STAT 131 – Lecture Review Week 7

Marginal distributions, independence, and conditional distributions

Antonio Aguirre

Spring 2026

Based on Lectures 6.1 and 6.2 topics, with midterm-style examples.

How to use this review

What this hour is for

This is not a second lecture. The goal is to turn joint distributions into a workflow: **marginalize, check independence, and condition by renormalizing**.

Four questions to ask every time

1. What is the joint object: a table, a density, or a conditional model?
2. What is the support? Which pairs (x, y) are possible?
3. Are we finding a marginal, a conditional distribution, or a probability over a region?
4. Does the problem give independence, or do we need to check it?

Main habit

Before computing, draw or describe the support. Most errors come from using the wrong region.

Big picture: three ways to look at a joint distribution

Start with the joint distribution

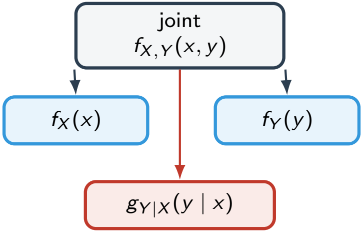
The joint distribution describes how probability is assigned to pairs:

$$(X, Y) \qquad \text{or} \qquad f_{X,Y}(x, y).$$

From it we can get marginals, conditionals, and probabilities over regions.

Core idea

Marginalize by removing one variable. Condition by fixing one variable and renormalizing.



Marginals: sum/integrate. Conditionals: restrict and divide.

$f_X(x)$: remove Y
$f_Y(y)$: remove X
$g_{Y|X}$: restrict and renormalize

Marginal distributions: what the word means

Discrete case

If X, Y have joint p.m.f. $f_{X,Y}(x,y)$, then

$$f_X(x) = \Pr(X = x) = \sum_y f_{X,Y}(x,y),$$

$$f_Y(y) = \Pr(Y = y) = \sum_x f_{X,Y}(x,y).$$

Continuous case

If X, Y have joint density $f_{X,Y}(x,y)$, then

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y,$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x.$$

Translation

A marginal distribution is what remains when we ignore the other variable.

Discrete marginals: row sums and column sums

Joint p.m.f. table

$x \backslash y$	1	2	3	4	$f_X(x)$
1	0.1	0	0.1	0	0.2
2	0.3	0	0.1	0.2	0.6
3	0	0.2	0	0	0.2
$f_Y(y)$	0.4	0.2	0.2	0.2	1

How to read it

$$f_X(2) = 0.3 + 0 + 0.1 + 0.2 = 0.6.$$

$$f_Y(1) = 0.1 + 0.3 + 0 = 0.4.$$

Check

All margins must be nonnegative and sum to 1.

Midterm-style discrete example: normalize first

Let (X, Y) have joint p.m.f.

$$f_{X,Y}(x,y) = \begin{cases} cxy, & x \in \{1,2\},\ y \in \{1,2,3\},\ x \neq y, \\ 0, & \text{otherwise.} \end{cases}$$

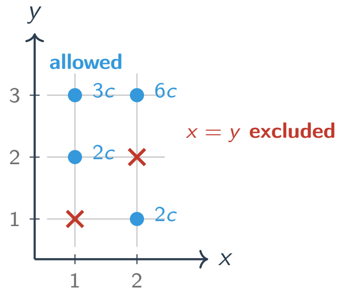
Allowed pairs

$$(1,2), \quad (1,3), \quad (2,1), \quad (2,3).$$

The diagonal points with $x = y$ are excluded.

Normalize over the support

$$1 = c(2+3+2+6) = 13c, \qquad \boxed{c = \frac{1}{13}}.$$



normalize only over blue points

Same example: marginals and a conditional

With $c = 1/13$, the nonzero probabilities are

$$f(1,2) = \frac{2}{13}, \quad f(1,3) = \frac{3}{13}, \quad f(2,1) = \frac{2}{13}, \quad f(2,3) = \frac{6}{13}.$$

Marginals

$$f_X(1) = \frac{5}{13}, \qquad f_X(2) = \frac{8}{13}.$$

$$f_Y(1) = \frac{2}{13}, \quad f_Y(2) = \frac{2}{13}, \quad f_Y(3) = \frac{9}{13}.$$

Condition by renormalizing

$$\Pr(X = 1 \mid Y = 3) = \frac{f(1,3)}{f_Y(3)} = \frac{3/13}{9/13} = \frac{1}{3}.$$

Interpretation

Given $Y = 3$, restrict to the column $y = 3$ and rescale it to total 1.

