More Random Variables Expectation, Independence

CSE 312 Autumn 25
Lecture 10

Two descriptions

PROBABILITY MASS FUNCTION

Defined for all $\mathbb{R}$ inputs.

Usually has "0 otherwise" as an extra case.

$\sum_x p_X(x) = 1$

$0 \leq p_X(x) \leq 1$

$\sum_{z:z\leq x} p_X(z) = F_X(x)$

CUMULATIVE DISTRIBUTION FUNCTION

Defined for all $\mathbb{R}$ inputs.

Often has "0 otherwise" and 1 otherwise" extra cases

Non-decreasing function

$0 \leq F_X(x) \leq 1$

$\lim_{x\to-\infty} F_X(x) = 0$

$\lim_{x\to\infty} F_X(x) = 1$

Try It Yourself

There are 20 balls, numbered 1,2,…,20 in an urn.

You'll draw out a size-three subset. (i.e. without replacement)

$\Omega = \{$size three subsets of $\{1, \dots, 20\}$ }, $\mathbb{P}()$ is uniform measure.

Let X be the largest value among the three balls.

If outcome is $\{4,2,10\}$ then $X = 10$.

Write down the PMF of X; Write down the CDF of X.

Try It Yourself

There are 20 balls, numbered 1,2,…,20 in an urn.

You'll draw out a size-three subset. (i.e. without replacement)

Let X be the largest value among the three balls.

$$p_X(x) = \begin{cases} \binom{x-1}{2} / \binom{20}{3} & \text{if } x \in \mathbb{N},\ 3 \leq x \leq 20 \\ 0 & \text{otherwise} \end{cases}$$

Good check: if you sum up $p_X(x)$ do you get 1?

Good check: is $p_X(x) \geq 0$ for all x? Is it defined for all x?

Try it yourself

What is the CDF of X where

X be the largest value among the three balls. (Drawing 3 of the 20 without replacement)

$$F_X(x) = \begin{cases} 0 & \text{if } x < 3 \\ \binom{\lfloor x \rfloor}{3} / \binom{20}{3} & \text{if } 3 \leq x \leq 20 \\ 1 & \text{otherwise} \end{cases}$$

Try it yourself

What is the CDF of X where

X be the largest value among the three balls. (Drawing 3 of the 20 without replacement)

$$F_X(x) = \begin{cases} 0 & \text{if } x < 3 \\ \binom{\lfloor x \rfloor}{3} / \binom{20}{3} & \text{if } 3 \le x \le 20 \\ 1 & \text{otherwise} \end{cases}$$

Good checks: Is $F_X(\infty) = 1$? If not, something is wrong.

Is $F_X(x)$ increasing? If not something is wrong.

Is $F_X(x)$ defined for all real number inputs? If not something is wrong.

More Random Variable Practice

Roll a fair die n times. Let Z be the number of rolls that are $5s$ or $6s$.

What is the pmf?

Don’t try to write the CDF…it’s a mess…

Or try for a few minutes to realize it isn’t nice.

More Random Variable Practice

Roll a fair die n times. Let Z be the number of rolls that are $5s$ or $6s$.

What’s the probability of getting exactly k 5’s/6’s? Well we need to know which k of the n rolls are 5’s/6’s. And then multiply by the probability of getting exactly that outcome

$$p_Z(z) = \begin{cases} \binom{n}{z} \cdot \left(\frac{1}{3}\right)^z \left(\frac{2}{3}\right)^{n-z} & \text{if } z \in Z, 0 \leq z \leq n \\ 0 & \text{otherwise} \end{cases}$$

Expectation

Expectation

The “expectation” (or “expected value”) of a random variable X is:

$$\mathbb{E}[X] = \sum_{k} k \cdot \mathbb{P}(X = k)$$

Intuition: The weighted average of values X could take on.

Weighted by the probability you actually see them.

Example 1

Flip a fair coin twice (independently)

Let X be the number of heads.

$\Omega = \{TT, TH, HT, HH\}$, $\mathbb{P}()$ is uniform measure.

$$\mathbb{E}[X] = \frac{1}{4} \cdot 0 + \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 = 0 + \frac{1}{2} + \frac{1}{2} = 1.$$

Example 2

You roll a biased die.

