STA 131A: Introduction to Probability Theory

Lecture 8: Expectation (cont'd) & Joint PMF of Multiple Random Variables

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
- Please review the homework problems early so you have time to ask questions if needed
- Feel free to ask questions during lecture, in office hours, or on Piazza

Midterm 1 is in class on Fri, Apr 24

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

I'll be holding my office hours today 2:30–3:30 PM at MSB 4220

- Bring your questions

Agenda

Last time:

- Functions of random variables
- Expectation and variance
- Examples

Today:

- More examples and applications of expectation
- Joint PMFs of two random variables (and more)

Recap: Expectation

Basic definitions:

$$\mathbb{E}[X] = \sum_x x\, p_X(x),$$
$$\mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right].$$

Useful identities:

$$\mathbb{E}\left[g(X)\right] = \sum_x g(x)\, p_X(x),$$
$$\mathbb{E}[aX + b] = a\,\mathbb{E}[X] + b,$$
$$\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$$

- In general, $\mathbb{E}\left[g(X)\right] \neq g\left(\mathbb{E}\left[(X)\right]\right)$, but they are equal for $g(x) = ax + b$
- $\mathrm{Var}(X) = \mathbb{E}[X^2] - \left(\mathbb{E}[X]\right)^2$

Example: Discrete uniform random variable

Example

Let X be the outcome of a fair die roll, so

$$P(X = x) = \frac{1}{6}, \qquad x = 1, 2, 3, 4, 5, 6.$$

Then

$$\mathbb{E}[X] = \frac{1+2+3+4+5+6}{6} = \frac{7}{2},$$

and

$$\mathbb{E}[X^2] = \frac{1^2+2^2+3^2+4^2+5^2+6^2}{6} = \frac{91}{6}.$$

Therefore,

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}.$$

Example: Poisson random variable

Example

If $X \sim \mathrm{Poisson}(\lambda)$, then

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \ldots$$

Then

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} k\, e^{-\lambda}\frac{\lambda^k}{k!} = \sum_{k=1}^{\infty} e^{-\lambda}\frac{\lambda^k}{(k-1)!} = \lambda \sum_{j=0}^{\infty} \underbrace{e^{-\lambda}\frac{\lambda^j}{j!}}_{=p_X(j)} = \lambda$$

Also,

$$\mathbb{E}[X(X-1)] = \sum_{k=2}^{\infty} k(k-1)e^{-\lambda}\frac{\lambda^k}{k!} = \lambda^2 \sum_{j=0}^{\infty} e^{-\lambda}\frac{\lambda^j}{j!} = \lambda^2,$$

so

$$\mathbb{E}[X^2] = \mathbb{E}[X(X-1)] + \mathbb{E}[X] = \lambda^2 + \lambda$$

Therefore,

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \lambda$$

Decision making with expected values

Expected value is often used to compare decisions with random payoffs

- To maximize long-run average payoff, we typically choose the larger *expected payoff*

Example (Choosing between two lotteries)

Suppose you must choose between

$$A: \text{ receive \$50 for sure}, \qquad B: \text{ receive \$100 with probability 0.6 and \$0 otherwise.}$$

Then

$$\mathbb{E}[\text{payoff under } A] = 50, \qquad \mathbb{E}[\text{payoff under } B] = 60.$$

If the goal is to maximize expected payoff (for example, in repeated play), you should choose B.

- Expected value is a useful criterion for repeated decisions
- It is not the only consideration though: two options may have the same mean but different risk (measured in variance)

Pop-up quiz

Two lotteries have the same expected payoff:

A : receive \$50 for sure, B : receive \$100 with probability $1/2$ and \$0 otherwise.

Question: Which statement is correct?

A) Lottery A has a larger expected payoff.

B) Lottery B has a larger expected payoff.

C) The two lotteries have the same expected payoff, so expected value alone does not distinguish them.

D) The two lotteries have the same variance.

Answer: C.

Both have expected payoff 50, but B is riskier. Here, mean alone does not tell the whole story.

Follow-up: What numerical summary might help distinguish the risk of the two options?

Why joint modeling?

Probabilistic models may involve several variables

- A medical diagnosis may involve several tests
- An investor may track the prices of multiple assets

These random variables are associated with the same experiment, sample space, and probability law, and their values may relate in interesting ways

Marginal PMFs alone do not describe how variables move together.

Example

Example: Let $X, Y \in \{0, 1\}$, with both marginally Bernoulli(1/2).

- Model 1: $Y = X$. Then $P(X = Y) = 1$.
- Model 2: $Y = 1 - X$. Then $P(X = Y) = 0$.