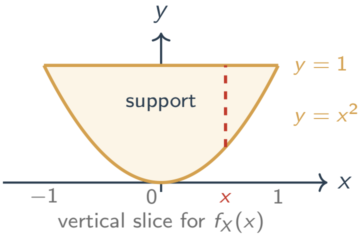
Continuous marginals: integrate out one coordinate

Example density

$$f_{X,Y}(x,y) = \begin{cases} \frac{21}{4}x^2y, & x^2 \le y \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

The support is

$$-1 < x < 1, \qquad x^2 < y < 1.$$



Reading the region

For a fixed x, y runs from x^2 to 1. For a fixed y, x runs from $-\sqrt{y}$ to $\sqrt{y}$.

Find the marginal density of X

Using

$$f_{X,Y}(x,y) = \frac{21}{4}x^2y, \qquad x^2 < y < 1,$$

we integrate over the compatible values of y.

Computation

For $-1 < x < 1$,

$$f_X(x) = \int_{x^2}^{1} \frac{21}{4}x^2y\,\mathrm{d}y.$$

Find the marginal density of X

Using

$$f_{X,Y}(x,y) = \frac{21}{4}x^2y, \qquad x^2 < y < 1,$$

we integrate over the compatible values of y.

Computation

For $-1 < x < 1$,

$$f_X(x) = \int_{x^2}^{1} \frac{21}{4}x^2y \,\mathrm{d}y.$$

$$= \frac{21}{4}x^2 \left[\frac{y^2}{2}\right]_{x^2}^{1} = \frac{21}{8}(x^2 - x^6).$$

Answer

$$f_X(x) = \begin{cases} \frac{21}{8}(x^2 - x^6), & -1 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Check the support

The marginal support of X is the projection of the joint support onto the x-axis.

Find the marginal density of Y

For a fixed $y \in (0,1)$, the condition $x^2 < y$ gives

$$-\sqrt{y} < x < \sqrt{y}.$$

Computation

$$f_Y(y) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{21}{4} x^2 y \, \mathrm{d}x.$$

Find the marginal density of Y

For a fixed $y \in (0, 1)$, the condition $x^2 < y$ gives

$$-\sqrt{y} < x < \sqrt{y}.$$

Computation

$$f_Y(y) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{21}{4} x^2 y \, \mathrm{d}x.$$

$$= \frac{21}{4} y \left[\frac{x^3}{3} \right]_{-\sqrt{y}}^{\sqrt{y}} = \frac{21}{4} y \cdot \frac{2y^{3/2}}{3}.$$

$$= \frac{7}{2} y^{5/2}.$$

Answer

$$f_Y(y) = \begin{cases} \frac{7}{2} y^{5/2}, & 0 < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Reliable habit

Use a horizontal slice for $f_Y(y)$.

Independence: what changes?

Factorization test

For both discrete and continuous random variables, independence means

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

for all relevant (x, y). Equivalently,

$$F_{X,Y}(x,y) = F_X(x)F_Y(y).$$

Interpretation

If $X \perp\!\!\!\perp Y$, knowing X does not change the distribution of Y:

$$g_{Y|X}(y \mid x) = f_Y(y).$$

Similarly,

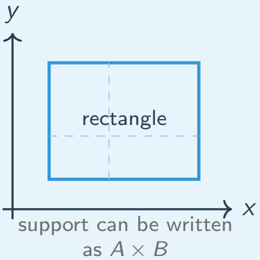
$$g_{X|Y}(x \mid y) = f_X(x).$$

Warning

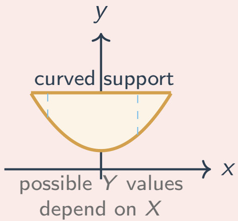
If the support is not a product set, independence is usually ruled out. Rectangular support alone does not prove independence.

