Random Variables

CSE 312 Su23
Lecture 7

Announcements

HW2 due today (11.00pm)

HW3 out tonight, due next Wednesday at 11.00pm

> Start early!

> There is a coding project on Naïve Bayes. The relevant content will be discussed in section tomorrow.

Outline

Random Variables

- PMFs
- CDFs
- Expectation
- LOTUS

Random Variables

Motivation

numeric

Last class, we looked at the following experiment:

Flip a fair coin until we see the first Tails. What is the probability we saw at least 3 Heads?

Using our existing tools, we could approach this in two ways:

1. $\mathbb{P}(3\ Heads) + \mathbb{P}(4\ Heads) + \mathbb{P}(5\ Heads) + \mathbb{P}(6\ heads) + \cdots$

2. $1 - (\mathbb{P}(0\ Heads) + \mathbb{P}(1\ Head) + \mathbb{P}(2\ Heads))$

$\mathbb{P}(<3H)$ $\mathbb{P}(\geq 20H)$

Motivation

1. $\mathbb{P}(3\ Heads) + \mathbb{P}(4\ Heads) + \mathbb{P}(5\ Heads) + \mathbb{P}(6\ heads) + \cdots$
2. $1 - (\mathbb{P}(0\ Heads) + \mathbb{P}(1\ Head) + \mathbb{P}(2\ Heads))$

But what happens if we want to know the probability of seeing at least 10 Heads? At least 20 Heads?

When we generalize this problem, we are interested in the Probability of seeing a particular "# of Heads".

$$\mathbb{P}(\#\text{ of Heads} \geq 3)$$

Random Variable

$\Omega \to \mathbb{R}$
ω

What's a random variable?

Formally:

probability measure

sample space

Random Variable (RV)

A RV for a probability space $(\Omega, \mathbb{P})$ is a $\boxed{\text{function}}$ $X: \Omega \to \mathbb{R}$. This function takes outcomes $\omega \in \Omega$ and maps them to real values.

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

HHHT $\to$ 3

HHT $\to$ 2

HT $\to$ 1

Random Variable

Random Variable (RV)

A RV for a probability space $(\Omega, \mathbb{P})$ is a function $X: \Omega \to \mathbb{R}$. This function takes outcomes $\omega \in \Omega$ and maps them to real values.

Example.

X: number of heads in two coin flips.

Outcome, ω	$X(\omega)$
HH	2
HT	1
TH	1
TT	0

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 (x, y)

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

of H in 2coin flip

$X=2$ $X=1$ $X=0$

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 $X = 2$

$X = x$

$f(x) = y$

Support

$\mathbb{P}(X = x) > 0$

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

D (difference of red and blue) has support $\{-5, -4, -3, \ldots, 4, 5\}$ $= \Omega_D$ ✓

S (sum) has support $\{2, 3, \ldots, 12\}$ $= \Omega_S$

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

D_1 1 2 3 4 5 6

D_2 1 2 3 4 5 6

$\Omega_M = \{1, \ldots, 6\}$

Support

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

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

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

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

{1, …, 6}

Probability Mass Function

values — P(values)

(PMF)

- Often, we're interested in the event $\{\omega : \boxed{X(\omega)} = x\}$ set of outcomes
- Which is the event… that $X = x$.

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

sum of 2 die rolls 6/36

So $\mathbb{P}(S = 2) = \frac{1}{36}$ (1+1), $\mathbb{P}(S = 7) = \frac{1}{6}$ (1+6, 2+5, 3+4, 4+3, 5+2, 6+1)

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.

$\mathbb{P}(X = x)$

Partition $\Omega_T = \{0, 1, 2\}$

$\mathbb{P}(T=t) = p_T(t)$

A random variable partitions Ω.

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

	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)

✓ $p_T(0) = 25/36$

✓ $p_T(1) = 10/36$

✓ $p_T(2) = 1/36$

Partition

→ exhaustive

→ mutually exclusive $\sum_{t \in \Omega_T} \mathbb{P}(T=t)$

A random variable partitions Ω.

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

$p_T(0) = 25/36$

$p_T(1) = 10/36$

$p_T(2) = 1/36$

$36/36 = 1$ $\sum_{t \in \Omega_T} \mathbb{P}(T=t) = 1$

	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)

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 $\longrightarrow$ $\mathbb{P}(X = x)$

Try It Yourself

$\{1,2,3\}$ $\quad \mathbb{P}(X=4) = \frac{1 \binom{3}{2}}{\binom{20}{3}}$

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.

$\Omega_X : \{3, \dots, 20\}$

$$p_X(x) = \mathbb{P}(X=x) = \begin{cases} \frac{\binom{x-1}{2}}{\binom{20}{3}} & x \in \underbrace{\{3, \dots, 20\}} \\ 0 & \text{otherwise} \end{cases}$$

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.

