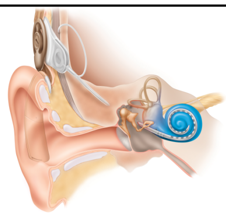


Stochastic processes – Lecture 1: Introduction

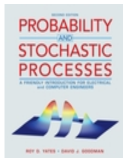
Dr. Ir. Richard C. Hendriks
21/04/16

TU Delft Delft University of Technology

Challenge the future

Organization (1)

- Plenary Lectures
- Book: R.D. Yates and D.J. Goodman, Probability and Stochastic Processes: A Friendly Introduction for Electrical and Computer Engineers (Second/third Edition)
- Slides, exams, etc: Blackboard
- (The least) preparation of the next lecture: Make sure to have
  - understood lecture material last lecture
  - read the corresponding chapters
  - Solved the selected exercises
- Your responsibility: Study and solve exercises before the next lecture.



TU Delft

Exam

The exam consists of two parts:

1. Part 1: Mid-term exam on 18th of May
2. Part 2: Final exam on 24th of June.

- Both parts will consist of exercises from signal processing and stochastic processes (50/50). The end result is the average of the mid-term and final exam.
- The result of mid-term exam is only valid in case you take the final exam in June. (Hence, it is not anymore valid for the re-exam)

TUDelft

Topics to be Discussed (2^{nd} edition)

1. Introduction and Refresher Chs.: 1, 2 and 3
2. Stochastic processes Chs.: 4, 5 and 10.1-10.8
3. Using the auto-correlation function Chs.: 10.9-10.12 and 11.1-11.3
4. Linear Prediction Chs.: 11.4-11.9
5. Statistical Estimators Ch.: 9
6. Sums of random variables Ch.: 6
7. Stochastic processes and signal processing: Understanding Quantization

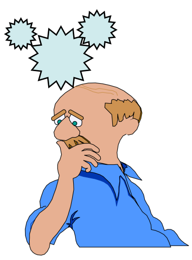
TUDelft

Chapters 1, 2, and 3

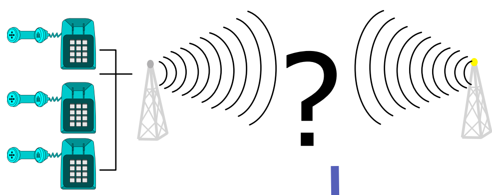
Motivation
Probability Models
Discrete and Continuous
Random Variables
PMF,PDF,expectation

TU Delft

The Engineer: Problem Solver



Reality



Simplification:

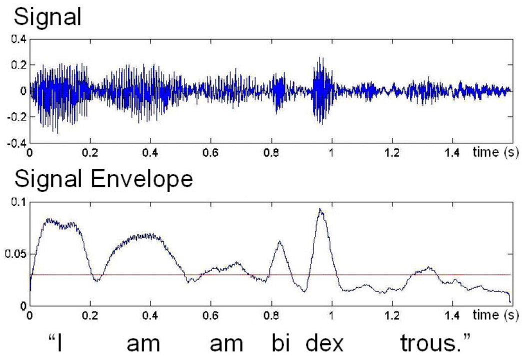
Model

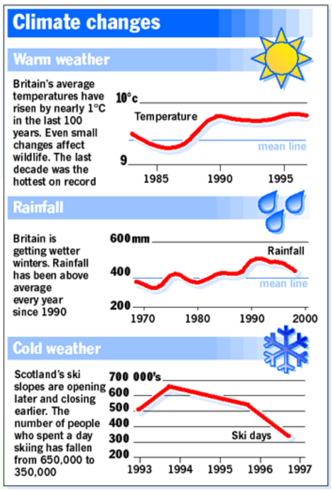
Solution

Tools:
- Analysis
- Algebra
- Physics
- Probability theory

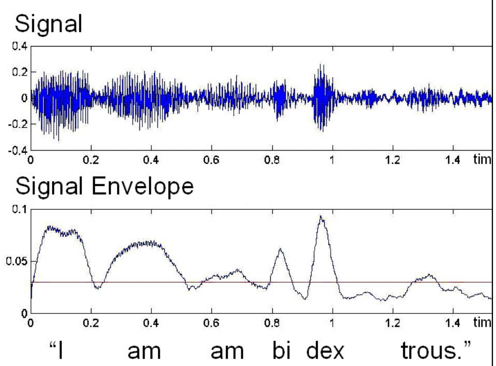
TU Delft

Voice signal, weather, stock





Stochastic processes

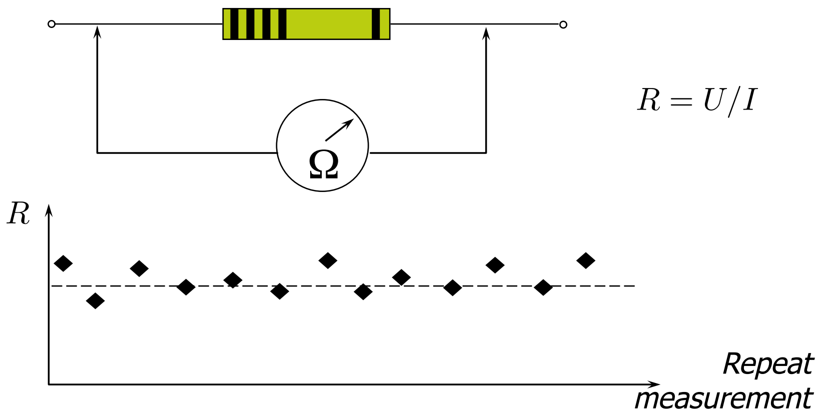


- The goal of this course: To learn how to **describe** and **analyze** these type of signals

TUDelft

Model: Abstraction of Reality

The Resistor



TU Delft

Random vs. Deterministic Modeling

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

TU Delft

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.
- Skillful usage of probability theory can greatly reduce complexity in some case.
- but dont use a saw if you need a screwdriver

TUDelft

10

Objectives – Today:

Random processes are build on concepts of probability theory:

Refresher on Probability Theory

- Modeling: Random versus deterministic processes
- Probability density functions
- Basic notions of probability
- Refresher random (stochastic) variables
- Probability Mass Function (PMF), Probability Density Function (PDF), Cumulative Distribution Function (CDF)
- Bayes' theorem
- Expected value, Variance

TUDelft

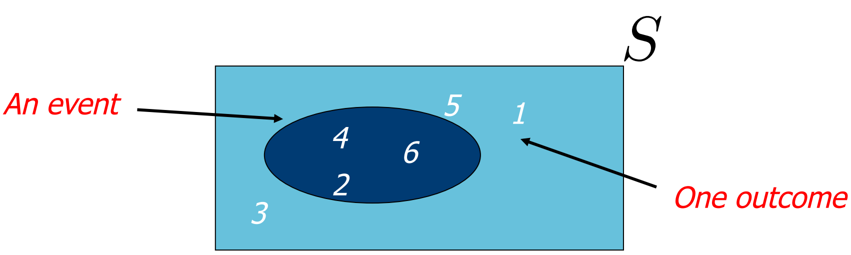
Probability (Stochastic) Model

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

TU Delft

Probability Model – "Dice" Example

- Experiment: to throw a die
- Identify the values that can be observed
  - Outcomes: The sample space is $S = \{1, 2, 3, 4, 5, 6\}$
  - Events (sets of outcomes): $A = \{\text{even}\}$
- Assign probabilities to events
  - For example: $P(A) = 1/2$, $P(S) = 1$



TU Delft

Axiomatic Approach to Probability (1)

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

1. A probability is never negative

$$P[A] \geq 0 \;\; \forall \; A$$

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:

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

Axiomatic Approach to Probability (2)

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?
- Based on relative frequency $f_k(n)$

Number of times

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

Outcome number

Relative Frequency for “Roll a 6”

$f_6(n)$

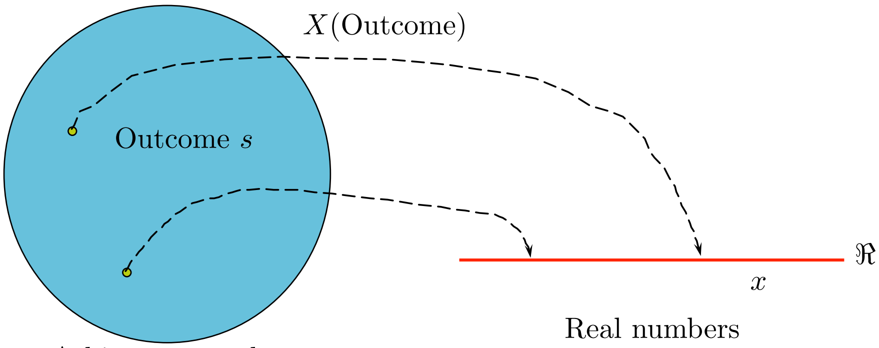
n

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)

