PROBABILITY

Foundational Sessions Housekeeping

- The use of disrespectful language is prohibited in the questions, this is a supportive, learning environment for all - please engage accordingly. **(FBV: Mutual Respect.)**
- No question is daft or silly - **ask them!**
- There are **Q&A sessions** midway and at the end of the session, should you wish to ask any follow-up questions. Moderators are going to be answering questions as the session progresses as well.
- If you have any questions outside of this lecture, or that are not answered during this lecture, please do submit these for upcoming Open Classes. You can submit these questions here: SE Open Class Questions **or** DS Open Class Questions

CoGrammar

Foundational Sessions Housekeeping cont.

- For all **non-academic questions**, please submit a query:
  www.hyperiondev.com/support

- Report a **safeguarding** incident:
  www.hyperiondev.com/safeguardreporting

- We would love your **feedback** on lectures: Feedback on Lectures

Reminders!

Guided Learning Hours

By now, ideally you should have 7 GLHs per week accrued. Remember to attend any and all sessions for support, and to ensure you reach 112 GLHs by the close of your Skills Bootcamp.

Progression Criteria

✅ **Criterion 1: Initial Requirements**

- Complete **15 hours of Guided Learning Hours** and the **first four tasks** within **two weeks**.

✅ **Criterion 2: Mid-Course Progress**

- **Software Engineering:** Finish 14 tasks by week 8.
- **Data Science: Finish** 13 tasks by week 8.

✅ **Criterion 3: Post-Course Progress**

- Complete all mandatory tasks by **24th March 2024**.
- Record an Invitation to Interview within 4 weeks of course completion, or by 30th March 2024.
- Achieve 112 GLH by 24th March 2024.

✅ **Criterion 4: Employability**

- Record a Final Job Outcome within 12 weeks of graduation, or by 23rd September 2024.

What does the probability value of an event indicate?

A. The frequency of the event in a series of trials
B. The likelihood of the occurrence of the event
C. The duration of the event
D. The impact of the event

What is a 'sample space' in probability?

A. The space where experiments are conducted
B. The set of all possible outcomes of a probability experiment
C. A type of probability distribution
D. The outcome with the highest probability.

Two events are independent if:

A. The occurrence of one affects the probability of the occurrence of the other
B. They occur simultaneously
C. The occurrence of one does not affect the probability of the occurrence of the other
D. They are mutually exclusive events

Recap of Linear Algebra

CoGrammar

Vectors, Matrices, and Operations

Vector: quantities having both magnitude and direction, represented as an array of numbers.

- Example: $\vec{v} = [3, 4]$ represents movement 3 units to the right and 4 units up

Matrices: rectangular arrays of numbers or expressions, used to represent complex data structures or transformations.

- A 2 x 2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ could represent a linear Transformation in a plane

Scalar Operations: multiplying a vector by a scalar changes its magnitude but not direction.

Dot Product: a measure of the similarity of two vectors, calculated as the sum of the products of their corresponding entries.

Probability Topics

1. Sample Space and Events
2. Basic Probability Theory
3. Addition and Multiplication Rules
4. Conditional Probability and Independence
5. Common Distributions

Predicting Customer Churn

Consider a telecommunications company that is experiencing a high rate of customer churn (customers leaving for competitors). We want to predict which customers are most likely to churn so the company can take action.

- How do we use Probability Theory to Predict Customer Churn?

Example: Coin Toss

- Sample Space: $S = \{Heads, Tails\}$.
- Probability of an Event: $P(E) = \frac{Number\ of\ favorable\ outcomes}{Total\ number\ of\ outcomes}$
- For a fair coin, $P(Heads) = \frac{1}{2}$

Sample Space and Events

- **Sample Space:** The set of all possible outcomes.
- **Events:** Specific outcomes or sets of outcomes from the sample space.

E.g. If $\{1, 2, 3, 4, 5, 6\}$ is the sample space,

then $\{2, 4, 6\}$ is one of the events.

CoGrammar

Basic Probability Theory

- Probability of an Event: $P(E) = \dfrac{Number\ of\ favorable\ outcomes}{Total\ number\ of\ outcomes}$

- If we roll a dice and want to know the probability of getting a 4, then $P(4) = \dfrac{1\ (since\ there\ is\ just\ one\ 4)}{6\ (since\ there\ are\ 6\ possible\ numbers)}$

Addition and Multiplication Rules

- **Addition Rule:** For mutually exclusive events A and B, $P(A \text{ } or \text{ } B) = P(A) + P(B)$. This cannot exceed 1.

