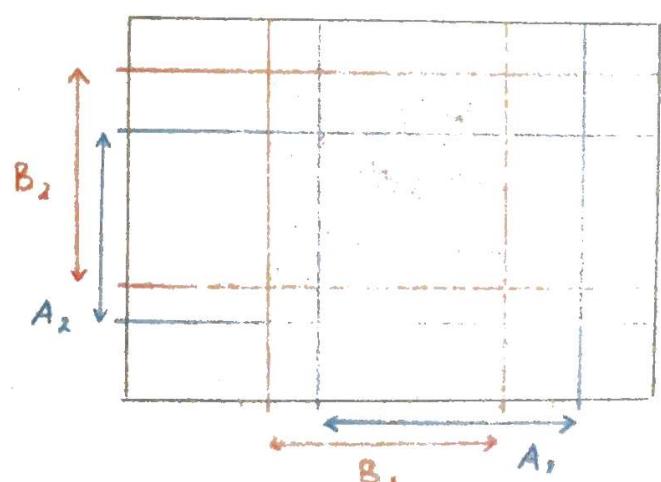
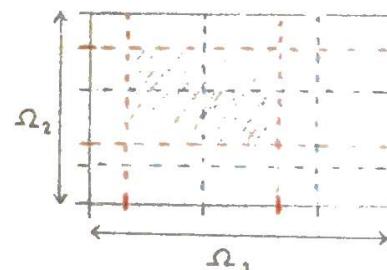
However it is a semi-ring

$$(A_1 \times A_2) \cap (B_1 \times B_2) = A_1 B_1 \times A_2 B_2$$

= rectangle

$$B \setminus A = \bigcup_{i=1}^n C_i$$

= union of disjoint rect.



$$f_1 \otimes f_2 = \sigma(R)$$

$[\Omega_1 \times \Omega_2, f_1 \otimes f_2]$ measurable space

$$\{w_1 \in A_1\} = A_1 \subseteq \Omega_1$$

$$\{w_1, w_2 \in ?\} = \{w_1 \in A_1, w_2 \in \Omega_2\}$$

$$\subseteq \Omega_2$$

$$\{w_2 \in A_2\} \in f_2 \Rightarrow \Omega_1 \times A_2 \in f_1 \otimes f_2$$



We want to define P on $f_1 = f_1 \otimes f_2$ s.t

$$\begin{aligned} P(w_1 \in A_1, w_2 \in A_2) &= P(A_1 \times A_2) = P(\Omega_1 \times \Omega_2) P(\Omega_1 \times A_2) \\ &= P_1(A_1) P_2(A_2) \end{aligned}$$

This defines P on \mathbb{R} .

By Kolmogorov's Extension Thm P extends to $f = f_1 \otimes f_2 = \sigma(R)$
Carathéodory's

Ex. Tossing 2 coins

$$\Omega_1 = \{H, T\} = \Omega_2$$

$$P_1(H) = p_1, \quad P_2(H) = p_2 \quad \text{need not be the same}$$

$$\Omega = \Omega_1 \times \Omega_2 = \{HH, HT, TH, TT\}$$

$$P(HH) = P(H \times \Omega_2 \cap \Omega_1 \times H) = P_1(H) P_2(H) = p_1 p_2$$

Ω	HH	HT	TH	TT
P	$p_1 p_2$	$p_1(1-p_2)$	$(1-p_1)p_2$	$(1-p_1)(1-p_2)$

n indep exp:

$$\Omega = \prod_{i=1}^n \Omega_i \quad \mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2 \otimes \dots \otimes \mathcal{F}_n \\ = \sigma(\{\prod_{i=1}^n A_i \mid A_i \in \mathcal{F}_i\}) = \mathcal{F}$$

$$P = \prod_{i=1}^n P_i \quad \text{where } P(A_1 \times \dots \times A_n) = \prod_{i=1}^n P_i(A_i) \quad A_i \in \mathcal{F}_i$$

(expansion to \mathcal{F} at \mathcal{C})

Let now $(\Omega_i, \mathcal{F}_i, P_i)$ be an int seq. of exp.

$$\Omega = \prod_{i=1}^{\infty} \Omega_i = \{(w_1, w_2, \dots) \mid w_i \in \Omega_i, i \in \mathbb{N}\}$$

Def: a cylinder event is an event of the form

$$\Omega_{i_1} \times \Omega_{i_2} \times \dots \times \Omega_{i_n} \times \dots$$

$i_1 \quad i_2 \quad i_3 \quad \dots \quad i_n$ finite r.

$$\Omega_{i_1} \times \Omega_{i_2} \times \dots \times \Omega_{i_{n-1}} \times A_{i_n} \times \Omega_{i_{n+1}} \times \dots \times \Omega_{i_{n+1}} \times A_{i_n} \times \dots \\ \times A_{i_m} \times \Omega_{i_{m+1}} \times \dots \in \mathcal{F}_{i_1} \subset \mathcal{F}_{i_2}$$

The system of cyl. events is a semiring \mathcal{C}

$$\mathcal{F} = \sigma(\mathcal{C}) \quad P\{w_i \in A_{i_j}, j=1, \dots, n\} = \prod_{j=1}^n P_{i_j}(A_{i_j})$$

Notice that independence only intervenes in (*)

So a general prob. measure P can be def first on \mathcal{C} and then extended to $\mathcal{F} = \sigma(\mathcal{C})$ provided it is consistent

$$P(w_{i_1} \in A_{i_1}, \dots, w_{i_n} \in A_{i_n}) = P(w_{i_1} \in A_{i_1}, \dots, w_{i_m} \in A_{i_m})$$

L could be any
not nec. last

EX. $\Omega_1 = \Omega_2 = \{H, T\}$ consistency

Consistent def. of joint prob. P

$$P(H \times \Omega_2) = p_1,$$

$$P(\Omega_1 \times H) = p_2$$

want to def: $P(HH) = p_{11}$

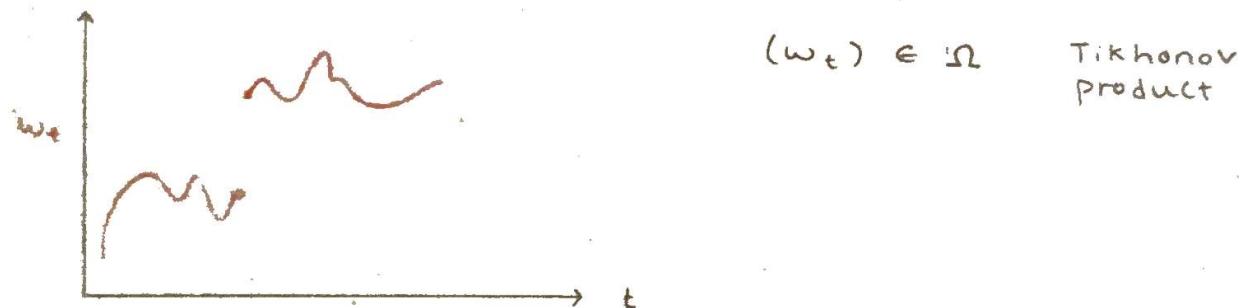
also: $P(HT) = p_{12}$ $P(TH) = p_{21}$ $P(TT) = p_{22} = 1 - p_{11} - p_{12} - p_{21}$

$$\Rightarrow P(H \times \Omega_2) = P(HH) + P(HT) = p_{11} + p_{12}$$

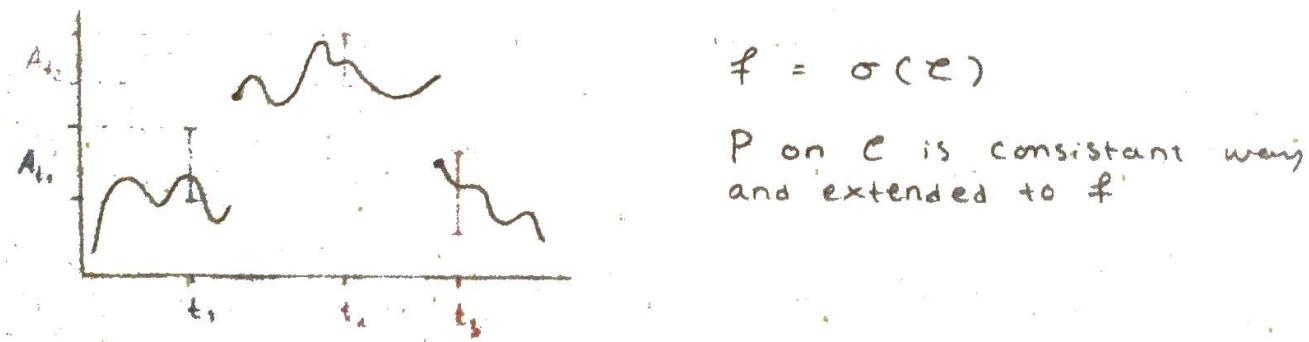
$$P(\Omega_1 \times H) = P(HH) + P(TH) = p_{11} + p_{21}$$

(w_1, w_2, \dots) seq. (w_t) $t \in \mathbb{N}$ countable int

(w_t) , $t \in \mathbb{R}_t = [0, \infty)$ uncount. int



$$\mathcal{C} = \{w_{t_i} \in A_{t_1}, \dots, w_{t_n} \in A_{t_n} \mid t_1, \dots, t_n \in I\}$$



Föreläsning 4

(cadlag fct : continuous from right, limits from left)

RANDOM VARIABLES

(Ω, \mathcal{F}, P)

EX. Ω HH HT TH TT

P $\frac{1}{4}$ $\frac{1}{4}$ $\frac{1}{4}$ $\frac{1}{4}$

ξ 0 1 1 2

How many Tails the experiment yields.

$\xi = \xi(\omega), \omega \in \Omega$

Def. for a discrete Ω a random variable is a fct $\Omega \rightarrow \mathbb{R}$

$\Omega = \{\omega_1, \omega_2, \dots\} \quad \xi : \xi(\omega_1), \xi(\omega_2), \dots$ discrete

$P(\xi = x) = P\{\omega : \xi(\omega) = x\}$

EX 1: $P(\xi = 1) = P\{HT, TH\} = \frac{1}{2}$

ξ	0	1	2
P_ξ	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Distribution
table

Def: $P_\xi(x) = P\{\xi = x\}, x \in \{\xi(\omega), \omega \in \Omega\}$

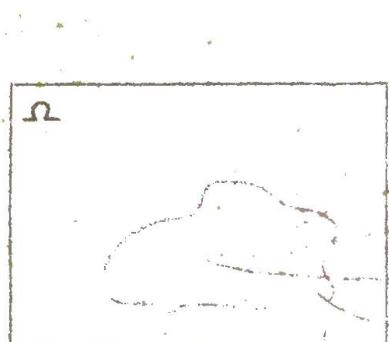
Viewed as a fct of x is called the probability mass fct
of ξ (pmf)

EX. $P_\xi(0) = P_\xi(2) = \frac{1}{4} \quad P_\xi(1) = \frac{1}{2}$

Ω NOT COUNTABLE

We should be able to evaluate the probabilities that ξ takes values in our sets of interest - e.g. Borel sets:

$$P\{\xi \in B\} = P\{\omega : \xi(\omega) \in B\}$$



$$P\{\omega : \xi(\omega) \in B\} = \xi^{-1}(B)$$

... inverse image

$P\{\xi^{-1}(B)\}$ is defined if $\xi^{-1}(B) \in \mathcal{F}$

Def: Let $[X, \mathcal{F}]$, $[Y, \mathcal{G}]$ be two measurable spaces.

A fct $f: X \rightarrow Y$ is called $[X, \mathcal{F}]$ -measurable if
 $f^{-1}(B) \in \mathcal{F} \quad \forall B \in \mathcal{G}$

Def: A random variable on $[\Omega, \mathcal{F}, P]$ is a $[\mathcal{F}, \mathcal{B}]$ -measurable mapping $\Omega \rightarrow \mathbb{R}$

For a discrete Ω , $\mathcal{F} = 2^\Omega$. $\forall B \in \mathcal{B}$. $\xi^{-1}(B)$ is a subset of Ω so it is also discrete $\in 2^\Omega$.

So any ξ on a disc. Ω is measurable, so a rand var

$$\Omega = \{HH, HT, TH, TT\}$$

$$\mathcal{F} = \{\emptyset, \Omega, \{HH, HT\}, \{TH, TT\}\}$$

ξ = no. of tails.
Is it \mathcal{F} meas?

$$\xi^{-1}\{\cdot\} = \{HT, TH\} \neq \emptyset$$

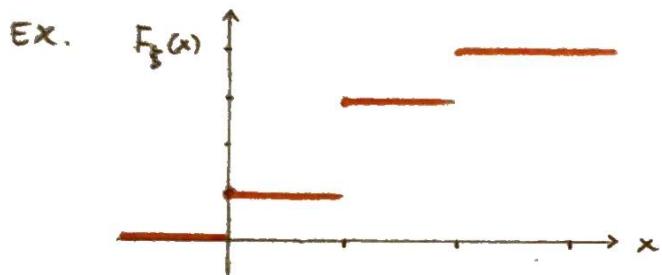
\Rightarrow NOT a random variable on (Ω, \mathcal{F})



$$\text{Ex: } \Omega = \{(\omega_1, \omega_2) \in \mathbb{R}^2 : \omega_1 > 0, \omega_2 > 0\} \quad P\{\xi \in B\} = \sum_{x \in \xi \cap B} P_\xi(x)$$

Def: Cumulative distribution function of a rand var ξ is

$$\begin{aligned} F_\xi(x) &= P\{\xi \leq x\} \quad x \in \mathbb{R} \\ &= P\{\omega : \xi(\omega) \leq x\} = P\{\xi^{-1}(-\infty, x]\} \end{aligned}$$



ξ : no of tails

$$F_\xi(0) = P\{\xi \leq 0\} = P(\xi = 0) = P(HH)$$

$$F_\xi(1) = P\{\xi \leq 1\} = P(HH \cup HT \cup TH)$$

- $F_\xi(-\infty) = P\{\xi \leq -\infty\} = 0$
- $F_\xi(+\infty) = P\{\xi \leq +\infty\} = 1$
- $F_\xi \nearrow$
- $x_2 > x_1 \Rightarrow \{\xi \leq x_1\} \subseteq \{\xi \leq x_2\} \Rightarrow F_\xi(x_1) \leq F_\xi(x_2)$

$F_\xi(x)$ is right cont so $\forall x \& x_n \downarrow x$ monotone $F_\xi(x_n) \rightarrow F_\xi(x)$
 $P(\xi \leq x_n)$ vs $P(\xi \leq x)$



$$\text{Tot: } \{\xi \leq x\} = \bigcap_n \{\xi \leq x_n\}$$

$$\{\xi \leq x_{n+k}\} \subseteq \{\xi \leq x_n\} \quad \forall n, \forall k \geq 1$$

By continuity of a probability measure:

$$P\{\xi \leq x_n\} = \lim P\{\xi \leq x_n\} = \lim F_f(x_n)$$

so $F_f(x) = P\{\xi \leq x\}$

$$y_n \uparrow x : \{\xi \leq x\} = \bigcup_n \{\xi \leq y_n\}$$

$$P\{\xi \leq x\} = \lim P\{\xi \leq y_n\} = \lim F_f(y_n) = F_f(x)$$

$$\begin{aligned} P\{\xi \in [a, b]\} &= P\{\xi \leq b, \text{not } \xi \leq a\} = P\{\xi \leq b\} - P\{\xi \leq a\} \\ &= F_f(b) - F_f(a^-) \end{aligned}$$



This way we define $P\{\xi \in [a, b]\}$ so we define a measure on semiring \mathcal{R} which extends to Borel set \mathbb{B}

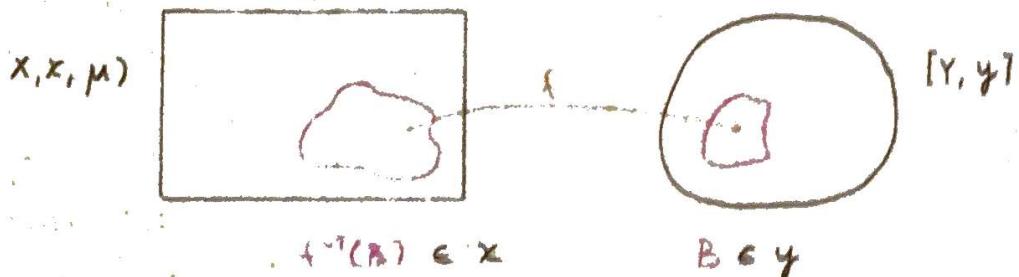
$$P\{\xi \in B\} = P\{\xi^{-1}(B)\} \quad \forall B \in \mathbb{B}$$

Def: Let (X, \mathcal{X}, μ) be a measurable space
 (Y, \mathcal{Y}) a meas. space

$f: (X, \mathcal{X})$ meas mapping $f: X \rightarrow Y$

A push forward measure $v = \mu f^{-1}$ is def by

$$v(B) = (\mu f^{-1})(B) = \mu(f^{-1}(B)) \quad \forall B \in \mathcal{Y}$$



$$\mu(f^{-1}(B)) = (\mu f^{-1})(B) = v(B)$$

Def : The distribution (measure) P_ξ of a
is the push-forward of P by ξ :

$$P_\xi(B) = P\{\xi^{-1}(B)\} = P(\omega : \xi(\omega) \in B) = P\{\xi \in B\}$$

$B \in \mathcal{B}$

We have just established that P_ξ is def by the cdf of F_ξ

on discrete spaces:

$$P\{\xi=x\} = P\{\xi \in [x, x]\} = F_\xi(x) - F_\xi(x^-)$$

the jump at x
zero if x continuity point

Def: A n.v is called continuous if F_ξ is a cont. fct.

