

# Lecture 11

Conditional Distributions and Independence

**Grace Tompkins (Presented by Dr. Geoff Pleiss)**

Last modified — 06 May 2026

# Learning Outcomes

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

- Calculate conditional distributions from a joint distribution
- Relate the notion of independence to conditional and joint distributions

# 1 Conditional Distributions

# Conditioning on Discrete RVs

## DEFINITION

Let  $Y$  be a random variable, where  $X$  is a discrete random variable. Suppose that  $\mathbb{P}(X = x) > 0$  for some  $x$ . Then the **conditional distribution** of  $Y$  given  $X = x$  assigns to each set  $A \subset \mathbb{R}$  the probability

$$\mathbb{P}(Y \in A \mid X = x) = \frac{\mathbb{P}(Y \in A, X = x)}{\mathbb{P}(X = x)}.$$

Here, we do not specify whether  $Y$  is continuous or discrete.

# Conditioning on Discrete RVs

## DEFINITION

If  $X$  and  $Y$  are *both* discrete random variables, then the **conditional PMF** of  $Y$  given  $X = x$  is defined by

$$p_{Y|X}(y | x) = \frac{\mathbb{P}(Y = y, X = x)}{\mathbb{P}(X = x)} = \frac{p_{X,Y}(x, y)}{p_X(x)}.$$

# Conditioning on Discrete RVs

- Let  $X$  and  $Y$  be the results of rolling two fair 6-sided dice.
- Let  $V = X + Y$  and  $W = \max\{X, Y\}$ .

The joint PMF of  $W$  and  $V$  is

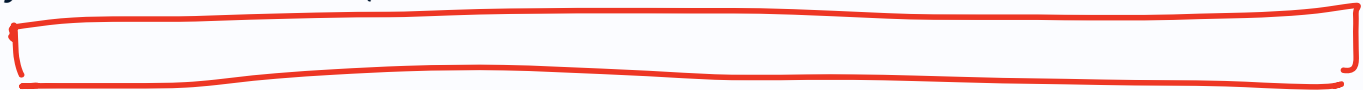
$W \setminus V$	2	3	4	5	6	7	8	9	10	11	12
1	1/36	0	0	0	0	0	0	0	0	0	0
2	0	2/36	1/36	0	0	0	0	0	0	0	0
3	0	0	2/36	2/36	1/36	0	0	0	0	0	0
4	0	0	0	2/36	2/36	2/36	1/36	0	0	0	0
5	0	0	0	0	2/36	2/36	2/36	2/36	1/36	0	0
6	0	0	0	0	0	2/36	2/36	2/36	2/36	2/36	1/36

sum  
 $P_W(w)$



To condition on  $W$ , we look at a particular **row**, while to condition on  $V$ , we look at a particular **column**. Then renormalize by the sum of the row/column.

$P_V(v)$



# Conditioning on Discrete RV

Handwritten notes: "SUM of 2 dice" (above the table), "1st dice" (with an arrow pointing to the first column).

$W \setminus V$	2	3	4	5	6	7	8	9	10	11	12
1	1/36	0	0	0	0	0	0	0	0	0	0
2	0	2/36	1/36	0	0	0	0	0	0	0	0
3	0	0	2/36	2/36	1/36	0	0	0	0	0	0
4	0	0	0	2/36	2/36	2/36	1/36	0	0	0	0
5	0	0	0	0	2/36	2/36	2/36	2/36	1/36	0	0
6	0	0	0	0	0	2/36	2/36	2/36	2/36	2/36	1/36

## EXAMPLE

Given the above table, what is  $\mathbb{P}(W = 5 \mid V = 8)$

$$P(W=5 \mid V=8) = \frac{P(W=5, V=8)}{P(V=8)}$$

$$= \frac{\frac{1}{36} + \frac{2}{36} + \frac{2}{36}}{\frac{5}{36}} = \frac{2/36}{5/36} = \frac{2}{5}$$

# Conditioning on Discrete RVs

# Conditioning on Continuous RVs

## DEFINITION

If  $X$  and  $Y$  are jointly absolutely continuous random variables, then the **conditional density** of  $Y$  given  $X = x$ , is the function

$$f_{Y|X}(y | x) = \frac{f_{X,Y}(x, y)}{f_X(x)},$$

valid for any  $y \in \mathbb{R}$ , and all  $x$  such that  $f_X(x) > 0$ .