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



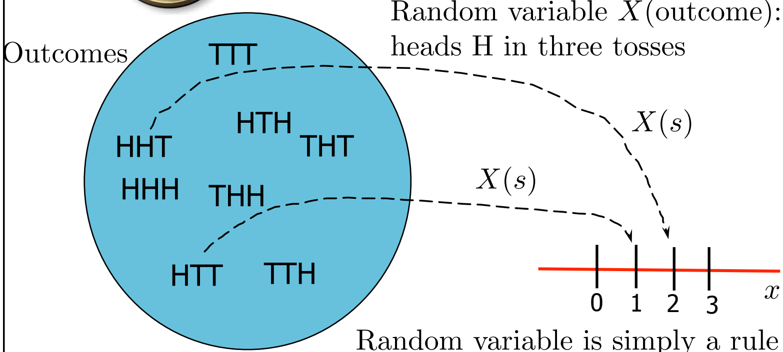
Arbitrary sample space
Often related to physical world

TU Delft

Definition of Random Variable (2)

Experiment:Head or tail
in 3 tosses

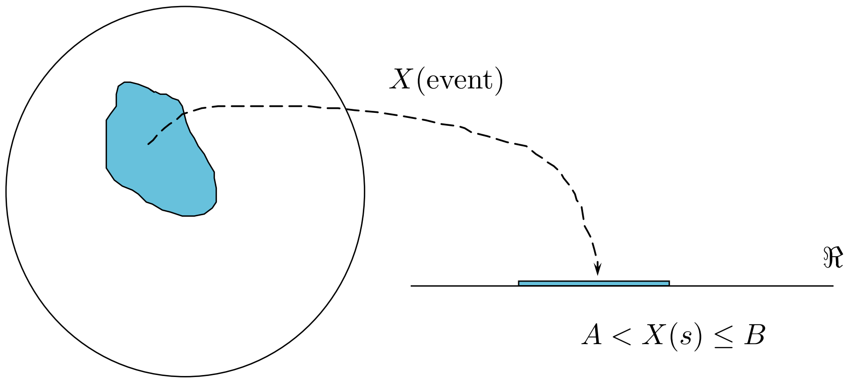
Random variable $X(\text{outcome})$: number of heads H in three tosses



Random variable is simply a rule to assign a number to each outcome in the sample space

TU Delft

Definition of Random Variable (3)



- roll even; grown up
- $\{2, 4, 6\}$; $[18, \infty)$

Notational Convention

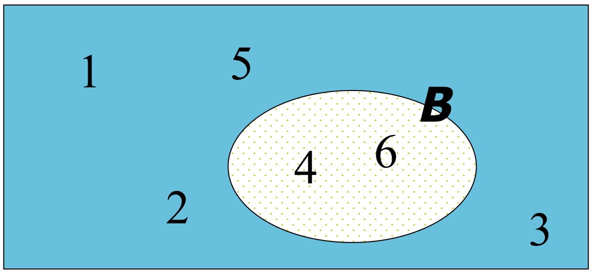
$A < X(s) \leq B$ is the same as $A < X \leq B$

- 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:

$$X = x$$

Conditional Probability (1)

Sometimes the occurrence of one event influences the probability of occurrence of other events



- P[odd number] ?
- P[odd number if we know that the outcome is in event B] ?

TU Delft

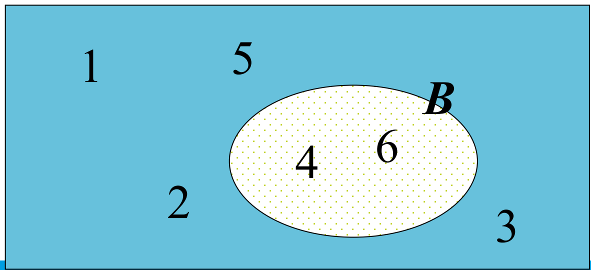
Conditional Probability (2)

Interpretation: The probability of A, given that the event B has already occurred.

