

STA 130A: Mathematical Statistics: Brief Course

Lecture 13: Continuous Conditioning and Independence

Dogyoon Song

Fall 2026, UC Davis

Announcements

Homework 4 is posted

- Due Thu, Oct 29, 11:59 PM PT
- Please submit on time and follow the submission instructions
- Review the problems early so you have time to ask questions

Agenda

Last time: Jointly continuous random variables

- Joint PDFs of two continuous random variables
- Marginal PDFs
- Joint CDFs
- Expectations with joint PDFs

Today: Conditioning and independence in the continuous case

- Conditioning on an event
- Conditioning on another random variable
- Conditional expectation and total expectation
- Independence of continuous random variables

Recap: Joint PDFs

Random variables X, Y are *jointly continuous* if there exists the joint PDF $f_{X,Y}$ such that

$$P((X, Y) \in A) = \iint_A f_{X,Y}(x, y) dx dy.$$

- Marginal PDFs are obtained by integrating out the other variable:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx.$$

- Expectations are weighted averages over the joint density:

$$\mathbb{E}[g(X, Y)] = \iint g(x, y) f_{X,Y}(x, y) dx dy.$$

Today: Conditioning on an event or another random variable.

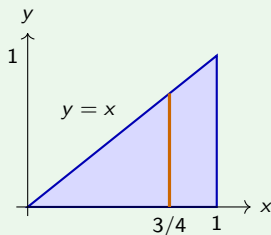
Recap example: Expected duration from a joint PDF

Example

A task starts at time Y and finishes at time X , measured in hours after 9 a.m. Assume the following joint model:

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

Question: What is the expected duration $D = X - Y$?



Support first: Fix $x \in [0, 1]$; then y ranges from 0 to x .

$$\mathbb{E}[D] = \int_0^1 \int_0^x (x - y) 2 \, dy \, dx = \int_0^1 x^2 \, dx = \frac{1}{3} \text{ hour} = 20 \text{ minutes.}$$

For an expectation, integrate *value* $\times$ *density* over the full support.

Next question: If the finish time is $X = 3/4$, which start times remain possible, and how should their density be normalized?

Conditioning a continuous random variable on an event

Definition

Let X be a continuous random variable and let A be an event with $P(A) > 0$. The **conditional PDF** of X given A is the density $f_{X|A}$ satisfying

$$P(X \in B \mid A) = \int_B f_{X|A}(x) dx.$$

If $A = \{X \in C\}$, then

$$f_{X|A}(x) = \begin{cases} \frac{f_X(x)}{P(A)}, & x \in C, \\ 0, & x \notin C. \end{cases}$$

Interpretation: For $A = \{X \in C\}$, restrict to C , then renormalize. Example: if $X \sim \text{Uniform}(0, 1)$ and $A = \{X > 1/2\}$, then $f_{X|A}(x) = 2$ on $(1/2, 1)$, and zero elsewhere.

Example: Memoryless property of exponential

Example

Let $X \sim \text{Exponential}(\lambda)$, so

$$f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0.$$

Fix $s > 0$, and condition on the event $A = \{X > s\}$.

$$f_{X|X>s}(x) = \begin{cases} \frac{f_X(x)}{P(X > s)} = \frac{\lambda e^{-\lambda x}}{e^{-\lambda s}} = \lambda e^{-\lambda(x-s)}, & x \geq s, \\ 0, & x < s. \end{cases}$$

Therefore, for $t \geq 0$,

$$P(X > s + t \mid X > s) = \int_{s+t}^{\infty} \lambda e^{-\lambda(x-s)} dx = e^{-\lambda t} = P(X > t).$$

Takeaway: Given that we have already waited s , the remaining waiting time $X - s$, conditional on $X > s$, behaves like a fresh exponential random variable.

Total probability theorem for PDFs

Total probability theorem (continuous RV)

Let $A_1, \dots, A_n$ be a partition of the sample space with $P(A_i) > 0$. Then

$$f_X(x) = \sum_{i=1}^n P(A_i) f_{X|A_i}(x).$$

Reason: Integrating either side over an interval gives the law of total probability.

