Stochastic Processes
Introduction to Probability

Stochastic Processes for
CS4070 multivariate data analysis

Dr. David Tax

Slides 'inspired' by:
Prof.dr.ir. Inald Lagendijk

Delft University of Technology

Organisation CS4070

- Stochastic Processes is the first part of:
- CS4070 **Multivariate Data Analysis**
  - together with Jakob Söhl

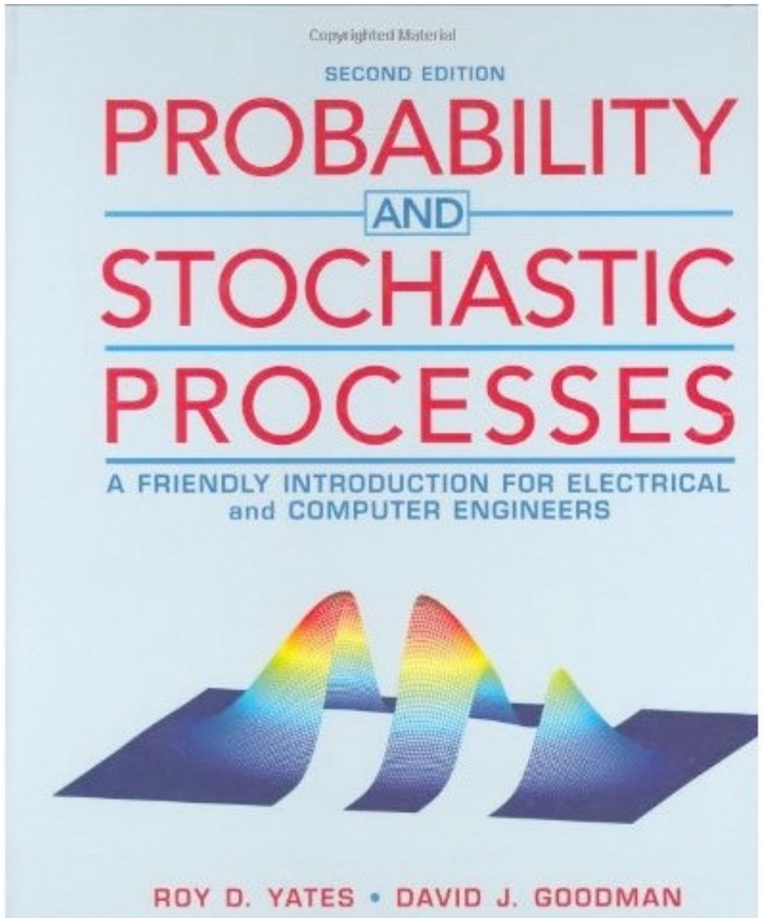
- Part I: Stochastic Processes, by me: fundaments of probability theory, extend to time series
- Part II: Statistical Learning, by Jakob: analysis of general data, frequentist- and Bayesian-approach

Organisation Part I

- Plenary Lectures (HoorCollege): 7x
- Working Groups (WerkCollege): 7x

- Book: R.D. Yates and D.J. Goodman, Probability and Stochastic Processes: A Friendly Introduction for Electrical and Computer Engineers (Second Edition)
- UHM: (There is a third edition now, also fine)

- Sheets, exercises, etc: Brightspace



More about Working Groups

- Plenary lectures:
  - Cover a lot of ground,
  - Don’t necessarily stick to the subject order in book
  - Should give you overview, red line, main issues
- In preparation of working group:
  - Read entire chapter(s) at home: (topics in ‘list.pdf’ on Brightspace)
  - Study the examples, formulate questions to ask
  - Be ready to answer questions from lecturer
- Working groups:
  - Reflect on plenary lectures and discuss tough issues or problems
  - Solve selected exercises
- Problems solved: See Brightspace

Final Computer Exercise

- At the end of the quarter on Stochastic Processes: a final computer exercise
- Pass or Fail
- You have to pass it before you can get the exam grade

- It will appear on Brightspace (Still constructing the exercise)

Final exam

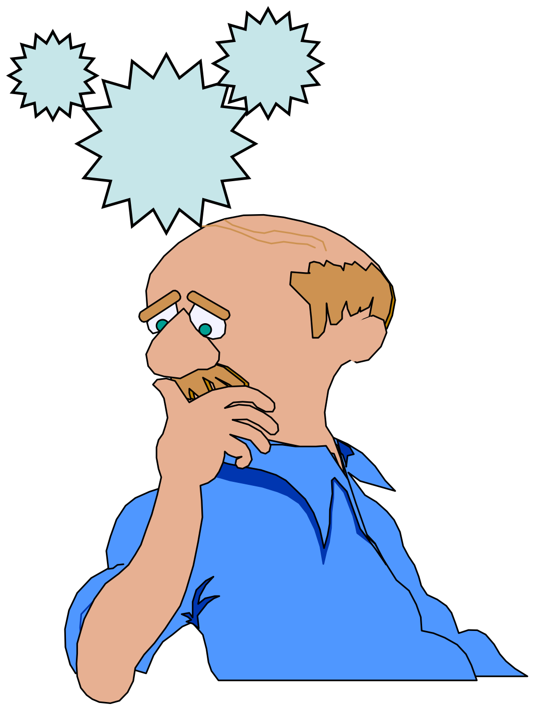
- Exam will be a physical written exam (but it depends on the Corona situation)
- I'll make a practice exam available.

- For the exam, you're allowed to bring one A4 paper with your own notes.

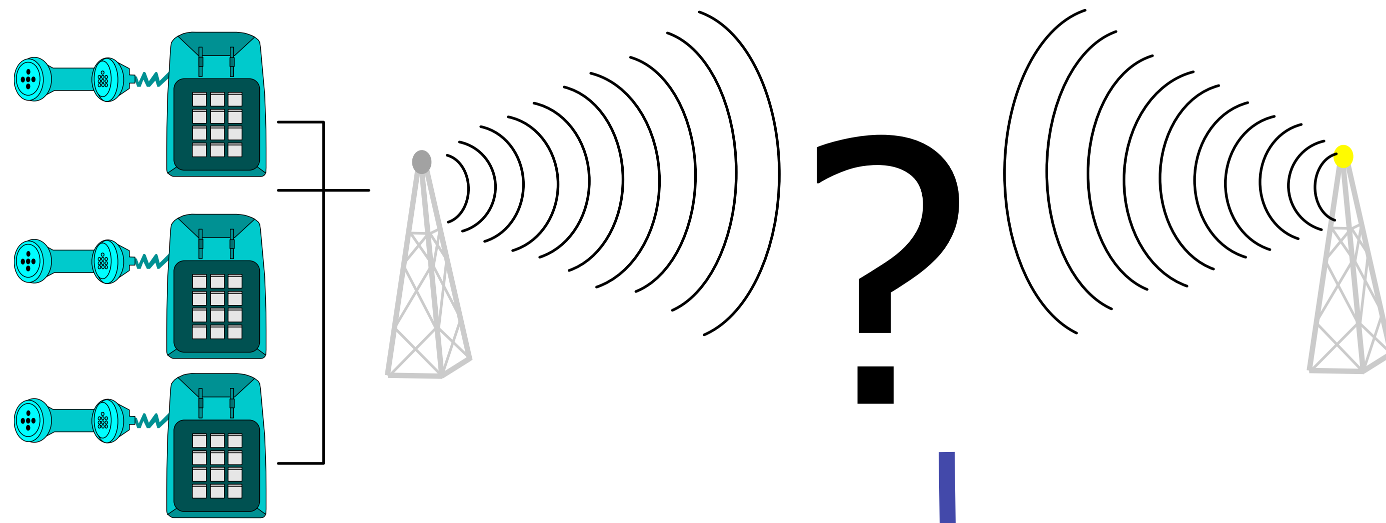
Chapters 1, 2, and 3

Motivation
Probability Models
Discrete and Continuous Random Variables
PMF,PDF,expectation

The Engineer: Problem Solver



Reality



Simplification:

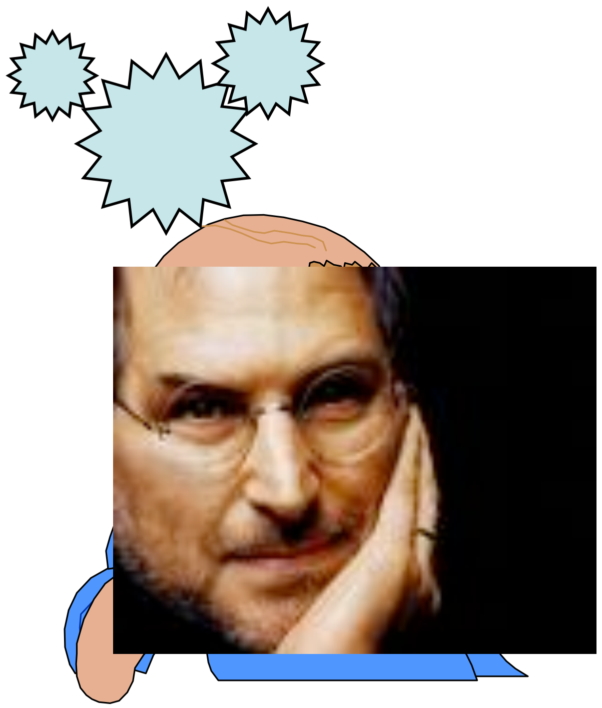
Model

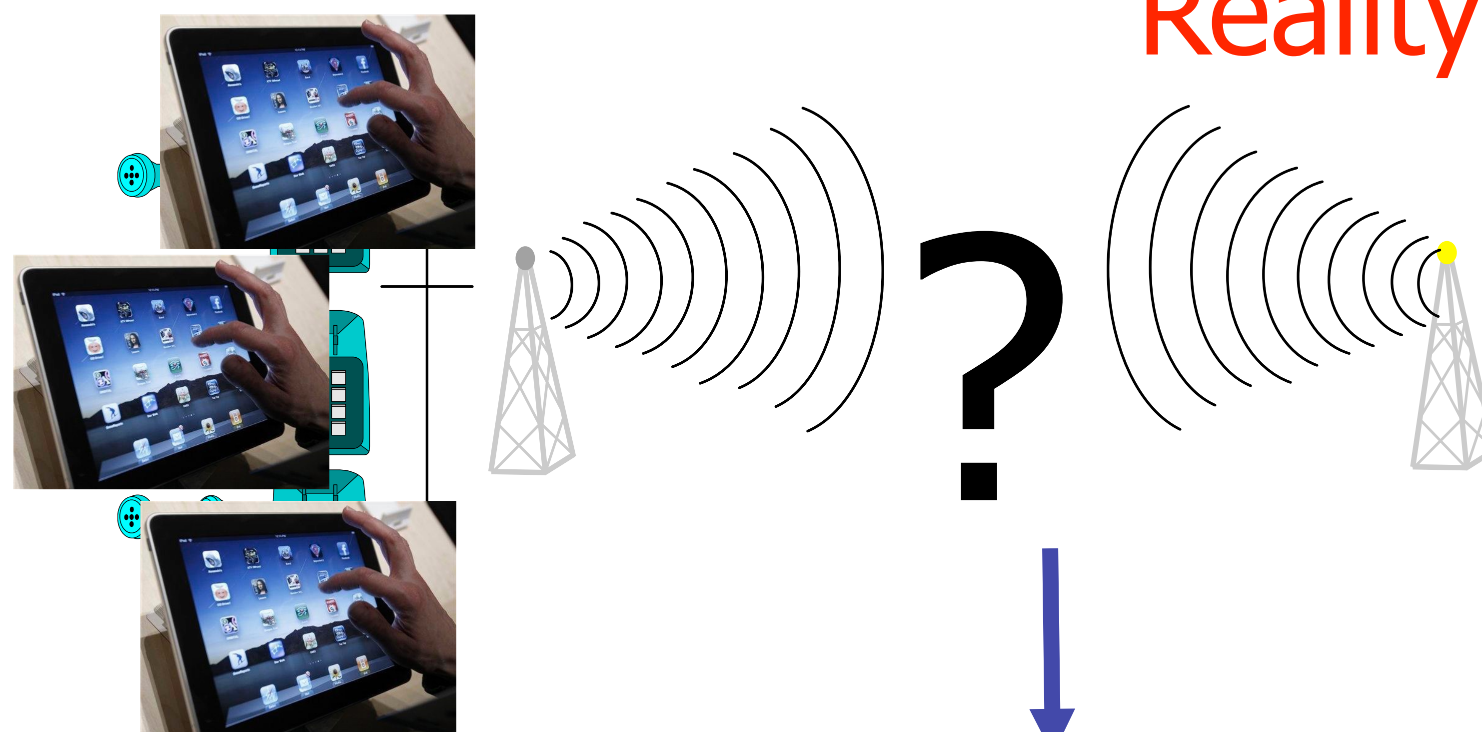
Solution

Tools:

Analysis
Algebra
Physics
Probability theory

The Engineer: Problem Solver





Reality

?

