EC5.102: Information and Communication

Module on

Introduction to probability and random variables

Arti D. Yardi

Email address: arti.yardi@iiit.ac.in

Introduction to probability

Why do we need probability?

- Let us consider some simple "events".
  - Event: Sunrise at 6am in the morning
  - Event: Party 'xyz' winning in an election
  - Event: Tossing a coin and getting 'HEAD' as its outcome
- We are not always "able" to say something about an outcome of an event in a "deterministic" manner!
- We live a "random" world but wish to predict/infer something about an event.
- "Under the condition" we have some side information, can we improve upon our previous inference?
- How to characterize/study such questions mathematically?

Applications of probability in real life

- Communication systems (noisy channel)
- Forecasting (weather, demand, election outcomes)
- Finance (stock market, online shopping, gambling)
- Machine learning (data analysis)
- Computer science (routing, scheduling)

Can you think of some applications?

Probability in communication engineering

- Communication channel
  - How does a channel "behaves"?
  - Mathematical modeling of a channel
- Analyze the performance of a system
  - Provide theoretical analysis of the system behaviour
  - Design guidelines
- Source encoding and channel encoding (to be studied in detail soon)
  - Data compression
  - Combat with the noise introduced by the channel

Recap of the previous class

Recap

- We now live a digital world!

  Analog signal $\rightarrow$ Sampling $\rightarrow$ Quantization $\rightarrow$ Bit-sequence

- **How to do further processing of this bit-sequence?**
  - Source encoding:

    Bit-sequence $\rightarrow$ Source encoding $\rightarrow$ Bit-sequence'
  - Channel encoding:

    Bit-sequence' $\rightarrow$ Channel encoding $\rightarrow$ Bit-sequence''

- We shall study source and channel coding soon.
- Digression: Basics of probability

Recap: Introduction to probability

- Need for probability: Our inability to say "something" about an outcome of an event in a "deterministic" manner!
- Applications: Communication systems, Forecasting, Finance, Machine learning, Computer science...
- Next agenda: Basics of probability and random variables
  - Probability space
  - Random variables (RVs)
  - Types of RVs: Discrete and continuous RVs
  - Joint and conditional RVs

About my teaching style

- I will be using a combination of slides and board.
- Be super interactive.. ask questions..
- Discuss learnings of the class with your friends.
- Refer to the suggested reference books.
- There will be breakout sessions to solve problems (very important).
- There will be self-quizes/surprize-quizes in the class.. :P
- NO LAPTOPS in the class!

Review: Set theory

Review: Set theory

- Universal set: $\mathcal{S}$
- Subset of set $\mathcal{S}$: $A \subseteq \mathcal{S}$
- Complement of A: A^c
- Intersection: $A \cap B$
- Union: $A \cup B$
- Empty (or null) set: Φ
- Element and simpleton sets

- $A \backslash B = A \cap B^c$
- Disjoint sets: $A \cap B = \Phi$
- $\Phi^c = \mathcal{S}$
- $(A \cup B)^c = A^c \cap B^c$
- $(A \cap B)^c = ?$
- $A \cap A^c = ?$
- $A \cup A^c = ?$

Review: Set theory

- Countably finite set
- Countably infinite set
- Uncountably infinite set
- Power set of a set with finite elements
- Can we talk about power set of a set with infinite elements?

Set theory and probability space

Set theory	Probability space (to be defined precisely soon)
Universe	Sample space
Subset	Event
Element	Outcome
Simpleton set	Simple event
Null set	Impossible event
Disjoint sets	Mutually exclusive events

- Probability:
  - Assign a value called as "probability" to an event.
  - Can you relate this to a function/map?
- **Probability space**: Formalize the notions of sample space, events, and probability.

Self-quiz

- When do two events A and B are said to be mutually exclusive?
- When do two events A and B are said to be independent?

Probability space

Motivation for defining a probability space

- Let us try to “formalize” a random experiment.
- Example: Rolling a dice
  - How to “describe” this experiment mathematically?
  - Make a list of things that one needs do towards this.
  - Suppose our focus is on only two events A and B. Can you simplify this description?
- Can you think something similar for a coin toss experiment?
- Can you think something similar for an experiment of randomly choosing a real number between an interval $[0, 1]$?
- Probability space provides a formal model of a random experiment.

Probability space $(\Omega, \mathcal{F}, P)$

- Probability space provides a formal model of a random experiment.
- A probability space consists of three elements $(\Omega, \mathcal{F}, P)$:
  1. Sample space Ω: Set of all possible outcomes
  2. Event space $\mathcal{F}$: Collection of sets outcomes in Ω such that
     - ★ Contains both empty set ϕ and Ω
     - ★ Closed under complement:
       $$\text{If } A \in \mathcal{F} \text{ then } A^c \in \mathcal{F}$$
     - ★ Closed under countable union
       $$\text{If } A_1, A_2, \ldots \in \mathcal{F} \text{ then } \cup_{i=1}^{\infty} A_i \in \mathcal{F}$$
  3. Probability measure $P(.)$: Function from event space $\mathcal{F}$ to $[0, 1]$

Event space: Example

- Event space $\mathcal{F}$: 1. Contains ϕ and Ω
  2. If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$
  3. If $A_1, A_2 \in \mathcal{F}$ then $A_1 \cup A_2 \in \mathcal{F}$

