Lecture 8

Cumulative Distribution Functions and Transformations

Grace Tompkins

Last modified — 06 May 2026

Learning Outcomes

By the end of this lecture, students are anticipated to be able to:

- Define a cumulative distribution function (CDF)
- Calculate CDFs from PDF/PMFs
- Perform change of variables to calculate CDFs and PDFs/PMFs for functions of random variables.

1 Cumulative Distribution Functions (CDFs)

Cumulative Distribution Functions (CDFs)

📖 **DEFINITION**

The cumulative distribution function of a random variable X is

$$F_X(x) = \mathbb{P}(X \le x), \qquad \text{for } x \in \mathbb{R}$$

📐 **THEOREM**

Let X be a random variable with CDF F_X. Then:

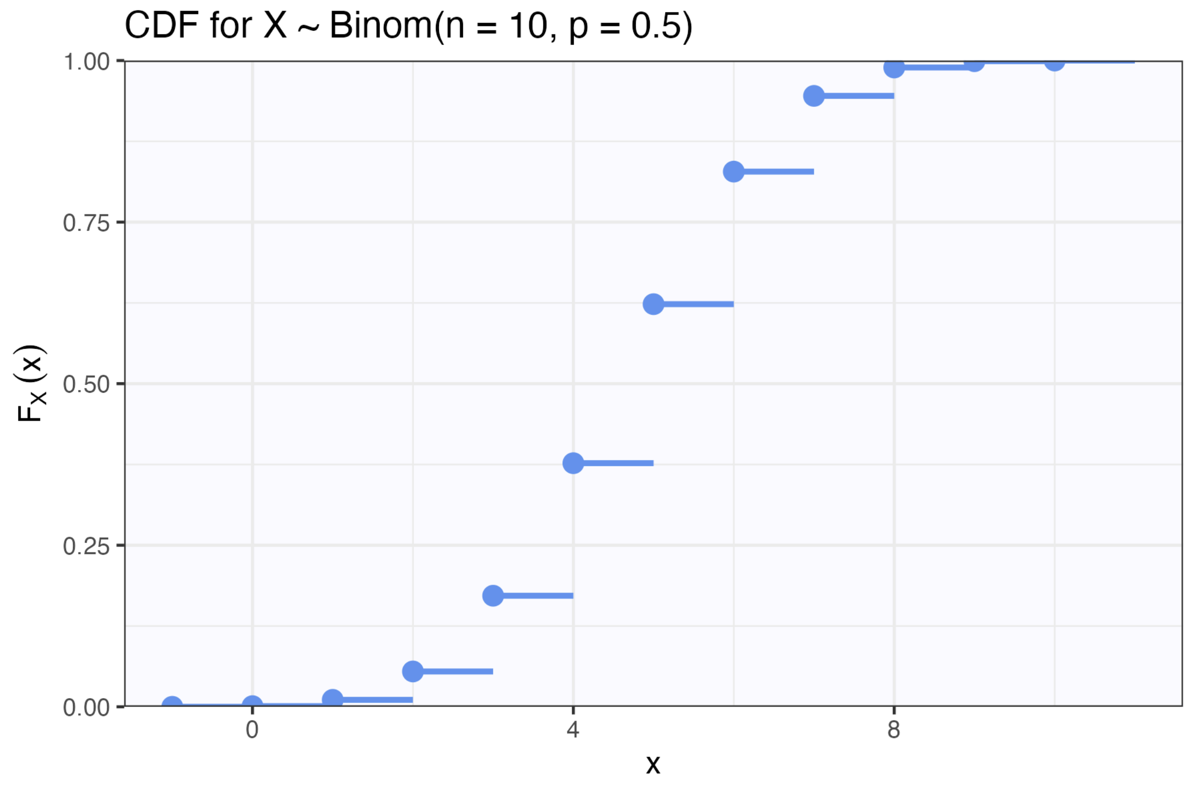
1. $0 \le F_X(x) \le 1$ for all $x \in \mathbb{R}$;
2. $F(x) \le F(y)$ for all $x \le y$;
3. $\lim_{a \to -\infty} F_X(a) = 0$,
4. $\lim_{a \to +\infty} F_X(a) = 1$

