Intro to Random Variables

CS 70, Summer 2019

Lecture 18, 7/24/19

Questions

- If I flip 20 coins, **how many are heads**?
- If I enter a raffle with 9 other people every day, **when will I first win**?
- If I pick a random woman from the US population, **what is her height**?
- If I mix up Alice, Bob, and Charlie's HW before returning them, **how many of them will get their own HW back**?

Example: Returning HW

(From notes.)
Let X_3 = the number of **fixed points**

outcomes after returning HW

Permutation	X_3
ABC	3
ACB	1
BAC	1
BCA	0
CAB	0
CBA	1

Alice gets Bob's
Bob gets Charlie's
Charlie gets A's

$$X_3 = \begin{cases} 0 & \text{wp } \frac{2}{6} = \frac{1}{3} \\ 1 & \text{wp } \frac{3}{6} = \frac{1}{2} \\ 3 & \text{wp } \frac{1}{6} = \frac{1}{6} \end{cases}$$

Definition: Random Variable

Let $\Omega, \mathbb{P}$ correspond to a probability space.

probabilities

outcomes

A **random variable** X is a **function**!

For every outcome, X assigns it a **real number.**

Discrete random variable:
X assigns a *countable* number of values.

Ω	$\mathbb{P}$	$X(\omega)$
$\omega \rightarrow$ ABC	$\frac{1}{6}$	3
ACB	$\frac{1}{6}$	1
BAC	$\frac{1}{6}$	1
$\vdots$		

Connections to Probability Intro

Probability Space:

- **Events** are sets of outcomes.
- $\mathbb{P}[A] = \sum_{\omega \in A} \mathbb{P}[\omega]$

Random Variable X:

- **Events** are sets of outcomes given the same value by X

  $\{\omega \in \Omega : X(\omega) = a\}$
- $\mathbb{P}[X = a] = \sum_{\omega \text{ if } X(\omega) = a} \mathbb{P}[\omega]$

EX: HW:
$\{X_3 = 1\}$ is an event
$\Rightarrow$ outcomes ACB, CBA, BAC.

Definition: Distribution

The **distribution** of a random variable X consists of two things:

- The **values** X can take on.
  HW example: X_3 = # of fixed points, 3 students
  $\{0, 1, 3\}$
- The **probability** of each value.
  HW example:
  $\mathbb{P}[X_3 = 0] = \frac{2}{6} = \frac{1}{3}$
  $\mathbb{P}[X_3 = 1] = \frac{1}{2}$
  $\mathbb{P}[X_3 = 3] = \frac{1}{6}$

Sanity Check!

- What should the probabilities sum to, across **all values** X can take on? 1
- Can X take on **negative** values? Yes!
- Can X take on an **infinite number** of values?
  - *Countable* values? Raffle: can win for the first time after any # of days
  - *Uncountable* values? - Height in a population - Darts!

continuous RVs (later)

Working with RVs

Let X be a random variable with the following distribution:

$$X = \begin{cases} 1 & \text{wp } 0.4 \\ \frac{1}{2} & \text{wp } 0.25 \\ -\frac{1}{2} & \text{wp } 0.35 \end{cases}$$

What is the probability that X is **positive**?

$\mathbb{P}[X>0] = \mathbb{P}[X=1] + \mathbb{P}[X=\frac{1}{2}]$

$= 0.4 + 0.25 = 0.65$

Functions of RVs

Same definition for X

$$X = \begin{cases} 1 & \text{wp } 0.4 \\ \frac{1}{2} & \text{wp } 0.25 \\ -\frac{1}{2} & \text{wp } 0.35 \end{cases}$$

Write the distribution of $f(X)$, where $f : \mathbb{R} \to \mathbb{R}$.

$$f(X) = \begin{cases} f(1) & \text{wp } 0.4 \\ f(\frac{1}{2}) & \text{wp } 0.25 \\ f(-\frac{1}{2}) & \text{wp } 0.35 \end{cases}$$

also a RV!!