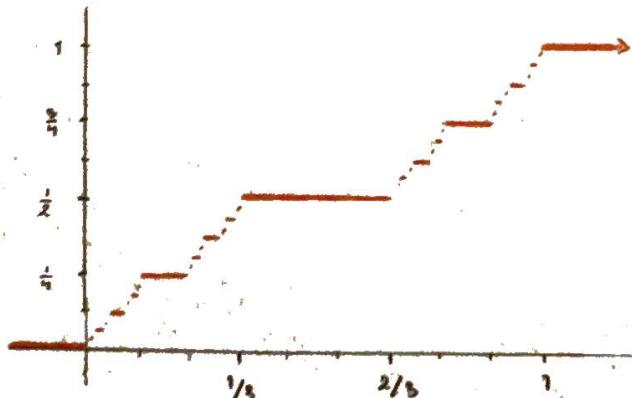
If in addition $\exists f'_\xi(x) = F'_\xi(x) = \int_{-\infty}^x f_\xi(t) dt$ Leibnitz
the n.v ξ (or its distribution) is called absolutely continuous
otherwise singular

such $f_\xi(x)$ is called the prob. density. fct (pdf)

$$P\{\xi \in (x - \frac{\alpha}{2}, x + \frac{\alpha}{2})\}$$

P { }

$$F_\xi \in C^\circ \text{ but } \int_{-\infty}^x F'(t) dt \neq F_\xi(x)$$



$F'(x) = 0$ almost everywhere
except countable set
so integral is zero

The sure thing principle

If you prefer x to y given C , and you prefer x to y given C^c then you always prefer x to y .

Each morning: do you prefer coffee (x) to tea (y)?

answer depends on if you slept well (c)

X : event that at random day you prefer coffee

$$P(X) = p(x|c) + p(x|c^c) = \frac{p(x|c)}{p(c)} + \frac{p(x|c^c)}{p(c^c)}$$

:

$$p(x|c) > p(y|c), \quad p(x|c^c) > p(y|c^c)$$

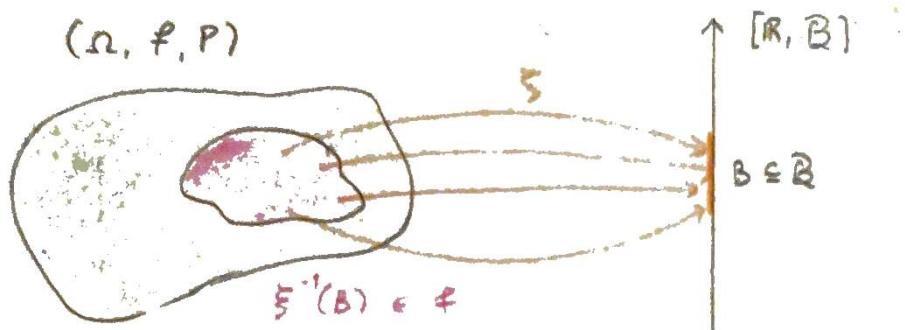
$$\frac{p(x|c)}{p(c)} > \frac{p(y|c)}{p(c)}, \quad \frac{p(x|c^c)}{p(c^c)} > \frac{p(y|c^c)}{p(c^c)}$$

Multiply & add together:

$$p(x|c) + p(x|c^c) > p(y|c) + p(y|c^c)$$

$$p(x) > p(y)$$

Föreläsning 5



ξ : a measurable mapping $\Omega \rightarrow \mathbb{R}$ \cong Random Variable

Distribution P of ξ is the push forward of P on $[\mathbb{R}, \mathcal{B}]$

$$P_\xi(B) = P(\xi^{-1}(B)) = P\{\omega : \xi(\omega) \in B\}$$

Different mappings $(\Omega_\alpha, \mathcal{F}_\alpha, P_\alpha) \xrightarrow{\xi_\alpha} \mathbb{R}$ may however have the same distribution

Even if the mapping ξ is not defined, but the distribution Q is given, one can always define ξ on $(\mathbb{R}, \mathcal{B}, Q)$ and $\xi(\omega) = \omega$ identical map

$$P_\xi(B) = Q(\xi^{-1}(B)) = Q(B)$$

$$\Rightarrow P_\xi = Q$$

Such rand. var. are called canonically defined

MIXTURES

Let $\xi_1, \xi_2, \dots, \xi_n$ be rand var with distributions $P_{\xi_1}, \dots, P_{\xi_n}$

Let v be another r.v. with ^{indep of all ξ_i} distr. Q s.t. $Q_n\{v=k\} = P_k$

$$\sum_{k=1, \dots, n} P_k = 1$$

A mixed r.v. η is realised as:

- 1) $v(\omega)$ is drawn
- 2) $\eta = \xi_{v(\omega)}$ is realised

$$\eta = \sum_{k=1}^n \xi_k \mathbb{1}\{v=k\}$$

$$\begin{aligned} F_\eta(x) &= P(\eta \leq x) = \sum_{k=1}^n P(\eta \leq x | v=k) P(v=k) \\ &= \sum_{k=1}^n P\{\xi_k \leq x | v=k\} Q\{k\} \\ &= \sum_{k=1}^n p_k F_{\xi_k}(x) \quad \text{CDF of mixture} \end{aligned}$$

Recall that cdf F_ξ is ↗ (monotonically growing) & cont. from right

Then they can be written in the form

$$F_\xi(x) = p_1 F_{\xi_1}(x) + p_2 F_{\xi_2}(x) + p_3 F_{\xi_3}(x)$$

where F_{ξ_1} stepwise const discrete

F_{ξ_2} absolutely cont. ie $F_{\xi_2}(b) - F_{\xi_2}(a) = \int_a^b f_{\xi_2}(t) dt$

F_{ξ_3} singular ie cont but not absolutely

⇒ Any ξ can be represented as a mixture

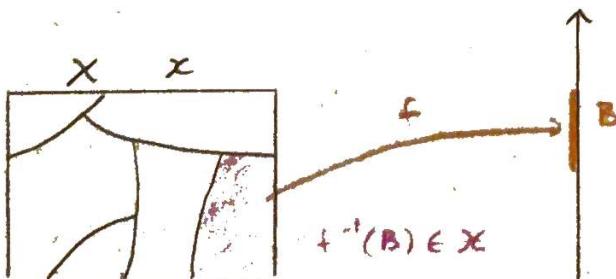
$$\xi = \xi_0 \mathbb{1}\{v=1\} + \xi_1 \mathbb{1}\{v=2\} + \xi_2 \mathbb{1}\{v=3\} \quad P\{v=k\} = p_k$$

LEBESGUE INTEGRAL

(X, \mathcal{X}, μ) finite measure space

$$\mu(X) < \infty$$

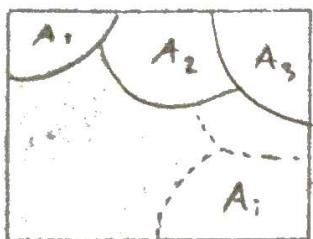
Let $f: X \rightarrow \mathbb{R}$, $\{x \in B\}$ measurable, $f^{-1}(B) \in \mathcal{X}$



Def: A fct. is called simple if it assumes only finite no values.

$$f(x) = \sum_{i=1}^n f_i \cdot 1_{A_i}(x)$$

$$1_B(x) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{otherwise} \end{cases}$$



$$X \setminus \bigcup_{i=1}^n A_i = A_0.$$

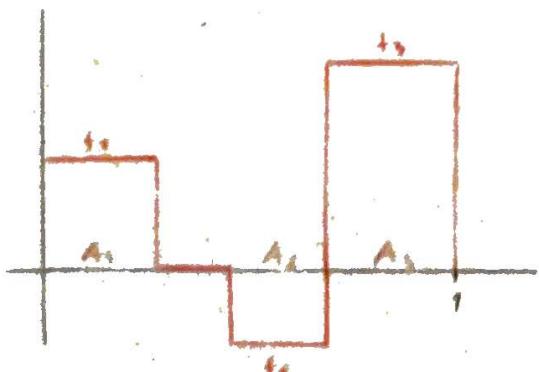
A_0, A_1, \dots, A_n form a partition of X

f is measurable iff $A_i \in \mathcal{X} \quad \forall i = 0, 1, \dots, n$ are meas

Def: The Lebesgue integral of a simple measurable f wrt μ

$$\int f(x) \mu(dx) = \int f d\mu = \mu(f) = \mu f = \langle f, \mu \rangle \text{ notation} \\ := \sum_{i=1}^n f_i \mu(A_i)$$

$$(X, \mathcal{X}, \mu) = ([0, 1], \mathcal{B}, \text{Lebesgue measure})$$

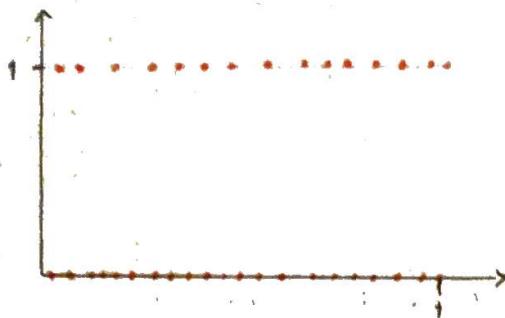


$$\int f dx = \sum f_i \ell(A_i) \\ = \int f(x) dx$$

* Riemann integral

EX. $f = \mathbb{1}_{[0,1] \setminus \mathbb{Q}}$

indicator of irrational points



$\int f dx$ does not exist

$\lim_{\substack{\sup |\Delta_j| \rightarrow 0 \\ x_i \in \Delta_j}} \sum f(x_j) |\Delta_j|$ does not exist
because $x_i \in \mathbb{R} \setminus \mathbb{Q}$

But f is a simple function:

$$f(x) = 1 \cdot \frac{\mathbb{1}_{[0,1] \setminus \mathbb{Q}}(x)}{A_1} + 0 \cdot \frac{\mathbb{1}_{\mathbb{Q} \cap [0,1]}(x)}{A_2}$$

$$\boxed{\int f dx = 1 \cdot \text{f}([0,1] \setminus \mathbb{Q}) = 1}$$

$$\text{f}([0,1] \setminus \mathbb{Q}) = \text{f}([0,1]) - \text{f}(\mathbb{Q} \cap [0,1]) = 1 - \sum \text{f}\{\text{q}_j\} = 1 - 0$$

Properties:

1) If $f \geq 0$ then $\mu(f) \geq 0$

$$\mu(f) = \sum f_i \mathbb{1}_{A_i} \geq 0 \text{ whenever } f_i \geq 0$$

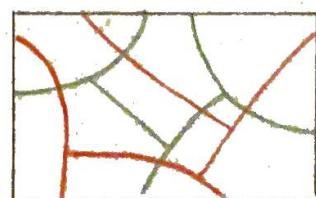
2) $\mu(cf) = c \mu(f)$

3) $\mu(f_1 + f_2) = \mu(f_1) + \mu(f_2)$

PROOF 3): $f_1 = \sum_{i=0}^n a_i \mathbb{1}_{A_i}, f_2 = \sum_{j=0}^m b_j \mathbb{1}_{B_j}$

two partitions

$A_i, B_j \in \mathcal{X}$ are measurable



$\{A_i; B_j\}$ is a partition of X

The value of $f_1(x) + f_2(x)$ for $x \in A_i; B_j$ is $a_i + b_j$

So $f_1 + f_2 = \sum_{i=0}^n \sum_{j=0}^m (a_i + b_j) \mathbb{1}_{A_i B_j}$ is a simple function!

$$\mu(f_1 + f_2) = \sum_i \sum_j (a_i + b_j) \mu(A_i B_j) = \sum_i a_i \sum_j \mu(A_i B_j) + \sum_j b_j \sum_i \mu(A_i B_j)$$

$$A_i = \bigcup_{j=0}^m A_i B_j$$

Since $\{B_j\}$ is a partition

$$B_j = \bigcup_{i=0}^n B_j A_i$$

Since $\{A_i\}$ is a partition

$$\Rightarrow \mu(A_i) = \sum_j \mu(A_i B_j)$$

$$\Rightarrow \mu(f_1 + f_2) = \sum_i a_i \mu(A_i) + \sum_j b_j \mu(B_j) = \mu(f_1) + \mu(f_2)$$

4) $f_1 \geq f_2 \Rightarrow \mu(f_1) \geq \mu(f_2)$

$$0 \leq \mu(f_1 - f_2) \stackrel{?}{=} \mu(f_1) + \mu(-f_2) \stackrel{?}{=} \mu(f_1) - \mu(f_2)$$

$$\Rightarrow \mu(f_1) \geq \mu(f_2)$$

5) $|\mu(f)| \leq \sup |f| \mu(X)$

$$|\mu(f)| = \left| \sum_i a_i \mathbb{1}_{A_i} \right| \leq \sum_i |a_i| \mu(A_i) = \mu(|f|)$$

$$|f| \leq \sup |f| \mathbb{1}_X, \text{ by 4 } |\mu(f)| \leq \sup |f| \mu(X)$$

Def Let $c \in X$

Define $\int_c f d\mu = \int f \mathbb{1}_c d\mu$



$$6) \text{ If } C_1, C_2 \neq \emptyset \text{ then } \int_{C_1 \cup C_2} f d\mu = \int_{C_1} f d\mu + \int_{C_2} f d\mu$$

$$f \mathbb{1}_{C_1 \cup C_2} \leq f \mathbb{1}_{C_1} + f \mathbb{1}_{C_2}$$

$$\mu(\dots) = \int_{C_1 \cup C_2} = \int_{C_1} + \int_{C_2}$$

Let $\Delta \in \mathcal{X}$ be s.t. $\mu(\Delta) = 0$. Then $\underbrace{\int f d\mu}_{\mu(f \mathbb{1}_\Delta)} = 0$

$$|\mu(f \mathbb{1}_\Delta)| \stackrel{s}{\leq} \sup |\mathbb{1}| \mu(\Delta) = 0$$

$$\int_{C \cup \Delta} f d\mu \stackrel{6}{=} \int_C f d\mu + \int_\Delta f d\mu = \int_C f d\mu \quad (\text{if } \Delta \subset \emptyset)$$

Take f and change it to g on Δ

$$f \mathbb{1}_{X \setminus \Delta} + g \mathbb{1}_\Delta$$

$$\int f \mathbb{1}_{X \setminus \Delta} + g \mathbb{1}_\Delta d\mu = \int f d\mu - \cancel{\int f d\mu} + \cancel{\int g d\mu}$$

so $\mu(f)$ is insensitive wrt change of f on a set of μ -measure 0

This is why $\int \mathbb{1}_{[0,1] \cap Q} dt = \int 1 dt = 1$ and $\lambda([Q \cap [0,1]])$

Step 2 :

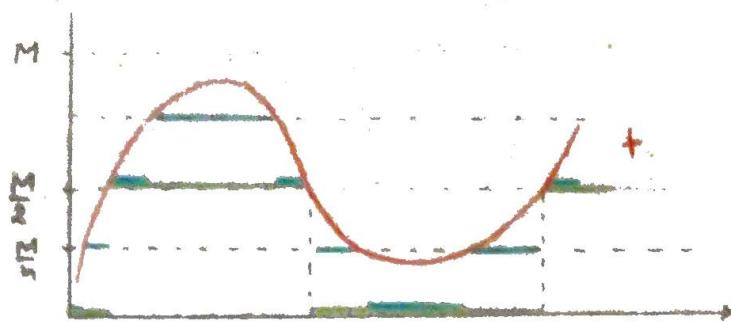
Let f be a measurable bounded fct $f \geq 0$

