Probability Theory - Lecture Notes

Part 1: Foundations of Probability

What is Probability?

Probability quantifies the likelihood of an event occurring after many repeated trials. For an event A in sample space S:

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

A **random experiment** is a process by which we observe something uncertain. For example, rolling a die is a random experiment because we don't know the result beforehand, but we know all possible outcomes.

Basic Examples

Example 1: Fair Six-Sided Die

- Sample space: S = $\{1, 2, 3, 4, 5, 6\}$
- P(rolling a 4) = 1/6 ≈ 0.167
- P(rolling an even number) = 3/6 = 1/2
- P(rolling a number > 4) = 2/6 = 1/3

Example 2: Coin Flip

- Sample space: S = $\{\text{Heads}, \text{Tails}\}$
- P(Heads) = 1/2 = 0.5
- P(not Heads) = P(Tails) = 1/2

Sample Spaces

The **sample space** (S) is the set of all possible outcomes of a random experiment.

Types of Sample Spaces:

- Rolling a die: S = {1, 2, 3, 4, 5, 6}
- Two coin flips: S = {HH, HT, TH, TT}
- Drawing two cards without replacement: S has $52 \times 51 = 2{,}652$ outcomes

Events

An **event** is a subset of the sample space.

Types of Events:

1. **Simple Event**: Contains exactly one outcome
   - Example: Rolling a 3 on a die = {3}
2. **Compound Event**: Contains multiple outcomes
   - Example: Rolling an even number = {2, 4, 6}
3. **Certain Event**: The entire sample space (S)
   - $P(S) = 1$
4. **Impossible Event**: The empty set (∅)
   - $P(\varnothing) = 0$

Example: Rolling Two Dice

- Event A: "Sum equals 7" = {(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)}
  - $P(A) = 6/36 = 1/6$
- Event B: "Both dice show same number" = {(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)}

    - $P(B) = 6/36 = 1/6$

- Event C: "Sum is even" = 18 outcomes out of 36
    - $P(C) = 18/36 = 1/2$

Standard Deck of Cards

A standard deck has 52 cards (4 suits × 13 ranks).

Key probabilities:

- P(Ace) = 4/52 = 1/13 ≈ 0.077
- P(Ace or King) = 8/52 = 2/13
- P(Heart) = 13/52 = 1/4 = 0.25
- P(Red card) = 26/52 = 1/2 = 0.5
- P(Face card) = 12/52 = 3/13 ≈ 0.231

Fundamental Properties of Probability

1. **Non-negativity**: For any event A, $P(A) \geq 0$
    - Probability cannot be negative
2. **Certainty**: $P(S) = 1$
    - The sample space covers all possible outcomes
3. **Additivity**: If $A_1, A_2, \ldots$ are disjoint events that cover the sample space, then $\sum P(A_i) = 1$
    - This implies that for any single event, $P(A) \leq 1$
4. **Complement Rule**: $P(A') = 1 - P(A)$
    - Where A' is the complement of A (all outcomes not in A)

Theoretical vs. Experimental Probability

Theoretical Probability:

- Based on mathematical reasoning and known outcomes
- Assumes ideal conditions (fair coin, balanced die)
- Formula: P(A) = favorable outcomes / total possible outcomes

Experimental Probability:

- Based on actual experiments or observations
- Relative frequency of event occurrence
- Formula: P(A) ≈ number of times A occurred / total number of trials

Law of Large Numbers: As the number of trials increases, experimental probability converges to theoretical probability.

Example: Coin Flip Experiment

Number of Flips	Heads	Experimental P(Heads)	Theoretical P(Heads)
10	4	0.40	0.50
100	47	0.47	0.50
1,000	512	0.512	0.50
10,000	4,998	0.4998	0.50

Key Insight: More trials → experimental probability gets closer to theoretical probability

---

Part 2: Set Operations and Probability

Basic Set Operations

Union ($A \cup B$): "A or B" - outcomes in A, B, or both

Intersection ($A \cap B$): "A and B" - outcomes in both A and B

Complement (A'): "not A" - outcomes not in A

Difference ($A - B$): Outcomes in A but not in B

Set Notation

- **$A \subseteq B$**: A is a subset of B
- **$A \cap B = \emptyset$**: A and B are disjoint (mutually exclusive)
- **$|A|$**: Cardinality (number of elements) in A

Example with Sets:

- $X = \{3, 12, 5, 13\}$
- $Y = \{14, 15, 16, 3\}$
- $X \cap Y = \{3\}$
- $X \cup Y = \{3, 5, 12, 13, 14, 15, 16\}$ — note that 3 appears only once in the union

Example with Probability:

Rolling a die:

- A = {rolling an even number} = $\{2, 4, 6\}$
- B = {rolling a number ≥ 5} = $\{5, 6\}$

- $A \cup B = \{2, 4, 5, 6\}$
- $A \cap B = \{6\}$

Therefore:

- P(A) = 3/6 = 1/2
- P(B) = 2/6 = 1/3
- $P(A \cap B)$ = 1/6
- $P(A \cup B)$ = 4/6 = 2/3

Addition Rule of Probability (General Form)

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Why subtract $P(A \cap B)$? To avoid double-counting the intersection.