A quick visual test for possible independence

Could be independent



Not independent



Why this helps

If the possible values of Y depend on the value of X, independence is ruled out for a positive density on that support.

Independence example: build the joint density

Suppose X and Y are independent and each has density

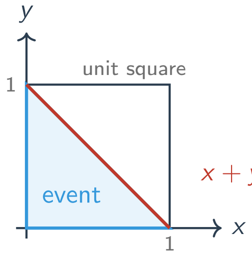
$$f(t) = \begin{cases} 2t, & 0 < t < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Find $\Pr(X + Y \leq 1)$.

Use independence

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) = 4xy,$$

for $0 < x < 1$ and $0 < y < 1$.



Independence example: integrate over the region

Region and integral

The event $X + Y \leq 1$ corresponds to

$$0 < x < 1, \qquad 0 < y < 1 - x.$$

Thus

$$\Pr(X + Y \leq 1) = \int_0^1 \int_0^{1-x} 4xy \, \mathrm{d}y \, \mathrm{d}x.$$

Independence example: integrate over the region

Region and integral

The event $X + Y \le 1$ corresponds to

$$0 < x < 1, \qquad 0 < y < 1 - x.$$

Thus

$$\Pr(X + Y \le 1) = \int_0^1 \int_0^{1-x} 4xy \, \mathrm{d}y \, \mathrm{d}x.$$

$$= \int_0^1 2x(1-x)^2 \, \mathrm{d}x = \frac{1}{6}.$$

Takeaway

Independence gives the joint density. The event still determines the integration region.

Common mistake

Do not multiply probabilities for $X + Y \le 1$. This is a region, not an intersection of separate events involving only X and only Y.

Conditional distributions: divide by the marginal

Discrete

If $f_Y(y) > 0$, then

$$g_{X|Y}(x \mid y) = \Pr(X = x \mid Y = y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$

Continuous

If $f_Y(y) > 0$, then

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

This is a density in x:

$$\int g_{X|Y}(x \mid y)\,\mathrm{d}x = 1.$$

Interpretation

Conditioning means: restrict to the slice, then renormalize that slice so total probability is 1.

Discrete conditioning: slice a column

Using the earlier table, condition on $Y = 1$.

Column $Y = 1$

x	$f(x, 1)$	$g_{X\mid Y}(x \mid 1)$
1	0.1	$0.1/0.4 = 0.25$
2	0.3	$0.3/0.4 = 0.75$
3	0	$0/0.4 = 0$
Total	0.4	1

Formula

$$g_{X\mid Y}(x \mid 1) = \frac{f(x, 1)}{f_Y(1)}.$$

Here

$$f_Y(1) = 0.1 + 0.3 + 0 = 0.4.$$

Meaning

Inside the $Y = 1$ column, $X = 2$ is three times as likely as $X = 1$.

Continuous conditioning: a vertical slice

Return to

$$f_{X,Y}(x,y) = \frac{21}{4}x^2 y, \qquad x^2 < y < 1.$$

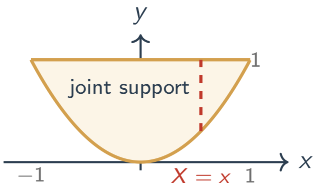
For $-1 < x < 1$,

$$f_X(x) = \frac{21}{8}(x^2 - x^6) = \frac{21}{8}x^2(1 - x^4).$$

Conditional density of Y given $X = x$

For $x \neq 0$ and $x^2 < y < 1$,

$$g_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{2y}{1 - x^4}.$$



Conditional probability from a conditional density

Find

$$\Pr\left(Y \geq \frac{3}{4} \,\middle|\, X = \frac{1}{2}\right).$$

Plug in $x = 1/2$

For $x = 1/2$,

$$g_{Y|X}\left(y \mid \frac{1}{2}\right) = \frac{2y}{1 - (1/2)^4} = \frac{32}{15}y,$$

