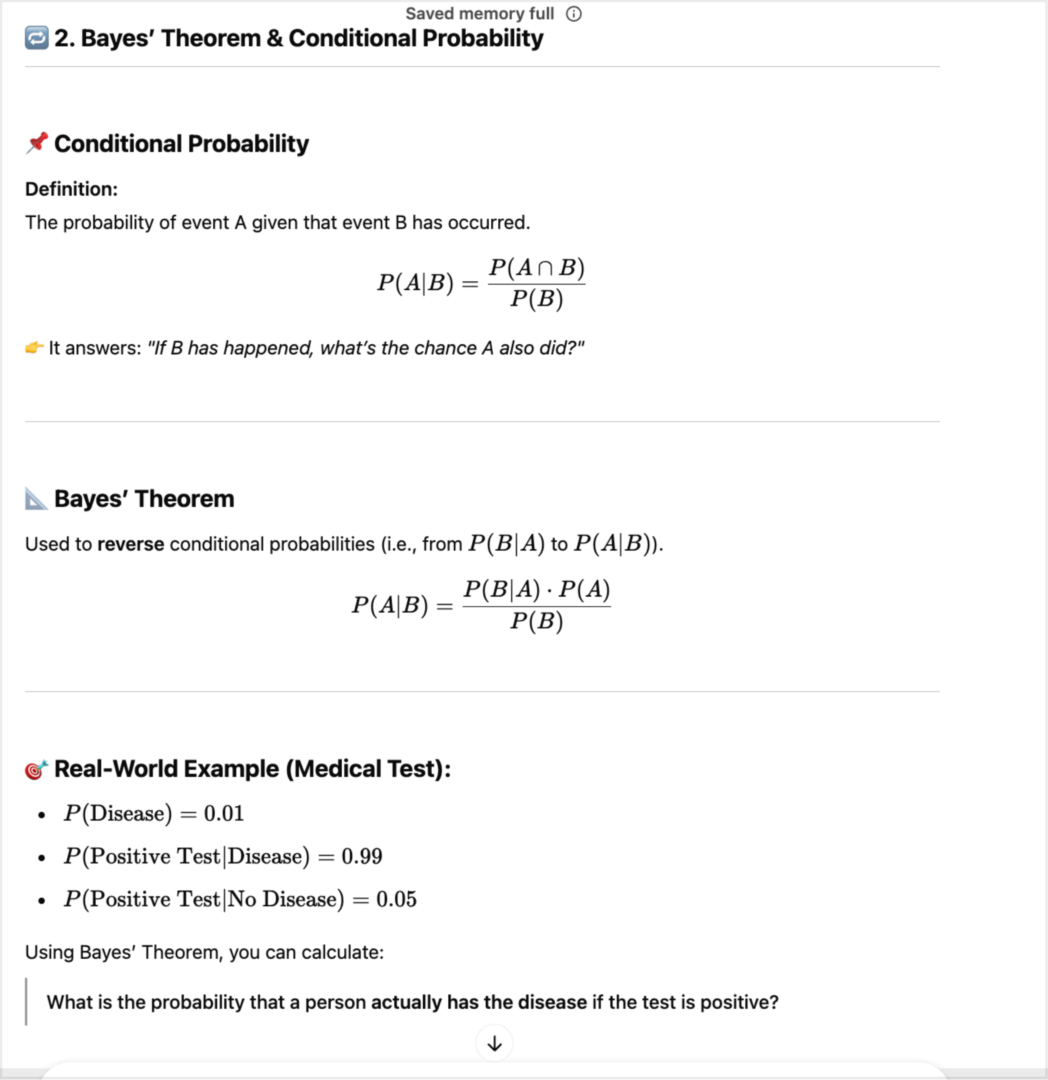
Probability and statistics

1. **Basic Probability**
2. **Bayes Theorem and Conditional Probability**
3. **Probability Distributions**
4. **Random Variables, Expectations, and Variance**
5. **Moments & CLT**
6. **Statistical Models: Regression**
7. **Hypothesis Testing, P-value**



🔁 2. Bayes' Theorem & Conditional Probability

📌 Conditional Probability

Definition:
The probability of event A given that event B has occurred.

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

👉 It answers: *"If B has happened, what's the chance A also did?"*

📐 Bayes' Theorem

Used to **reverse** conditional probabilities (i.e., from $P(B|A)$ to $P(A|B)$).

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

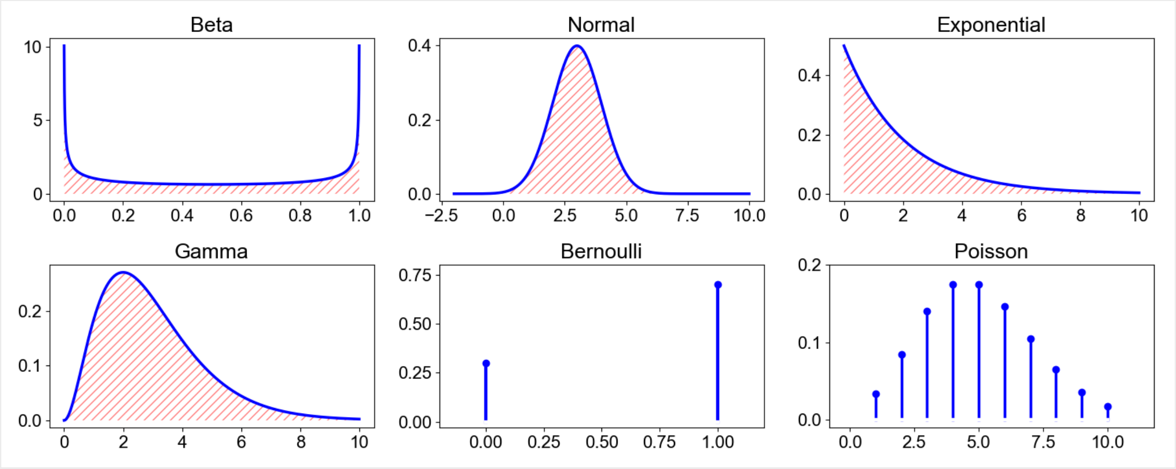
🎯 Real-World Example (Medical Test):