Simplification:

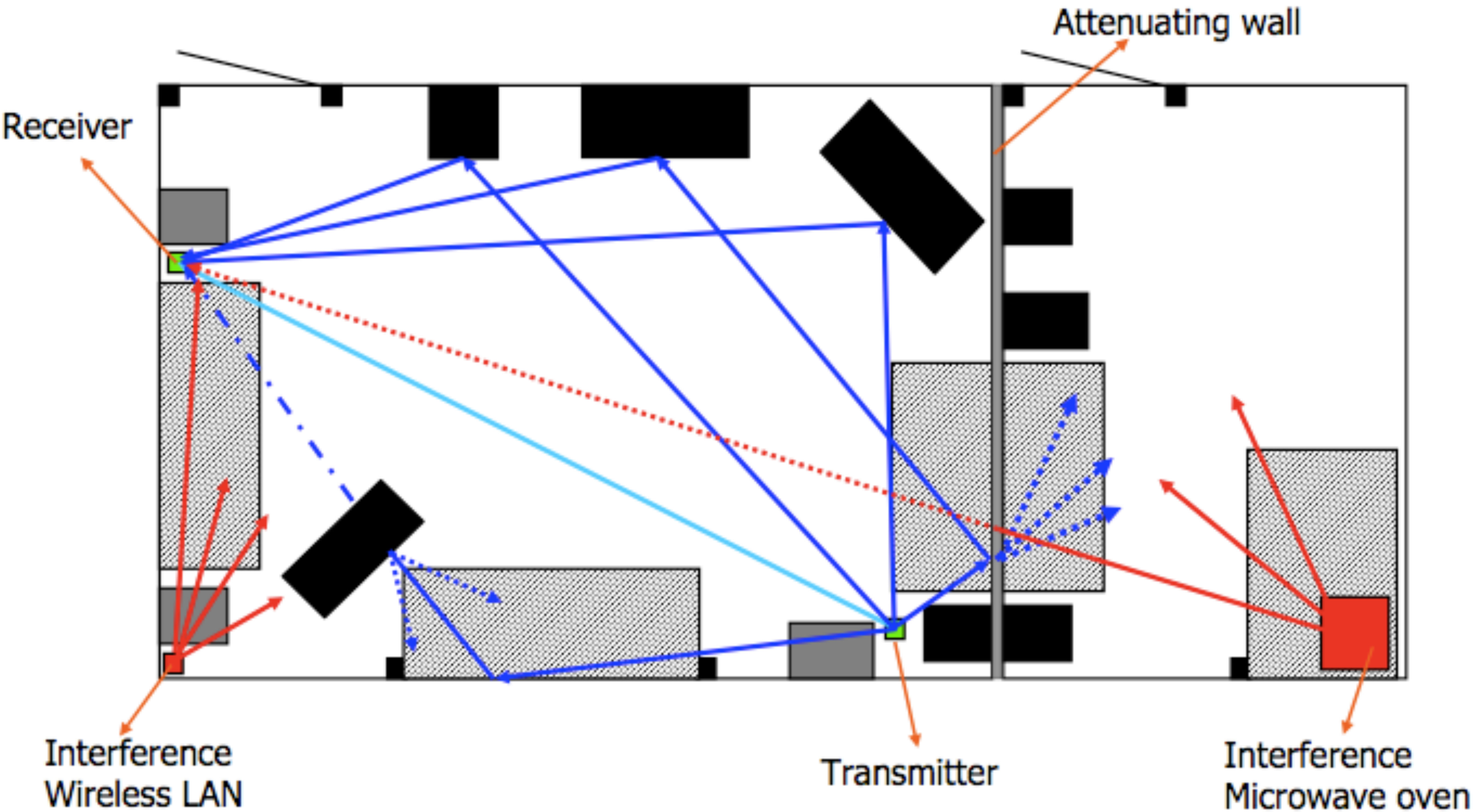
Tools:

Analysis
Algebra
Physics
Probability theory

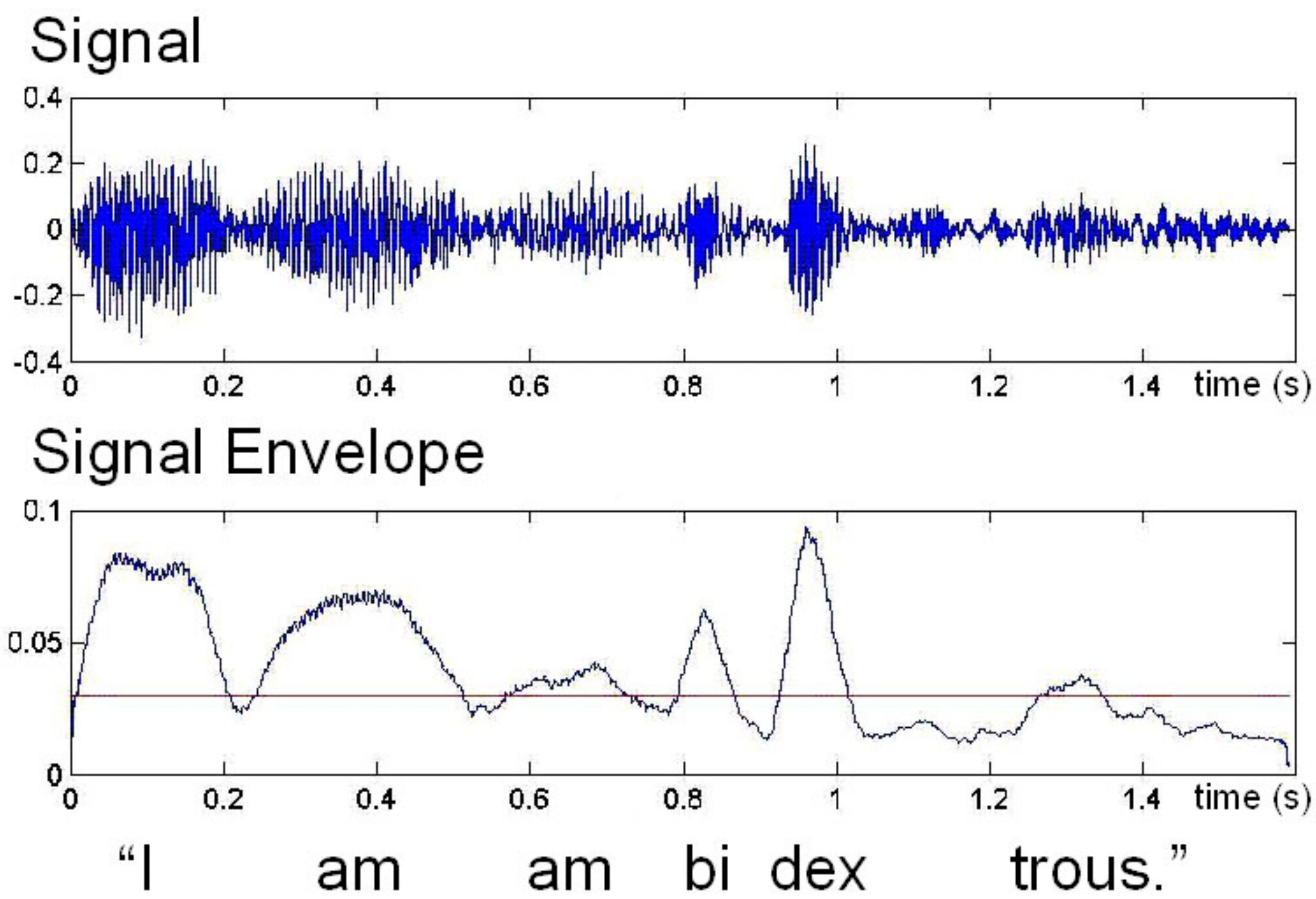
Solution

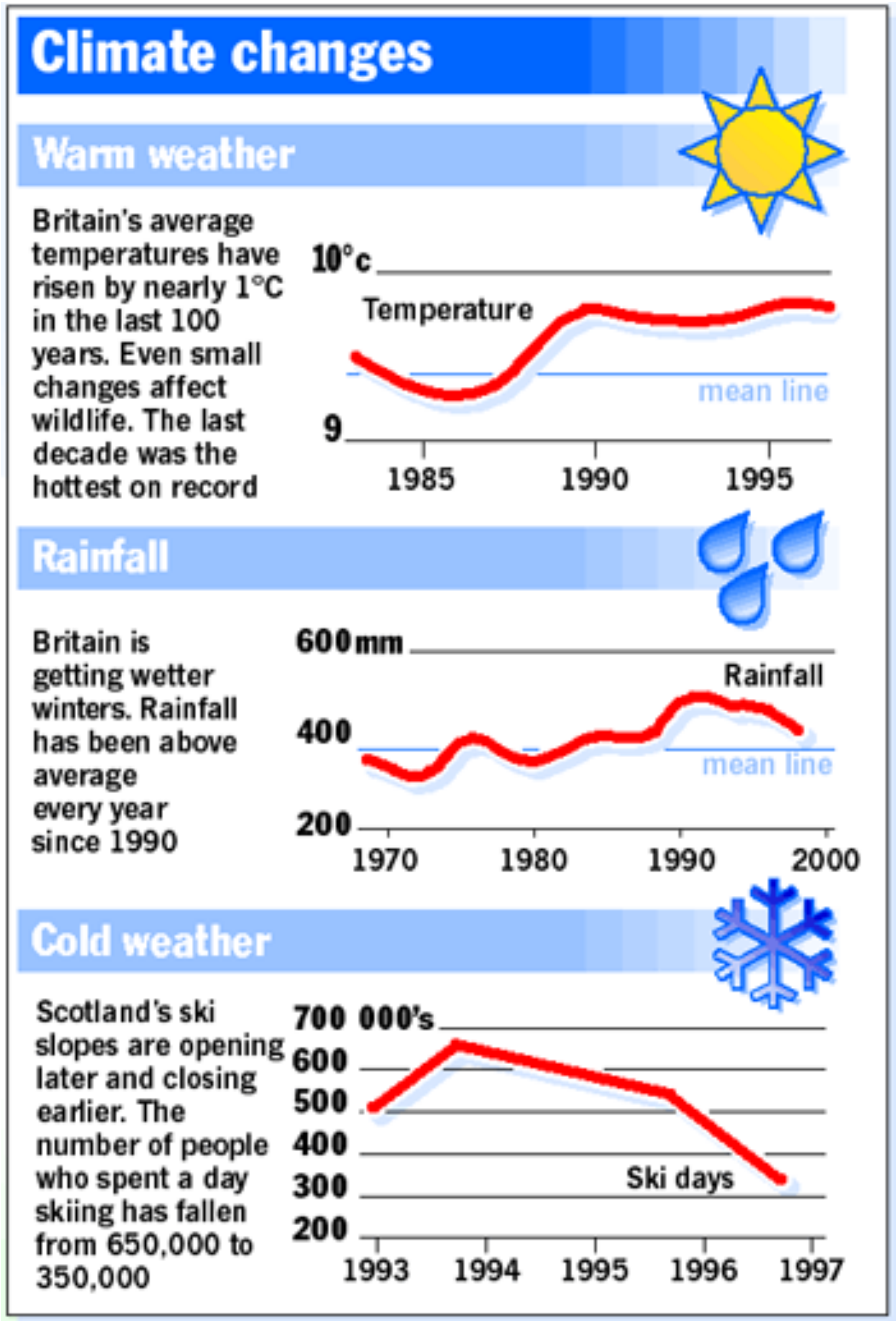
Model

Radio signals in an indoor environment

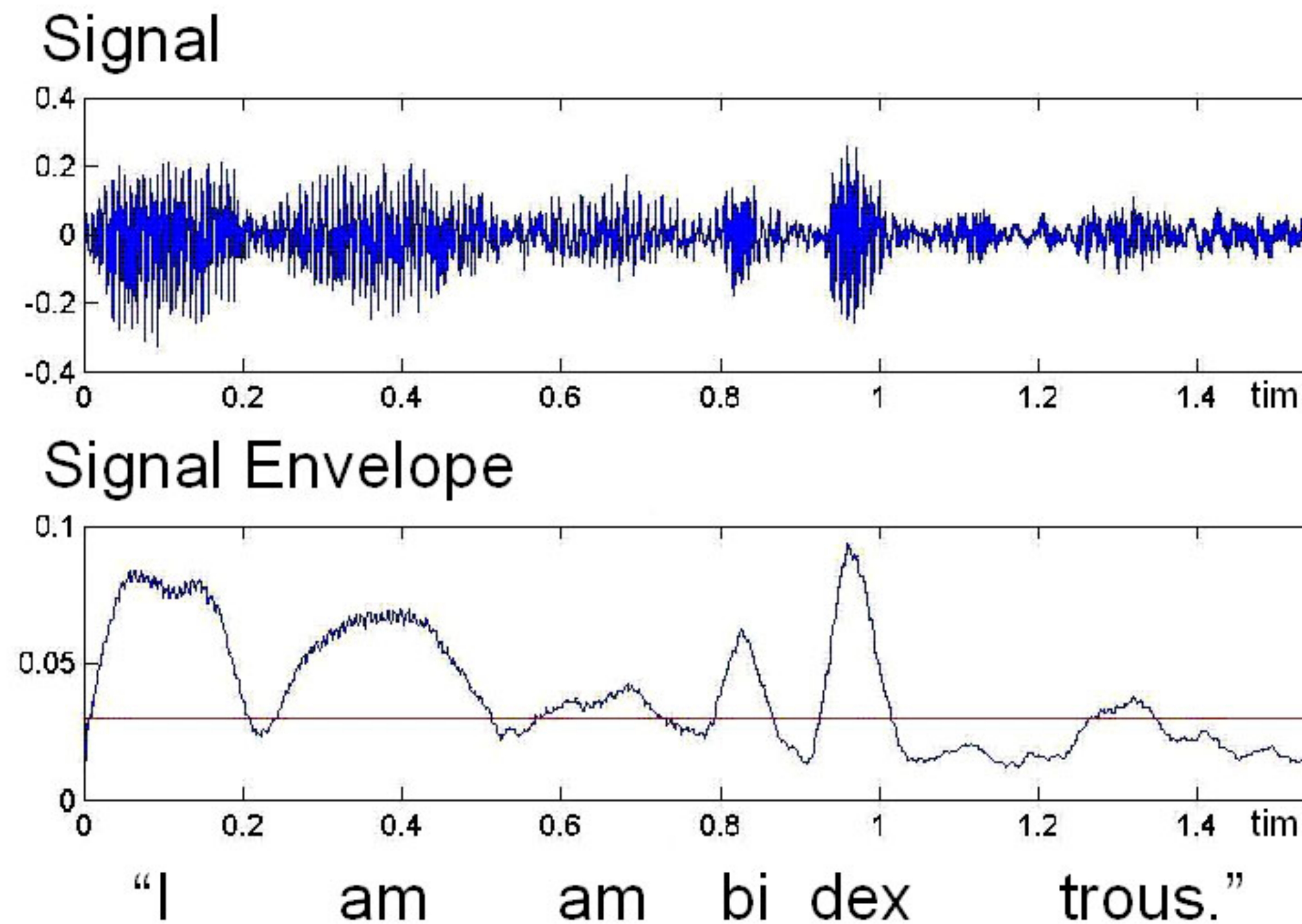


Voice signal, weather, stock market





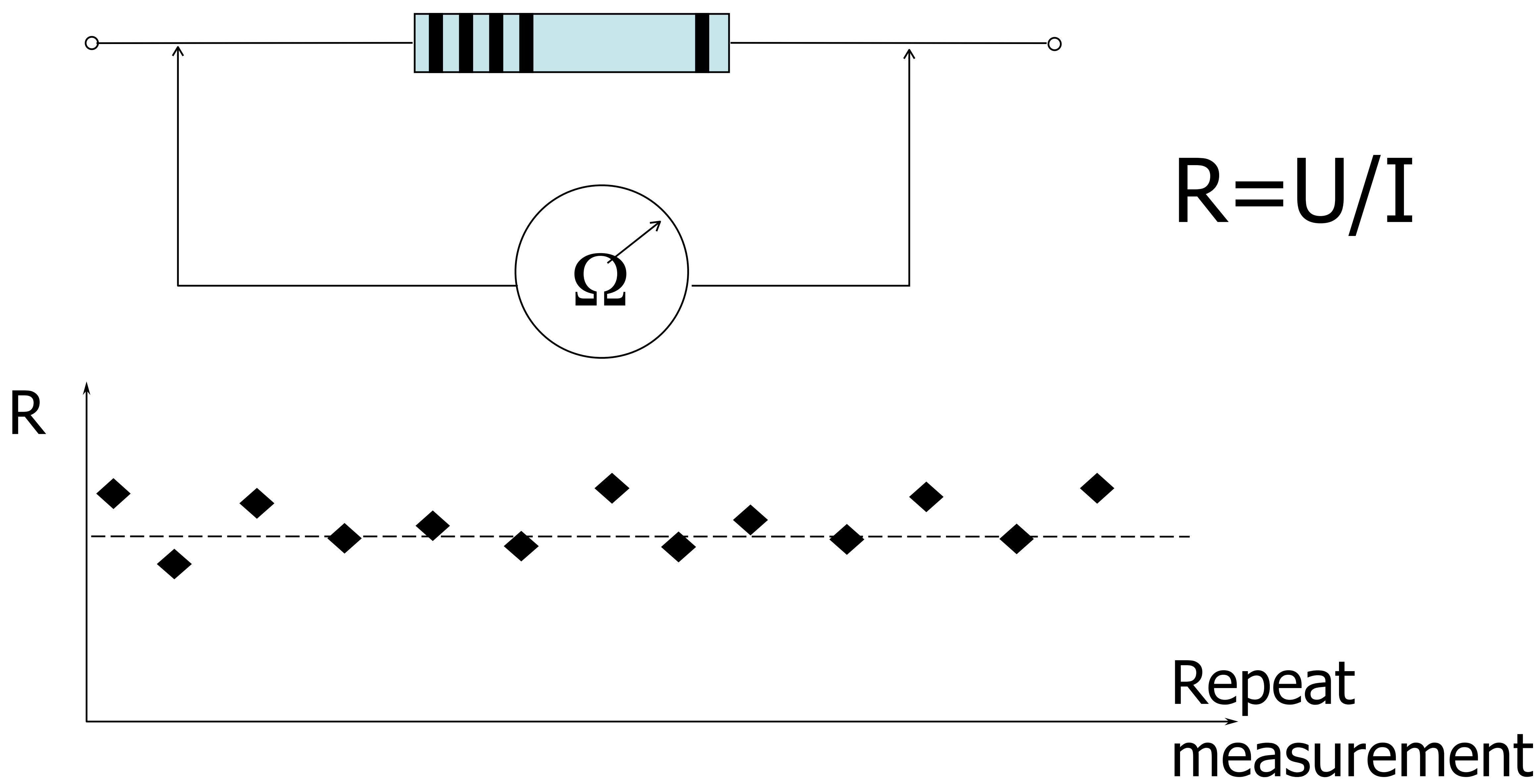
Stochastic processes



- The goal of this course:
  describe these types of signal (and get rich?!!?)

Model: Abstraction of Reality

The Resistor



$R=U/I$

Models

- Deterministic model
  - R = U / I
  - R = U / I + f(T)
  - R = U / I + f(T) + g(measuring equipment)
  - ...
  - increasing complexity to describe behaviour
- Stochastic (=probability) model
  - R = U / I + N
  - N is unpredictable (random, stochastic) component ~ Obtain a different value when measurement is repeated

Model: Abstraction of Reality

- Sometimes a stochastic (probability) model leads to a **simpler** description, analysis, and design of a problem or solution than a deterministic model

- Stochastic models enhance the toolkit of the modern engineer
- Skilful usage of probability theory can greatly reduce complexity in some case

- … but don't use a saw if you need a screwdriver

Applications in CS

But applicability is much much broader

- Computer and network performance analysis
- Prediction of internet traffic
- Security and reliability analysis

- Prediction of energy consumption
- Decision support systems

- Design of digital filters for speech, audio, images, communication signal processing
- Multimedia processing
- Data analysis, data mining
- Character and pattern recognition

Question

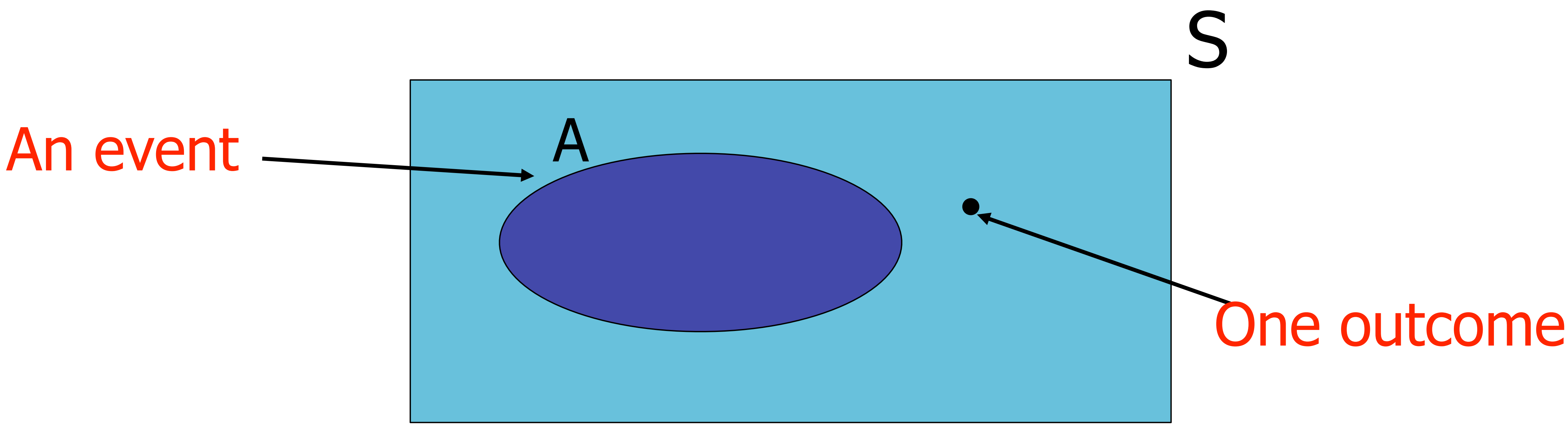
- Can you come up with something that is **not** a stochastic process?

Probability Model

- What does a model describe?
- What do we actually mean by “probability”?

Probability (Stochastic) Model

- Describe the situation (reality, experiment)
- Identify the values that can be observed
  - Outcomes
  - Events (sets of outcomes)



- Assign probabilities to events $P(A)$, $P(S)$

Notion of Probability

- How to assign probabilities to outcomes or events in a stochastic model?
- Intuitively: How frequent does an outcome/event occur when the experiment is repeated many times?

Number of times

$$P_k = \lim_{n \to \infty} f_k(n) = \lim_{n \to \infty} \frac{N_k(n)}{n}$$

Outcome number

Axiomatic Approach to Probability

- Axiomatic: Pose (three) basic assumptions (axioms) and build the theory on top of that.

1. A probability is never negative

   $P(A) \geq 0$ for any A

2. The probability of the sample space is 1

   $$P(S) = 1$$

3. The probability of events A and B that are mutually exclusive (disjoint) can be added:

$$P(A \cup B) = P(A \,\mathrm{or}\, B) = P(A) + P(B)$$

Some consequences: See Section 1.4

Mutually exclusive/disjoint

- Take the standard, silly, example: throwing with a dice

- (1) what are the possible outcomes?

Mutually exclusive/disjoint

- Take the standard, silly, example: throwing with a dice
- (1) what are the possible outcomes?
- (2) are the outcomes disjoint?

Mutually exclusive/disjoint

- Take the standard, silly, example: throwing with a dice
- (1) what are the possible outcomes?
- (2) are the outcomes disjoint?
- (3) are events A={1,3,5} and B={3,4,5} disjoint?

Examples of Probability Models

- Bernoulli probability model
  - Probability of success in binary situation
- Geometric probability model
  - Probability of n repeats of Bernoulli experiment needed before success
- Binomial probability model
  - Probability of k successes in n repeats of Bernoulli experiment
- Poisson probability model
  - Probability of k events in certain period of time

Have a look in the book!

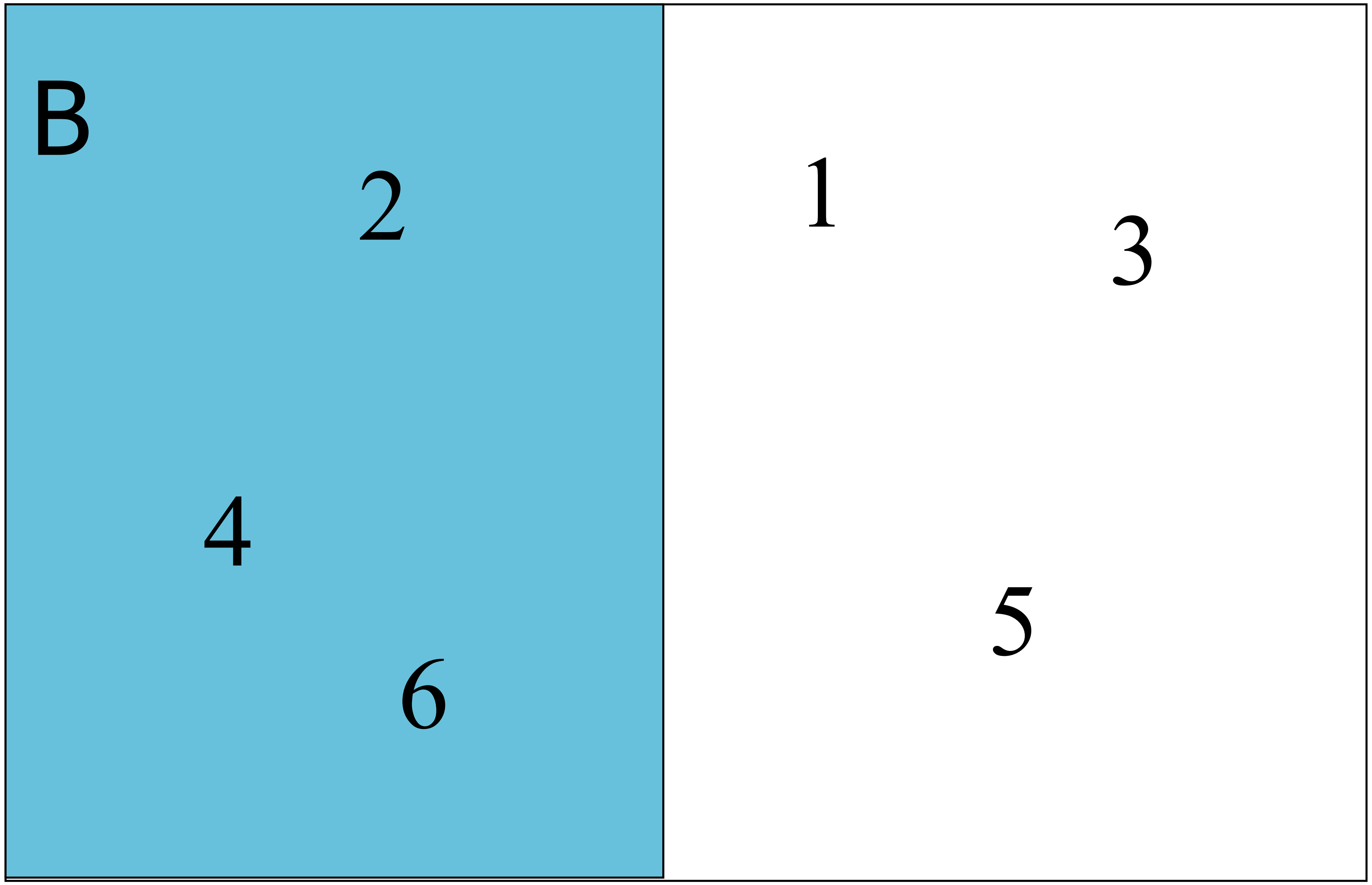
Conditional Probability

- Notation P(A / B) or P(A | B)
- Interpretation: The probability of A, given that we know that the outcome is in event B
- Probability of A "given B"
- $$P(A|B) \overset{?}{\substack{< \\ = \\ >}} P(A)$$