with support

$$\frac{1}{4} < y < 1.$$

Conditional probability from a conditional density

Find

$$\Pr\left(Y \geq \frac{3}{4} \;\middle|\; X = \frac{1}{2}\right).$$

Plug in $x = 1/2$

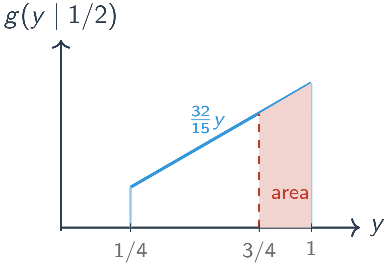
For $x = 1/2$,

$$g_{Y|X}\left(y \mid \frac{1}{2}\right) = \frac{2y}{1-(1/2)^4} = \frac{32}{15}y,$$

with support

$$\frac{1}{4} < y < 1.$$

$$\Pr\left(Y \geq \frac{3}{4} \mid X = \frac{1}{2}\right) = \int_{3/4}^{1} \frac{32}{15} y \,\mathrm{d}y = \frac{7}{15}.$$



Mixed discrete–continuous example: setup

Sometimes one variable is continuous and the other is discrete.

Defect-rate model

Suppose the unknown defect rate is random:

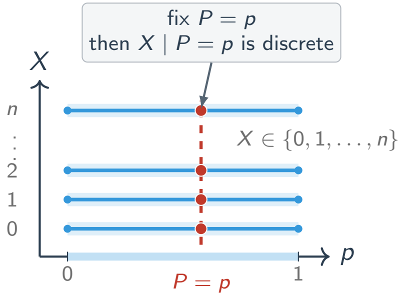
$$P \sim \mathrm{Unif}(0,1).$$

Given $P = p$, sample n items and let

$$X \mid P = p \sim \mathrm{Bin}(n, p).$$

Build the joint object

$$f_{X,P}(x,p) = g_{X|P}(x \mid p) f_P(p).$$



Key idea

Use the conditional model to construct the joint, then integrate out p.

Mixed example: integrate out the unknown rate

For $x = 0, 1, \ldots, n$ and $0 < p < 1$,

$$f_{X,P}(x, p) = \binom{n}{x} p^x (1-p)^{n-x}.$$

Marginal p.m.f. of X

$$f_X(x) = \int_0^1 \binom{n}{x} p^x (1-p)^{n-x} \, \mathrm{d}p.$$

Using

$$\int_0^1 p^x (1-p)^{n-x} \, \mathrm{d}p = \frac{x!(n-x)!}{(n+1)!},$$

we get

$$f_X(x) = \frac{1}{n+1}.$$

Interpretation

Before seeing the defect rate, every count

$$0, 1, \ldots, n$$

is equally likely under this model.

Takeaway

A conditional distribution plus a marginal distribution determines the joint.

Independence through conditionals

Equivalent tests

Assuming the relevant marginals are positive, the following are equivalent:

$$X \perp\!\!\!\perp Y, \qquad f_{X,Y}(x,y) = f_X(x)f_Y(y),$$

$$g_{X|Y}(x \mid y) = f_X(x), \qquad g_{Y|X}(y \mid x) = f_Y(y).$$

How to say it in words

If X and Y are independent, conditioning on $Y = y$ does not change the distribution of X.

Exam use

To disprove independence, one counterexample is enough: find one pair where

$$f_{X,Y}(x,y) \neq f_X(x)f_Y(y).$$

Common mistakes to avoid

Marginals

- Summing/integrating over the wrong variable.
- Forgetting the support changes after projection.
- Treating $f_X(x)$ and $f_Y(y)$ as the same function just because both are called "f."

Conditionals

- Forgetting to divide by the marginal.
- Ignoring where the marginal is positive.
- In the continuous case, treating $g_{Y|X}(y \mid x)$ as a point probability.

Reliable habit

Write the support first, then choose the sum or integral.

Final checklist

Before computing

- Identify whether the joint object is a p.m.f., p.d.f., or mixed model.
- Draw or describe the support.
- Marginal of X: remove Y by summing/integrating over y.
- Marginal of Y: remove X by summing/integrating over x.
- Conditional distribution: divide the joint by the relevant marginal.
- Independence: check whether the joint factors into the product of marginals.

Main message

Joint distributions are not just formulas. They are maps of where probability can live and how it accumulates.