EE 325: Probability and Random Processes Module 2: Random Variables and Probability Distributions

D Manjunath

Networking Lab
Department of Electrical Engineering
Indian Institute of Technology

September 14, 2020

Topics in Module 2

- Definition of a random variable (RV).
- Cumulative distribution functions (cdf) and its properties.
- Discrete random variables, probability mass function (pmf) with examples.
- Continuous random variables, probability density function (pdf) with examples.
- Joint and conditional distributions with examples.
- Mostly from Chapter 4 of the text; Sections 4.1–4.4.

Topics in Module 2

- Definition of a random variable (RV).
- Cumulative distribution functions (cdf) and its properties.
- Discrete random variables, probability mass function (pmf) with examples.
- Continuous random variables, probability density function (pdf) with examples.
- Joint and conditional distributions with examples.
- Mostly from Chapter 4 of the text; Sections 4.1–4.4.

Topics in Module 2

- Definition of a random variable (RV).
- Cumulative distribution functions (cdf) and its properties.
- Discrete random variables, probability mass function (pmf) with examples.
- Continuous random variables, probability density function (pdf) with examples.
- Joint and conditional distributions with examples.
- Mostly from Chapter 4 of the text; Sections 4.1–4.4.

Topics in Module 2

- Definition of a random variable (RV).
- Cumulative distribution functions (cdf) and its properties.
- Discrete random variables, probability mass function (pmf) with examples.
- Continuous random variables, probability density function (pdf) with examples.
- Joint and conditional distributions with examples.
- Mostly from Chapter 4 of the text; Sections 4.1–4.4.

Topics in Module 2

- Definition of a random variable (RV).
- Cumulative distribution functions (cdf) and its properties.
- Discrete random variables, probability mass function (pmf) with examples.
- Continuous random variables, probability density function (pdf) with examples.
- Joint and conditional distributions with examples.
- Mostly from Chapter 4 of the text; Sections 4.1–4.4.

Topics in Module 2

- Definition of a random variable (RV).
- Cumulative distribution functions (cdf) and its properties.
- Discrete random variables, probability mass function (pmf) with examples.
- Continuous random variables, probability density function (pdf) with examples.
- Joint and conditional distributions with examples.
- Mostly from Chapter 4 of the text; Sections 4.1–4.4.

Random Variables: Motivation and Background

- Recall that in a random experiment, the outcome can be a concrete object—a person, a lot of a manufactured item, a mango from a box of mangoes, a scooter on the road, ...
- These objects will possess some qualities that are measurable—height or weight of a person, the number of defective components in the lot, the amount pulp that you can get from the mango, the number of people on the scooter, ...
- We want to model these measurable attributes, i.e., we would like to deal with numbers associated with the objects, rather than the objects themselves.
- Examples
  - On a coin toss, say outcome is 1 if it is a Head and 0 if it is a Tail.
  - Alternately, *map* outcome of Head to +1 and Tail to -1.
  - Throw two dice and consider only sums of the two throws.

Random Variables: Motivation and Background

- Recall that in a random experiment, the outcome can be a concrete object—a person, a lot of a manufactured item, a mango from a box of mangoes, a scooter on the road, ...
- These objects will possess some qualities that are measurable—height or weight of a person, the number of defective components in the lot, the amount pulp that you can get from the mango, the number of people on the scooter, ...
- We want to model these measurable attributes, i.e., we would like to deal with numbers associated with the objects, rather than the objects themselves.
- Examples
  - On a coin toss, say outcome is 1 if it is a Head and 0 if it is a Tail.
  - Alternately, *map* outcome of Head to +1 and Tail to -1.
  - Throw two dice and consider only sums of the two throws.

Random Variables: Motivation and Background

- Recall that in a random experiment, the outcome can be a concrete object—a person, a lot of a manufactured item, a mango from a box of mangoes, a scooter on the road, ...
- These objects will possess some qualities that are measurable—height or weight of a person, the number of defective components in the lot, the amount pulp that you can get from the mango, the number of people on the scooter, ...
- We want to model these measurable attributes, i.e., we would like to deal with numbers associated with the objects, rather than the objects themselves.
- Examples
  - On a coin toss, say outcome is 1 if it is a Head and 0 if it is a Tail.
  - Alternately, *map* outcome of Head to +1 and Tail to -1.
  - Throw two dice and consider only sums of the two throws.

Random Variables: Motivation and Background

- Recall that in a random experiment, the outcome can be a concrete object—a person, a lot of a manufactured item, a mango from a box of mangoes, a scooter on the road, ...
- These objects will possess some qualities that are measurable—height or weight of a person, the number of defective components in the lot, the amount pulp that you can get from the mango, the number of people on the scooter, ...
- We want to model these measurable attributes, i.e., we would like to deal with numbers associated with the objects, rather than the objects themselves.
- Examples
  - On a coin toss, say outcome is 1 if it is a Head and 0 if it is a Tail.
  - Alternately, *map* outcome of Head to +1 and Tail to -1.
  - Throw two dice and consider only sums of the two throws.

Random Variables: Motivation and Background

- Recall that in a random experiment, the outcome can be a concrete object—a person, a lot of a manufactured item, a mango from a box of mangoes, a scooter on the road, ...
- These objects will possess some qualities that are measurable—height or weight of a person, the number of defective components in the lot, the amount pulp that you can get from the mango, the number of people on the scooter, ...
- We want to model these measurable attributes, i.e., we would like to deal with numbers associated with the objects, rather than the objects themselves.
- Examples
    - On a coin toss, say outcome is 1 if it is a Head and 0 if it is a Tail.
    - Alternately, *map* outcome of Head to +1 and Tail to -1.
    - Throw two dice and consider only sums of the two throws.

Random Variables: Motivation and Background

- Recall that in a random experiment, the outcome can be a concrete object—a person, a lot of a manufactured item, a mango from a box of mangoes, a scooter on the road, ...
- These objects will possess some qualities that are measurable—height or weight of a person, the number of defective components in the lot, the amount pulp that you can get from the mango, the number of people on the scooter, ...
- We want to model these measurable attributes, i.e., we would like to deal with numbers associated with the objects, rather than the objects themselves.
- Examples
  - On a coin toss, say outcome is 1 if it is a Head and 0 if it is a Tail.
  - Alternately, *map* outcome of Head to +1 and Tail to -1.
  - Throw two dice and consider only sums of the two throws.

Random Variables: Motivation and Background

- Recall that in a random experiment, the outcome can be a concrete object—a person, a lot of a manufactured item, a mango from a box of mangoes, a scooter on the road, ...
- These objects will possess some qualities that are measurable—height or weight of a person, the number of defective components in the lot, the amount pulp that you can get from the mango, the number of people on the scooter, ...
- We want to model these measurable attributes, i.e., we would like to deal with numbers associated with the objects, rather than the objects themselves.
- Examples
  - On a coin toss, say outcome is 1 if it is a Head and 0 if it is a Tail.
  - Alternately, *map* outcome of Head to +1 and Tail to -1.
  - Throw two dice and consider only sums of the two throws.

Random Variables: Motivation and Background

- Examples (contd)
  - Pick a coin, not necessarily a fair coin, and toss it a number of times and note the sequence. Pick an arbitray point in the sequence, say ω, and count the number of tosses till the next Head, $\mathsf{H}(\omega)$.
  - For a video, say ω, downloaded from the 'Net,' the time that it takes for you to download it, $\mathsf{T}(\omega)$. A different map could be its size $\mathsf{S}(\omega)$.
  - Pick a person from the population, say ω, (outcome) and note her/his weight ($\mathsf{W}(\omega)$), height ($\mathsf{H}(\omega)$), income $\mathsf{I}(\omega)$, number of members in the person's family ($\mathsf{N}(\omega)$), ...
- In all of the above, we have converted outcomes of a random experiment to numbers (integers or real), i.e., we have mapped an outcome of an experiment to a numerical value.
- In other words, the random variable is the numerical value of the outcome of a random experiment.

- Examples (contd)
  - Pick a coin, not necessarily a fair coin, and toss it a number of times and note the sequence. Pick an arbitray point in the sequence, say ω, and count the number of tosses till the next Head, $\mathsf{H}(\omega)$.
  - For a video, say ω, downloaded from the 'Net,' the time that it takes for you to download it, $\mathsf{T}(\omega)$. A different map could be its size $\mathsf{S}(\omega)$.
  - Pick a person from the population, say ω, (outcome) and note her/his weight ($\mathsf{W}(\omega)$), height ($\mathsf{H}(\omega)$), income $\mathsf{I}(\omega)$, number of members in the person's family ($\mathsf{N}(\omega)$), ...
- In all of the above, we have converted outcomes of a random experiment to numbers (integers or real), i.e., we have mapped an outcome of an experiment to a numerical value.
- In other words, the random variable is the numerical value of the outcome of a random experiment.

