The Law of Total Probability, Bayes' Rule, and Random Variables (Oh My!)

Administrivia

- Homework 2 is posted and is due two Friday's from now
- If you didn't start early last time, please do so this time. Good **Milestones**:
  - Finish Problems 1-3 this week. More math, some programming.
  - Finish Problems 4-5 next week. Less math, more programming.

Administrivia

Reminder of the course **Collaboration Policy**:

- **Inspiration is Free**: you may discuss homework assignments with anyone. You are especially encouraged to discuss solutions with your instructor and your classmates.
- **Plagiarism is Forbidden**: the assignments and code that you turn in should be written entirely on your own.
- **Do NOT Search for a Solution On-Line**: You may not actively search for a solution to the problem from the internet. This includes posting to sources like StackExchange, Reddit, Chegg, etc
- **Violation** of **ANY** of the above will result in an F in the course / trip to **Honor Council**

Previously on CSCI 3022

- **Conditional Probability**: The probability that A occurs given that C occurred

$$P(A \mid C) = \frac{P(A \cap C)}{P(C)}$$

- **Multiplication Rule**: $P(A \cap C) = P(A \mid C)\ P(C)$
- **Independence**: Events A and B are independent if

  1. $P(A \mid B) = P(A)$
  2. $P(B \mid A) = P(B)$
  3. $P(A \cap B) = P(A)P(B)$

Law of Total Probability

Example: Suppose I have two bags of marbles. The first bag contains 6 white marbles and 4 black marbles. The second bag contains 3 white marbles and 7 black marbles. Now suppose I put the two bags in a box. If I close my eyes, grab a bag from the box, and then grab a marble from the bag, what is the probability that it is black?

$$\frac{11 \text{ BLACK}}{20 \text{ TOTAL}} = \frac{11}{20} = 0.55$$

$$\frac{1}{2} \cdot \frac{4}{10} + \frac{1}{2} \frac{7}{10} = \frac{4}{20} + \frac{7}{20} = \frac{11}{20}$$

Law of Total Probability

Example: Same scenario as before, but now suppose that the first bag is much larger than the second bag, so that when I reach into the box I'm twice as likely to grab the first bag as the second. What is the probability of grabbing a black marble?

$P(B_1) = 2/3$ $\quad$ $P(B_2) = 1/3$

$$\frac{2}{3} \cdot \frac{4}{10} + \frac{1}{3} \cdot \frac{7}{10} = \frac{8}{30} + \frac{7}{30} = \frac{15}{30} = \frac{1}{2}$$