Functions of RVs

Same definition for X

$$X = \begin{cases} 1 & \text{wp } 0.4 \\ \frac{1}{2} & \text{wp } 0.25 \\ -\frac{1}{2} & \text{wp } 0.35 \end{cases}$$

Write the distribution of X^2.

$$X^2 = \begin{cases} 1 & \text{wp. } 0.4 \\ \frac{1}{4} & \text{wp. } 0.25 \\ \frac{1}{4} & \text{wp. } 0.35 \end{cases} = \begin{cases} 1 & \text{wp } 0.4 \\ \frac{1}{4} & \text{wp } 0.6 \end{cases}$$

Bernoulli Random Variable

Models whether **one biased coin flip** is a head.
$\Longrightarrow$ Models a **yes/no**-type question or event

Possible values of X: $\{0, 1\}$ "no" "yes"

$$\mathbb{P}[X = 1] = p$$

$$\mathbb{P}[X=0] = 1-p$$

Parameters: p

Notation: $X \sim \text{Bernoulli}(p)$

Bernoulli Example: Indicators

If $X \sim \text{Bernoulli}(p)$, and $X = 1$ corresponds to an event A in an experiment:
$\Longrightarrow$ We say that X is an **indicator** for A.

Each day, if it is sunny in Berkeley with probability 0.8 and cloudy with probability 0.2.
Indicator for a sunny day?

S = indicator for a sunny day

$$S = \begin{cases} 1 & \text{wp. } 0.8 \\ 0 & \text{wp. } 0.2 \end{cases}$$

Binomial Random Variable

Models how many heads are in n $\underbrace{\textbf{biased coin flips}}_{\mathbb{P}[H]=p}$

$\Longrightarrow$ Models a sum of independent, identically distributed (**i.i.d**) Bernoulli(p) RVs.

Possible values of X: $\{0, 1, \ldots, n\}$

$$\mathbb{P}[X = i] = \mathbb{P}\left[\begin{array}{l}\text{\# heads after } n \text{ flips}\\ \text{is } i\end{array}\right] = \binom{n}{i} p^i (1-p)^{n-i}$$

Parameters: n, p

Notation: $X \sim \text{Bin}(n, p)$

Binomial Example: Weather I

Each day, it is sunny in Berkeley with probability 0.8 and cloudy with probability 0.2.
Weather across days is independent.

What is the probability that over a 10 day period, there are exactly 5 sunny days?

$S = \#$ sunny over 10 days.

$S \sim \text{Bin}(n,p)$ $\quad n = 10$, $p = 0.8$

$$\mathbb{P}[\underset{i}{S=5}] = \underbrace{\binom{10}{5}}_{\binom{n}{i}} \underbrace{(0.8)^5}_{p^i} \underbrace{(0.2)^5}_{(1-p)^{n-i}}$$

Let $S_i \sim \text{Ber}(0.8)$. $S_i = 1$ if day i is sunny. $S = S_1 + S_2 + \ldots + S_n$

Binomial Example: Weather II

What is the probability that over a 10 day period, there are at least two sunny days?

$\mathbb{P}[S \geq 2] = 1 - \mathbb{P}[S \leq 1]$ ← complement

$= 1 - \underbrace{\mathbb{P}[S=0]}_{\text{use}} - \underbrace{\mathbb{P}[S=1]}_{\text{distribution}}$ $\text{Bin}(10, 0.8)$

$\mathbb{P}[S=0] = \underbrace{\binom{10}{0}}_{=1}\underbrace{(0.8)^0}_{=1}(0.2)^{10} = (0.2)^{10}$

$\mathbb{P}[S=1] = \binom{10}{1}(0.8)^1(0.2)^9 = 10 \cdot 0.8 \cdot 0.2^9$

$\mathbb{P}[S \geq 2] = 1 - (0.2)^{10} - 10(0.8)(0.2)^9$

Break

What is your real favorite movie, and what movie do you pretend is your favorite to sound cultured?

Geometric Random Variable

Models how many **biased coin flips** I need until my first head. $\mathbb{P}[H] = p$.

