STAT 131 – Lecture Review Week 6

Joint distributions, probability regions, and joint CDFs

Antonio Aguirre

Spring 2026

Based on Lectures 5.1 and 5.2, with midterm-style examples.

How to use this review

What this hour is for

This is not a second lecture. The goal is to practice the main setup moves: **read the event**, **draw the region**, and **choose the right sum or integral**.

Four questions to ask every time

1. What are the possible values of (X, Y)? **Write the support.**
2. Is the joint distribution **discrete** or **continuous**?
3. Is the event a set of **cells** or a **region in the plane**?
4. Are we given a joint p.m.f./p.d.f. or a joint c.d.f.?

Main habit

Before computing: draw or list the event. Most errors are bounds errors.

From one random variable to two

Univariate distribution

A single random variable assigns probabilities to values of X:

$$f_X(x) = \Pr(X = x) \quad \text{or} \quad f_X(x)\,\mathrm{d}x.$$

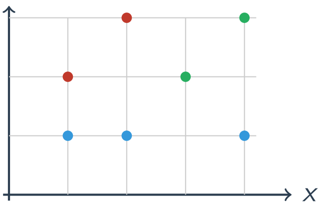
Bivariate distribution

Two random variables assign probabilities to **pairs**:

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

for discrete variables, or probability density over the plane for continuous variables.

(X, Y) lives in the plane



support = allowed pairs

Joint p.m.f. and joint p.d.f.: same idea, different language

Discrete joint p.m.f.

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

A valid joint p.m.f. satisfies

$$f_{X,Y}(x,y) \ge 0, \qquad \sum_x \sum_y f_{X,Y}(x,y) = 1.$$

For a set A of ordered pairs,

$$\Pr((X,Y) \in A) = \sum_{(x,y)\in A} f_{X,Y}(x,y).$$

Continuous joint p.d.f.

A valid joint density satisfies

$$f_{X,Y}(x,y) \ge 0, \qquad \iint_{\mathbb{R}^2} f_{X,Y}(x,y)\,\mathrm{d}A = 1.$$

For a region A in the plane,

$$\Pr((X,Y) \in A) = \iint_A f_{X,Y}(x,y)\,\mathrm{d}A.$$

Reading cue

Discrete $\Rightarrow$ sum cells. Continuous $\Rightarrow$ integrate over a region.

Discrete joint distribution: a table is a map

Example table

$x \backslash y$	1	2	3	4
1	0.1	0	0.1	0
2	0.3	0	0.1	0.2
3	0	0.2	0	0

Event probability

For an event such as

$$A = \{X \geq 2,\ Y \geq 2\},$$

add the cells in that region:

$$\Pr(A) = 0 + 0.2 + 0.1 + 0 + 0.2 + 0 = 0.5.$$

$X \backslash Y$	1	2	3	4
1	0.1	0	0.1	0
2	0.3	0	0.1	0.2
3	0	0.2	0	0

Add the shaded cells

Midterm-style: normalize a joint p.m.f.

Let (X, Y) be discrete with joint p.m.f.

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

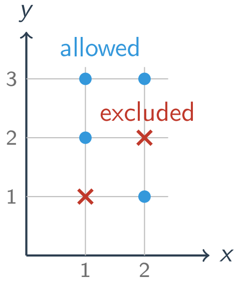
Normalize over allowed pairs

The support excludes the diagonal $x = y$:

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

$$1 = \sum_x \sum_y f_{X,Y}(x,y) = c\{1\cdot 2 + 1\cdot 3 + 2\cdot 1 + 2\cdot 3\}.$$

$$1 = 13c, \qquad \boxed{c = \frac{1}{13}}.$$



Pitfall

Normalize only over pairs in the support.

Midterm-style: compute an event from the table

For the same joint p.m.f., compute

$$\Pr(Y > X).$$

Read the event

The event $Y > X$ includes the allowed pairs

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

Therefore

$$\Pr(Y > X) = \sum_{(x,y):y>x} f_{X,Y}(x,y).$$

Midterm-style: compute an event from the table

For the same joint p.m.f., compute

$$\Pr(Y > X).$$

Read the event

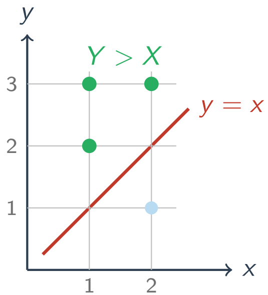
The event $Y > X$ includes the allowed pairs

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

Therefore

$$\Pr(Y > X) = \sum_{(x,y):y>x} f_{X,Y}(x,y).$$

$$= \frac{1}{13}(1 \cdot 2 + 1 \cdot 3 + 2 \cdot 3) = \boxed{\frac{11}{13}}.$$



sum only the green points

Continuous joint densities: main idea

Joint density

For a continuous random vector (X, Y) with joint density $f_{X,Y}(x, y)$

How to read this

- The **support** tells us where $f_{X,Y}(x, y)$ can be positive.
- The event A is a **region** inside that support.
- The density gives a **height** above each point (x, y).
- Probability is the **volume under the density surface** above A.

