Lecture 26

Produced by Dr. Worldwide

Welcome to the 305

Probabilistic Models

- An experiment is an event whose outcome is not known with certainty
- The set of possible outcomes of an experiment is called the sample space which we will denote S
- The outcomes themselves are called sample points
- Examples of experiments
  - Flipping a coin → $S = \{H, T\}$
  - Tossing a die → $S = \{1,2,3,4,5,6\}$
  - Flipping a coin 10 times → $S = \{strings\ of\ length\ 10\ with\ letters\ H\ \&\ T\}$
  - Time waiting on phone for airline to answer → $S = [0, \infty)$
  - Score in the next UNC basketball game → $S = \{(x, y): x, y \geq 0\}$
- Probability is a measure of how likely an event is to occur

Population vs. Samples

- The total population of an experiment is a set containing all observations
- A sample consists of a subset (usually randomly selected) of total population
- Total population of a random experiment that can be repeated an infinite number of times cannot be observed
- Q: What is an example of an experiment that can be infinitely repeated?
- If the total population is known, we can introduce randomness by considering the experiment of selecting one element (observation) of the population at a time
- The probability of selected an observation exhibiting "property x" is

$$P(property\ x) = \frac{\#\ of\ elements\ exhibiting\ property\ x}{total\ population\ size}$$

Interpretations of Probability

- Frequentist approach (classic)
  - Suppose we can repeat an experiment, under the exact conditions as many times as we want
  - We want to assign a value to how likely a specific outcome is
  - Compute the relative frequency of the desired outcome

$$\frac{\#\ of\ times\ outcome\ occurs}{\#\ of\ experiments}$$

  - We can think of the probability as the limit of its relative frequency as the number of repetitions grows to infinity

$$P(Outcome) = \lim_{n\to\infty} \frac{\#\ of\ times\ outcome\ occurs}{n}$$

where n is the number of times we repeat the experiment

Interpretations of Probability

- Bayesian approach
  - Define probability as the degree of belief rather than the long-run frequency
  - Degree of belief is based off prior probability (subjective probability) and the relative frequency from observed data
  - Posterior probability is the updated belief on the probability of an event happening given the prior and data observed

- Difference between frequentist and Bayesian approach
  - Consider the experiment where we flip a coin
  - We want to find the probability of heads
  - Frequentist concludes probability is 0.5 under the belief that the relative frequency would get closer to 50% the more the coin is flipped
  - There is an assumption that out of the two outcomes both are equally likely
  - Bayesian would take the 50% as a prior belief with a lot of uncertainty until data has been gathered to back up the claim

Probability Laws

- A probability law P assigns to each event $A \subseteq S$ a value in $[0,1]$
- Let $\Omega = S$ be the universe and $\emptyset$ denote the empty set
- Notation: $\cup$= "or“, $\cap$= "and“, and A^c="A complement”/”not A”
- Axioms: Let $A, B \subseteq \Omega$
  - $P(A) \geq 0$
  - If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$
  - $P(\Omega) = 1$
- Properties proven from axioms
  - If $A \subseteq B = \emptyset$, then $P(A) \leq P(B)$
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - $P(\emptyset) = 0$
  - $P(A^c) = 1 - P(A)$

Ex: Letter Grades in School

- School collected records of its 3,000 students
- Students in the science class have the following grade distribution (probability law)

Grade	Number of students	Probability
A	300	.10
B	600	.20
C	1500	.50
D	450	.15
F	150	.05

- Experiment = choose at random one of the 3000 students
- Q: What is the probability the student's grade in the science class is an A?
- Q: What is the probability the student's grade is C or higher?

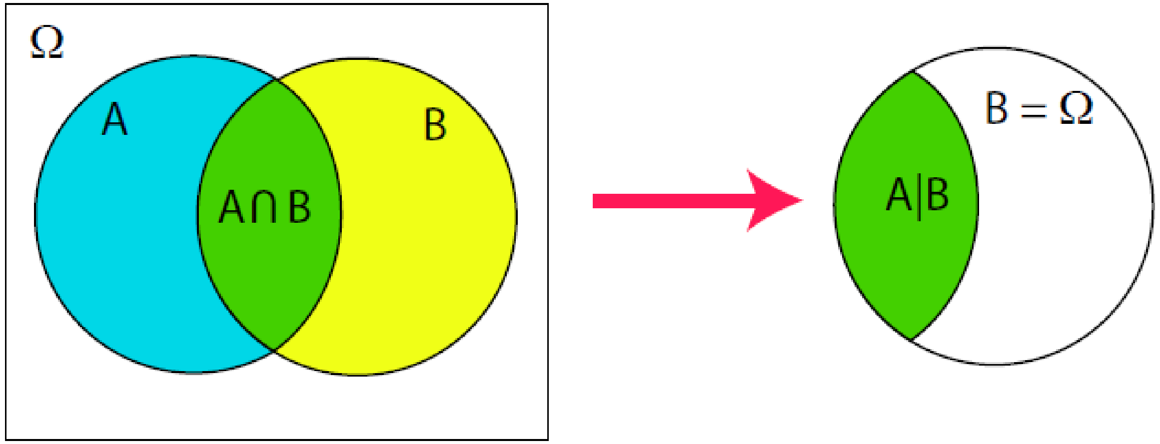
$$\mathrm{P}(A \cup B \cup C) = P(A) + P(B) + P(C) = 0.1 + 0.2 + 0.5 = 0.8 = 1 - P(D \cup F)$$