Let f_n be a sequence of simple fcts $f_n \geq 0$ monotonely uniformly converging to f

$$f_n \nearrow f \Rightarrow f_{n+k} \geq f_n \quad \forall n, k$$

$$\sup_x (f - f_n) \xrightarrow{n \rightarrow \infty} 0$$

Define $\int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu$



$$A_{n,k} = \left\{ x : M \frac{k-1}{2^n} \leq f(x) \leq M \frac{k}{2^n} \right\} \quad k=1, \dots, 2^n$$

$$f_1 = \sum M \frac{k-1}{M} \mathbb{1}_{A_{n,k}} \quad f_2$$

$$\mu(f_n) \leq \sup f_n \mu(x) \leq \sup f \mu(x) \leq M \mu(x)$$

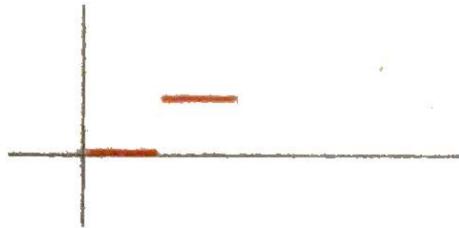
$\mu(f_n) \nearrow$ because $f_n \nearrow$

$$\Rightarrow \lim \mu(f_n) \text{ exists} \leq M \mu(x)$$

Föreläsning 6

1) $\int f d\mu(x) = \int f d\mu = \mu(f) = \mu^f = \langle f, \mu \rangle$

$$f(x) = \sum_{i=1}^n f_i \mathbb{1}_{A_i}(x) \quad \text{Simple}$$



$$\mu(x) < \infty$$

$$\mu(f) = \sum_{i=1}^n f_i \mu(A_i)$$

2) $f \geq 0 \quad |f| \leq M \quad f_n \nearrow f \quad \sup |f_n - f| = 0$
 $|f_n| \leq 2M \quad \leftarrow \text{Simple facts}$

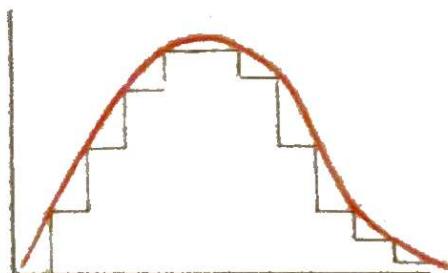
$$\mu(f) = \lim_{n \rightarrow \infty} \mu(f_n) \quad \mu(f_n) \nearrow \quad \mu(f_n) \leq M$$

$$g_n \nearrow f \quad |g| \leq 2M$$

Simple

$$\begin{aligned} |\mu(f_n) - \mu(g_n)| &= |\mu(f_n - g_n)| \leq \sup |f_n - g_n| \mu(x) \\ &\leq (\sup |f_n - f| + \sup |g_n - f|) \mu(x) \rightarrow 0 \end{aligned}$$

Def. does not depend on the covering sequence $f_n \nearrow f$

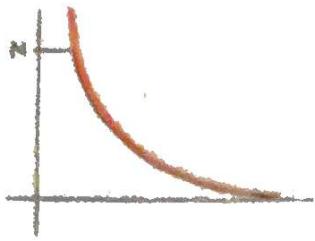


$$(x - x_n) \leq \frac{M}{2^n}$$

$$f_n \nearrow f$$

3) If f not bounded, consider $f_N(x) = \begin{cases} \inf(x) = M & \\ \text{Otherwise} & \end{cases} = \min \{f, N\}$

$$f_N \nearrow f \quad \text{Define } \mu(f) = \lim_{N \rightarrow \infty} \mu(f_N)$$



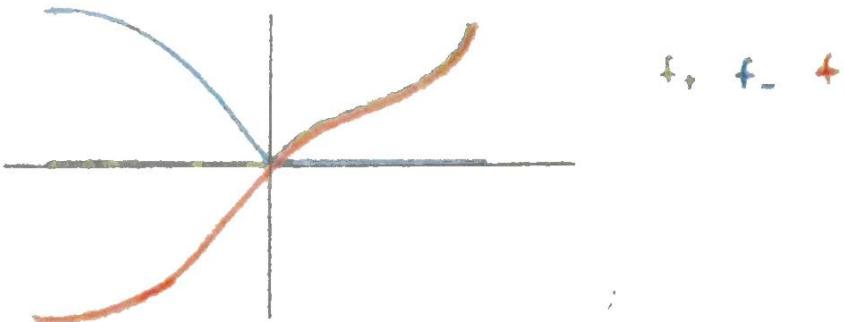
4) If f is not necessarily non-neg, consider

$$f = f_+ - f_-$$

$$f_+ = \max\{f, 0\} \geq 0$$

$$f_- = -\min\{f, 0\} \geq 0$$

Define $\mu(f) = \mu(f_+) - \mu(f_-)$ it exists



5) $\mu(x) = +\infty$ but $X = \bigcup_{n=1}^{\infty} X_n$

where $X_n \subseteq X_{n+1}$ s.t $\mu(X_n) < \infty$

$$\mu(f) = \lim_{n \rightarrow \infty} \mu(f \mathbf{1}_{X_n}) \quad \text{it exists}$$

EX. 1: Lebesgue measure on \mathbb{R}

$$R = \bigcup_{n=1}^{\infty} [-n, n] \quad \ell([-n, n]) = 2n < \infty$$

$$\int f d\ell = \int f \mathbf{1}_{[-n,n]} d\ell = \int_{[-n,n]} f d\ell = \int_n f d\ell$$

1) $\mu(f) \geq 0 \quad \forall f \geq 0$ simple

Take $f_n \nearrow f$ bounded

$$0 \leq \mu(f_n) \nearrow \mu(f)$$

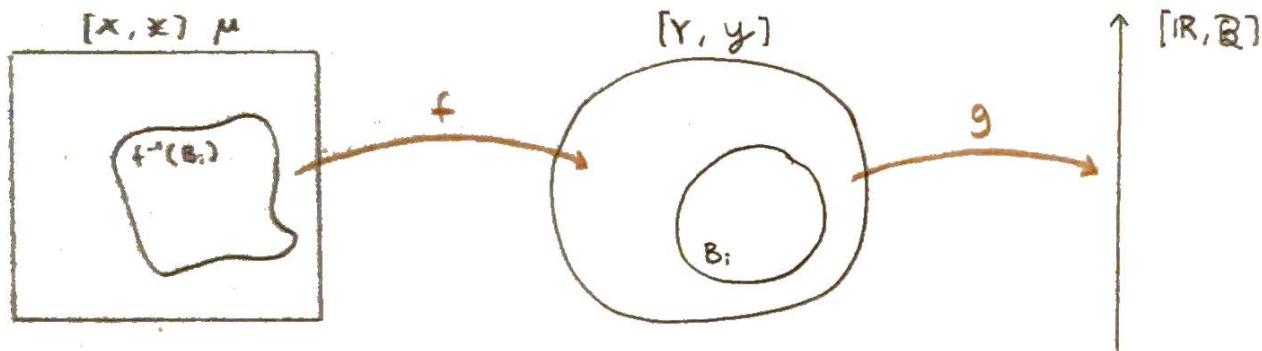
Similarly take \lim in steps 3, 5 we get $\mu(f) \geq 0 \quad \forall f \geq 0$

We have shown linearity $\mu(cf + g) = c\mu(f) + \mu(g)$ for $f, g \in \text{Simple}$

But \lim in steps 2-5 is a linear operation, so taking \lim in 2-5 we get $\mu(f)$ for all functions where the integrals exist

This is the Monotone Class Argument (MCA)

CHANGE OF MEASURE



$$g(f(x)) = g(f(x))$$

$$\int g(f(x)) \mu(dx)$$

$$\text{Let } g = \sum_{i=1}^n g_i \mathbb{1}_{B_i}(y) \quad \text{or simpler: } g(y) = \mathbb{1}_{B_1}(y)$$

$$g(f(x)) = \mathbb{1}_B(f(x)) = \mathbb{1}_{f^{-1}(B_1)}(x)$$

$$= \int \mathbb{1}_{f^{-1}(B_1)}(x) \mu(dx) = \mu(f^{-1}(B_1)) = (\mu f^{-1})(B_1) = \int \mathbb{1}_{B_1} d(\mu f^{-1})$$

$$= \int g(y) (\mu f^{-1})(dy)$$

True for indicators \rightarrow True for simple fct
 \rightarrow True for all fct

By MCA : $\int (gf)(x) \mu(dx) = \int g(x) (\mu + \nu)(dx)$

$$\int_A f d\mu = \int_A f \mathbb{1}_{A \cup B} d\mu = \int_A f d\mu + \int_B f d\mu$$

$$\int_A f d\mu = 0 \quad \text{if} \quad \mu(A) = 0 \quad f = \sum_i f_i \mathbb{1}_{A_i}, \quad \mu = \sum_i f_i \underline{\mu(A_i, A)} = 0 \quad \text{By MCA}$$

$$\therefore \mu(A) = 0$$

Def: functions f and g are called μ -equivalent if

$$\mu \{x : f(x) \neq g(x)\} = 0$$

\Rightarrow

$$\mathbb{1}(x) \underset{\mu(\{x\})=0}{\cong} \mathbb{1}_{R \setminus Q}(x) \text{ on } \mathbb{R} \quad \text{L. Lebesgue because } \ell(Q) = 0$$

$$\mu(f) - \mu(g) = \mu(f-g) = \mu \underbrace{(f-g) \mathbb{1}_A}_{\mu(\{x\})=0} + \mu \underbrace{(f-g) \mathbb{1}_{X \setminus A}}_{=0} = 0$$

Let μ be an atomic (discrete) measure

$$\mu = \sum_i m_i \delta_{x_i}$$

$m_i > 0$, \sum finite or inf, $\{x_i\}$ at most countable

$$\delta_x : \text{Dirac measure}, \quad \delta_x(B) = \mathbb{1}_B(x) = \begin{cases} 1 & x \in B \\ 0 & x \notin B \end{cases}$$

$$\int f d\mu = \int \underset{U\{x_i\}}{f d\mu} + \int \underset{X \setminus U\{x_i\}}{f d\mu}^0 \quad \text{since } \mu(X \setminus U\{x_i\}) = 0 \text{ zero measure}$$

$$= \sum_i \int_{\{x_i\}} f(x) d(m_i \delta_{x_i}) = \sum_i f(x_i) m_i = \sum f(x_i) m_i$$

Simple taking values
 $f(x_i)$ on $\{x_i\}$

may not exist.

Def : Let v be some 'reference' measure on $[x, x]$

A measure μ is absolutely continuous wrt v if \exists density fct $\varphi(x)$ s.t

$$\mu(B) = \int_B \varphi(x) v(dx) \quad \forall B \in \mathcal{X}$$

Such φ is def up to a set of v -measure 0

Ex. A uniform distribution $u(dx)$ on $[a, b] \subset \mathbb{R}$ is def by the density $f_v(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$ wrt Leb. measure (v)
 $a < c < b < d$

$$U[c, d] = \int_{[c,d]} f_v(x) u(dx) = \int_c^d \frac{1}{b-a} dx = \frac{d-c}{b-a}$$

Ex. $N(0, 1)$ - density

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Let φ_v be the density of μ wrt v ($\varphi_v = \frac{d\mu}{dv}$)

$$\text{Then } \int_x g d\mu = \int_x g \varphi_\mu dv$$

proof Take $g(x) = \mathbb{1}_B(x)$

$$\int g d\mu = \mu(B) = \int_B \varphi_\mu dv = \int \mathbb{1}_B \varphi_\mu dv = \int g \varphi_\mu dv$$

MCA \Rightarrow True $\forall g$

LEBESGUE THEOREMS

Thm 1 (Monotone convergence)

If $f_n \nearrow f$ then $\mu(f_n) \nearrow \mu(f)$ technically part of def.

Thm 2 (Dominated/bounded/majorated conv)

$|f_n| \leq g$, $f_n \rightarrow f$, $\mu(g) < \infty$

Then $\mu(f_n) \rightarrow \mu(f) < \infty$

Föreläsning 7

NUMERIC CHARACTERISTICS OF RAND VAR

$$\xi : \Omega \rightarrow \mathbb{R} \quad \text{nv}$$

$$(\Omega, \mathcal{F}, P) \xrightarrow{\xi} [\mathbb{R}, \mathcal{B}]$$

$$P_\xi(B) = P(\xi^{-1}(B)) = P(\xi \in B) \quad \text{on } B \quad B \in \mathcal{B}$$

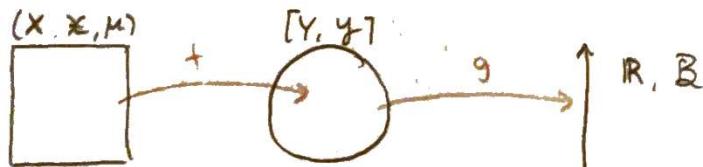
distribution of ξ

Def: The expectation $E\xi$ of rand. var. ξ is

$$E\xi = \int \xi(\omega) P(d\omega) \quad (*)$$

Leb. int.

Change of measure

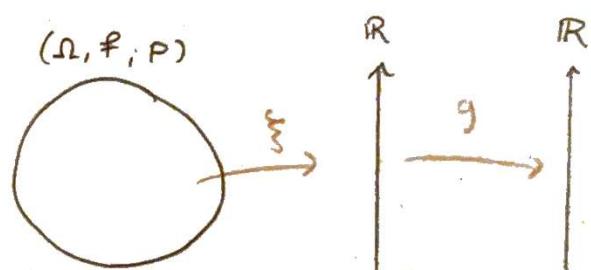


$$\int f(g(x)) \mu(dx) = \int g(y) (\mu f^{-1})(dy)$$

$$(g \circ f^{-1})(A) = \mu(f^{-1}(A)) \quad A \in \mathcal{Y}$$

$$\text{Take } [Y, y] = [\mathbb{R}, \mathcal{B}]$$

$$(X, X, \mu) = (\Omega, \mathcal{F}, P)$$



$$\begin{aligned} \int g(\xi(\omega)) P(d\omega) &= \int g(y) P \xi^{-1}(dy) = \\ &= \int g(y) P_\xi(dy) = \end{aligned}$$

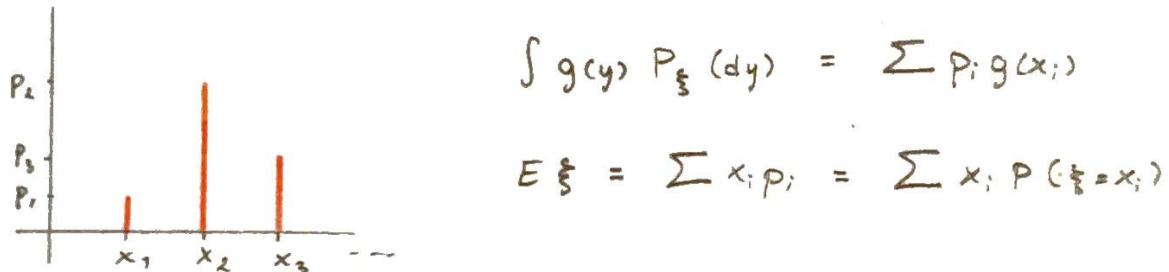
$$\text{Take } g(y) = y \quad E\xi = \int_R y P_\xi(dy)$$

1) ξ discrete

ξ	x_1	x_2	\dots
P_ξ	p_1	p_2	\dots

$$P_\xi = \sum p_i \delta_{x_i} + \text{atomic measure}$$

$$\delta_{x_i}(B) = \mathbb{1}_B(x_i) = \begin{cases} 1 & x_i \in B \\ 0 & x_i \notin B \end{cases}$$



Ex.

M	0	3	4	5
Fr.	2	4	3	1
P_M	$2/10$	$4/10$	$3/10$	$1/10$

$$\bar{M} = \frac{2 \cdot 0 + 4 \cdot 3 + 3 \cdot 4 + 1 \cdot 5}{10} = 0 \cdot \frac{2}{10} + 3 \cdot \frac{4}{10} + 4 \cdot \frac{3}{10} + 5 \cdot \frac{1}{10} = EM$$

Ex.

