STAT 24400 Lecture 1 Probability

Yibi Huang Department of Statistics University of Chicago

Outline

Coverage: Section 1.1-1.4 of Rice's Book

- 1.2 Sample Space and Events
 - intersection, union, complement, empty set, subset, disjoint
 - Venn diagrams
 - De Morgan's Laws
- 1.3 Probability Measure
- 1.4 Counting Methods
 - Permutation
 - Combination

Warmup Puzzle

A fair coin is flipped repeatedly until HH or TH comes up for the first time.

- Player A wins if HH comes up first.
- Player B wins if TH comes up first.

What are the odds of winning for each player?

Warmup Puzzle

A fair coin is flipped repeatedly until HH or TH comes up for the first time.

- Player A wins if HH comes up first.
- Player B wins if TH comes up first.

What are the odds of winning for each player?

Answer:

- Player A wins only if the first two coins are HH (25% chance).
- Otherwise, player B wins (75% chance).

Sample Space and Events

Sample Spaces & Events (1)

When we perform an experiment or observe a random process,

- the sample space Ω is the set of all possible outcomes.
- an event is any subset A of the sample space Ω
- $\mathbf{Ex} \ \mathbf{1}$. Rolling a die once.
 - The sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$
 - Events:
 - ▶ {an ace} = {1}
 - {an even number} = $\{2, 4, 6\}$
 - $\{5 \text{ or more}\} = \{5, 6\}$

Sample Spaces & Events (2)

Ex 2 Tossing a coin repeatedly until heads comes up

- The sample space is $\Omega = \{H, TH, TTH, TTTH, \ldots\} = \{1, 2, 3, 4, \ldots\}$ • Events:
 - $A = \{ \text{getting heads in the first try} \} = \{ H \} = \{ 1 \}$
 - ▶ B = {no heads in the first 5 tosses} = {6, 7, ...} ={all integers > 5}
- Ex 3 How many inches of rain today?
 - The sample space is Ω = [0,∞)
 Events:

•
$$A = \{$$
It rains today $\} = (0, \infty)$

• $B = \{ \text{over 3 inches} \} = (3, \infty)$

There may be **multiple sensible sample spaces** for one random phenomenon.

Example. Drawing two balls at random without replacement from the box below



- \blacktriangleright When only interested in the number of reds in the 2 draws. . . \blacktriangleright $\Omega_1=\{0,1,2\}$
- When interested in the color of two balls,
 - $\Omega_2 = \{RR, RG, GR, GG\}$ if the order matters
 - $\Omega_3 = \{RR, RG, GG\} = \{2, 1, 0\} = \Omega_1$ if the order doesn't matter

When interested in the numbers shown on two balls...

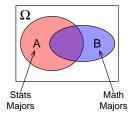
$$\Omega_{4} = \underbrace{ \left\{ \begin{array}{ccccc} (1,2) & (1,3) & (1,4) & (1,5) \\ (2,1) & (2,3) & (2,4) & (2,5) \\ (3,1) & (3,2) & (3,4) & (3,5) \\ (4,1) & (4,2) & (4,3) & (4,5) \\ (5,1) & (5,2) & (5,3) & (5,4) \end{array} \right\}}_{\text{if the order doesn't matter}} \text{ or } \Omega_{5} = \underbrace{ \left\{ \begin{array}{c} (2,1) \\ (3,1) & (3,2) \\ (4,1) & (4,2) & (4,3) \\ (5,1) & (5,2) & (5,3) & (5,4) \end{array} \right\}}_{\text{if the order doesn't matter}}$$

if the order matters

Empty Set, Intersection, Union, Complement

Consider the experiment of choosing one student from the class, at random

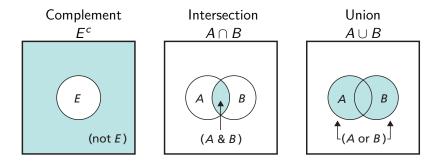
- Ω = set of all students in the class
- A = all stat majors
- \blacktriangleright B = all math majors



Notation	Definition	Example/Interpretation
Ø	the empty set	$B = \varnothing \Leftrightarrow$ there are no math
		majors in the class
$A \cap B$	intersection of A and B	stats&math double majors
	(both A and B occur)	
$A \cup B$	union of A and B	the set of students who are
	(A and/or B occur)	majoring in stats and/or math
Bc	complement of B	the set of all students that are
	(<i>B</i> does not occur)	not math majors

