STA 131A: Introduction to Probability Theory

Lecture 9: Conditioning

Dogyoon Song

Spring 2026, UC Davis

Announcements

Homework 3 is posted (Due: Tue, Apr 21, 11:59 PM)

- Please submit on time and follow the submission instructions
- You may discuss the homework with classmates, but
  - All submitted work must be your own, and
  - You must clearly list the names of all students you collaborated with

Midterm 1 is in class on Fri, Apr 24

- You may bring *one* ***hand-written*** *letter-sized (8.5 $\times$ 11 inches), double-sided sheet of paper* with formulas and brief notes
- **Calculator**: Simple (non-graphing) calculators only
- **No textbooks** or other materials beyond the single cheat sheet
- **SDC accommodations:** Confirm scheduling with AES online

Agenda

Last time:

- Joint PMFs of two random variables (and more)

Today:

- Conditional PMFs
- Conditional expectation
- Total expectation theorem (a.k.a. law of total expectation)

Recap: Joint and Marginal PMFs

If X and Y are discrete random variables, their **joint PMF** is

$$p_{X,Y}(x, y) = P(X = x,\ Y = y), \qquad x, y \in \mathbb{R}.$$

- $p_{X,Y}(x, y) \geq 0$ for all x, y
- For any set $A \subseteq \mathbb{R}^2$, $P((X, Y) \in A) = \sum_{(x,y)\in A} p_{X,Y}(x, y)$

The marginal PMFs are obtained by summing out the other variable:

$$p_X(x) = \sum_y p_{X,Y}(x, y), \qquad p_Y(y) = \sum_x p_{X,Y}(x, y)$$

The linearity of expectation naturally extends:

$$\mathbb{E}\left[aX + bY + c\right] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c$$

These ideas further extend to more than two random variables

Conditioning a random variable on an event

Definition

The **conditional PMF** of a random variable X given an event A with $P(A) > 0$ is

$$p_{X|A}(x) = P(X = x \mid A) = \frac{P\left(\{X = x\} \cap A\right)}{P(A)}$$

- This is the same conditioning idea as before, now applied to the event $\{X = x\}$
- For each fixed event A with $P(A) > 0$, $p_{X|A}$ is a valid PMF

Example

Let X be the roll of a fair die, and let A be the event that the roll is an even number

$$p_{X|A}(x) = \frac{P(X = k \text{ and } X \text{ is even})}{P(\text{roll is even})} = \begin{cases} 1/3, & \text{if } k \in \{2,4,6\}, \\ 0, & \text{otherwise.} \end{cases}$$

Conditioning keeps only the probability mass compatible with A, then renormalizes it.

Conditioning a random variable on another

Let X, Y be random variables associated with the same experiment

Knowing $Y = y$ may provide partial knowledge about the value of X

Definition

For any y with $p_Y(y) > 0$, the **conditional PMF** of X given $Y = y$ is

$$p_{X|Y}(x \mid y) = P(X = x \mid Y = y)$$

- This is a specialization of the previous definition applied to the event $A = \{Y = y\}$

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

- For each fixed y with $p_Y(y) > 0$, the function $x \mapsto p_{X|Y}(x \mid y)$ is a valid PMF
- Operationally: take the slice $Y = y$ of the joint table and renormalize it

Illustration: Conditional PMF from a joint table

Example

Suppose the joint PMF of X and Y is

$p_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$	$p_X(x)$
$x=0$	$1/8$	$1/8$	$1/4$	$1/2$
$x=1$	$1/4$	$1/8$	$1/8$	$1/2$
$p_Y(y)$	$3/8$	$1/4$	$3/8$	

Condition on the event $\{Y=2\}$. Since $p_Y(2)=3/8$,

$$p_{X|Y}(0 \mid 2) = \frac{1/4}{3/8} = \frac{2}{3}, \qquad p_{X|Y}(1 \mid 2) = \frac{1/8}{3/8} = \frac{1}{3}.$$

Given $Y=2$, the conditional PMF of X is

$$p_{X|Y}(x \mid 2) = \begin{cases} 2/3, & x=0, \\ 1/3, & x=1, \\ 0, & \text{otherwise.} \end{cases}$$

Illustration: Building a joint PMF sequentially

Example (Short Q&A session)