ξ	1	2	3	\dots
P_ξ	p_1	p_2	p_3	\dots

$$P_k = P\{\xi = k\} = c \frac{1}{k^2} \quad c = \left(\sum_{k=1}^{\infty} \frac{1}{k^2} \right)^{-1}$$

$$E\xi = \sum_{k=1}^{\infty} k c \frac{1}{k^2} = c \sum_{k=1}^{\infty} \frac{1}{k} = \infty$$

$$\xi = \xi^+ - \xi^-$$

P_{ξ} abs. cont.

$$P_{\xi}(B) = \int_B f_{\xi}(x) dx$$

Lebesgue integral

$$E\xi = \int x P_{\xi}(dx)$$

$$Eg(\xi) = \int g(x) P_{\xi}(dx) = \int x f_{\xi}(x) dx$$

$$\begin{aligned} \text{Take } g = \sum g_i \mathbb{1}_{B_i} &= \sum g_i P_{\xi}(B_i) = \sum g_i \int_{B_i} f_{\xi}(x) dx \\ &= \int \sum g_i \mathbb{1}_{B_i}(x) f_{\xi}(x) dx = \int g(x) f_{\xi}(x) \end{aligned}$$

Exercise : Show that $E\xi = \int x \underbrace{df_{\xi}}_{\substack{\text{non-decreasing fct of finite variation} \\ \text{shifts int } 1}}(dx)$

$$\text{Tot var } f = \sup_{\{x_1, \dots, x_n\} \subset \mathbb{R}} \sum_{i=1}^{n-1} |f(x_{i+1}) - f(x_i)|$$

$$\int x e^{-x} dx = - \int x de^{-x} = \int x \underbrace{d(1-e^{-x})}_{\text{cat}}$$

Properties

1) $E\xi \geq 0$ if $\xi \geq 0$ almost surely a.s : 'almost everywhere wrt P_{ξ} '

2) $E(c\xi) = c E\xi$

3) $E(\xi_1 + \xi_2) = E\xi_1 + E\xi_2$

4) Def: Two r.v. ξ_1, ξ_2 are called independent if

$$P(\xi_1 \in B_1, \xi_2 \in B_2) = P\{\xi_1 \in B_1\} P\{\xi_2 \in B_2\} \quad \forall B_1, B_2 \in \mathcal{B}$$

Def 2: Rvs ξ_1, ξ_2, \dots are called mutually indep if $\forall i_1, \dots, i_k$

$$P(\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_k}) = \prod_{j=1}^k P(\xi_{i_j} \in B_{i_j}) \quad \forall B_{i_1}, B_{i_2}, \dots \in \mathcal{B}$$

For fixed $A_1, A_2, \dots \in \mathcal{F}$ take $\xi_i(w) = \mathbb{1}_{A_i}(w)$

$$\xi_i = \begin{cases} 1 & w \in A_i \\ 0 & w \notin A_i \end{cases}$$

$$E\xi_i = 0 \cdot P\{\xi_i=0\} + 1 \cdot P\{\xi_i=1\} = P(A_i)$$

$$P(\cap A_{ij}) = P(\cap \{\xi_{ij}=1\}) = \prod_j P(\xi_{ij}=1) = \prod_j P(A_{ij})$$

4) If ξ_1, ξ_2 are indep ($\xi_1 \perp\!\!\!\perp \xi_2$) then

$$E(\xi_1 \xi_2) = E\xi_1 E\xi_2$$

proof: $\xi_1 = \mathbb{1}_{A_1}, \quad \xi_2 = \mathbb{1}_{A_2}$

$$\begin{aligned} E(\xi_1 \xi_2) &= 0 \cdot P(\dots) + 1 \cdot P\{\xi_1=1, \xi_2=1\} = P\{\xi_1=1\} P\{\xi_2=1\} \\ &= E\xi_1 E\xi_2 \end{aligned}$$

By Monotone Class Argument this extends to lin comb of indicators
ie simple fct, then to monotone limits etc.

$\xi - E\xi$: deviation from expectation

$$E[\xi - E\xi] = E\xi + (-1) E[E\xi] = E\xi - E\xi = 0$$

Variance of ξ : $E[(\xi - E\xi)^2] = \text{Var } \xi$

Standard deviation: $\sigma_\xi = \sqrt{\text{Var } \xi}$

Properties

1) $\text{Var}(c\xi) = c^2 \text{Var } \xi$

$$E(c\xi - E(c\xi))^2 = E(c(\xi - E\xi))^2 = c^2 E(\xi - E\xi)^2$$

$$2) \text{Var } c = E(c - Ec)^2 = 0$$

$$3) \xi_1 \perp\!\!\!\perp \xi_2 \Rightarrow \text{Var}(\xi_1 + \xi_2) = \text{Var } \xi_1 + \text{Var } \xi_2$$

$$\begin{aligned}\text{Var } \xi &= E(\xi - E\xi)^2 = E(\xi^2 - 2E\xi \cdot \xi + (E\xi)^2) \\ &= E\xi^2 - 2E\xi \cdot E\xi + (E\xi)^2 \\ &= E\xi^2 - (E\xi)^2\end{aligned}$$

$$\begin{aligned}\text{Var}(\xi_1 + \xi_2) &= E(\xi_1 + \xi_2)^2 - (E\xi_1 + E\xi_2)^2 \\ &= E\xi_1^2 + 2E\xi_1 \cdot E\xi_2 + E\xi_2^2 - (E\xi_1)^2 - 2E\xi_1 \cdot E\xi_2 - (E\xi_2)^2 \\ &= E\xi_1^2 - (E\xi_1)^2 + E\xi_2^2 - (E\xi_2)^2 \\ &= \text{Var } \xi_1 + \text{Var } \xi_2\end{aligned}$$

In general not true for dependent variables

Take $\xi_2 = -\xi_1$, with $\text{Var } \xi_1 > 0$

$$\Rightarrow \text{Var}(\xi_1 + \xi_2) = \text{Var } 0 = 0 \neq \text{Var } \xi_1 + \text{Var } \xi_2 = 2 \text{Var } \xi,$$

$$4) \text{Var } \xi = 0 \Leftrightarrow \xi = c \text{ a.s. i.e. } P\{\xi \neq c\} = 0$$

Exercise: Show this, Leb. int.

$$\text{Var } \xi = E\xi^2 - (E\xi)^2, \quad E\xi = \int x \cdot P_\xi(dx) \quad E\xi^2 = \int x^2 \cdot P_\xi(dx)$$

FURTHER CHARACTERISTICS

$E\xi^\alpha$: moments of ξ

$$\xi \geq 0 \text{ a.s.} \Rightarrow E\xi^\alpha = \int x^\alpha P_\xi(dx)$$

$\alpha \in \mathbb{N}$; ok $E\xi^\alpha$ well def

$\alpha = 1$ corresponds to Expectation
 \equiv 1st Moment

$E|\xi|^\alpha$: absolute moments

$E(\xi - E\xi)^\alpha$: central moments

$E|\xi - E\xi|^\alpha$: central abs. moments

$\text{Var } \xi$ is 2nd central moment

$$E(\xi - E\xi)^\alpha = \int (x - E\xi)^2 P_\xi dx$$

$\xi - E\xi$: centered r.v.

$E e^{-\alpha \xi}$ exponential moments Laplace transform

$$\text{If } \xi \text{ has density } f_\xi \quad E e^{-\alpha \xi} = \int e^{-\alpha x} f_\xi(x) dx \\ = L[f_\xi](\alpha)$$

If $\xi \in \mathbb{Z}^+$ $P\{\xi = k\} = p_k$

$$\Rightarrow E e^{-\alpha \xi} = \sum_{k=0}^{\infty} e^{-\alpha k} p_k = \sum_{k=0}^{\infty} (e^{-\alpha})^k p_k = E Z^\xi$$

$Z = e^{-\alpha}$: probability generating func of ξ

Föreläsning 8

MOST COMMON DISTRIBUTIONS

Discrete

Bernoulli distr. $\xi \sim \text{Bern}(p)$ $p \in [0, 1]$

$$P\{\xi=1\} = p \quad P\{\xi=0\} = 1-p$$

$$E\xi = p \cdot 1 + (1-p) \cdot 0 = p \quad E\xi^2 = E\xi = p$$

$$\text{Var } \xi = E\xi^2 - (E\xi)^2 = p - p^2 = p(1-p)$$

Ex 1: Toss of coin

$$\text{Head} = 1 \quad \text{Tail} = 0 \quad P(\text{Head}) = p$$

ξ	0	1
P_ξ	$1-p$	p

Ex 2: $E \in \mathcal{F}$ an event

$$\mathbb{1}_E(w) = \begin{cases} 1 & w \in A \\ 0 & w \notin A \end{cases} \quad \text{indicator of event } A$$

$$\mathbb{1}_E \sim \text{Bern}(P(A))$$

Binomial distr. $\xi \sim \text{Bin}(n, p)$ $n \in \mathbb{N}$ $p \in [0, 1]$

$$P\{\xi=k\} = \binom{n}{k} p^k (1-p)^{n-k}, \quad k=0, 1, \dots, n$$

ξ corresponds to no Heads in n tosses of a coin s.t. $P(\text{Head}) = p$

$$\frac{H}{1} \frac{H}{2} \frac{T}{1} \frac{H}{4} \dots \frac{H}{n!} \frac{T}{n} \rightarrow P P (1-p) P \dots p (1-p) = p^{*n} (1-p)^{*T}$$

$$P\{\text{no heads} = k\} = P\{w : \# H(w) = k\} = \binom{n}{k} p^k (1-p)^{n-k}$$

n indep. realisations of a Bernoulli r.v are called Bernoulli trials.
often we talk about 'Successes' and 'Failures'

$\xi \sim \text{Bin}(n, p)$: no successes in n Bernoulli trials
w. $p(\text{success}) = p$

$$E\xi = \sum_{k=0}^{\infty} k P\{\xi = k\} = \sum_{k=0}^{\infty} k \binom{n}{k} p^k (1-p)^{n-k}$$

$$\frac{H}{1} \quad \frac{H}{1} \quad \frac{T}{0} \quad \dots \quad \frac{H}{1} \quad \frac{T}{0}$$

$X_i = \mathbb{1}\{\text{i-th trial is success}\}$ $X_i \sim \text{Bern}(p)$ ← iid rv

$$\sum_{i=1}^n X_i = 5 \quad \text{no successes}$$

$$\Rightarrow E\xi = E \sum_{i=1}^n X_i = \sum_{i=1}^n EX_i = np$$

$$\text{Var } \xi = \text{Var} \sum_{i=1}^n X_i \stackrel{\text{indep}}{=} \sum_{i=1}^n \text{Var } X_i = np(1-p)$$

Discrete uniform over a set $S = \{s_1, \dots, s_n\}$ $n < \infty$

$$\xi \sim \text{Unif}(S)$$

$$P(\xi = s_k) = \frac{1}{n} \quad E\xi^k = \frac{1}{n} \sum_{k=1}^n s_k^k$$

Geometrical distr $\xi \sim \text{Geom}(p) \quad p \in (0, 1]$

No Bernoulli trials with prob. success p , to get first success.

$$\Omega = \{\omega : \underbrace{FF\dots FS}_{K-1} \} \quad \xi(\omega) = k$$

$$P\{\xi = k\} = (1-p)^{k-1} p \quad k = 1, 2, \dots$$

$$E\xi = \sum_{k=1}^{\infty} k (1-p)^{k-1} p$$

$$F(q) = \sum_{k=1}^{\infty} q^k = \frac{q}{1-q} \quad F'(q) = \sum_{k=1}^{\infty} k q^{k-1} = \frac{1}{(1-q)^2}$$

$$\Rightarrow E\xi = p F'(q) = p \frac{1}{p^2} = \frac{1}{p}$$

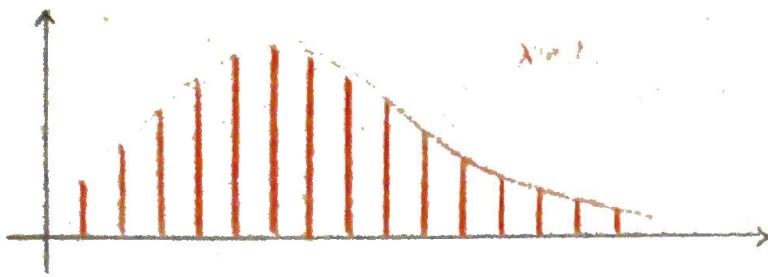
2nd def: $\xi_0 \sim \text{Geom}(p)$ No failures until first success

$$P\{\xi_0 = k\} = (1-p)^k p = E(\xi - 1) = \frac{1}{p} - 1$$

Poisson distribution

$$\xi \sim Po(\lambda) \quad \lambda > 0$$

$$P\{\xi = k\} = \frac{\lambda^k}{k!} e^{-\lambda} \quad k = 0, 1, 2, \dots$$



$$E\xi = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = e^{-\lambda} \lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = \lambda$$

$$\text{Var } \xi = \dots = \lambda$$

Poisson Limit Theorem

Consider $\xi_n \sim \text{Bin}(n, p_n)$ $n p_n \rightarrow \lambda \in (0, \infty)$

Then $P\{\xi_n = k\} \xrightarrow{n \rightarrow \infty} P(\pi = k)$

where $\pi \sim Po(\lambda)$

Proof $p_n = \frac{\lambda^n}{n!} + O\left(\frac{1}{n}\right)$

$$\begin{aligned} P\{\xi_n = k\} &= \frac{n!}{(n-k)! k!} \left(\frac{\lambda}{n} + O\left(\frac{1}{n}\right)\right)^k \left(1 - \frac{\lambda}{n} + O\left(\frac{1}{n}\right)\right)^{n-k} \cdot \frac{\lambda^k}{n^k} \\ &\stackrel{(*)}{=} \frac{1}{k!} (\lambda + O(n))^k \frac{n!}{(n-k)! n^k} \left(1 - \frac{\lambda}{n} + O\left(\frac{1}{n}\right)\right)^{n-k} \\ &= \frac{(\lambda + O(1))^k}{k!} \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-k+1}{n} \left(1 - \frac{\lambda}{n} + O\left(\frac{1}{n}\right)\right)^{n-k} \\ &= \frac{(\lambda + O(1))^k}{k!} \underbrace{1 \cdot \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right)}_{\rightarrow 1} \frac{\left(1 - \frac{\lambda}{n} + O\left(\frac{1}{n}\right)\right)^n}{\left(1 - \frac{\lambda}{n} + O\left(\frac{1}{n}\right)\right)^k} e^{-\lambda} \\ &\rightarrow \frac{\lambda^k}{k!} \end{aligned}$$

$$\left(1 - \frac{\lambda}{n} + O\left(\frac{1}{n}\right)\right)^n = \exp\{-n \ln(1 - \frac{\lambda}{n} + O\left(\frac{1}{n}\right))\} \rightarrow \exp\{-\lambda\}$$

$$\text{so } (*) \rightarrow \frac{\lambda^k}{k!} e^{-\lambda} = P\{\pi = k\} \quad \text{for } \pi \sim Po(\lambda)$$

Continuous (absolutely)

Uniform $\xi \sim \text{Unit } [a, b]$

$f_\xi(x) = \text{const on } [a, b], 0 \text{ otherwise}$

$$E\xi = \int_a^b x f_\xi(x) dx = \frac{1}{b-a} \int_a^b x dx = \frac{1}{b-a} \left[\frac{x^2}{2} \right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{1}{2}(a+b)$$

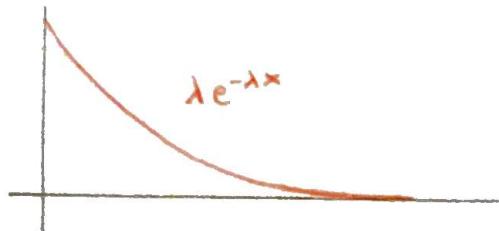
$$\text{Var } \xi = \frac{1}{12} (b-a)^2$$

Exponential $\xi \sim \text{Exp } (\lambda) \quad \lambda > 0$

$$F_\xi(x) = P\{\xi \leq x\} = \begin{cases} 1 - e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

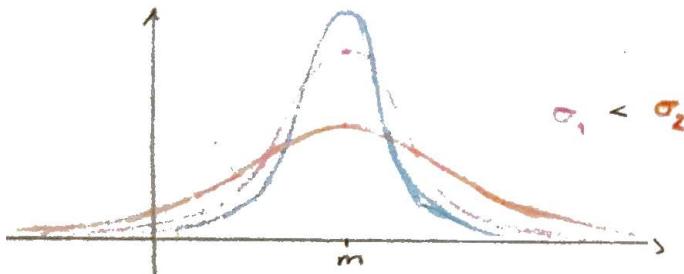
$$f_\xi(x) = \lambda e^{-\lambda x} \mathbb{1}_{x \geq 0}$$

$$E\xi = \frac{1}{\lambda} \quad \text{Var } \xi = \frac{1}{\lambda^2}$$



Normal/Gaussian $\xi \sim N(m, \sigma^2) \quad m \in \mathbb{R} \quad \sigma \in \mathbb{R}_+$

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \frac{(x-m)^2}{\sigma^2}}$$



Standard Normal: $Z \sim N(0, 1) \quad f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad x \in \mathbb{R}$

$$\text{Def: } F_Z(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt = \Phi(x) \quad \text{Laplace Int}$$

Standardisation: $\xi \sim N(m, \sigma^2) \Rightarrow \frac{\xi - m}{\sigma} \sim N(0, 1)$

Föreläsning 9

Normal / Gaussian distr

$$\xi \sim N(m, \sigma^2) \quad f_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad x \in \mathbb{R}$$

$$z \sim N(0, 1) \quad f_z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$\xi \sim f_\xi(x) \quad g = a\xi + b \quad a > 0$$

$$F_g(x) = P\{\xi \leq x\} = P\{a\xi + b \leq x\} = F_\xi\left(\frac{x-b}{a}\right)$$

$$f_z(x) = \frac{1}{a} f_\xi\left(\frac{x-b}{a}\right)$$

ξ has the same distribution as $\sigma z + m$

$$\xi = \sigma z + m$$

$$E\xi = E(\sigma z + m) = m + \sigma E z$$

$$\text{Var } \xi = \text{Var}(\sigma z + m) = \sigma^2 \text{Var } z$$

$$EZ = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 0 \quad \text{odd fct}$$

$$\text{Var } z = EZ^2 - (EZ)^2 = \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \right)^2 = \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \int_{-\infty}^{\infty} y^2 \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \\ &= \iint_{\mathbb{R}^2} x^2 y^2 e^{-\frac{x^2+y^2}{2}} dx dy \quad x = p \cos \varphi \quad y = p \sin \varphi \quad p \in \mathbb{R}^+ \\ &\quad xy = p^2 \cos^2 \varphi \sin^2 \varphi = \frac{1}{4} p^4 \sin 2\varphi \quad \varphi \in [0, 2\pi] \end{aligned}$$

$$= \frac{1}{4} \int_0^{2\pi} \int_0^\infty p^5 e^{-p^2/2} dp d\varphi$$

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2/2} dx \right)^2 = \dots = \int_0^\infty \int_0^{2\pi} p e^{-p^2/2} d\varphi dp = 2\pi \int_0^\infty p e^{-p^2/2} dp/2 \\ &= 2\pi (-e^{-p^2/2}) \Big|_0^\infty = 2\pi \end{aligned}$$