It shows a 6 with probability $\frac{1}{3}$, and 1,…,5 with probability $2/15$ each.

Let X be the value of the die. What is $\mathbb{E}[X]$?

$$\frac{1}{3} \cdot 6 + \frac{2}{15} \cdot 5 + \frac{2}{15} \cdot 4 + \frac{2}{15} \cdot 3 + \frac{2}{15} \cdot 2 + \frac{2}{15} \cdot 1$$
$$= 2 + \frac{2(5+4+3+2+1)}{15} = 2 + \frac{30}{15} = 4$$

$\mathbb{E}[X]$ is not just the most likely outcome!

Try it yourself

Let X be the result shown on a fair die. What is $\mathbb{E}[X]$?

Let Y be the sum of two (independent) fair die rolls. What is $\mathbb{E}[Y]$?

Try it yourself

Let X be the result shown on a fair die. What is $\mathbb{E}[X]$

$6 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 1 \cdot \frac{1}{6}$

$= \frac{21}{6} = 3.5$

$\mathbb{E}[X]$ is not necessarily a possible outcome!

That's ok, it's an average!

Try it yourself

$\mathbb{E}[Y] =$

$\frac{1}{36} \cdot 2 + \frac{2}{36} \cdot 3 + \frac{3}{36} \cdot 4 + \frac{4}{36} \cdot 5 + \frac{5}{36} \cdot 6 + \frac{6}{36} \cdot 7 + \frac{5}{36} \cdot 8 + \frac{4}{36} \cdot 9 + \frac{3}{36} \cdot 10 + \frac{2}{36} \cdot 11 + \frac{1}{36} \cdot 12$

$= 7$

$\mathbb{E}[Y] = 2\mathbb{E}[X]$. That's not a coincidence…we'll talk about why on Friday.

Subtle but Important

X is random. You don't know what it is (at least until you run the experiment).

$\mathbb{E}[X]$ is not random. It's a number.

You don't need to run the experiment to know what it is.

More Independence

Independence of Random Variables

That's for events…what about random variables?

Independence (of random variables)

X and Y are independent if for all k, ℓ

$$\mathbb{P}(X = k, Y = \ell) = \mathbb{P}(X = k)\mathbb{P}(Y = \ell)$$

We'll often use commas instead of ∩ symbol.

Independence of Random Variables

The “for all values” is important.

We say that the event “the sum is 7” is independent of “the red die is 5”

What about $S =$“the sum of two dice” and $R =$“the value of the red die”

Independence of Random Variables

The “for all values” is important.

We say that the event “the sum is 7” is independent of “the red die is 5”

What about $S =$“the sum of two dice” and $R =$“the value of the red die”

NOT independent.

$\mathbb{P}(S = 2, R = 5) \neq \mathbb{P}(S = 2)\mathbb{P}(R = 5)$ (for example)

Independence of Random Variables

Flip a coin independently $2n$ times.

Let X be "the number of heads in the first n flips."

Let Y be "the number of heads in the last n flips."

X and Y are independent.

Mutual Independence for RVs

A little simpler to write down than for events

Mutual Independence (of random variables)

$X_1, X_2, \ldots, X_n$ **are mutually independent if for all** $x_1, x_2, \ldots, x_n$

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \mathbb{P}(X_1 = x_1)\mathbb{P}(X_2 = x_2) \cdots \mathbb{P}(X_n = x_n)$$

DON'T need to check all subsets for random variables…

But you do need to check all values (all possible x_i) still.

What does Independence give you?

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

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

What does XY mean? I rolled two dice, let X be the red die, Y the blue die. XY is the random variable that tells you the product of the two dice.

That’s a function that takes in an outcome and gives you a number back…so a random variable!! (Same for $X + Y$).

Functions of a random variable

Let X, Y be random variables defined on the same sample space.

Functions of X and/or Y like

$X + Y$

X^2

$2X + 3$

Etc.

Are random variables! (Say what the outcome is, and these functions give you a number. They're functions from $\Omega \to \mathbb{R}$. **That's the definition of a random variable!**

Expectations of functions of random variables

Let's say we have a random variable X and a function g. What is $\mathbb{E}[g(X)]$?