$\Longrightarrow$ Models time until a "success" when performing **i.i.d.** trials with **success** probability p

Possible values of X: $1, 2, 3, \ldots$ positive integers.

$$\mathbb{P}[X = i] = \mathbb{P}\left[\begin{array}{l} i-1 \text{ fail,}\\ \text{then one}\\ \text{success}\end{array}\right] = \overbrace{(1-p)(1-p)\ldots(1-p)}^{i-1}p = (1-p)^{i-1}p.$$

Parameters: **Success** probability p

Notation: $X \sim \text{Geometric}(p)$

Probabilities Sum To 1? $\quad X \sim \text{Geom}(p)$

Not obvious that the probabilities sum to 1.

$$\sum_{i=1}^{\infty} \overbrace{(1-p)^{i-1}p}^{\mathbb{P}[X=i]} = p + (1-p)p + (1-p)^2 p + \ldots$$

$\times(1-p)$ $\quad \times(1-p)$

Each term is the previous $\times$ the **same** multiplier.
$\{a, ar, ar^2, ar^3, \ldots\}$ is a **geometric sequence**.

starting value $\quad$ "ratio"

Need $\boxed{|r| < 1}$

$$\sum_{i=1}^{\infty} ar^{i-1} = \frac{a}{1-r} = \frac{p}{1-(1-p)} = \frac{p}{p} = 1$$

$r = 1-p$, $a = p$

Aside: The Formula

An easy way to recreate the formula?

Let $S = a + ar + ar^2 + \dots$ ← sum of all terms.

Key Idea: distributing r

$$r \cdot S = ar + ar^2 + ar^3 + \dots = S - a$$

Solve for S!

$rS = S - a$

$a = S - Sr$

$a = S(1-r)$ → $S = \frac{a}{1-r}$

If $|r| \geq 1$, "$S = \infty$" → doesn't work.

Geometric Example: Raffle I

I enter a raffle with 9 other people every day.
Each day, a winner is chosen independently, and with equal probability.

What is the probability that I win **for the first time** on the 5th day?

W = day of first win.

$p = \frac{1}{10}$

$\mathbb{P}[W=5] = \left(\frac{9}{10}\right)^4\left(\frac{1}{10}\right)$

Geometric Example: Raffle II

What is the probability that I win the raffle for the first time some time **on or before** the 8th day??

$\mathbb{P}[W \leq 8) = 1 - \mathbb{P}[W \geq 9]$

$= 1 - \mathbb{P}[\text{lose for first 8 days}]$

$= 1 - \left(\frac{9}{10}\right)^8$

If $X \sim \text{Geometric}(p)$, then:

$$\mathbb{P}[X \geq i] = \mathbb{P}[i-1 \text{ losses}] = (1-p)^{i-1}$$

Poisson Random Variable

Models number of **rare events** over a time period
$\Longrightarrow$ Use the "rate" of event per unit time.

Possible values of X: $0, 1, 2, \dots$ non-neg. integers.

$$\mathbb{P}[X = i] = \frac{\lambda^i}{i!} e^{-\lambda}$$

λ

Parameters: **Rate** λ

Notation: $X \sim \text{Poisson}(\lambda)$

Probabilities Sum To 1?

Again, not obvious that the probabilities sum to 1.

$$\sum_{i=0}^{\infty} \underbrace{\frac{\lambda^i}{i!} e^{-\lambda}}_{\mathbb{P}(X=i)} = e^{-\lambda} \underbrace{\sum_{i=0}^{\infty} \frac{\lambda^i}{i!}}_{e^{\lambda}}$$

role of x

$= e^{\lambda} \cdot e^{-\lambda} = 1$

Taylor Series for e^x:

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$$

Poisson Example: Typos

(Notes.) We make on average 1 typo per page.
The number of typos per page is modeled by a Poisson(λ) random variable.

What is the probability that a single page has exactly 5 typos?

T = # typos per page.

$\lambda = 1$

$\mathbb{P}[T=5] = \frac{\lambda^5}{5!} e^{-\lambda} = \frac{1}{120} e^{-1} = \frac{1}{120e}$

Poisson Example: Typos

We type 200 pages. **The pages are all independent.** What is the probablity that at least one page has **exactly** 5 typos?

$$\mathbb{P}\left[\substack{\text{at least 1 pg.} \\ T=5}\right] = 1 - \mathbb{P}\left[\substack{\text{no pages} \\ T=5}\right]$$

$$= 1 - \left[\mathbb{P}\left[\substack{\text{1 page} \\ T\neq 5}\right]\right]^{200} \quad \text{all pages.}$$

$$= 1 - \left[1 - \frac{1}{120e}\right]^{200}$$

previous page

Summary

- Random variables **assign numbers** to outcomes.
- Treat $X = i$ as any ordinary event.
- Bernoulli, Binomial, and Geometric RVs have nice interpretations via **biased coin flips**.
- Practice **modeling** real world events as Bernoulli, Binomial, Geometric, Poisson.