Thus $f_z(x)$ is a density

Then $f_g(x) = f_{\sigma z + m}(x)$ is also a density

$$\int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx = [-x \cdot e^{-x^2/2}]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-x^2/2} dx$$

$$\Rightarrow \text{Var } z = 1$$

$$\xi \sim N(m, \sigma^2)$$

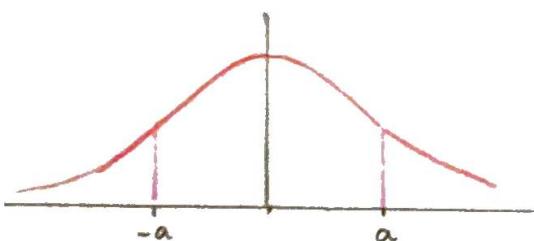
$$E\xi \quad \text{Var } \xi$$

Stability of the normal distr:

$$\xi_1, \xi_2 \text{ indep Normal} \Rightarrow \xi_1 + \xi_2 \text{ Normal}$$

$$\text{we will show that } \xi_1 + \xi_2 \sim N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$$

THE RULE OF SIGMAS



$$Z \sim f_Z(x)$$

$$\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$$F_2(a) - F_2(-a) = 1 - 2F_2(-a) = 1 - 2\Phi(-a) = P\{|z| \leq a\}$$

$$P\{|z| \geq 2\} \leq 0.05$$

$$P\{|z| \geq 3\} \leq 0.01$$

$$P\left\{|\xi - m| \geq \frac{2\sigma}{3}\right\} \leq \frac{0.05}{0.01}$$

We will later prove Central Limit Theorem

ξ_1, ξ_2, \dots are indep identically distr. RV
with $E \xi_i = m$ $\text{Var } \xi_i = \sigma^2 < \infty$
and $S_n = \sum_{i=1}^n \xi_i$

$$\Rightarrow P \left\{ \frac{S_n - nm}{\sqrt{n}\sigma} \leq x \right\} \Rightarrow \phi(x) = P(Z \leq x)$$
$$Z = \frac{S_n - ES_n}{\sqrt{\text{Var } S_n}}$$

RANDOM VECTORS

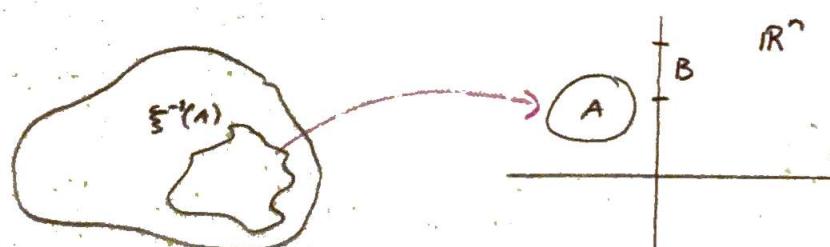
RV $(\Omega, \mathcal{F}, P) \xrightarrow{F} [\mathbb{R}, \mathcal{B}]$

Def. A random vector (n-dimensional) is a $[\mathcal{F}, \mathcal{B}(\mathbb{R}^n)]$ - measure mapping $\Omega \rightarrow \mathbb{R}^n$

$\mathcal{B}(\mathbb{R}^n)$: Borel σ -field of subsets of \mathbb{R}^n
which is generated by open sets in \mathbb{R}^n
but also by rectangles $\prod_{i=1}^n [x_i; a_i, b_i]$
and also by $\prod_{i=1}^n (-\infty, x_i] \quad (*)$

The distribution P_F of $\bar{\xi}$ is the push-forward of P

$$P_{\bar{\xi}}(A) = P(\xi^{-1}(A)) = P\{\omega: \bar{\xi}(\omega) \in A\}$$
$$A \in \mathcal{B}(\mathbb{R}^n) \quad \mathcal{F}$$



Ω, \mathcal{F}, P

$\xi(\omega)$ in \mathbb{R}^n can be written as $(\xi_1(\omega), \dots, \xi_n(\omega))$

Then $\xi_i : \Omega \rightarrow \mathbb{R}$ are called the coordinates

Show that ξ_i are rv's

Because of (*) and the extension theorem, P_{ξ} is def by its values on $\prod_{i=1}^n (-\infty, x_i]$ ie

$$P_{\xi}(\prod_{i=1}^n (-\infty, x_i]) = P\{\xi \in \prod_{i=1}^n (-\infty, x_i]\} = P\{\xi_1 \leq x_1, \dots, \xi_n \leq x_n\}$$
$$= F_{\xi}(x_1, \dots, x_n)$$

is called joint (or multivariate) cdf

To define the distr. of a random vector it is sufficient to know $F_{\xi}(x_1, \dots, x_n)$

Def: P_{ξ} defines the marginal distr.

ie the distr. of individual coordinates

$$P_{\xi_1}, P_{\xi_2}, \dots, P_{(\xi_1, \xi_2)}, \dots, P_{(\xi_1, \dots, \xi_k)}$$

$$F_{\xi_1}(x_1) = P\{\xi_1 \leq x_1\} = P\{\xi_1 \leq x_1, \xi_2 < \infty, \dots, \xi_n < \infty\}$$
$$= F_{\xi}(x_1, \infty, \dots, \infty)$$

$$F_{(\xi_1, \dots, \xi_k)}(x_1, \dots, x_k) = F_{\xi}(\infty, x_1, \dots, x_k, \infty)$$

DISCRETE MULTIVAR. DISTR.

$$P_{\xi} = \sum_i p_i \delta_{x_i}, \quad \sum p_i = 1$$

(Absolutely) cont distr. $\exists f_{\xi}(x_1, \dots, x_n)$ s.t

$$P_{\xi}(B) = \int_B f_{\xi}(x_1, \dots, x_n) dx_1 \dots dx_n \quad \forall B \in \mathcal{B}(\mathbb{R}^n)$$

Ex. $B = \bigcap_{i=1}^n (-\infty, x_i]$

$$F_B(x_1, \dots, x_n) = P_B(B) = \int_{-\infty}^{x_1} dx_1 \dots \int_{-\infty}^{x_n} dx_n f_B(x_1, \dots, x_n)$$

$$\text{so } f_B(x_1, \dots, x_n) = \frac{\partial^n}{\partial x_1 \dots \partial x_n} F_B(x_1, \dots, x_n)$$

INDEP. COORDINATES

Assume that ξ_1, \dots, ξ_n are mutually indep.

$$\xi = (\xi_1, \dots, \xi_n)$$

$$\begin{aligned} F_\xi(x_1, \dots, x_n) &= P\{\xi_1 \leq x_1, \dots, \xi_n \leq x_n\} \\ &= F_{\xi_1}(x_1) \dots F_{\xi_n}(x_n) \end{aligned}$$

In particular if all ξ_i 's are rv. then

$$\begin{aligned} f_\xi(x_1, \dots, x_n) &\stackrel{(con)}{=} \frac{\partial^n}{\partial x_1 \dots \partial x_n} F_{\xi_1}(x_1) \dots F_{\xi_n}(x_n) \\ &= f_{\xi_1}(x_1) \dots f_{\xi_n}(x_n) \end{aligned}$$

0-1 vectors

$$\xi = (\xi_1, \dots, \xi_n) \text{ where } \xi_i \in \{0, 1\}$$

$$P_{\xi_i} \sim \text{Bern}(p_i)$$

But ξ_i need not be indep.

P_ξ is a measure on $\{0, 1\}^n \subset \mathbb{R}^n$

If ξ_1, \dots, ξ_n are indep then $P_\xi = P_{\xi_1} \times \dots \times P_{\xi_n}$

If $p_i \in \mathbb{R}^+$ (any) then P_ξ is $\text{Unif}\{0, 1\}^n$

f_ξ is uniform on $S \subset \mathbb{R}^n$ $|S| > 0$

$f_\xi(x) \equiv \text{const}$ on S

Since $\int f_\xi(x) d\bar{x} = 1 \Rightarrow \text{Const} = \frac{1}{|S|} = n\text{-volume of } S$

$\xi_1, \dots, \xi_n \sim N(0, 1)$

indep Standard Multivariate Normal

$\xi \sim MVN(\vec{0}, \mathbb{I}) \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$$f_\xi(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}(x_1^2 + \dots + x_n^2)}$$

Foreläsning 10

MULTIVARIABLE NORMAL DISTR

$$\bar{\xi} \sim MVN(\bar{m}, V)$$

$\bar{\xi} \in \mathbb{R}^n$, $\bar{m} \in \mathbb{R}^n$, V : $n \times n$ matrix pos. def. ; $\det V > 0$

$$\forall \bar{x} \in \mathbb{R}^n, \bar{x}^T V \bar{x} = \sum_{ij} v_{ij} x_i x_j > 0$$

$$f_{\bar{\xi}}(\bar{x}) = \frac{1}{(2\pi \det V)^{n/2}} \exp \left\{ -\frac{1}{2} (\bar{x} - \bar{m})^T V^{-1} (\bar{x} - \bar{m}) \right\}$$

V pos def $\Rightarrow \exists$ orthogonal matrix O $O^T = O^{-1}$, $O^T = O^{-1}$ $\det O = 1$

$$\text{s.t. } OVO^T = \text{diag}(\sigma_1^2 \sigma_2^2 \dots \sigma_n^2)$$

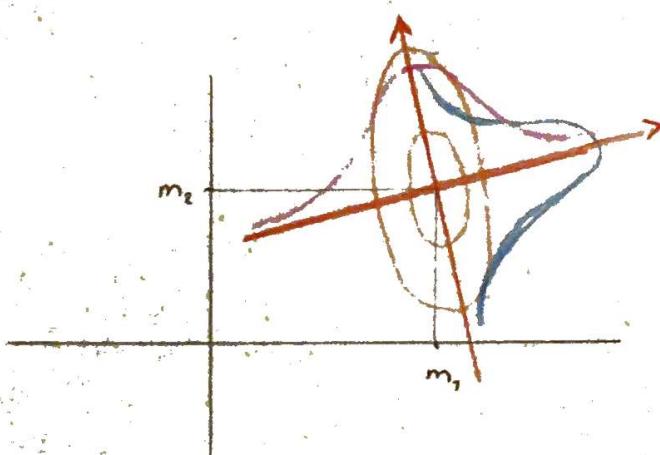
$$(OVO^T)^{-1} = \text{diag}\left(\frac{1}{\sigma_1^2} \dots \frac{1}{\sigma_n^2}\right) = (O^T)^{-1} V^{-1} O^{-1} = OV^{-1}O^T$$

Change of variable:

$$\bar{y} = (\bar{x} - \bar{m}) O^T \quad \bar{x} - \bar{m} = \bar{y} O$$

$$K = y^T O V^{-1} O^T y = \sum_{i=1}^n \frac{y_i^2}{\sigma_i^2}$$

$$\begin{aligned} f_{\bar{\xi}}(\bar{x}) &= f_{\bar{\xi}}(\bar{m} + \bar{y} O) = \frac{1}{(2\pi \prod_{i=1}^n \sigma_i^2)} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \frac{y_i^2}{\sigma_i^2} \right\} \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi} \sigma_i} e^{-\frac{1}{2} \frac{y_i^2}{\sigma_i^2}} \\ &\sim N(0, \sigma_i^2) \end{aligned}$$



New axes

Normal

Normal

σ_1, σ_2 ellipses

The level sets of f_{ξ} are the (different) level sets of k .

$$k(\bar{y}) = \text{const} \sum_{i=1}^n \frac{y_i^2}{\sigma_i^2}$$

n -dim ellipses with semiaxes σ_i .

$$E\xi = (E\xi_1, \dots, E\xi_n) = \bar{m}$$

If V not pos def, one of σ_i could be 0

\Rightarrow ellipse with semiaxes 0 = line

\Rightarrow go down one dimension --

COVARIANCE

Let ξ, η be two r.v. with joint distr. $P(\xi, \eta)$

Def. Covariance between ξ and η :

$$\text{Cov}(\xi, \eta) = E[(\xi - E\xi)(\eta - E\eta)]$$

$$\text{Note } \text{Cov}(\xi, \xi) = \text{Var}(\xi)$$

Def. Correlation coef.

$$\text{Cor}(\xi, \eta) = \frac{\text{Cov}(\xi, \eta)}{\sqrt{\text{Var}\xi \cdot \text{Var}\eta}} \quad -1 \leq \text{Cor} \leq 1$$

$$\text{Cor}(\xi, \eta) = \pm 1 \quad \text{iff} \quad \exists a, b \text{ s.t. } \eta = a\xi + b$$

$$a > 0 : \text{Cor} = 1 \quad a < 0 : \text{Cor} = -1$$

Properties cov:

$$\begin{aligned} \text{Cov}(\xi, \eta) &= E(\xi\eta) - E\xi E\eta - \cancel{E\xi E\eta} + \cancel{E\xi E\eta} \\ &= E(\xi\eta) - E\xi E\eta \end{aligned}$$

- $\text{Cov}(\xi, \eta) = \text{Cov}(\eta, \xi)$
- $\text{Cov}(t\xi, \eta) = t \cdot \text{Cov}(\xi, \eta)$
- $\text{Cov}(\xi + \zeta, \eta) = \text{Cov}(\xi, \eta) + \text{Cov}(\zeta, \eta)$

→ Cov is a bilinear operation

$$\text{Cov}(\xi + c, \eta) = \text{Cov}(\xi, \eta) + \cancel{\text{Cov}(c, \eta)} = \text{Cov}(\xi, \eta)$$

Let's now consider a linear space L_2^0 of rv with $E\xi=0$ and $\text{Var } \xi = E\xi^2 < \infty$

$$\text{Cov}(\xi, \eta) = E\xi\eta = \iint xy P_{\xi\eta}(dx, dy)$$

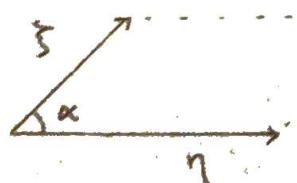
$$\text{scalar product: } (\xi, \eta) \quad \| \xi \|^2 = (\xi, \xi) = 0 \Rightarrow \xi = 0$$

$$E\xi^2 = 0 \Rightarrow \xi = 0 \text{ a.s}$$

Covariance has the meaning of scalar product

Variance has meaning square of the norm

$$\text{Cor}(\xi, \eta) = \frac{(\xi, \eta)}{\|\xi\| \|\eta\|} \quad (\xi, \eta) = \|\xi\| \|\eta\| \cos \alpha$$



$$(\xi, \eta)^2 \leq \|\xi\|^2 \|\eta\|^2 \quad \text{Cauchy-Schwarz ineq.}$$

$$\text{corresponds to: } (E\xi\eta)^2 \leq E\xi^2 E\eta^2 \quad \xi, \eta \in L_2^0$$

$$\text{proof: } 0 = E(a\xi + b\eta)^2 = a^2 E\xi^2 + 2ab E\xi\eta + b^2 E\eta^2$$

$$0 = E[\alpha\xi + \beta\eta]^2 = \underline{\alpha^2} E\xi^2 + 2\alpha\beta E\xi\eta + \underline{\beta^2} E\eta^2$$

$$\text{discriminant} \quad \Delta^2 = 4\alpha^2\beta^2 (E(\xi\eta))^2 - 4\alpha^2 (E\xi^2) \beta^2 (E\eta^2)$$

CONDITIONAL DISTR OF RV

$$\xi \text{ - disc. rv} \quad \xi(w) = \sum x_i \mathbb{1}_{A_i}(w_i)$$

$$P_\xi = \sum_i p_i \delta_{x_i} \quad P\{\xi = x_i\} = p_i \quad \sum p_i = 1$$

Consider event A s.t. $P(A) > 0$

$$P\{\xi = x_i | A\} = \sum_i P\{\xi = x_i | A\} = \sum_i \frac{P\{\{w : \xi = x_i\} \cap A\}}{P(A)} = \frac{1}{P(A)} =$$