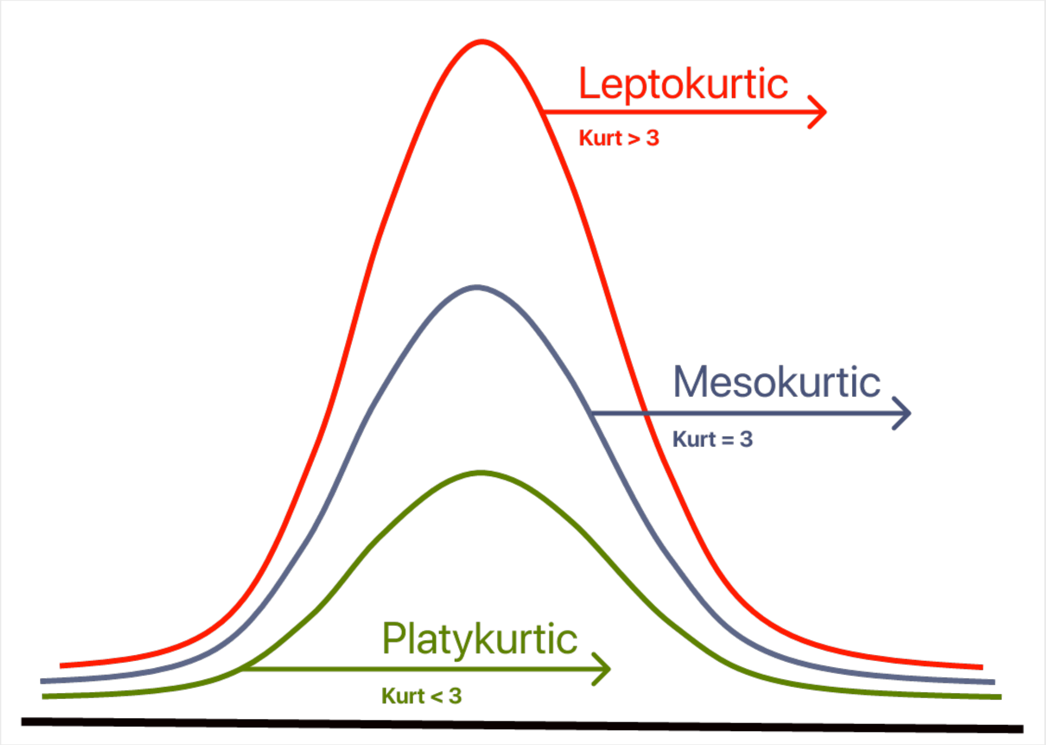
- $P(\text{Disease}) = 0.01$
- $P(\text{Positive Test}|\text{Disease}) = 0.99$
- $P(\text{Positive Test}|\text{No Disease}) = 0.05$

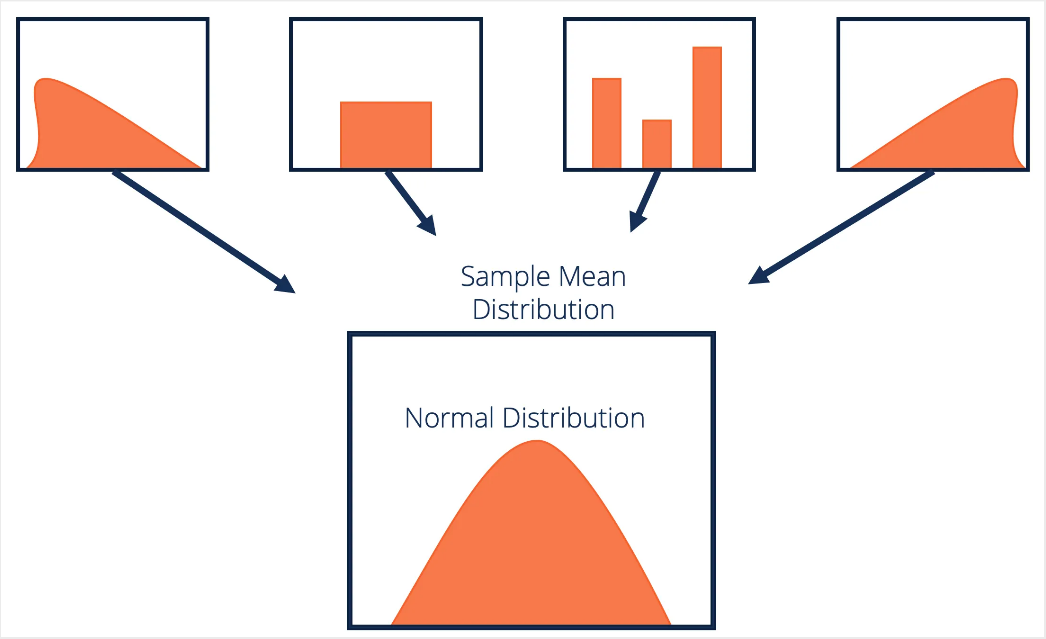
Using Bayes' Theorem, you can calculate:

> What is the probability that a person **actually has the disease** if the test is positive?