$\mathbb{E}[g(X)] = \sum_{\omega \in \Omega} g\big(X(\omega)\big) \cdot \mathbb{P}(\omega)$

Equivalently: $\mathbb{E}[g(X)] = \sum_{k \in \Omega_{\mathrm{g(X)}}} k \cdot \mathbb{P}(g(X) = k)$

Notice that $\mathbb{E}[g(X)]$ might not be $g(\mathbb{E}[X])$.

Extra Practice

More Practice

Suppose you flip a coin until you see a heads for the first time.

Let X be the number of trials (including the heads)

What is the pmf of X?

The cdf of X?

$\mathbb{E}[X]$?

More Practice

Suppose you flip a coin until you see a heads for the first time.

Let X be the number of trials (including the heads)

What is the pmf of X? $f_X(x) = 1/2^x$ for $x \in \mathbb{Z}^+$, 0 otherwise

The cdf of X? $F_X(x) = 1 - 1/2^{\lfloor x \rfloor}$ for $x \geq 0$, 0 for $x < 0$.

$\mathbb{E}[X]$? $\sum_{i=1}^{\infty} \frac{i}{2^i} = 2$

More Random Variable Practice

Roll a fair die n times. Let X be the number of rolls that are $5s$ or $6s$.

What is the pmf?

Don't try to write the CDF…it's a mess…

Or try for a few minutes to realize it isn't nice.

What is the expectation?

More Random Variable Practice

Roll a fair die n times. Let Z be the number of rolls that are $5s$ or $6s$.

What's the probability of getting exactly k 5's/6's? Well we need to know which k of the n rolls are 5's/6's. And then multiply by the probability of getting exactly that outcome

$$f_Z(z) = \begin{cases} \binom{n}{z} \cdot \left(\frac{1}{3}\right)^z \left(\frac{2}{3}\right)^{n-z} & \text{if } z \in Z, 0 \leq z \leq n \\ 0 & \text{otherwise} \end{cases}$$

Expectation formula is a mess. If you plug it into a calculator you'll get a nice, clean simplification: $n/3$.

More Practice: Infinite sequential processes

Infinite sequential process

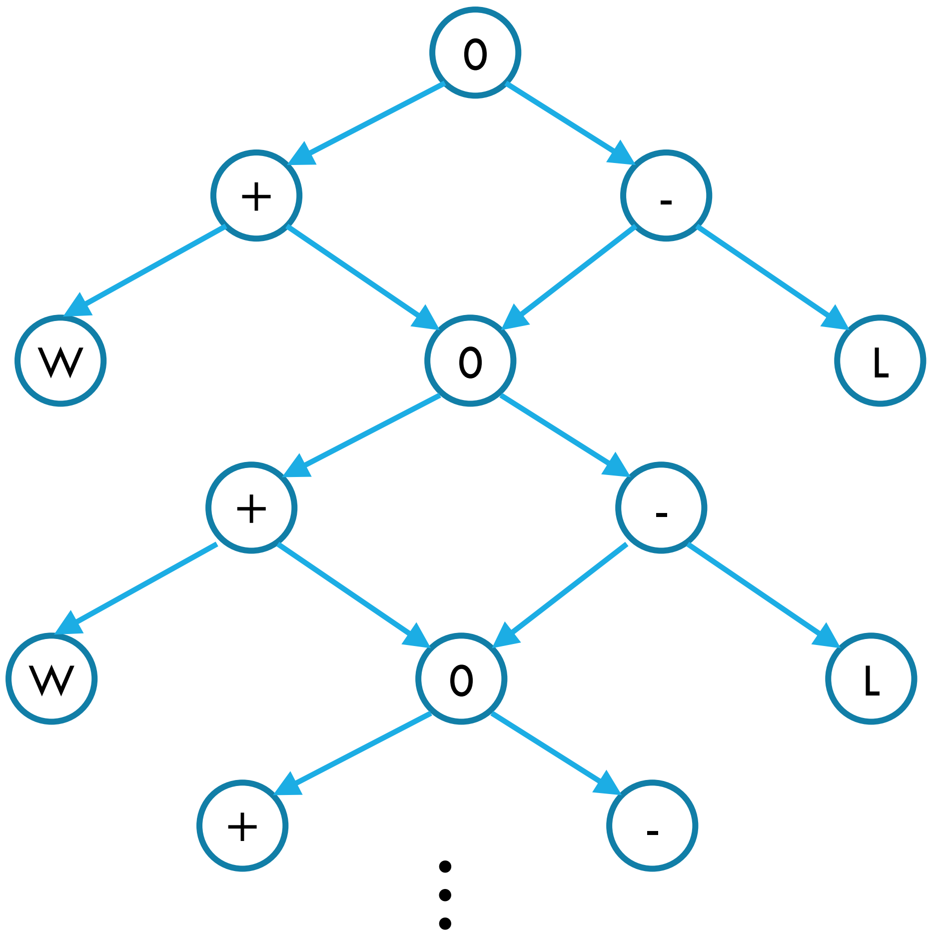
In volleyball, sets are played first team to