Random Variables: Motivation and Background

- Examples (contd)
  - Pick a coin, not necessarily a fair coin, and toss it a number of times and note the sequence. Pick an arbitray point in the sequence, say ω, and count the number of tosses till the next Head, $\mathsf{H}(\omega)$.
  - For a video, say ω, downloaded from the 'Net,' the time that it takes for you to download it, $\mathsf{T}(\omega)$. A different map could be its size $\mathsf{S}(\omega)$.
  - Pick a person from the population, say ω, (outcome) and note her/his weight ($\mathsf{W}(\omega)$), height ($\mathsf{H}(\omega)$), income $\mathsf{I}(\omega)$, number of members in the person's family ($\mathsf{N}(\omega)$), ...
- In all of the above, we have converted outcomes of a random experiment to numbers (integers or real), i.e., we have mapped an outcome of an experiment to a numerical value.
- In other words, the random variable is the numerical value of the outcome of a random experiment.

- Examples (contd)
  - Pick a coin, not necessarily a fair coin, and toss it a number of times and note the sequence. Pick an arbitray point in the sequence, say ω, and count the number of tosses till the next Head, $\mathsf{H}(\omega)$.
  - For a video, say ω, downloaded from the 'Net,' the time that it takes for you to download it, $\mathsf{T}(\omega)$. A different map could be its size $\mathsf{S}(\omega)$.
  - Pick a person from the population, say ω, (outcome) and note her/his weight ($\mathsf{W}(\omega)$), height ($\mathsf{H}(\omega)$), income $\mathsf{I}(\omega)$, number of members in the person's family ($\mathsf{N}(\omega)$), ...
- In all of the above, we have converted outcomes of a random experiment to numbers (integers or real), i.e., we have mapped an outcome of an experiment to a numerical value.
- In other words, the random variable is the numerical value of the outcome of a random experiment.

- Examples (contd)
    - Pick a coin, not necessarily a fair coin, and toss it a number of times and note the sequence. Pick an arbitray point in the sequence, say ω, and count the number of tosses till the next Head, $\mathsf{H}(\omega)$.
    - For a video, say ω, downloaded from the 'Net,' the time that it takes for you to download it, $\mathsf{T}(\omega)$. A different map could be its size $\mathsf{S}(\omega)$.
    - Pick a person from the population, say ω, (outcome) and note her/his weight ($\mathsf{W}(\omega)$), height ($\mathsf{H}(\omega)$), income $\mathsf{I}(\omega)$, number of members in the person's family ($\mathsf{N}(\omega)$), ...
- In all of the above, we have converted outcomes of a random experiment to numbers (integers or real), i.e., we have mapped an outcome of an experiment to a numerical value.
- In other words, the random variable is the numerical value of the outcome of a random experiment.

- Examples (contd)
    - Pick a coin, not necessarily a fair coin, and toss it a number of times and note the sequence. Pick an arbitray point in the sequence, say ω, and count the number of tosses till the next Head, $\mathsf{H}(\omega)$.
    - For a video, say ω, downloaded from the 'Net,' the time that it takes for you to download it, $\mathsf{T}(\omega)$. A different map could be its size $\mathsf{S}(\omega)$.
    - Pick a person from the population, say ω, (outcome) and note her/his weight ($\mathsf{W}(\omega)$), height ($\mathsf{H}(\omega)$), income $\mathsf{I}(\omega)$, number of members in the person's family ($\mathsf{N}(\omega)$), ...
- In all of the above, we have converted outcomes of a random experiment to numbers (integers or real), i.e., we have mapped an outcome of an experiment to a numerical value.
- In other words, the random variable is the numerical value of the outcome of a random experiment.

- May have more complex forms
  - $Z = aH(\omega) + bW(\omega)$
  - Let $N(\omega)$ be the number of members in the family of ω and $I(\omega)$ the family income. Interested in

$$Z := I(\omega)/N(\omega).$$

- In the preceding examples, we have defined two separate random variables on Ω and then the random variable of interest is a function of the two random variables–H and W in the first case, and I and N in the second case.

Random Variables: Motivation and Background

- May have more complex forms
  - $\mathsf{Z} = a\mathsf{H}(\omega) + b\mathsf{W}(\omega)$
  - Let $\mathsf{N}(\omega)$ be the number of members in the family of ω and $\mathsf{I}(\omega)$ the family income. Interested in

$$\mathsf{Z} := \mathsf{I}(\omega)/\mathsf{N}(\omega).$$

- In the preceding examples, we have defined two separate random variables on Ω and then the random variable of interest is a function of the two random variables–H and W in the first case, and I and N in the second case.

- May have more complex forms
  - $\mathsf{Z} = a\mathsf{H}(\omega) + b\mathsf{W}(\omega)$
  - Let $\mathsf{N}(\omega)$ be the number of members in the family of ω and $\mathsf{I}(\omega)$ the family income. Interested in

$$\mathsf{Z} := \mathsf{I}(\omega)/\mathsf{N}(\omega).$$

- In the preceding examples, we have defined two separate random variables on Ω and then the random variable of interest is a function of the two random variables–H and W in the first case, and I and N in the second case.

Random Variables: Motivation and Background

- May have more complex forms
  - $\mathsf{Z} = a\mathsf{H}(\omega) + b\mathsf{W}(\omega)$
  - Let $\mathsf{N}(\omega)$ be the number of members in the family of ω and $\mathsf{I}(\omega)$ the family income. Interested in

$$\mathsf{Z} := \mathsf{I}(\omega)/\mathsf{N}(\omega).$$

- In the preceding examples, we have defined two separate random variables on Ω and then the random variable of interest is a function of the two random variables–H and W in the first case, and I and N in the second case.

Random Variables: Definition

- A **random variable** is a numerical valued function defined on Ω, i.e., a random variable maps the elements of Ω to a number.

$$\omega \in \Omega : \quad \omega \to \mathsf{X}(\omega)$$
$$\mathsf{X}(\omega) : \omega \to \mathcal{Z}$$
$$\mathsf{X}(\omega) : \omega \to \Re$$

- $\Re$ is the set of real numbers and $\mathcal{Z}$ is the set of integers.
- Can define **joint functions** on Ω :

$$\omega \in \Omega : \quad \omega \to (\mathsf{X}(\omega), \mathsf{Y}(\omega))$$

- We will usually omit the reference to ω and just say X, Y, etc.
- We will be interested in events generated by a random variable of the form $\{\mathsf{X} \leq x\}$.

Random Variables: Definition

- A **random variable** is a numerical valued function defined on Ω, i.e., a random variable maps the elements of Ω to a number.

$$\omega \in \Omega : \quad \omega \rightarrow \mathsf{X}(\omega)$$
$$\mathsf{X}(\omega) : \omega \rightarrow \mathcal{Z}$$
$$\mathsf{X}(\omega) : \omega \rightarrow \Re$$

- $\Re$ is the set of real numbers and $\mathcal{Z}$ is the set of integers.
- Can define **joint functions** on Ω :

$$\omega \in \Omega : \quad \omega \rightarrow (\mathsf{X}(\omega), \mathsf{Y}(\omega))$$

- We will usually omit the reference to ω and just say X, Y, etc.
- We will be interested in events generated by a random variable of the form $\{\mathsf{X} \leq x\}$.

Random Variables: Definition

- A **random variable** is a numerical valued function defined on Ω, i.e., a random variable maps the elements of Ω to a number.

$$\omega \in \Omega : \quad \omega \rightarrow \mathsf{X}(\omega)$$
$$\mathsf{X}(\omega) : \omega \rightarrow \mathcal{Z}$$
$$\mathsf{X}(\omega) : \omega \rightarrow \Re$$

- $\Re$ is the set of real numbers and $\mathcal{Z}$ is the set of integers.
- Can define **joint functions** on Ω :

$$\omega \in \Omega : \quad \omega \rightarrow (\mathsf{X}(\omega), \mathsf{Y}(\omega))$$

- We will usually omit the reference to ω and just say X, Y, etc.
- We will be interested in events generated by a random variable of the form $\{\mathsf{X} \leq x\}$.

Random Variables: Definition

- A **random variable** is a numerical valued function defined on Ω, i.e., a random variable maps the elements of Ω to a number.

$$\begin{aligned} &\omega \in \Omega: \quad \omega \to \mathsf{X}(\omega) \\ &\mathsf{X}(\omega) : \omega \to \mathcal{Z} \\ &\mathsf{X}(\omega) : \omega \to \Re \end{aligned}$$

- $\Re$ is the set of real numbers and $\mathcal{Z}$ is the set of integers.
- Can define **joint functions** on Ω :

$$\omega \in \Omega: \quad \omega \to (\mathsf{X}(\omega), \mathsf{Y}(\omega))$$

- We will usually omit the reference to ω and just say X, Y, etc.
- We will be interested in events generated by a random variable of the form $\{\mathsf{X} \leq x\}$.