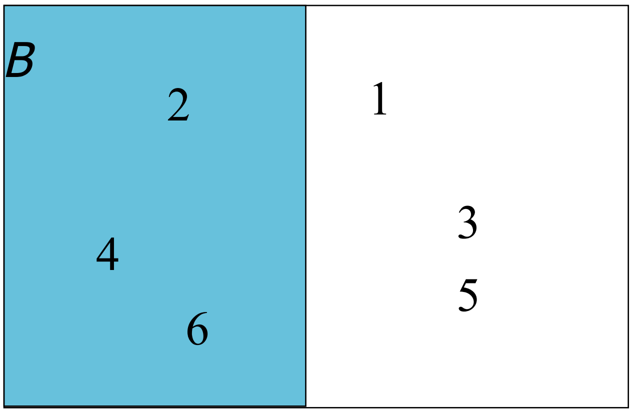
$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A,B)}{P(B)}$$

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

Known as:
Bayes theorem



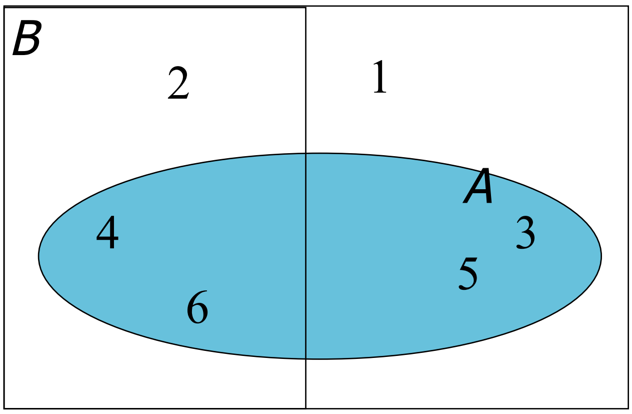
Example (1)



B: "Even outcome" when rolling the dice

TU Delft

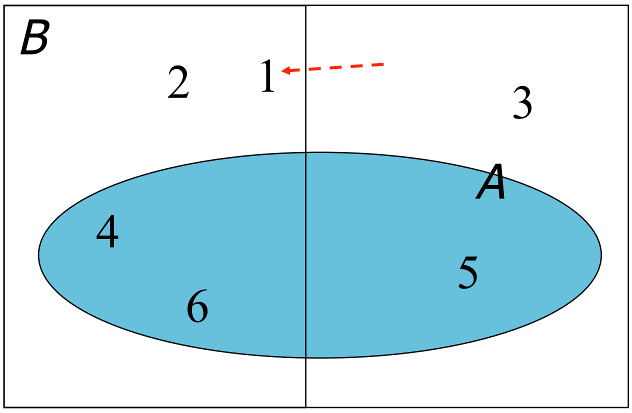
Example (2)



A: "3 or more" when rolling the dice.
How large is $P(A|B)$?

TU Delft

Example (3)



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

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: $A = \{2, 4, 6\}$ $B = \{1, 3, 5\}$

$$P(A, B) = 0 \neq P(A)P(B) = \frac{1}{2}\frac{1}{2} = \frac{1}{4}$$

So, mutual exclusive, but not independent!

Discrete vs. continuous Random Variables

Discrete random variable:

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

Continuous random variable:

- continuous sample space
  - Have an infinite number of outcomes
  - Example: Pick a real-valued number between 3 and 5.
    * How many outcomes?
    * Every outcome has probability zero: $P[X = x] = 0$.
      Therefore we will consider events instead, e.g., $4 \leq X \leq 5$.

TU Delft

PDF versus PMF

- Discrete RV's are usually described by their Probability Mass Function (PMF)
- Continuous RV's are usually described by their Probability Density Function (PDF)
- For both cases the CDF can also be used
- For Discrete RV's the PDF also exists . . .
  - Useful for mixed random variables
  - PDF becomes a sum of delta functions (section 3.6)

TU Delft

Properties of PDF and PMF

For a continuous random variable X with PDF $f_X(x)$:

- $f_X(x) \geq 0 \; \forall \; x$
- $F_X(x) = \int_{-\infty}^{x} f_X(u)du$
- $\int_{-\infty}^{+\infty} f_X(x)dx = 1$

For a discrete random variable X with PMF $P_X(x)$ and range S_X:

- $P_X(x) \geq 0 \; \forall \; x$
- $\sum_{x \in S_X} P_X(x) = 1$
- For an event $B \subset S_X$, the probability that X is in the set B is

$$P[B] = \sum_{x \in B} P_X(x)$$

TUDelft

Properties of the CDF

For continuous AND discrete RVs:

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

Properties:

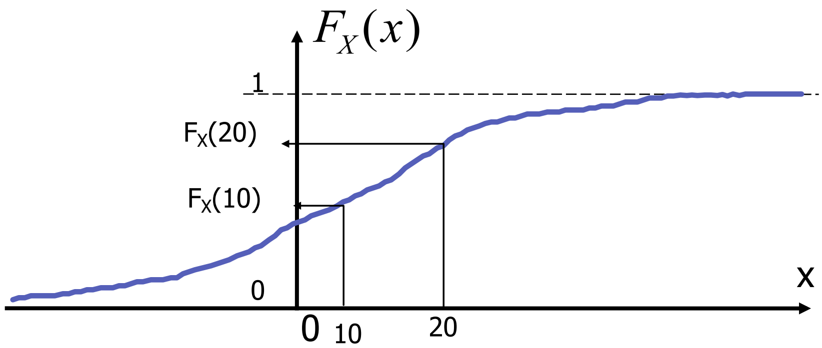
- $F_X(-\infty) = 0$
- $F_X(\infty) = 1$
- $P[x_1 < X \leq X_2] = F_X(x_2) - F_X(x_1)$
- For all $x' \geq x$, $F_X(x') \geq F_X(x)$

For continuous RVs: $F_X(x) = P[X \leq x] = \int_{-\infty}^{x} f_X(u)du$

TUDelft

Example CDF

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



$$P(10 < X \le 20) = F_X(20) - F_X(10)$$

Properties of the CDF

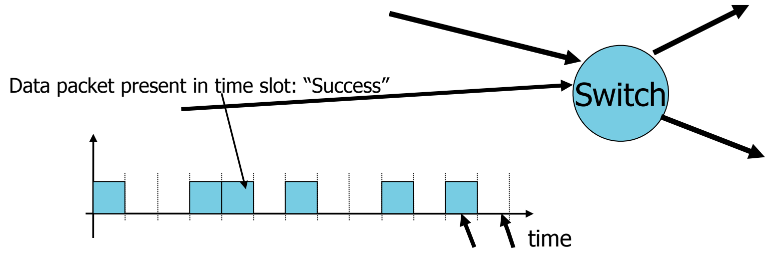
Important difference between continuous AND discrete RVs:

- CDF of a continuous RV is always a continuous function
- CDF of discrete RV shows jumps at values of x with on-zero probability.
- 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!

Probability Mass Function - example



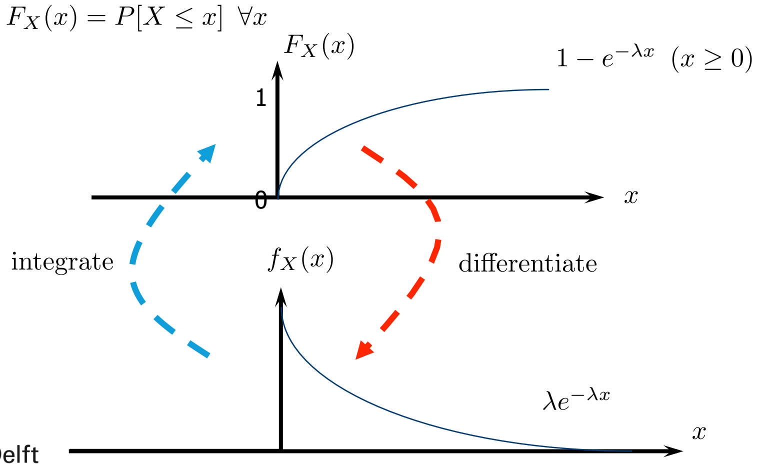
$$P_X(x) = \begin{cases} 1-p & x=0 \\ p & x=1 \\ 0 & \text{otherwise} \end{cases}$$

$P_X(x)$ is called the Probability Mass Function (PMF) of a Bernoulli RV. CDF?