CDFs for Discrete and Continuous RVs

Discrete RV

$$F_X(x) = \sum_{t \le x} p_X(t)$$

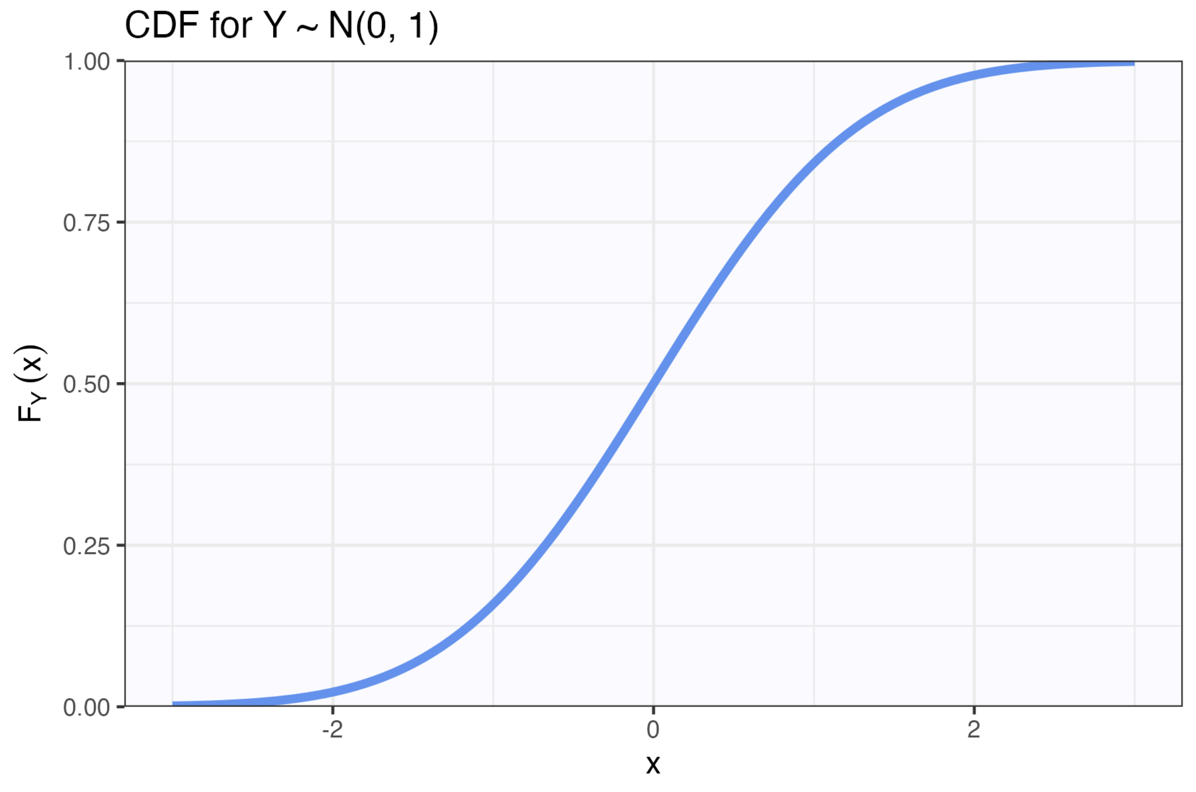
- The CDF is a step function and right-continuous.



Continuous RV

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

- The CDF is continuous (left and right limits are equal).