- Score 25 points
- Lead by at least 2

At the same time wins a set.

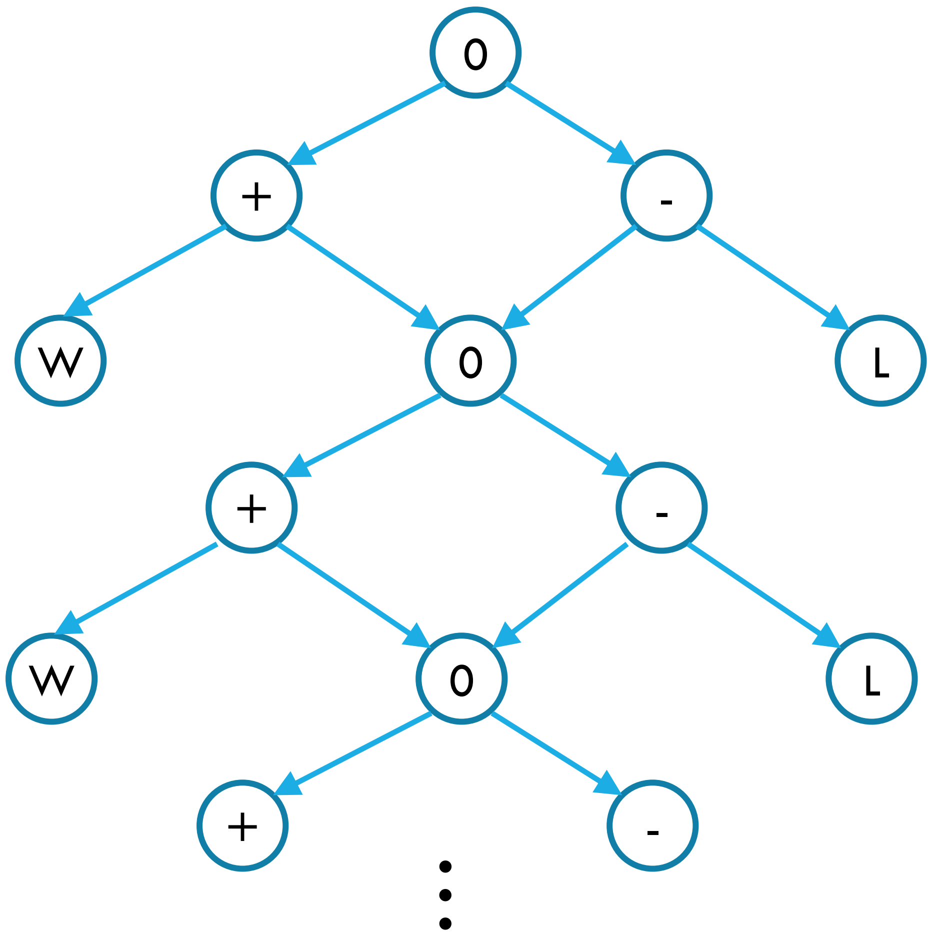
Suppose a set is 23-23. Your team wins each point independently with probability p.
What is the probability your team wins the set?

Sequential Process



$$\mathbb{P}(win\ from\ even) = p^2 + 2p(1-p)\mathbb{P}(win\ from\ even)$$

Sequential Process



$$\mathbb{P}(win\ from\ even) = p^2 + 2p(1-p)\mathbb{P}(win\ from\ even)$$

$$x - x[2p - p^2] = p^2$$
$$x[1 - 2p + p^2] = p^2$$

$$x = \frac{p^2}{p^2 - 2p + 1}$$