Random Variables: Definition

- A **random variable** is a numerical valued function defined on Ω, i.e., a random variable maps the elements of Ω to a number.

$$\omega \in \Omega : \quad \omega \to \mathsf{X}(\omega)$$
$$\mathsf{X}(\omega) : \omega \to \mathcal{Z}$$
$$\mathsf{X}(\omega) : \omega \to \Re$$

- $\Re$ is the set of real numbers and $\mathcal{Z}$ is the set of integers.
- Can define **joint functions** on Ω :

$$\omega \in \Omega : \quad \omega \to (\mathsf{X}(\omega), \mathsf{Y}(\omega))$$

- We will usually omit the reference to ω and just say X, Y, etc.
- We will be interested in events generated by a random variable of the form $\{\mathsf{X} \leq x\}$.

Cumulative Distribution Function

- Let X be a random variable and A a subset of the number line.

$$\mathsf{Prob}(\mathsf{X} \in A) = \mathsf{Prob}(\omega : \mathsf{X}(\omega) \in A)$$

- Define

$$F_{\mathsf{X}}(x) = \mathsf{Prob}(\mathsf{X} \leq x)$$

- $F_{\mathsf{X}}(x)$ is called the **Cumulative Distribution Function (CDF)** of X.
- $F_{\mathsf{X}}(x)$ is also called the **Probability Distribution Function (PDF)** of X, or simply the distribution, of X.

Cumulative Distribution Function

- Let X be a random variable and A a subset of the number line.

$$\mathsf{Prob}(\mathsf{X} \in A) = \mathsf{Prob}(\omega : \mathsf{X}(\omega) \in A)$$

- Define

$$F_{\mathsf{X}}(x) = \mathsf{Prob}(\mathsf{X} \leq x)$$

- $F_{\mathsf{X}}(x)$ is called the **Cumulative Distribution Function (CDF)** of X.
- $F_{\mathsf{X}}(x)$ is also called the **Probability Distribution Function (PDF)** of X, or simply the distribution, of X.

Cumulative Distribution Function

- Let X be a random variable and A a subset of the number line.

$$\mathsf{Prob}(\mathsf{X} \in A) = \mathsf{Prob}(\omega : \mathsf{X}(\omega) \in A)$$

- Define

$$F_{\mathsf{X}}(x) = \mathsf{Prob}(\mathsf{X} \leq x)$$

- $F_{\mathsf{X}}(x)$ is called the **Cumulative Distribution Function (CDF)** of X.
- $F_{\mathsf{X}}(x)$ is also called the **Probability Distribution Function (PDF)** of X, or simply the distribution, of X.

Cumulative Distribution Function

- Let X be a random variable and A a subset of the number line.

$$\mathsf{Prob}(\mathsf{X} \in A) = \mathsf{Prob}(\omega : \mathsf{X}(\omega) \in A)$$

- Define

$$F_{\mathsf{X}}(x) = \mathsf{Prob}(\mathsf{X} \leq x)$$

- $F_{\mathsf{X}}(x)$ is called the **Cumulative Distribution Function (CDF)** of X.
- $F_{\mathsf{X}}(x)$ is also called the **Probability Distribution Function (PDF)** of X, or simply the distribution, of X.

Cumulative Distribution Function

- Let X be a random variable and A a subset of the number line.

$$\mathsf{Prob}(\mathsf{X} \in A) = \mathsf{Prob}(\omega : \mathsf{X}(\omega) \in A)$$

- Define

$$F_{\mathsf{X}}(x) = \mathsf{Prob}(\mathsf{X} \le x)$$

- $F_{\mathsf{X}}(x)$ is called the **Cumulative Distribution Function (CDF)** of X.
- $F_{\mathsf{X}}(x)$ is also called the **Probability Distribution Function (PDF)** of X, or simply the distribution, of X.

Cumulative Distribution Function

- Let X be a random variable and A a subset of the number line.

$$\text{Prob}(\mathsf{X} \in A) = \text{Prob}(\omega : \mathsf{X}(\omega) \in A)$$

- Define

$$F_{\mathsf{X}}(x) = \text{Prob}(\mathsf{X} \leq x)$$

- $F_{\mathsf{X}}(x)$ is called the **Cumulative Distribution Function (CDF)** of X.
- $F_{\mathsf{X}}(x)$ is also called the **Probability Distribution Function (PDF)** of X, or simply the distribution, of X.

Properties of CDF

- $F_{\mathsf{X}}(-\infty) = 0.$
- $F_{\mathsf{X}}(\infty) = 1.$
- If $x_1 \leq x_2$, then $F_{\mathsf{X}}(x_2) \geq F_{\mathsf{X}}(x_2)$
  Follows because
  $$\{\mathsf{X} \leq x_1\} \subseteq \{\mathsf{X} \leq x_2\}$$
  This means that $F_{\mathsf{X}}(x)$ is a non decreasing function
- $\mathsf{Prob}(x_1 < \mathsf{X} \leq x_2) = F_{\mathsf{X}}(x_2) - F_{\mathsf{X}}(x_1).$
- $\mathsf{Prob}(\mathsf{X} > x) = 1 - F_{\mathsf{X}}(x)$
- $F_{\mathsf{X}}(x)$ is continuous from the right, i.e., $\lim_{x \downarrow x_0} F_{\mathsf{X}}(x) = F_{\mathsf{X}}(x_0)$.
  In other words, $F_{\mathsf{X}}(x^+) = F_{\mathsf{X}}(x)$.

Properties of CDF

- $F_{\mathsf{X}}(-\infty) = 0.$
- $F_{\mathsf{X}}(\infty) = 1.$
- If $x_1 \leq x_2$, then $F_{\mathsf{X}}(x_2) \geq F_{\mathsf{X}}(x_2)$
  Follows because
  $$\{\mathsf{X} \leq x_1\} \subseteq \{\mathsf{X} \leq x_2\}$$
  This means that $F_{\mathsf{X}}(x)$ is a non decreasing function
- $\mathsf{Prob}(x_1 < \mathsf{X} \leq x_2) = F_{\mathsf{X}}(x_2) - F_{\mathsf{X}}(x_1).$
- $\mathsf{Prob}(\mathsf{X} > x) = 1 - F_{\mathsf{X}}(x)$
- $F_{\mathsf{X}}(x)$ is continuous from the right, i.e., $\lim_{x \downarrow x_0} F_{\mathsf{X}}(x) = F_{\mathsf{X}}(x_0)$.
  In other words, $F_{\mathsf{X}}(x^+) = F_{\mathsf{X}}(x)$.

Properties of CDF

- $F_{\mathsf{X}}(-\infty) \; = \; 0.$
- $F_{\mathsf{X}}(\infty) \; = \; 1.$
- If $x_1 \leq x_2$, then $F_{\mathsf{X}}(x_2) \geq F_{\mathsf{X}}(x_2)$
  Follows because
  $$\{\mathsf{X} \leq x_1\} \subseteq \{\mathsf{X} \leq x_2\}$$
  This means that $F_{\mathsf{X}}(x)$ is a non decreasing function
- $\mathsf{Prob}(x_1 < \mathsf{X} \leq x_2) \; = \; F_{\mathsf{X}}(x_2) - F_{\mathsf{X}}(x_1).$
- $\mathsf{Prob}(\mathsf{X} > x) = 1 - F_{\mathsf{X}}(x)$
- $F_{\mathsf{X}}(x)$ is continuous from the right, i.e., $\lim_{x \downarrow x_0} F_{\mathsf{X}}(x) = F_{\mathsf{X}}(x_0)$.
  In other words, $F_{\mathsf{X}}(x^+) = F_{\mathsf{X}}(x)$.

Properties of CDF

- $F_{\mathsf{X}}(-\infty) \; = \; 0.$
- $F_{\mathsf{X}}(\infty) \; = \; 1.$
- If $x_1 \leq x_2$, then $F_{\mathsf{X}}(x_2) \geq F_{\mathsf{X}}(x_2)$
  Follows because
  $$\{\mathsf{X} \leq x_1\} \subseteq \{\mathsf{X} \leq x_2\}$$
  This means that $F_{\mathsf{X}}(x)$ is a non decreasing function
- $\mathsf{Prob}(x_1 < \mathsf{X} \leq x_2) \; = \; F_{\mathsf{X}}(x_2) - F_{\mathsf{X}}(x_1).$
- $\mathsf{Prob}(\mathsf{X} > x) = 1 - F_{\mathsf{X}}(x)$
- $F_{\mathsf{X}}(x)$ is continuous from the right, i.e., $\lim_{x \downarrow x_0} F_{\mathsf{X}}(x) = F_{\mathsf{X}}(x_0)$.
  In other words, $F_{\mathsf{X}}(x^+) = F_{\mathsf{X}}(x)$.