Example: Drawing a Card

- A = {drawing a Jack}, P(A) = 4/52
- B = {drawing a Heart}, P(B) = 13/52
- $A \cap B$ = {Jack of Hearts}, $P(A \cap B)$ = 1/52
- $P(A \cup B)$ = 4/52 + 13/52 - 1/52 = 16/52 = 4/13

Addition Rule for Mutually Exclusive Events

Mutually exclusive events are events that CANNOT happen together (e.g., rolling a 3 means a 5 cannot be rolled on the same roll).

If $A \cap B = \varnothing$, then:

$$P(A \cup B) = P(A) + P(B)$$

Example: Mutually Exclusive Events

P(rolling a 2 or a 5) = ?

- These are mutually exclusive (can't roll both on one die)
- P(2 or 5) = P(2) + P(5) = 1/6 + 1/6 = 2/6 = 1/3

More Addition Rule Examples

Example 1: Student Activities

In a class of 100 students:

- 60 play sports
- 45 play music
- 25 do both

P(student plays sports OR music) = ?

$P(S \cup M) = 60/100 + 45/100 - 25/100 = 80/100 = 0.80$

Exercise: Rolling a Die

P(rolling a 2 or an even number) = ?

- $A = \{2\}$, $P(A) = 1/6$
- $B = \{2, 4, 6\}$, $P(B) = 3/6$
- Note: $A \subseteq B$, so $A \cap B = A$
- $P(A \cup B) = P(B) = 3/6 = 1/2$

Part 3: Compound Probability

Compound probability refers to the probability of a combination of two or more events.

Independence

Definition: Events A and B are **independent** if the occurrence of one does not affect the probability of the other. This is NOT to be confused with mutually exclusive events (which cannot occur together and are by definition dependent).

Mathematical Definition: Events A and B are independent if:

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

Equivalently: $P(A|B) = P(A)$

Examples of Independent Events:

- Flipping a coin twice: $P(H_1)$ and $P(H_2)$ are independent
- Rolling two dice: outcome of first die doesn't affect second die

Multiplication Rule for Independent Events

$$P(A \text{ and } B) = P(A) \times P(B)$$

Example Calculation:

P(two heads in two flips) = $P(H_1) \times P(H_2) = (1/2) \times (1/2) = 1/4$

Extended to n independent events:

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \times P(A_2) \times \cdots \times P(A_n)$$

"At Least One" Problems

For "at least one" problems, it's often easier to use the complement.

Example: Three Coin Flips

Sample space: {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}

P(at least 1 head) = 1 - P(no heads) = 1 - P(TTT) = 1 - 1/8 = 7/8

Example: Rolling a Die Twice

P(at least one 6)?

- P(no sixes) = $(5/6)^2$ = 25/36
- P(at least one 6) = 1 - 25/36 = 11/36 ≈ 0.306

Dependence

Definition: Events A and B are **dependent** if the occurrence of one affects the probability of the other.

Example: Drawing 2 Aces from a Deck Without Replacement

- P(1st Ace) = 4/52
- P(2nd Ace | 1st Ace) = 3/51
- P(both Aces) = (4/52) × (3/51) = 12/2652 = 1/221 ≈ 0.0045

Example: Selecting Without Replacement

A box contains 5 red balls and 3 blue balls. Draw 2 balls without replacement.

P(both red) = ?

- P(1st red) = 5/8
- P(2nd red | 1st red) = 4/7
- P(both red) = (5/8) × (4/7) = 20/56 = 5/14 ≈ 0.357

The General Multiplication Rule

For any two events:

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

where P(B|A) is the conditional probability of B given A.

The vertical bar "|" means "given," so P(B|A) reads as "the probability that B occurs given that A has occurred."

This formula tells us we can multiply the probabilities of two events, but we need to account for the first event when considering the probability of the second event.

Special case: If events are independent, then P(B|A) = P(B), which gives us back the simple multiplication rule: P(A ∩ B) = P(A) × P(B).

