Introduction to Statistical Methods in Political Science

Lecture 5: Random Variables, PMF, PDF, and CDF

Ignacio Urbina
Ph.D. Candidate in Political Science

Random Variables and their Probability Distributions

From Events to Numbers

- Previously, we studied the probability of discrete events, such as:
  - Probability of two countries going to war.
  - Probability that a Senator is a woman.
  - Probability that a randomly sampled person a bachelor's degree.
  - Probability of getting an even number when throwing a die.
- These probabilities help us understand the likelihood of events.
- To perform deeper statistical analysis, we need to quantify these events numerically.

Transforming Educational Attainment into Numbers

- For example, consider educational attainment levels:
  - No education
  - Incomplete high school
  - Complete high school
  - Some college
  - Bachelor's degree
  - Graduate degree
- Instead of discrete categories, we can measure the number of years of schooling.
- This allows us to calculate statistics like the mean number of years of education.

Continuous Numerical Variables: Feeling Thermometers

- Consider the feeling thermometer used in surveys like the American National Elections Study (ANES):
  - Rate feelings toward political figures on a scale from 0 to 100.
  - $0 =$ Very cold or unfavorable
  - $50 =$ Neutral
  - $100 =$ Very warm or favorable
- This allows us to calculate statistics like mean favorability.

Introduction to Random Variables

- We use **random variables** to numerically quantify events.
- A random variable assigns a numerical value to each outcome in a sample space.

> **Random Variable** (*Def.*)
>
> A random variable is a function that maps each outcome in the sample space S to a specific numeric value.

- Allows us to quantify events and analyze them numerically.
- **In practice, we treat Random Variables as numeric variables** and manipulate them using specific algebra rules.
- Example: Tossing a coin, where heads are assigned a 1, and tails a 0.

Types of Random Variables

- **Discrete Random Variables**
  - Take on a finite or countable number of values.
  - Example: Number of children in a family, number of votes received by a candidate.
- **Continuous Random Variables**
  - Take on an infinite number of values within a given range.
  - Example: Income level, time spent on social media.

Random Variables for One Die

- Define the sample space S for throwing one die as:

$$S = \{1, 2, 3, 4, 5, 6\}$$

- Define a random variable X as a function

$$X : S \to \{1, 2, 3, 4, 5, 6\}$$

- X maps each outcome in the sample space to a numerical value.
- Let $X = x$ be a specific value of X. Examples:
  - If the die shows 2, then $x = 2$.
  - If the die shows 5, then $x = 5$.
  - If the die shows 1, then $x = 1$.

Random Variables for Two Dice

- Define the sample space S for two dice as:

$$\begin{aligned} S &= \{(i,j) \mid i,j \in \{1,2,3,4,5,6\}\} \\ &= \{(1,1),(1,2),(1,2),\cdots,(2,1),(2,2),\cdots,(6,5),(6,6)\} \end{aligned}$$

- The sum of the two dice numbers can be represented by a random variable Z.
- Let $Z = z$ be a specific value of Z. Examples:
  - If the dice show (2, 3), then $z = 2 + 3 = 5$.
  - If the dice show (6, 1), then $z = 6 + 1 = 7$.
  - If the dice show (4, 4), then $z = 4 + 4 = 8$.
  - If the dice show (3, 5), then $z = 3 + 5 = 8$.

Random Variables for Two Dice

- Depending on **how we define an event, the function used as a random variable will change**.
- Define the sample space S for two dice as:

$$\begin{aligned} S &= \big\{(i,j) \mid i,j \in \big\{1,2,3,4,5,6\big\}\big\} \\ &= \big\{(1,1),(1,2),(1,3),\cdots,(2,1),(2,2),\cdots,(6,5),(6,6)\big\} \end{aligned}$$

- The **total number of even dice obtained from throwing two dice** can be represented by a random variable V.
- Let $V = v$ be a specific value of V. Examples:
  - If the dice show (2, 3), then $v = 1 + 0 = 1$.
  - If the dice show (1, 6), then $v = 0 + 1 = 1$.
  - If the dice show (4, 4), then $v = 1 + 1 = 2$.
  - If the dice show (3, 5), then $v = 0 + 0 = 0$.

Random Variables for One Coin

- Define the sample space S for throwing a coin as:

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

- Define a random variable W as a function $W : S \to \{0, 1\}$, where W represents the number of heads.
- W maps each outcome in the sample space to the number of heads.
- Examples:
  - If the coin shows H, then $w = 1$.
  - If the coin shows T, then $w = 0$.

