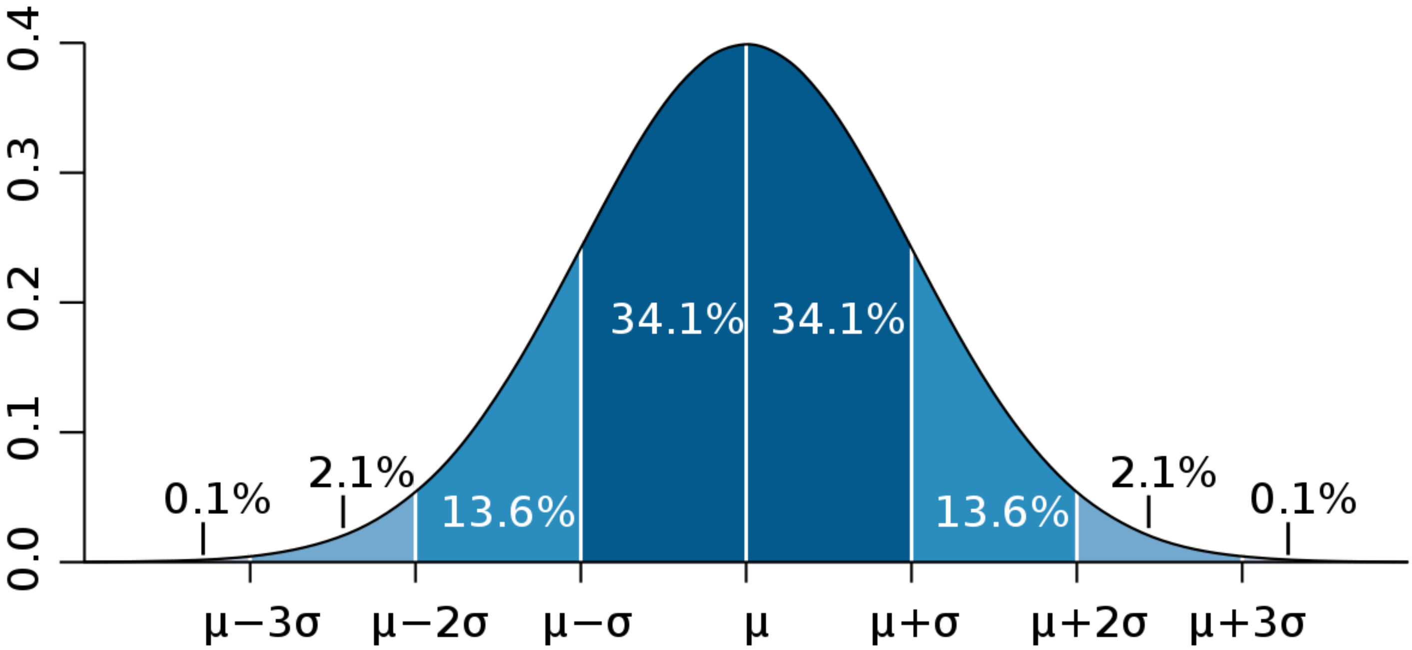
Bytes Of Intelligence
Exploring AI's Secrets

Email ahammadmejbah@gmail.com | GitHub @BytesOfIntelligences | LinkedIn Mejbah Ahammad
Website Bytes of Intelligence | YouTube BytesofIntelligence | ResearchGate Mejbah Ahammad
Phone +8801874603631 | Hackerrank ahammadmejbah



Probability Theory

Image Credit: Probability theory - Wikipedia

1. Foundations of Probability Theory

2. Combinatorics and Probability

- 2.1. Basic Counting Principles
- 2.2. Permutations and Combinations
- 2.3. Applications in Probability

2.1. Definition of Probability

Probability measures the likelihood that an event will occur. It is quantified as a number between 0 and 1, where 0 indicates impossibility, and 1 indicates certainty. The probability of an event (A), denoted as $P(A)$, can be defined in classical, frequentist, or axiomatic terms.

- **Classical Definition**: When all outcomes of a sample space are equally likely, the probability of an event is the ratio of the number of favorable outcomes to the total number of possible outcomes.

$P(A)$ = $\frac{\text{Number of outcomes favorable to} A}{\text{Total number of outcomes in the sample space}}$

- **Frequentist Definition**: Probability is the limit of the relative frequency of an event after many trials.

- **Axiomatic Approach**: Proposed by Kolmogorov, where probabilities are assigned to events based on a set of axioms.