Continuous CDF Example

Let $X \sim \mathrm{Exp}(\lambda)$, with pdf

$$f_X(x) = \left(\lambda e^{-\lambda x}\right) I_{[0,\infty)}(x).$$

💡 EXAMPLE

Find $F_X(x)$. Use the CDF to calculate $\mathbb{P}(x \leq 10)$ when $\lambda = 1/5$.

$$F_X(x) = \int_{-\infty}^{x} \lambda e^{-\lambda t} I_{[0,\infty)}(t)\,dt = \int_{0}^{x} \lambda e^{-\lambda t}\,dt = \frac{\lambda e^{-\lambda t}}{-\lambda}\Bigg|_{t=0}^{t=x} = 1 - e^{-\lambda x}$$

$$F_X(x) = \left(1 - e^{-\lambda x}\right) I_{[0,\infty)}(x)$$

$$P(X \leq 10) = \left(1 - e^{-1/5(10)}\right) I_{[0,\infty)}(10) = 0.864$$

Continuous CDF Example

Discrete CDF Example

💡 EXAMPLE

Consider rolling one fair six-sided die, so that $\Omega = \{1, 2, 3, 4, 5, 6\}$, with $\mathbb{P}(\omega) = 1/6$ for each $\omega \in \Omega$. Let X be the number showing on the die. What is $F_X(x)$? Use this to find $\mathbb{P}(X \le 5)$.

$$F_X(x) = P(X \le x) = \sum_{\omega \in \Omega, \omega \le x} p(\omega) = \sum_{\omega \in \Omega, \omega \le x} 1/6 = \frac{1}{6} \left| \{\omega \in \Omega, \ \omega \le x\} \right|$$

number of elements

$$F_X(x) = P(X \le x) = \begin{cases} 0 & x < 1 \\ 1/6 & 1 \le x < 2 \\ 2/6 & 2 \le x < 3 \\ 3/6 & 3 \le x < 4 \\ \vdots & \\ 1 & x \ge 6 \end{cases}$$

$$P(X \le 3) = F_X(3) = 3/6$$

$$P(X \le 100) = F_X(100) = 1$$

CDFs and Distributions

📐 THEOREM

Let X be any random variable with CDF F_X. Let B be any subset of $\mathbb{R}$. Then $\mathbb{P}(X \in B)$ can be determined solely from F_X.

- The CDF contains all the information about the distribution of X.
- It doesn't matter whether X is discrete, continuous, or anything else.

📌 COROLLARY

Let X be an absolutely continuous random variable with cumulative distribution function F_X. Let

$$f_X(x) = \frac{d}{dx} F_X(x) = F'(x).$$

Then, $f(x)$ is the density function for X.

Aside: The discrete case involves finding the differences between consecutive CDF values.

Discrete example: assume we're given CDF $F_X(X)$.

$$P(X=5) = P(X \le 5) - P(X < 5)$$
$$= P(X \le 5) - P(X \le 4)$$
$$= F_X(5) - F_X(4)$$

(try this with dice example!)

Mixture Distributions

📚 **PROPOSITION**

Let $X_1, X_2, \ldots$ be a random variables with CDFs $F_{X_1}, F_{X_2}, \ldots$.

For any constants p_i such that $p_i \geq 0$ and $\sum_{i=1}^{k} p_i = 1$, $F_G(x) = \sum_{i=1}^{k} p_i F_{X_i}(x)$ is the CDF of the **mixture** of F_{X_i}.

(p_i are like weights)

$G = X_1 + X_2$

$F_G(x) = p_1 F_{X_1}(x) + p_2 F_{X_2}(x)$

Mixture Distributions

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

💡 EXAMPLE

Suppose a bag contains two coins. You choose one at random and flip it once. Let X be the number of heads.

Coin A is a fair coin, with $\mathbb{P}(\text{flip Heads}) = 0.5$ Coin B is a loaded coin, with $\mathbb{P}(\text{flip Heads}) = 0.9$

Write out the CDF for the number of heads flipped.

Each coin has a probability of being chosen = 0.5 $\Rightarrow$ $p_A = 0.5$
$p_B = 0.5$

$\boxed{\text{Case 0: } x \leq 0}$ If a coin is flipped, cannot have <0 heads.