Moments & CLT


Beta
10
5
0
0.0
0.2
0.4
0.6
0.8
1.0
Normal
0.4
0.2
0.0
−2.5
0.0
2.5
5.0
7.5
10.0
Exponential
0.4
0.2
0.0
0
2
4
6
8
10
Gamma
0.2
0.1
0.0
0
2
4
6
8
10
Bernoulli
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00
Poisson
0.2
0.1
0.0
0.0
2.5
5.0
7.5
10.0


Leptokurtic
Kurt > 3
Mesokurtic
Kurt = 3
Platykurtic
Kurt < 3



📈 Statistical Models: Regression

🔍 What is Regression?

Regression is a **statistical technique** used to model the relationship between a **dependent variable** and one or more **independent variables**.

🧠 Types of Regression:

Type	Description	Example Use Case
Linear Regression	Models a straight-line relationship	Predicting house prices 🏠
Multiple Regression	Uses multiple predictors	Predicting sales using ads, season 📊
Logistic Regression	Predicts binary outcomes (Yes/No)	Spam detection 📧
Polynomial Regression	Models curves (non-linear relationships)	Growth modeling 🌱

🔢 Linear Regression Formula:

$y=\beta0 +\beta1\ x+\varepsilon$

- y: Dependent variable
- x: Independent variable
- $\beta0$: Intercept
- $\beta1$: Slope
- ε: Error term

📊 Why Regression is Useful in ML & Data Science:

- Predicting outcomes (sales, grades, etc.)
- Understanding the influence of variables
- Feature selection and model interpretation

📊 Probability Distributions

A **probability distribution** describes how the values of a random variable are distributed — essentially, how likely each outcome is.

🔄 Types of Probability Distributions

A. Discrete Distributions (Countable outcomes)

- **Binomial Distribution**: Success/failure (e.g., flipping a coin)
- **Poisson Distribution**: Number of events in a fixed time (e.g., calls per hour)

B. Continuous Distributions (Infinite outcomes in an interval)

- **Normal Distribution (Gaussian)**: Bell-shaped curve (e.g., height, IQ scores)
- **Exponential Distribution**: Time between events (e.g., waiting time)

🎯 What is Hypothesis Testing?

A statistical method used to make **inferences** or **decisions** about a population based on sample data.

🧪 Key Steps:

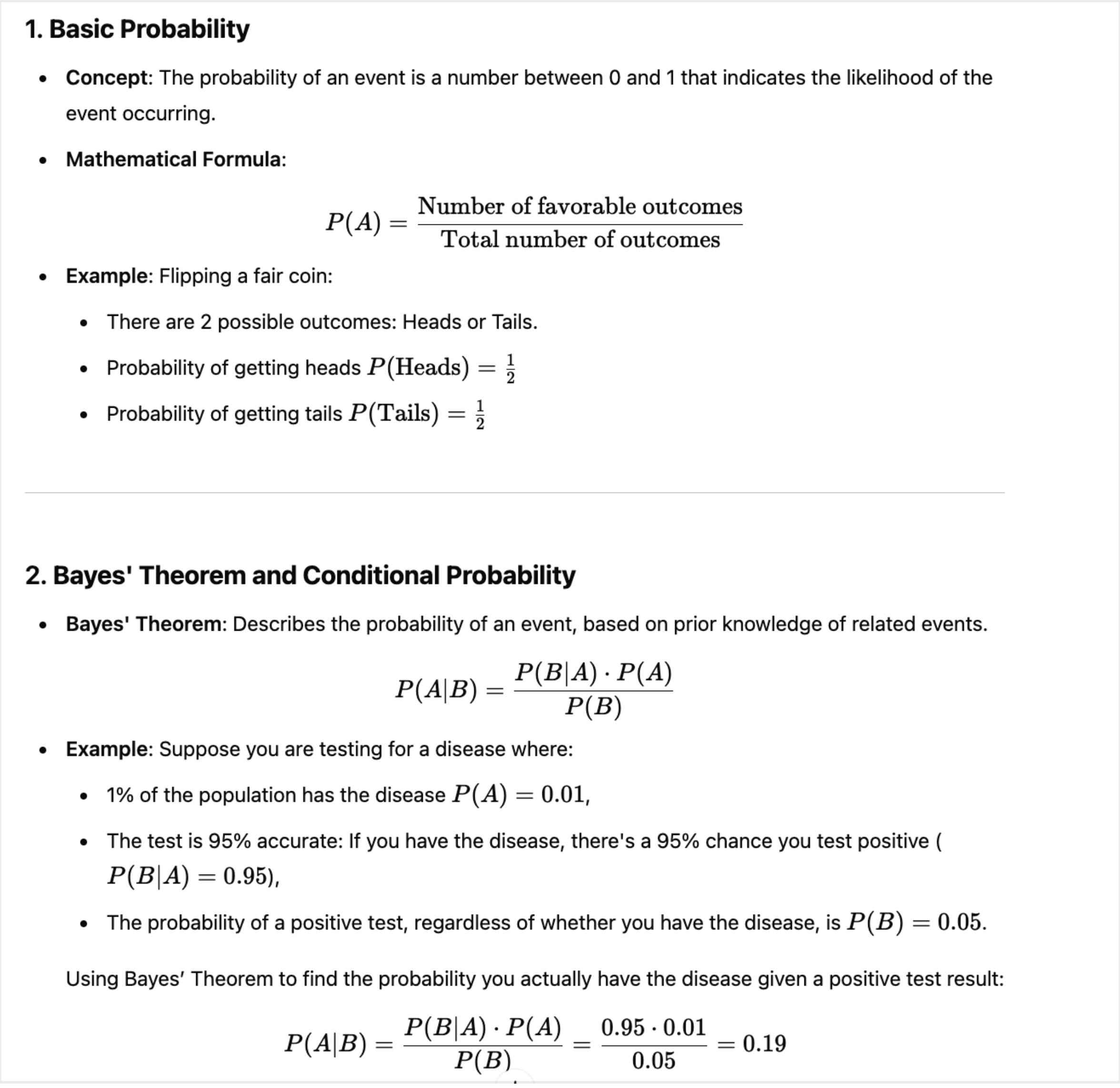
1. **State Hypotheses:**
    - **Null Hypothesis (H0)**: No effect or no difference
    - **Alternative Hypothesis (H1 or Ha)**: There *is* an effect or difference
2. **Choose Significance Level (α):**
    - Common values: **0.05**, **0.01**, **0.10**
3. **Select and Compute Test Statistic:**
    - e.g., **t-test**, **z-test**, **chi-square**, etc.
4. **Find p-value** or compare with **critical value**
5. **Make Decision:**
    - If **p ≤ α** → Reject H0
    - If **p > α** → Fail to reject H0
6. **Draw Conclusion:**
    - Relate the decision to the context of the problem