Conditional Probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{where } P(B) > 0$$

Intuition:

- We're restricting our sample space to only those outcomes where B occurred

- Then we find the proportion of those outcomes where A also occurs

Example: Drawing Cards

Given that you drew a face card, what's the probability it's a King?

- B = {face card}, |B| = 12, P(B) = 12/52
- A = {King}, A ∩ B = {King}, P(A ∩ B) = 4/52
- P(A|B) = (4/52) / (12/52) = 4/12 = 1/3

Alternative approach: Think in the reduced sample space

- Given it's a face card, there are 12 possibilities
- 4 of them are Kings
- P(King | face card) = 4/12 = 1/3

Example: Student Survey

Survey of 100 students about exercise and diet:

	Healthy Diet	**Unhealthy Diet**	**Total**
Exercises	45	15	60
No Exercise	20	20	40
Total	65	35	100

Questions:

1. P(Exercises) = 60/100 = 0.6
2. P(Healthy Diet | Exercises) = 45/60 = 0.75
3. P(Exercises | Healthy Diet) = 45/65 ≈ 0.692

Important Note: $P(A|B) \neq P(B|A)$ in general! Compare questions 2 and 3 above.

However: $P(A \cap B) = P(B \cap A)$ (intersection is commutative)

General Multiplication Rule Extended

$$P(A \cap B \cap C) = P(A) \times P(B|A) \times P(C|A \cap B)$$

Example: Drawing 3 Cards Without Replacement, All Red

P(all 3 red) = $P(R_1) \times P(R_2|R_1) \times P(R_3|R_1 \cap R_2)$
= $(26/52) \times (25/51) \times (24/50)$

Law of Total Probability

Step 1: Using a Venn diagram, we can see that:

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

Step 2: Using the multiplication rule:

$$P(A) = P(A|B)P(B) + P(A|B')P(B')$$

General Rule: If $B_1, B_2, ..., B_n$ partition the sample space (disjoint and cover everything), then:

$$P(A) = \sum_i P(A|B_i)P(B_i)$$

Bayes' Rule

Bayes' Rule is one of the most powerful results in probability theory, derived from the multiplication rule and the law of total probability.

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

Components:

- **P(A)**: Prior probability (before considering B)
- **P(B|A)**: Likelihood of B given A
- **P(A|B)**: Posterior probability (after considering B)

Example: False Positive Paradox

A disease affects 1 in 10,000 people. A test for this disease has:

- 2% false positive rate (positive result when person is healthy)
- 1% false negative rate (negative result when person is sick)

A random person tests positive. What is the probability they actually have the disease?

Solution:

Let D = person has disease, T = test is positive

Given:

- P(D) = 1/10,000 = 0.0001
- $P(T|D^c) = 0.02$
- $P(T^c|D) = 0.01$

Find: P(D|T)

Using Bayes' rule:

$$P(D|T) = \frac{P(T|D) \times P(D)}{P(T|D) \times P(D) + P(T|D^c) \times P(D^c)}$$

$$P(D|T) = \frac{(0.99) \times (0.0001)}{(0.99) \times (0.0001) + (0.02) \times (0.9999)}$$

$$P(D|T) \approx 0.0049 = 0.49\%$$

Answer: There is less than 0.5% chance the person has the disease despite testing positive.

This counterintuitive result occurs because the disease is so rare (low prior probability) that even with a fairly accurate test, most positive results are false positives.

Part 4: Counting Principles

Many probability problems come down to counting: P(A) = |A| / |S|

Multiplication Principle

Multiplication Principle: Suppose we perform r experiments such that the kth experiment has n_k possible outcomes, for k = 1, 2, ..., r. Then there are a total of $n_1 \times n_2 \times n_3 \times ... \times n_r$ possible outcomes for the sequence of r experiments.

Example: Computer Purchase

Suppose you want to purchase a tablet computer. You can choose:

- Screen: large or small (2 options)

- Storage: 64 GB, 128 GB, or 256 GB (3 options)
- Cover: black or white (2 options)

Total options = $2 \times 3 \times 2 = 12$

Example: Outfits

With 5 shirts, 3 pants, and 2 pairs of shoes:
Total outfits = $5 \times 3 \times 2 = 30$

Factorial

Definition: n! (n factorial) is the product of all positive integers from 1 to n.

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1$$

Examples:

- $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$
- $3! = 3 \times 2 \times 1 = 6$
- $0! = 1$ (by definition)

Useful property: $n! = n \times (n-1)!$

Permutations

Permutations count arrangements where **order matters** and we select **without replacement**.

Example: How many ways can we seat 5 people (A, B, C, D, E) in 5 chairs?