$F_G(x) = 0$ when $x \leq 0$

$\boxed{\text{Case 1: } 0 \leq x < 1}$

$F_A(x) = P(X < 1) = P(X=0) = 0.5$

$F_B(x) = P(X<1) = P(X=0) = 1-0.9 = 0.1$

$F_G(x) = p_A F_A(x) + p_B F_B(x) = 0.5(0.5) + 0.5(0.1) = 0.30$

when $0 \leq x < 1$

Mixture Distributions

$\boxed{\text{Case 2: } x \geq 1}$

The number of heads can only be 0, or 1.
All possible outcomes are captured once
$x \geq 1$. $F_A(x) = 1$ when $x \geq 1$

$F_B(x) = 1$ when $x \geq 1$

$\therefore$

$$F_G(x) = 0.5(1) + 0.5(1) = 1$$

$$F_G(x) = \begin{cases} 0 & \text{if } x < 0 \\ 0.30 & \text{if } 0 \leq x < 1 \\ 1 & \text{if } x \geq 1 \end{cases}$$

Mixture Distributions

Consider the scores of a midterm. Suppose that there are three types of students, modeled as follows:

1. Poorly prepared students who did not study much (30%): $X_1 \sim \mathcal{N}(\mu = 50, \sigma = 10)$.
2. Well prepared students who studied a lot (60%): $X_2 \sim \mathcal{N}(\mu = 80, \sigma = 8)$.
3. Students who didn't take the exam at all (10%): $X_3 = 0$ with probability 1.

✏️ **EXERCISE: MIXTURE DISTRIBUTION FOR MIDTERM**

Find the CDF of the overall score distribution X. Write your answer in terms of the standard normal CDF $\Phi(\cdot)$.

Hint: If $Y \sim \mathcal{N}(\mu, \sigma)$, then $F_Y(y) = \Phi\left(\frac{y-\mu}{\sigma}\right)$.

Mixture Distributions

$$F_X(x) = 0.3F_{X_1}(x) + 0.6F_{X_2}(x) + 0.10F_{X_3}(x)$$

$$= 0.3\,\Phi\left(\frac{x-50}{10}\right) + 0.6\,\Phi\left(\frac{x-80}{8}\right) + 0.1\,I_{[0,\infty)}(x)$$

2 Transformations of CDFs

One-Dimensional Change of Variables

Sometimes we want to perform transformations on a random variable. This is when we can turn to **change of variables**.

- Let X be a random variable with some distribution
- Let $Y = g(X)$ for some function $g : \mathbb{R} \to \mathbb{R}$.
- We want to find the distribution of Y.

There are two ways to do this: the **distribution method** and the **Jacobian method**. We will present both.

One-Dimensional Change of Variables: Distribution Method

The **distribution method** just uses the definition:

$$\mathbb{P}(Y \in A) = \mathbb{P}(g(X) \in A) = \mathbb{P}(\{x : g(x) \in A\}).$$

If we can characterize these sets, we can find the distribution of Y. This method always works, and is easy for discrete random variables.

> **THEOREM**
>
> Let X be a discrete random variable with PMF $p_X(x)$. Let $Y = g(X)$ for some function $g : \mathbb{R} \to \mathbb{R}$. Then the PMF of Y is given by
>
> $$p_Y(y) = \sum_{x:g(x)=y} p_X(x) = \sum_{x \in g^{-1}(y)} p_X(x).$$

One-Dimensional Change of Variables: Distribution Method

Let's make this method a little more algorithmic.

If you want to find the distribution of $Y = g(X)$, then you can do the following:

$$\mathbb{P}(Y \in A) = \mathbb{P}(g(X) \in A) = \mathbb{P}(\{x : g(x) \in A\}) = \cdots$$

1. Start with the definition of the distribution of Y.
2. Use the definition of Y in terms of X to rewrite the probability in terms of X.
3. Use the distribution of X to compute the probability that $g(X) \in A$.
4. Then play algebra games to get it into a nice form; possibly "match" a known PDF or CDF.