- Example: Rolling a six-sided dice ($\Omega = \{1, 2, 3, 4, 5, 6\}$)
  - Consider two events: $A = \{1 \cup 2 \cup 3\}$ and $B = \{1\}$
  - Event space $\mathcal{F}_A$ generated by A will be

$$\mathcal{F}_A = \{\Omega, \phi, A, A^c\} = \Big\{\Omega, \phi, \{1 \cup 2 \cup 3\}, \{4 \cup 5 \cup 6\}\Big\}$$

  - Event space $\mathcal{F}_B$ generated by B will be

$$\mathcal{F}_B = \{\Omega, \phi, B, B^c\} = \Big\{\Omega, \phi, \{1\}, \{2 \cup 3 \cup 4 \cup 5 \cup 6\}\Big\}$$

  - Event space $\mathcal{F}_{AB}$ generated by A and B will be

$$\begin{aligned}\mathcal{F}_{AB} = \Big\{&\Omega, \phi, A, A^c, B, B^c, (A \cup B), (A \cup B)^c, (A \cup B^c), (A \cup B^c)^c,\\ &(A^c \cup B), (A^c \cup B)^c, (A^c \cup B^c), (A^c \cup B^c)^c\Big\}\\ = \Big\{&\Omega, \phi, A, A^c, B, B^c, A^c \cup B, (A^c \cup B)^c\Big\}\\ = \Big\{&\Omega, \phi, A, A^c, B, B^c, \{1 \cup 4 \cup 5 \cup 6\}, \{2 \cup 3\}\Big\}\end{aligned}$$

Event space: Example

- Example: Rolling a six-sided dice ($\Omega = \{1, 2, 3, 4, 5, 6\}$)
- Power set of the sample space will be

$$
\begin{aligned}
\mathcal{F} = \Big\{ & \Phi, \{1 \cup 2 \cup 3 \cup 4 \cup 5 \cup 6\}, \\
& \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \\
& \{1 \cup 2\}, \{1 \cup 3\}, \ldots, \{5 \cup 6\}, \\
& \{1 \cup 2 \cup 3\}, \ldots, \{4 \cup 5 \cup 6\}, \\
& \{1 \cup 2 \cup 3 \cup 4\}, \ldots, \{3 \cup 4 \cup 5 \cup 6\}, \\
& \{1 \cup 2 \cup 3 \cup 4 \cup 5\}, \ldots, \{2 \cup 3 \cup 4 \cup 5 \cup 6\} \Big\}
\end{aligned}
$$

- Power set $\mathcal{F}$ is a valid event set: Homework
- How many elements will be there in power set $\mathcal{F}$?

Axioms of probability

Probability measure

- Probability measure $P(.)$: Function from event space $\mathcal{F}$ to $[0, 1]$
- Axioms of probability:
  - For all $A \in \mathcal{F}$, $0 \leq P(A) \leq 1$
  - $P(\phi) = 0$
  - $P(\Omega) = 1$
  - If $A \cap B = \phi$ then $P(A \cup B) = P(A) + P(B)$
- Example: Rolling a six-sided (biased) dice

$$P(\{1\}) = 0.1, \ P(\{2\}) = 0.2, \ P(\{3\}) = 0.1,$$
$$P(\{4\}) = 0.3, \ P(\{5\}) = 0.15, \ P(\{6\}) = 0.15,$$

Various concepts in probability

- Conditional probability
- Independent events
- Mutually exclusive events
- Total probability theorem
- Bayes' theorem

Recap of the previous class

Recap: Probability space

- Review: Set theory and elementary concepts in probability (Bayes theorem, total probability law, Axioms of probability etc)
- A probability space consists of three elements $(\Omega, \mathcal{F}, P)$:
  1. Sample space Ω: Set of all possible outcomes
  2. Event space $\mathcal{F}$: Collection of sets outcomes in Ω such that
     - ★ Contains both empty set ϕ and Ω
     - ★ Closed under complement: If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$
     - ★ Closed under countable union: If $A_1, A_2, \ldots \in \mathcal{F}$ then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$
  3. Probability measure $P(.)$: Function from event space $\mathcal{F}$ to $[0, 1]$
- **Think: Probability space does capture "essential" model of a random experiment!**

Motivation to random variables

Example of rolling two dice

- Example of rolling two dice where we are interested in the sum of two dice.
- Suppose $X =$ sum of two dice. Then we have

$$\Omega = \Big\{(1,1),(1,2),\ldots,(1,6) \\ (2,1),(2,2),\ldots,(2,6) \\ \vdots \\ (6,1),(6,2),\ldots,(6,6)\Big\} \xrightarrow{X} \Omega' = \{2,3,\ldots,12\}$$

- Suppose $\mathcal{F}$ and $\mathcal{F}'$ are power sets of Ω and Ω' respectively.
- Can you write down P'?
- We have a map X given by

$$X : (\Omega, \mathcal{F}, P) \to (\Omega', \mathcal{F}', P')$$

- For our application, it is more convenient to work with $(\Omega', \mathcal{F}', P')$ than $(\Omega, \mathcal{F}, P)$.

Example of choosing two real numbers from $[0, 1]$

- Choose any two real numbers from $[0, 1]$
  - What will be Ω?
  - What will be $\mathcal{F}$? (to be discussed soon)
  - What will be P? (e.g. choose any number with uniform probability)
- We are interested only in addition of these two numbers.
  - What will be Ω'?
  - What will be $\mathcal{F}'$?
  - What will be P'? (Will it be uniform? Yes/no?)
- We have a map X given by

$$X : (\Omega, \mathcal{F}, P) \to (\Omega', \mathcal{F}', P')$$

- It is more convenient to work with $(\Omega', \mathcal{F}', P')$ than $(\Omega, \mathcal{F}, P)$.