Properties of CDF

- $F_{\mathsf{X}}(-\infty) \;=\; 0.$
- $F_{\mathsf{X}}(\infty) \;=\; 1.$
- If $x_1 \leq x_2$, then $F_{\mathsf{X}}(x_2) \geq F_{\mathsf{X}}(x_2)$
  Follows because
  $$\{\mathsf{X} \leq x_1\} \subseteq \{\mathsf{X} \leq x_2\}$$
  This means that $F_{\mathsf{X}}(x)$ is a non decreasing function
- $\mathsf{Prob}(x_1 < \mathsf{X} \leq x_2) \;=\; F_{\mathsf{X}}(x_2) - F_{\mathsf{X}}(x_1).$
- $\mathsf{Prob}(\mathsf{X} > x) = 1 - F_{\mathsf{X}}(x)$
- $F_{\mathsf{X}}(x)$ is continuous from the right, i.e., $\lim_{x \downarrow x_0} F_{\mathsf{X}}(x) = F_{\mathsf{X}}(x_0)$.
  In other words, $F_{\mathsf{X}}(x^+) = F_{\mathsf{X}}(x)$.

Properties of CDF

- $F_{\mathsf{X}}(-\infty) \;=\; 0.$
- $F_{\mathsf{X}}(\infty) \;=\; 1.$
- If $x_1 \leq x_2$, then $F_{\mathsf{X}}(x_2) \geq F_{\mathsf{X}}(x_2)$
  Follows because
  $$\{\mathsf{X} \leq x_1\} \subseteq \{\mathsf{X} \leq x_2\}$$
  This means that $F_{\mathsf{X}}(x)$ is a non decreasing function
- $\mathsf{Prob}(x_1 < \mathsf{X} \leq x_2) \;=\; F_{\mathsf{X}}(x_2) - F_{\mathsf{X}}(x_1).$
- $\mathsf{Prob}(\mathsf{X} > x) = 1 - F_{\mathsf{X}}(x)$
- $F_{\mathsf{X}}(x)$ is continuous from the right, i.e., $\lim_{x \downarrow x_0} F_{\mathsf{X}}(x) = F_{\mathsf{X}}(x_0)$.
  In other words, $F_{\mathsf{X}}(x^+) = F_{\mathsf{X}}(x)$.

Properties of CDF

- $F_X(-\infty) \;=\; 0.$
- $F_X(\infty) \;=\; 1.$
- If $x_1 \leq x_2$, then $F_X(x_2) \geq F_X(x_2)$
  Follows because
  $$\{X \leq x_1\} \subseteq \{X \leq x_2\}$$
  This means that $F_X(x)$ is a non decreasing function
- $\text{Prob}(x_1 < X \leq x_2) \;=\; F_X(x_2) - F_X(x_1).$
- $\text{Prob}(X > x) = 1 - F_X(x)$
- $F_X(x)$ is continuous from the right, i.e., $\lim_{x \downarrow x_0} F_X(x) = F_X(x_0)$.
  In other words, $F_X(x^+) = F_X(x)$.

- $F_{\mathsf{X}}(-\infty) \; = \; 0.$
- $F_{\mathsf{X}}(\infty) \; = \; 1.$
- If $x_1 \leq x_2$, then $F_{\mathsf{X}}(x_2) \geq F_{\mathsf{X}}(x_2)$
  Follows because
  $$\{\mathsf{X} \leq x_1\} \subseteq \{\mathsf{X} \leq x_2\}$$
  This means that $F_{\mathsf{X}}(x)$ is a non decreasing function
- $\mathsf{Prob}(x_1 < \mathsf{X} \leq x_2) \; = \; F_{\mathsf{X}}(x_2) - F_{\mathsf{X}}(x_1).$
- $\mathsf{Prob}(\mathsf{X} > x) = 1 - F_{\mathsf{X}}(x)$
- $F_{\mathsf{X}}(x)$ is continuous from the right, i.e., $\lim_{x \downarrow x_0} F_{\mathsf{X}}(x) = F_{\mathsf{X}}(x_0)$.
  In other words, $F_{\mathsf{X}}(x^+) = F_{\mathsf{X}}(x)$.

Discrete Random Variables

- X is a **discrete** random variable if it takes values from the integer set, $\mathcal{Z}$.
- For discrete random variables, we define a **probability mass function**

$$p_X(k) = \text{Prob}(X = k) = F_X(k) - F_X(k-1)$$

- Properties of a pmf

$$0 \leq p_X(k) \leq 1$$
$$\sum_k p_X(k) = 1$$

$$\text{Prob}(x_1 \leq X \leq x_2) \leq = \sum_{k=\lceil x_1 \rceil}^{\lfloor x_2 \rfloor} p_X(k)$$

Discrete Random Variables

- X is a **discrete** random variable if it takes values from the integer set, $\mathcal{Z}$.
- For discrete random variables, we define a **probability mass function**

$$p_{\mathsf{X}}(k) \;=\; \mathsf{Prob}(\mathsf{X}=k) \;=\; F_{\mathsf{X}}(k) - F_{\mathsf{X}}(k-1)$$

- Properties of a pmf

$$0 \leq p_{\mathsf{X}}(k) \;\leq\; 1$$
$$\sum_k p_{\mathsf{X}}(k) \;=\; 1$$

$$\mathsf{Prob}(x_1 \leq \mathsf{X} \leq x_2) \leq= \sum_{k=\lceil x_1 \rceil}^{\lfloor x_2 \rfloor} p_{\mathsf{X}}(k)$$

Discrete Random Variables

- X is a **discrete** random variable if it takes values from the integer set, $\mathcal{Z}$.
- For discrete random variables, we define a **probability mass function**

$$p_{\mathsf{X}}(k) \;=\; \mathsf{Prob}(\mathsf{X} = k) \;=\; F_{\mathsf{X}}(k) - F_{\mathsf{X}}(k-1)$$

- Properties of a pmf

$$0 \leq p_{\mathsf{X}}(k) \;\leq\; 1$$

$$\sum_{k} p_{\mathsf{X}}(k) \;=\; 1$$

$$\mathsf{Prob}(x_1 \leq \mathsf{X} \leq x_2) \leq= \sum_{k=\lceil x_1 \rceil}^{\lfloor x_2 \rfloor} p_{\mathsf{X}}(k)$$

Example pmf

- Two fair, six-sided dice are thrown; X is sum of the values outcomes. $\mathsf{X} \in \{2, \ldots, 12\}$

$p_X \downarrow \quad \mathsf{X} \rightarrow$	2	3	4	5	6	7
	1/36	2/36	3/36	4/36	5/36	6/36

$p_X \downarrow \quad \mathsf{X} \rightarrow$	8	9	10	11	12
	5/36	4/36	3/36	2/36	1/36

Example pmf: Bernoulli

- Discrete **uniform** random variable takes values in set $\{x_1, x_2, \ldots, x_N\}$ and for $1 \leq i \leq N$,

$$\mathsf{Prob}(\mathsf{X} = x_i) = \frac{1}{N}$$

- **Bernouilli** random variable takes values in set $\{0, 1\}$.

$$\begin{aligned} &\mathsf{X} \in \{0,1\} \\ &\mathsf{Prob}(\mathsf{X}=1) = p_{\mathsf{X}}(1) = \alpha \\ &\mathsf{Prob}(\mathsf{X}=0) = p_{\mathsf{X}}(0) = 1-\alpha \end{aligned}$$

- Best visualised as a coin toss with head mapped to 1 and tail to 0.
- The **indicator** variable I_A is defined as

$$I_A = \begin{cases} 1 & \text{if event } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

is another example of a Bernouilli random variable.

Example pmf: Bernoulli

- Discrete **uniform** random variable takes values in set $\{x_1, x_2, \dots, x_N\}$ and for $1 \leq i \leq N$,

$$\mathsf{Prob}(\mathsf{X} = x_i) = \frac{1}{N}$$

- **Bernouilli** random variable takes values in set $\{0, 1\}$.

$$\begin{aligned} &\mathsf{X} \in \{0, 1\} \\ &\mathsf{Prob}(\mathsf{X} = 1) = p_{\mathsf{X}}(1) = \alpha \\ &\mathsf{Prob}(\mathsf{X} = 0) = p_{\mathsf{X}}(0) = 1 - \alpha \end{aligned}$$

- Best visualised as a coin toss with head mapped to 1 and tail to 0.
- The **indicator** variable I_A is defined as

$$I_A = \begin{cases} 1 & \text{if event } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

is another example of a Bernouilli random variable.