Law of Total Probability

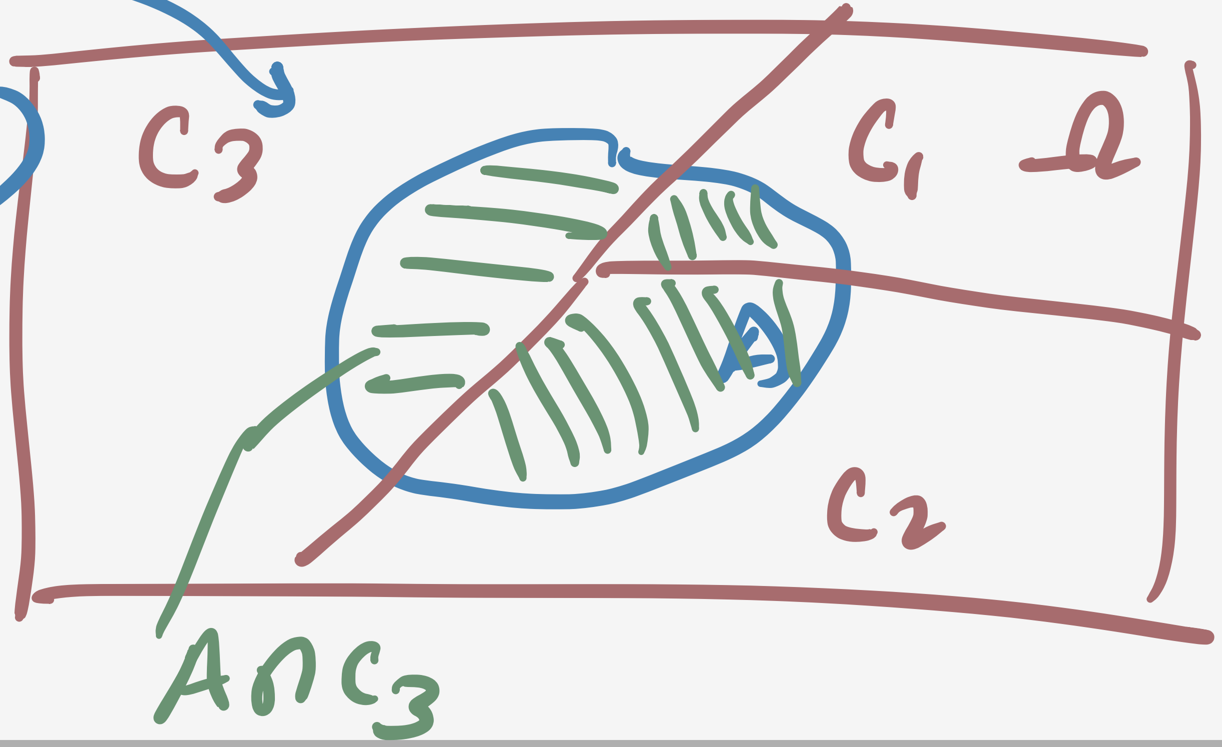
Def: Suppose $C_1, C_2, \ldots, C_m$ are disjoint events such that $C_1 \cup C_2 \cup \cdots \cup C_m = \Omega$. The probability of an arbitrary event A can be expressed as:

$$P(A) = P(A \mid C_1)P(C_1) + P(A \mid C_2)P(C_2) + \cdots + P(A \mid C_m)P(C_m)$$

$P(A|C_3)P(C_3)$

$\left(\dfrac{P(A \cap C_3)}{\cancel{P(C_3)}}\right)\cancel{P(C_3)}$

$= P(A \cap C_3)$



Let's Flip Things Around

$P(B_1) = P(B_2)$

Example: Suppose I have two bags of marbles. The first bag contains 6 white marbles and 4 black marbles. The second bag contains 3 white marbles and 7 black marbles. Now suppose I put the two bags in a box. If I close my eyes, grab a bag from the box, reach into the bag and pull out a white marble. What is the probability that I picked Bag 1?

compute $P(B_1 \mid \text{WHITE}) = \dfrac{P(B_1 \cap \text{WHITE})}{P(\text{WHITE})}$

$P(B_1) = P(B_2) = 1/2$

$P(\text{WHITE} \mid B_1) = \dfrac{6}{10}$

$P(\text{WHITE} \mid B_2) = 3/10$

$P(B_1 \cap \text{WHITE}) = P(\text{WHITE} \mid B_1)\, P(B_1)$

Let's Flip Things Around

Example: Suppose I have two bags of marbles. The first bag contains 6 white marbles and 4 black marbles. The second bag contains 3 white marbles and 7 black marbles. Now suppose I put the two bags in a box. If I close my eyes, grab a bag from the box, reach into the bag and pull out a white marble. What is the probability that I picked Bag 1?