Motivation to random variables

- One can write down the event space $\mathcal{F}$, when Ω has finite entries.
- What to do when Ω has infinite (countable/uncountable) entries?
- Given a random experiment with associated $(\Omega, \mathcal{F}, P)$, it is sometimes difficult to deal directly with $\omega \in \Omega$. Example: Rolling a dice twice
- Depending on the experiment, Ω will be different.
- Can we come up with a general platform which is independent of the choice of a particular Ω?
- It is desirable and convenient to study probability space and related advanced concepts using this general platform!

Answer: Random Variables (rv or RV)

Definition of random variables

Random variables

- Recall example of rolling two dice.
  - Suppose we are only interested in 'sum of two dice'.
  - It is convenient to consider the map $X : (\Omega, \mathcal{F}, P) \to (\Omega', \mathcal{F}', P')$
- For random variables, we consider "special" $\Omega', \mathcal{F}'$ and the corresponding induced probability measure P'.
  - Ω' will be the set of real numbers, denoted by $\mathbb{R}$.
  - $\mathcal{F}'$ will be Borel σ-algebra, denoted by $\mathcal{B}(\mathbb{R})$. (This is an advanced level topic. We will not go into the details).
  - P' will be the corresponding induced probability measure, denoted by P_X.
- A random variable X is a map given by

$$X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$$

(Map X needs to satisfy some more conditions! To be discussed soon.)

Borel σ-algebra

- Borel σ-algebra $\mathcal{B}(\mathbb{R})$:

  If $\Omega = \mathbb{R}$, then $\mathcal{B}(\mathbb{R})$ is the event space generated by open sets of the form (a, b) where $a \leq b$ and $a, b \in \mathbb{R}$.

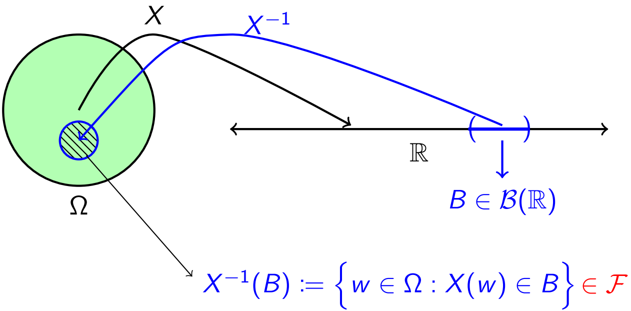
- $\mathcal{B}(\mathbb{R})$ contains intervals of the form

$$
\begin{gathered}
[a, b] \\
[a, b) \\
(a, \infty) \\
[a, \infty) \\
(-\infty, b] \\
(-\infty, b)
\end{gathered}
$$

Simple examples of random variables

- A single coin toss
- Three coin tosses
- Choose a real number in the interval $[0, 100]$

Random variables



- $\Omega \xrightarrow{X} \mathbb{R}$, $\mathcal{F} \xrightarrow{X} \mathcal{B}(\mathbb{R})$, and $P(.) \xrightarrow{X} P_X(.)$
- Care must be taken such that the events you consider in the new event space $\mathcal{B}(\mathbb{R})$ are also valid events included in $\mathcal{F}$.
- $X^{-1}(B)$ is called as the preimage or the inverse image of B.

Definition of a random variable

- **Main idea:** The events you consider in $\mathcal{B}(\mathbb{R})$ are also valid events included in $\mathcal{F}$.

- **Definition of a random variable:**

  A random variable X is a map $X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ such that for each $B \in \mathcal{B}(\mathbb{R})$, the inverse image $X^{-1}(B) := \{w \in \Omega : X(w) \in B\}$ satisfies the condition

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

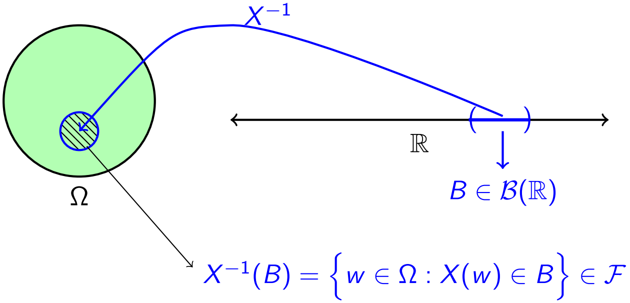
- P_X denotes the probability measure induced by X on $\mathcal{B}(\mathbb{R})$ and is given by

$$P_X(B) = \Pr(w \in \Omega : X(w) \in B)$$

- **General definition of a random variable**:

  A random variable X is a map $X : (\Omega, \mathcal{F}, P) \to (\Omega', \mathcal{F}', P')$ s.t. for each $f \in \mathcal{F}'$, $X^{-1}(f) \in \mathcal{F}$.