Example pmf: Bernoulli

- Discrete **uniform** random variable takes values in set $\{x_1, x_2, \ldots, x_N\}$ and for $1 \leq i \leq N$,

$$\mathsf{Prob}(\mathsf{X} = x_i) = \frac{1}{N}$$

- **Bernouilli** random variable takes values in set $\{0, 1\}$.

$$\begin{aligned} &\mathsf{X} \in \{0, 1\} \\ &\mathsf{Prob}(\mathsf{X} = 1) = p_{\mathsf{X}}(1) = \alpha \\ &\mathsf{Prob}(\mathsf{X} = 0) = p_{\mathsf{X}}(0) = 1 - \alpha \end{aligned}$$

- Best visualised as a coin toss with head mapped to 1 and tail to 0.
- The **indicator** variable I_A is defined as

$$I_A = \begin{cases} 1 & \text{if event } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

is another example of a Bernouilli random variable.

Example pmf: Bernoulli

- Discrete **uniform** random variable takes values in set $\{x_1, x_2, \ldots, x_N\}$ and for $1 \leq i \leq N$,
$$\mathsf{Prob}(\mathsf{X} = x_i) = \frac{1}{N}$$
- **Bernouilli** random variable takes values in set $\{0, 1\}$.
$$\begin{aligned} &\mathsf{X} \in \{0,1\} \\ &\mathsf{Prob}(\mathsf{X}=1) = p_{\mathsf{X}}(1) = \alpha \\ &\mathsf{Prob}(\mathsf{X}=0) = p_{\mathsf{X}}(0) = 1-\alpha \end{aligned}$$
- Best visualised as a coin toss with head mapped to 1 and tail to 0.
- The **indicator** variable I_A is defined as
$$I_A = \begin{cases} 1 & \text{if event } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$
is another example of a Bernouilli random variable.

Example pmf: Geometric

- **Geometric** Random Variable: Count the number of independent tosses till the first head.

$$
\begin{aligned}
& \mathsf{X} \in \{1, \ldots, \} \\
& \mathsf{Prob}(\mathsf{X} = k) \;=\; p_{\mathsf{X}}(k) \;=\; \begin{cases} \alpha^{k-1}(1-\alpha) & \text{for } k \geq 1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

- Geometric random variables are also defined as the number of tails before the first head. In this

$$
\begin{aligned}
& \mathsf{X} \in \{0, \ldots, \} \\
& \mathsf{Prob}(\mathsf{X} = k) \;=\; p_{\mathsf{X}}(k) \;=\; \begin{cases} \alpha^{k}(1-\alpha) & \text{for } k \geq 0 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

- Let X be the number of coin tosses needed to obtain r heads.

$$
\mathsf{Prob}(\mathsf{X} = k) = \binom{k-1}{r-1} p^{r-1}(1-p)^{(k-1)-(r-1)}p \;=\; \binom{k-1}{r-1} p^{r}(1-p)^{k-r}
$$

Example pmf: Geometric

- **Geometric** Random Variable: Count the number of independent tosses till the first head.

$$\mathsf{X} \in \{1, \ldots, \}$$

$$\mathsf{Prob}(\mathsf{X} = k) \;=\; p_{\mathsf{X}}(k) \;=\; \begin{cases} \alpha^{k-1}(1-\alpha) & \text{for } k \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

- Geometric random variables are also defined as the number of tails before the first head. In this

$$\mathsf{X} \in \{0, \ldots, \}$$

$$\mathsf{Prob}(\mathsf{X} = k) \;=\; p_{\mathsf{X}}(k) \;=\; \begin{cases} \alpha^{k}(1-\alpha) & \text{for } k \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- Let X be the number of coin tosses needed to obtain r heads.

$$\mathsf{Prob}(\mathsf{X} = k) = \binom{k-1}{r-1} p^{r-1}(1-p)^{(k-1)-(r-1)}p \;=\; \binom{k-1}{r-1} p^{r}(1-p)^{k-r}$$

Example pmf: Geometric

- **Geometric** Random Variable: Count the number of independent tosses till the first head.

$$\mathsf{X} \in \{1, \ldots, \}$$

$$\mathsf{Prob}(\mathsf{X} = k) \;=\; p_{\mathsf{X}}(k) \;=\; \begin{cases} \alpha^{k-1}(1-\alpha) & \text{for } k \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

- Geometric random variables are also defined as the number of tails before the first head. In this

$$\mathsf{X} \in \{0, \ldots, \}$$

$$\mathsf{Prob}(\mathsf{X} = k) \;=\; p_{\mathsf{X}}(k) \;=\; \begin{cases} \alpha^{k}(1-\alpha) & \text{for } k \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- Let X be the number of coin tosses needed to obtain r heads.

$$\mathsf{Prob}(\mathsf{X} = k) = \binom{k-1}{r-1} p^{r-1}(1-p)^{(k-1)-(r-1)}p \;=\; \binom{k-1}{r-1} p^{r}(1-p)^{k-r}$$

Example pmf: Binomial

- **Binomial** Random Variable: Sum of N independent Bernouilli random variables, e.g., number of Heads from N independent coin tosses; equivalently, count the number of heads in N tosses of a coin with bias α.

$$\mathsf{X} \in \{0, 1, \ldots, N\}$$

$$\mathsf{Prob}(\mathsf{X} = k) \;=\; p_{\mathsf{X}}(k) \;=\; \begin{cases} \binom{N}{k}\alpha^k(1-\alpha)^{N-k} & \text{for } 0 \leq k \leq N \\ 0 & \text{otherwise} \end{cases}$$

Example pmf: Poisson

- Consider the Binomial random variable with parameters N and p.
- Increase N (the number of coin tosses) and decrease α the probability of Heads (success, etc.) such that the expected number of successes is a constant.
- That is let $N \to \infty$, $\alpha \to 0$ such that $N\alpha = \lambda$ is a constant.
- See what happens to the binomial pmf.

$$
\begin{aligned}
p_X(k) &= \lim_{\substack{N\to\infty \\ N\alpha=\lambda}} \frac{N!}{k!\,(N-k)!}\alpha^k\,(1-\alpha)^{N-k} \\
&= \lim_{\substack{N\to\infty \\ N\alpha=\lambda}} \frac{N!}{k!\,(N-k)!}\left(\frac{\lambda}{N}\right)^k\left(1-\frac{\lambda}{N}\right)^{N-k} \\
&= \frac{\lambda^k}{k!}\lim_{\substack{N\to\infty \\ N\alpha=\lambda}}\left[\frac{N!}{(N-k)!}\left(\frac{1}{N}\right)^k\right]\left(1-\frac{\lambda}{N}\right)^{-k}\left(1-\frac{\lambda}{N}\right)^{N} \\
&= \;??
\end{aligned}
$$

Example pmf: Poisson

- Consider the Binomial random variable with parameters N and p.
- Increase N (the number of coin tosses) and decrease α the probability of Heads (success, etc.) such that the expected number of successes is a constant.
- That is let $N \to \infty$, $\alpha \to 0$ such that $N\alpha = \lambda$ is a constant.
- See what happens to the binomial pmf.

$$\begin{aligned}
p_X(k) &= \lim_{\substack{N\to\infty \\ N\alpha=\lambda}} \frac{N!}{k!\,(N-k)!}\alpha^k\,(1-\alpha)^{N-k} \\
&= \lim_{\substack{N\to\infty \\ N\alpha=\lambda}} \frac{N!}{k!\,(N-k)!}\left(\frac{\lambda}{N}\right)^k\left(1-\frac{\lambda}{N}\right)^{N-k} \\
&= \frac{\lambda^k}{k!}\lim_{\substack{N\to\infty \\ N\alpha=\lambda}}\left[\frac{N!}{(N-k)!}\left(\frac{1}{N}\right)^k\right]\left(1-\frac{\lambda}{N}\right)^{-k}\left(1-\frac{\lambda}{N}\right)^{N} \\
&= \;??
\end{aligned}$$

Example pmf: Poisson

- Consider the Binomial random variable with parameters N and p.
- Increase N (the number of coin tosses) and decrease α the probability of Heads (success, etc.) such that the expected number of successes is a constant.
- That is let $N \to \infty$, $\alpha \to 0$ such that $N\alpha = \lambda$ is a constant.
- See what happens to the binomial pmf.