# Conditioning on Continuous RVs

## DEFINITION

Let  $X$  and  $Y$  be jointly absolutely continuous random variables with joint PDF  $f_{X,Y}$ . The **conditional distribution** of  $Y$  given  $X = x$  assigns to each set  $A \subset \mathbb{R}$  the probability

$$\mathbb{P}(a \leq Y \leq b \mid X = x) = \int_a^b f_{Y|X}(y \mid x) \, dy,$$

valid for all  $x$  such that  $f_X(x) > 0$ .

# Conditioning on Continuous RVs

## EXAMPLE

Let  $X$  and  $Y$  be jointly continuous random variables with joint PDF

$$f_{X,Y}(x, y) = \frac{1}{x} e^{-x} I_{\{0 \leq y \leq x\}}(x, y).$$

a. Find the marginal PDF of  $X$ . What distribution does  $X$  have?

b. Find the conditional PDF of  $Y$  given  $X = x$ . What distribution does  $Y | X$  have?

$$f_X(x) = \int f_{X,Y}(x, y) dy$$

$$\begin{aligned} f_X(x) &= \int f_{X,Y}(x, y) dy \\ &= \int_0^x \frac{1}{x} e^{-x} dy \\ &= \frac{1}{x} e^{-x} \int_0^x dy = \frac{1}{x} e^{-x} x = e^{-x} \end{aligned}$$

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)}$$

$$= \frac{\frac{1}{x} e^{-x} I_{\{0 \leq y \leq x\}}}{e^{-x}}$$

$$= \frac{1}{x} I_{\{0 \leq y \leq x\}}$$

# Conditioning on Continuous RVs

# Conditioning on Continuous RVs

 EXERCISE: TRY THIS ONE

Let  $X$  and  $Y$  be independent with  $X \sim \text{Exp}(1)$  and  $Y \sim \text{Exp}(2)$ . Let  $S = X + Y$ . Find the conditional density  $f_{X|S}(X = x | S = s)$

# Conditioning on Continuous RVs

# Conditioning on Continuous RVs

# Heuristics Types of Distributions of Random Variables

- Suppose that  $X$  and  $Y$  have a joint distribution  $\mathbb{P}(X \in A, Y \in B)$ .
- We can find the joint CDF, PMF, or PDF of  $X$  and  $Y$ .
- The **marginal** distribution of  $X$  is the distribution of  $X$  when we “ignore”  $Y$ .
- The **conditional** distribution of  $X$  given  $Y = y$  is the distribution of  $X$  when we “fix”  $Y$  to be  $y$ .
- The **marginal** comes from “summing out” or “integrating out” the other variable along the rows or columns of the table.
- The **conditional** comes from “dividing out” the other variable. Fix a row or a column, and then renormalize (divide out).
- When we say  $Y \mid X = x$ , we usually specify a formula. Meaning, we write down a function that gives the conditional PMF or PDF of  $Y$  given  $X = x$  for all  $x$  such that  $p_X(x) > 0$  or  $f_X(x) > 0$ . Renormalizing the row of a table gives the conditional PMF for a specific value of  $x$ .

# 2 Independence

# Independent Random Variables

## THEOREM

$X$  and  $Y$  are independent if and only if for any sets  $A$  and  $B$  we have

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A) \mathbb{P}(Y \in B).$$

This is the definition of independence. It says that the joint distribution factors into the product of the marginals.

But this has immediate consequences for the joint CDF, PMF, and PDF.

Choosing  $A = (-\infty, x]$  and  $B = (-\infty, y]$ , we have that, if  $X$  and  $Y$  are independent, then

$$F_{X,Y}(x, y) = \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x) \mathbb{P}(Y \leq y) = F_X(x) F_Y(y).$$

The converse is also true: if  $F_{X,Y}(x, y) = F_X(x) F_Y(y)$  for all  $x, y$ , then  $X$  and  $Y$  are independent.

$$\begin{aligned} & \mathbb{P}(X \in (-\infty, x], Y \in (-\infty, y]) \\ &= F_{X,Y}(x, y) \end{aligned}$$

# Independent Random Variables, Using PMFs

## THEOREM

If  $X$  and  $Y$  are discrete random variables, then  $X$  and  $Y$  are independent if and only if

$$\underline{p_{X,Y}(x,y) = p_X(x)p_Y(y)},$$

for all  $x, y \in \mathbb{R}$ .

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

# Independent Random Variables, Using PDFs

## THEOREM

If  $X$  and  $Y$  are jointly continuous random variables, then  $X$  and  $Y$  are independent if and only if their joint density can be chosen such that

$$f_{X,Y}(x, y) = f_X(x)f_Y(y),$$

for all  $x, y \in \mathbb{R}$ .

