STA 131A: Introduction to Probability Theory

Lecture 16: Conditioning Continuous Random Variables

Dogyoon Song

Spring 2026, UC Davis

Announcements

Homework 5 is posted

- Please review the problems early and ask questions as needed.
- Submit on time and follow the submission instructions.

Office hours today

- Instructor office hours: 2:30–3:30 PM, MSB 4220.
- You are welcome to bring questions about Homework 5 or recent lecture material.

Agenda

Last time: Jointly continuous random variables

- Joint PDFs of two continuous random variables
- Marginal PDFs
- Joint CDFs
- Expectations with joint PDFs
- Extension to more than two random variables

Today: Conditioning and independence for continuous random variables

- 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:

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

- Expectations are weighted averages over the joint density:

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

These extend to the case where there are more than two random variables

Today: conditioning means taking a slice and renormalizing it.

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

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

Interpretation: conditioning keeps only the density compatible with A, then renormalizes.

Example: Memoryless property of exponential

Example

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

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

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

$$f_{X|X>s}(x) = \frac{f_X(x)}{P(X > s)} = \frac{\lambda e^{-\lambda x}}{e^{-\lambda s}} = \lambda e^{-\lambda(x-s)}.$$

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 behaves like a fresh exponential random variable.

Total probability theorem for PDFs

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

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

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 density

Example

A population consists of two groups:

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

Given the group, X has density

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

Then by total probability for PDFs,

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

Thus,

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

Conditioning on another continuous random variable

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

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

cannot be defined by the elementary ratio formula.

Nevertheless, the conditional density is defined by slicing the joint density at $Y = y$ and renormalizing:

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

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

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

Example: Conditional density from a joint density

Example

Let

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

From earlier,

$$f_X(x) = 2x, \qquad 0 \le x \le 1.$$

Thus, for $0 \le y \le x \le 1$,

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

So

$$Y \mid X = x \sim \mathrm{Uniform}(0, x).$$

Interpretation: once $X = x$ is fixed, the possible values of Y are uniformly spread over $[0, x]$.

Modeling using conditional PDFs

Conditional densities can be used to build a joint model:

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

Example

Suppose

$$f_X(x) = 2x, \qquad 0 \le x \le 1,$$

and

$$Y \mid X = x \sim \mathrm{Uniform}(0, x).$$

Then

$$f_{Y|X}(y \mid x) = \frac{1}{x}, \qquad 0 \le y \le x.$$

Therefore,

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

Pop-up quiz

Suppose

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

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

Question: Which expression gives $f_{X|Y}(x \mid y)$ for $0 \le y \le 1$?

A) $\frac{1}{x}, \qquad 0 \le y \le x \le 1$

B) $\frac{2x}{1-y^2}, \qquad y \le x \le 1$

C) $3x, \qquad y \le x \le 1$

D) $2(1-y), \qquad 0 \le x \le y$

Answer: B.

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

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

Follow-up: Why are the bounds $y \le x \le 1$, not $0 \le x \le 1$?

Conditional expectation

For a conditional density $f_{X|Y}(x \mid y)$, define

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

More generally,

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

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

Total expectation theorem

For jointly continuous X, Y,

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} \mathbb{E}[X \mid Y = y] \, f_Y(y) \, dy.$$

This is the continuous analogue of

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

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

Example: Total expectation for a piecewise PDF

Example

Let

$$f_X(x) = \begin{cases} \frac{1}{3}, & 0 \le x < 1, \\ \frac{2}{3}, & 1 \le x \le 2, \\ 0, & \text{otherwise.} \end{cases}$$

Let

$$A_1 = \{0 \le X < 1\}, \qquad A_2 = \{1 \le X \le 2\}.$$

Then

$$P(A_1) = \frac{1}{3}, \qquad P(A_2) = \frac{2}{3}.$$

Given A_1, $X \sim \mathrm{Uniform}(0, 1)$, so

$$\mathbb{E}[X \mid A_1] = \frac{1}{2}.$$

Example: Continuing the variance calculation

Example

For $X \mid A_1 \sim \mathrm{Uniform}(0,1)$,

$$\mathbb{E}[X^2 \mid A_1] = \frac{1}{3}.$$

For $X \mid A_2 \sim \mathrm{Uniform}(1,2)$,

$$\mathbb{E}[X^2 \mid A_2] = \frac{1^2 + 1 \cdot 2 + 2^2}{3} = \frac{7}{3}.$$

By total expectation applied to X^2,

$$\mathbb{E}[X^2] = \frac{1}{3} \cdot \frac{1}{3} + \frac{2}{3} \cdot \frac{7}{3} = \frac{5}{3}.$$

Thus,

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{5}{3} - \left(\frac{7}{6}\right)^2 = \frac{11}{36}.$$

Independence of continuous random variables

Definition

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

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) \qquad \text{for all } x, y.$$

Equivalent statements:

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

and for $f_Y(y) > 0$,

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

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

Example: Independent normal random variables

Example

Let

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

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

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

Caution: linearity of expectation does not require independence:

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

always holds when the expectations exist.

Pop-up quiz

Suppose

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

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 $0 \le x \le 1$ and $f_Y(y) = 3y^2$ for $0 \le y \le 1$, and

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

Follow-up: What changes if the support is triangular instead of rectangular?

Wrap-up

Conditional PDFs on an event keeps the compatible density and renormalizes it

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

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

Conditional expectation is an average under the conditional density:

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

- Total expectation averages conditional means: $\mathbb{E}[X] = \int \mathbb{E}[X \mid Y = y] \, f_Y(y) \, dy$.

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

- 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]

References

Dimitri Bertsekas and John N Tsitsiklis.
Introduction to probability, volume 1.
Athena Scientific, 2nd edition, 2008.