Lecture 11

Conditional Distributions and Independence

Grace Tompkins (Presented by Dr. Geoff Pleiss)

27 May 2026

0.1 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

1.1 Conditioning on Discrete RVs

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.

1.2 Conditioning on Discrete RVs

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

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

1.3 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

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.

1.4 Conditioning on Discrete RV

$W \setminus V$	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

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

Solution

$$\begin{aligned} p_{W|V}(W = 5|V = 8) &= \frac{p_{W,V}(5, 8)}{p_V(8)} \\ &= \frac{2/36}{1/36 + 2/36 + 2/36} \\ &= \frac{2/36}{5/36} \\ &= 2/5 \end{aligned}$$

1.5 Conditioning on Discrete RVs

1.6 Conditioning on Continuous RVs

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

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

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

1.7 Conditioning on Continuous RVs

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

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

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

1.8 Conditioning on Continuous RVs

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 $Y \mid X$ have?

ANSWERS

a. We have that

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) \mathsf{d}y = \int_0^x \frac{1}{x} e^{-x} \mathsf{d}y = e^{-x} \qquad \text{for } x \geq 0,$$

and $f_X(x) = 0$ for $x < 0$. Therefore, $X \sim \text{Exp}(1)$.

b. We have that

$$f_{Y|X}(y; x) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \frac{e^{-x}/x}{e^{-x}} = \frac{1}{x} \qquad \text{for } 0 \leq y \leq x,$$

and $f_{Y|X}(y; x) = 0$ otherwise. Note that this is the PDF of a $\text{Unif}(0, x)$ distribution, so $Y|X = x \sim \text{Unif}(0, x)$.

Note: x operates as a parameter in the conditional distribution of Y given $X = x$. It is no longer random, and we can treat it as a constant when working with the conditional distribution of Y given $X = x$.

1.9 Conditioning on Continuous RVs

1.10 Conditioning on Continuous RVs

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 $f_{X|S}(X = x|S = s)$

Give students 10 minutes to try this first

Solution

Step 1: find joint density

Since X and Y are independent, the joint density is:

$$f_{X,Y}(x, y) = e^{-x} 2e^{-2y} I_{(0,\infty)}(x) I_{(0,\infty)}(y)$$

Then, recall $X + Y = S$, which implies $Y = S - X$. This means conditioning on $S = s$ means we're fixing $y = s - x$. Thus, we can subsitute S - X in for Y and get the joint density of X and S:

$$f_{X,S}(x, s) = e^{-x} 2e^{-2(s-x)} I_{(0,s)}(x) = 2e^{x-2s} I_{(0,s)}(x)$$

The support changes as having $y > 0$ impies $s - x > 0$ which implies $s > x$ We still also have $x > 0$, so we can combine this into a single indicator. .

Step 2: find the marginal density of S = X + Y

Now, X + Y = S implies that Y = S - X. Further, y > 0 implies s - x >0 which implies x > s. so we can substitute this into our expression.

$$\begin{aligned} f_S(s) &= \int_0^s e^{-x} 2e^{-2(s-x)} dx \\ &= 2e^{-2s} \int_0^s e^x dx \\ &= 2e^{-2s}(e^s - 1) \\ &= 2(e^{-s} - e^{-2s}) \end{aligned}$$

Step 3: Solve for the conditional density

$$
\begin{aligned}
f_{X|S}(X = x|S = s) &= \frac{f_{X,S}(x, s)}{f_S(s)} \\
&= \frac{2e^{x-2s}}{2(e^{-s} - e^{-2s})} I_{(0,s)}(x) \\
&= \frac{e^{x-s}}{1 - e^{-s}} I_{(0,s)}(x) \text{ after some factoring}
\end{aligned}
$$

You can integrate this as a sanity check from x = 0 to s to ensure it's a valid density.

1.11 Conditioning on Continuous RVs

1.12 Conditioning on Continuous RVs

1.13 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

2.1 Independent Random Variables

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

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

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.

2.2 Independent Random Variables, Using PMFs

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

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

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

2.3 Independent Random Variables, Using PDFs

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".