2.2. Probability Axioms

There are three axioms of probability:

1. **Non-negativity**: $(P(A) \geq 0) for any event (A)$.
2. **Normalization**: $(P(S) = 1) for the sample space (S)$.
3. **Additivity**: For any two mutually exclusive events: $(A) and (B), (P(A \cup B) = P(A) + P(B))$

These axioms lay the foundation for all probability theory.

2.3. Conditional Probability and Independence

Conditional probability is the probability of an event given that another event has occurred, denoted as (P(A|B)) and calculated by:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, provided (P(B) > 0.$$

Two events A and B are independent if and only if:

$$P(A \cap B) = P(A)P(B)$$

Independence means the occurrence of one event does not affect the probability of the other.

2.4. Bayes' Theorem

Bayes' Theorem relates the conditional and marginal probabilities of random events:

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

It is especially useful for updating probabilities upon obtaining new information.

2.5. Probability Distributions

2.5.1. Discrete Distributions

Discrete distributions describe the probability of outcomes of discrete random variables (e.g., Bernoulli, Binomial, Poisson distributions).

- **Bernoulli Distribution**: Probability distribution of a random variable which takes the value 1 with probability (p) and the value 0 with probability (1-p).
- **Binomial Distribution**: Describes the number of successes in a sequence of independent yes/no experiments, each of which yields success with probability (p).
- **Poisson Distribution**: Expresses the probability of a given number of events happening in a fixed interval of time/space, given the events occur with a known constant mean rate and independently of the time since the last event.

2.5.2. Continuous Distributions

Continuous distributions describe the probability of outcomes of continuous random variables (e.g., Uniform, Normal, Exponential distributions).

- **Normal Distribution**: Described by its mean ((\mu)) and standard deviation ((\sigma)), it shows the distribution of a continuous variable where most observations cluster around the central peak.
- **Exponential Distribution**: Describes the time between events in a Poisson point process, i.e., a process in which events occur continuously and independently at a constant average rate.

2.6. Expected Value and Variance

- **Expected Value** ((E[X])) of a random variable (X) is a measure of the center of the distribution of (X). For a discrete random variable, it's calculated as:

$$E[X] = \sum_i x_i P(x_i)$$

For a continuous random variable:

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx$$

- **Variance** ((Var(X))) measures the spread of a distribution. It's the expected value of the squared deviation of (X) from its mean:

$$Var(X) = E[(X - E[X])^2]$$

For each subsection, let's implement a Python example to illustrate these concepts. Due to the extensive material, we'll start with 2.1 and 2.2, and then proceed to the others in follow-up responses. Let's begin with the basics: calculating simple probabilities and applying the axioms.

Here are the basic calculations and applications of probability axioms using a simple dice roll example:

- **Probability of Rolling a 4**: The probability of rolling a 4 with a fair six-sided dice is $\frac{1}{6}$ or approximately 0.167.
- **Probability of Rolling an Even Number (2, 4, 6)**: The probability is $\frac{3}{6}$ or 0.5, since three out of the six outcomes are even numbers.
- **Axiom 1 (Non-negativity)**: Confirmed, as both calculated probabilities are non-negative.
- **Axiom 2 (Normalization)**: The probability of the sample space, which includes all outcomes (1 through 6), is 1, fulfilling the normalization axiom.
- **Axiom 3 (Additivity)**: The probability of rolling either a 4 or a 5 (considering them as mutually exclusive for the sake of this example) is $\frac{1}{6}$ + $\frac{1}{6}$ = $\frac{2}{6}$ or approximately 0.333.

```
#Calculation of probability and application of axioms
# Let's consider a simple dice roll (a fair six-sided dice)
total_outcomes = 6  # Total possible outcomes (1, 2, 3, 4, 5, 6)

# Probability of rolling a 4
P_roll_4 = 1 / total_outcomes

# Probability of rolling an even number (2, 4, 6)
P_even = 3 / total_outcomes

# Axiom 1: Non-negativity
non_negativity = P_roll_4 >= 0 and P_even >= 0

# Axiom 2: Normalization
# Considering the sample space S includes all outcomes of the dice roll
P_sample_space = 1  # By definition, since all possible outcomes are covered

# Axiom 3: Additivity
```

```
    # Probability of rolling a 4 or an even number
    #(mutually exclusive in this context is not strictly correct since 4 is even)
    # However, let's apply it to rolling a 4 or a 5 for clarity on mutual exclusivity
    P_roll_4_or_5 = 1/total_outcomes + 1/total_outcomes

    P_roll_4, P_even, non_negativity, P_sample_space, P_roll_4_or_5
```

Conditional Probability, independence, and Bayes' theorem examples.