This is the continuous analogue of

$$p_X(x) = \sum_i P(A_i) p_{X|A_i}(x).$$

Interpretation: The unconditional density is a weighted average of conditional densities.

Example: A mixture of densities

Example

A population consists of two groups:

$$P(G = 1) = 0.4, \quad P(G = 2) = 0.6.$$

Given the group, X has density

$$f_{X|G=1}(x) = \begin{cases} 1, & 0 \leq x < 1, \\ 0, & \text{otherwise,} \end{cases} \quad \text{and} \quad f_{X|G=2}(x) = \begin{cases} 1, & 1 \leq x \leq 2, \\ 0, & \text{otherwise.} \end{cases}$$

Then by total probability for PDFs,

$$f_X(x) = 0.4 f_{X|G=1}(x) + 0.6 f_{X|G=2}(x).$$

Thus,

$$f_X(x) = \begin{cases} 0.4, & 0 \leq x < 1, \\ 0.6, & 1 \leq x \leq 2, \\ 0, & \text{otherwise.} \end{cases}$$

Self-study: Train waiting time

Example (Train waiting time)

A train arrives every 15 minutes. You arrive uniformly at random between 7:10 AM and 7:30 AM.
Let

$$X = \text{your arrival time in minutes after 7:10 AM.} \quad \text{and} \quad X \sim \text{Uniform}(0, 20).$$

Let

$$A = \{0 \leq X \leq 5\} = \{\text{you catch the 7:15 train}\},$$

$$B = \{5 < X \leq 20\} = \{\text{you catch the 7:30 train}\}.$$

Thus,

$$P(A) = \frac{5}{20} = \frac{1}{4}, \quad P(B) = \frac{15}{20} = \frac{3}{4}.$$

Conditional on catching the 7:15 train and catching the 7:30 train respectively,

$$f_{X|A}(x) = \begin{cases} 1/5, & 0 \leq x \leq 5, \\ 0, & \text{otherwise,} \end{cases} \quad f_{X|B}(x) = \begin{cases} 1/15, & 5 < x \leq 20, \\ 0, & \text{otherwise.} \end{cases}$$

We obtain $f_X(x) = P(A)f_{X|A}(x) + P(B)f_{X|B}(x)$.

Conditioning on another continuous random variable

Definition

Let X and Y be continuous random variables with joint PDF $f_{X,Y}$. For any y with $f_Y(y) > 0$, the **conditional PDF** of X given that $Y = y$, is defined by

$$f_{X|Y}(x | y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}, \quad f_Y(y) > 0.$$

For each fixed y , $x \mapsto f_{X|Y}(x | y)$ is a valid density:

$$\int_{-\infty}^{\infty} f_{X|Y}(x | y) dx = 1.$$

Wait...! For jointly continuous X, Y , the event $\{Y = y\}$ has probability 0, so

$$P(X \in B | Y = y)$$

cannot be defined by the elementary ratio formula as before... no?

Conditioning on another continuous random variable (cont'd)

Question: For jointly continuous X, Y , the event $\{Y = y\}$ has probability 0, so

$$P(X \in B \mid Y = y)$$

cannot be defined by the elementary ratio formula as before. Why is the definition valid?

Think of conditioning on a thin strip:

$$y \leq Y \leq y + \Delta y.$$

Then

$$\begin{aligned} P(X \in B \mid y \leq Y \leq y + \Delta y) &= \frac{P(X \in B \text{ \& } y \leq Y \leq y + \Delta y)}{P(y \leq Y \leq y + \Delta y)} = \frac{\int_B \int_y^{y+\Delta y} f_{X,Y}(x, t) dt dx}{\int_y^{y+\Delta y} f_Y(t) dt} \\ &\approx \frac{\Delta y \int_B f_{X,Y}(x, y) dx}{\Delta y f_Y(y)} = \int_B \frac{f_{X,Y}(x, y)}{f_Y(y)} dx. \end{aligned}$$