Conditional probability

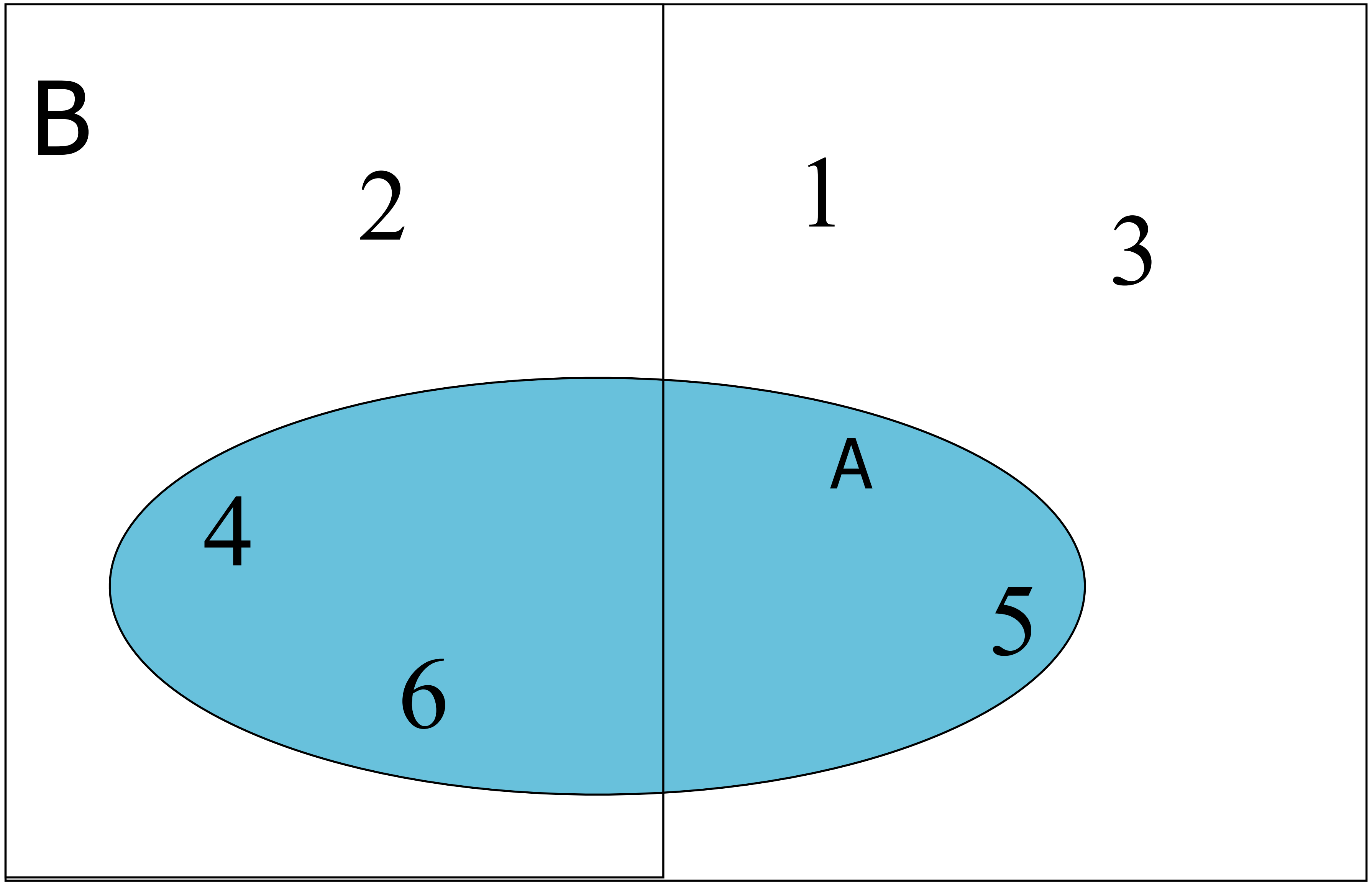
A priori probability

Example (1)



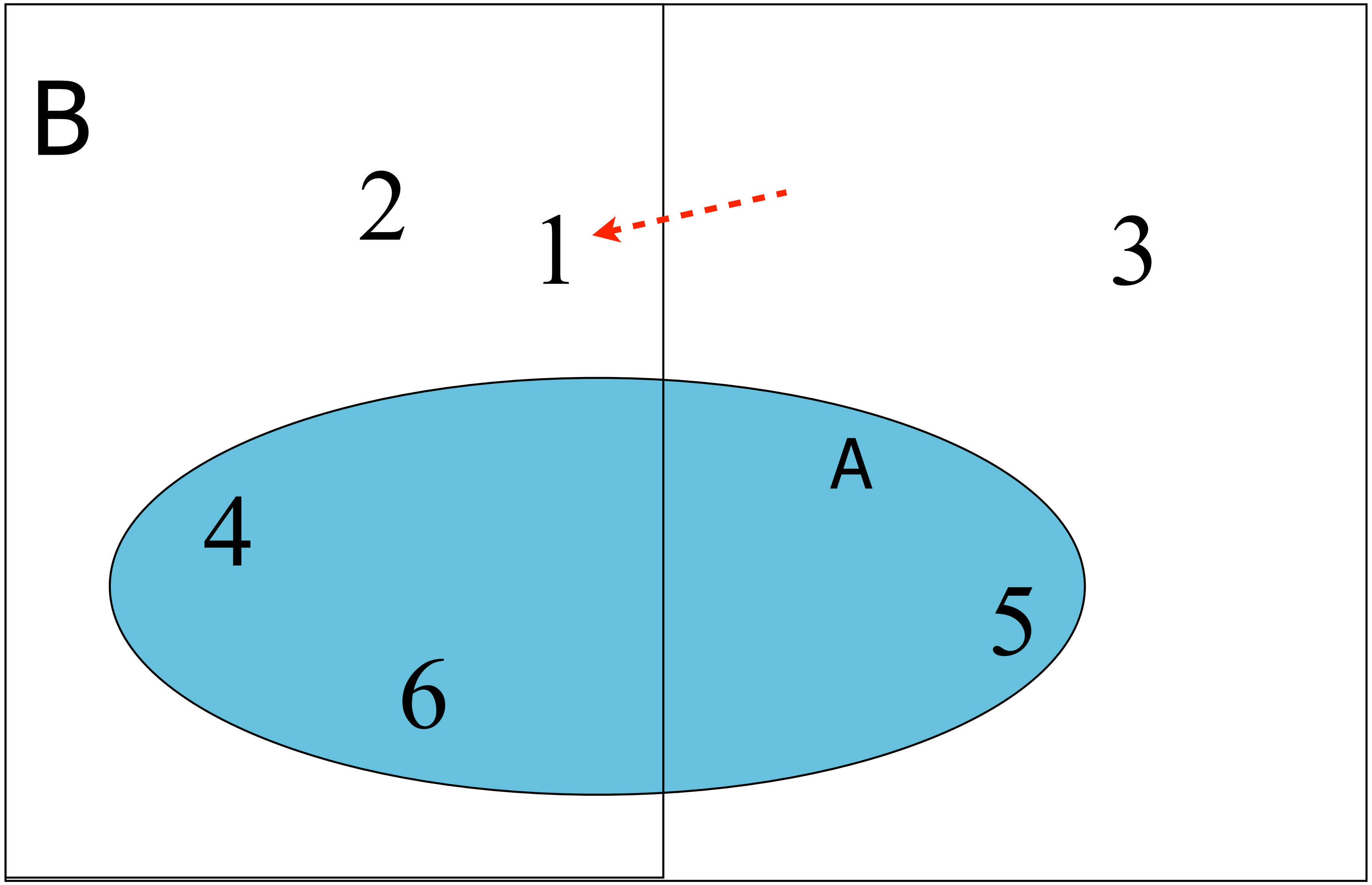
- B: “Even outcome” when rolling the dice

Example (2)



- A: “4 or more” when rolling the dice: what is P(A)?
- How large is P(A|B)?

Example (3)



- Different B! How large is P(A|B) now?

Definition Conditional Probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad P(B) > 0$$

Notation:

$$P(A \,\mathrm{and}\, B) = P(A, B) = P(A \cap B) = P(AB)$$

Alternative formulation

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

Bayes' theorem

- A consequence of the definition of conditional probabilities: Bayes' theorem
- Due to time limitations: study it yourself, section 1.5 pg.16-21

Independent Events

- If $P(A|B) = P(A)$ then A and B are independent events

- Hence if A and B are independent events, then

$$P(A, B) = P(A|B)P(B) = P(A)P(B)$$

- Independence is a special case and can never be assumed to by true by default
- Be careful: Independence and mutually exclusive are different concepts
- Example: A={2,4,6}   B={1,3,5}.

Random Variables

Outcomes, events
Probability mass functions
Probability density functions
Cumulative distribution functions
Expected value
Variance

Definition of Random Variable (1)

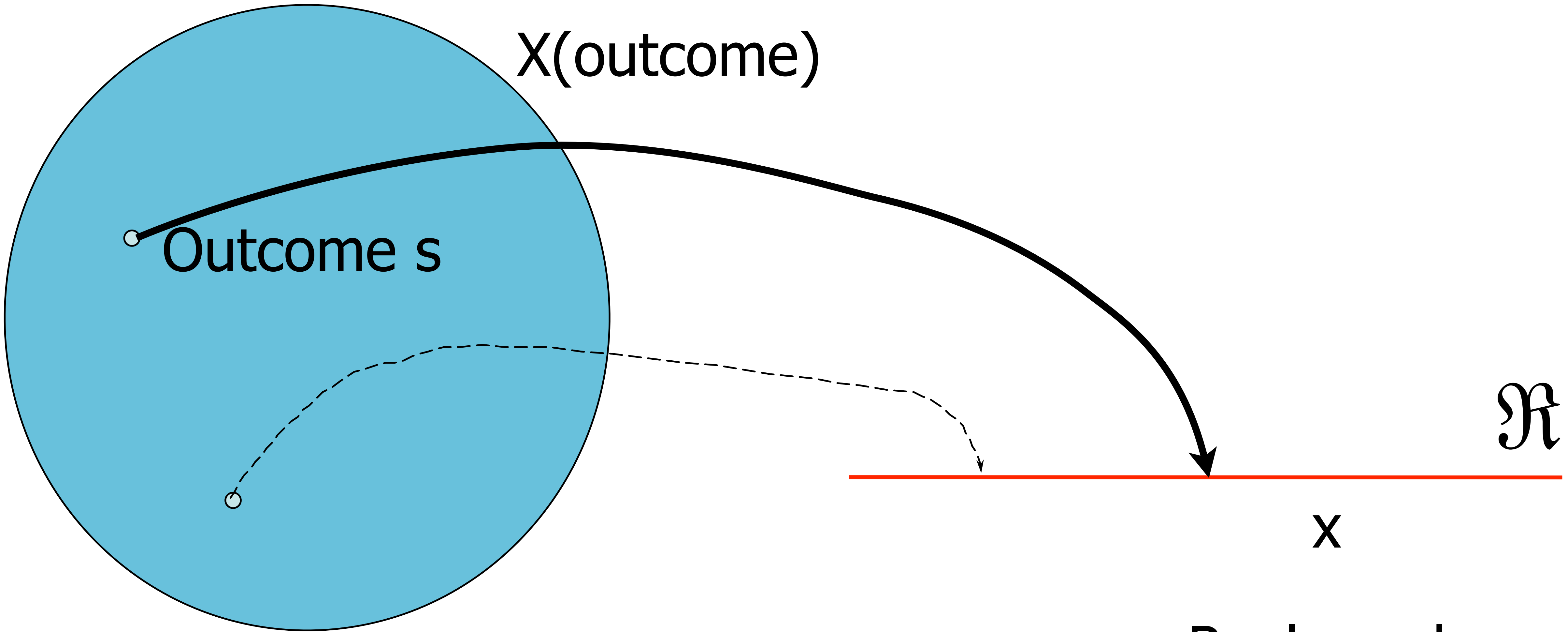
- Experiments often take place in a physical world
  - We observe physical quantities

- Probability theory works in a mathematical world, with mathematical tools.