TUDelft

Calculating Probabilities (1)

Cumulative Distribution Function (CDF):
$F_X(x) = P[X \leq x] \ \forall x$



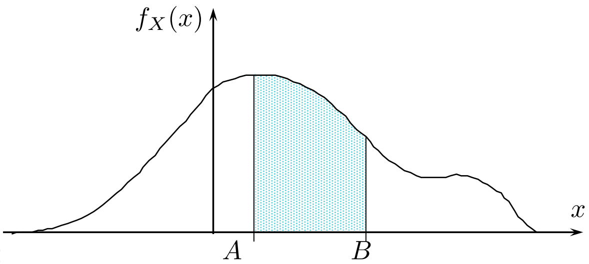
TUDelft

Calculating Probabilities (2)

Usually, the PDF is used to calculate probabilities, i.e.,

$$P(X \leq A) = F_X(A) = \int_{-\infty}^{A} f_X(x)dx$$

$$P(A \leq X \leq B) = F_X(B) - F_X(A) = \int_{A}^{B} f_X(x)dx$$



TU Delft

Useful Random Variables

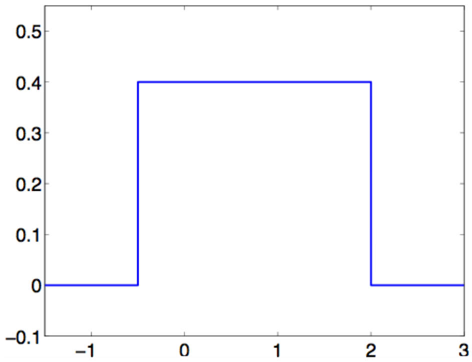
- Discrete random variables
  - Bernoulli, Binomial, Geometric, Poisson
  - Uniform
  - See book for equations and properties
- Continuous random variables
  - Gaussian
  - Exponential
  - Erlang
  - Uniform
  - See book for equations and properties.

TU Delft

The Uniform Distribution

- Uniform distribution exists in discrete and continuous version!

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



$$\begin{aligned} a &= ? \\ b &= ? \end{aligned}$$

TU Delft

39

The Uniform Distribution

- Uniform distribution exists in discrete and continuous version!

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

What is the corresponding CDF?

$F_X(x)$ is 1 if $x \geq b$ and 0 if $x < a$.
What is $F_X(x)$ for $a \leq x < b$?

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} f_X(u)du \\ &= \int_{a}^{x} \frac{1}{b-a} du = \left[\frac{u}{b-a}\right]_a^x \\ &= \frac{x-a}{b-a} \end{aligned}$$

$$F_X(x) = \begin{cases} 0 & x \leq a \\ \frac{x}{b-a} & a \leq x < b \\ 1 & x > b \end{cases}$$

TU Delft

40

Derived Random Variables (1)

- The PDf of an arbitrary derived random variable $Y = g(X)$ is often difficult to calculate. A general procedure is to
  1. Find the CDF of $F_Y(y) = P[Y \leq y]$
  2. Compute the PDF by calculating the $f_Y(y) = \frac{dF_Y(y)}{dy}$
- Special case is a linear transformation:

$$Y = aX + b \Leftrightarrow f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

TUDelft

Derived Random Variables (2)

Proof: $Y = aX + b$

$a > 0$

$$F_Y(y) = P(Y \leq y) = P(aX + b \leq y) = P\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right)$$

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$$

$a < 0$

$$F_Y(y) = P(Y \leq y) = P(aX + b \leq y) = P\left(X \geq \frac{y-b}{a}\right) = 1 - F_X\left(\frac{y-b}{a}\right)$$

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{-a} f_X\left(\frac{y-b}{a}\right)$$

$$\Rightarrow f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

TUDelft

Derived Random Variables (3)

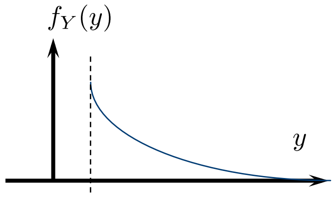
Example:

$$Y = aX + b \Leftrightarrow f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

$$Y = 2X + 1$$

$$f_X(x) = 3e^{-3x} x \geq 0$$

$$f_Y(y) = \frac{1}{2} f_X\left(\frac{y-1}{2}\right) = \frac{3}{2} e^{-3\left(\frac{y-1}{2}\right)}$$



Derived Random Variables (4)

Example:

Let X have an exponential PDF: $f_X(x) = \lambda e^{-\lambda x}$ $x \geq 0$.

What is the pdf of the random variable $Y = \sqrt{X}$?