How could we compute this?

$$P(B_1 \mid \text{WHITE}) = \frac{P(\text{WHITE} \mid B_1)\,P(B_1)}{P(\text{WHITE})}$$

$$= \frac{P(\text{WHITE} \mid B_1)\,P(B_1)}{P(\text{WHITE} \mid B_1)\,P(B_1) + P(\text{WHITE} \mid B_2)\,P(B_2)} = \frac{6/10 \cdot 1/2}{\frac{6}{10}\cdot\frac{1}{2} + \frac{3}{10}\cdot\frac{1}{2}}$$

$$= \frac{6/20}{9/20} = 6/9 = \boxed{2/3}$$

Let's Flip Things Around

Example: Suppose I have two bags of marbles. The first bag contains 6 white marbles and 4 black marbles. The second bag contains 3 white marbles and 7 black marbles. Now suppose I put the two bags in a box. If I close my eyes, grab a bag from the box, reach into the bag and pull out a white marble. What is the probability that I picked Bag 1?

How could we compute this? Let's write down literally everything we know…

Let's Flip Things Around

Example: Suppose I have two bags of marbles. The first bag contains 6 white marbles and 4 black marbles. The second bag contains 3 white marbles and 7 black marbles. Now suppose I put the two bags in a box. If I close my eyes, grab a bag from the box, reach into the bag and pull out a white marble. What is the probability that I picked Bag 1?

How could we compute this? Let's write down literally everything we know…

Let's Flip Things Around

Example: Suppose I have two bags of marbles. The first bag contains 6 white marbles and 4 black marbles. The second bag contains 3 white marbles and 7 black marbles. Now suppose I put the two bags in a box. If I close my eyes, grab a bag from the box, reach into the bag and pull out a white marble. What is the probability that I picked Bag 1?

How could we compute this? Let's write down literally everything we know...

Bayes' Rule

The notion of using evidence (the marble is White) to update our belief about an event (that we selected Box 1 from the box) is the cornerstone of a statistical framework called **Bayesian Reasoning**.