- How to map “physical observations” to numbers we can do mathematics on?
  - E.g. Flip a coin: Outcomes are head/tail
  - E.g. Observe flooding because of frequent rainfall. Outcomes are yes/no

Definition of Random Variable (2)

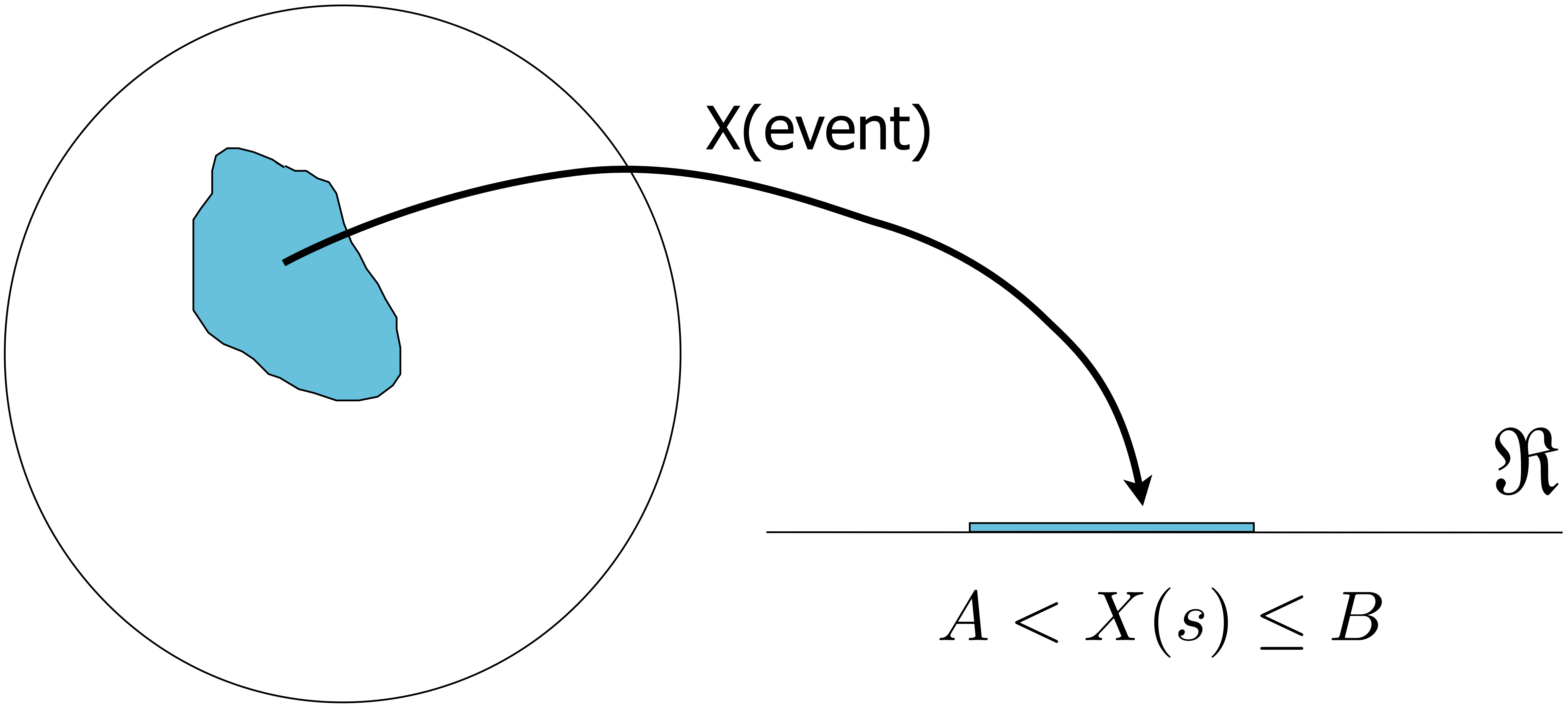
$$x = X(\text{outcome}) = X(s)$$



Arbitrary sample space
Often related to physical world

Real numbers

Definition of Random Variable (3)



- draw red ball; roll even; grown up
- $\{1\}$ $\{2, 4, 6\}$ $[18, \infty\rangle$

Notational Convention

$$a < X(s) < b$$

- is the same as

$$a < X < b$$

- But with X (upper case) we mean a function of the outcomes of the experiment
- The outcomes themselves are denoted by x (lower case), so we get:

$$X = x$$

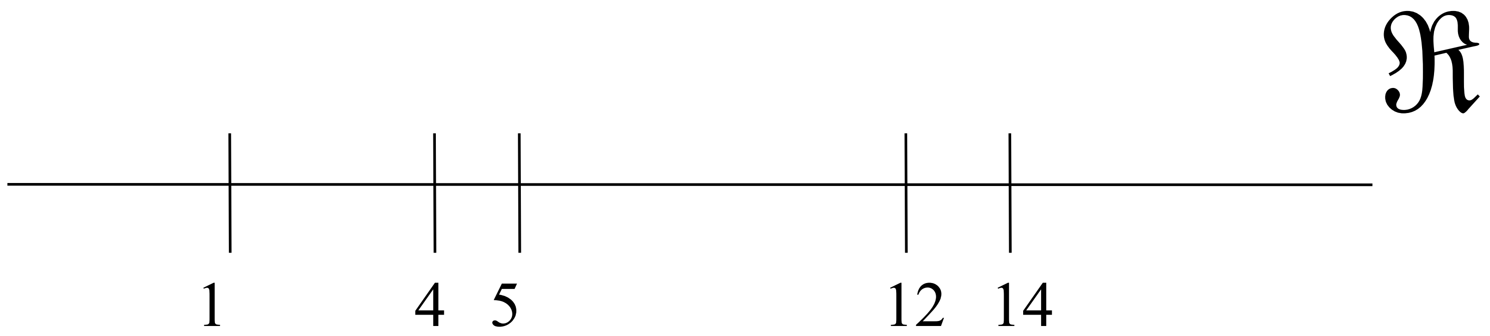
Why Numerical Function X?

- Many engineering problems consider numerical problems
- Numbers in $\Re$ can be ranked
- With numerical functions one can carry out calculations, such as “averages”
- Model behaviour of the abstract random variables, not of particular experiments

- Consequence: the values that X takes on, are now random

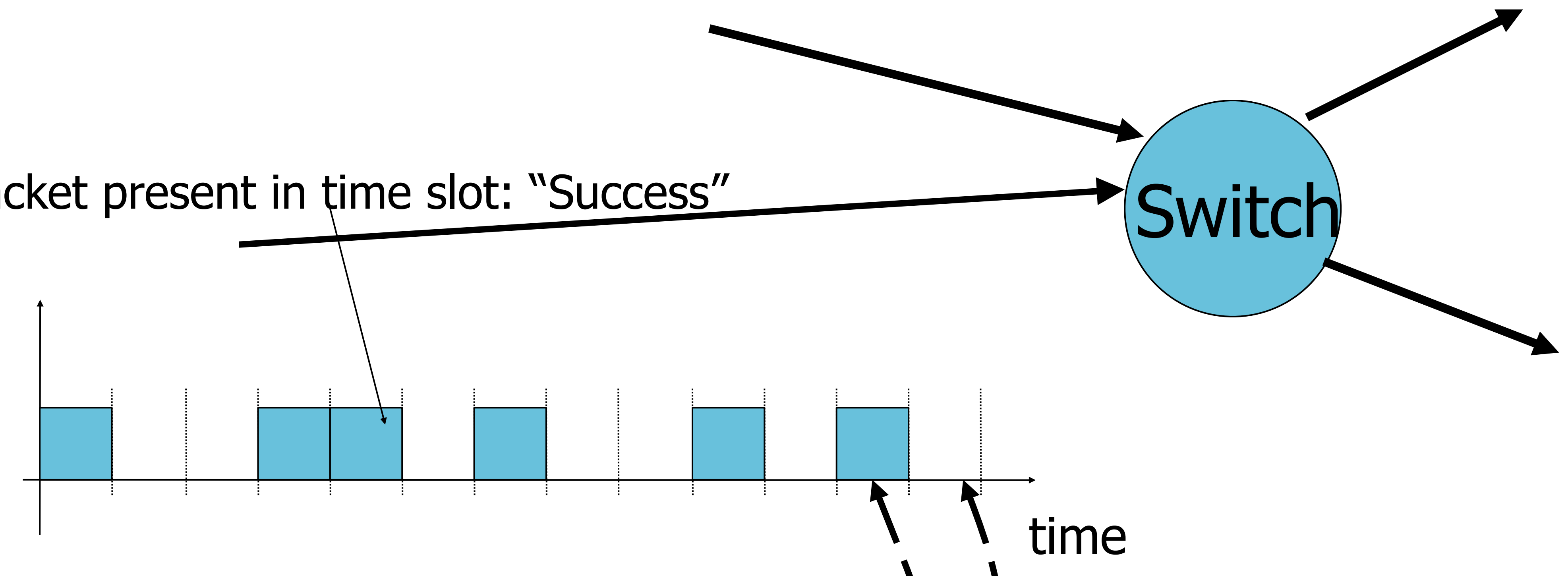
Definition Discrete RV

- If an experiment has a countable number of outcomes, the random variable is discrete
- Example:



- Other example:
  - Number of active speakers X
  - $X \in \{0,1,2,3,4,5\}$

Probability Mass Function

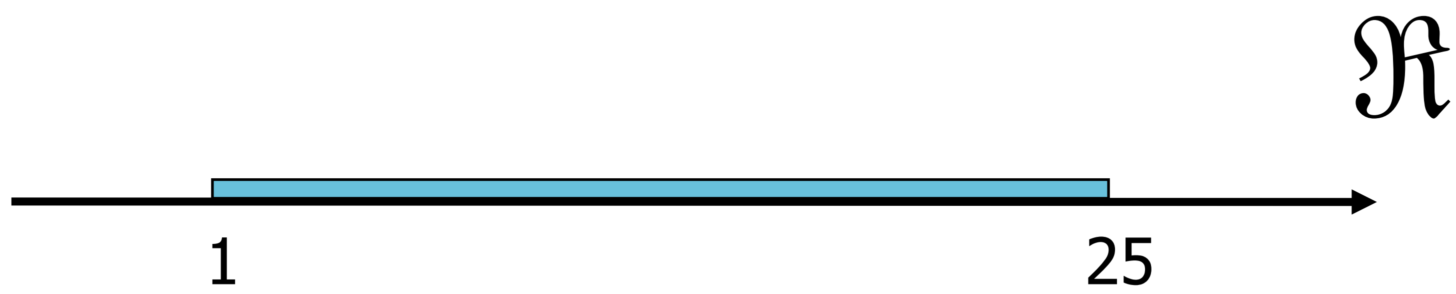


$$P_X(x) = \begin{cases} 1-p & x=0 \\ p & x=1 \\ 0 & \text{otherwise} \end{cases}$$

- $P_X(x)$ is called the **Probability Mass Function** (PMF)

Continuous Random Variables

- Have as range the entire Real Axis
- Events are always intervals on the real axis
  - $A = \{x | 1 < X < 25\}$



- Theory is comparable to discrete random variables, with one additional concept namely PDF (probability density function)
- Problems with discrete RV's are more insightful
- … but calculations with continuous RV's are usually a little easier in practical cases of interest.

TU Delft

What is Different …

- Discrete sample space
  - Countable number of outcomes
  - Each outcome has a non-zero probability of occurrence

- But now … continuous sample space
  - Have an infinite number of outcomes
  - Example: Pick a real-valued number between 3 and 5.
    - How many outcomes?
    - What is probability of each outcome?

Two Consequences

- For continuous RV's it is meaningless to describe its behaviour in terms of the probability of an outcome
  - Probability mass function (PMF) is not a useful way of describing continuous random variables
  - P[X=x] = 0
  - We will consider events instead

- The probability "0" apparently does not mean that an outcome can not occur …
  - It occurs statistically too infrequent to yield a non-zero relative frequency.

Cumulative Distribution Function

- We consider a particular event in $\Re$ namely
  - event $X \leq x$

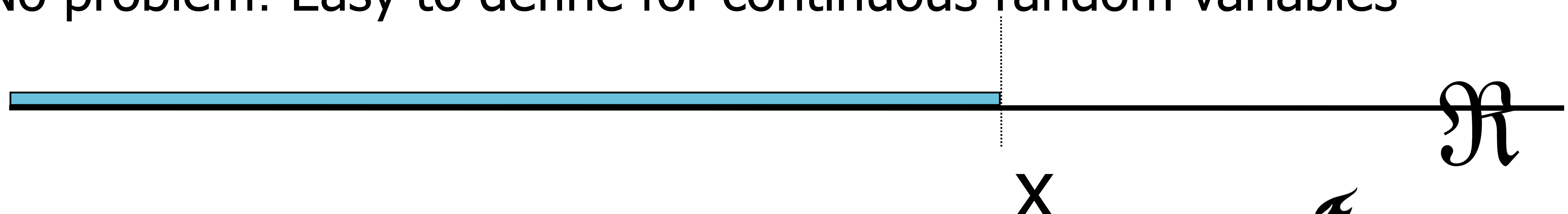
upper value of X

name of random variable

- $F_X(x) = P[X \leq x]$ for **ALL** x

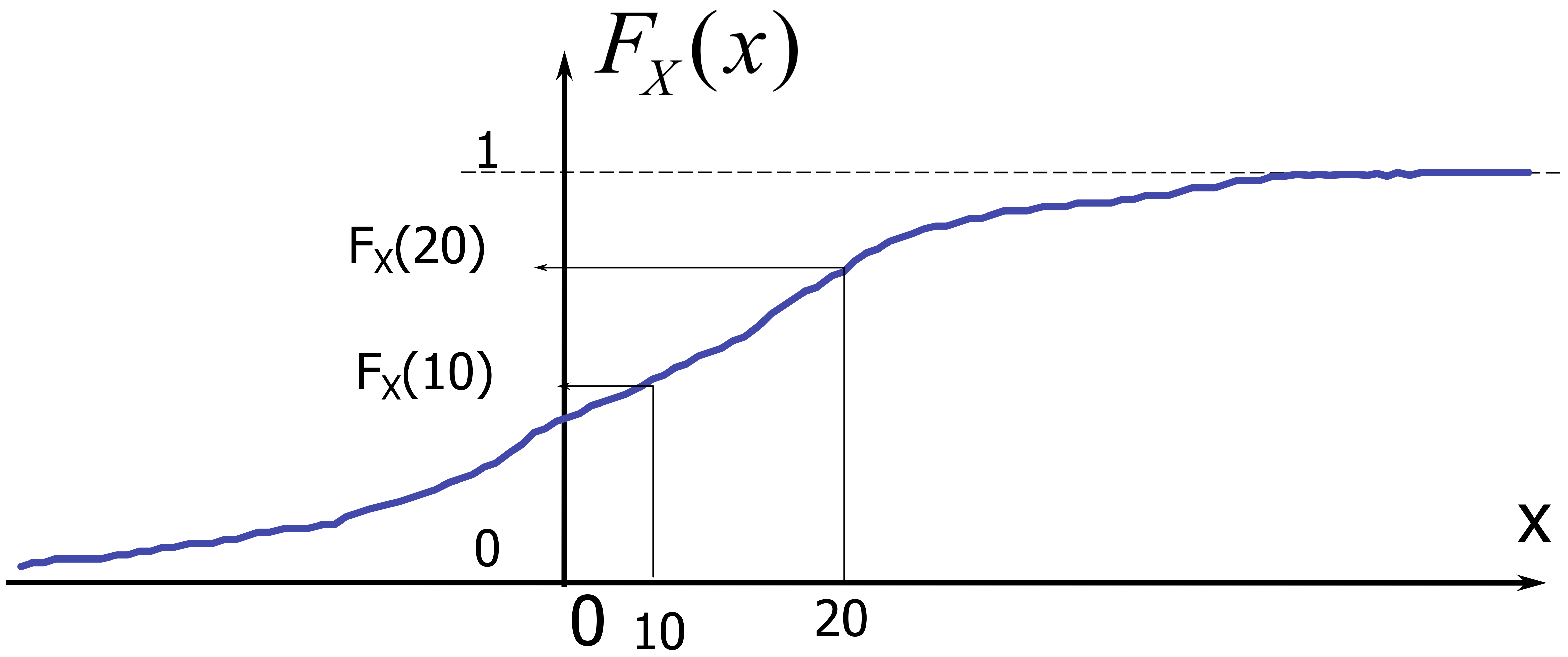
is called the **cumulative distribution function** (CDF)