P-Value

The p-value is a crucial concept in statistical hypothesis testing.

A small p-value (typically ≤ 0.05) indicates strong evidence against the null hypothesis. This means that it's unlikely to observe the data if the null hypothesis were true.

A large p-value indicates weak evidence against the null hypothesis. This means that the data is consistent with the null hypothesis.



1. Basic Probability

- **Concept**: The probability of an event is a number between 0 and 1 that indicates the likelihood of the event occurring.
- **Mathematical Formula**:

$$P(A) = \frac{\text{Number of favorable outcomes}}{\text{Total number of outcomes}}$$

- **Example**: Flipping a fair coin:
  - There are 2 possible outcomes: Heads or Tails.
  - Probability of getting heads $P(\text{Heads}) = \frac{1}{2}$
  - Probability of getting tails $P(\text{Tails}) = \frac{1}{2}$

---

2. Bayes' Theorem and Conditional Probability

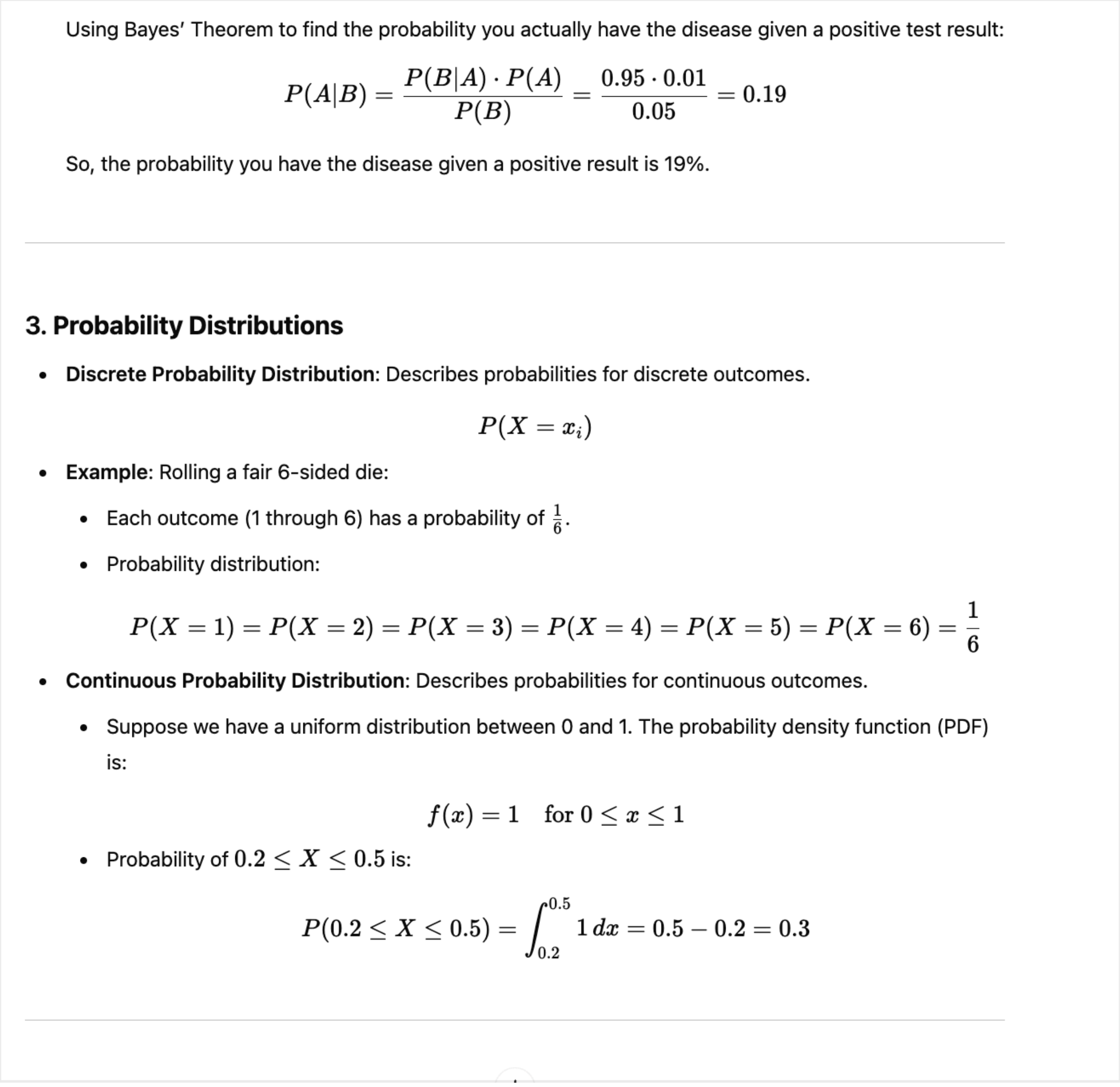
- **Bayes' Theorem**: Describes the probability of an event, based on prior knowledge of related events.