Example: Conditional density from a joint density

Example

Consider the uniform distribution over a triangle:

$$f_{X,Y}(x,y) = 2, \quad 0 \leq y \leq x \leq 1.$$

From earlier lecture,

$$f_X(x) = 2x, \quad 0 \leq x \leq 1.$$

Thus, for $0 < x \leq 1$ and $0 \leq y \leq x$,

$$f_{Y|X}(y | x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{2}{2x} = \frac{1}{x}.$$

So, for $0 < x \leq 1$,

$$Y | X = x \sim \text{Uniform}(0, x).$$

Concrete check: Given $X = 1/2$, $Y \sim \text{Uniform}(0, 1/2)$, so $P(Y \leq 1/4 | X = 1/2) = 1/2$.

Modeling using conditional PDFs

Conditional densities reconstruct a joint model:

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

Example (The same triangle)

Suppose $f_X(x) = 2x$ for $0 < x < 1$ (zero elsewhere), and

$$Y | X = x \sim \text{Uniform}(0, x), \quad 0 < x < 1.$$

Then

$$f_{X,Y}(x,y) = \begin{cases} 2x \cdot (1/x) = 2, & 0 < x < 1, 0 < y < x, \\ 0, & \text{otherwise.} \end{cases}$$

Therefore,

$$f_{X,Y}(x,y) = 2x \cdot \frac{1}{x} = 2, \quad 0 \leq y \leq x \leq 1.$$

This reconstructs the uniform density on the triangular region.

Pop-up quiz

Suppose

$$f_{X,Y}(x, y) = 3x, \quad 0 \leq y \leq x \leq 1,$$

and $f_{X,Y}(x, y) = 0$ otherwise.

Question: For a fixed $y \in [0, 1)$, which expression gives $f_{X|Y}(x | y)$?

A) $\frac{1}{x}, \quad 0 \leq y \leq x \leq 1$

B) $\frac{2x}{1 - y^2}, \quad y \leq x \leq 1$

C) $3x, \quad y \leq x \leq 1$

D) $2(1 - y), \quad 0 \leq x \leq y$

Answer: B.

For fixed y , the allowed values are $y \leq x \leq 1$, and

$$f_Y(y) = \int_y^1 3x \, dx = \frac{3}{2}(1 - y^2), \quad \text{therefore,} \quad f_{X|Y}(x | y) = \frac{3x}{(3/2)(1 - y^2)} = \frac{2x}{1 - y^2}.$$

Support check: Once y is fixed, the triangle allows only $y \leq x \leq 1$.

Conditional expectation

Definition

For a conditional density $f_{X|Y}(x | y)$ at a valid conditioning value y , define the **conditional expectation**

$$\mathbb{E}[X | Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x | y) dx.$$

More generally,

$$\mathbb{E}[g(X) | Y = y] = \int_{-\infty}^{\infty} g(x) f_{X|Y}(x | y) dx.$$

Interpretation: $\mathbb{E}[X | Y = y]$ is the average value of X under the conditional density after $Y = y$ is fixed.

Total expectation theorem

Law of total expectation

For jointly continuous X, Y ,

$$\mathbb{E}[X] = \int_{\{y: f_Y(y) > 0\}} \mathbb{E}[X \mid Y = y] f_Y(y) dy.$$

This is the continuous analogue of

$$\mathbb{E}[X] = \sum_{y: p_Y(y) > 0} \mathbb{E}[X \mid Y = y] p_Y(y).$$

Interpretation: The unconditional mean is a weighted average of conditional means.

Example: Total expectation in the same triangle

Example

For $f_{X,Y}(x,y) = 2$ on $0 \leq y \leq x \leq 1$, we found

$$f_X(x) = 2x, \quad Y \mid X = x \sim \text{Uniform}(0, x), \quad 0 < x \leq 1.$$

Thus

$$\mathbb{E}[Y \mid X = x] = \int_0^x y \frac{1}{x} dy = \frac{x}{2}.$$

Average these conditional means using the marginal density of X :

$$\mathbb{E}[Y] = \int_0^1 \frac{x}{2} (2x) dx = \int_0^1 x^2 dx = \frac{1}{3}.$$

Check: This agrees with the joint-density-based calculation in Lecture 12.