- No problem! Easy to define for continuous random variables



Example

$$F_X(x) = P[X \leq x]$$



P (10 < X ≤ 20) = F_X (20) – F_X (10)

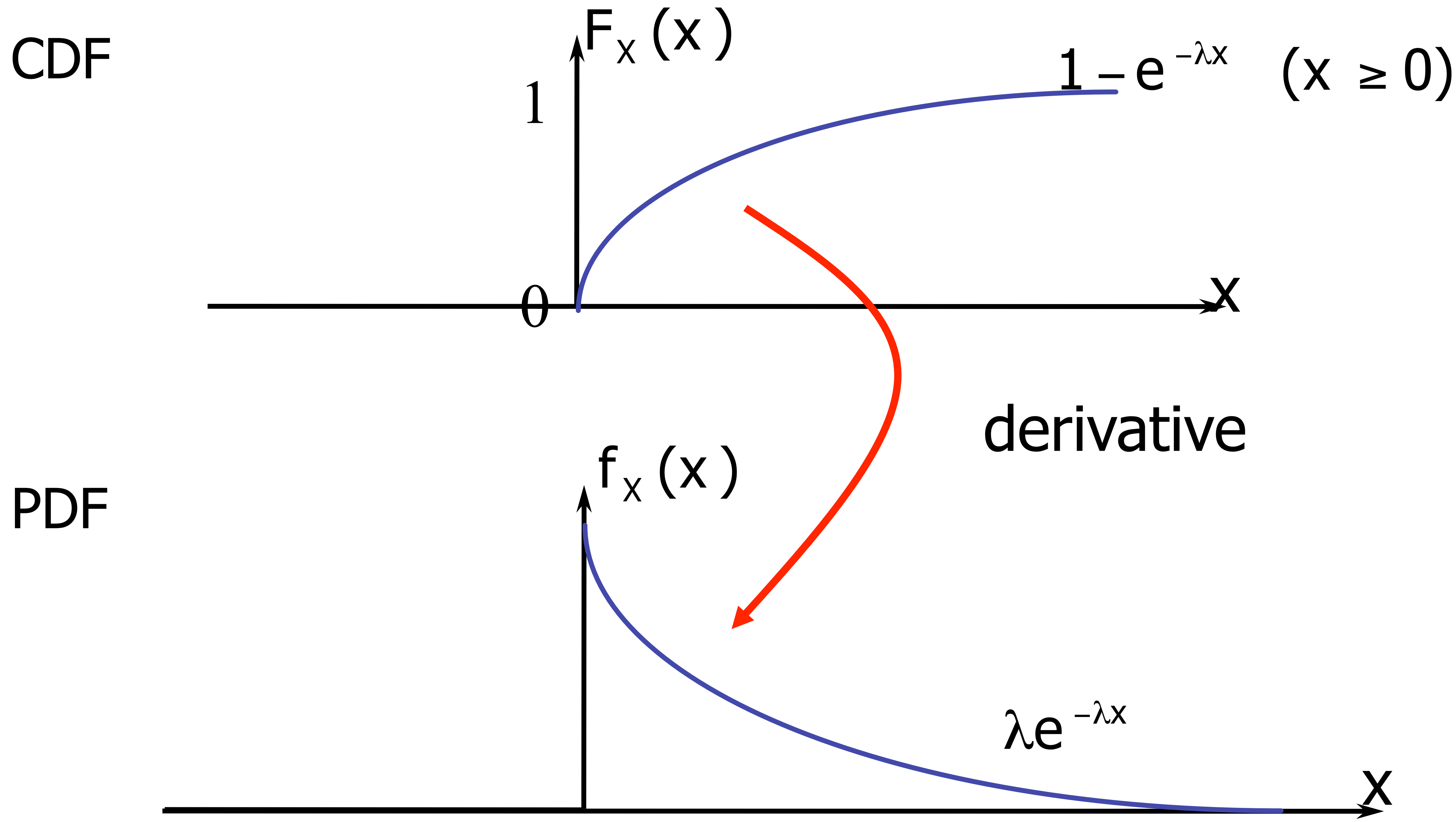
Probability Density Function

- Instead of a PMF (for discrete RVs), we have a probability density function, or pdf, for continuous RVs

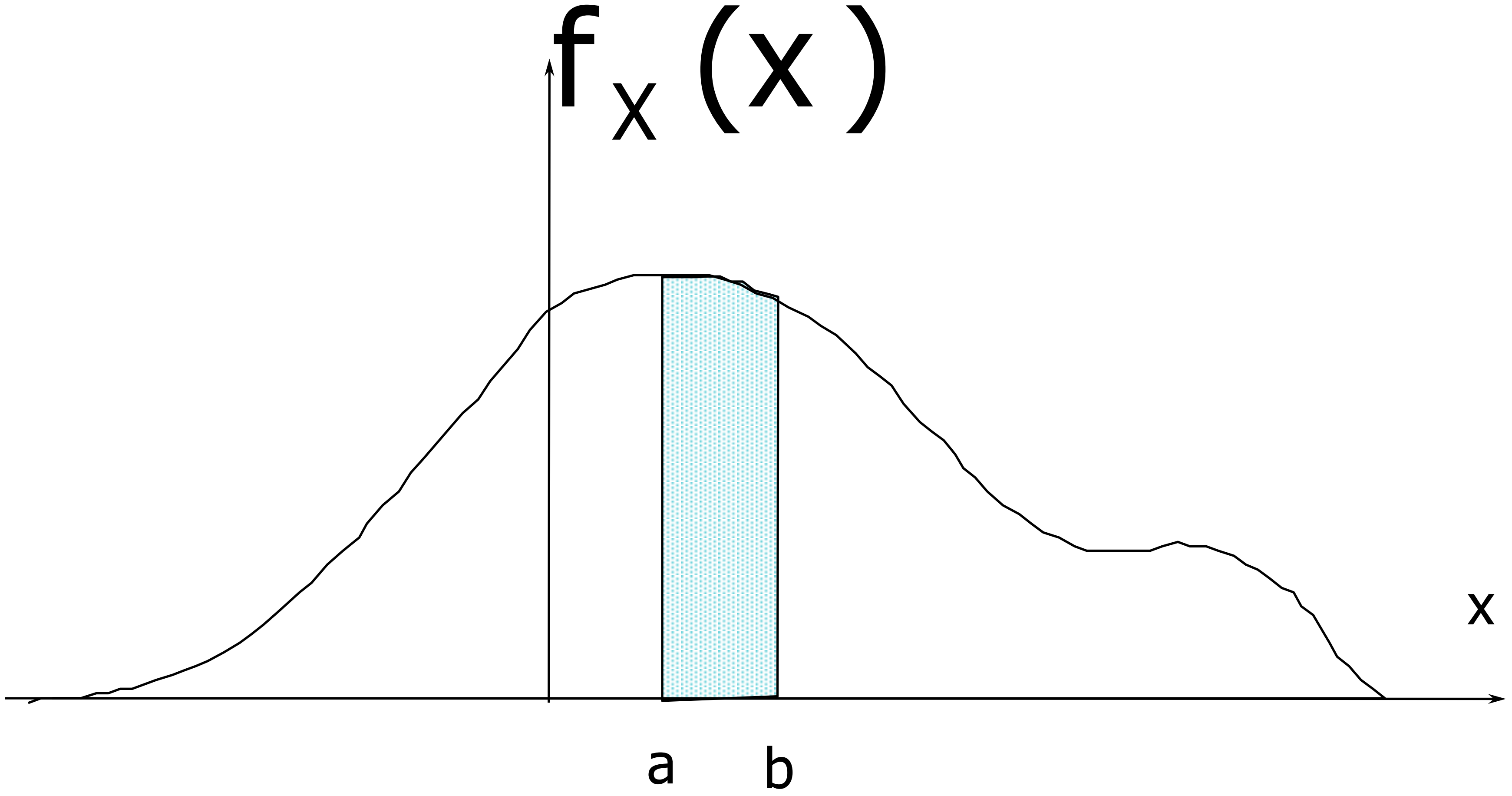
$$f_X(x) = \frac{dF_X(x)}{dx}$$

- A probability density is NOT a probability
  It can even be larger than one!

Example


CDF
$F_X(x)$
1
$1-e^{-\lambda x}$ $(x \geq 0)$
0
x
derivative
PDF
$f_X(x)$
$\lambda e^{-\lambda x}$
x

To get a probability…



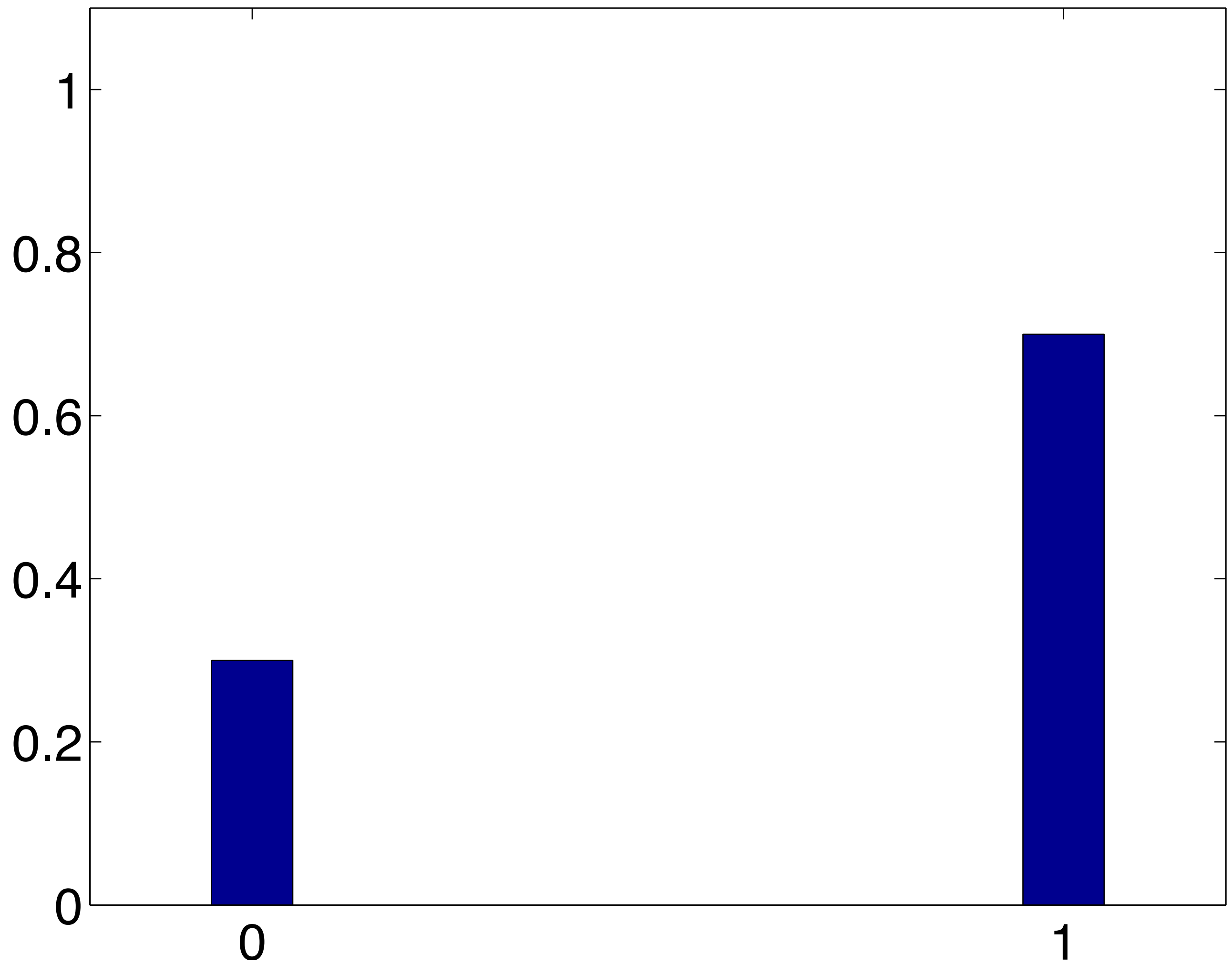
$$P[a < X < b] = \int_a^b f_X(x)dx$$

- You need to integrate the pdf to get to the probability

The Bernoulli distribution

- The Bernoulli RV is a discrete RV:

$$P_X(x) = \begin{cases} 1-p & x=0 \\ p & x=1 \\ 0 & \text{otherwise} \end{cases}$$