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

- **Example**: Suppose you are testing for a disease where:
  - 1% of the population has the disease $P(A) = 0.01$,
  - The test is 95% accurate: If you have the disease, there's a 95% chance you test positive ($P(B|A) = 0.95$),
  - The probability of a positive test, regardless of whether you have the disease, is $P(B) = 0.05$.

Using Bayes’ Theorem to find the probability you actually have the disease given a positive test result:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} = \frac{0.95 \cdot 0.01}{0.05} = 0.19$$



Using Bayes' Theorem to find the probability you actually have the disease given a positive test result:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} = \frac{0.95 \cdot 0.01}{0.05} = 0.19$$

So, the probability you have the disease given a positive result is 19%.

3. Probability Distributions

- **Discrete Probability Distribution**: Describes probabilities for discrete outcomes.

$$P(X = x_i)$$

- **Example**: Rolling a fair 6-sided die:
  - Each outcome (1 through 6) has a probability of $\frac{1}{6}$.
  - Probability distribution:

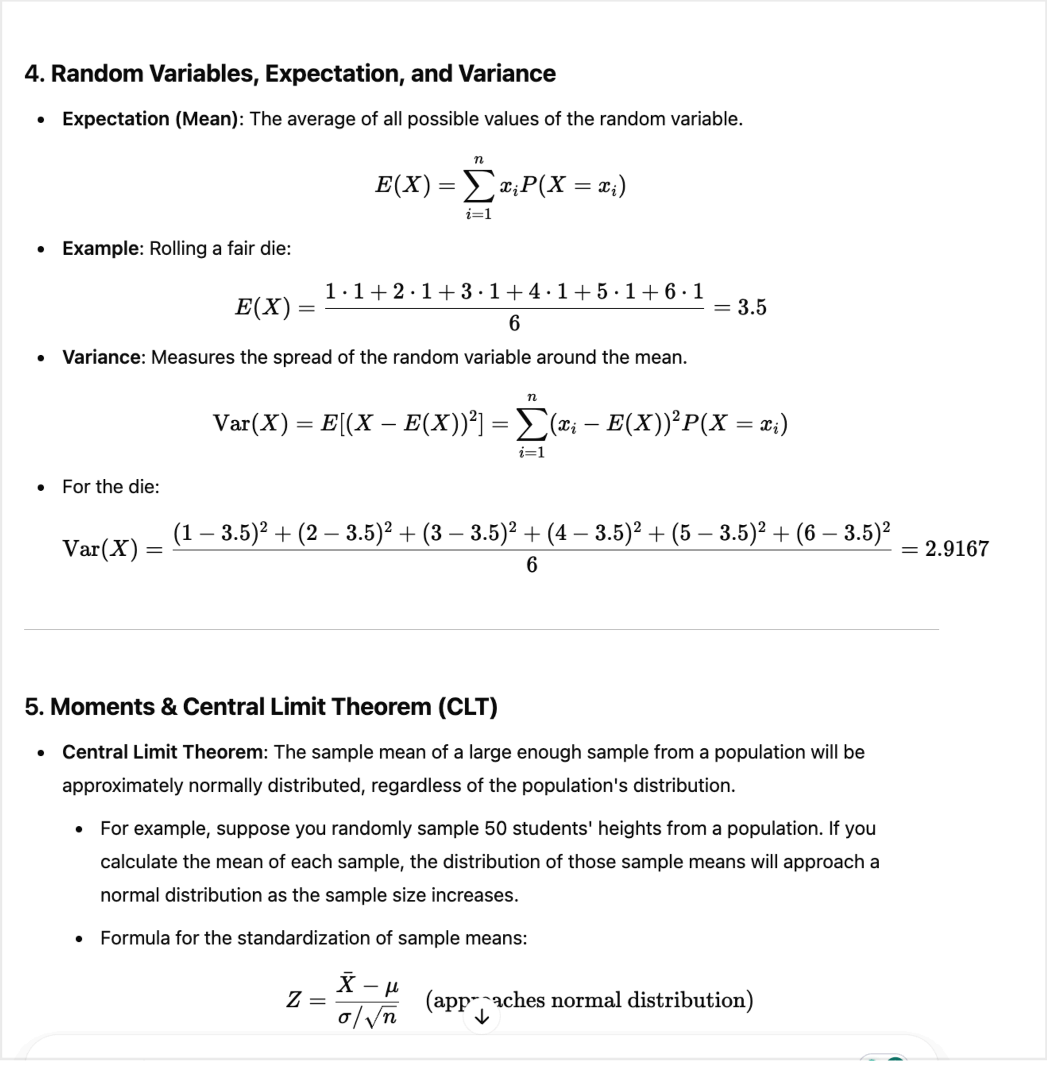
$$P(X = 1) = P(X = 2) = P(X = 3) = P(X = 4) = P(X = 5) = P(X = 6) = \frac{1}{6}$$

- **Continuous Probability Distribution**: Describes probabilities for continuous outcomes.
  - Suppose we have a uniform distribution between 0 and 1. The probability density function (PDF) is:

$$f(x) = 1 \quad \text{for } 0 \leq x \leq 1$$

  - Probability of $0.2 \leq X \leq 0.5$ is:

$$P(0.2 \leq X \leq 0.5) = \int_{0.2}^{0.5} 1\,dx = 0.5 - 0.2 = 0.3$$



4. Random Variables, Expectation, and Variance

- **Expectation (Mean)**: The average of all possible values of the random variable.