Thus $P\{\xi_i = x_i | A\}$ $i=1, \dots$
is a prob distr $P_{\xi|A}$

ξ	x_1	x_2
$P_{\xi A}$	$P\{\xi = x_1 A\}$	$P\{\xi = x_2 A\}$

$$E[\xi | A] = \sum f(x_i) P\{\xi_i | A\}$$

Let η be another disc. r.v

$$P\{\eta = y_i\} = q_i > 0$$

one gets conditional distr of ξ given $\{\eta = y_i\}$

In particular:

$$E[\xi | \eta = y_i] = \sum_i x_i P\{\xi = x_i | \eta = y_i\}$$

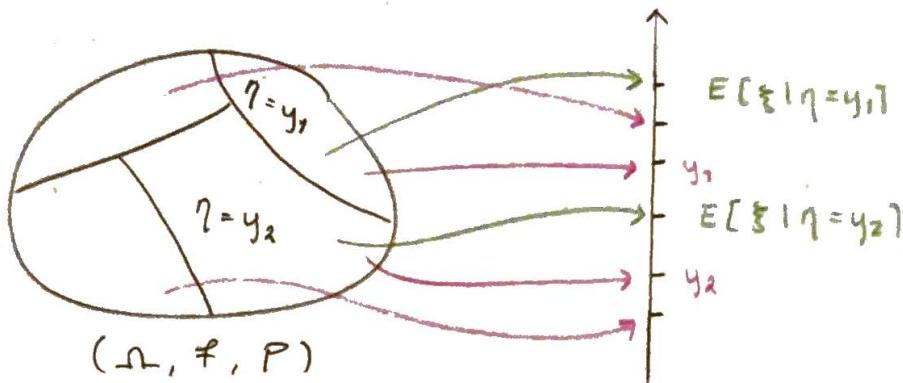
This is a fct of y_i 's

More exactly: it is a fct on Ω taking values

$E[\xi | \eta = y_i]$ on the sets $\{w : \eta(w) = y_i\} \in \mathcal{F}$

So it is a random variable:

$$E[\xi | \eta] = \sum_i E[\xi | \eta = y_i] \mathbf{1}_{\eta^{-1}\{y_i\}}(w)$$



$E[\xi | \eta]$ is measurable wrt $\sigma(\eta) = \sigma(\eta^{-1}(B), B \in \mathbb{R})$

Ex. $\Omega = \{0, 1\}^2 : 0,0 \quad 0,1 \quad 1,0 \quad 1,1$

counting table

ξ	0	1		$\xi, \eta \in \{0, 1\}$
0	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{2}$	
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	
	$\frac{3}{8}$	$\frac{5}{8}$		

$$E[\xi | \eta = 0] = 1 \cdot E[\xi = 1 | \eta = 0] + 0 \dots = \frac{3/8}{1/2} = \frac{6}{8}$$

$$E[\xi | \eta = 1] = E[\xi = 1 | \eta = 1] = \frac{1/4}{1/2} = \frac{1}{2}$$

$E[\xi \eta]$	$\frac{1}{4}$	$\frac{1}{2}$
P_η	$P\{\eta=0\}$	$P\{\eta=1\}$

Föreläsning 11

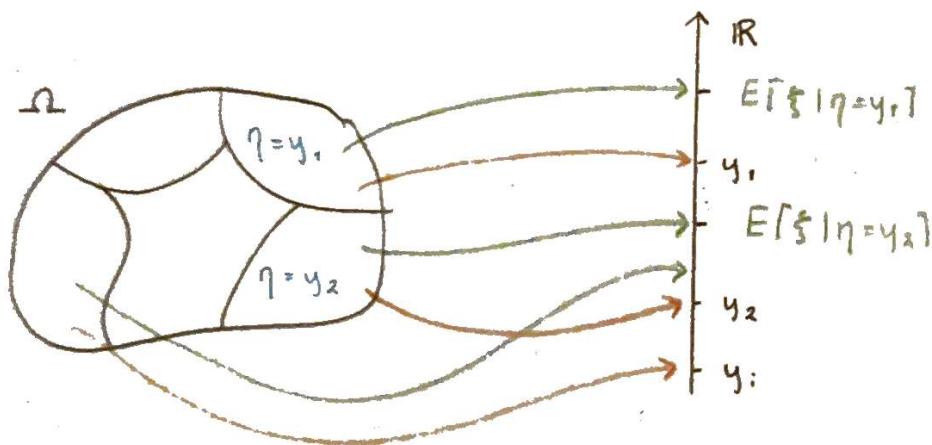
$(\xi, \eta) : \Omega \rightarrow \mathbb{R}^2$ discrete rv

$$P\{\xi = x_i, \eta = y_j\} = p_{ij} > 0$$

$$P_{\xi|\eta=y_j}(x_i) = P\{\xi = x_i | \eta = y_j\}$$

$$E[\xi | \eta = y_j] = \sum x_i P\{\xi = x_i | \eta = y_j\}$$

$$\psi(\eta(\omega)) = \sum E[\xi | \eta = y_i] \mathbf{1}_{\eta^{-1}(y_i)}(\omega)$$



$$\psi(\eta) = E[\xi | \eta] \quad \text{a r.v. measurable wrt } \sigma(\{\eta^{-1}(B), B \in \mathcal{B}\})$$

$$\sigma(\eta) =$$

$$E\psi(\eta) = E[E[\xi | \eta]] = \sum_j E[\xi | \eta = y_j] P(\eta^{-1}(y_j))$$

$$= \sum_i \sum_j x_i P\{\xi = x_i | \eta = y_j\} P\{\eta = y_j\}$$

$$= \sum_i x_i \sum_j \frac{P\{\xi = x_i, \eta = y_j\}}{P\{\eta = y_j\}} P\{\eta = y_j\} = \sum_i x_i P\{\xi = x_i\}$$

$$= E\xi$$

FULL EXPECTATION FORMULA

$$EE[\xi | \eta] = E\xi$$

$$\xi = \mathbb{1}_A \{H_j\} \quad \text{path of } \Omega$$

$$\eta = \sum_j j \underbrace{\mathbb{1}_{H_j}}_{\text{distinct values}}$$

$$E[\xi | \eta = y_i] = P\{A | \eta = y_i\} = P\{A | H_j\}$$

$$EE[\xi | \eta] = \sum_j P\{A | H_j\} P(H_j) = P(A) = E \mathbb{1}_A \quad \text{last prob from}$$

(ξ, η) - abs cont

$$P\{(\xi, \eta) \in B\} = \int_B \underbrace{f_{(\xi, \eta)}(x, y)}_{Pdt} dx dy \quad \forall B$$

We want to define 'conditional density' $f_{\xi|\eta}(x|y)$

$$P\{\xi \in (x - \frac{\Delta x}{2}, x + \frac{\Delta x}{2}) \mid \eta \in (y - \frac{\Delta y}{2}, y + \frac{\Delta y}{2})\}$$

$$(P\{\xi \in (x - \frac{\Delta x}{2}, x + \frac{\Delta x}{2})\} + o(\Delta_x)) = f_\xi(x) \Delta_x$$

$$= \frac{P\{\xi \in u_x, \eta \in u_y\}}{P\{\eta \in u_y\}} + o(\Delta_x) = \frac{\frac{f_{(\xi, \eta)}(x, y) \Delta_x \Delta_y}{f_\eta(y) \Delta y}}{f_\eta(y) \Delta y} + o(\Delta_x)$$

The conditional density of ξ given $\eta = y$ is

$$f_{\xi|\eta}(x|y) = \frac{f_{(\xi, \eta)}(x, y)}{f_\eta(y)}$$

$$f_{\eta}(y) = \int f_{(\xi, \eta)}(x, y) dx$$

$$\int f_{\xi|\eta}(x|y) dx = \frac{1}{f_{\eta}(y)} \int f_{(\xi, \eta)}(x, y) dy = \frac{1}{f_{\eta}(y)} f_{\eta}(y) = 1$$

Thus $f_{\xi|\eta}(x|y)$ is a density (as fct of x , not y) at a r.v. $\xi | \eta = y$

$$E[\xi | \eta = y] = \int x f_{\xi|\eta}(x|y) dx =$$

This is a fct of y (the values of η)

So it is a fct $\Psi(\eta) - r.v.$

$$\begin{aligned} E\Psi(\eta) &= \int E[\xi | \eta = y] f_{\eta}(y) dy dx = \iint x f_{\xi|\eta}(x|y) f_{\eta}(y) dy dx \\ &= \iint x \frac{f_{(\xi, \eta)}(x, y)}{f_{\eta}(y)} f_{\eta}(y) dy dx = \int x \underbrace{\int f_{(\xi, \eta)}(x, y) dy}_{f_{\xi}(x)} dx \\ &= \int x f_{\xi}(x) dx = E\xi \quad \text{full exp f'la} \end{aligned}$$

when $\xi \perp \eta$:

$$P\{\xi = x; \eta = y\} = \frac{P\{\xi = x; \eta = y\}}{P\{\eta = y\}} = P\{\xi = x\}$$

+

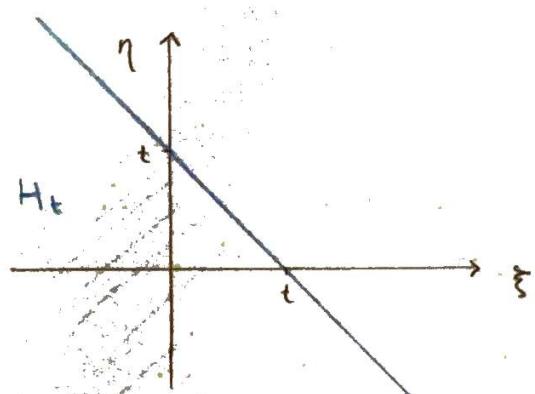
DISTRIBUTION OF A SUM

$$(\xi, \eta) \sim P_{(\xi, \eta)}$$

What is $F_{\xi+\eta}(t)$?

$$H_t = \{(x, y) : x + y \leq t\}$$

$$F_{\xi+\eta}(t) = P\{\xi + \eta \leq t\} = P_{\xi, \eta}(H_t) = (*)$$



For discrete (ξ, η) :

$$(*) = \sum_{(x_i, y_j) \in H_t} P\{\xi = x_i, \eta = y_j\}$$

Ex. $\xi \sim Po(\lambda)$ $\eta \sim Po(\nu)$, $\xi \perp \eta$

pmt: $f_{\xi+\eta}(k) = P\{\xi + \eta = k\} = \sum_{i=0}^k P\{\xi = i\} P\{\eta = k-i\}$
 $= \dots = ?$

$$\Rightarrow \xi + \eta \sim Po(\lambda + \nu)$$

Do calculation,
Show this

$$\xi, \eta \sim f_{\xi, \eta}(x, y) = f_\xi(x) f_\eta(y)$$

$$F_{\xi+\eta}(t) = P_{\xi, \eta}(H_t) = \iint_{H_t} f_{\xi, \eta}(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{t-x} f_\xi(x) f_\eta(y) dy dx$$

$$f_{\xi+\eta}(t) = \frac{d}{dt} F_{\xi+\eta}(t) = \int_{-\infty}^{\infty} f_\xi(x) f_\eta(t-x) dx$$

convolution $(f_\xi * f_\eta)(t)$

Ex. $\xi_i \sim N(u_i, \sigma_i^2)$ $i=1, 2$ $\xi_1 \perp \xi_2$

What is the distribution of $\xi_1 + \xi_2$

$$\xi_1, \xi_2 \sim N(0, 1) \text{ indep} \quad \xi_i = m_i + \sigma_i \zeta_i \text{ standardization}$$

$$\begin{aligned} \xi_1 + \xi_2 &= m_1 + m_2 + \sigma_1 \zeta_1 + \sigma_2 \zeta_2 \\ &= m_1 + m_2 + \sigma_1 (\zeta_1 + \sigma_2 \zeta_2) \quad \sigma = \frac{\sigma_1}{\sigma_2} \end{aligned}$$

$$\begin{aligned} f_{\xi_1 + \sigma \xi_2}(t) &= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2 + \frac{(t-x)^2}{\sigma^2})} dx = \dots \\ &= \frac{1}{\sqrt{2\pi} (1+\sigma^2)} \exp\left(-\frac{x^2}{2(1+\sigma^2)}\right) \end{aligned}$$

$$\Rightarrow \xi_1 + \sigma \xi_2 \sim N(0, 1+\sigma^2)$$

→ Standardisation. $\xi_1 + \xi_2 \sim N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$

Recall: a family of distributions is called stable if the sum of indep. variables with the distr. from this family belongs to this family.

So: Poisson distr & Normal family of distr are stable.

THE MAIN INEQUALITIES

ξ : rv with finite $E\xi$

1) Markov ineq. $P\{\xi \geq \varepsilon\} \leq \frac{E\xi}{\varepsilon}$ for $\xi \geq 0$ a.s. $\forall \varepsilon > 0$

proof: $E\xi = E(\xi(1_{\xi < \varepsilon} + 1_{\xi \geq \varepsilon})) = E\xi \underbrace{1_{\xi < \varepsilon}}_{> 0} + E\xi \underbrace{1_{\xi \geq \varepsilon}}_{\geq \varepsilon 1_{\xi \geq \varepsilon}}$
 $\geq 0 + E\varepsilon 1_{\xi \geq \varepsilon} = \varepsilon P\{\xi \geq \varepsilon\}$

2) Chebychev ineq. $P\{| \xi - E\xi | > \varepsilon\} \leq \frac{\text{Var } \xi}{\varepsilon^2}$

proof $= P\{(\xi - E\xi)^2 \geq \varepsilon^2\} \stackrel{\text{Markov}}{\leq} \frac{E(\xi - E\xi)^2}{\varepsilon^2} = \frac{\text{Var } \xi}{\varepsilon^2}$

3) Chernoff ineq. $P\{\xi \geq \varepsilon\} \leq \inf_{\lambda > 0} e^{-\lambda\varepsilon} Ee^{\lambda\xi}$

proof $= P\{e^{\lambda\xi} \geq e^{\lambda\varepsilon}\} \stackrel{\text{Markov}}{\leq} \frac{Ee^{\lambda\xi}}{e^{\lambda\varepsilon}} = e^{-\lambda\varepsilon} Ee^{\lambda\xi} \quad \forall \lambda > 0$

Take inf over λ

Föreläsning 12

Chebyshov i mag.

$$P\{|\xi - E\xi| \geq \varepsilon\} \leq \frac{\text{Var } \xi}{\varepsilon^2}$$

Chernoff bound:

$$P\{\xi \geq \varepsilon\} \leq \inf_{\lambda > 0} e^{-\lambda \varepsilon} E e^{\lambda \xi}$$

LAW OF LARGE NUMBERS (LLN)

A - event , $A \in \mathcal{F}$, (Ω, \mathcal{F}, P)

N repetitions of experiment has $(\Omega^n, \mathcal{F}^{\otimes n}, P^n)$ as prob. space

$$\Omega = \{(w_1, \dots, w_n) | w_i \in \Omega\}$$

$\mathbb{1}_A(w_i)$ r.v on $(\Omega^n, \mathcal{F}^{\otimes n}, P^n)$

Frequency of A : $\frac{1}{n} \sum_{i=1}^n \mathbb{1}_A(w_i) = \Phi_n(A) \xrightarrow{?} P(A)$

Def: A sequence ξ_n of r.v def on the same prob space
is said to converge to a r.v ξ (on same space)
in probability if

$$P\{|\xi_n - \xi| > \varepsilon\} \xrightarrow{n \rightarrow \infty} 0 \quad \forall \varepsilon > 0$$

Theorem (LLN)

$$\Phi_n(A) \xrightarrow{P} P(A) \quad (\text{in probability})$$

proof $\Phi_n(A) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_A(w_i)$

$\mathbb{1}_A \sim \text{Bern}(P(A))$ and $\mathbb{1}_A(w_i)$ are iid . $E \mathbb{1}_A = P(A)$

$$\text{Var } \mathbb{1}_A = P(A)(1-P(A))$$

$$P\{|Y_n - P(A)| > \epsilon\} = E[Y_n]$$

$$E Y_n = \frac{1}{n} n P(A) = P(A)$$

$$\text{Var } Y_n = \frac{1}{n^2} n P(A)(1-P(A)) = \frac{1}{n} P(A)(1-P(A))$$

$$P\{|Y_n - P(A)| > \epsilon\} \leq \frac{\text{Var } Y_n}{\epsilon^2} = \frac{P(A)(1-P(A))}{n \epsilon^2} \xrightarrow[n \rightarrow \infty]{} 0$$

$$\text{So } Y_n \xrightarrow{P} P(A) (= E Y_n)$$

Def: ξ_n on the same prob. space

$\xi_n \rightarrow \xi$ almost surely ($\xi_n \xrightarrow{\text{a.s.}} \xi$) if

$$P\{\omega : \xi_n(\omega) \not\rightarrow \xi(\omega)\} = 0$$

Thm (strong law of large numbers)

$$\text{SLLN: } Y_n(A) \xrightarrow{\text{a.s.}} P(A)$$

Def Given a seq. of events A_1, A_2, \dots the event that they are happening infinitely often is

$$\{A_n \text{ i.o.}\} = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \overline{\lim}_{n \rightarrow \infty} A_n$$

A_n happens finitely many times:

$$\{A_n \text{ i.o.}\}^c = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c = \underline{\lim}_{n \rightarrow \infty} A_n$$

Theorem Borel-Cantelli Lemma

1. If $\sum P(A_n) < \infty$ then $P\{A_n \text{ i.o.}\} = 0$

2. If (A_n) mutually indep then $P\{A_n \text{ i.o.}\} = 1 \Leftrightarrow \sum P(A_n) = \infty$

so for indep A_n 's there is only two possibilities: $P\{A_n \text{ i.o.}\} = \begin{cases} 1 \\ 0 \end{cases}$