Random Variables for Three Coins

- Define the sample space S for three coins as:

$$S = \{(H,H,H),(H,H,T),(H,T,H),(H,T,T),(T,H,H),\\ (T,H,T),(T,T,H),(T,T,T)\}$$

- Define a random variable Y as a function $Y : S \to \{0,1,2,3\}$, where $Y = y$ represents the number of heads.
- Y maps each outcome in the sample space to the number of heads.
- Examples:
  - If the coins show (H, H, T), then $y = 2$.
  - If the coins show (T, T, T), then $y = 0$.
  - If the coins show (H, H, H), then $y = 3$.
  - If the coins show (T, H, T), then $y = 1$.

Random Variables for Election Results

- Consider the results of the last Senatorial election in three battleground states. Define the sample space S for these results as:
  $$S = \{(R,R,R),(R,R,D),(R,D,R),(R,D,D),(D,R,R),\\ (D,R,D),(D,D,R),(D,D,D)\}$$
- Define a random variable Z as a function $Z : S \to \{0,1,2,3\}$, where $Z = z$ represents the number of battleground states won by the Republican Party.
- Z maps each outcome in the sample space to the number of states won by the Republican Party.
- Examples:
  - If the results are (R, R, D), then $z = 2$.
  - If the results are (D, D, D), then $z = 0$.
  - If the results are (R, R, R), then $z = 3$.
  - If the results are (D, R, D), then $z = 1$.

Definition of Probability for Discrete Random Variables

Probability for Discrete Random Variables (*Def.*)

When each outcome in the sample space S is equally likely, the probability that a discrete random variable X takes the value x is given by the ratio of the number of outcomes in S where $X = x$ to the total number of outcomes in S.

- For a discrete random variable X and a specific value x:

$$P(X = x) = \frac{\text{Number of favorable outcomes in which } X = x}{\text{Total number of possible outcomes}}$$

- This is based on the classical definition of probability

Properties of the Probability Distribution for a Discrete Random Variable

A function can serve as the probability distribution for a discrete random variable X if and only if its values, $P(X = x)$, satisfy the conditions:

- $P(X = x) \geq 0$ for each value within its domain
- $\sum_{i=1}^{k} P(X = x_i) = 1$, where the summation extends over all the values within its domain ($\mathcal{D}_X = \{x_1, x_2, \cdots, x_k\}$)

Note that a *discrete* random variable's probability distribution is often referred to as its **Probability Mass Function (PMF)**.

Probability Distribution for Number of Heads in Three Coin Tosses

- Sample space S for three coin tosses:

$$S = \{(H, H, H), (H, H, T), (H, T, H), (H, T, T), (T, H, H), (T, H, T), (T, T, H), (T, T, T)\}$$

- Define random variable X as the number of heads.
- Probability Distribution (PMF):

Number of Heads: x	**Probability:** $P(X = x)$
0	$1/8$
1	$3/8$
2	$3/8$
3	$1/8$

Example: Simplified Feeling Thermometer

- Suppose a survey firm implemented a 10-point feeling thermometer to measure people's feelings towards political leaders or public figures.
- Ratings range from 0 (very unfavorable) to 10 (very favorable).
- Let's say this pollster was interested in the feelings towards the current President among the public.
- Define the random variable X as the response of one subject on this thermometer.

PMF for the 10-Point Feeling Thermometer Towards the Current President (x)

PMF for X

x	$P(X = x)$
0	0.05
1	0.10
2	0.10
3	0.15
4	0.10
5	0.15
6	0.10
7	0.10
8	0.05
9	0.05
10	0.05

Probability Questions

What is the probability that X is...

1. Greater than or equal to 8?

$$P(X \geq 8) = P(X = 8) + P(X = 9) + P(X = 10)$$
$$= 0.05 + 0.05 + 0.05 = 0.15.$$

2. Less than 5?

$$P(X < 5) =$$
$$= p_X(0) + p_X(1) + p_X(2) + p_X(3) + p_X(4)$$
$$= 0.05 + 0.10 + 0.10 + 0.15 + 0.10 = 0.50.$$

3. Between 3 and 7 inclusive?

$$P(3 \leq X \leq 7) = P(X = 3) + P(4) + \cdots + P(7)$$
$$= 0.15 + 0.10 + 0.15 + 0.10 + 0.10 = 0.60.$$

Common PMF For Discrete Random Variables

Common PMFs for Discrete Random Variables

- There are several common probability mass functions (PMFs) used to model discrete random variables.
- We will discuss three of them:
  - Uniform Distribution
  - Bernoulli Distribution
  - Binomial Distribution
- But note that there are several more: Negative Binomial, Geometric, Poisson, Multinomial, etc (see: Link).
- The following is a strongly recommended reference: `https://uw-statistics.github.io/Stat311Tutorial/discrete-distributions.html`.