$$E(X) = \sum_{i=1}^{n} x_i P(X = x_i)$$

- **Example**: Rolling a fair die:

$$E(X) = \frac{1 \cdot 1 + 2 \cdot 1 + 3 \cdot 1 + 4 \cdot 1 + 5 \cdot 1 + 6 \cdot 1}{6} = 3.5$$

- **Variance**: Measures the spread of the random variable around the mean.

$$\text{Var}(X) = E[(X - E(X))^2] = \sum_{i=1}^{n} (x_i - E(X))^2 P(X = x_i)$$

- For the die:

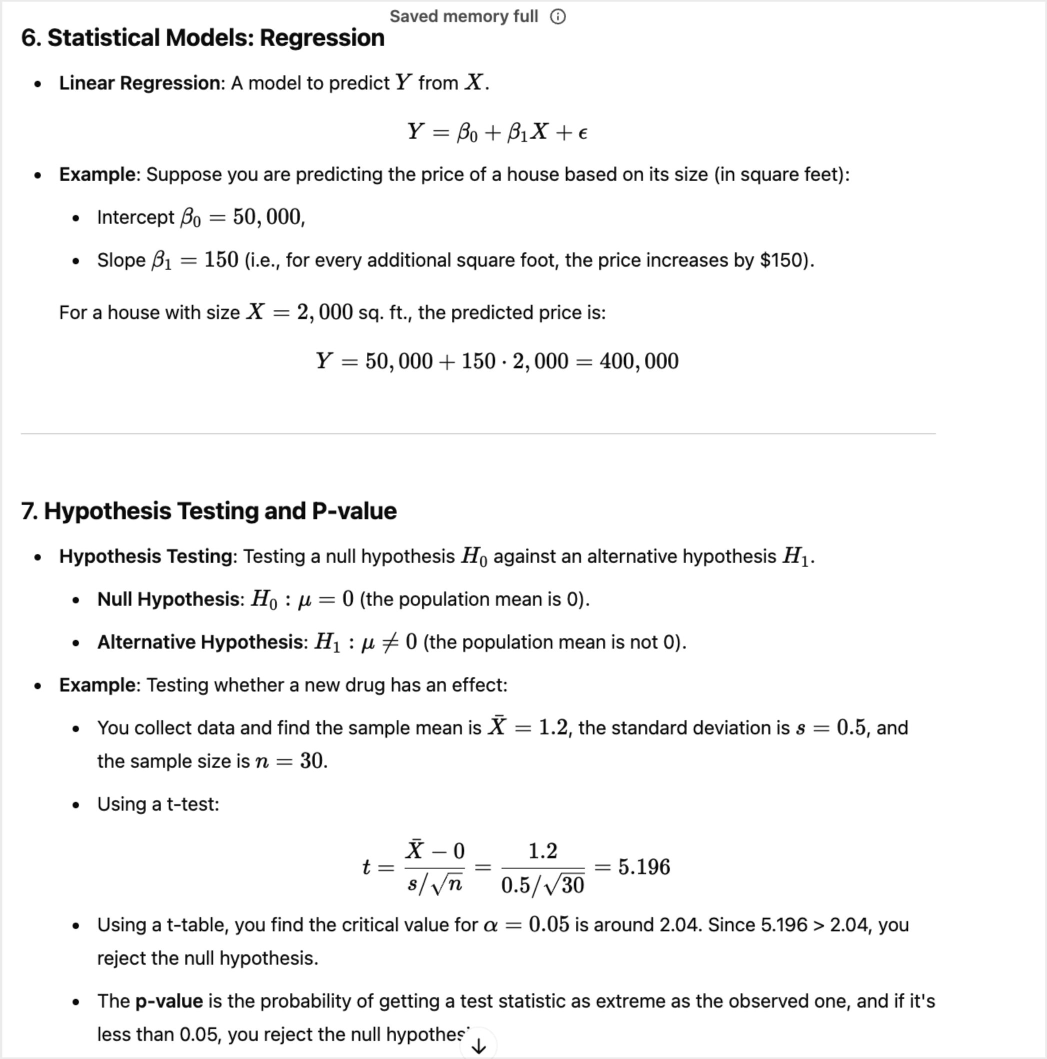
$$\text{Var}(X) = \frac{(1-3.5)^2 + (2-3.5)^2 + (3-3.5)^2 + (4-3.5)^2 + (5-3.5)^2 + (6-3.5)^2}{6} = 2.9167$$

---

5. Moments & Central Limit Theorem (CLT)

- **Central Limit Theorem**: The sample mean of a large enough sample from a population will be approximately normally distributed, regardless of the population's distribution.
  - For example, suppose you randomly sample 50 students' heights from a population. If you calculate the mean of each sample, the distribution of those sample means will approach a normal distribution as the sample size increases.
  - Formula for the standardization of sample means:

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \quad \text{(approaches normal distribution)}$$



Saved memory full ⓘ

6. Statistical Models: Regression

- **Linear Regression**: A model to predict Y from X.

$$Y = \beta_0 + \beta_1 X + \epsilon$$

- **Example**: Suppose you are predicting the price of a house based on its size (in square feet):
  - Intercept $\beta_0 = 50,000$,
  - Slope $\beta_1 = 150$ (i.e., for every additional square foot, the price increases by $150).