- First chair: 5 choices
- Second chair: 4 choices

- Third chair: 3 choices
- Fourth chair: 2 choices
- Fifth chair: 1 choice
- Total: 5! = 5 × 4 × 3 × 2 × 1 = 120

Example: How many ways can we seat 5 people in only 3 chairs?

- First chair: 5 choices
- Second chair: 4 choices
- Third chair: 3 choices
- Total: 5 × 4 × 3 = 60

This equals 5!/(5-3)! = 5!/2! = 120/2 = 60

General Formula for Permutations:

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

This gives the number of ways to arrange r objects selected from n objects.

Example: How many 3-letter "words" can we create from the 26-letter alphabet if:

- Letters can repeat: 26 × 26 × 26 = 17,576
- Letters must be different: 26 × 25 × 24 = 15,600 = 26!/23!

Combinations

Combinations count selections where **order does NOT matter** and we select **without replacement**.

Example: How many ways can we pick 3 people from 6 people (A, B, C, D, E, F)?

Using permutations, there are 6P3 = 6 × 5 × 4 = 120 ways to arrange 3 people in 3 positions.

But {A, B, C}, {B, A, C}, {C, A, B}, etc. are all the same group of 3 people.

Each group of 3 has been counted 3! = 6 times.

So the number of combinations = 120/6 = 20

General Formula for Combinations:

$$ {}_nC_r = \binom{n}{r} = \frac{n!}{(n-r)! \cdot r!} $$

This gives the number of ways to choose r objects from n objects (where order doesn't matter).

Key Difference:

- **Permutations**: Order matters (ABC ≠ BAC)
- **Combinations**: Order doesn't matter (ABC = BAC = CAB)

To convert from permutations to combinations: divide by r! to remove the ordering of the r selected items.

Probability and Combinatorics

Example: Coin Flips

What is the probability of getting exactly 3 heads if we flip a coin 8 times?

P(\text{exactly 3 heads}) = \frac{\text{# of ways to get 3 heads}}{\text{total outcomes}}

- Denominator: 2^8 = 256 (each flip has 2 outcomes)
- Numerator: ${}_8C_3$ = 8!/(5!×3!) = 56 (choose which 3 of the 8 flips are heads)

$$P(\text{exactly 3 heads}) = \frac{56}{256} = \frac{7}{32} \approx 0.219$$

Example: Choosing Officers

A club of 9 people wants to choose a board of three officers: President, Vice President, and Secretary. Assuming the officers are chosen at random, what is the probability that the officers are Marsha for President, Sabita for Vice President, and Robert for Secretary?

P(\text{specific arrangement}) = \frac{1}{\text{total # of possibilities}}

Total number of ways to choose 3 officers from 9 people (order matters):

$${}_9P_3 = \frac{9!}{6!} = 9 \times 8 \times 7 = 504$$

There's only 1 way to get the specific arrangement (Marsha-Sabita-Robert).

$$P(\text{Marsha, Sabita, Robert}) = \frac{1}{504} \approx 0.002$$

Summary of Key Formulas

Basic Probability

- P(A') = 1 - P(A)
- P(A ∪ B) = P(A) + P(B) - P(A ∩ B)
- If A and B are mutually exclusive: P(A ∪ B) = P(A) + P(B)

Independent Events

- $P(A \cap B) = P(A) \times P(B)$
- $P(A|B) = P(A)$

Dependent Events (General)

- $P(A \cap B) = P(A) \times P(B|A)$
- $P(A|B) = P(A \cap B) / P(B)$

Bayes' Rule

- $P(A|B) = [P(B|A) \times P(A)] / P(B)$

Counting

- Permutations: ${}_nP_r = n!/(n-r)!$
- Combinations: ${}_nC_r = n!/[(n-r)! \times r!]$
- Multiplication Principle: $n_1 \times n_2 \times ... \times n_r$

Important Distinctions to Remember

1. **Mutually Exclusive vs. Independent**
    - Mutually exclusive: Cannot happen together, $P(A \cap B) = 0$
    - Independent: One doesn't affect the other, $P(A \cap B) = P(A) \times P(B)$
    - These are DIFFERENT concepts!
2. **P(A|B) vs. P(B|A)**
    - These are generally NOT equal

- P(A|B) = probability of A given B happened
- P(B|A) = probability of B given A happened

3. **Permutations vs. Combinations**
   - Permutations: Order matters
   - Combinations: Order doesn't matter
4. **With Replacement vs. Without Replacement**
   - With replacement: Probabilities stay the same for each draw
   - Without replacement: Probabilities change after each draw (dependent events)