Definition of a random variables



A random variable X is a map $X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ such that for each $B \in \mathcal{B}(\mathbb{R})$, the inverse image $X^{-1}(B) := \{w \in \Omega : X(w) \in B\}$ satisfies

$$X^{-1}(B) \in \mathcal{F} \text{ and}$$
$$P_X(B) = \Pr(w \in \Omega : X(w) \in B)$$

Examples

Definition of a random variable: Example-1

- Definition of a random variable:

> A rv X is a map $X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ s.t. for each $B \in \mathcal{B}(\mathbb{R})$, the inverse image $X^{-1}(B) := \{w \in \Omega : X(w) \in B\}$ satisfies
>
> $$X^{-1}(B) \in \mathcal{F} \;\text{ and }\; P_X(B) = \Pr(w \in \Omega : X(w) \in B)$$

- Example: Consider the following map corresponding to two coin tosses.

$$HH \xrightarrow{X} 0,\; HT \xrightarrow{X} 1,\; TH \xrightarrow{X} 2,\; TT \xrightarrow{X} 3$$

- For which of the following event spaces can the map X defined above will be a valid random variable and why?
    - $\mathcal{F} = \Big\{\phi, \Omega, \{HH \cup TT\}, \{HT \cup TH\}\Big\}$
    - $\mathcal{F}$ is a power set of Ω

Definition of a random variable: Example-2 (Homework)

- Definition of a random variable:

> A rv X is a map $X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ s.t. for each $B \in \mathcal{B}(\mathbb{R})$, the inverse image $X^{-1}(B) := \{w \in \Omega : X(w) \in B\}$ satisfies
>
> $$X^{-1}(B) \in \mathcal{F} \;\text{ and }\; P_X(B) = \Pr(w \in \Omega : X(w) \in B)$$

- Example: Consider the following map corresponding to three coin tosses.

$$HHH \xrightarrow{X} 0,\; HHT \xrightarrow{X} 1,\; HTH \xrightarrow{X} 2,\; HTT \xrightarrow{X} 3$$
$$THH \xrightarrow{X} 4,\; THT \xrightarrow{X} 5,\; TTH \xrightarrow{X} 6,\; TTT \xrightarrow{X} 7$$

- For which of the following event spaces can the map X defined above will be a valid random variable and why?
  - $\mathcal{F} = \big\{\phi, \Omega, \{HHH, HTT, THT, TTH\}, \{HHT, HTH, THH, TTT\}\big\}$
  - $\mathcal{F}$ is a power set of Ω

Definition of a random variable: Example-3

- General definition of a random variable:

> A rv X is a map $X : (\Omega, \mathcal{F}, P) \to (\Omega', \mathcal{F}', P')$ s.t. for each $f \in \mathcal{F}'$, the inverse image $X^{-1}(f) := \{w \in \Omega : X(w) \in f\}$ satisfies
>
> $$X^{-1}(f) \in \mathcal{F} \;\text{ and }\; P_X(f) = \Pr(w \in \Omega : X(w) \in f)$$

- Example:
  - $\Omega = \{1, 2, 3, 4\}$ and $\Omega' = \{a, b, c\}$
  - $\mathcal{F}$ and $\mathcal{F}'$ are power sets of Ω and Ω' respectively.
  - Consider the following map.

$$X(1) = a, X(2) = b, X(3) = c, X(4) = a$$

  - Is $X : (\Omega, \mathcal{F}, P) \to (\Omega', \mathcal{F}', P')$ a random variable?

Definition of a random variable: Example-4 (Homework)

- General definition of a random variable:

> A rv X is a map $X : (\Omega, \mathcal{F}, P) \rightarrow (\Omega', \mathcal{F}', P')$ s.t. for each $f \in \mathcal{F}'$, the inverse image $X^{-1}(f) \coloneqq \{w \in \Omega : X(w) \in f\}$ satisfies
>
> $$X^{-1}(f) \in \mathcal{F} \text{ and } P_X(f) = \Pr(w \in \Omega : X(w) \in f)$$

- Example: Rolling two dice where we are interested in the sum of two dice.
  - Suppose X denotes sum of two dice.
  - Write down Ω, P, Ω', P'
  - $\mathcal{F}$ = Power set of Ω and $\mathcal{F}'$ = Power set of Ω'
  - Is $X : (\Omega, \mathcal{F}, P) \rightarrow (\Omega', \mathcal{F}', P')$ a random variable?

Definition of a random variable: Example-5

- General definition of a random variable:

> A rv X is a map $X : (\Omega, \mathcal{F}, P) \rightarrow (\Omega', \mathcal{F}', P')$ s.t. for each $f \in \mathcal{F}'$, the inverse image $X^{-1}(f) \coloneqq \{w \in \Omega : X(w) \in f\}$ satisfies
>
> $$X^{-1}(f) \in \mathcal{F} \text{ and } P_X(f) = \Pr(w \in \Omega : X(w) \in f)$$

- Example:
  - $\Omega = \{a, b, c\}$ and $\Omega' = \{0, 1\}$
  - $\mathcal{F} = \big\{\phi, \Omega, \{a\}, \{b \cup c\}\big\}$ and $\mathcal{F}'$ is power set of Ω'.
  - Consider the following map.

$$X(a) = 1, \quad X(b) = 0, \quad X(c) = 0$$

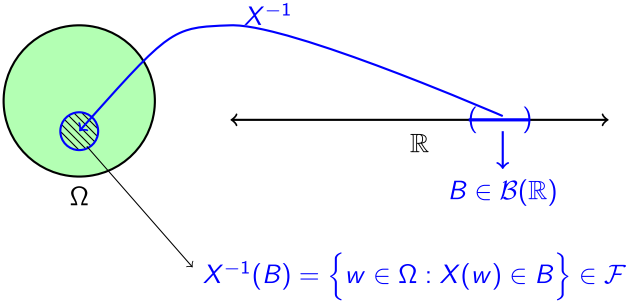
  - Is $X : (\Omega, \mathcal{F}, P) \rightarrow (\Omega', \mathcal{F}', P')$ a random variable?
  - This is an indicator random variable.