Conditional probability and independence, using a deck of 52 cards:

- **Probability of Drawing a Heart (A)**: $P(A)$ = $\frac{13}{52}$ = 0.25
- **Probability of Drawing a Red Card (B)**: $P(B)$ = $\frac{26}{52}$ = 0.5
- **Conditional Probability of Drawing a Heart Given It's a Red Card (P(A|B))**: Since all hearts are red, $P(A|B)$ = $\frac{P(A)}{P(B)}$ = 0.5 . This indicates that if we know the card is red, the probability of it being a heart (which is one of the two red suits) is 0.5.
- **Independence Check**: The events A (drawing a heart) and B (drawing a red card) are not independent, as confirmed by the independence check. The occurrence of B (drawing a red card) increases the probability of A (drawing a heart), making $P(A \cap B) \neq P(A)P(B)$.

This demonstrates that knowing a card is red (event B) directly influences the likelihood of the card being a heart (event A), hence they are not independent.

Next, let's illustrate Bayes' Theorem with an example.

In the context of using Bayes' Theorem with a deck of cards:

- **Probability of Drawing a Face Card (C)**: $P(C)$ = $\frac{12}{52}$ \approx 0.231), which represents the probability of drawing a Jack, Queen, or King from any suit.
- **Conditional Probability of Drawing a Face Card Given It's a Heart (P(C|A))**: This calculation gives us (P(C|A) $\approx$ 0.231 , indicating that the probability of a card being a face card, given that it's a heart, is the same as the overall probability of drawing a face card from the deck.

```
# Bayes' Theorem

# Continuing with the card example, let's introduce a new event:
# C = Drawing a card that is a face card (Jack, Queen, King of any suit)

# Total face cards in a deck
total_face_cards = 12  # 3 face cards * 4 suits
```

```
# Probability of drawing a face card (C)
P_C = total_face_cards / total_cards

# Probability of drawing a heart that is also a face card (A ∩ C)
P_A_intersect_C = 3 / total_cards  # There are 3 hearts that are face cards

# Probability of drawing a face card given it's a heart
# Using Bayes' Theorem: P(C|A) = P(A ∩ C) / P(A)
P_C_given_A = P_A_intersect_C / P_A

# Let's also calculate P(C) for general knowledge
P_C, P_C_given_A
```

3. Combinatorics and Probability

3.1. Basic Counting Principles

The basic counting principles, including the **Rule of Sum** and the **Rule of Product**, are foundational to understanding how to count the number of ways events can occur, which is crucial in probability and combinatorics.

- **Rule of Sum (Addition Principle)**: If event A can occur in m ways and event B can occur in n ways, and if events A and B cannot occur at the same time, then there are $m + n$ ways for either event A or B to occur.
- **Rule of Product (Multiplication Principle)**: If an event A can occur in (m) ways and for each way that A occurs, a second event B can occur in n ways, then the sequence of two events can occur in $m \times n$ ways.

Example: Suppose you have 2 pairs of jeans (blue and black) and 3 shirts (white, gray, and red). How many different outfits can you create?

Using the Rule of Product, the total number of outfits is 2 (jeans) $\times 3$ (shirts) = 6 outfits.

Imagine you're planning a week-long trip and have to decide on two main things: the destination and the type of accommodation.

1. **Destination**: You can choose from 5 different cities.
2. **Accommodation**: You have 2 types of accommodation to choose from in any city.

Objective:

Calculate the total number of unique trip plans you can make based on these choices, using the Rule of Product (Multiplication Principle).

Solution:

Step 1: Identify the Number of Choices

- **Number of Destinations (D)**: 5 cities
- **Number of Accommodation Types (A)**: 2 types

Step 2: Apply the Rule of Product

The Rule of Product states that if you have (m) ways of making one choice and n ways of making a second choice, then there are $m \times n$ ways of making both choices.

Here, the first choice is the city destination with 5 options, and the second choice is the type of accommodation with 2 options. Applying the Rule of Product:

$$\text{Total Trip Plans} = D \times A$$

$$\text{Total Trip Plans} = 5 \times 2$$

Step 3: Calculate the Total Number of Trip Plans

$$\text{Total Trip Plans} = 10$$

Based on the Basic Counting Principles, you can create 10 unique trip plans, considering the combination of 5 possible destinations and 2 types of accommodations. This simple calculation illustrates how the Multiplication Principle is applied to determine the total number of outcomes for a series of decisions, providing a foundational understanding of combinatorial calculations in probability theory.