Uniform Distribution

- In a uniform distribution, each outcome is equally likely.
- For a discrete random variable X with n possible outcomes:

$$p_X(x) = \frac{1}{n} \quad \text{for all } x$$

- Note that: $p_X(x) \geq 0$ and $\sum_x p_X(x) = \frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n} = n \times \frac{1}{n} = 1$
- This means that $p_X(x)$ is a well-defined PMF.

Example: Uniform Distribution

- Consider rolling a fair six-sided die:

$$S = \{1, 2, 3, 4, 5, 6\}$$

- The PMF for each outcome is:

$$p_X(x) = \frac{1}{6} \quad \text{for } x = 1, 2, 3, 4, 5, 6$$

- This means each number has an equal probability of $\frac{1}{6}$.

Calculating Probabilities for Events Involving a Fair Six-Sided Die

- **Single Value Event:**
  - Probability of rolling a 4:

$$P(X = 4) = \frac{1}{6}$$

- **Composite Event:**
  - Probability of rolling an even number (2, 4, or 6):

$$P(x \text{ is even}) = P(X = 2) + P(X = 4) + P(X = 6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6}$$

- **Multiple Disjoint Events:**
  - Probability of rolling a number less than 4 (1, 2, or 3):

$$P(x < 4) = P(X = 1) + P(X = 2) + P(X = 3) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$$

- **Non-occurring Event:**
  - Probability of rolling a 7:

$$P(x = 7) = 0$$

Bernoulli Distribution

- The Bernoulli distribution models a single trial with two outcomes: success (1) and failure (0).
- For a random variable X with success probability p:

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

- Which of the two outcomes is labeled as a success or failure is generally arbitrarily defined.

Example: Bernoulli Distribution

- Consider flipping a fair coin:

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

- Let X be 1 if heads (H) and 0 if tails (T). If the coin is fair, the PMF is:

$$p_X(x) = \begin{cases} 0.5 & \text{if } x = 1 \\ 0.5 & \text{if } x = 0 \end{cases}$$

Example: Bernoulli Distribution for Sociodemographic Variable

- **Example**: If we randomly select one person among the American public, what is the probability they identify as Latino/a/x?
- Let X be a random variable where $X = 1$ if the person identifies as Latino/a/x and $X = 0$ otherwise.
- According to 2020 US Census data, the probability p that a randomly selected person identifies as Latino/a/x is 0.189.
- We can represent X as a Bernoulli random variable :

$$p_X(x) = \begin{cases} 0.189 & \text{if } x = 1 \\ 0.811 & \text{if } x = 0 \end{cases}$$

Binomial Distribution

- The binomial distribution models the number of successes in a fixed number of independent Bernoulli trials.
- For a random variable X representing the number of successes in n trials with success probability p:

$$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

- Example: Number of heads in 10 flips of a fair coin.

Example: Binomial Distribution

- Consider flipping a fair coin 10 times:

$$n = 10, \quad p = 0.5$$

- The PMF for the number of heads X is:

$$p_X(x) = \binom{10}{x}(0.5)^x(0.5)^{10-x}$$

- This gives the probability of getting exactly x heads in 10 flips.

Example: Given a random sample of 10 people, what is the probability that two or fewer identify as Latino/a/x?

- Define X as the number of Latinos in a sample of 10. This follows a binomial distribution: $X \sim \text{Binomial}(N = 10, p = 0.189)$.

$$P(X = 0) = \binom{10}{0} \times 0.189^0 \times 0.811^{10} \approx 0.1231$$

$$P(X = 1) = \binom{10}{1} \times 0.189^1 \times 0.811^9 \approx 0.2868$$

$$P(X = 2) = \binom{10}{2} \times 0.189^2 \times 0.811^8 \approx 0.3008$$

- Adding these probabilities:

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) \approx 0.7107$$

- This formula calculates the total probability of getting 0, 1, or 2 Latinos in the sample, which is approximately 71.07%.

Example: Probability of at least 1 or at least 3 Latinos in a sample of 10

- Define X as the number of Latinos in a sample of 10. This follows a binomial distribution: $X \sim \text{Binomial}(N = 10, p = 0.189)$.
- **Probability of at least 1 Latino:**

$$P(X \geq 1) = 1 - P(X = 0)$$

$$P(X \geq 1) = 1 - 0.1231 \approx 0.8769$$

- **Probability of at least 3 Latinos:**

$$P(X \geq 3) = 1 - P(X \leq 2)$$

$$P(X \geq 3) = 1 - 0.7107 \approx 0.2893$$

- These probabilities show that there is an 87.69% chance of having at least one Latino in the sample and a 28.93% chance of having at least three.