Summary

- Motivation for random variables
- Definition of a random variable
- Examples

Recap of the previous class

Definition of a random variables



A random variable X is a map $X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ such that for each $B \in \mathcal{B}(\mathbb{R})$, the inverse image $X^{-1}(B) := \{w \in \Omega : X(w) \in B\}$ satisfies

$$X^{-1}(B) \in \mathcal{F} \text{ and}$$
$$P_X(B) = \Pr(w \in \Omega : X(w) \in B)$$

Random variables

- A random variable X is a map $X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ s.t. for each $B \in \mathcal{B}(\mathbb{R})$, the inverse image $X^{-1}(B) \in \mathcal{F}$.
- Irrespective of which random experiment we are conducting, our probability space will be $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$!
  - $\mathbb{R}$: Sample space
  - $\mathcal{B}(\mathbb{R})$: Event space
  - P_X: Probability measure (Note: $P_X : \mathcal{B}(\mathbb{R}) \to [0, 1]$)
- **Focus: Let X be a random variable.**
  - X: Random variable (rv or RV or r.v.)
  - $\mathcal{X}$: "Support set" of rv X
  - x: "Realization" of rv X (Note: $x \in \mathcal{X}$)

Discrete random variables

Discrete random variable: Example

- Support set of a rv: The set of values that it can take.
- A random variable is called "discrete" if its support set consists of finite or countable finite elements.
- Example of discrete rv: Tossing two coins
  - $\Omega = \{HH, HT, TH, TT\}$
  - Suppose $P[HH] = 0.2, P[HT] = 0.3, P[TH] = 0.35, P[TT] = 0.15$
  - Consider example map: In class
- For a discrete rv X, we shall next study:
  - Probability mass function (PMF or pmf) of a discrete rv
  - Cumulative distribution function (CDF or cdf) of a discrete rv

Probability mass function (PMF or pmf) of a random variable

Examples of discrete rv

- Bernoulli random variable
- Binomial random variable
- Geometric random variable
- Poisson random variable

Cumulative Distribution Function (CDF) of a random variable

Definition of a random variable

- Consider a random variable $X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$.
- Random variable X translates the probability law defined on the events in the sample space to a probability law on corresponding to events on the real line $\mathbb{R}$.
- What corresponds to events on the real line?
- Recall: $\mathcal{B}(\mathbb{R})$ consists of intervals of the type (a, b), $(-\infty, a), (a, \infty)$ and so on.
- We shall now study cumulative distribution function (CDF) of a random variable, where we will focus on the events of the form $(-\infty, x]$, where $x \in \mathbb{R}$.

Cumulative distribution function (CDF) of a random variable: Definition

- In CDF of a r.v., we will focus on the events of the form $(-\infty, x]$, where $x \in \mathbb{R}$.
- We are thus intersted in events of the form $\{X \leq x, x \in \mathbb{R}\}$.
- CDF of a random variable X, denoted by $F_X(.)$ is defined as

$$F_X(x) := P_X[X \leq x]$$

$$= \text{Probability of event } (-\infty, x]$$

$$= P\Big(\Big\{w \in \Omega \text{ such that } X(w) \leq x\Big\}\Big)$$

Properties of CDF $F_x(.)$

- $F_x(.)$ is monotonically nondecreasing.
- $F_x(-\infty) = 0$ and $F_x(\infty) = 1$.
- $F_x(.)$ is right continuous.

Proofs: In class

Summary

Summary

- A random variable is called discrete if its support set consists of finite or countable infinite elements.

- CDF of a random variable X, denoted by $F_X(.)$ is defined as

$$F_X(x) := P[X \leq x]$$
$$= P\left(\left\{w \middle| X(w) \leq x\right\}\right)$$

- Examples:
  - Bernoulli r.v.
  - Binomial r.v.
  - Geometric r.v.
  - Poisson r.v.

Recap of the previous class

Discrete random variables

- A random variable (RV or rv) is called discrete if its support set consists of finite or countable infinite elements.
- Let X be a discrete rv with the support set $\mathcal{X}$.
- Probability mass function (PMF or pmf) of X: $\left\{P_X[X = x] \ \forall \ x \in \mathcal{X}\right\}$
- Cumulative distribution function (CDF or cdf) of X, denoted by $F_X(.)$ is defined as

$$F_X(x) := P_X[X \leq x] \quad \text{where} \quad x \in \mathbb{R}$$

- Examples:
  - Bernoulli(p): Toss a coin s.t. probability of heas is p
  - Binomial(n, p): Toss a coin n times independently with k number of heads
  - Geometric(p): Keep tossing a coin till head occurs
  - Poisson(λ): Number of people standing in a queue

Properties of CDF $F_x(.)$

- $F_x(.)$ is monotonically nondecreasing.
- $F_x(-\infty) = 0$ and $F_x(\infty) = 1$.
- $F_x(.)$ is right continuous.

Proofs: In class

Continuity of a function

- Left limit of a function $f(x)$ at a point x_0 is defined as

$$\text{Left limit} := \lim_{x_n \uparrow x_0} f(x_n)$$

- Right limit of a function $f(x)$ at a point x_0 is defined as

$$\text{Right limit} := \lim_{x_n \downarrow x_0} f(x_n)$$

- When $f(x)$ is said to be left-continuous, right-continuous, continuous, and discontinuous at x_0?

- Show that $F_X(x) := P_X[X \leq x]$ is right continuous.

Expected value (or mean) and Variance of a discrete r.v.

Expected value of a discrete r.v.