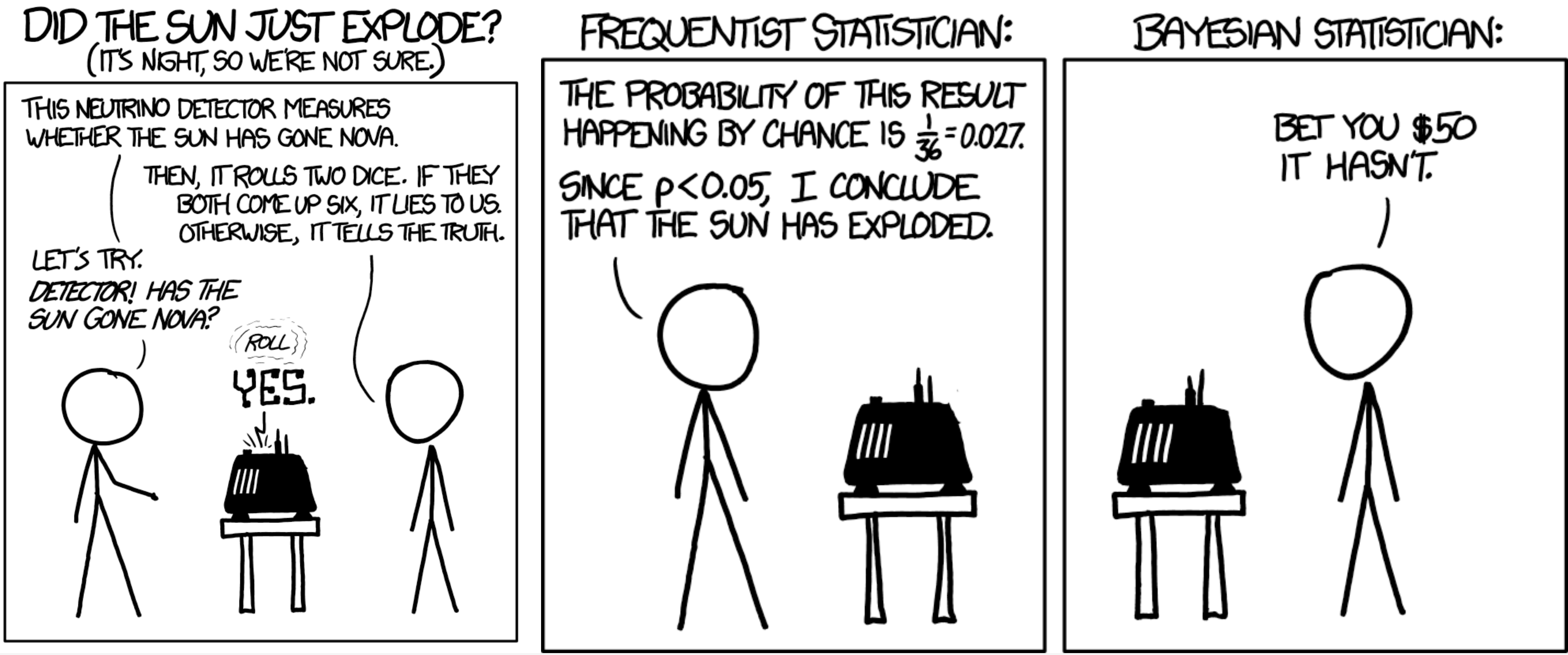
LIKELIHOOD

The formulas we derived in the previous example are called **Bayes' Rule** or **Bayes' Theorem**

PRIOR

$$P(A \mid C) = \frac{P(C \mid A)P(A)}{P(C)} = \frac{P(C \mid A)P(A)}{P(C \mid A)P(A) + P(C \mid A^c)P(A^c)}$$

PROB OF EVIDENCE


DID THE SUN JUST EXPLODE?
(IT'S NIGHT, SO WE'RE NOT SURE.)
THIS NEUTRINO DETECTOR MEASURES WHETHER THE SUN HAS GONE NOVA.
THEN, IT ROLLS TWO DICE. IF THEY BOTH COME UP SIX, IT LIES TO US. OTHERWISE, IT TELLS THE TRUTH.
LET'S TRY.
DETECTOR! HAS THE SUN GONE NOVA?
ROLL
YES.
FREQUENTIST STATISTICIAN:
THE PROBABILITY OF THIS RESULT HAPPENING BY CHANCE IS $\frac{1}{36}$=0.027.
SINCE $p<0.05$, I CONCLUDE THAT THE SUN HAS EXPLODED.
BAYESIAN STATISTICIAN:
BET YOU $50 IT HASN'T.

Bayes' Rule

The notion of using evidence (the marble is White) to update our belief about an event (that we selected Box 1 from the box) is the cornerstone of a statistical framework called **Bayesian Reasoning**.

The formulas we derived in the previous example are called **Bayes' Rule** or **Bayes' Theorem**

$$P(A \mid C) = \frac{P(C \mid A)P(A)}{P(C)} = \frac{P(C \mid A)P(A)}{P(C \mid A)P(A) + P(C \mid A^c)P(A^c)}$$

Bayes' Rule

Bayes' Rule has applications all over science.

- Should we test men for prostate cancer?
- Bayes' Rule allows us to write down the probability that someone who tests positive for prostate cancer **actually has** prostate cancer.
- False positives may cause huge amounts of stress, heartache, and pain.
- On the other hand, if you don't test for cancer, you may not discover it until it's too late
- Things are slightly more complicated than this: Other factors are age, PSA cutoffs, etc.

Bayes' Rule

Classic Example: Suppose that 1% of men over the age of 40 have prostate cancer. Also suppose that a test for prostate cancer exists with the following properties: 90% of people **have cancer** will test positive and 8% of people who do not have cancer will also test positive. What is the probability that a person who tests positive for cancer **actually has cancer**?