Both models have the same individual PMFs for X and Y, but very different joint behavior

Joint probability mass function

Definition (Joint PMF)

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
- $\sum_x \sum_y p_{X,Y}(x,y) = 1$
- 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:

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

Example: Joint PMF and marginal PMF

Example

$p_{X,Y}(x,y)$	$y=1$	$y=2$	$y=3$	$y=4$	$p_X(x)$
$x=1$	$1/20$	$1/10$	$3/20$	0	$3/10$
$x=2$	0	$1/20$	$1/10$	$1/5$	$7/20$
$x=3$	$1/10$	$1/10$	$1/10$	$1/20$	$7/20$
$p_Y(y)$	$3/20$	$1/4$	$7/20$	$1/4$	

The marginal PMFs are given by the row sums and the column sums

We can also compute probabilities of events involving both variables, for example:

$$\begin{aligned} P(X=Y) &= p_{X,Y}(1,1) + p_{X,Y}(2,2) + p_{X,Y}(3,3) \\ &= 1/5, \\ P(X<Y) &= p_{X,Y}(1,2) + p_{X,Y}(1,3) + p_{X,Y}(1,4) + p_{X,Y}(2,3) + p_{X,Y}(2,4) + p_{X,Y}(3,4) \\ &= 3/5. \end{aligned}$$

Function of multiple random variables

A function $Z = g(X, Y)$ defines another random variable

- Its PMF can be calculated from the joint PMF by

$$p_Z(z) = \sum_{\{(x,y) | g(x,y) = z\}} p_{X,Y}(x, y)$$

The expected value rule for functions naturally extends:

$$\mathbb{E}\left[g(X, Y)\right] = \sum_{x,y} g(x, y)\, p_{X,Y}(x, y)$$

In the special case where g is an affine function,

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

Example: The distribution of $Z = X + Y$

Example

Let $Z = X + Y$ where

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

Then

$$P(Z=0) = 1/8, \qquad P(Z=1) = 1/8 + 1/4 = 3/8,$$
$$P(Z=2) = 1/4 + 1/8 = 3/8, \qquad P(Z=3) = 1/8.$$

Hence

$$\mathbb{E}[Z] = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{3}{2}.$$

Also,

$$\mathbb{E}[X] + \mathbb{E}[Y] = \frac{1}{2} + 1 = \frac{3}{2} = \mathbb{E}[Z],$$

which illustrates the linearity of expectation

Pop-up quiz

Using the joint PMF table from the previous example,

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

which statement is correct?

A) $p_X(0) = 3/8$

B) $p_Y(2) = 3/8$ and $P(X < Y) = 1/2$

C) $p_{X,Y}(0,2) = p_X(0)p_Y(2)$

D) $\mathbb{E}[X+Y] = 1$

Answer: B.

The column sum gives $p_Y(2) = 3/8$, and summing the three cells with $x < y$ gives $1/2$.

More than two random variables

The joint PMF of three random variables X, Y, Z is defined analogously as

$$p_{X,Y,Z}(x, y, z) = P(X = x, Y = y, Z = z), \qquad \forall (x, y, z) \in \mathbb{R}^3$$

Marginal PMFs are obtained by summing out variables, for example,

$$p_{X,Y}(x, y) = \sum_z p_{X,Y,Z}(x, y, z), \qquad p_X(x) = \sum_y \sum_z p_{X,Y,Z}(x, y, z)$$

More generally, for any positive integer n and any random variables $X_1, \dots, X_n$,

$$p_{X_1,\dots,X_n}(x_1, \cdots, x_n) = P(X_1 = x_1, \dots, X_n = x_n), \qquad \forall (x_1, \dots, x_n) \in \mathbb{R}^n$$

- $\mathbb{E}\left[g(X_1, \dots, X_n)\right] = \sum_{x_1,\dots,x_n} g(x_1, \dots, x_n) p_{X_1,\dots,X_n}(x_1, \dots, x_n)$
- In particular,

$$\mathbb{E}\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i \, \mathbb{E}[X_i]$$

Example: Mean of the binomial

Example (Mean of a binomial random variable)

Let $X \sim \mathrm{Binomial}(n, p)$. Write

$$X = X_1 + \cdots + X_n,$$

where

$$X_i = \begin{cases} 1, & \text{if trial } i \text{ is a success,} \\ 0, & \text{otherwise.} \end{cases}$$

Then each $X_i \sim \mathrm{Bernoulli}(p)$, so $\mathbb{E}[X_i] = p$.

By linearity of expectation,

$$\mathbb{E}[X] = \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n] = np.$$

Example: The hat problem

Example (Expected number of people who get their own hats back)

Suppose n people throw their hats into a box, and then each person picks one hat at random.

Question: What is $\mathbb{E}[X]$, where X is the number of people who get back their own hat?

For each $i = 1, \ldots, n$, define

$$X_i = \begin{cases} 1, & \text{if person } i \text{ gets back their own hat,} \\ 0, & \text{otherwise.} \end{cases}$$

Observe that $P(X_i = 1) = 1/n$, so $\mathbb{E}[X_i] = \frac{1}{n}$.

Since $X = X_1 + \cdots + X_n$, by linearity of expectation,

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n \cdot \frac{1}{n} = 1.$$

Although the number of correct matches is random, its expected value is always 1, regardless of n.

Writing complicated counts as sums of indicators may make $\mathbb{E}[X]$ easy to compute

Wrap-up

Joint PMFs

- The joint PMF is
$$p_{X,Y}(x, y) = P(X = x,\ Y = y).$$
- Marginal PMFs are obtained by summing out the other variable.
- Joint PMFs let us compute probabilities of events involving both variables.

Functions of several random variables

- For any function g,
$$\mathbb{E}[g(X, Y)] = \sum_{x,y} g(x, y)\, p_{X,Y}(x, y).$$
- In particular,
$$\mathbb{E}[aX + bY + c] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c,$$

Suggested reading: [BT08, Ch. 2.5]

References

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