Conditional expectation as a random variable

For each value x , let $m(x) = \mathbb{E}[Y \mid X = x]$. Then $\mathbb{E}[Y \mid X] = m(X)$ is a **random variable**, because it depends on X .

Example (The triangular model)

$$m(x) = \frac{x}{2}, \quad \mathbb{E}[Y \mid X] = \frac{X}{2}.$$

For example, $X = 1/2$ gives conditional mean $1/4$; $X = 1$ gives conditional mean $1/2$.

The law of total expectation can therefore be written compactly as

$$\mathbb{E}[\mathbb{E}[Y \mid X]] = \mathbb{E}[Y].$$

Distinguish: $\mathbb{E}[Y \mid X = x]$ is a number; $\mathbb{E}[Y \mid X]$ is a function of the random X .

Independence of continuous random variables

Definition

Jointly continuous random variables X and Y are **independent** if

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) \quad \text{almost everywhere.}$$

Equivalent characterizations (density equalities ignore null-set exceptions):

- CDF:

$$F_{X,Y}(x,y) = F_X(x)F_Y(y) \quad \text{for all } x,y,$$

- Conditional PDF: For almost every y with $f_Y(y) > 0$,

$$f_{X|Y}(x | y) = f_X(x) \quad \text{for almost every } x.$$

Interpretation: knowing $Y = y$ does not change the density of X .

Example: Independent normal random variables

Example

Let

$$X \sim N(\mu_X, \sigma_X^2), \quad Y \sim N(\mu_Y, \sigma_Y^2),$$

and suppose X and Y are independent.

Then the joint PDF is

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

That is,

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y} \exp \left\{ -\frac{(x - \mu_X)^2}{2\sigma_X^2} - \frac{(y - \mu_Y)^2}{2\sigma_Y^2} \right\}.$$

The density contours are ellipses aligned with the coordinate axes.

Consequences of independence

If X and Y are independent, then for suitable functions g, h ,

$$\mathbb{E}[g(X) h(Y)] = \mathbb{E}[g(X)] \mathbb{E}[h(Y)].$$

In particular,

$$\mathbb{E}[XY] = \mathbb{E}[X] \mathbb{E}[Y].$$

If X and Y have finite variances, then

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

Caution: linearity of expectation does not require independence:

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

holds whenever both expectations are finite.

Pop-up quiz

Suppose

$$f_{X,Y}(x,y) = 6xy^2, \quad 0 \leq x \leq 1, \quad 0 \leq y \leq 1.$$

Which statement is correct?

- A) X and Y are independent.
- B) X and Y are not independent because the joint density contains xy^2 .
- C) The function is not a valid joint PDF.
- D) $f_X(x) = 6x$.

Answer: A.

The marginals are $f_X(x) = 2x$ for $0 \leq x \leq 1$ and $f_Y(y) = 3y^2$ for $0 \leq y \leq 1$, and

$$f_X(x)f_Y(y) = 2x \cdot 3y^2 = 6xy^2.$$

Support matters: Our triangle example is not independent: $f_{Y|X}(y | x)$ changes with x .

Wrap-up

Conditional PDFs: For $A = \{X \in C\}$, restrict to C and renormalize

- Conditioning on $Y = y$ is defined by slicing the joint density:

$$f_{X|Y}(x | y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}, \quad 0 < f_Y(y) < \infty.$$

Conditional expectation is an average under the conditional density:

$$\mathbb{E}[X | Y = y] = \int x f_{X|Y}(x | y) dx.$$

- Total expectation averages conditional means: $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X | Y]]$.

Independence: X and Y are independent iff (almost everywhere) $f_{X,Y}(x, y) = f_X(x) f_Y(y)$.

- Under independence, expectations factor and variances add.
- Geometry still matters: the support can prevent factorization even if it looks separable.

Suggested reading: [BT08, Ch. 3.5 & 4.3]

Next lecture: Transformations and convolution

References



Dimitri Bertsekas and John N Tsitsiklis.

Introduction to probability, volume 1.

Athena Scientific, 2nd edition, 2008.