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

$x - \mathbb{P}(X = x)$

Cumulative Distribution Function (CDF)

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

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

$$\mathbb{P}(X \leq x) = \mathbb{P}(X = x) + \mathbb{P}(X = x-1) + \mathbb{P}(X = x-2) + \dots$$

$F_X(x) = \sum_{i: i \leq x} p_X(i)$ - i.e., the CDF at x is the sum of all PMF values until x

Try it yourself

What is the CDF of X where

X is the largest value among the three balls (when drawing 3 of the 20 without replacement)

$$F_X(x) = \mathbb{P}(X \le x) = \ \le$$

$$\mathbb{P}(X \le 15) =$$

$$\mathbb{P}(\text{all values} \le 15)$$

$$= \frac{\binom{15}{3}}{\binom{20}{3}}$$

Try it yourself

What is the CDF of X where

X is the largest value among the three balls (when drawing 3 of the 20 without replacement)

$x = 15.5$

$15.0 >$

≤ 15.5

$\mathbb{P}(X \leq 15.5)$

$= \mathbb{P}(X \leq 15)$

$$\mathbb{P}(X \leq x) = \begin{cases} 0 & x < 3 \\ \dfrac{\binom{\lfloor x \rfloor}{3}}{\binom{20}{3}} & 3 \leq x \leq 20 \\ 1 & x > 20 \end{cases}$$

$$\sum_{x \in \Omega_X} p_X(x) = 1$$

Try it yourself

What is the CDF of X where

X is the largest value among the three balls (when drawing 3 of the 20 without replacement)

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

$x > 20$

$\mathbb{R}$

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.

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

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

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

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

Expectation

Expectation

Expectation

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

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

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)

What is the expected $\boxed{\text{number of heads}}$ we see?

X : # of H in 2 coin flips

$\Omega_X : \{0, 1, 2\}$

$$p_X(x) = \begin{cases} \frac{1}{4} & x=0 \\ 1/2 & x=1 \\ 1/4 & x=2 \end{cases}$$

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

Example 1

Flip a fair coin twice (independently)

What is the expected number of heads we see?

Let X be the number of heads.

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

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

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 $\mathbb{E}[X]$?

$$1 \cdot \frac{2}{15} + 2 \cdot \frac{2}{15} + 3 \cdot \frac{2}{15} + 4 \cdot \frac{2}{15} + 5 \cdot \frac{2}{15} + 6 \cdot \frac{1}{3}$$

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

$\mathbb{E}[X]$ 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 $\mathbb{E}[Y]$?

Fill out the poll everywhere so Shreya knows how long to explain
Go to pollev.com/312

Try it yourself

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

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

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

$\mathbb{E}[X]$ 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 next time.

Subtle but Important

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

values/weight.

$\mathbb{E}[X]$ is not random. It's a number.
You don't need to run the experiment to know what it is.

fixed

What about $\mathbb{E}(g(X))$?

In general, $\mathbb{E}[g(X)] \neq g(\mathbb{E}[X])$.

$\hat{\mathbb{E}}[X]$

$\underbrace{\hat{\mathbb{E}}[X^2]}_{} \neq \underbrace{\left(\hat{\mathbb{E}}[X]\right)^2}_{}$

> For e.g., is $\mathbb{E}[X^2] = (\mathbb{E}[X])^2$?

Consider this counter example,

$$X = \begin{cases} 1, with\ probability\ ½ \\ -1, with\ probability\ ½ \end{cases}$$

$X^2 = 1$ w/ prob 1

$\hat{\mathbb{E}}[X] = 1 \times \frac{1}{2} + (-1) \times \frac{1}{2}$

$= 0$

$\hat{\mathbb{E}}[X^2] = 1 \cdot 1 = 1$

$0^2 \neq 1$

What about $\mathbb{E}\left(g(X)\right)$?

In general, $\mathbb{E}[g(X)] \neq g(\mathbb{E}[X])$.

> For e.g., is $\mathbb{E}[X^2] = (\mathbb{E}[X])^2$?

Consider this counter example,

$$X = \begin{cases} 1, with\ probability\ ½ \\ -1, with\ probability\ ½ \end{cases}$$

Then:

$X^2 = 1$, with probability 1

So: $\mathbb{E}[X^2] = 1 \cdot 1 \neq 0 = \left(1 \cdot \frac{1}{2} - 1 \cdot \frac{1}{2}\right)^2 = (\mathbb{E}[X])^2$

How do we compute $\mathbb{E}[\overset{x}{\uparrow}g(X)]$?

$$\mathbb{E}[X^2] = \sum_{x \in \Omega_X} x^2 \cdot \mathbb{P}(X = \underline{\underline{x}})$$

Law of the Unconscious Statistician (LOTUS)

Given a discrete RV $X: \Omega \to \mathbb{R}$, then the expected value of g(X) is

$$\mathbb{E}[g(\mathrm{X})] = \sum_{x \in \Omega_X} g(x) \cdot \mathbb{P}(X = x) = \sum_{x \in \Omega_X} g(x) \cdot p_X(x)$$

Example

Suppose we roll a fair, 6-sided die. You win the cube of the number rolled, times 10, in dollars. How much do you expect to win?

Let X be the result of the die roll.

Example

Suppose we roll a fair, 6-sided die. You win the cube of the number rolled, times 10, in dollars. How much do you expect to win?

Let X be the result of the die roll. Then, our winnings are $10X^3$.

$\mathbb{E}[10X^3] = \sum_{x=1}^{6} 10x^3 \cdot \mathbb{P}(X = x)$

$= 10 \cdot \frac{1}{6} + 10 \cdot 2^3 \cdot \frac{1}{6} + 10 \cdot 3^3 \cdot \frac{1}{6} + 10 \cdot 4^3 \cdot \frac{1}{6} + 10 \cdot 5^3 \cdot \frac{1}{6} + 10 \cdot 6^3 \cdot \frac{1}{6}$