  For a house with size $X = 2,000$ sq. ft., the predicted price is:

$$Y = 50,000 + 150 \cdot 2,000 = 400,000$$

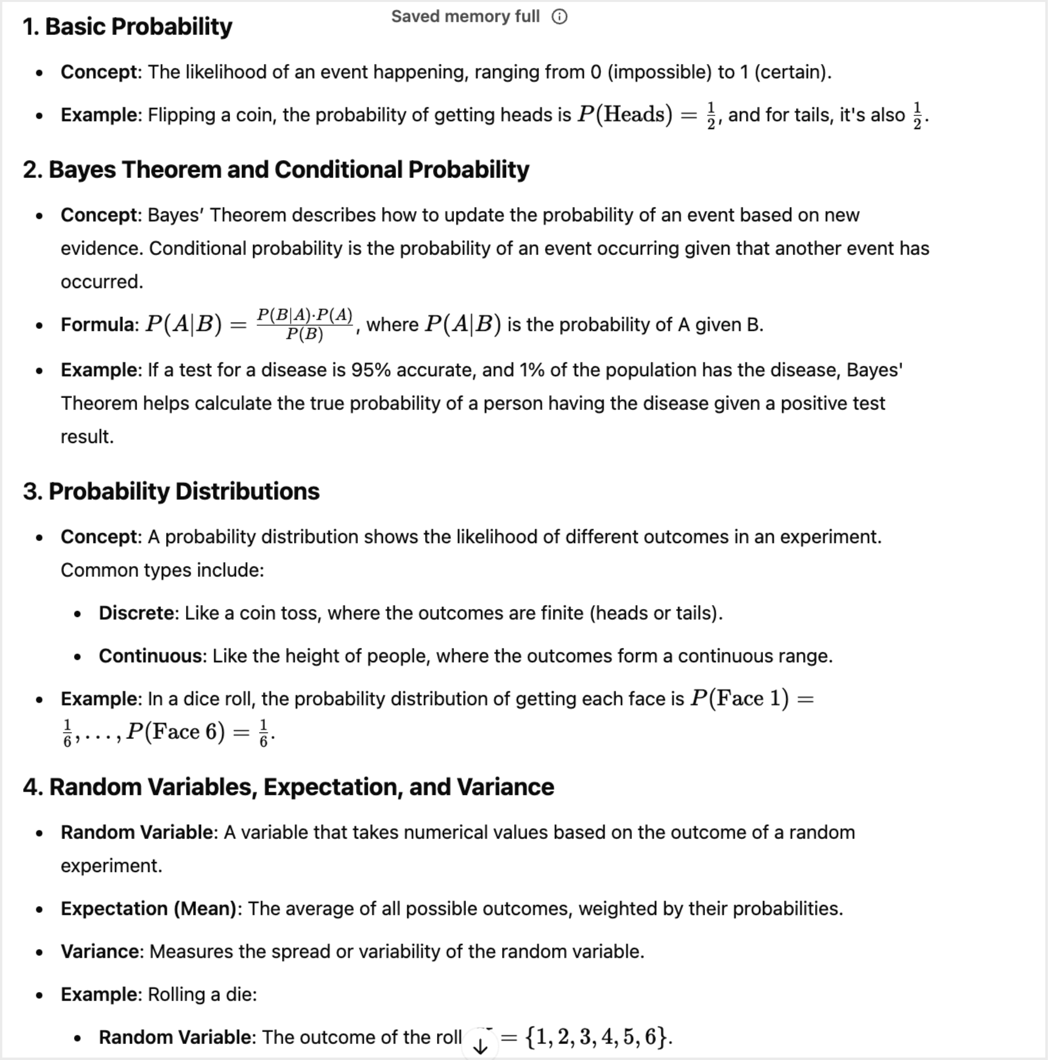
---

7. Hypothesis Testing and P-value

- **Hypothesis Testing**: Testing a null hypothesis H_0 against an alternative hypothesis H_1.
  - **Null Hypothesis**: $H_0 : \mu = 0$ (the population mean is 0).
  - **Alternative Hypothesis**: $H_1 : \mu \neq 0$ (the population mean is not 0).
- **Example**: Testing whether a new drug has an effect:
  - You collect data and find the sample mean is $\bar{X} = 1.2$, the standard deviation is $s = 0.5$, and the sample size is $n = 30$.
  - Using a t-test:

$$t = \frac{\bar{X} - 0}{s/\sqrt{n}} = \frac{1.2}{0.5/\sqrt{30}} = 5.196$$

  - Using a t-table, you find the critical value for $\alpha = 0.05$ is around 2.04. Since 5.196 > 2.04, you reject the null hypothesis.
  - The **p-value** is the probability of getting a test statistic as extreme as the observed one, and if it's less than 0.05, you reject the null hypothes



1. Basic Probability

- **Concept**: The likelihood of an event happening, ranging from 0 (impossible) to 1 (certain).
- **Example**: Flipping a coin, the probability of getting heads is $P(\text{Heads}) = \frac{1}{2}$, and for tails, it's also $\frac{1}{2}$.

