Recap ch. 1-3

Probability:

Experiment: An act that produces list of outcomes.

$\omega = \Omega$ $\quad \Omega$ = sample space.

Event: an event is a subset of Ω

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

$A = \{H\}$ we get a head.

Trial: One repetition of an experiment
"Doing the experiment"

Def Ω is a set then we say that $\mathcal{F}$ is
σ-algebra if the following holds:

1) $\Omega \in \mathcal{F}$ "something happened"

2) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

3) $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow A_1 \cup A_2 \cup \ldots \in \mathcal{F}$

$\wedge$ = "and" $\Leftrightarrow \cap$

$\vee$ = "or" $\Leftrightarrow \cup$

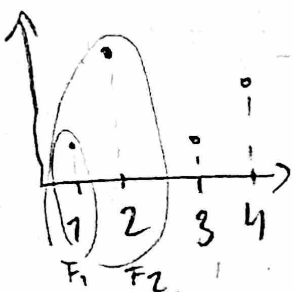
$a \leq b \underbrace{\wedge}_{min} c$

$a \leq b \underbrace{\vee}_{max} c$

Skapad med Tiny Scanner

Ex. Let $F_1, F_2, \dots$ be the family of σ-algebras associated to a Markov process X_t, then X_t is observable in $\mathcal{F}_t$ but not $\mathcal{F}_s$ when $s \le t$.

$$\mathcal{F}_t : \{\{\omega : X_1(\omega) = x_1, \dots, X_t(\omega) = x_t \,;\, (x_1, \dots, x_t) \in \underset{\mathbb{R}^t}{X^t}$$



σ-algebra: set of things that are observable.

Def. $(\Omega, \mathcal{F})$ a $\overbrace{\text{probability measure}}^{\text{assigns values to events}}$ is a function $\mathbb{P} : \mathcal{F} \to [0,1]$

Rules.

1) $\mathbb{P}(\Omega) = 1$ "something happens"
2) $A, B \in \mathcal{F}$ $\quad A \cap B = \emptyset$
   then $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$
   "the Addition rule"

2*) $A_1, A_2, \dots \in \mathcal{F}$
   $A_i \cap A_j = \emptyset \quad i \neq j \Rightarrow \mathbb{P}(\bigcup A_i) = \sum_i \mathbb{P}(A_i)$

Skapad med Tiny Scanner

Group Ass. 7:

We don't know if $\lim_{x \to \infty} \mathbb{P}(X \leq x) = 1$

The counterexample would be:

Maybe $\mathbb{P}(X = \infty) \neq 0$

We say that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability triple

Condition:

* $\mathcal{F}$ defines what is <u>observable</u> and $\mathbb{P}$ prescribes its <u>values</u>.

in our course we want to infer the values of this triple, which is unknown. We want to infer the triplet based of our data

$\Omega \in \mathcal{F}$ as we want to make sure that any outcome we might get belongs to the sample space.

<u>Conditional probability:</u>

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be given and let $A, B \in \mathcal{F}$ and $\mathbb{P}(A) > 0$

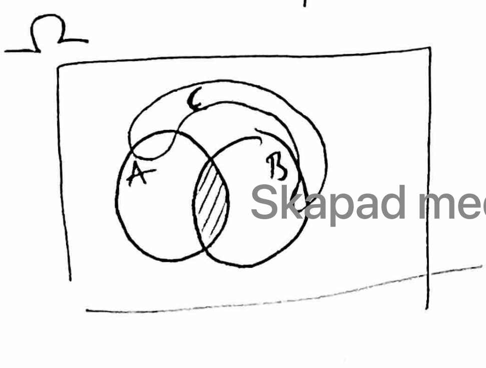
$$\mathbb{P}(B|A) = \frac{P(A \cap B)}{P(A)}$$

spam vs. not spam example.
B: is this spam or not?
A: word "free" or "prize"

"A is evidence. Use that to change the" distribution of the prediction of B.