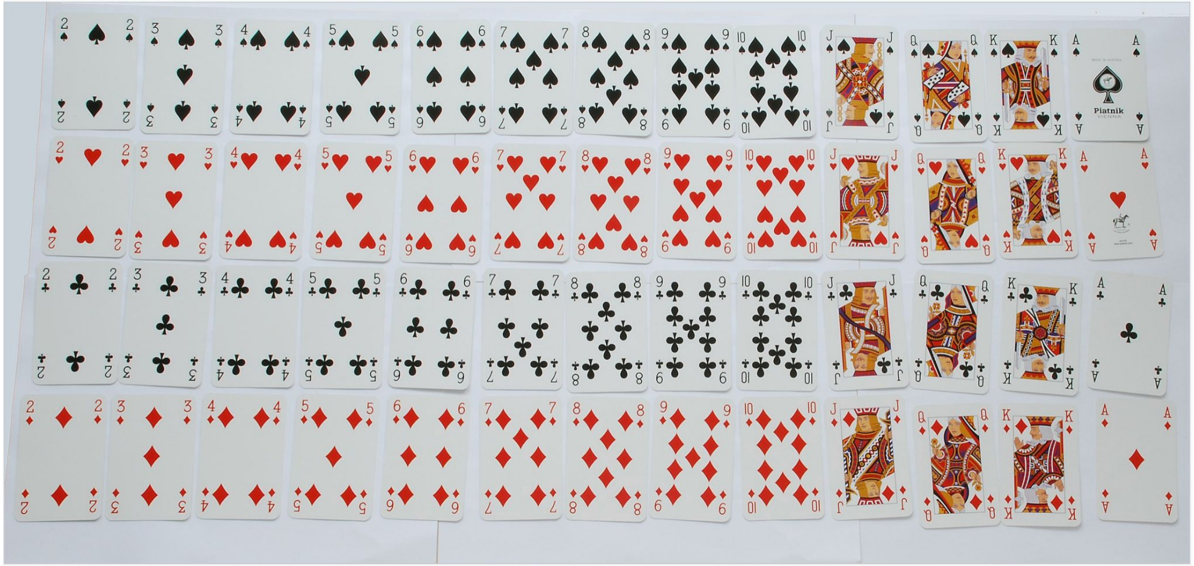
- For example, to find the probability of landing on a 4 or 5 with a fair dice: $P(4 \text{ } or \text{ } 5) = P(4) + P(5) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$

- **Multiplication Rule:** For independent events A and B, $P(A \text{ } and \text{ } B) = P(A) \times P(B)$. Try finding P(4 and 5).

Conditional Probability and Independence

- **Condition Probability:** $P(A|B) = \frac{P(A\ and\ B)}{P(B)}$ is the probability that A happened given that B already happened. E.g. $P(Heart|Red) = \frac{13}{26} = \frac{1}{2}$

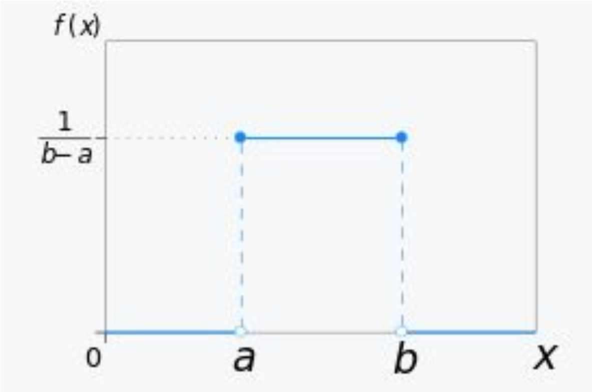
- **Independence:** Events A and B are independent if $P(A|B) = P(A)$ And $P(B|A) = P(B)$.



- For reference, this is what a standard deck of 52 playing cards look like.

Uniform Distribution

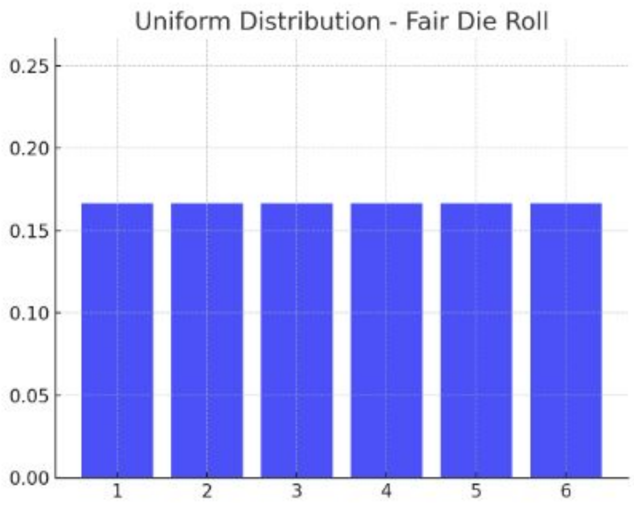
- In a uniform distribution all outcomes are equally likely.