2. Bayes Theorem and Conditional Probability

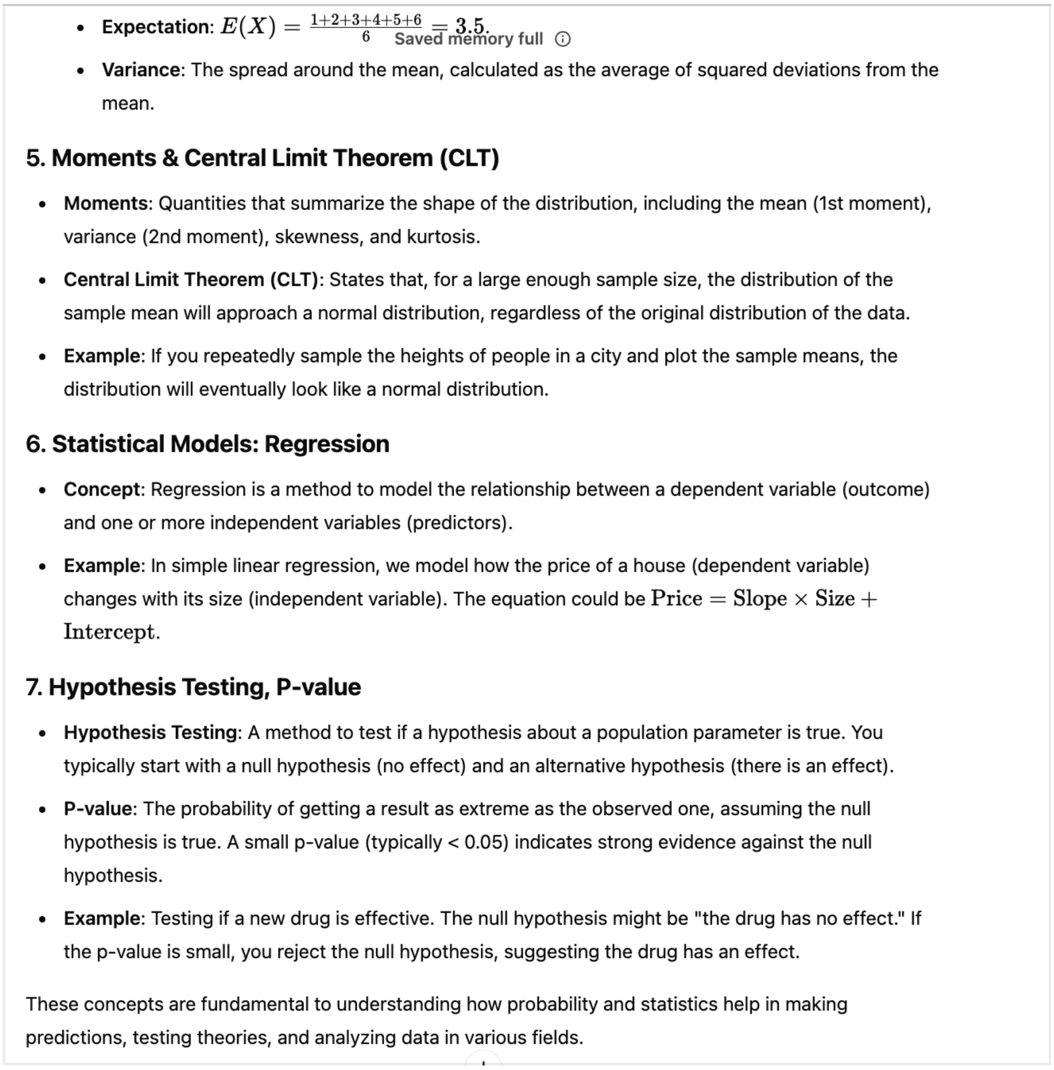
- **Concept**: Bayes' Theorem describes how to update the probability of an event based on new evidence. Conditional probability is the probability of an event occurring given that another event has occurred.
- **Formula**: $P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$, where $P(A|B)$ is the probability of A given B.
- **Example**: If a test for a disease is 95% accurate, and 1% of the population has the disease, Bayes' Theorem helps calculate the true probability of a person having the disease given a positive test result.

3. Probability Distributions

- **Concept**: A probability distribution shows the likelihood of different outcomes in an experiment. Common types include:
  - **Discrete**: Like a coin toss, where the outcomes are finite (heads or tails).
  - **Continuous**: Like the height of people, where the outcomes form a continuous range.
- **Example**: In a dice roll, the probability distribution of getting each face is $P(\text{Face 1}) = \frac{1}{6}, \ldots, P(\text{Face 6}) = \frac{1}{6}$.

4. Random Variables, Expectation, and Variance

- **Random Variable**: A variable that takes numerical values based on the outcome of a random experiment.
- **Expectation (Mean)**: The average of all possible outcomes, weighted by their probabilities.
- **Variance**: Measures the spread or variability of the random variable.
- **Example**: Rolling a die:
  - **Random Variable**: The outcome of the roll $= \{1, 2, 3, 4, 5, 6\}$.



  - **Expectation**: $E(X) = \frac{1+2+3+4+5+6}{6} = 3.5$.
  - **Variance**: The spread around the mean, calculated as the average of squared deviations from the mean.

5. Moments & Central Limit Theorem (CLT)

- **Moments**: Quantities that summarize the shape of the distribution, including the mean (1st moment), variance (2nd moment), skewness, and kurtosis.
- **Central Limit Theorem (CLT)**: States that, for a large enough sample size, the distribution of the sample mean will approach a normal distribution, regardless of the original distribution of the data.
- **Example**: If you repeatedly sample the heights of people in a city and plot the sample means, the distribution will eventually look like a normal distribution.

6. Statistical Models: Regression

- **Concept**: Regression is a method to model the relationship between a dependent variable (outcome) and one or more independent variables (predictors).
- **Example**: In simple linear regression, we model how the price of a house (dependent variable) changes with its size (independent variable). The equation could be $\text{Price} = \text{Slope} \times \text{Size} + \text{Intercept}$.

7. Hypothesis Testing, P-value

- **Hypothesis Testing**: A method to test if a hypothesis about a population parameter is true. You typically start with a null hypothesis (no effect) and an alternative hypothesis (there is an effect).
- **P-value**: The probability of getting a result as extreme as the observed one, assuming the null hypothesis is true. A small p-value (typically < 0.05) indicates strong evidence against the null hypothesis.
- **Example**: Testing if a new drug is effective. The null hypothesis might be "the drug has no effect." If the p-value is small, you reject the null hypothesis, suggesting the drug has an effect.

These concepts are fundamental to understanding how probability and statistics help in making predictions, testing theories, and analyzing data in various fields.