"Prediction problem" A = evidence
B = "what you want to predict"

Ω

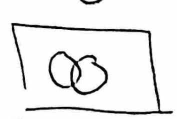


▨ $= A \cap B$

$\Omega^* = A$

$\mathbb{P}(\Omega^* | A) = 1$

group ass. 1
verify all axioms



$\mathcal{F}^* = \{E \cap A, E \in \mathcal{F}\}$

$\mathbb{P}^*(E) = \mathbb{P}(E|A)$ then verify all axioms

$B \cap C \neq 0$ then we might have

$$(B \cap A) \cap (C \cap A) = \emptyset$$

$$\mathbb{P}^* (C^* \cup B^*) = \mathbb{P}^*(C^*) + \mathbb{P}^*(B^*)$$

$$C^* = A \cap C \qquad B^* = B \cap A$$

Random variables

Def. A $\mathbb{R}$-valued RV is a function

$X : \Omega \to \mathbb{R}$

"$(\Omega, \mathcal{F}, \mathbb{P})$ is given."

such that

$\{\omega : X(\omega) \leq x\} \in \mathcal{F}$

$\Rightarrow \mathbb{P}(\{\omega \ \ X(\omega) \leq x\})$ well defined

$\mathbb{P}(X \leq x) = F_X(x)$ the cumulative distribution function

CDF.

1) Discrete: Discrete outcomes $\{0, 1, 2, \ldots\}$

$\{\pi, 2\pi, 3, \pi\}$

X is discrete

Q: $\mathbb{P}(X = 0) = 1/2$?

yes.

Continuous RV

$$\mathbb{P}(X \le x) = \int_{-\infty}^{x} f(s)\, ds$$

↑ density

Q: $\mathbb{P}(X = 0) = 1/2$?

No

In the notes we $\mathbb{P}(X \le x) = \int_{-\infty}^{x} dF_X$

Expectation

Let X be discrete $0, 1, 2, 3$

$$\mathbb{P}(X=0) = P_0$$
$$\mathbb{P}(X=0) = P_1$$
$$\vdots$$

$$E[X] = \sum_{i=0}^{3} i \cdot P_i$$

General case:

$$E[X] = \int_{-\infty}^{\infty} x \cdot dF_X$$

In continuous case:

$$E[X] = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

In discrete case: → PMF $P(X=x)$

$$E[X] = \sum_{x \in \mathbb{X}} x \cdot f(x)$$

$\mathbb{X}$ ← set of all possible outcomes

Properties

1) $E[aX+b]$ $\quad a, b \in \mathbb{R}$

$= aE[X]+b$

2) $\underbrace{E[X+Y]}_{\text{joint dist. function of } X \& Y} = E[X]+E[Y]$

3) $E[XY] = E[X]E[Y]$

this is not always true.

It is true if X & Y are independent.

Rule 3 $\quad Var(X+Y) = E[(X+Y-E(X+Y))^2]$

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

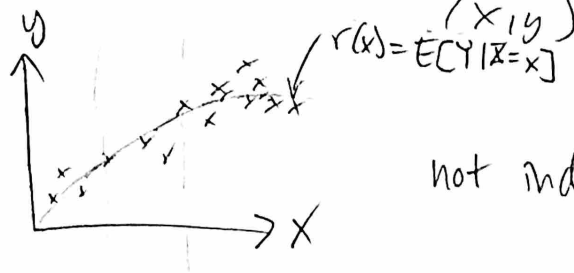
as the covariance (dot product) is zero.

Def. X and Y are independent if the joint dist. func. can be written as a product.

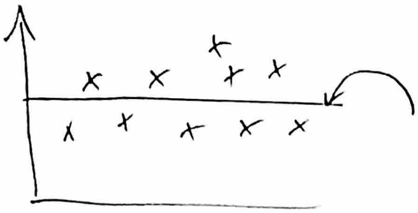
$F_{XY}(x,y) = F_X(x) F_Y(y)$

for densities $f_{XY}(x,y) = f_{Y|X}(y|x) f_X(x)$

Example



not independence



independence

Concentration

tail bounds imply concentration.

Concentration: whatever you RV is is concentrated around a specific value. The better it is concentrated, the better it is.

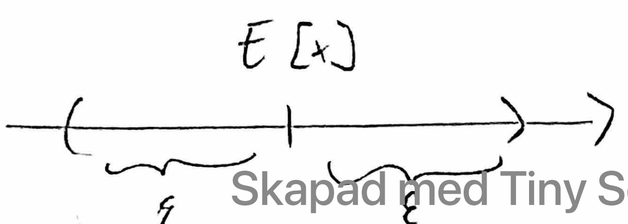
How much does the empirical mean "concentrate" around its expectation?

Theorem Hoeffding

$X \in \{0,1\}$

$$\underbrace{\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - E[X]\right| > \varepsilon\right)} \leq 2e^{-2n\varepsilon^2} = \alpha$$

the prob. that you deviate from the expectation more than epsilon is something small



$\mathbb{P}(E[X] \in I_n) \geq 1 - \alpha$