$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \le x \le b, \\ 0 & \text{for } x < a \ \text{ or } \ x > b. \end{cases}$$

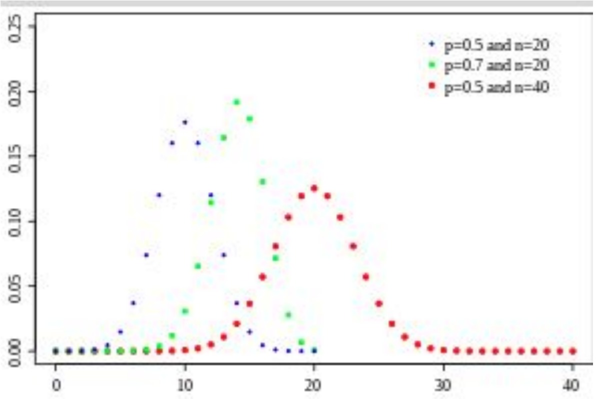
Source: https://en.wikipedia.org/wiki/Continuous_uniform_distribution

- An example is a fair 6-sided die, which has P(x)=⅙ for all sides.



Binomial Distribution

- Number of success in a fixed number of trials.



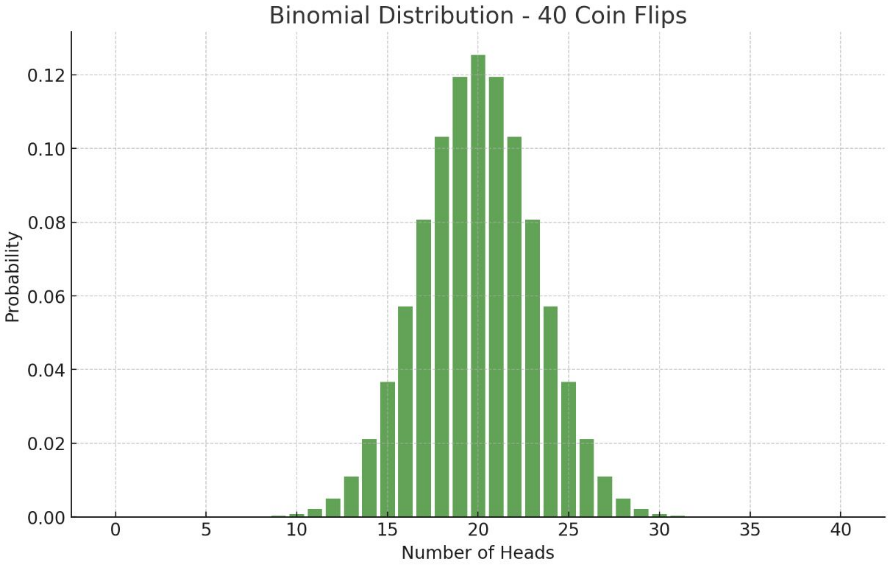
$$f(k,n,p) = \Pr(k;n,p) = \Pr(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

for $k = 0, 1, 2, \ldots, n$, where

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

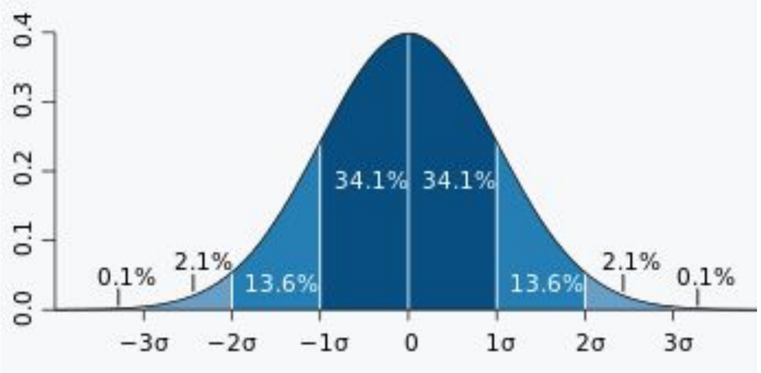
Source: https://en.wikipedia.org/wiki/Binomial_distribution