0.1 law ↗

proof: $B_n = \bigcup_{k \geq n} A_k \rightarrow$ i.e. $B_{n+1} \subseteq B_n$

By continuity of prob.

$$\begin{aligned} P\{A_n \neq 0\} &= P\{\bigcap_{n \geq 1} B_n\} = \lim_{n \rightarrow \infty} P\{B_n\} = \lim_{n \rightarrow \infty} P\left(\bigcup_{k \geq n} A_k\right) \\ &\leq \lim_{n \rightarrow \infty} \sum_{k \geq n} P(A_k) \xrightarrow{n \rightarrow \infty} 0 \quad \text{if } \sum_k P(A_k) < \infty \end{aligned}$$

2. Let's show that $P\left\{\bigcup_{n \geq 1} \bigcap_{k \geq n} A_k^c\right\} = 0$

$$C_n \supseteq C_{n+1}, C_n \nearrow C_\infty$$

$$P\{A_n \neq 0\}^c = \lim_{n \rightarrow \infty} P(C_n) = \lim_{n \rightarrow \infty} \prod_{k \geq n} P(A_k^c)$$

$$1-x \leq e^{-x} \quad x \in [0,1]$$

$$\begin{aligned} &\leq \lim_{n \rightarrow \infty} \prod_{k \geq n} e^{-P(A_k)} = \lim_{n \rightarrow \infty} e^{-\sum_{k \geq n} P(A_k)} \\ &= \begin{cases} 0 & \text{if } \sum P(A_k) = \infty \\ 1 & \text{if } \sum P(A_k) < \infty \end{cases} \end{aligned}$$

proof of SLLN:

$$\varphi_n(A) \xrightarrow{a.s.} P(A) \quad \text{i.e.} \quad P\{\omega : \varphi_n(A, \omega) \xrightarrow{a.s.} P(A)\} = 0 \quad (*)$$



$$\forall m \exists n \text{ s.t. } a_n \rightarrow a \quad |a_n - a| \leq \frac{1}{m}$$

$$\exists m : \forall n \forall k \geq n |a_k - a| > \frac{1}{m}$$

$$\{\varphi_n \rightarrow P\} = \bigcup_m \underbrace{\{|\varphi_n - P| > \frac{1}{m} : \omega\}}_{D_m} \quad D_m \subseteq D_{m+1}, D_m \nearrow$$

$$P\{\varphi_n \rightarrow P\} = \lim_{m \rightarrow \infty} P\{|\varphi_n - P| > \frac{1}{m} : \omega\}$$

We aim to apply B.C lemma, to show that it is 0 provided

$$\sum_n P\left\{|\varphi_n - P| > \frac{1}{m}\right\} < \infty$$

$$P\{|Q_n - p| > \varepsilon\} \leq \frac{P(1-p)}{\varepsilon^2 n} \quad \text{Want work}$$

$$\text{Recall } Q_n = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_A(w_i)$$

$$B_n = n Q_n = \sum_{i=1}^n \mathbb{1}_A(w_i) \sim \text{Bin}(n, p)$$

$$\begin{aligned} E e^{\lambda B_n} &= E e^{\lambda \sum \mathbb{1}_A(w_i)} = E \prod_{i=1}^n e^{\lambda \mathbb{1}_A(w_i)} = \prod_{i=1}^n E e^{\lambda \mathbb{1}_A} \\ &= \prod_{i=1}^n (1(1-p) + p e^\lambda) = (1-p+pe^\lambda)^n \end{aligned}$$

$$P\{|Q_n - p| > \varepsilon\} \Rightarrow P\{|B_n - np| > n\varepsilon\} = P\{B_n > n(p+\varepsilon)\}$$

$$+ P\{B_n < n(p-\varepsilon)\}$$

$$1. \leq \inf_{\lambda > 0} e^{-\lambda n(p+\varepsilon)} E e^{\lambda B_n} \quad \text{Chernoff}$$

$$= \inf_{\lambda > 0} \exp \left\{ -n[\lambda(p+\varepsilon) - \log(1-p+pe^\lambda)] \right\} = \inf_{\lambda > 0} e^{-n f_1(\lambda)} =$$

$$f'(\lambda) = p+\varepsilon - \frac{pe^\lambda}{1-p+pe^\lambda} = 0$$

$$(p+\varepsilon)(1-p+px) = px \quad x = e^\lambda$$

$$(p+\varepsilon)(1-p) = px(1-p-\varepsilon)$$

$$x = \frac{(p+\varepsilon)/(1-p)}{p/(1-p-\varepsilon)} > 1 \quad e^\lambda > 1 \Rightarrow \lambda_1 > 0$$

$$\lambda_1 = \log \left[\frac{p+\varepsilon}{p} \frac{1-p}{1-p-\varepsilon} \right] > 0$$

... explanation ... λ_1 is max not min

$$= e^{-n f_1(\lambda_1)}$$

$$\Rightarrow \sum P(Q_n - p > \varepsilon) < \infty$$

2. Apply Chernoff ... exercise

proof finished

Actually, the same method can be applied
to show that

$$\xi_n \text{ iid rv} : \frac{1}{n} \sum_{k=1}^n \xi_k \xrightarrow{\text{a.s.}} E\xi$$

provided $Ee^{|\xi|} < \infty$

Actually: $\frac{1}{n} \sum_{k=1}^n \xi_k \xrightarrow{\text{a.s.}} E\xi$ if $E|\xi| < \infty$

Föreläsning 13

MODES OF CONVERGENCE OF RV

(ξ_n) , ξ on the same prob. space

1. $(\xi_n) \xrightarrow{P} \xi$ (in probability)

$$\lim_{n \rightarrow \infty} P\{| \xi_n - \xi | > \varepsilon\} = 0 \quad \forall \varepsilon > 0$$

2. $\xi_n \xrightarrow{\text{a.s.}} \xi$

$$P\{\omega : \xi_n(\omega) \rightarrow \xi(\omega)\} = 0 = P\left(\bigcup_{\varepsilon > 0} \bigcap_{n=1}^{\infty} \{|\xi_n - \xi| < \varepsilon\}\right) = (*)$$

3. Convergence in R-mean ($R \geq 1$)

$$\xi_n \xrightarrow{L_R} \xi \quad \text{if} \quad E |\xi_n - \xi|^R \xrightarrow{n \rightarrow \infty} 0$$

$$= \|\xi_n - \xi\|_{L_R} \quad \|\cdot\|_{L_R} : L_R \text{ norm on } (\Omega, \mathcal{F}, P)$$

$$\|\xi\|_{L_R} = \left(\int |\xi(\omega)|^R P d\omega \right)^{1/R}$$

(ξ_n) A ξ not necessarily on same prob. space

4. $\xi_n \xrightarrow{d} \xi$ (in distribution) if

$$F_n(x) \xrightarrow{n \rightarrow \infty} F(x) \quad \forall x \text{ points of continuity of } F$$

$$F_n(x) = F_{\xi_n}(x) = P\{\xi_n \leq x\}$$

$$F(x) = F_\xi(x) = P\{\xi \leq x\}$$

5. $\xi_n \xrightarrow{w} \xi$ (weakly, $\xi \Rightarrow \xi$) if

$$E f(\xi_n) \rightarrow E f(\xi) \quad \forall \text{ cont fct } f \text{ bounded}$$

L, conv: convergence in mean ($E|\xi_n - \xi| \rightarrow 0$)

l.i.m $\xi_n \rightarrow \xi$ notation

Theorem

$$\begin{array}{c} \xi_n \xrightarrow{\text{a.s.}} \xi \xrightarrow{2.} \xi_n \xrightarrow{P} \xi \xrightarrow{3.} \xi_n \xrightarrow{a.s.} \xi \xrightarrow{4.} \xi_n \xrightarrow{w} \xi \\ \xi_n \xrightarrow{\text{A.s.}} \xi \xrightarrow{1.} \end{array}$$

Almost Surely \Rightarrow in probability \Rightarrow in Distribution \Leftrightarrow weakly
in R-mean

No implication can in general be reversed

$$1. P\{|\xi_n - \xi| > \varepsilon\} = P\{|(\xi_n - \xi)^R| > \varepsilon^R\} \stackrel{\text{Markov}}{\leq} \frac{E|(\xi_n - \xi)^R|}{\varepsilon^R} \xrightarrow{R \rightarrow 0} 0 \quad \text{so } \xi_n \xrightarrow{P} \xi$$

2. Fix $\varepsilon_0 > 0$

$$\begin{aligned} 0 &\geq P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{|\xi_k - \xi| > \varepsilon_0\}\right) \quad \text{decreasing sequence} \\ &= \lim_{n \rightarrow \infty} P\left(\bigcup_{k=n}^{\infty} \{|\xi_k - \xi| > \varepsilon_0\}\right) \geq \lim_{n \rightarrow \infty} P\{|\xi_n - \xi| > \varepsilon_0\} = 0 \\ \text{so } \xi_n &\xrightarrow{\text{a.s.}} \xi \quad \forall \varepsilon_0 > 0 \end{aligned}$$

$$\begin{aligned} 3. F_n(x) &= F_{\xi_n}(x) = P\{\xi_n \leq x\} = \\ &= P\{\xi_n \leq x, \xi \leq x + \varepsilon\} + P\{\xi_n \leq x, \xi > x + \varepsilon\} \\ &\leq P\{\xi \leq x + \varepsilon\} + \underbrace{P\{|\xi_n - \xi| > \varepsilon\}}_{P_n} = F(x + \varepsilon) + P_n \\ \text{swap } \xi_n &\leftrightarrow \xi \quad P_n \xrightarrow{\text{a.s.}} 0 \\ x + \varepsilon &\rightarrow y \quad \text{since } \xi_n \xrightarrow{P} \xi \end{aligned}$$

$$F(y - \varepsilon) \leq F_n(y) + P_n \quad (y \rightarrow x)$$

$$F(x - \varepsilon) - P_n \leq F_n(x) \leq F(x - \varepsilon) + P_n$$

$$F(x - \varepsilon) \leq \liminf F_n(x) \leq \limsup F_n(x) \leq F(x + \varepsilon)$$

for x points of continuity of F , $F(x + \varepsilon) - F(x - \varepsilon)$ is arb small

Thus for such $x \exists \lim_{n \rightarrow \infty} F_n(x) = F(x)$

$$\text{so } \xi_n \xrightarrow{D} \xi$$

$$H. E f(\xi_n) - E f(\xi) = E [f(\xi_n) - f(\xi)]$$

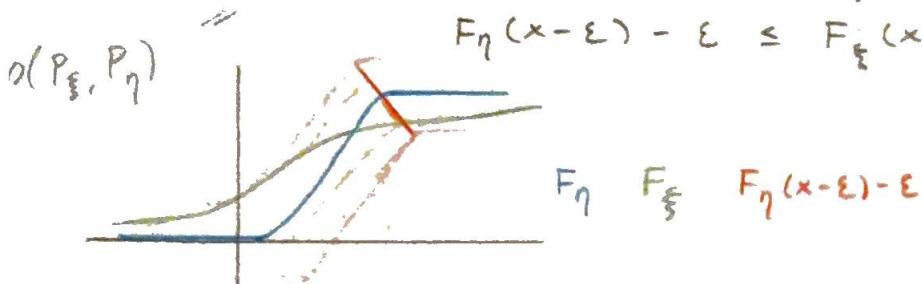
$$= \int f(x) F_n(dx) - \int f(x) F(dx).$$

$= \int f(x) ((F_n - F)(dx))$ small whenever F_n close to F

the weak conv. is metrisable with so called Lévy-Prokhorov metric

$$\rho(\xi, \eta) = \inf_{\varepsilon > 0} \{ F_\xi(x-\varepsilon) - \varepsilon \leq F_\eta(x) \leq F_\xi(x+\varepsilon) + \varepsilon$$

$$F_\eta(x-\varepsilon) - \varepsilon \leq F_\xi(x) \leq F_\eta(x+\varepsilon) + \varepsilon \}$$



Note: when $\rho = 0$ the fct. conv. de

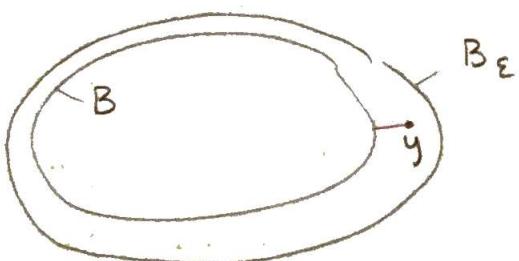
$$\xi_n \xrightarrow{\text{w}} \xi \Leftrightarrow \rho(\xi_n, \xi) \rightarrow 0$$

Moreover with ρ , the space of distr. is Polish
(= complete separable metric space)

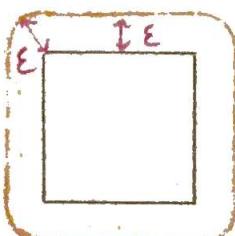
This is just idea of proof \rightarrow

$\xi : \Omega \rightarrow X$ loc. compact separable metric space:

$$\rho(P_\xi, P_\eta) = \inf_{\varepsilon > 0} \{ P_\xi(B_\varepsilon) \leq P_\eta(B) + \varepsilon, P_\eta(B_\varepsilon) \leq P_\xi(B) + \varepsilon \}$$

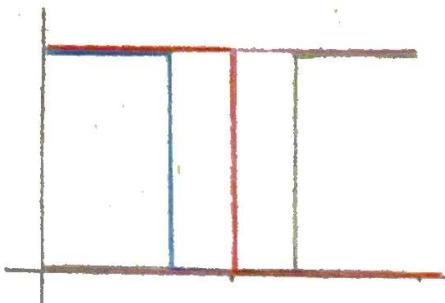


$$\begin{aligned} d(y, B) &\leq \varepsilon \\ &= \inf_{x \in B} d(y, x) \end{aligned}$$



Counter ex.: $\xi_n \xrightarrow{P} \xi \not\Rightarrow \xi_n \xrightarrow{as} \xi$

$\Omega = (\Omega_0, \Omega_1, \mathcal{B}, \mathcal{E})$



$$\begin{aligned}\xi_n^{(1)} &= \mathbb{1}_{[0, 1/2]}, & \xi_n^{(2)} &= \mathbb{1}_{[1/2, 1]} \\ \xi_n^{(3)} &= \mathbb{1}_{[0, 1]}, & \xi_n^{(4)} &= \mathbb{1}_{[1/2, 1]} \\ &\vdots \\ \xi_n^{(K)} &= \mathbb{1}_{[\frac{n-1}{K}, \frac{n}{K}]}\end{aligned}\quad n=1, 2, \dots, K$$

$P\{| \xi_n - 0 | > \varepsilon\} = \text{trivalue}$

$= |\{\omega : \xi_n(\omega) > \varepsilon\}| = \text{width of spike} \rightarrow 0$

so $\xi_n \xrightarrow{P} 0$

in contrast: $\xi_n(\omega) \rightarrow 0$

so $\xi_n \xrightarrow{as} 0$

3. Cannot be reversed because $\xi_n \xrightarrow{w} \xi$ does not assume that $(\xi_n), \xi$ are on the same Ω , BUT:

Skorokhod embedding theorem

$\exists (\Omega', \mathcal{F}', P')$ on which one may define $\xi'_n, \xi' : \Omega' \rightarrow \mathbb{R}$
s.t. $\xi'_n \stackrel{D}{=} \xi_n$, $\xi' \stackrel{D}{=} \xi$ & $\xi' \xrightarrow{as} \xi$

If ξ has (strictly increasing) F_ξ (pdf)

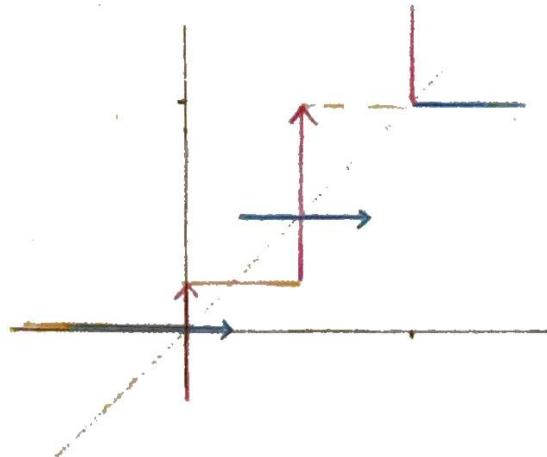
Then $u \sim \text{unit } [0, 1] \rightarrow F_\xi^{-1}(u) \stackrel{D}{=} \xi$

$$\begin{aligned}F_{\xi'}(x) &= P\{F_\xi^{-1}(\omega) \leq x\} = P\{u \leq F_\xi(x)\} \\ &= F_u(F_\xi(x)) = F_\xi(x)\end{aligned}$$

Generalized inverse :

$$\xi' = F_{\xi}^{-1}(u) \quad \text{where } F_{\xi}(x) = \inf$$

$$\text{where } F_{\xi}^{-1}(x) = \inf_{\omega} \{\omega : \omega \leq F_{\xi}(x)\}$$



$$(\Omega', \mathcal{F}', P') = ([0, 1], \mathcal{B}, \ell) \quad \xi_n'(w) = F_{\xi_n}^{-1}(w)$$

Foreläsning 14

CHARACTERISTIC FCT

Def: A char. fct. (chf) of a rv ξ is

$$\begin{aligned}\Phi_\xi(t) &= E e^{it\xi} = \int_{-\infty}^{\infty} e^{itx} P_\xi(dx) \quad t \in \mathbb{R} \\ &= E \cos(tx) + i E \sin(tx)\end{aligned}$$

$$\Phi_\xi(0) = E \underbrace{e^{i0\xi}}_1 = 1 \quad (1)$$

$$|\Phi_\xi(t)| = |E e^{it\xi}| \leq E |e^{it\xi}| = 1 \quad (2)$$

$$\begin{aligned}|\Phi_\xi(t+s) - \Phi_\xi(t)| &= |E [e^{i(t+s)\xi} - e^{it\xi}]| \leq E |e^{it\xi} (e^{is\xi} - 1)| \\ &= E |e^{it\xi}| |e^{is\xi} - 1| \quad e^{is\xi} - 1 = s + iS\xi - 1 + o(s\xi) \\ &\quad |e^{is\xi} - 1| \leq 2 \Rightarrow\end{aligned} \quad (3)$$

By Lebesgue dominated convergence theorem:

$$\lim_{s \rightarrow 0} E |e^{is\xi} - 1| = E \lim_{s \rightarrow 0} \underbrace{|e^{is\xi} - 1|}_{} = 0$$

Thus $\Phi_\xi(t)$ is uniformly cont.