Let X be the number of questions asked in a short Q&A session, and assume

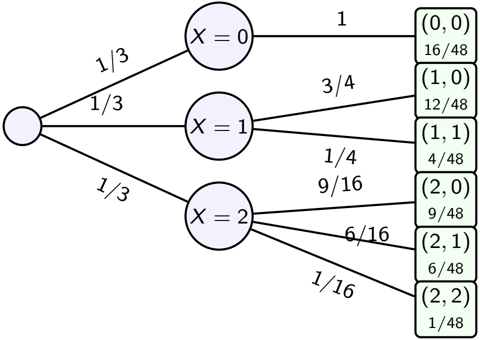
$$P(X=0) = P(X=1) = P(X=2) = \frac{1}{3},$$

Suppose that given $X = x$, each question is answered incorrectly with probability $1/4$, independently of the others. Let Y be the number answered incorrectly. Thus

$$Y \mid X = x \sim \mathrm{Binomial}(x, 1/4),$$

and therefore,

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



$p_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$
$x=0$	$16/48$	0	0
$x=1$	$12/48$	$4/48$	0
$x=2$	$9/48$	$6/48$	$1/48$

Pop-up quiz

Using the joint PMF table

$p_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$
$x=0$	$1/8$	$1/8$	$1/4$
$x=1$	$1/4$	$1/8$	$1/8$

which statement is correct?

A) $P(X=1 \mid Y=0) = 1/3$
B) $P(X=1 \mid Y=2) = 1/3$
C) $P(X=1 \mid Y=2) = 1/2$
D) $P(X=1 \mid Y=0) = P(X=1)$

Answer: B.
Because $P(Y=2) = 3/8$, so

$$P(X=1 \mid Y=2) = \frac{P(X=1, Y=2)}{P(Y=2)} = \frac{1/8}{3/8} = \frac{1}{3}.$$

Conditional expectation

Definition

The **conditional expectation** of X given an event A with $P(A) > 0$ is

$$\mathbb{E}[X \mid A] = \sum_x x \, p_{X|A}(x).$$

The conditional expectation of X given a value y of Y is

$$\mathbb{E}[X \mid Y = y] = \sum_x x \, p_{X|Y}(x \mid y).$$

Example

$p_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$
$x=0$	$1/8$	$1/8$	$1/4$
$x=1$	$1/4$	$1/8$	$1/8$

From the joint PMF table

$$\mathbb{E}[X \mid Y = 2] = 0 \cdot \frac{2}{3} + 1 \cdot \frac{1}{3} = \frac{1}{3}.$$

Total expectation theorem

Theorem

If $A_1, \ldots, A_n$ form a partition of the sample space and $P(A_i) > 0$ for all i, then

$$\mathbb{E}[X] = \sum_{i=1}^{n} P(A_i)\, \mathbb{E}[X \mid A_i]$$

Proof. For each x,

$$p_X(x) = \sum_{i=1}^{n} P(A_i)\, p_{X|A_i}(x),$$

by the PMF version of the law of total probability. Therefore,

$$\mathbb{E}[X] = \sum_x x\, p_X(x) = \sum_x x \sum_{i=1}^{n} P(A_i)\, p_{X|A_i}(x) = \sum_{i=1}^{n} P(A_i) \sum_x x\, p_{X|A_i}(x)$$

$$= \sum_{i=1}^{n} P(A_i)\, \mathbb{E}[X \mid A_i].$$

Total expectation theorem (cont'd)

Theorem

If $A_1, \ldots, A_n$ form a partition of the sample space and $P(A_i) > 0$ for all i, then

$$\mathbb{E}[X] = \sum_{i=1}^{n} P(A_i)\, \mathbb{E}[X \mid A_i]$$

As corollaries,

- For any event B with $P(A_i \cap B) > 0$ for all i,

$$\mathbb{E}[X \mid B] = \sum_{i=1}^{n} P(A_i \mid B)\, \mathbb{E}[X \mid A_i \cap B]$$

- For discrete random variables X and Y,

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

Total expectation theorem: Toy example

Example

In the Q&A example, given $X = x$, we have

$$Y \mid X = x \sim \mathrm{Binomial}(x, 1/4),$$

so

$$\mathbb{E}[Y \mid X = x] = \frac{x}{4}.$$

Therefore,

$$\begin{aligned}\mathbb{E}[Y] &= \sum_x \mathbb{E}[Y \mid X = x]\, p_X(x) = \sum_x \frac{x}{4}\, p_X(x) = \frac{1}{4}\,\mathbb{E}[X] \\ &= \frac{1}{4} \cdot \frac{0+1+2}{3} = \frac{1}{4}.\end{aligned}$$

Total expectation theorem: Portfolio example (1/4)

Example (State-dependent investment returns)

Suppose the market state after one year is $M \in \{\text{boom}, \text{normal}, \text{crash}\}$, with probabilities

$$P(M = \text{boom}) = 0.20, \qquad P(M = \text{normal}) = 0.75, \qquad P(M = \text{crash}) = 0.05.$$