2.4 Independence

Let X and Y have joint pdf

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

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

Give students 10 min if time allows to try this first

Solution

$$f_X(x) = \int_x^1 8xy\,\mathsf{d}y = 8x \cdot \frac{y^2}{2}\Big|_{y=x}^{y=1} = 4x(1-x^2), \quad 0 < x < 1.$$

$$\Rightarrow f_X(x) = 4x(1-x^2)I_{(0,1)}(x) \qquad (X^2 \sim \text{Beta}(1,2)).$$

$$f_Y(y) = \int_0^y 8xy\,\mathsf{d}x = 8y \cdot \frac{x^2}{2}\Big|_{x=0}^{x=y} = 4y^3, \quad 0 < y < 1.$$

$$\Rightarrow f_Y(y) = 4y^3 I_{(0,1)}(y) \qquad (Y \sim \text{Beta}(4,1)).$$

No, they are not independent, because $f_{X,Y}(x,y) \neq f_X(x)f_Y(y)$.

2.5 Independence

2.6 Maximum of Independent RVs

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.

Solution

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

2.7 Maximum of Independent RVs

2.8 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

$$F_W(w) = \mathbb{P}(W \leq w) = \mathbb{P}(X_1 \leq w, X_2 \leq w, \dots, X_n \leq w) = (F_X(w))^n,$$

$$\Longrightarrow f_W(w) = \frac{\mathsf{d}}{\mathsf{d}w}F_W(w) = n(F_X(w))^{n-1}f_X(w). \qquad \text{(by the chain rule if } X \text{ is continuous)}$$

2.9 Minimum of Two Independent Random Variables

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 $z > 0$.

Give students 10 min if time allows to try this first

Solution

We have that for $u > 0$,

$$\begin{aligned}
\mathbb{P}(U < u) &= \mathbb{P}(\min\{X, Y\} < u) = 1 - \mathbb{P}(\min\{X, Y\} \geq u) \\
&= 1 - \mathbb{P}(X \geq u, Y \geq u) \\
&= 1 - \mathbb{P}(X \geq u)\mathbb{P}(Y \geq u) \\
&= 1 - e^{-\lambda u} e^{-\mu u} \\
&= 1 - e^{-(\lambda+\mu)u}.
\end{aligned}$$

Therefore, U has CDF $F_U(u) = 1 - e^{-(\lambda+\mu)u}$ for $u > 0$, so $U \sim \text{Exp}(\lambda + \mu)$.

2.10 Minimum of Two Independent Random Variables

2.11 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.

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

$$p_U(u) = \sum_w p_X(u - w) p_Y(w) = \sum_w p_Y(u - w) p_X(w).$$

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

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

2.12 Sums of Independent Random Variables

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

Solution

$$\begin{aligned}
f_S(s) &= \int_{-\infty}^{\infty} f_X(w) f_Y(s-w)\mathsf{d}w = \int_{-\infty}^{\infty} I_{(0,1)}(w) I_{(0,1)}(s-w)\mathsf{d}w \\
&= \int_{-\infty}^{\infty} I_{(0,1)}(w) I_{(s-1,s)}(w)\mathsf{d}w \\
&= \begin{cases} \int_0^s \mathsf{d}w = s & 0 < s < 1 \\ \int_{s-1}^1 \mathsf{d}w = 2-s & 1 \leq s < 2 \\ 0 & \text{otherwise.} \end{cases} \\
&= sI_{(0,1)}(s) + (2-s)I_{[1,2)}(s).
\end{aligned}$$

This is sometimes called the triangular distirbution on $(0, 2)$.

2.13 Sum of Independent Random Variables

2.14 Sum of Independent Random Variables

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

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

Give students 10 min if time allows to try this first

Solution Using the convolution method, we have that for $u < 0$, $f_U(u) = 0$, and for $u > 0$, we have that

$$\begin{aligned}
f_U(u) &= \int_{-\infty}^{\infty} f_X(w) f_Y(u-w)\mathsf{d}w \\
&= \int_0^u \lambda e^{-\lambda w} \lambda e^{-\lambda(u-w)}\mathsf{d}w \\
&= \lambda^2 e^{-\lambda u} \int_0^u \mathsf{d}w \\
&= \lambda^2 u e^{-\lambda u}.
\end{aligned}$$

Therefore, U has PDF $f_U(u) = \lambda^2 u e^{-\lambda u} I_{(0,\infty)}(u)$, so $U \sim$ Gam$(2, \lambda)$.

2.15 Sum of Independent Random Variables

2.16 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.