One-Dimensional Change of Variables: Distribution Method (Discrete)

Let's start with a straightforward example:

💡 EXAMPLE

- Let $X \sim \mathrm{Binom}(n, \theta)$, for some $\theta \in (0, 1)$.
- Let $Y = n - X$. $x = n - y$

Find the PMF of Y.

$$f_X(x) = \binom{n}{x} \theta^x (1-\theta)^{n-x} I_{\{0,1,\dots,n\}}(x)$$

One-Dimensional Change of Variables: Distribution Method (Discrete)

$X \sim \text{Binom}(n, \theta)$
$Y = n - X$

$$p_Y(y) = \sum_{x: g(x) = y} p_X(x) = \sum_{x: (n-x) = y} p_X(x)$$

$Y = n - X$
$\Rightarrow X = n - Y$

$$= p_X(n-y) = \binom{n}{n-y} \theta^{n-y} (1-\theta)^{y} I_{\{0,1,\dots,n\}}(n-y)$$

Note: Symmetry of binomial coeff $\binom{n}{k} = \binom{n}{n-k}$ $\binom{10}{9} = \binom{10}{1}$

$$= \binom{n}{y} (1-\theta)^y (\theta)^{n-y} I_{\{0,1,\dots,n\}}(y)$$

$(1-(1-\theta))$

$Y \sim \text{Bin}(n, 1-\theta)$

One-Dimensional Change of Variables: Distribution Method (Discrete)

Addendum: **Note** on why the indicator changed:

We have $x \in \{0, 1, 2, \ldots, n-1, n\}$

$Y = n - X$ implies $y \in \{n-0, n-1, \ldots, 0\} = \{n, n-1, \ldots, 2, 1, 0\}$

$\equiv y \in \{0, 1, 2, \ldots, n-1, n\}$

One-Dimensional Change of Variables: Distribution Method (Continuous)

💡 EXAMPLE

Let $X \sim \mathrm{Exp}(\lambda)$, find the distribution of $Y = 3X$. (CDF)

$X: (0, \infty)$

$3X: (0, \infty)$

$$P(Y \le y) = P(3X \le y)$$
$$= P(X \le y/3)$$
$$= F_X(y/3)$$
$$= 1 - e^{-\lambda(y/3)} I_{[0,\infty)}(y)$$

→ CDF of $\mathrm{Exp}(\lambda)$

$$= 1 - e^{-(\lambda/3)y} I_{[0,\infty)}(y)$$

$Y \sim \mathrm{Exp}(\lambda/3)$

One-Dimensional Change of Variables: Distribution Method (Continuous)

> 💡 **EXAMPLE**
>
> If $X \sim \mathrm{Exp}(\lambda)$, what is the CDF of $Y = X + 5$?

$$
\begin{aligned}
P(Y \le y) &= P(X + 5 \le y) \\
&= P(X \le y - 5) \\
&= F_X(y-5) \\
&= 1 - e^{-\lambda(y-5)} I_{[5,\infty)}
\end{aligned}
$$

Support changes!

$X: [0, \infty)$

$y = x+5: [5, \infty)$

One-Dimensional Change of Variables: Useful Results

The following useful results follow from the previous examples

📚 **PROPOSITION**

If X be a random variables with CDFs F_X Then the the following hold:

1. The RV $Y = X + c$ has CDF $F_X(x - c)$;
2. The RV $Y = kX$ has CDF $F_X(x/k)$ for any $k > 0$;

One-Dimensional Change of Variables: Jacobian Method

For continuous random variables, you can also use the distribution method, and sometimes this is the easiest way. The other common method is the **Jacobian method**.

THEOREM

Let X be an (absolutely) continuous random variable, with density function f_X. Let $Y = h(X)$, where $h : \mathbb{R} \to \mathbb{R}$ is a function that is differentiable and monotonic.

Then Y is also absolutely continuous, and its density function f_Y is given by

$$f_Y(y) = f_X(h^{-1}(y)) \left| \frac{\mathrm{d}}{\mathrm{d}y}(h^{-1}(y)) \right|,$$

where $h^{-1}(y)$ is the unique number x such that $h(x) = y$.

One-Dimensional Change of Variables: Jacobian Method

💡 EXAMPLE

- Let $X \sim \text{Unif}(0, 1)$.
- Let $Y = -\log(X)$. → monotonic and diff. on $(0, \infty)$. Can use Jac. method ⸮

Find the PDF of Y.