Conditional Probability

- For any events A and B in the sample space, with $\mathrm{P}(\mathrm{B}) > 0$, the conditional probability of event A given B is defined according to the formula

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{or} \quad P(A|B)P(B) = P(A \cap B)$$

- Visual understanding of conditional probability



Relationships Between Events

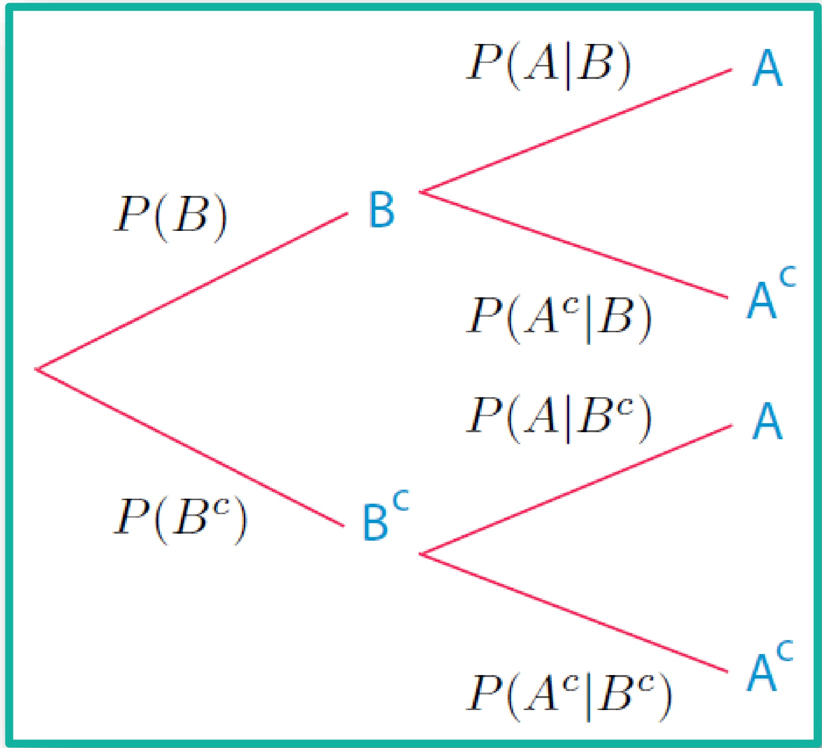
- Two events A and B are mutually exclusive if $A \cap B = \emptyset$
- Mutually exclusive refers to events that cannot occur simultaneously
- Events of getting a 3 on a die roll and 4 on the same die roll are mutually exclusive
- Two events A and B are independent if $\mathrm{P}(A \cap B) = P(A) \times P(B)$
- If two events are independent,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \times P(B)}{P(B)} = P(A)$$

- Independence implies that the probability of a random event is not impacted at all by the occurrence of another event
- Events of getting a 3 on a die roll and a 4 on another die roll are independent

Probability Trees

- A probability tree is a diagram used to represent a probability space from a series of experiments (different or repetitive)
- Each path leads to a different outcome
- Numbers on path indicate probability
- Visualization of conditional probability
- Multiply probabilities along path to find the probabilities of different outcomes