C = HAVE CANCER   + = POS TEST   − = NEG TEST

$$P(C \mid +) = \frac{P(+ \mid C)\,P(C)}{P(+)}$$

$$= \frac{\overset{0.90}{P(+ \mid C)}\,P(C)\ 0.01}{\underset{0.90}{P(+ \mid C)}\,\underset{0.01}{P(C)} + \underset{.08}{P(+ \mid C^c)}\,\underset{0.99}{P(C^c)}} = \boxed{\begin{array}{l}0.102 \\ = P(C \mid +)\end{array}}$$

Bayes' Rule

Classic Example: Suppose that 1% of men over the age of 40 have prostate cancer. Also suppose that a test for prostate cancer exists with the following properties: 90% of people **have cancer** will test positive and 8% of people who do not have cancer will also test positive. What is the probability that a person who tests positive for cancer **actually has cancer**?

Bayes' Rule

Classic Example: Suppose that 1% of men over the age of 40 have prostate cancer. Also suppose that a test for prostate cancer exists with the following properties: 90% of people **have cancer** will test positive and 8% of people who do not have cancer will also test positive. What is the probability that a person who tests positive for cancer **actually has cancer**?

$$P(+) = P(+|c)\,P(c) + P(+|c^c)\,P(c^c)$$

$$\frac{90}{100}\cdot\frac{1}{100} + \frac{8}{100}\cdot\frac{99}{100}$$

$$= 8.8\%$$

Random Variables

Suppose that I roll two dice

- What is the most combination?
- What is the most likely sum?