- Consider a discrete r.v. X with support set $\mathcal{X}$.
- Expected value (or mean) of X, denoted by $\mathbb{E}[X]$, is defined as

$$\mathbb{E}[X] := \sum_{x \in \mathcal{X}} x P_X(X = x)$$

  - Find $\mathbb{E}[X]$ for Bernoulli(p) r.v. ($P_X(X = 1) = p$ and $P_X(X = 0) = 1 - p$).
  - Find $\mathbb{E}[X]$ of a r.v. X with $\mathcal{X} = \{1, 2, 3\}$ and PMF $\{0.25, 0.5, 0.25\}$
- Expected value of function $f(X) := f \circ X$ of X is defined as

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

  - Find $\mathbb{E}[X^2]$ for above examples.
- For r.v.s X, Y and constants a, b, we have $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$.

Variance of a discrete r.v.

- Consider a discrete r.v. X with support set $\mathcal{X}$.
- Variance of X, denoted by $V[X]$, is defined as

$$V[X] := \mathbb{E}\Big[\big(X - \mathbb{E}[X]\big)^2\Big] = \sum_{x \in \mathcal{X}} \big(x - \mathbb{E}[X]\big)^2 P_X(X = x)$$

- We will show that (in class)

$$V[X] = \mathbb{E}[X^2] - (\mathbb{E}(X))^2$$

- Examples:
  - Find $V[X]$ for Bernoulli(p) r.v. ($P_X(X = 1) = p$ and $P_X(X = 0) = 1 - p$).
  - Find $V[X]$ of a r.v. X with $\mathcal{X} = \{1, 2, 3\}$ and PMF $\{0.25, 0.5, 0.25\}$

Continuous random variables

Definition of a continuous r.v.

- A random variable X is called continuous if its probability law can be described in terms of a nonnegative function $f_X(.)$, called the probability density function (PDF) of X, which satisfies

$$P_X\Big[X \in B\Big] = \int_B f_X(x)\mathrm{d}x,$$

for every subset B of the real line.

- Thus the probability that $X \in [a, b]$ will be

$$P_X\Big[X \in [a, b]\Big] = \int_a^b f_X(x)\mathrm{d}x,$$

- What will be $P_X[X = a]$?
- When do I say that $f_X(.)$ is a valid PDF?

Relationship between PDF and CDF

- Recall: For a continuous r.v. X with PDF $f_X(x)$ we have

$$P_X\Big[X \in [a, b]\Big] = \int_a^b f_X(x)\mathrm{d}x,$$

- CDF of X will be

$$F_X(x) \coloneqq P_X\Big[X \leq x\Big] = P_X\Big[X \in (-\infty, x]\Big] = \int_{-\infty}^{x} f_X(x)\mathrm{d}x,$$

- PDF can be expressed in terms of CDF as

$$f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x}$$

Mean and Variance of a continuous r.v.

- Consider a continuous r.v. X with PDF $f_X(x)$.
- Expected value of X is defined as

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) \mathrm{d}x,$$

- Variance of X is defined as

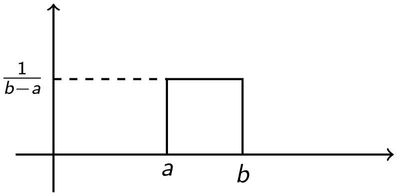
$$\begin{aligned} V[X] &:= \mathbb{E}\left[\left(X - \mathbb{E}[X]\right)^2\right] \\ &= \int_{-\infty}^{\infty} \left(x - \mathbb{E}[X]\right)^2 f_X(x) \mathrm{d}x \\ &= \int_{-\infty}^{\infty} \left(x - \mu\right)^2 f_X(x) \mathrm{d}x \\ &= \mathbb{E}[X^2] - \mu^2 \\ &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \end{aligned}$$

Examples of continuous r.v.s

Uniform random variable

- PDF of a uniform r.v. with the support set $[a, b]$ is given by

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

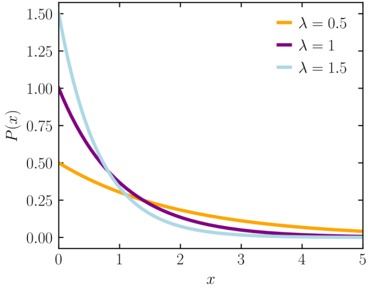


- Is this a a valid PDF?
- Find mean and variance.

Exponential random variable

- PDF of an exponential r.v. with parameter λ is given by

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if} \quad x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

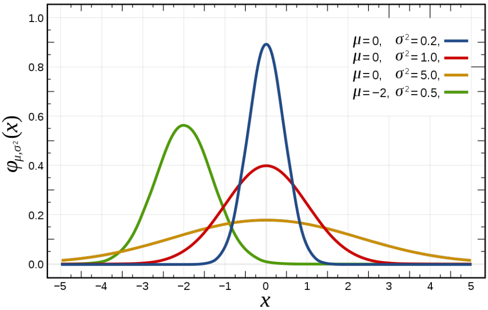


- Is this a a valid PDF?
- Find mean and variance.

Gaussian random variable

- PDF of a Gaussian r.v. with parameters μ and σ is given by

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



- Is this a a valid PDF?
- Find mean and variance.

  Nice link: https://www.youtube.com/watch?v=I27xKSNad2Y

Summary

Summary

- Definition of a continuous random variable
- Expected value (or mean) and variance of a random variable
- Examples:
  - Uniform r.v.
  - Exponential r.v.
  - Gaussian r.v.