1) compute CDF $F_X(x)$

$$\int_0^x f_X(u)du = \int_0^x \lambda e^{-\lambda u} du = \left[-e^{-\lambda u}\right]_0^x = 1 - e^{-\lambda x}$$

2) Compute CDF $F_Y(y)$

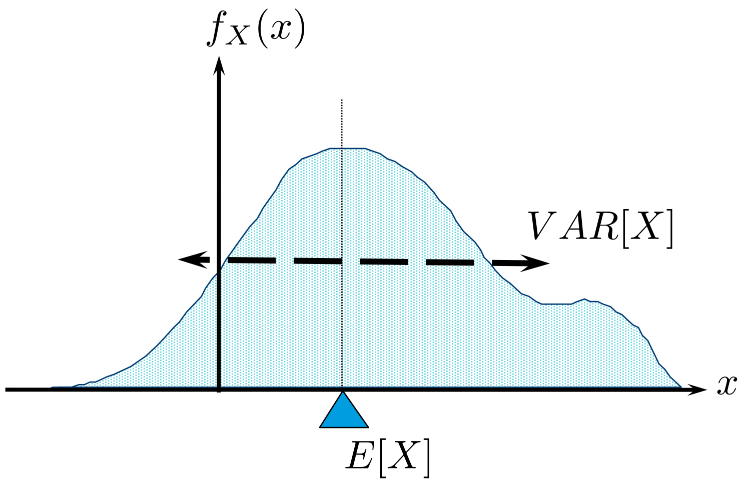
$$F_Y(y) = P(Y \leq y) = P(\sqrt{X} \leq y) = P(X \leq y^2) = 1 - e^{-\lambda y^2}$$

3) Compute PDF $f_Y(y)$

Rayleigh PDF

$$f_Y(y) = \frac{d}{dy} 1 - e^{-\lambda y^2} = 2y\lambda e^{-\lambda y^2}$$

Characterizing the PDF/PMF



TU Delft

Expected Value

- The precise and total behavior of a random variable is described by the PMF/PDF 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 practice.

  Two measures for average behavior:
  - Expected value (expectation, mean)
  - Variance (standard deviation)

TU Delft

Definition of Expected Value

- Expected value is one number that characterizes the PMF (and CDF)
- Discrete RV:
$$E[X] = \mu_X = \sum_{\forall} x P_X(x)$$
- Continuous RV:
$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$
- $E[\,\cdot\,]$ is a linear operator:
$$E[Z] = E[X + Y] = E[X] + E[Y]$$
$$E[aX] = aE[X]$$

TU Delft

Expectation Operator E[.]

- Use $E[\,\cdot\,]$ as much as possible, and stay away from the integrals/summations as long as you can.
- $E[\,\cdot\,]$ is a linear operator:
$$E[Z] = E[X + Y] = E[X] + E[Y]$$
$$E[aX] = aE[X]$$
- Example:
  - X and Y are Bernoulli RVs with $p = 0.4$ and $p = 0.7$.
  - Find $E[Z] = E[3X + 6Y]$
  - Long route: Find $P_Z(z)$
  - Short route: $E[Z] = E[3X + 6Y] = 3E[X] + 6E[Y]$

TU Delft

Variance: Var[X]

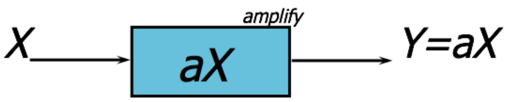
- Variance is (again) one number to characterize the behavior of the PMF.
- Measure for "width", or "dispersion" of X around expected value of PMF.
- Discrete RV:

$$\begin{aligned} VAR[X] &= \sigma_X^2 = E[(X - E[X])^2] \\ &= \sum_{\forall x} (x - E[X])^2 P_X(x) \\ &= E[X^2] - E[X]^2 \end{aligned} \tag{1}$$

- Continuous RV:

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

Example



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

Variance scales quadratically

Moments

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

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

TUDelft

Covered Today

- Chapter 1, 2, and 3
- Key terms
  - Stochastic (probability) model
  - Outcome, event
  - Axioms of probability
  - Conditional probability
  - Independence of events
  - Discrete and continuous random variables
  - Bernoulli, Uniform, Gaussian
  - Derived random variables
  - Expectation, variance
  - Moments

TUDelft

What is expected from you before next lecture?

Before next Lecture:

- Make sure to have studied Chapters 1, 2 and 3 and understand the theory.
- Make sure to have made the following exercises:
  Lecture 1: refresher random variables

Event space	1.4.1, 1.5.1
Conditional probability	1.5.5
Bayes; rule	1.7.6
CDF	2.4.3
Expected value	2.5.5
Variance	2.8.4
Conditional PMF	2.9.3
Uniform distribution	3.4.5

TU Delft