Galton Board and the Binomial Distribution

- Binomial PMF graph:
  https://shiny.rit.albany.edu/stat/binomial/
- Galton Board:
  https://www.mathsisfun.com/data/quincunx.html

Probability for Continuous Random Variables

Concept of Probability for Continuous Random Variables

- Continuous random variables take on infinite possible values within a given range.
- Therefore, the probability of *exactly* obtaining a specific value $X = x$ after a random trial is 0 due to this definition.
- Hence, when dealing with **continuous random variables**, we need to work with **probabilities of intervals**, i.e., $\Pr(X \in [a, b])$.
- So, instead of asking what's the probability that someone's height is precisely 5 feet 7.16 inches, we can ask what is the probability that someone's height is between 5 feet 5 inches and 5 feet 10 inches.

Introduction to Probability Density Function (PDF)

- The **Probability Density Function (PDF)**, denoted as $f_X(x)$, describes the likelihood of a continuous random variable X taking on a particular value.
- Note: the PDF represents the *relative likelihood* of X being near x. $f_X(x)$ **IS NOT** the probability of x.
- To find the probability that X lies within an interval $[a, b]$, we calculate the area under the PDF curve between a and b.
- Mathematically, this is formally expressed using definite integrals:

$$P(X \in [a, b]) = P(a \leq X \leq b) = \int_a^b f_X(x) \cdot dx$$

Notation: Definite Integrals (*optional slide*)

- A definite integral represents the area under the graph of a function and the x-axis.
- The notation for a definite integral is:

$$\int_a^b f(x) \cdot dx$$

- Here:
  - a is the lower limit of integration.
  - b is the upper limit of integration.
  - $f(x)$ is the function being integrated.
  - dx indicates that the integration is with respect to x.
- The definite integral calculates the net area between the function $f(x)$ and the x-axis from $x = a$ to $x = b$.

Definition of a Well-Defined Probability Density Function (PDF)

Probability Density Function (PDF)

A function $f_X(x)$ is a valid probability density function (PDF) of a continuous random variable X if and only if it satisfies the non-negativity and normalization conditions.

1. **Non-Negativity:** For all possible values of x,

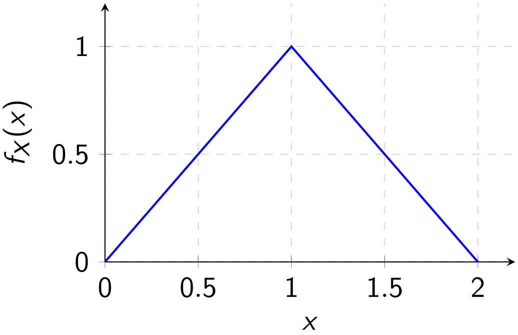
$$f_X(x) \geq 0, \quad \text{for all } x.$$

2. **Normalization:** The total probability over all possible values of X must sum to 1, meaning:

$$\int_{-\infty}^{\infty} f_X(x)\, dx = 1.$$

Connecting the Dots: Example

- Define a random variable X that takes values from 0 to 2.
- The PDF of X follows a *triangular distribution* with mode at $c = 1$:



$$f_X(x) = \begin{cases} x, & 0 \leq x \leq 1 \\ (2 - x), & 1 \leq x \leq 2 \end{cases}$$

Area Under the Curve of the PDF

- The area under the PDF consists of two right triangles.
- The left triangle (from 0 to 1) has:
  - Base $= 1 - 0 = 1$
  - Height $= 1$
- The right triangle (from 1 to 2) has:
  - Base $= 2 - 1 = 1$
  - Height $= 1$
- Using the formula for the area of a triangle:

$$\left(\frac{1}{2} \times 1 \times 1\right) + \left(\frac{1}{2} \times 1 \times 1\right) = 0.5 + 0.5 = 1$$

- This confirms the total area under the curve is 1, making it a valid PDF.