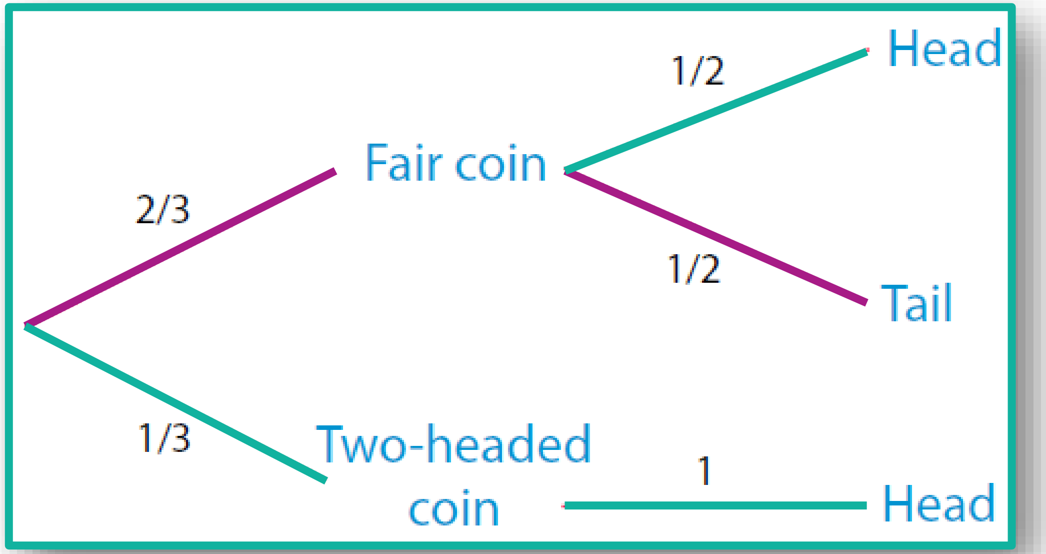


Ex: Flippin' Unfair Coins

- A friend of yours has 3 coins in her pocket, two fair coins and one two-headed
- The two of you are trying to decide whether to watch “The Greatest Showman” or “Pitch Perfect” tonight
- You decide to flip a coin and go see “The Greatest Showman” if it is heads
- Your friend takes out one of the coins without looking and flips it
- Q: What is the probability that you go see “Pitch Perfect” ?
- Q: What is the probability that you go see “The Greatest Showman” ?

Ex: Flippin' Unfair Coins

- Diagram of this example



- Purple indicates the path to watching "Pitch Perfect"
- Teal indicates the path to watching "The Greatest Showman"

Ex: Flippin' Unfair Coins

- Probability of "Pitch Perfect"

$$P(Pitch\ Perfect) = P(Tails) = P(Tails \cap Fair\ Coin)$$
$$= P(Tails|Fair\ Coin)P(Fair\ Coin) = \frac{1}{2} \times \frac{2}{3} = \frac{1}{3} = 0.3333$$

- Probability of "The Greatest Showman"

$$P(The\ Greatest\ Showman) = P(Heads)$$
$$= P(Heads \cap Fair\ Coin) + P(Heads \cap Unfair\ Coin)$$
$$= P(Heads|Fair\ Coin)P(Fair\ Coin) + P(Heads|Unfair\ Coin)P(Unfair\ Coin)$$
$$= \frac{1}{2} \times \frac{2}{3} + 1 \times \frac{1}{3} = \frac{1}{3} + \frac{1}{3} = \frac{2}{3} = 0.6667 = 1 - 0.3333$$

Binomial Probability

- Consider tossing a coin with probability of heads equal to p a total of 6 times
- Q: What is the probability that we get exactly 3 heads?
- We could express all possible outcomes of tossing a coin 6 times using a tree diagram that goes on forever but we all have lives
- Let's consider a few of the outcomes (sequences) where we get exactly 3 heads
- If $\mathrm{A} = \text{Event of Exactly 3 Heads, then }\ \mathrm{A} = \{HHHTTT, TTTHHH, HTHTHT, \cdots\}$
- For each outcome where A occurs, the probability is $F^3(1-p)^3$ because each coin flip is independent
- Q: How many such sequences exist where A occurs?

Binomial Probability

- Number of ways in which we can choose *k* items out of *n* distinct things is called n choose k, denoted by $\binom{n}{k}$, and computed by

$$\binom{n}{k} = \frac{\mathrm{n!}}{(\mathrm{n-k})!\,\mathrm{k!}}$$

where $\mathrm{n! = n \times (n-1) \times (n-2) \cdots 3 \times 2 \times 1}$ (n factorial)

- The numbers $\binom{n}{0}, \binom{n}{1}, \binom{n}{2}, \cdots, \binom{n}{n-1}, \binom{n}{n}$ are called binomial coefficients, since

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

- From coin example, $\mathrm{P(Exactly\ 3\ Heads) = P(A)} = \binom{6}{3} p^3 (1-p)^3$

Binomial Probability

- From coin example,

$$\mathrm{P(Exactly\ 3\ Heads)} = \mathrm{P(A)} = \binom{6}{3} p^3 (1-p)^3$$

- Bernouilli process is a repetition of fixed number of independent trials with a binary outcome where the probability of each outcome remains constant

- Each trial/experiment is called a Bernouilli trial

- For a Bernouilli process, the probability of k successes in n trials is

$$\binom{n}{k} p^n (1-p)^{n-k}$$

- These probabilities build the binomial distribution

- Excel formula is $= BINOM.DIST(n, k, p, FALSE)$

The End

Dale