Important

For continuous variables,

$$\Pr(X = x, Y = y) = 0.$$

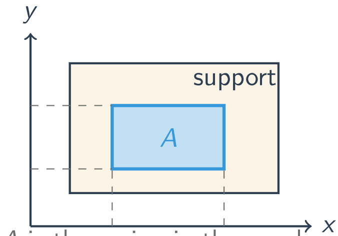
A density value is *not* a probability. Probability comes from integrating over a region.

Setup checklist

1. Draw the support, and the event region A.
2. Write the bounds for A and integrate.

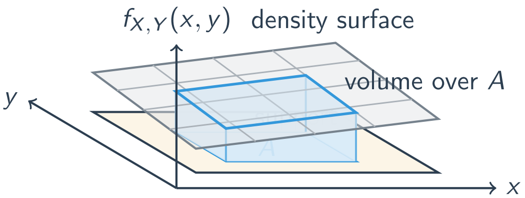
Continuous joint densities: support, event, and volume

Top view: region of integration



A is the region in the xy-plane over which we integrate.

3D view: probability as volume



The probability is the volume under the density surface and above the region A.

$$\Pr\left((X,Y) \in A\right) = \iint_A f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

Midterm-style: integrate over the region

Let (X, Y) have joint density

$$f_{X,Y}(x,y) = \begin{cases} 4y(1-x), & 0 < x < 1, \quad 0 < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Compute $\Pr(X > Y)$.

Region from the previous slide

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

Midterm-style: integrate over the region

Let (X, Y) have joint density

$$f_{X,Y}(x,y) = \begin{cases} 4y(1-x), & 0 < x < 1, \quad 0 < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Compute $\Pr(X > Y)$.

Region from the previous slide

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

Inner integral:

$$\int_0^x 4y(1-x)\,\mathrm{d}y = 4(1-x)\left[\frac{y^2}{2}\right]_0^x = 2(1-x)x^2.$$

Finish

$$\Pr(X > Y) = \int_0^1 2(1-x)x^2\,\mathrm{d}x$$

$$= \int_0^1 (2x^2 - 2x^3)\,\mathrm{d}x = \frac{2}{3} - \frac{1}{2} = \boxed{\frac{1}{6}}.$$

The bounds carry the meaning; the integral only evaluates the setup.

Same region, two equivalent orders

Vertical slices

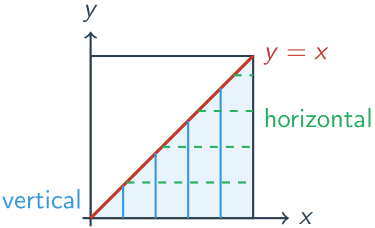
$$0 < x < 1, \qquad 0 < y < x.$$

$$\Pr(X > Y) = \int_0^1 \int_0^x f_{X,Y}(x, y)\,\mathrm{d}y\,\mathrm{d}x.$$

Horizontal slices

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

$$\Pr(X > Y) = \int_0^1 \int_y^1 f_{X,Y}(x, y)\,\mathrm{d}x\,\mathrm{d}y.$$



Lecture-style: read a curved support

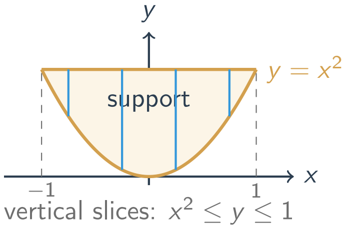
Suppose

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

Support first

The condition $x^2 \le y \le 1$ means the point (x, y) lies above the parabola and below the line:

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



Setup

To normalize, integrate over the entire shaded support:

$$1 = \int_{-1}^{1} \int_{x^2}^{1} cx^2 y \, \mathrm{d}y \, \mathrm{d}x.$$

Lecture-style: normalize on a curved support

Normalize the joint density

$$1 = \int_{-1}^{1} \int_{x^2}^{1} c x^2 y \, \mathrm{d}y \, \mathrm{d}x = c \int_{-1}^{1} x^2 \left[\frac{y^2}{2} \right]_{x^2}^{1} \mathrm{d}x.$$

$$1 = c \int_{-1}^{1} \left(\frac{x^2}{2} - \frac{x^6}{2} \right) \mathrm{d}x = c \left(\frac{1}{3} - \frac{1}{7} \right) = c \cdot \frac{4}{21}.$$

$$\boxed{c = \frac{21}{4}}.$$

What to remember

The bounds come from the support. The constant c is chosen so the total probability is 1.

Event inside a curved support: $\Pr(X \geq Y)$

Use the same density with $c = 21/4$. Compute $\Pr(X \geq Y)$.