Calculating Probability for an Interval

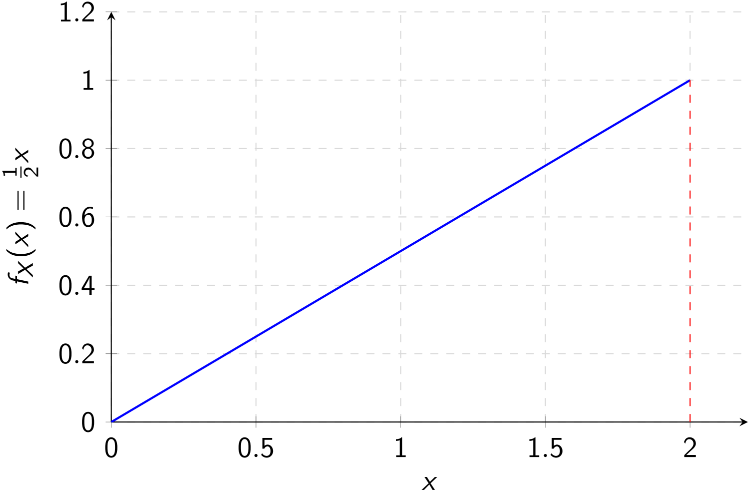
Problem: Calculate the probability that X lies between 0.5 and 1.

- The probability is the area under the PDF between 0.5 and 1, which can be broken down as the area of a small triangle and a rectangle under $f_X(x)$ from 0.5 to 1.0.
- Rectangle area:
  - Base $= 1 - 0.5 = 0.5$
  - Height $= 0.5 - 0 = 0.5$
  - Area $= \mathit{Base} \times \mathit{Height} = 0.5 \times 0.5 = 0.25$
- Small triangle area:
  - Base $= 1 - 0.5 = 0.5$
  - Height $= 1 - 0.5 = 0.5$
  - Area $= \frac{1}{2} \times \mathit{Base} \times \mathit{Height} = \frac{1}{2} \times 0.5 \times 0.5 = 0.125$
- Thus,

$$P(0.5 \leq X \leq 1) = 0.25 + 0.125 = 0.375$$

Example 2

- Define a random variable X that takes on values from 0 to 2.
- Suppose the PDF of X is $f_X(x) = \frac{1}{2}x$ for $0 \leq x \leq 2$.
- Graph of $f_X(x) = \frac{1}{2}x$ from 0 to 2:



Area Under the Curve of the PDF

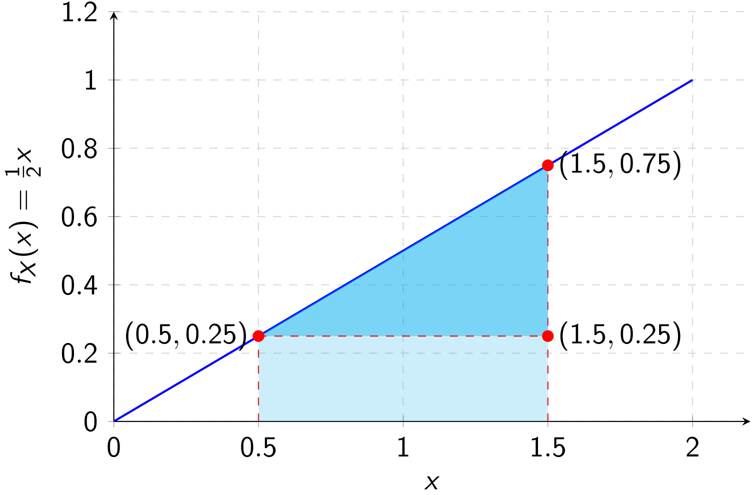
- The area under the curve $f_X(x) = \frac{1}{2}x$ from $x = 0$ to $x = 2$ forms a right triangle.
- The base of the triangle is $2 - 0 = 2$.
- The height of the triangle is $\frac{1}{2} \times 2 = 1$.
- Using the formula for the area of a triangle:

$$\text{Area} = \frac{1}{2} \times \text{base} \times \text{height} = \frac{1}{2} \times 2 \times 1 = 1$$

- This means the total area under the PDF is 1, making it a well-defined PDF.

Calculating Probability for an Interval

- To calculate the probability that X lies between 0.5 and 1.5:
- We use one rectangle and one triangle to represent the area.