$\Phi_\xi(t)$ is pos def: (4)

$$\sum_{k,l} \Phi_\xi(t_k - t_l) z_k \bar{z}_l \geq 0 \quad \forall t_k, t_l \in \mathbb{R} \quad z_k, z_l \in \mathbb{C}$$

$$= \sum_{k,l} E e^{i(t_k - t_l)\xi} z_k \bar{z}_l = E \sum_{k,l} e^{it_k \xi} \overline{e^{it_l \xi}} z_k \bar{z}_l$$

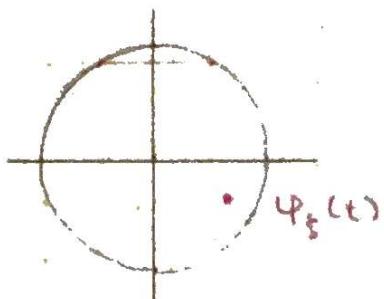
$$= E \sum_k e^{it_k \xi} z_k \sum_l \overline{e^{it_l \xi}} z_l = E \left| \sum_k e^{it_k \xi} z_k \right|^2$$

Theorem (Bochner)

Any fct satisfying (1)-(4) is a Fourier transform
of a probability measure μ

$$\text{i.e. } \varphi(t) = \int e^{itx} \mu(dx)$$

Equivalently define a rv ξ distr as μ $\varphi(t) = E e^{it\xi}$



point on unit circle

meaning: center of mass

Theorem

If $E|\xi|^n < \infty$ then φ_ξ is n times differentiable at 0 ($\varphi_\xi \in D^n(0)$) and

$$\frac{d^n}{dt^n} \varphi_\xi(0) = i^n E \xi^n$$

proof: $e^{itz} = \sum_{k=0}^{\infty} \frac{(it\xi)^k}{k!} + \frac{(it\xi)^n}{n!} (\cos \theta_1 t\xi + i \sin \theta_2 t\xi) \quad 0 < \theta_1, \theta_2 \leq 1$

taylor $= \sum_{k=0}^n \frac{(it\xi)^k}{k!} + \frac{(it\xi)^n}{n!} (\cos \theta_1 t\xi + i \sin \theta_2 t\xi - 1)$

$E|\cdot|$

$$|E\xi(t)| \leq |\xi|^n \cdot 3$$

$$\text{By assumption: } E|\xi(t)| \leq 3 E|\xi|^n < \infty$$

By Lebesgue dominated conv. we may pass to the $\lim_{t \rightarrow 0}$ under E to get that $E\xi(t) \rightarrow 0$

$$\text{Thus } Ee^{itz} = 1 + \sum_{k=1}^n \frac{(it\xi)^k}{k!} E\xi^k + \frac{it^n}{n!} E\xi^n \xrightarrow{O(t^n)} 0$$

$$\Rightarrow \frac{d^k}{dt^k} \varphi_\xi(0) = i E \xi^k \quad \forall k=1, \dots, n$$

So by differentiating you can recover all moments!

Examples :

$$1) \xi \sim \text{Bern}(p) : E e^{it\xi} = 1(1-p) + e^{it} \cdot p \\ = 1 - p + p e^{it}$$

Notice that $it = -\lambda \Rightarrow$

$$\Phi_\xi(t) = E e^{-\lambda t} \sim \text{Laplace transform of } \xi$$

$$E e^{\lambda x} = \int_{-\infty}^{\infty} e^{-\lambda x} f_\xi(x) dx$$

If ξ has a pdt $f_\xi(|y|) \sim \frac{1}{y^2}$ $y \rightarrow \infty$

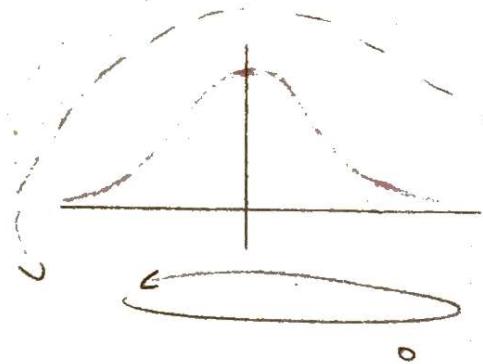
$$\text{then } \int_{-\infty}^{\infty} e^{-\lambda x} f_\xi(x) dx = \infty$$

it belongs to neg x values

But char tct is analytic, def on $m(z)$ but not nec. on whole no. derivatives up to moment exist. complex plane

$$2) \xi \sim N(0, 1)$$

$$\begin{aligned} \Phi_\xi(t) &= E e^{it\xi} = \frac{1}{\sqrt{2\pi}} \int e^{itx - \frac{x^2}{2}} dx = \|it = s\| \\ &= e^{\frac{s^2}{2}} \frac{1}{\sqrt{2\pi}} \int e^{-(x-s)^2/2} dx = e^{\frac{s^2}{2}} = e \end{aligned}$$



Shift Theorem

$$\Phi_{a\xi+b}(t) = E e^{it(a\xi+b)} = E e^{itb} e^{ita} = e^{itb} \Phi_\xi(at)$$

$$3) \xi \sim N(m, \sigma^2)$$

$\xi \stackrel{D}{=} -\sigma \cdot m$ first scale, then shift.

$$\Psi_{\xi}(t) = e^{itm} \Psi_{\xi}(0t) = e^{itm - \sigma^2 t^2/2}$$

CHAR. FCT OF SUM OF R.V

$\xi_1 + \xi_2$

$$\Psi_{\xi_1 + \xi_2}(t) = E e^{it(\xi_1 + \xi_2)} = E e^{it\xi_1} E e^{it\xi_2} = \Psi_{\xi_1}(t) \Psi_{\xi_2}(t)$$

In particular: if ξ_1, ξ_2 are abs cont with pdf. f_1, f_2
then $\xi_1 + \xi_2$ has pdt $f_1 * f_2$

$$\begin{aligned}\Psi_{\xi_1 + \xi_2}(t) &= \int f_1 * f_2(x) e^{itx} dx = F[f_1 * f_2] \\ &= F[f_1] F[f_2] = \int f_1(x) e^{itx} dx \int f_2(x) e^{itx} dx \\ &= \Psi_{\xi_1}(t) \Psi_{\xi_2}(t)\end{aligned}$$

1) $\beta_n \sim \text{Bin}(n, p)$

complicate.

$$\Psi_{\beta_n}(t) = E e^{it\beta_n} = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} e^{itk} = ?$$

$$\beta_n = X_1 + \dots + X_n$$

$$X_k = \begin{cases} 0 & k\text{'th Bern trial is fail} \\ 1 & \text{success} \end{cases}$$

$$X_k \text{ iid } \text{Bern}(p) \quad \Psi_{X_k}(t) = 1-p + pe^{it}$$

$$\Psi_{\beta_n}(t) = \prod_{k=1}^n \Psi_{X_k}(t) = (1-p + pe^{it})^n$$

Recall that if $F[f] = \int e^{itx} f(x) dx = \Psi(t)$

then $f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \Psi(t) dt$

* points x where f is cont

The inverse formula:

If Φ_{ξ} is char fct of ξ with cdf F_{ξ} then

$$F_{\xi}(b) - F_{\xi}(a) = \frac{1}{2\pi} \lim_{c \rightarrow \infty} \int_{-c}^c \frac{e^{-iat} - e^{-ibt}}{it} \Phi_{\xi}(t) dt$$

increments received from

$\forall a, b$ pts of cont
of F_{ξ}

$$= P\{\xi \in (a, b)\}$$

otherwise

$$\frac{1}{2}(F(b) - F(a)) + F(a) + F(b)$$

Thus P is completely determined by $\Phi_{\xi}(t)$

Continuity thm

Assume that ξ_1, ξ_2, \dots with char fct $\Phi_{\xi_1}, \Phi_{\xi_2}, \dots$
and $\Phi_{\xi_n}(t) \rightarrow \Phi_{\xi}(t) \quad \forall t$

$$F_{\xi_n}(b) - F_{\xi_n}(a) \rightarrow F_{\xi}(b) - F_{\xi}(a) \quad \forall \text{ pts } a, b \text{ of cont. of } F_{\xi}$$

$$\Rightarrow \xi_n \xrightarrow{\text{dist}} \xi \quad \text{convergence in distribution} \\ = \text{weak conv.}$$

Moreover if $\Phi_{\xi_n}(t) \rightarrow \Phi(t) \quad \forall t$

s.t. ~~$\Phi(0) = 1$~~ & cont at $t=0$

then $\exists \xi$ s.t. $\Phi(t) = \Phi_{\xi}(t)$

Poisson limit thm

(revisited) if $n p_n \rightarrow \lambda \in (0, +\infty)$

$$\text{then } P\{\xi_n = k\} \rightarrow \frac{\lambda^k}{k!} e^{-\lambda} = P\{\Pi_{\lambda} = k\}$$

$$\Pi_{\lambda} \sim Po(\lambda) \quad B_n \sim Bin(n, p_n)$$



$$P_n(k) = (k) p_n^k (1-p_n)^{n-k}$$

poisson similar but cont. to int.

Equivalently $\beta_n \xrightarrow{D} \pi_\lambda$ or $\beta_n \xrightarrow{\text{weak}} \pi_\lambda$

weak: $E F(\beta_n) \rightarrow E F(\pi_\lambda)$ & F bounded cont. fct.

Thus we prove the theorem if we show that

$\Phi_{\beta_n}(t) \rightarrow \Phi_{\pi_\lambda}(t) \quad \forall t$ by cont. thm

$$\pi_\lambda \sim P_0(\lambda)$$

$$\begin{aligned}\Phi_{\pi_\lambda} &= E e^{it\pi_\lambda} = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} \\ &= e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}\end{aligned}$$

$$\begin{aligned}\Phi_{\beta_n}(t) &= (1 - p_n + p_n e^{it})^n \\ &= \exp \left\{ n \log \left(1 - \frac{\lambda}{n} + \frac{\lambda}{n} e^{it} + O\left(\frac{1}{n}\right) \right) \right\} \\ &= \exp \left\{ n \left(\frac{\lambda}{n} (e^{it}-1) + O\left(\frac{1}{n}\right) \right) \right\} \rightarrow \exp \{ \lambda (e^{it}-1) \}\end{aligned}$$

CENTRAL LIMIT THEOREM

Theorem (CLT)

Let $\xi_1, \xi_2, \xi_3, \dots$ be iid r.v.

with $E \xi_k = m$ & $\text{Var } \xi_k = \sigma^2$

Let $S_n = \sum_{k=1}^n \xi_k$

Then $\frac{S_n - E S_n}{\sqrt{\text{Var } S_n}} = \frac{S_n - nm}{\sigma \sqrt{n}} \xrightarrow[\text{conv. in dist.}]{D} \zeta \sim N(0, 1)$

proof: consider $\eta_k = \frac{1}{\sigma} (\xi_k - m)$

$$E \eta_k = \frac{1}{\sigma} [E \xi_k - m] = 0$$

$$\text{Var } \eta_k = \frac{1}{\sigma^2} \text{Var} [\xi_k - m] = \frac{1}{\sigma^2} \sigma^2 = 1$$

$$\frac{S_n - nm}{\sigma \sqrt{n}} = \left[\xi_k = \sigma \eta_k + m \right] = \frac{\sigma \sum_{k=1}^n \eta_k + nm - nm}{\sigma \sqrt{n}}$$
$$= \frac{\sum_{k=1}^n \eta_k}{\sqrt{n}} = Z_n$$

Let $\Phi(t) = E e^{it\eta_k} = 1 + it E\eta_k + \frac{(it)^2}{2} E\eta_k^2 + o(t^2)$

because $E|\eta_k|^2 = E\eta_k^2 = \text{Var } \eta_k = 1 < \infty$

$$\Rightarrow \Phi_\eta(t) = 1 - \frac{t^2}{2} + o(t^2)$$

$$\begin{aligned} \Phi_\xi(t) &\stackrel{\text{shift}}{=} \Phi_{\sum \eta_k} \left(\frac{t}{\sigma} \right) = \left[\Phi_\eta \left(\frac{t}{\sigma} \right) \right]^n = \exp \{ n \log \Phi_\eta \left(\frac{t}{\sigma} \right) \} \\ &= \exp \{ n \log (1 - \frac{t^2}{2\sigma^2} + o(\frac{t^2}{\sigma^2})) \} \\ &= \exp \{ n (-\frac{t^2}{2\sigma^2} + o(\frac{t^2}{\sigma^2})) \} \rightarrow \exp \{ -t^2/2 \} \\ &= \Phi_\xi(t) \end{aligned}$$

What if second moment doesn't exist?

What if not iid?

;

etc

1. If ξ_k are indep but not identically distr. with
 $E\xi_k = m_k$, $\text{Var } \xi_k = \sigma_k^2$

$$\frac{S_n - E S_n}{\sqrt{\text{Var } S_n}} = \frac{S_n - \sum_{k=1}^n m_k}{\sqrt{\sum \sigma_k^2}} \xrightarrow{w} \zeta$$

provided additional Lindeberg condition

$$B_n = \sum_{k=1}^n \sigma_k^2 \quad \forall \varepsilon \quad \sum_{k=1}^n E[(\xi_k - m_k) \mathbb{1}_{|\xi_k - m_k| > \varepsilon B_n}] \rightarrow 0$$

Sum of variances should behave like n

2. If ξ_k are iid but $\text{Var } \xi_k = +\infty$

If $E\xi_k = m < \infty$

$$\frac{S_n - nm}{n^{1/\alpha}} \xrightarrow{w} \theta \quad (*) \quad \alpha < 2$$

By changing ξ_k can always consider

$$\xi'_k = \xi_k - m_k \quad \text{if exist.} \quad \xi'_k = \xi_k \quad \text{otherwise}$$

$$S_n = \sum_{k=1}^n \xi'_k$$

$$\underbrace{\xi'_1 + \xi'_2 + \dots + \xi'_n}_{np} = S_n \quad \underbrace{\dots}_{(1-p)n}$$

take $p \in (0, 1)$

$$\theta \xleftarrow{w} \frac{S_n}{n^{1/\alpha}} = \frac{S_{pn}}{n^{1/\alpha}} + \frac{S_{(1-p)n}}{n^{1/\alpha}} = p^{1/\alpha} \underbrace{\frac{S_{pn}}{(pn)^{1/\alpha}}}_{\theta'} + (1-p)^{1/\alpha} \underbrace{\frac{S_{(1-p)n}}{((1-p)n)^{1/\alpha}}}_{\theta''} \quad \theta' = \theta \quad \theta'' = \theta$$

If $(*)$ for some $\alpha \leq 2$ then the r.v. Θ
(or its distr.) satisfies

$$\Theta = p^{1/\alpha} \theta' + (1-p)^{1/\alpha} \theta'' \quad \text{where } \begin{matrix} \theta' \stackrel{\mathcal{D}}{=} \theta'' \stackrel{\mathcal{D}}{=} \Theta \\ \text{indep.} \end{matrix}$$

$\alpha = 1/2$: CLT, $\Theta = \xi$

$$\sqrt{p} \xi' + \sqrt{1-p} \xi'' \sim N(0, p+1-p) = N(0, 1) \sim \xi$$

$N(0, p) \sim N(0, 1-p)$

α -stable distr. have $\Phi_{\theta_\alpha}(t) = e^{-\gamma|t|^\alpha}$

$$\alpha=2 \Rightarrow \gamma=1/2$$

$\alpha=1$ cauchy distr

Self Decomposable: $\Theta \stackrel{\mathcal{D}}{=} p^{1/\alpha} \theta' + \Xi$

random vectors same

etc. —