- Regardless of continuous or discrete, when you hear “independent”, think “joint factors into the product of the marginals”.

$$f_{y|x}(y|x) = \frac{f_{x,y}(x,y)}{f_x(x)} = \frac{\cancel{f_x(x)} f_y(y)}{\cancel{f_x(x)}} = f_y(y)$$

# Independence

## EXERCISE: EVALUATING INDEPENDENCE USING MARGINAL DISTRIBUTIONS

Let  $X$  and  $Y$  have joint pdf

$$f_{X,Y}(x,y) = \begin{cases} 8xy & \text{if } 0 \leq x < y < 1 \\ 0 & \text{else.} \end{cases}$$

Find the marginal PDFs of  $X$  and  $Y$ . Are  $X$  and  $Y$  independent?

$$f_X(x) = \int_x^1 8xy \, dy = 8x \left. \frac{y^2}{2} \right|_{y=x}^{y=1} = 4x(1-x^2) I_{[0,1]}(x)$$

$$f_Y(y) = \int_0^y 8xy \, dx = 8y \left. \frac{x^2}{2} \right|_{x=0}^{x=y} = 4y^3 I_{(0,1)}(y)$$

$$f_X(x)f_Y(y) = 4x(1-x^2) 4y^3 I_{[0,1]}(x) I_{(0,1)}(y)$$

# Independence

# Maximum of Independent RVs

## 💡 EXAMPLE

Let  $X$  and  $Y$  be independent random variables with CDFs  $F_X$  and  $F_Y$ .

Let  $W = \max\{X, Y\}$ .

Find the CDF of  $W$ .

$$\begin{aligned} F_W(w) &= \mathbb{P}(W \leq w) = \mathbb{P}(\max\{X, Y\} \leq w) \\ &= \mathbb{P}(X \leq w, Y \leq w) = F_{X,Y}(w, w) \\ &= \underbrace{F_X(w)}_{\mathbb{P}(X \leq w)} \underbrace{F_Y(w)}_{\mathbb{P}(Y \leq w)} \end{aligned}$$

# Maximum of Independent RVs

# Maximum of Independent RVs

The previous example also extends to the maximum of  $n$  independent random variables.

Suppose that  $X_1, X_2, \dots, X_n$  are independent random variables with **common** CDF  $F_X$ . Let  $W = \max\{X_1, X_2, \dots, X_n\}$ . Then

$$F_W(w) = \mathbb{P}(W \leq w) = \mathbb{P}(X_1 \leq w, X_2 \leq w, \dots, X_n \leq w) = (F_X(w))^n,$$
$$\implies f_W(w) = \frac{d}{dw} F_W(w) = n(F_X(w))^{n-1} f_X(w). \quad (\text{by the chain rule if } X \text{ is continuous})$$

# Minimum of Two Independent Random Variables

## EXERCISE: MINIMUMS

Let  $X \sim \text{Exp}(\lambda)$  and  $Y \sim \text{Exp}(\mu)$  be independent random variables. Find the distribution of  $U = \min\{X, Y\}$ .

Hints:

- If  $Z \sim \text{Exp}(\theta)$ , then  $F_Z(z) = \mathbb{P}(Z \leq z) = 1 - e^{-\theta z}$  for  $z > 0$ .
- Therefore,  $\mathbb{P}(Z > z) = e^{-\theta z}$  for  $z > 0$ .

$$\begin{aligned}\mathbb{P}(X > x, Y > y) &= \mathbb{P}(X > x) \mathbb{P}(Y > y) \\ &= e^{-\lambda x} e^{-\mu y}\end{aligned}$$

# Minimum of Two Independent Random Variables

$$U = \min\{X, Y\}$$

$$\begin{aligned} P(U \leq u) &= P(\min\{X, Y\} \leq u) \\ &= 1 - P(\min\{X, Y\} \geq u) \\ &= 1 - P(X > u, Y > u) \\ &= 1 - \underbrace{P(X > u)}_{e^{-\lambda u}} \underbrace{P(Y > u)}_{e^{-\mu u}} \\ &= 1 - e^{-\lambda u} e^{-\mu u} = 1 - e^{-(\lambda + \mu)u} \end{aligned}$$

# Sums of Independent Random Variables

To find the distribution of a sum of independent random variables, we could use the distribution method or the Jacobian method.

For this specific case where RVs are independent, there's also a third method called [convolution](#).

$$X + Y = U$$

## THEOREM

Let  $X$  and  $Y$  be independent random variables.