Calculating Probability for an Interval

- The area of the rectangle (base = 1.5 - 0.5 = 1, height = 0.25):

$$\text{Area} = \text{base} \times \text{height} = 1 \times 0.25 = 0.25$$

- The area of the triangle (base = 1.5 - 0.5 = 1, height = 0.75 - 0.25 = 0.5):

$$\text{Area} = \frac{1}{2} \times \text{base} \times \text{height} = \frac{1}{2} \times 1 \times 0.5 = 0.25$$

- Total area from 0.5 to 1.5:

$$0.25 + 0.25 = 0.5$$

- This means:

$$P(0.5 \leq X \leq 1.5) = \int_{0.5}^{1.5} \frac{1}{2}x \cdot dx = 0.5$$

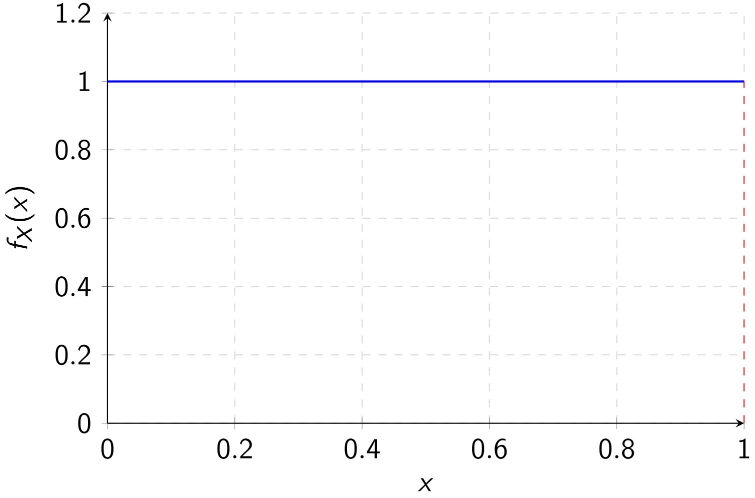
Uniform Distribution for Continuous Random Variables

- Consider the uniform distribution on the interval [0, 1].
- The PDF is:

$$f_X(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- The total area under the PDF curve is 1.

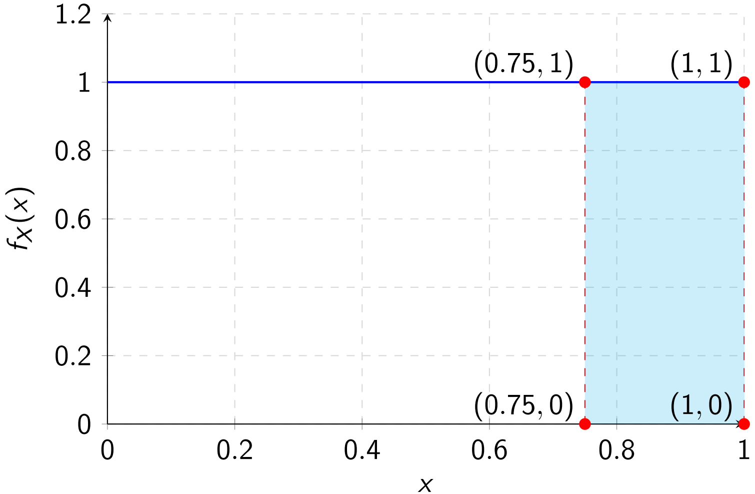
Graph of the Uniform Distribution PDF



Probability Calculation for Uniform Distribution

- The probability that X is greater than 0.75 in a uniform distribution from 0 to 1 is calculated as follows:

$$P(0.75 < X) = 1 \times (1 - 0.75) = 0.25$$



Introduction to the Normal Distribution

Normal Distribution

- The normal distribution is a continuous probability distribution.
- Defined by its mean μ and standard deviation σ.
- The PDF of the normal distribution is:

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

- Main features of the normal distribution:
  - It is a symmetrical distribution.
  - Values close to the mean are more likely than values very far from it.

Example: Standard Normal Distribution

- Consider a normal distribution with $\mu = 0$ and $\sigma = 1$ (standard normal distribution).
- The PDF is:
$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$
- The total area under the PDF curve is 1.