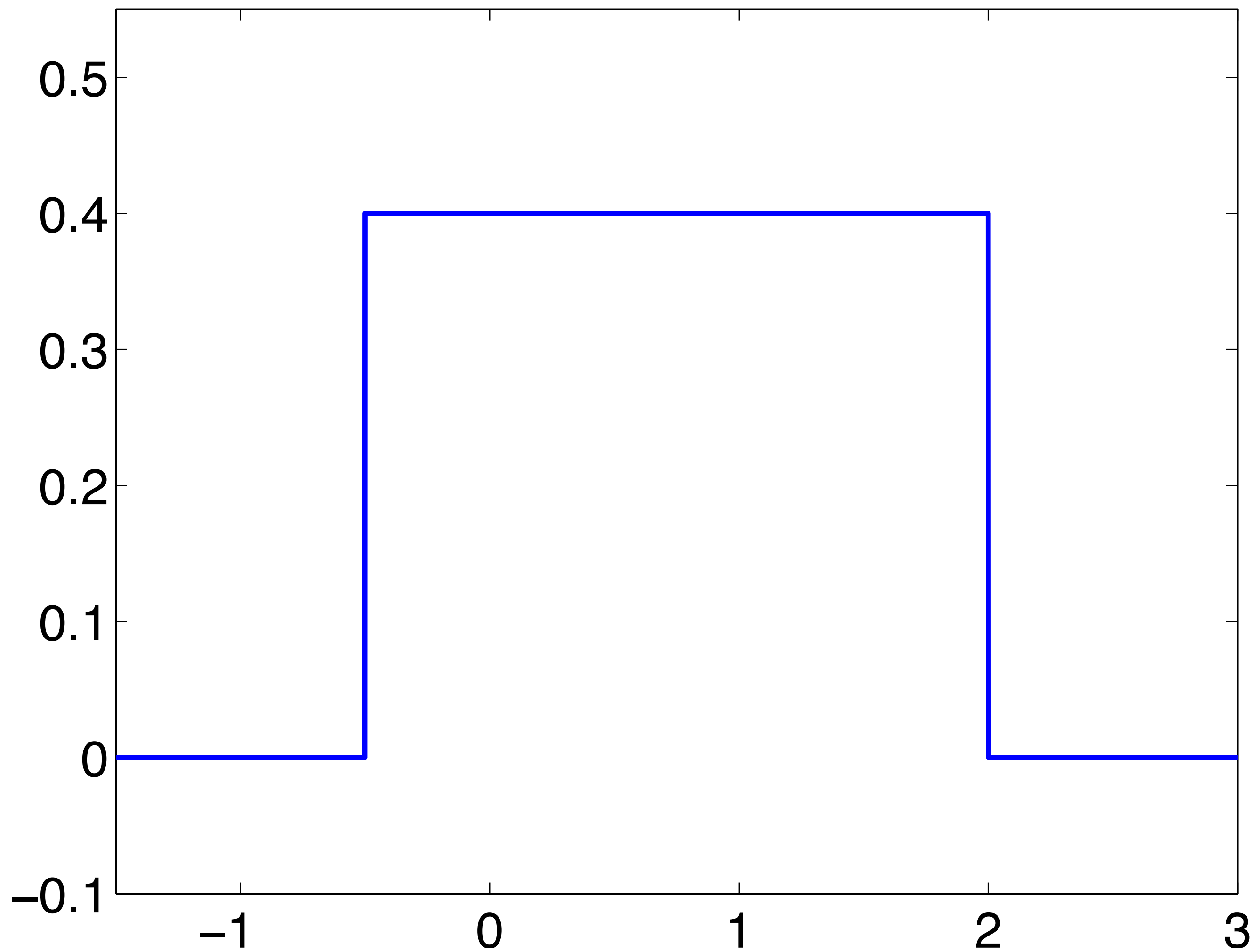


$p = ?$

The Uniform Distribution

- Uniform distribution is (almost) always **continuous**!

$$f_X(x) = \begin{cases} 1/(b-a) & a \leq x < b \\ 0 & \text{otherwise} \end{cases}$$



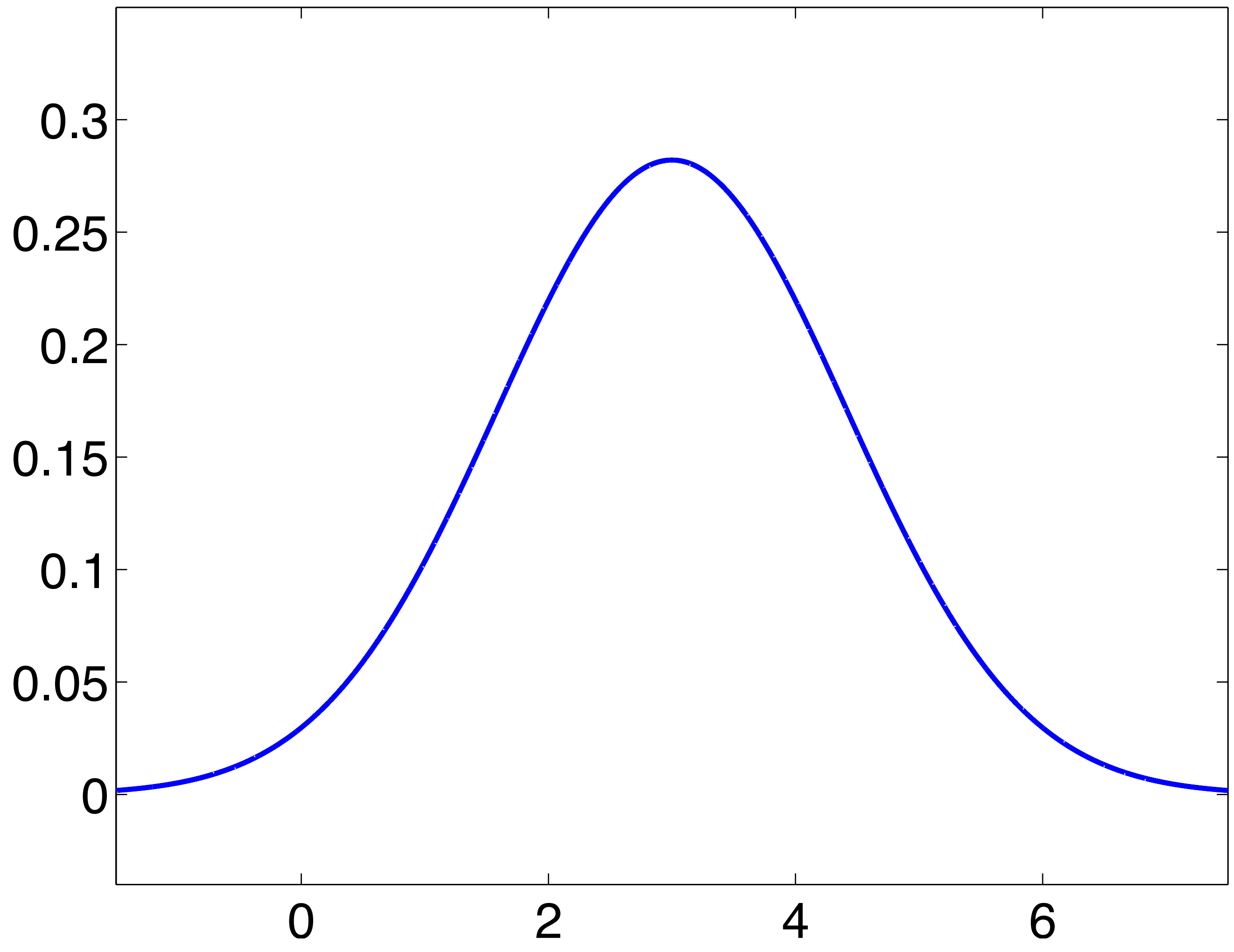
$$a = ?$$

$$b = ?$$

The Gaussian distribution

- The Gaussian distribution is the same as the normal distribution

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



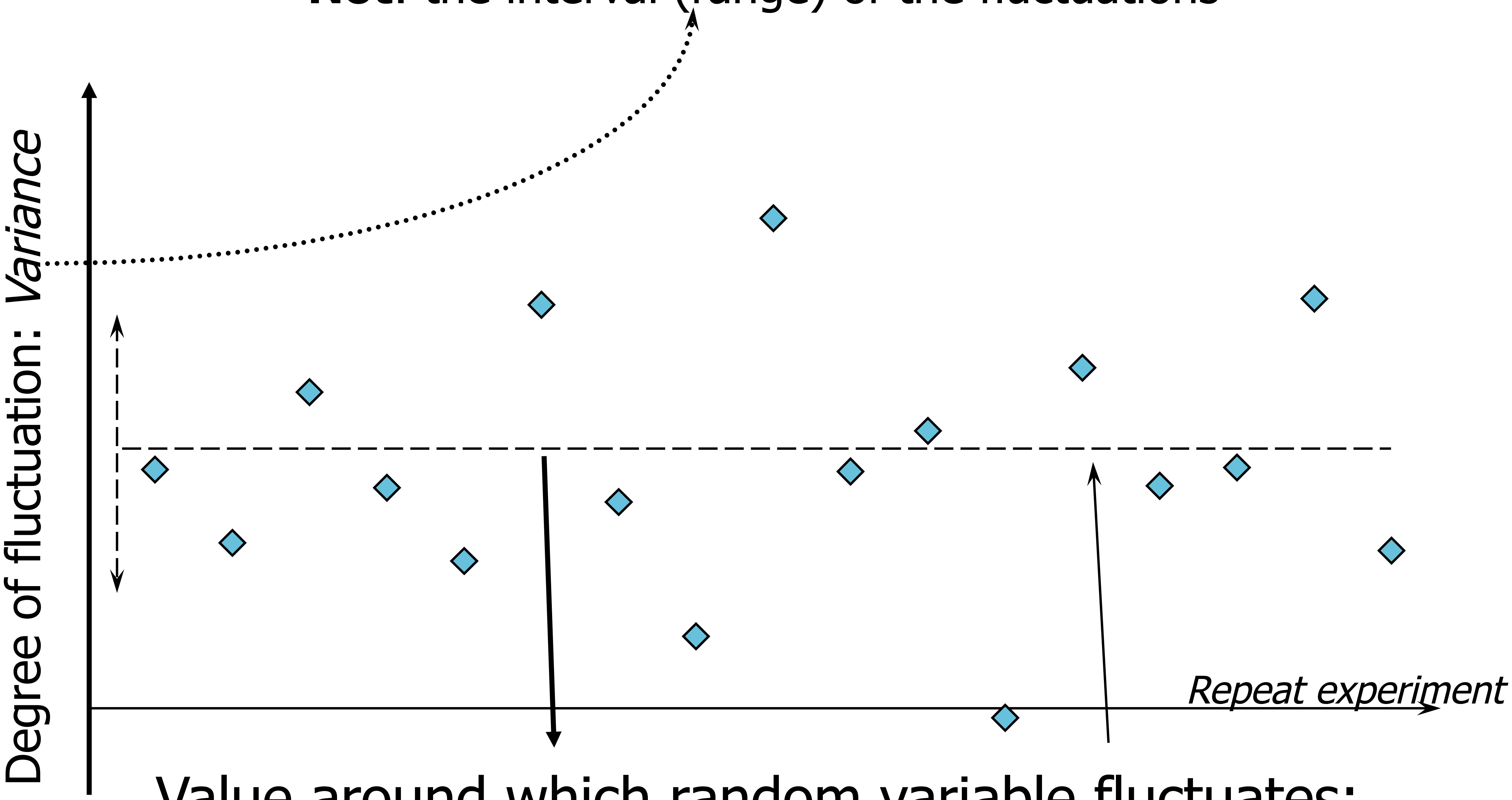
$\mu = ?$

Expected value

- The precise and total behaviour of a random variable is described by the PMF or CDF
- Sometimes we use an approximating description of its behavior because
  - this is sufficient for the application studied
  - this is the only information we can obtain in a practical application
- Two measures for average behavior:
  - expected value (expectation, mean)
  - variance (standard deviation)

Characterisation of a RV

Not: the interval (range) of the fluctuations



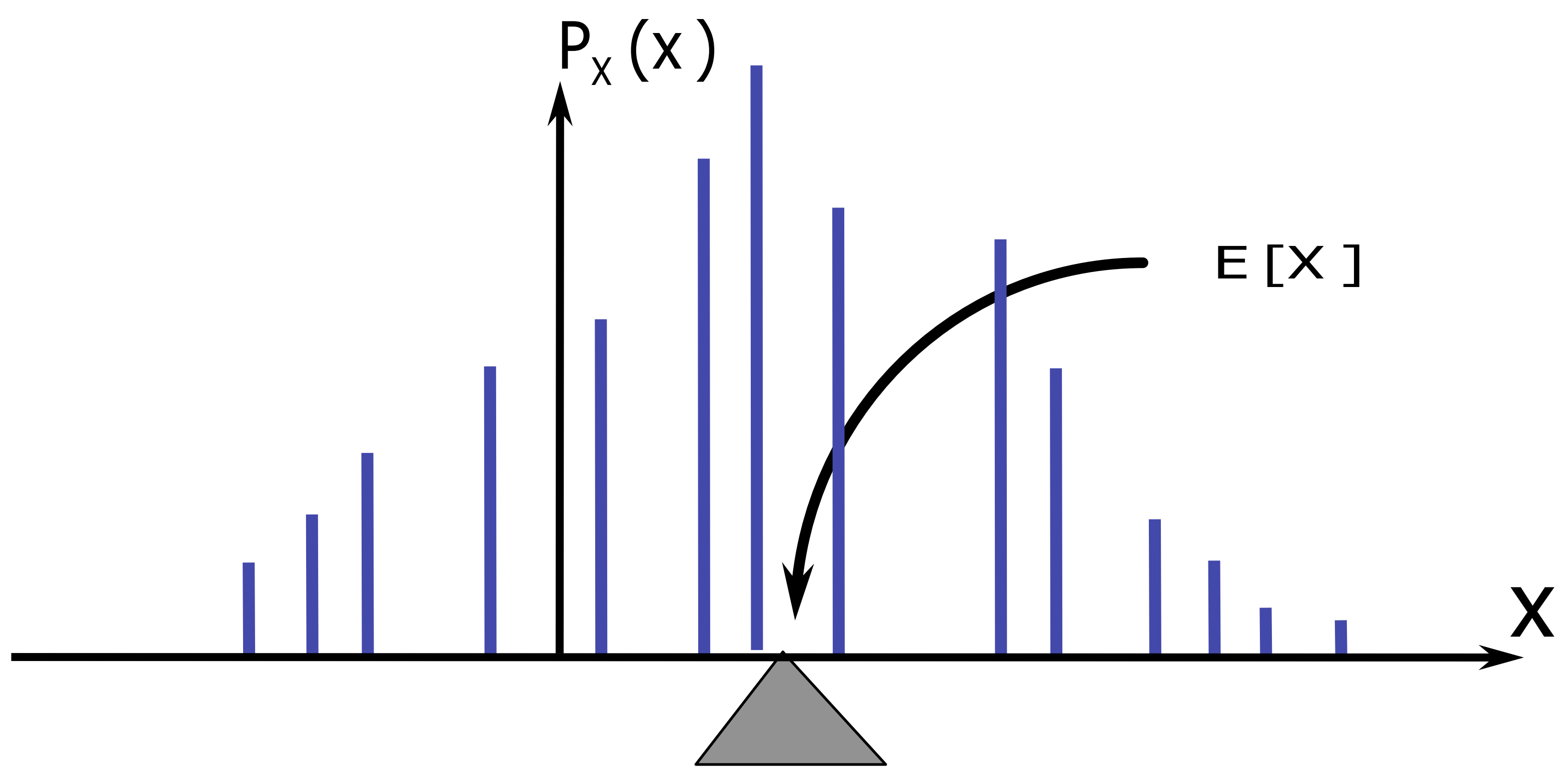
Value around which random variable fluctuates:
Expected Value