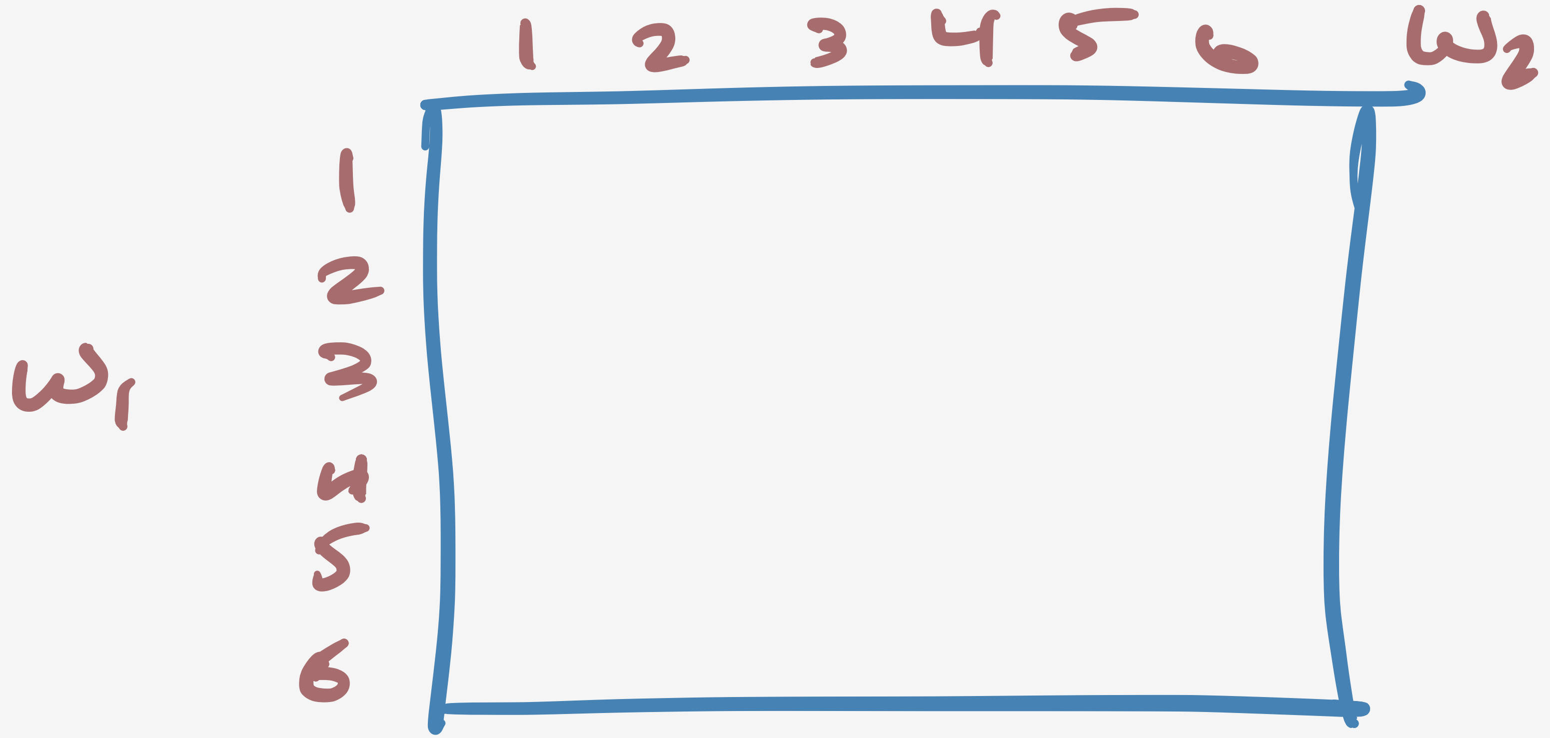
Random Variables

Suppose that I roll two dice

- What is the most combination?
- What is the most likely sum?

Random Variables

What is the sample space? $\omega_1 = 1^{st}$ Roll $\omega_2 = 2^{nd}$ Roll



Random Variables

What is the sample space?

The Key: the dice are random, so the sum is random. Let's sidestep the sample space entirely and just go straight to the thing we care about: the sum.

We call the sum of the dice a **random variable**.

Random Variables

What is the sample space?

The Key: the dice are random, so the sum is random. Let's sidestep the sample space entirely and just go straight to the thing we care about: the sum.

We call the sum of the dice a **random variable**.

Def: a discrete random variable is a function that maps the elements of the sample space Ω To a finite number of values $a_1, a_2, \ldots, a_n$ or an infinite number of values $a_1, a_2, \ldots$

Examples:

- Sum of the dice, difference of the dice, maximum of the dice
- Number of coin flips until we get a heads

Probability Mass Function

Def: a $\boxed{\text{probability mass function}}$ is the map between the random variable's values and the probabilities of those values

$$f(a) = P(X = a)$$

R.V.

- Called a "probability mass function" (PMF) because each of the random variable' values has some probability mass (or weight) associated with it

- Because the PMF is a probability function, the sum of all the masses must be what?

$$\sum_{i=1}^{n} f(a_i) = 1$$

Probability Mass Function

$P(H) = p$

Question: what is the probability mass function for the number of coin flips until a biased coin comes up heads?

$$\Omega = \{ H, TH, TTH, TTTH, \ldots \}$$

$$X = 1 \quad 2 \quad 3 \quad 4 \quad \ldots$$

$$f = p \quad (1-p)p \quad (1-p)^2 p \quad (1-p)^3 p \quad \ldots$$

Cumulative Distribution Function

Def: a **cumulative distribution function** (CDF) is **a** function whose value at a point a is the cumulative sum of probability masses up until **a**.

$$F(a) = P(X \leq a)$$

F(6) = prob of rolling a sum ≤ 6

Question: What is the relationship between the PMF and the CDF?

$$F(a) = \sum_{x \leq a} f(a)$$

Cumulative Distribution Function

Question: what is the probability that I roll two dice and they add up to at least 9?

$$P(X = \{9, 10, 11, 12\}) = P(X \geq 9)$$

$$= 1 - P(X < 9) = 1 - P(X \leq 8)$$

$$= 1 - F(8) = \frac{10}{36}$$

OK! Let's Go to Work!

Get in groups, get out laptop, and open the Lecture 6 In-Class Notebook

Let's:

- Get some more practice with the Law of Total Probability and Bayes' Rule
- Look at a famous Bayesian example called **the Monte Hall Problem**