Consider two one-year investment options with state-dependent returns:

Market state	boom	normal	crash
Stock fund (R_S)	30%	10%	−40%
Hedge option (R_H)	−10%	−5%	100%

The numbers in the table represent the conditional expectation of the return under each market state:

$$R_S(m) = \mathbb{E}[R_S \mid M = m], \qquad R_H(m) = \mathbb{E}[R_H \mid M = m].$$

By the total expectation theorem,

$$\mathbb{E}[R_S] = 0.20(30\%) + 0.75(10\%) + 0.05(-40\%) = 11.5\%,$$
$$\mathbb{E}[R_H] = 0.20(-10\%) + 0.75(-5\%) + 0.05(100\%) = -0.75\%.$$

Total expectation theorem: Portfolio example (2/4)

Example (Maximizing expected return)

Suppose you invest a fraction $w \in [0, 1]$ in the stock fund and the remaining $1 - w$ in the hedge.

The portfolio return is

$$R_w = wR_S + (1 - w)R_H.$$

By linearity of expectation,

$$\mathbb{E}[R_w] = w\,\mathbb{E}[R_S] + (1 - w)\,\mathbb{E}[R_H] = w(11.5\%) + (1 - w)(-0.75\%).$$

Equivalently,

$$\mathbb{E}[R_w] = 12.25\%\,w - 0.75\%.$$

Since this is increasing in w, the expected return is maximized at

$$w = 1.$$

Thus, if the goal is to maximize expected return *alone*, the best choice is to invest everything in the stock:

$$\max_{0 \le w \le 1} \mathbb{E}[R_w] = 11.5\%.$$

Total expectation theorem: Portfolio example (3/4)

Example (A simple hedging strategy)

Now suppose instead that you want the portfolio return to be *nonnegative in every market state*. Observe

$$R_w = wR_S + (1-w)R_H = \begin{cases} 40\%\, w - 10\%, & \text{boom}, \\ 15\%\, w - 5\%, & \text{normal}, \\ 100\% - 140\%\, w, & \text{crash}. \end{cases}$$

To keep the return nonnegative in every state, we need

$$40\%\, w - 10\% \geq 0, \qquad 15\%\, w - 5\% \geq 0, \qquad 100\% - 140\%\, w \geq 0.$$

These inequalities imply

$$w \geq \frac{1}{4}, \qquad w \geq \frac{1}{3}, \qquad w \leq \frac{5}{7}, \qquad \text{and thus,} \qquad \frac{1}{3} \leq w \leq \frac{5}{7}$$

Since $\mathbb{E}[R_w] = 12.25\%\, w - 0.75\%$ is increasing in w, the best hedged portfolio is at $w = \frac{5}{7}$.

Total expectation theorem: Portfolio example (4/4)

Example (A simple hedging strategy)

The portfolio with $w = \frac{5}{7}$ gives state-by-state returns

$$R_{5/7} = \begin{cases} \frac{13}{70} \approx 18.57\%, & \text{boom}, \\ \frac{2}{35} \approx 5.71\%, & \text{normal}, \\ 0, & \text{crash}, \end{cases}$$

and expected return

$$\mathbb{E}[R_{5/7}] = 12.25\% \cdot \frac{5}{7} - 0.75\% = 8\%.$$

Takeaway: The law of total expectation computes expected returns from market states. A nontrivial portfolio-allocation problem appears once we add a hedging/risk constraint.

Pop-up quiz

Which identity correctly expresses the law of total expectation for discrete random variables X and Y?

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

B) $\mathbb{E}[X] = \mathbb{E}[Y \mid X = x]$

C) $\mathbb{E}[X] = \sum_y \mathbb{E}[X \mid Y = y]$

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

Answer: D.

The unconditional mean is the weighted average of the conditional means, with weights $p_Y(y)$.

Follow-up: Why do we need the weights $p_Y(y)$?

Wrap-up

Conditional PMFs

- Conditioning on an event or on $Y = y$ produces a new PMF for X.
- For $p_Y(y) > 0$,

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

- Joint, marginal, and conditional PMFs are linked by

$$p_{X,Y}(x,y) = p_Y(y)\, p_{X|Y}(x \mid y), \qquad p_X(x) = \sum_y p_Y(y)\, p_{X|Y}(x \mid y).$$

Conditional expectation

- $\mathbb{E}[X \mid Y = y]$ is the average of X under the conditional PMF.
- The total expectation theorem states

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

Suggested reading: [BT08, Ch. 2.6]

References

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