Definition of Expected Value

- Expected value is one number that characterises the PMF (and CDF)

$$E[X] = \mu_X = \sum_{\text{all } x} x P_X(x)$$

- Comparable to “Center of Mass” in physics



Properties of E[X]

- We introduce the following shorthand notation

$$E[.] = \sum_{\text{all } x} \{.\} P_X(x)$$

- Work as much as possible with E[.]
- E[.] is a linear operator:

$$E[Z] = E[X+Y] = E[X] + E[Y]$$

$$E[aX] = aE[X]$$

- Expected value scales linearly

Variance: Var[X]

- Variance is (again) one number to characterise the behaviour of the PMF

$$Var[X] = \sigma_X^2 = E\left[(\underline{X - E[X]})^2\right]$$

$$= \sum_{\text{all } x} (\underline{x - E[X]})^2 P_X(x)$$

- Measure for "width", or "dispersion" of X around expected value of PMF

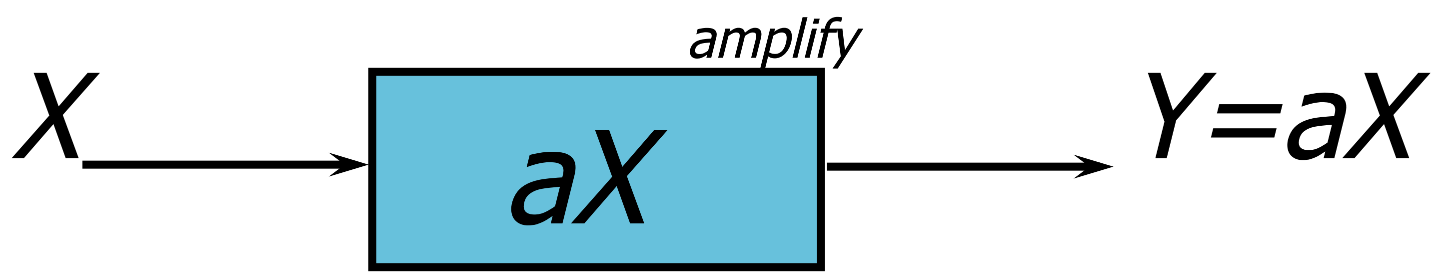
Usual Way to Calculate Var[X]

- Because E[.] is a linear operator, we get:

$$
\begin{aligned}
Var[X] &= E\left[(X - E[X])^2\right] \\
&= E[X^2] - E[X]^2
\end{aligned}
$$

- Use this form for calculations (when possible)
- Square root of Var[X] is called standard deviation σ_X

Example



$$\begin{aligned} Var[aX] &= E[(aX)^2] - E[aX]^2 \\ &= E[a^2X^2] - (aE[X])^2 \\ &= a^2E[X^2] - a^2E[X]^2 \\ &= a^2(E[X^2] - E[X]^2) = a^2Var(X) \end{aligned}$$

- Variance scales quadratically

Expected Value Continuous RV

- Same concept as for discrete RV
  - “Replace” summation by integration
  - “Replace” PMF by PDF

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

$$Var[X] = \int_{-\infty}^{\infty} (x - E[X])^2 f_X(x) dx$$

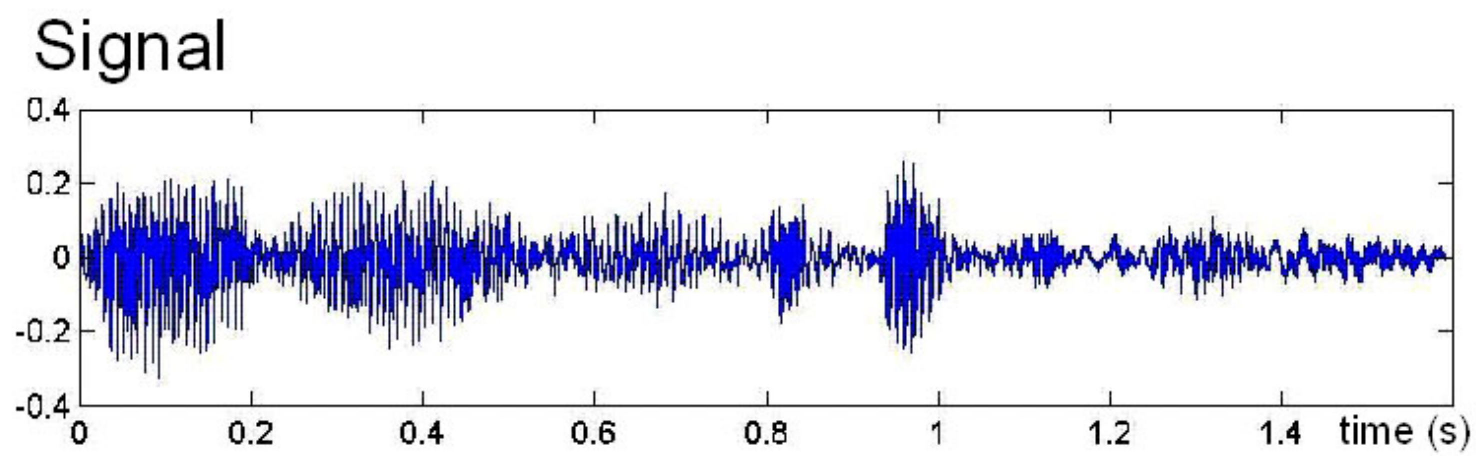
Moments

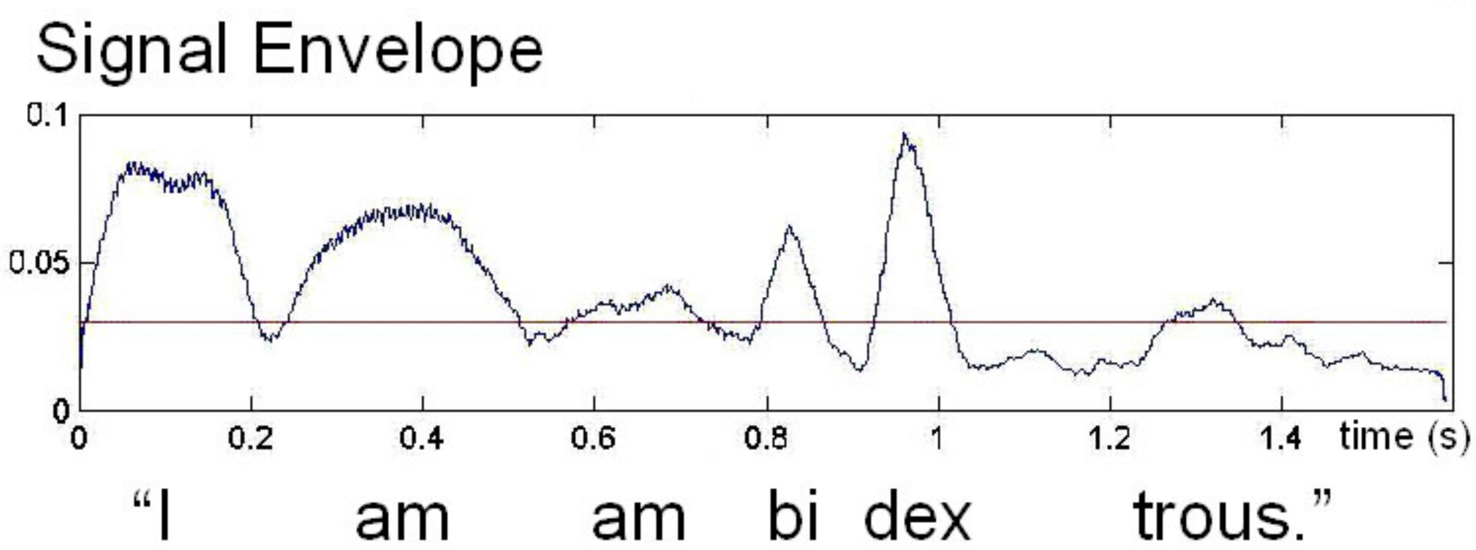
- We have seen
  - $E[X]$ Expected value or first moment
  - $E[X^2]$ Second moment
  - $E[(X-E[X])^2]$ Second central moment

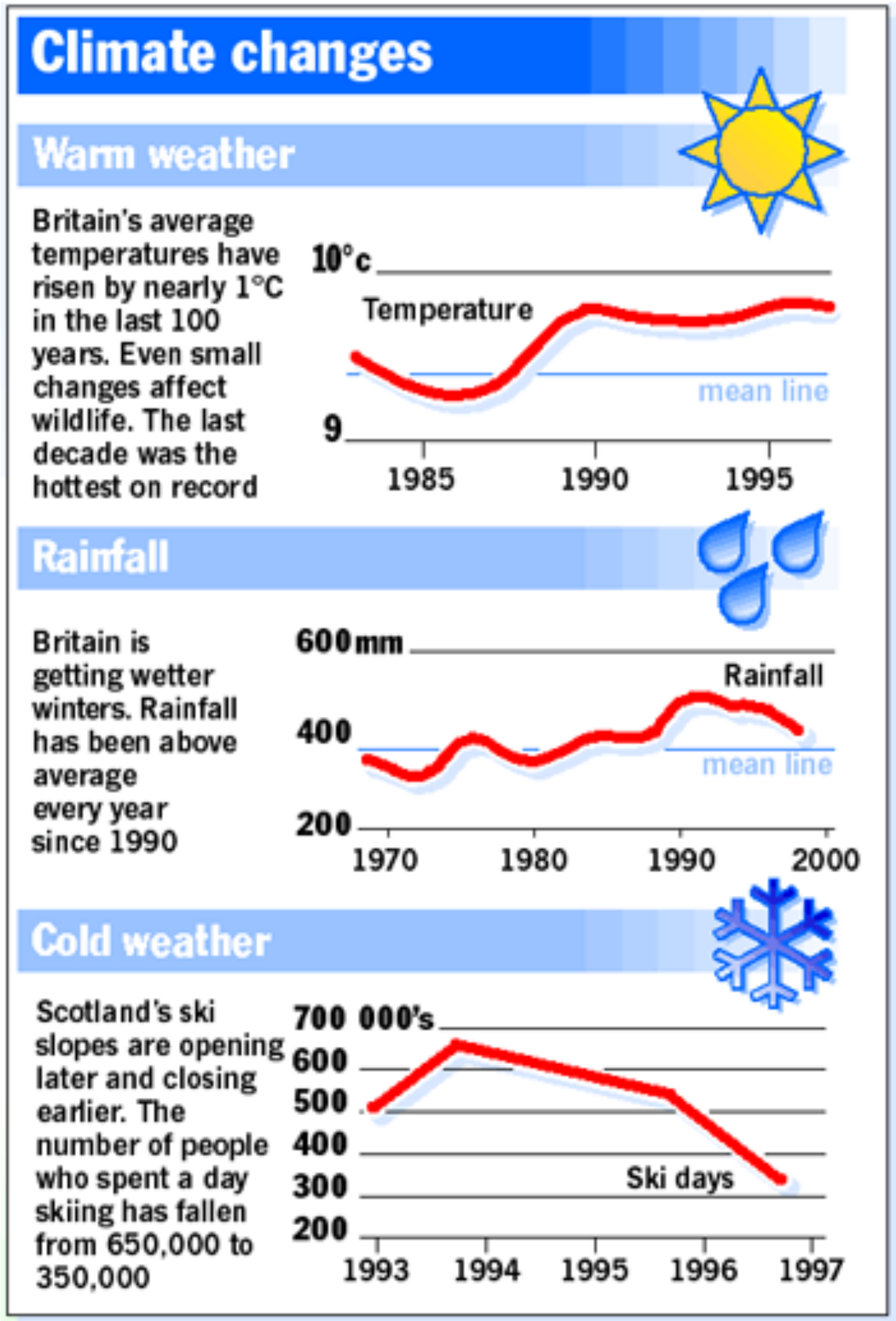
- Generalising
  - $E[X^n]$ n^{th} moment
  - $E[(X-E[X])^n]$ n^{th} central moment

- First and second moment are used a lot. Other moments can be useful in certain applications to characterize the PMF

Still a long way to go...







Covered Today

- Chapter 1, 3, and 4
- Key terms
  - Stochastic (probability) model
  - Outcome, event
  - Axioms of probability
  - Conditional probability
  - Independence of events
  - Discrete and continuous random variables
  - Bernoulli, Uniform, Gaussian
  - Expectation, variance
  - Moments