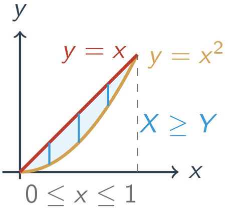
Intersect the event with the support

The support has $x^2 \leq y \leq 1$. The event $X \geq Y$ means $y \leq x$.
This forces

$$0 \leq x \leq 1, \qquad x^2 \leq y \leq x.$$

Therefore

$$\Pr(X \geq Y) = \int_0^1 \int_{x^2}^{x} \frac{21}{4} x^2 y \, \mathrm{d}y \, \mathrm{d}x.$$



Result

$$\Pr(X \geq Y) = \boxed{\frac{3}{20}}.$$

Joint CDF: accumulating down-left

Definition

For any two random variables,

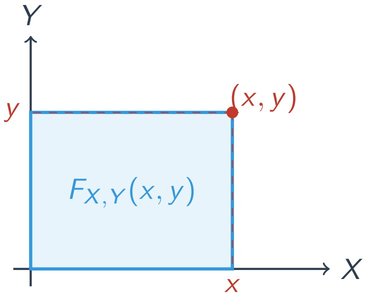
$$F_{X,Y}(x,y) = \Pr(X \le x,\ Y \le y).$$

For a continuous joint density,

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

Interpretation

The joint c.d.f. accumulates probability in the southwest rectangle ending at (x,y).



all points with $X \le x$ and $Y \le y$

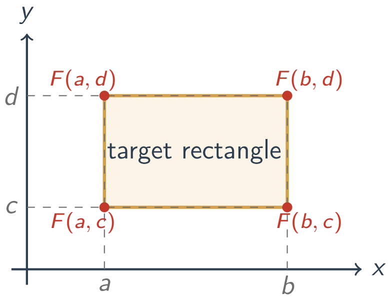
Rectangle probabilities from a joint CDF

For $a < b$ and $c < d$,

$$\Pr(a < X \leq b,\ c < Y \leq d) = F(b,d) - F(a,d) - F(b,c) + F(a,c).$$

Why the signs alternate

- $F(b,d)$ includes everything down-left of (b,d).
- subtract the part left of a.
- subtract the part below c.
- add back the overlap counted twice.



CDF example: rectangle probability

Suppose (X, Y) is supported on $0 \le x \le 3$, $0 \le y \le 4$, and

$$F(x,y) = \frac{1}{156} xy(x^2 + y).$$

Compute

$$\Pr(1 \le X \le 2, \quad 1 \le Y \le 2).$$

Use the four corners

$$\Pr(1 \le X \le 2, 1 \le Y \le 2)$$
$$= F(2,2) - F(1,2) - F(2,1) + F(1,1).$$

Compute

$$= \frac{1}{156}\{2 \cdot 2(2^2 + 2) - 1 \cdot 2(1^2 + 2)$$
$$-2 \cdot 1(2^2 + 1) + 1 \cdot 1(1^2 + 1)\}$$
$$= \frac{24 - 6 - 10 + 2}{156} = \boxed{\frac{5}{78}}.$$

From joint CDF to joint density

For a continuous pair (X, Y),

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x\,\partial y},$$

where the mixed partial derivative exists.

Example

$$F(x,y) = \frac{1}{156}xy(x^2+y) = \frac{1}{156}(x^3y+xy^2).$$

Differentiate with respect to y:

$$\frac{\partial F}{\partial y} = \frac{1}{156}(x^3+2xy).$$

Then with respect to x

$$f_{X,Y}(x,y) = \frac{\partial}{\partial x}\left[\frac{1}{156}(x^3+2xy)\right].$$

$$f_{X,Y}(x,y) = \boxed{\frac{3x^2+2y}{156}}.$$

On the support:

$$0 \le x \le 3,$$
$$0 \le y \le 4.$$

What a good setup looks like

For discrete joint problems

1. Write the support.
2. Identify the cells in the event.
3. Sum only those cells.

$$\Pr((X, Y) \in A) = \sum_{(x,y)\in A} f_{X,Y}(x, y).$$

For continuous joint problems

1. Draw the support and the event region.
2. Choose vertical or horizontal slices.
3. Write bounds before doing algebra.

Common mistakes to avoid

Discrete mistakes

- Forgetting excluded pairs in the support.
- Summing over a rectangle when the event is triangular.
- Normalizing over all pairs instead of only allowed pairs.

Continuous mistakes

- Integrating over the support, not the event.
- Using bounds that describe the wrong region.
- Treating $f_{X,Y}(x, y)$ as $\Pr(X = x, Y = y)$.

Reliable habit

Write one sentence before calculating: "The event corresponds to the region/cells ..."

Final checklist

Before computing any joint probability

- Define the pair: $(X, Y) =$ what two numbers?
- Write the support.
- If discrete: list or mark the cells in the event.
- If continuous: draw the region and choose bounds.
- If normalizing: sum/integrate over the full support.
- If using a joint c.d.f.: use the four-corner formula for rectangles.
- If finding a joint p.d.f. from a c.d.f.: take the mixed partial derivative.

Main message

Joint distributions are mostly geometry. The calculation is only as good as the region or cells you wrote down.