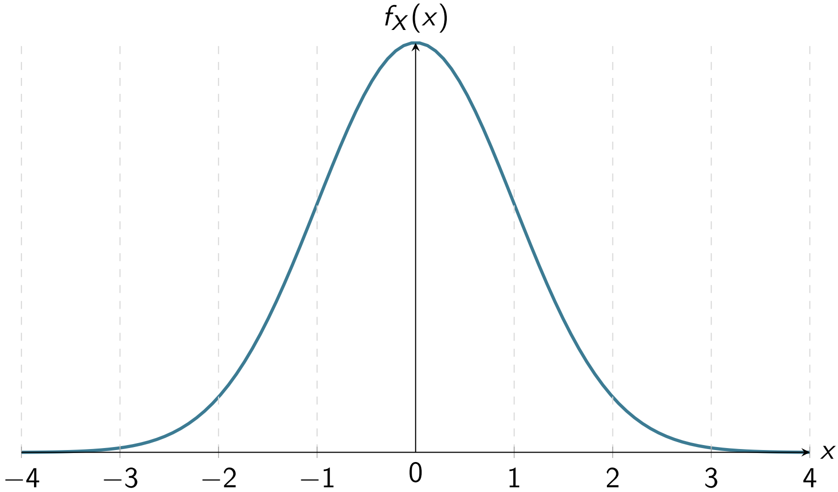


Figure: Standard Normal Distribution

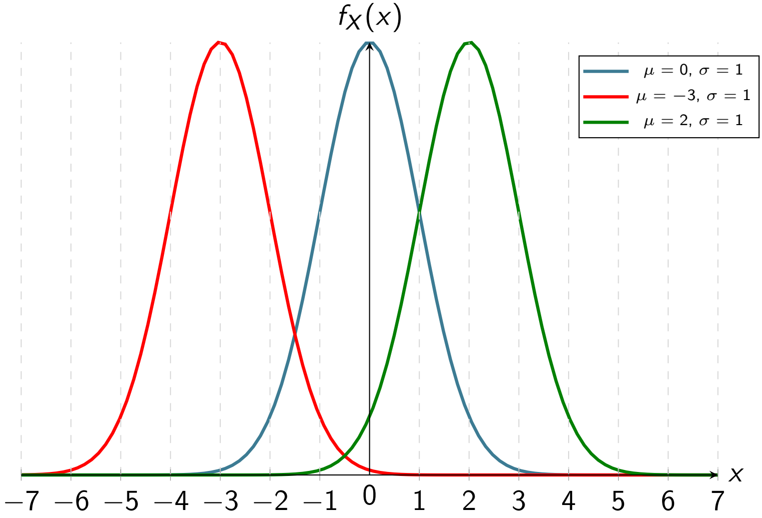


Figure: Comparison of Normal Distributions with Different Means

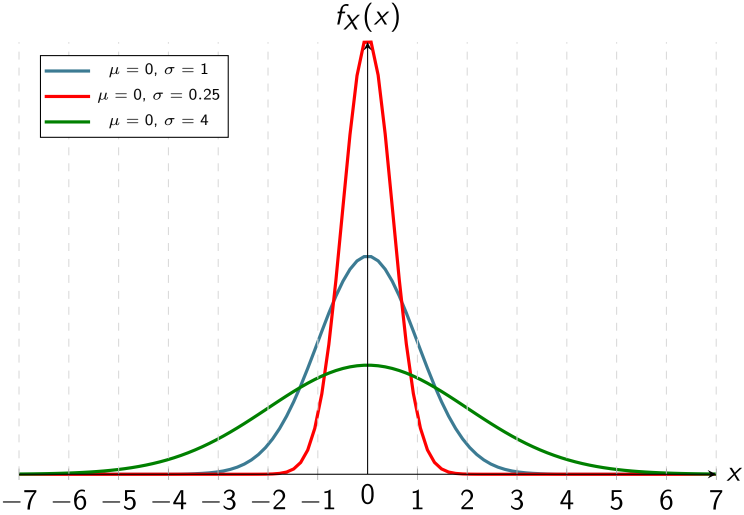


Figure: Comparison of Normal Distributions with Same Mean and Different Standard Deviations

Cumulative Distribution Function

Cumulative Distribution Function (CDF): Concept

- The Cumulative Distribution Function (CDF) is a function that describes the probability that a random variable takes a value less than or equal to a specific value.
- It provides a complete description of the probability distribution of a random variable.
- The CDF is fundamental in probability theory and statistics as it gives an integral overview of the distribution and helps in understanding the likelihood of different outcomes.

Cumulative Distribution Function (CDF): Formal Definition

Discrete Random Variable (RV)

- Definition:

$$F_X(x) = P(X \leq x) = \sum_{k:k\leq x} p_X(k)$$

- Here, $p_X(k)$ represents the probability mass function (PMF) of X at k.
- For here and after we will denote $F_X(x)$ by $CDF_X(x)$.

Continuous Random Variable (RV)

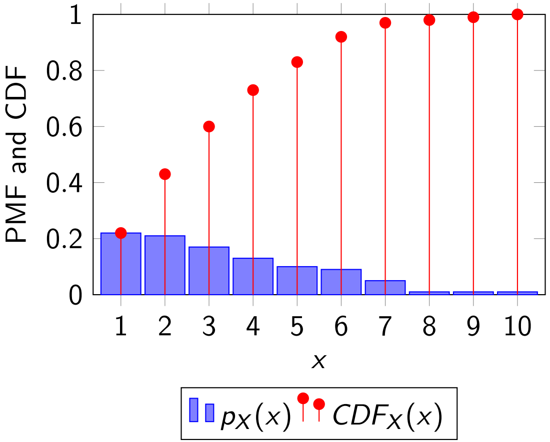
- Definition:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)\, dt$$

- $f_X(t)$ is the probability density function (PDF) of X.
- For here and after we will denote $F_X(x)$ by $CDF_X(x)$.