3.2. Permutations and Combinations Example

3.2.1 Permutations

Permutations are used when the order of selection matters. The number of permutations of (n) objects taken (r) at a time is given by:

$$P(n, r) = \frac{n!}{(n-r)!}$$

Example: Permutations

Question: In a race with 8 runners, how many different ways can the first three places be awarded?

Solution:

- Here, (n = 8) (total runners), and r = 3 (positions to be awarded).
- We want to find $P(8, 3)$.

$$P(8,3) = \frac{8!}{(8-3)!} = \frac{8!}{5!}$$
$$P(8,3) = \frac{8 \times 7 \times 6 \times 5!}{5!}$$
$$P(8,3) = 8 \times 7 \times 6$$
$$P(8,3) = 336$$

So, there are 336 different ways the first three places in the race can be awarded.

3.2.2 Combinations

Combinations are used when the order of selection does not matter. The number of combinations of (n) objects taken r at a time is given by:

$$C(n,r) = \frac{n!}{r!(n-r)!}$$

Example: Combinations

Question: From a class of 30 students, how many ways can a committee of 4 students be formed?

Solution:

- Here, (n = 30) (total students), and (r = 4) (students to be selected for the committee).
- We want to find $C(30,4)$

$$C(30,4) = \frac{30!}{4!(30-4)!} = \frac{30!}{4! \cdot 26!}$$
$$C(30,4) = \frac{30 \times 29 \times 28 \times 27 \times 26!}{4 \times 3 \times 2 \times 1 \times 26!}$$
$$C(30,4) = \frac{30 \times 29 \times 28 \times 27}{4 \times 3 \times 2 \times 1}$$
$$C(30,4) = \frac{657720}{24}$$
$$C(30,4) = 27405$$

Therefore, there are 27,405 different ways to form a committee of 4 students from a class of 30.

The following examples illustrate the mathematical procedures that are required to compute permutations and combinations. They also illustrate the many applications of these mathematical operations, which vary according on whether or not the order of selection counts.

3.3. Applications in Probability

Applications of combinatorics in probability involve using permutations and combinations to calculate the likelihood of various outcomes. These principles allow us to determine the total number of possible outcomes and the number of ways specific events can occur, which are crucial for calculating probabilities.

Example Problem: Drawing Cards from a Deck

Question: What is the probability of drawing a flush (five cards of the same suit) in a five-card poker hand from a standard 52-card deck?

Solution:

Step 1: Calculate Total Number of Possible Hands

The total number of ways to draw 5 cards from a 52-card deck, regardless of order, is given by combinations:

$$C(52,5) = \frac{52!}{5!(52-5)!}$$
$$C(52,5) = \frac{52!}{5! \cdot 47!}$$
$$C(52,5) = \frac{52 \times 51 \times 50 \times 49 \times 48}{5 \times 4 \times 3 \times 2 \times 1}$$
$C(52,5) = 2,598,960$ (Total possible hands)

Step 2: Calculate Number of Ways to Draw a Flush

To draw a flush, all five cards must be of the same suit. There are 4 suits, and for each suit, there are 13 cards. The number of ways to draw 5 cards from a single suit is:

$$C(13,5) = \frac{13!}{5!(13-5)!}$$
$$C(13,5) = \frac{13!}{5! \cdot 8!}$$
$$C(13,5) = \frac{13 \times 12 \times 11 \times 10 \times 9}{5 \times 4 \times 3 \times 2 \times 1}$$
$C(13,5) = 1,287$ (Ways to draw a flush from one suit)

Since there are 4 suits, the total number of flush hands is:

$4 \times C(13,5) = 4 \times 1,287 = 5,148$(Total flush hands)

Step 3: Calculate the Probability of Drawing a Flush

The probability of drawing a flush is the ratio of the number of flush hands to the total number of possible hands:

$$P(\text{Flush}) = \frac{\text{Number of flush hands}}{\text{Total possible hands}}$$
$$P(\text{Flush}) = \frac{5,148}{2,598,960}$$
$$P(\text{Flush}) \approx 0.00198$$

The probability of drawing a flush in a five-card poker hand from a standard 52-card deck is approximately 0.198%, demonstrating a specific application of combinatorics in calculating probabilities for complex card games. This approach highlights the importance of understanding both the total number of outcomes and the number of favorable outcomes to accurately determine probabilities.

Continue ....