$$\begin{aligned}
p_X(k) &= \lim_{\substack{N\to\infty \\ N\alpha=\lambda}} \frac{N!}{k!\,(N-k)!}\alpha^k\,(1-\alpha)^{N-k} \\
&= \lim_{\substack{N\to\infty \\ N\alpha=\lambda}} \frac{N!}{k!\,(N-k)!}\left(\frac{\lambda}{N}\right)^k\left(1-\frac{\lambda}{N}\right)^{N-k} \\
&= \frac{\lambda^k}{k!}\lim_{\substack{N\to\infty \\ N\alpha=\lambda}}\left[\frac{N!}{(N-k)!}\left(\frac{1}{N}\right)^k\right]\left(1-\frac{\lambda}{N}\right)^{-k}\left(1-\frac{\lambda}{N}\right)^N \\
&= ??
\end{aligned}$$

Example pmf: Poisson

- Consider the Binomial random variable with parameters N and p.
- Increase N (the number of coin tosses) and decrease α the probability of Heads (success, etc.) such that the expected number of successes is a constant.
- That is let $N \to \infty$, $\alpha \to 0$ such that $N\alpha = \lambda$ is a constant.
- See what happens to the binomial pmf.

$$
\begin{aligned}
p_{\mathsf{X}}(k) &= \lim_{\substack{N\to\infty \\ N\alpha=\lambda}} \frac{N!}{k!\,(N-k)!}\alpha^{k}\,(1-\alpha)^{N-k} \\
&= \lim_{\substack{N\to\infty \\ N\alpha=\lambda}} \frac{N!}{k!\,(N-k)!}\left(\frac{\lambda}{N}\right)^{k}\left(1-\frac{\lambda}{N}\right)^{N-k} \\
&= \frac{\lambda^{k}}{k!}\lim_{\substack{N\to\infty \\ N\alpha=\lambda}} \left[\frac{N!}{(N-k)!}\left(\frac{1}{N}\right)^{k}\right]\left(1-\frac{\lambda}{N}\right)^{-k}\left(1-\frac{\lambda}{N}\right)^{N} \\
&= \;??
\end{aligned}
$$

Continuous Random Variables

- X is a **continuous** random variable if it takes real values.
- The cdf of a continuous random variable is also continuous.
- For continuous random variables, we can define a **probability density function.** Informally,

$$\text{Prob}(x < \mathsf{X} \leq x + \Delta x) \approx f_{\mathsf{X}}(x)\Delta x$$

$$f_{\mathsf{X}}(x) = \frac{dF_{\mathsf{X}}(x)}{dx}$$

- The mass at any point is zero. Thus the probability that a continuous random variable takes a specific value is zero.

Continuous Random Variables

- X is a **continuous** random variable if it takes real values.
- The cdf of a continuous random variable is also continuous.
- For continuous random variables, we can define a **probability density function.** Informally,

$$\mathsf{Prob}(x < \mathsf{X} \leq x + \Delta x) \approx f_{\mathsf{X}}(x)\Delta x$$

$$f_{\mathsf{X}}(x) = \frac{dF_{\mathsf{X}}(x)}{dx}$$

- The mass at any point is zero. Thus the probability that a continuous random variable takes a specific value is zero.

Continuous Random Variables

- X is a **continuous** random variable if it takes real values.
- The cdf of a continuous random variable is also continuous.
- For continuous random variables, we can define a **probability density function.** Informally,

$$\mathsf{Prob}(x < \mathsf{X} \leq x + \Delta x) \approx f_{\mathsf{X}}(x)\Delta x$$

$$f_{\mathsf{X}}(x) = \frac{dF_{\mathsf{X}}(x)}{dx}$$

- The mass at any point is zero. Thus the probability that a continuous random variable takes a specific value is zero.

Continuous Random Variables

- X is a **continuous** random variable if it takes real values.
- The cdf of a continuous random variable is also continuous.
- For continuous random variables, we can define a **probability density function.** Informally,

$$\mathsf{Prob}(x < \mathsf{X} \leq x + \Delta x) \approx f_{\mathsf{X}}(x)\Delta x$$

$$f_{\mathsf{X}}(x) = \frac{dF_{\mathsf{X}}(x)}{dx}$$

- The mass at any point is zero. Thus the probability that a continuous random variable takes a specific value is zero.

Continuous Random Variables

- X is a **continuous** random variable if it takes real values.
- The cdf of a continuous random variable is also continuous.
- For continuous random variables, we can define a **probability density function.** Informally,

$$\mathsf{Prob}(x < \mathsf{X} \leq x + \Delta x) \approx f_{\mathsf{X}}(x)\Delta x$$

$$f_{\mathsf{X}}(x) = \frac{dF_{\mathsf{X}}(x)}{dx}$$

- The mass at any point is zero. Thus the probability that a continuous random variable takes a specific value is zero.

Continuous Random Variables

- X is a **continuous** random variable if it takes real values.
- The cdf of a continuous random variable is also continuous.
- For continuous random variables, we can define a **probability density function.** Informally,

$$\mathsf{Prob}(x < \mathsf{X} \leq x + \Delta x) \approx f_{\mathsf{X}}(x)\Delta x$$

$$f_{\mathsf{X}}(x) = \frac{dF_{\mathsf{X}}(x)}{dx}$$

- The mass at any point is zero. Thus the probability that a continuous random variable takes a specific value is zero.

Properties of a probability density function

- $F_{\mathsf{X}}(a) \;=\; \int_{-\infty}^{a} f_{\mathsf{X}}(x)\,dx$
- $\mathsf{Prob}(x_1 < \mathsf{X} \leq x_2) \;=\; \int_{x_1}^{x_2} f_{\mathsf{X}}(x)\,dx$
- $\mathsf{Prob}(x < \mathsf{X} \leq x + \Delta x) \;=\; F_{\mathsf{X}}(x + \Delta x) - F_{\mathsf{X}}(x)$
- $\mathsf{Prob}(x < \mathsf{X} \leq x + \Delta x) \approx f_{\mathsf{X}}(x)\Delta x$
- $\mathsf{Prob}(\mathsf{X} \in A) \;=\; \int_{A} f_{\mathsf{X}}(x)\,dx$
- $\int_{-\infty}^{\infty} f_{\mathsf{X}}(x)\,dx \;=\; 1$

Properties of a probability density function

- $F_{\mathsf{X}}(a) \;=\; \int_{-\infty}^{a} f_{\mathsf{X}}(x)\, dx$
- $\text{Prob}(x_1 < \mathsf{X} \leq x_2) \;=\; \int_{x_1}^{x_2} f_{\mathsf{X}}(x)\, dx$
- $\text{Prob}(x < \mathsf{X} \leq x + \Delta x) \;=\; F_{\mathsf{X}}(x + \Delta x) - F_{\mathsf{X}}(x)$
- $\text{Prob}(x < \mathsf{X} \leq x + \Delta x) \approx f_{\mathsf{X}}(x)\Delta x$
- $\text{Prob}(\mathsf{X} \in A) \;=\; \int_{A} f_{\mathsf{X}}(x)\, dx$
- $\int_{-\infty}^{\infty} f_{\mathsf{X}}(x)\, dx \;=\; 1$

Properties of a probability density function

- $F_{\mathsf{X}}(a) \;=\; \int_{-\infty}^{a} f_{\mathsf{X}}(x)\, dx$
- $\mathsf{Prob}(x_1 < \mathsf{X} \le x_2) \;=\; \int_{x_1}^{x_2} f_{\mathsf{X}}(x)\, dx$
- $\mathsf{Prob}(x < \mathsf{X} \le x + \Delta x) \;=\; F_{\mathsf{X}}(x + \Delta x) - F_{\mathsf{X}}(x)$
- $\mathsf{Prob}(x < \mathsf{X} \le x + \Delta x) \approx f_{\mathsf{X}}(x)\Delta x$
- $\mathsf{Prob}(\mathsf{X} \in A) \;=\; \int_{A} f_{\mathsf{X}}(x)\, dx$
- $\int_{-\infty}^{\infty} f_{\mathsf{X}}(x)\, dx \;=\; 1$

Properties of a probability density function

- $F_{\mathsf{X}}(a) \;=\; \int_{-\infty}^{a} f_{\mathsf{X}}(x)\,dx$
- $\mathsf{Prob}(x_1 < \mathsf{X} \le x_2) \;=\; \int_{x_1}^{x_2} f_{\mathsf{X}}(x)\,dx$
- $\mathsf{Prob}(x < \mathsf{X} \le x + \Delta x) \;=\; F_{\mathsf{X}}(x+\Delta x) - F_{\mathsf{X}}(x)$
- $\mathsf{Prob}(x < \mathsf{X} \le x + \Delta x) \approx f_{\mathsf{X}}(x)\Delta x$
- $\mathsf{Prob}(\mathsf{X} \in A) \;=\; \int_{A} f_{\mathsf{X}}(x)\,dx$
- $\int_{-\infty}^{\infty} f_{\mathsf{X}}(x)\,dx \;=\; 1$