Recap of the previous class

Random variables: Discrete vs continuous

- Discrete random variables:
  - Probability mass function (PMF or pmf) of X: $\Big\{ P_X[X = x] \ \forall \ x \in \mathcal{X} \Big\}$
  - $\mathbb{E}[X]$ = Expected value (or mean) of $X := \sum_{x \in \mathcal{X}} x P_X(X = x)$
  - $V[X]$ = Variance of $X := \mathbb{E}\Big[\big(X - \mathbb{E}[X]\big)^2\Big] \ = \mathbb{E}[X^2] - (\mathbb{E}(X))^2$
- Continuous random variables:
  - Probability law of a continuous r.v. X is described using a funtion $f_X(.)$, called as probability density function (PDF or pdf) of X, which satisfies

$$P_X\Big[X \in B\Big] = \int_B f_X(x)\mathrm{d}x, \quad \text{where } B \in \mathcal{B}(\mathbb{R})$$

  - $\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\mathrm{d}x$
  - $V[X] = \mathbb{E}[X^2] - (\mathbb{E}(X))^2 \ = \int_{-\infty}^{\infty} x^2 f_X(x)\mathrm{d}x - \Big(\int_{-\infty}^{\infty} x f_X(x)\mathrm{d}x\Big)^2$

Cumulative distribution function

- Cumulative distribution function (CDF or cdf) of X, denoted by $F_X(.)$ is defined as

$$F_X(x) := P_X[X \leq x] \quad \text{where} \quad x \in \mathbb{R}$$

- Discrete random variable:

$$F_X(x) = \sum_{i=-\infty}^{x} P_X[X = i]$$

- Continuous random variable:

$$F_X(x) = \int_{-\infty}^{x} f_X(x) \mathrm{d}x$$

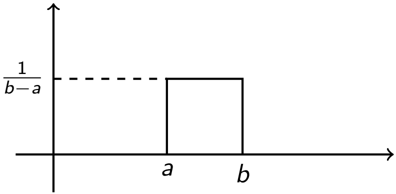
- PDF can be expressed in terms of CDF as $f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x}$.

Examples of continuous r.v.s

Uniform random variable

- PDF of a uniform r.v. with the support set $[a, b]$ is given by

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

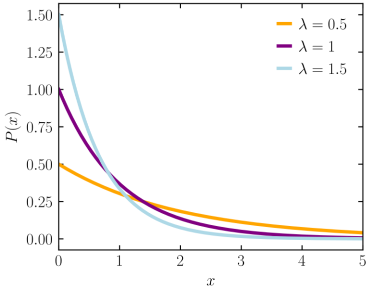


- Mean $\mathbb{E}[X] = (a + b)/2$ and variance $V[X] = (b - a)^2/12$.

Exponential random variable

- PDF of an exponential r.v. with parameter λ is given by

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if} \quad x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

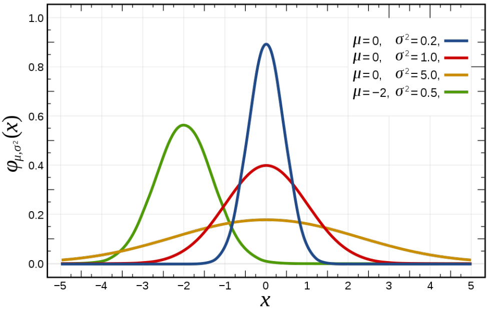


- Mean $\mathbb{E}[X] = 1/\lambda$ and variance $V[X] = 1/\lambda^2$.

Gaussian random variable

- PDF of a Gaussian r.v. with parameters μ and σ is given by

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



- Is this a a valid PDF?
- Mean $\mathbb{E}[X] = \mu$ and variance $V[X] = \sigma^2$. Link

Joint random variables

Joint random variables (discrete)

- Consider two discrete random variables X and Y associated with the probability space $(\Omega, \mathcal{F}, P)$.
    - Definition of joint random variables X and Y
    - Joint PMF
    - Joint CDF
    - Marginal PMFs from joint PMF
    - Independence
- Can you extend these concepts to n random variables $X_1, X_2, \ldots, X_n$?
- Simplified notation

Joint random variables: Example

- Find the marginal PMFs for the following examples.

$X \backslash Y$	0	1
0	$\frac{1}{3}$	$\frac{1}{3}$
1	0	$\frac{1}{3}$

$Y \backslash X$	1	2
1	0	$\frac{3}{4}$
2	$\frac{1}{8}$	$\frac{1}{8}$

$Y \backslash X$	1	2	3	4
1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$
2	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$
3	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
4	$\frac{1}{4}$	0	0	0

Joint random variables (continuous)

- Consider two continuous random variables X and Y associated with the probability space $(\Omega, \mathcal{F}, P)$.
  - Definition of joint random variables X and Y
  - Joint PDF
  - Joint CDF
  - Marginal PMFs from joint PMF
  - Independence
- Can you extend these concepts to n random variables $X_1, X_2, \ldots, X_n$?
- Simplified notation
- Revise basics of integration!!

Self quiz

- Consider a random vector $\underline{X} = \begin{bmatrix} X_1 & X_2 & X_3 & X_4 & X_5 & X_6 & X_7 \end{bmatrix}$ with PDF $f(\underline{x})$. Which of the following statement is/are incorrect?

1. Suppose $g(\cdot) = \int_{x_1}\int_{x_3}\int_{x_7} f(\underline{x})\mathrm{d}x_1\mathrm{d}x_3\mathrm{d}x_7$. Then $g(\cdot)$ is a function of variables x_2, x_4, x_5, x_6.