- To get the probability of getting 20 heads in a coin toss when doing 40 trials, substitute in p=½ , n=40, k=20, to get P(40,20,½) = 0.125



Normal Distribution

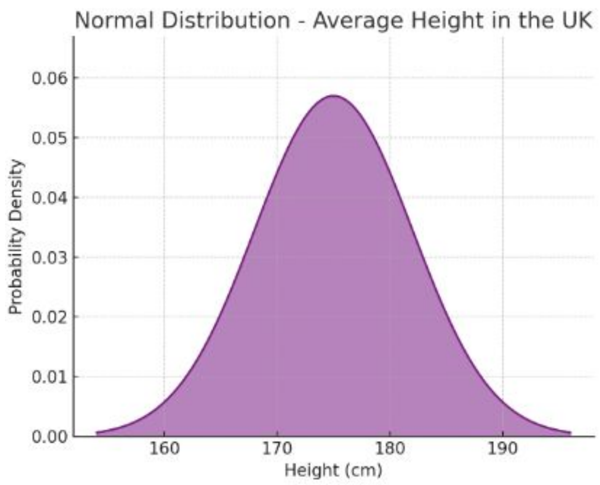
- Describes data in clusters around a mean. It is the most common distribution in statistics since it tends to represent natural phenomena more accurately than most other distributions most of the time.



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Source: https://en.wikipedia.org/wiki/Normal_distribution

- An example is the height of people. The probability of a male in the UK being between 168 cm (one standard deviation below the mean) and 182 cm (one standard deviation above the mean) is approximately 0.683.



- We get this by calculating the area underneath the curve with P(182)-P(168) where the mean is 175 cm and the standard deviation is 7 cm.

Worked Example

A telecom company identifies two customer types: those on month-to-month plans (Type A), likely to churn for better deals, and those on long-term contracts (Type B), prioritizing service quality. Historical data suggests different churn probabilities for each type.

Historical Data Probabilities:

P(churn | Type A) = 0.25

P(churn | Type B) = 0.10

P(Type A) = 0.60

P(Type B) = 0.40

1. What's the overall probability of a customer churning?

Worked Example

A telecom company identifies two customer types: those on month-to-month plans (Type A), likely to churn for better deals, and those on long-term contracts (Type B), prioritizing service quality. Historical data suggests different churn probabilities for each type.

Historical Data Probabilities:
P(churn | Type A) = 0.25
P(churn | Type B) = 0.10
P(Type A) = 0.60
P(Type B) = 0.40

1. What's the overall probability of a customer churning?

P(Type A churns) = P(Type A and churns) = P(Type A) x P(Type A churns) = P(Type A) x P(churn | Type A) = 0.6 x 0.25 = 0.15
P(Type B churns) = P(Type B and churns) = P(Type B) x P(Type B churns) = P(Type B) x P(churn | Type B) = 0.4 x 0.1 = 0.04

P(Type A churns or Type B churns) = P(Type A churns) + P(Type B churns) = 0.19

Thus, there is a **19% chance** of a random customer churning.

Note: for those curious you could research **Law of Total Probability**, which is used to solve this easily

CoGrammar

Summary

Sample Space and Events

★ The set of all possible outcomes of an experiment.
★ An event is a subset of the sample space that we are interested in.

Basic Probability

★ The likelihood of an event occurring, calculated as favorable outcomes divided by total outcomes.

Conditional Probability and Independence

★ Probability of an event given another has occurred.
★ Independence when one event does not influence another.

Summary

Probability Distributions

★ Uniform: Equal probability for all outcomes.
★ Binomial: Probability of 'success' in 'n' trials.
★ Normal: Bell-curved distribution, common in natural data.

Further Learning

- Khan Academy - Basic Probability
- LibreTexts - Basic Probability more examples
- Coursera - Basic to Advanced Probability (for the VERY curious)

What is conditional probability?

A. The probability of two events occurring together
B. The probability of an event, given that another event has occurred
C. The likelihood of an event occurring after a series of other events
D. The probability of an event occurring without any conditions.

Which statement is true about the normal distribution?

- **A.** It is skewed to the left or right.
- **B.** It is a distribution where all outcomes are equally likely.
- **C.** It is symmetric around its mean.
- **D.** It applies only to discrete random variables.

Questions and Answers

Questions around Probability