Example of CDF for a Discrete RV

- X = consecutive elections won by current members of the House of Representatives, with values from 1 to 10.
- We define a left-skewed PMF for X where each $p_X(x)$ decreases as x increases, reflecting a higher probability for fewer election wins.



Probability Rules Involving CDFs

- We know that every PMF satisfies:

$$\sum_{x \in D_X} p_X(x) = 1$$

- Let x_1 be the minimum, and x_n the maximum values in the domain of X, D_X. For an $a \in D_X$ such that $x_1 < a < x_n$, the sum of probabilities can be split at a:

$$\sum_{x=x_1}^{x_n} p_X(x) = \sum_{x=x_1}^{a} p_X(x) + \sum_{x=a+1}^{x_n} p_X(x) = 1$$

- This leads to the rule:

$$P(X \leq a) + P(a < X) = CDF_X(a) + P(a < X) = 1$$

- Therefore:

$$\boldsymbol{P(a < X) = 1 - CDF_X(a)}$$

Probability Rules with CDFs and Intervals

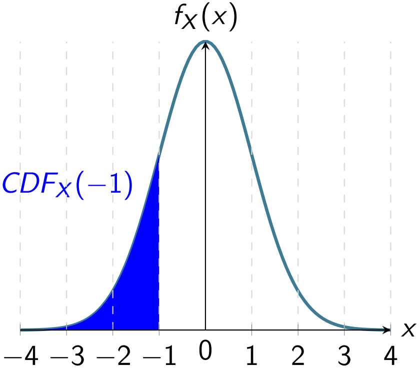
The previous derivation (proof) was shown *just as an illustration*. Proceeding similarly, one can show the following rules apply.

- $1 = P(X \leq a) + P(X > a)$
- $\boldsymbol{P(X > a) = 1 - CDF_X(a)}$
- $1 = P(X \geq a) + P(a > X > b) + P(X \leq b)$
- $P(a > X > b) = 1 - [P(X \geq a) + P(X \leq b)]$
- For $b < a$, $P(X \leq a) = P(X \leq b) + P(b < X \leq a)$
- $\boldsymbol{P(b < X \leq a) = CDF_X(a) - CDF_X(b)}$

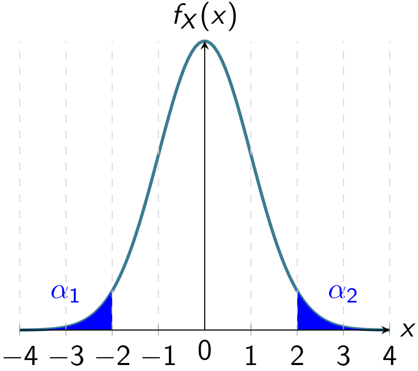
Note: These rules apply to both discrete and continuous random variables. For continuous random variables, the symbols $>$ and $<$, as well as $\geq$ and $\leq$, are used interchangeably because there is no practical distinction between them.

Probabilities with CDF: Example

- Let X be a standard normal random variable. Thus, $X \sim N(0, 1)$.
- Then, $\Pr(X \leq -1) = CDF_X(-1) = \Phi(-1) = 0.1586$.
- The "blue area" corresponds to a probability of 15.9%.
- In other words, the probability that X is lower or equal to -1 is 15.9%.



CDFs for Symmetric Distributions with $\mu_X = 0$



Symmetry and CDF

- Assume $a > 0$. For a symmetric PDF (or PMF) with $\mu_X = 0$, the following holds:

$$P(X \leq -a) = P(X \geq a)$$

- Hence:

$$CDF_X(-a) = 1 - CDF_X(a)$$

- e.g., if x is a Std. Normal, then:

$$0.023 = \alpha_1 = CDF_X(-2) = 1 - CDF_X(2) = \alpha_2 = 0.023$$

Symmetrical Distribution of a Discrete Random Variable

X	$p_X(x)$	$CDF_X(x)$
1	0.1	0.1
2	0.15	0.25
3	0.25	0.5
4	0.25	0.75
5	0.15	0.9
6	0.1	1.0