Properties of a probability density function

- $F_{\mathsf{X}}(a) \;=\; \int_{-\infty}^{a} f_{\mathsf{X}}(x)\,dx$
- $\mathsf{Prob}(x_1 < \mathsf{X} \le x_2) \;=\; \int_{x_1}^{x_2} f_{\mathsf{X}}(x)\,dx$
- $\mathsf{Prob}(x < \mathsf{X} \le x + \Delta x) \;=\; F_{\mathsf{X}}(x + \Delta x) - F_{\mathsf{X}}(x)$
- $\mathsf{Prob}(x < \mathsf{X} \le x + \Delta x) \approx f_{\mathsf{X}}(x)\Delta x$
- $\mathsf{Prob}(\mathsf{X} \in A) \;=\; \int_{A} f_{\mathsf{X}}(x)\,dx$
- $\int_{-\infty}^{\infty} f_{\mathsf{X}}(x)\,dx \;=\; 1$

Properties of a probability density function

- $F_{\mathsf{X}}(a) \;=\; \int_{-\infty}^{a} f_{\mathsf{X}}(x)\,dx$
- $\mathsf{Prob}(x_1 < \mathsf{X} \leq x_2) \;=\; \int_{x_1}^{x_2} f_{\mathsf{X}}(x)\,dx$
- $\mathsf{Prob}(x < \mathsf{X} \leq x+\Delta x) \;=\; F_{\mathsf{X}}(x+\Delta x) - F_{\mathsf{X}}(x)$
- $\mathsf{Prob}(x < \mathsf{X} \leq x+\Delta x) \approx f_{\mathsf{X}}(x)\Delta x$
- $\mathsf{Prob}(\mathsf{X} \in A) \;=\; \int_{A} f_{\mathsf{X}}(x)\,dx$
- $\int_{-\infty}^{\infty} f_{\mathsf{X}}(x)\,dx \;=\; 1$

- Uniform Distribution

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

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

- Exponential Distribution

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

$$F_X(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ 1 - e^{-\lambda x} & \text{for } x \geq 0 \end{cases}$$

Some ‘Continuous’ Distributions

- Uniform Distribution

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

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

- Exponential Distribution

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

$$F_X(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ 1 - e^{-\lambda x} & \text{for } x \geq 0 \end{cases}$$

An Important Continuous Distribution

- Gaussian Distribution

$$
\begin{aligned}
f_X(x) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\
F_X(x) &= ??
\end{aligned}
$$

- This is also called the Normal distribution.
- If $\mu = 0$ and $\sigma = 1$, then this is called the 'unit Normal' distribution.
- Often used to measure 'noise' and 'errors'.
- Many more distributions using 'functions' of Gaussian are widely used. See text for names and descriptions.

An Important Continuous Distribution

- Gaussian Distribution

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

- This is also called the Normal distribution.
- If $\mu = 0$ and $\sigma = 1$, then this is called the 'unit Normal' distribution.
- Often used to measure 'noise' and 'errors'.
- Many more distributions using 'functions' of Gaussian are widely used. See text for names and descriptions.

An Important Continuous Distribution

- Gaussian Distribution

$$
\begin{aligned}
f_{\mathsf{X}}(x) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\
F_{\mathsf{X}}(x) &= ??
\end{aligned}
$$

- This is also called the Normal distribution.
- If $\mu = 0$ and $\sigma = 1$, then this is called the 'unit Normal' distribution.
- Often used to measure 'noise' and 'errors'.
- Many more distributions using 'functions' of Gaussian are widely used. See text for names and descriptions.

An Important Continuous Distribution

- Gaussian Distribution

$$
\begin{aligned}
f_X(x) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\
F_X(x) &= ??
\end{aligned}
$$

- This is also called the Normal distribution.
- If $\mu = 0$ and $\sigma = 1$, then this is called the 'unit Normal' distribution.
- Often used to measure 'noise' and 'errors'.
- Many more distributions using 'functions' of Gaussian are widely used. See text for names and descriptions.

An Important Continuous Distribution

- Gaussian Distribution

$$
\begin{aligned}
f_X(x) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\
F_X(x) &= ??
\end{aligned}
$$

- This is also called the Normal distribution.
- If $\mu = 0$ and $\sigma = 1$, then this is called the 'unit Normal' distribution.
- Often used to measure 'noise' and 'errors'.
- Many more distributions using 'functions' of Gaussian are widely used. See text for names and descriptions.

Conditional distributions and densities

- Conditional distribution $F_{\mathsf{X}|A}(x|A)$ is defined exactly like conditional probability.

$$F_{\mathsf{X}|A}(x|A) = \frac{\mathsf{Prob}(\mathsf{X} \le x,\ A)}{\mathsf{Prob}(A)}$$

- Properties of cdf also hold for conditional cdf, i.e.,
  - $F_{\mathsf{X}|A}(-\infty|A) = 0;\quad F_{\mathsf{X}|A}(\infty|A) = 1;$
  - $\mathsf{Prob}(x_1 < \mathsf{X} \le x_2|A) = F_{\mathsf{X}|A}(x_2|A) - F_{\mathsf{X}|A}(x_1|A)$
  - $F_{\mathsf{X}|A}(x|A)$ is non decreasing.
- Can also define conditional pdf and conditional pmf

$$f_{\mathsf{X}|A}(x|A) = \frac{dF_{\mathsf{X}|A}(x|A)}{dx} = \lim_{\Delta x \to 0} \frac{\mathsf{Prob}(x \le \mathsf{X} \le x + \Delta x \mid A)}{\Delta x}$$

$$p_{\mathsf{X}|A}(x|A) = \mathsf{Prob}(\mathsf{X} = x \mid A) = \frac{\mathsf{Prob}(\mathsf{X} = x,\ A)}{\mathsf{Prob}(A)}$$

Conditional distributions and densities

- Conditional distribution $F_{\mathsf{X}|A}(x|A)$ is defined exactly like conditional probability.

$$F_{\mathsf{X}|A}(x|A) = \frac{\mathsf{Prob}(\mathsf{X} \leq x,\ A)}{\mathsf{Prob}(A)}$$

- Properties of cdf also hold for conditional cdf, i.e.,
  - $F_{\mathsf{X}|A}(-\infty|A) = 0; \quad F_{\mathsf{X}|A}(\infty|A) = 1;$
  - $\mathsf{Prob}(x_1 < \mathsf{X} \leq x_2|A) \ = F_{\mathsf{X}|A}(x_2|A) - F_{\mathsf{X}|A}(x_1|A)$
  - $F_{\mathsf{X}|A}(x|A)$ is non decreasing.
- Can also define conditional pdf and conditional pmf

$$f_{\mathsf{X}|A}(x|A) \quad = \quad \frac{dF_{\mathsf{X}|A}(x|A)}{dx} = \lim_{\Delta x \to 0} \frac{\mathsf{Prob}(x \leq \mathsf{X} \leq x + \Delta x \mid A)}{\Delta x}$$

$$p_{\mathsf{X}|A}(x|A) \quad = \quad \mathsf{Prob}(\mathsf{X} = x \mid A) = \frac{\mathsf{Prob}(\mathsf{X} = x,\ A)}{\mathsf{Prob}(A)}$$

Conditional distributions and densities

- Conditional distribution $F_{\mathsf{X}|A}(x|A)$ is defined exactly like conditional probability.

$$F_{\mathsf{X}|A}(x|A) = \frac{\mathsf{Prob}(\mathsf{X} \leq x,\ A)}{\mathsf{Prob}(A)}$$

- Properties of cdf also hold for conditional cdf, i.e.,
  - $F_{\mathsf{X}|A}(-\infty|A) = 0;\ \ F_{\mathsf{X}|A}(\infty|A) = 1;$
  - $\mathsf{Prob}(x_1 < \mathsf{X} \leq x_2|A)\ = F_{\mathsf{X}|A}(x_2|A) - F_{\mathsf{X}|A}(x_1|A)$
  - $F_{\mathsf{X}|A}(x|A)$ is non decreasing.
- Can also define conditional pdf and conditional pmf

$$f_{\mathsf{X}|A}(x|A) = \frac{dF_{\mathsf{X}|A}(x|A)}{dx} = \lim_{\Delta x \to 0} \frac{\mathsf{Prob}(x \leq \mathsf{X} \leq x + \Delta x \mid A)}{\Delta x}$$

$$p_{\mathsf{X}|A}(x|A) = \mathsf{Prob}(\mathsf{X} = x \mid A) = \frac{\mathsf{Prob}(\mathsf{X} = x,\ A)}{\mathsf{Prob}(A)}$$

Conditional distributions and densities

- Conditional distribution $F_{X|A}(x|A)$ is defined exactly like conditional probability.