2. $f(x_1, x_2, x_4, x_5, x_7) = \int_{x_6}\int_{x_3} f(\underline{x})\mathrm{d}x_6\mathrm{d}x_3$

3. $f(x_3, x_7) = \int_{x_1}\int_{x_2}\int_{x_4}\int_{x_5}\int_{x_6} f(\underline{x})\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_4\mathrm{d}x_5\mathrm{d}x_6$

4. $f(x_4) = \int_{x_4} f(\underline{x})\mathrm{d}x_4$

Properties of joint CDF

- Consider two random variables X and Y associated with the same experiment.
- The joint CDF is defined as

$$\begin{aligned} F_{X,Y}(x,y) &:= P_{X,Y}\left[X \leq x, Y \leq y\right] \\ &= P\left[w \middle| X(w) \leq x, Y(w) \leq y\right]\right] \end{aligned}$$

- Properties of joint CDF
  - $F_{X,Y}(.)$ should be non-decreasing and right continuous in both variables.
  - $F_{X,Y}(\infty,\infty) =?$, $F_{X,Y}(-\infty,-\infty) =?$

    $F_{X,Y}(-\infty,y) =?$, $F_{X,Y}(x,-\infty) =?$

    $F_{X,Y}(\infty,y) =?$, $F_{X,Y}(x,\infty) =?$

Joint random variables: Examples

Joint RVs (Discrete): Example

- Find $\mathbb{E}[X + Y]$

$X \backslash Y$	0	1
0	$\frac{1}{3}$	$\frac{1}{3}$
1	0	$\frac{1}{3}$

$Y \backslash X$	1	2	3	4
1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$
2	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$
3	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
4	$\frac{1}{4}$	0	0	0

Joint RVs (Discrete): Example

- Random variables X and Y are independent and identically distribution according to Bernoulli(p) distribution. Find joint PMF.

Joint RVs (Continuous): Example

- Consider the following joint PDF.

$$f(x, y) = \begin{cases} c & \text{if } \; 2000 \leq x, y \leq 2200 \\ 0 & \text{otherwise} \end{cases}$$

- Is this a valid joint PDF? What should be the value of constant c?

- Find $P_{XY}\Big(|X - Y| \leq 20\Big)$

Joint RVs (Continuous): Homework/Tutorial

- Consider the following joint PDF.

$$f(x,y) = \begin{cases} 0.64e^{-0.8y} & \text{if } \; 0 < x < y \\ 0 & \text{otherwise} \end{cases}$$

- Is this a valid joint PDF?

- Find $P_{XY}\left[1 < X < 2 \;\text{ and }\; 1 < Y < 2\right]$

- Does this PDF look familiar? BONUS applause!

Recap of the previous class

Recap

- Notation: $P_X(X=x)$, $P(X=x)$, $P_X(x)$, $p(x)$ are the same. Similarly, $f_X(x)$ and $f(x)$ are the same.
- Joint distribution (pmf/pdf) of two RVs X and Y
- Marginal distribution of two RVs X and Y, given their joint distribution
- Let $\underline{X} = \begin{bmatrix} X_1 & X_2 & \dots & X_n \end{bmatrix}$ be a random vector and $\underline{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}$ be its realization. Note that $X_1, X_2, \dots, X_n$ are a set of n RVs.
  - For discrete RVs, the joint pmf is given by

$$P\Big(\underline{X} = \underline{x}\Big) = P\Big(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\Big) \forall \underline{x} \in \mathcal{X}_1 \times \mathcal{X}_2 \times \dots \times \mathcal{X}_n$$

  - For continuous RVs, suppose $f(\underline{x})$ denotes the joint pdf. Then we have

$$P\Big(X_1 \in [a_1, b_1], \dots, X_n \in [a_n, b_n]\Big) = \int_{x_1=a_1}^{b_1} \dots \int_{x_n=a_n}^{b_n} f(x_1, \dots, x_n) \mathrm{d}x_1 \dots \mathrm{d}x_n$$

Discrete uniform RV

- Let X be a discrete RV with support set $\mathcal{X} = \{a_1, a_2, \ldots, a_n\}$. Suppose the pmf of X is given by

$$P_X(X = a_i) = \frac{1}{n} \text{ for any } \; a_i \in \mathcal{X}$$

- X is said to follow uniform distribution.
- Cardinality of a set $\mathcal{X}$ is defined as the number of elements in the set, denoted by $|\mathcal{X}|$.
- For a uniform RV we have,

$$P_X(X = x) = \frac{1}{|\mathcal{X}|} \text{ for any } \; x \in \mathcal{X}$$

Conditional distribution

Conditioning one random variable on another

- Conditional distribution of X given $Y = y$:

$$P_{X|Y}(X = x|Y = y) = \frac{P_{X,Y}(X = x, Y = y)}{P_Y(Y = y)}$$

- Note: $P_{X|Y}(x|y)$ is defined only for those values of y s.t. $P_Y(Y = y) > 0$.
- What will be the value of $\sum_{x \in \mathcal{X}} P_{X|Y}(X = x|Y = y)$?
- Can you define conditional expectation?
- Example: Find the conditional distribution of $X|Y = 0$ and $X|Y = 1$.

$X \backslash Y$	0	1
0	$\frac{1}{3}$	$\frac{1}{3}$
1	0	$\frac{1}{3}$

Conditional RVs: Example

- Four dice are rolled independently. Let X be the number of 1's and Y be the number of 2's. Find the joint PMF of X and Y.

My office hours

- Office hours:
  - Tuesday (5 to 6pm)
  - Friday (2 to 3pm)
- Email: arti.yardi@iiit.ac.in
- Office: Vindhya A2, 1st floor, SPCRC + CQST lab, Office no. 4

Reference books

- “Probabiilty and measure” by P. Billingsley
- “Probabiilty, Random variables, and Stochastic processes”, by A. Papoulis