If  $X$  and  $Y$  are both discrete random variables, then the PMF of  $U = X + Y$  is given by

$$p_U(u) = \sum_w \underbrace{p_X(u-w)}_{X=u-w} \underbrace{p_Y(w)}_{Y=w} = \sum_w p_Y(u-w)p_X(w).$$

If  $X$  and  $Y$  are both continuous random variables, then the PDF of  $U = X + Y$  is given by

$$f_U(u) = \int_{-\infty}^{\infty} f_X(u-w)f_Y(w)dw = \int_{-\infty}^{\infty} f_Y(u-w)f_X(w)dw.$$

# Sums of Independent Random Variables

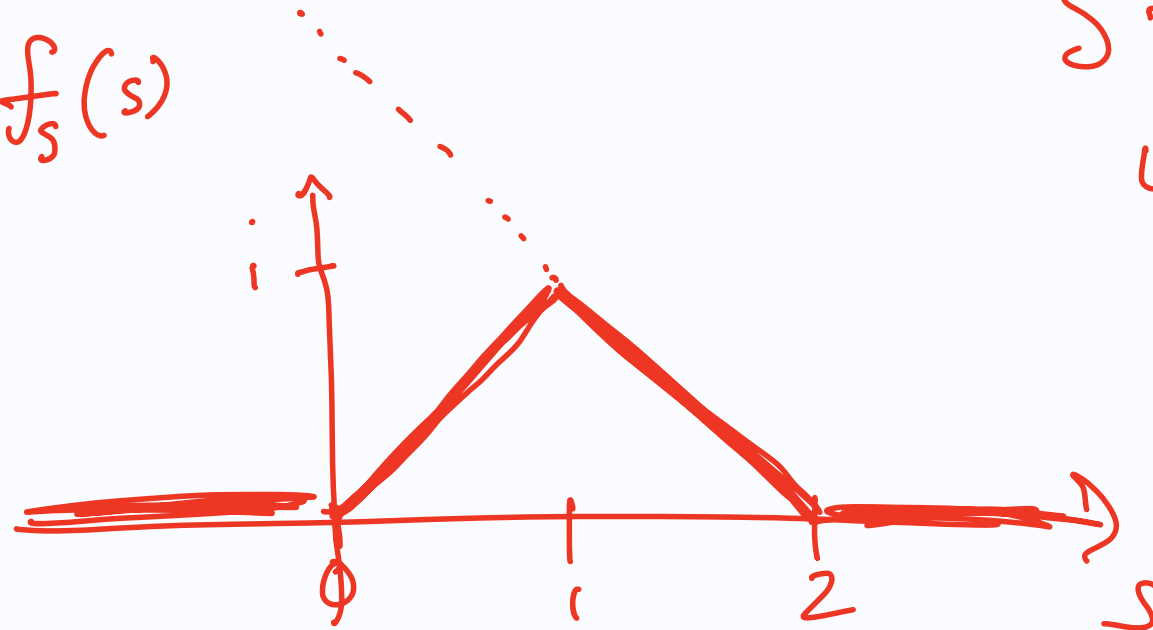
## 💡 EXAMPLE

Let  $X$  and  $Y$  be independent  $\text{Unif}(0, 1)$  random variables. Let  $S = X + Y$ . Find the distribution of  $S$ .

$$\begin{aligned} f_S(s) &= \int_{-\infty}^{\infty} f_X(\omega) f_Y(s-\omega) d\omega = \int_{-\infty}^{\infty} I_{(0,1)}(\omega) I_{(0,1)}(s-\omega) d\omega \\ &= \int_{-\infty}^{\infty} I_{(0,1)}(\omega) I_{(s-1,s)}(\omega) d\omega \\ &= \begin{cases} \int_0^s d\omega = s & 0 < s < 1 \\ \int_{s-1}^1 d\omega = 2-s & 1 \leq s < 2 \\ 0 & \text{if } s \leq 0, s \geq 2 \end{cases} \end{aligned}$$

# Sum of Independent Random Variables

$f_S(s)$



$$S = \underbrace{X}_{\text{Unif}(0,1)} + \underbrace{Y}_{\text{Unif}(0,1)}$$

Triangle Distribution

# Sum of Independent Random Variables


## EXERCISE: INDEPENDENT EXPONENTIALS

Let  $X$  and  $Y$  be independent  $\text{Exp}(\lambda)$  random variables,. Find the PDF of  $U = X + Y$  using the convolution method.

Hint: be careful with the limits of integration. Recall that if  $Z \sim \text{Exp}(\theta)$ , then  $f_Z(z) = \theta e^{-\theta z} I_{(0, \infty)}(z)$ .

# Sum of Independent Random Variables

# To Do

- Work on Assignment 3, due Wednesday June 10, 11:59pm on Gradescope.
- Read **Chapter 3.1 and 3.2**  before next class.
- Grace will be back for your next class! Please save your questions for her and/or the TAs where possible.