$$F_{X|A}(x|A) = \frac{\text{Prob}(X \le x,\ A)}{\text{Prob}(A)}$$

- Properties of cdf also hold for conditional cdf, i.e.,
  - $F_{X|A}(-\infty|A) = 0;\ \ F_{X|A}(\infty|A) = 1;$
  - $\text{Prob}(x_1 < X \le x_2|A) \ = F_{X|A}(x_2|A) - F_{X|A}(x_1|A)$
  - $F_{X|A}(x|A)$ is non decreasing.
- Can also define conditional pdf and conditional pmf

$$f_{X|A}(x|A) \quad = \quad \frac{dF_{X|A}(x|A)}{dx} = \lim_{\Delta x \to 0} \frac{\text{Prob}(x \le X \le x + \Delta x \mid A)}{\Delta x}$$

$$p_{X|A}(x|A) \quad = \quad \text{Prob}(X = x \mid A) = \frac{\text{Prob}(X = x,\ A)}{\text{Prob}(A)}$$

Conditional distributions and densities

- Conditional distribution $F_{\mathsf{X}|A}(x|A)$ is defined exactly like conditional probability.
$$F_{\mathsf{X}|A}(x|A) = \frac{\text{Prob}(\mathsf{X} \leq x,\ A)}{\text{Prob}(A)}$$
- Properties of cdf also hold for conditional cdf, i.e.,
  - $F_{\mathsf{X}|A}(-\infty|A) = 0;\ \ F_{\mathsf{X}|A}(\infty|A) = 1;$
  - $\text{Prob}(x_1 < \mathsf{X} \leq x_2|A)\ = F_{\mathsf{X}|A}(x_2|A) - F_{\mathsf{X}|A}(x_1|A)$
  - $F_{\mathsf{X}|A}(x|A)$ is non decreasing.
- Can also define conditional pdf and conditional pmf

$$f_{\mathsf{X}|A}(x|A) = \frac{dF_{\mathsf{X}|A}(x|A)}{dx} = \lim_{\Delta x \to 0} \frac{\text{Prob}(x \leq \mathsf{X} \leq x + \Delta x \mid A)}{\Delta x}$$

$$p_{\mathsf{X}|A}(x|A) = \text{Prob}(\mathsf{X} = x \mid A) = \frac{\text{Prob}(\mathsf{X} = x,\ A)}{\text{Prob}(A)}$$

Conditional distributions and densities

- Conditional distribution $F_{\mathsf{X}|A}(x|A)$ is defined exactly like conditional probability.

$$F_{\mathsf{X}|A}(x|A) = \frac{\mathsf{Prob}(\mathsf{X} \le x,\ A)}{\mathsf{Prob}(A)}$$

- Properties of cdf also hold for conditional cdf, i.e.,
  - $F_{\mathsf{X}|A}(-\infty|A) = 0;\ \ F_{\mathsf{X}|A}(\infty|A) = 1;$
  - $\mathsf{Prob}(x_1 < \mathsf{X} \le x_2|A)\ = F_{\mathsf{X}|A}(x_2|A) - F_{\mathsf{X}|A}(x_1|A)$
  - $F_{\mathsf{X}|A}(x|A)$ is non decreasing.
- Can also define conditional pdf and conditional pmf

$$f_{\mathsf{X}|A}(x|A) = \frac{dF_{\mathsf{X}|A}(x|A)}{dx} = \lim_{\Delta x \to 0} \frac{\mathsf{Prob}(x \le \mathsf{X} \le x + \Delta x \mid A)}{\Delta x}$$

$$p_{\mathsf{X}|A}(x|A) = \mathsf{Prob}(\mathsf{X} = x \mid A) = \frac{\mathsf{Prob}(\mathsf{X} = x,\ A)}{\mathsf{Prob}(A)}$$

Conditional distributions and densities

- Total probability theorem

$$\mathsf{Prob}(A) = \sum_i \mathsf{Prob}(A \mid \mathsf{X} = i)\ p_{\mathsf{X}}(x)\ dx$$

$$\mathsf{Prob}(A) = \int_{-\infty}^{\infty} \mathsf{Prob}(A \mid \mathsf{X} = x)\ f_{\mathsf{X}}(x)\ dx$$

- Bayes Theorem

$$f_{\mathsf{X}|A}(x|A) = \frac{\mathsf{Prob}(A|\mathsf{X}=x)}{\mathsf{Prob}(A)} f_{\mathsf{X}}(x) = \frac{\mathsf{Prob}(A|\mathsf{X}=x) f_{\mathsf{X}}(x)}{\int_{-\infty}^{\infty} \mathsf{Prob}(A|\mathsf{X}=x)\ f_{\mathsf{X}}(x)\ dx}$$

$$p_{\mathsf{X}|A}(x|A) = \frac{\mathsf{Prob}(A|\mathsf{X}=x)}{\mathsf{Prob}(A)} p_{\mathsf{X}}(x) = \frac{\mathsf{Prob}(A|\mathsf{X}=x) p_{\mathsf{X}}(x)}{\sum_x \mathsf{Prob}(A|\mathsf{X}=x)\ p_{\mathsf{X}}(x)}$$

- $f_{\mathsf{X}}(x)$ (resp. $p_{\mathsf{X}}(x)$) is also called the unconditional **prior** pdf (resp. pmf) of X and $f_{\mathsf{X}|A}(x|A)$ (resp. $p_{\mathsf{X}|A}(x|bA)$) is called the **posterior** pdf (resp. pmf).

Conditional distributions and densities

- Total probability theorem

$$\mathsf{Prob}(A) = \sum_i \mathsf{Prob}(A \mid \mathsf{X} = i) \; p_{\mathsf{X}}(x) \; dx$$

$$\mathsf{Prob}(A) = \int_{-\infty}^{\infty} \mathsf{Prob}(A \mid \mathsf{X} = x) \; f_{\mathsf{X}}(x) \; dx$$

- Bayes Theorem

$$f_{\mathsf{X}|A}(x|A) = \frac{\mathsf{Prob}(A|\mathsf{X}=x)}{\mathsf{Prob}(A)} f_{\mathsf{X}}(x) = \frac{\mathsf{Prob}(A|\mathsf{X}=x) f_{\mathsf{X}}(x)}{\int_{-\infty}^{\infty} \mathsf{Prob}(A|\mathsf{X}=x) \; f_{\mathsf{X}}(x) \; dx}$$

$$p_{\mathsf{X}|A}(x|A) = \frac{\mathsf{Prob}(A|\mathsf{X}=x)}{\mathsf{Prob}(A)} p_{\mathsf{X}}(x) = \frac{\mathsf{Prob}(A|\mathsf{X}=x) p_{\mathsf{X}}(x)}{\sum_x \mathsf{Prob}(A|\mathsf{X}=x) \; p_{\mathsf{X}}(x)}$$

- $f_{\mathsf{X}}(x)$ (resp. $p_{\mathsf{X}}(x)$) is also called the unconditional **prior** pdf (resp. pmf) of X and $f_{\mathsf{X}|A}(x|A)$ (resp. $p_{\mathsf{X}|A}(x|bA)$) is called the **posterior** pdf (resp. pmf).

Conditional distributions and densities

- Total probability theorem

$$\mathsf{Prob}(A) = \sum_i \mathsf{Prob}(A \mid \mathsf{X} = i)\ p_{\mathsf{X}}(x)\ dx$$

$$\mathsf{Prob}(A) = \int_{-\infty}^{\infty} \mathsf{Prob}(A \mid \mathsf{X} = x)\ f_{\mathsf{X}}(x)\ dx$$

- Bayes Theorem

$$f_{\mathsf{X}|A}(x|A) = \frac{\mathsf{Prob}(A|\mathsf{X}=x)}{\mathsf{Prob}(A)} f_{\mathsf{X}}(x) = \frac{\mathsf{Prob}(A|\mathsf{X}=x) f_{\mathsf{X}}(x)}{\int_{-\infty}^{\infty} \mathsf{Prob}(A|\mathsf{X}=x)\ f_{\mathsf{X}}(x)\ dx}$$

$$p_{\mathsf{X}|A}(x|A) = \frac{\mathsf{Prob}(A|\mathsf{X}=x)}{\mathsf{Prob}(A)} p_{\mathsf{X}}(x) = \frac{\mathsf{Prob}(A|\mathsf{X}=x) p_{\mathsf{X}}(x)}{\sum_x \mathsf{Prob}(A|\mathsf{X}=x)\ p_{\mathsf{X}}(x)}$$

- $f_{\mathsf{X}}(x)$ (resp. $p_{\mathsf{X}}(x)$) is also called the unconditional **prior** pdf (resp. pmf) of X and $f_{\mathsf{X}|A}(x|A)$ (resp. $p_{\mathsf{X}|A}(x|bA)$) is called the **posterior** pdf (resp. pmf).