$Y = h(X) = -\log(X)$

↳ $h^{-1}(x) = e^{-x}$

↳ $h^{-1}(y) = e^{-y}$

↳ $\frac{d}{dy} h^{-1}(y) = -e^{-y}$

$f_X(x) = 1 \cdot I_{[0,1]}(x)$

$$f_Y(y) = f_X(h^{-1}(y)) \left| \frac{d}{dy}(h^{-1}(y)) \right|$$

$$= f_X(e^{-y}) \left| \frac{d}{dy}(e^{-y}) \right|$$

$$= (1)\left(I_{[0,1]}(e^{-y})\right) \left| -e^{-y} \right|$$

$$= I_{[0,\infty)}(y) \cdot e^{-y} \Rightarrow Y \sim \text{Exp}(1)$$

One-Dimensional Change of Variables: Jacobian Method

✏️ **EXERCISE:**

1. Let $X \sim \mathrm{Unif}\,(-1, 1)$. Use the distribution method to find the PDF of $Z = X^2$.
2. Let $X \sim \mathrm{Gam}(\alpha, \lambda)$. Use the Jacobian method to find the PDF of $Y = 1/X$.

Recall the Gamma PDF:

$$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} I_{(0,\infty)}(x).$$

If $X \sim \mathrm{Unif}(L, R)$ $\quad f_X(x) = \frac{1}{R-L} I_{[L,R]}(x)$ $\quad$ where $L < R$

One-Dimensional Change of Variables: Jacobian Method

1. $h(x)$ is not monotonic → need to use Dist. method.

$Z = h(X) = X^2$: range of X: $[-1, 1]$

Z: $[0, 1]$

For $z \in (0,1)$ $F_Z(z) = P(Z \le z) = P(X^2 \le z)$

$= P(-\sqrt{z} \le X \le \sqrt{z})$

$= \int_{-\sqrt{z}}^{\sqrt{z}} \left(\frac{1}{1-(-1)}\right) dt = \frac{1}{2} t \Big|_{t=-\sqrt{z}}^{t=\sqrt{z}} = \sqrt{z}\, I_{[0,1]}(z)$

PDF: $f_Z(a) = \frac{d}{dz} F_Z(z) = \frac{1}{2} z^{-1/2} I_{(0,1]}(z)$

← cannot $\div$ 0

One-Dimensional Change of Variables: Jacobian Method

2. $h(x) = 1/x$ is monotonic

$$f_Y(y) = f_X(h^{-1}(y)) \left| \frac{d}{dy} h^{-1}(y) \right|$$

$Y = h(x) = 1/x$

$\hookrightarrow h^{-1}(y) = 1/y$

$\hookrightarrow \frac{d}{dy}(h^{-1}(y)) = \frac{-1}{y^2}$

$$f_Y(y) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \left(\frac{1}{y}\right)^{\alpha-1} e^{-(1/y)\lambda} \cdot \left| \frac{-1}{y^2} \right| I_{(0,\infty)}(y)$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{-\alpha-1} e^{-\lambda/y} I_{(0,\infty)}(y)$$

$\hookrightarrow$ "Inverse Gamma"

This is the end of the "testable" content for your midterm.

- Your midterm will cover materials from Lectures 1 - 8
- DATE: in class on Tuesday June 2
- LENGTH: 1 hour and 50 minutes in length.
- You may bring in one (1) "cheat sheet":
  - Must be HAND WRITTEN with pen/pencil on said sheet of paper (not typed, not photo copied, not printed, not written on an iPad)
  - Must be on 8.5 by 11 inch sheet of paper or smaller
  - You may write on both sides
  - No magnifying glasses or anything else silly
  - **I will confiscate cheatsheets that do not follow these rules** 🥀
  - I do not care what is written on it
- Exam is hand written on paper, bring something to write with
- You may bring a non-programmable, non-graphing calculator.

To do:

- Read Chapter 2.7 before next class
- Assignment 2 due tonight, May 27th @ 11:59pm.
- Start reviewing for your midterm. No assignment due next week.