Random Variables

CSE 312 Autumn 25
Lecture 9

Random Variable

What's a random variable?

Formally

Random Variable

$X: \Omega \to \mathbb{R}$ **is a random variable**

$X(\omega)$ **is the summary of the outcome** ω

Informally: A random variable is a way to **summarize** the important (numerical) information from your outcome.

The sum of two dice

EVENTS

We could define

E_2 ="sum is 2"

E_3 ="sum is 3"

…

E_{12} ="sum is 12"

And ask "which event occurs"?

RANDOM VARIABLE

$X: \Omega \to \mathbb{R}$

X is the sum of the two dice.

More random variables

From one sample space, you can define many random variables.

Roll a fair red die and a fair blue die

Let D be the value of the red die minus the blue die $D(4,2) = 2$

Let S be the sum of the values of the dice $S(4,2) = 6$

Let M be the maximum of the values $M(4,2) = 4$

…

Notational Notes

We will always use capital letters for random variables.

It's common to use lower-case letters for the values they could take on.

Formally random variables are functions, so you'd think we'd write

$X(H,H,T) = 2$

But we nearly never do. We just write $X = 2$

Support (Ω_X)

The "support" (aka "the range") is the set of values X can actually take.

We called this the "image" in 311.

D (difference of red and blue) has support $\{-5, -4, -3, \dots, 4, 5\}$

S (sum) has support $\{2, 3, \dots, 12\}$

What is the support of M (max of the two dice)

Probability Mass Function

Often, we're interested in the event $\{\omega: X(\omega) = x\}$

Which is the event…that $X = x$.

We'll write $\mathbb{P}(X = x)$ to describe the probability of that event

So $\mathbb{P}(S = 2) = \frac{1}{36}$, $\mathbb{P}(S = 7) = \frac{1}{6}$

The function that tells you $\mathbb{P}(X = x)$ is the "probability mass function"

We'll often write $\mathrm{p}_X(x)$ for the pmf.

Partition

A random variable partitions Ω.

Let T be the number of twos in rolling a (fair) red and blue die.

$p_T(0) = 25/36$

$p_T(1) = 10/36$

$p_T(2) = 1/36$

	D2=1	D2=2	D2=3	D2=4	D2=5	D2=6
D1=1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1.6)
D1=2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
D1=3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
D1=4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
D1=5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
D1=6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Describing a Random Variable (differently)

The most common way to describe a random variable is the PMF.

But there's a second representation:

The cumulative distribution function (CDF) gives the probability $X \leq x$

More formally, $\mathbb{P}(\{\omega: X(\omega) \leq x\})$

Often written $F_X(x) = \mathbb{P}(X \leq x)$

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

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 $\{1, \ldots, 20\}$ }, $\mathbb{P}()$ 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.

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

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)

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

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)

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

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.

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

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

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

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

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

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.