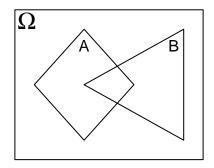
Venn Diagrams

Complements, intersections, unions of events can be represented visually using *Venn diagrams*:



Venn Diagram Exercise 1

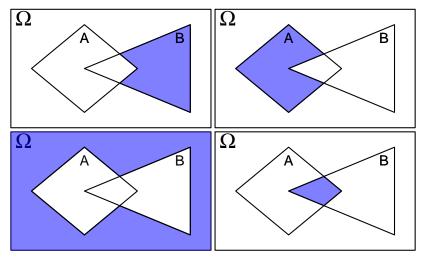
Sketch the region that represents $A \cap B^c$ in the Venn Diagram below



Exercise 2 : Please match the events

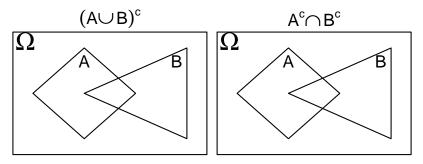
 $A \cap B$, $A \cap B^c$, $A^c \cap B$, $A^c \cap B^c$

with the corresponding Venn diagrams.



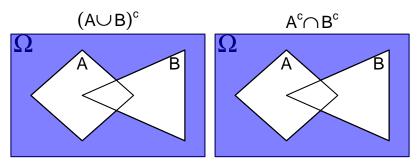
Venn Diagram Exercise 3

Sketch the regions that represent $(A \cup B)^c$ and $A^c \cap B^c$ in the Venn Diagrams below



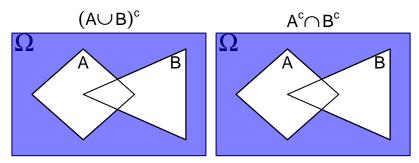
Venn Diagram Exercise 3 (Answers)

Sketch the regions that represent $(A\cup B)^c$ and $A^c\cap B^c$ in the Venn Diagrams below



Venn Diagram Exercise 3 (Answers)

Sketch the regions that represent $(A\cup B)^c$ and $A^c\cap B^c$ in the Venn Diagrams below



This proves one of the De Morgan's Laws

$$(A\cup B)^c = A^c \cap B^c$$

De Morgan's Laws

Exercise: Convince yourself the validity of the second De Morgan's Law using Venn Diagrams.

De Morgan's Laws Exercises

Randomly select an individual from the population and ask whether he has Visa Card or a MasterCard.

- 1. Which of the following is the complement of the randomly selected individual has both a Visa Card and a MasterCard?
 - A. he has either a Visa Card or a MasterCard
 - B. he has neither Visa Card nor a MasterCard
- 2. Which of the following is the complement of the randomly selected individual has neither a Visa Card nor a MasterCard?
 - A. he has both a Visa Card and a MasterCard
 - B. he has a Visa Card or a MasterCard

De Morgan's Laws For 3 or More Events

$$(A_1 \cup A_2 \cup A_3 \cup \cdots)^c = A_1^c \cap A_2^c \cap A_3^c \cap \cdots$$
$$(A_1 \cap A_2 \cap A_3 \cap \cdots)^c = A_1^c \cup A_2^c \cup A_3^c \cup \cdots$$

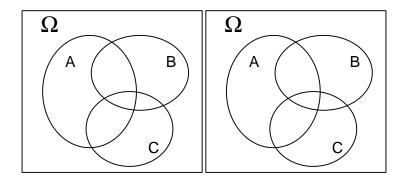
Exercises: Tossing a fair coin 5 times,

- which of the following is the complement of {the coin lands heads at least once}?
 - A. the coin lands never lands heads (all tails in 5 tosses)
 - B. the coin lands tails at least once
- 2. which of the following is the complement of {the coin lands heads in all 5 tosses}?
 - A. the coin lands never lands heads (all tails in 5 tosses)
 - B. the coin lands tails at least once

Distributive Laws

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

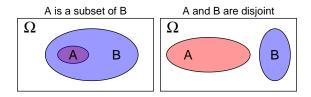
Exercise: Convince yourself the validity of the Distributive Laws using Venn Diagrams.



Set Notation — Subset & Disjoint

Recall the example of randomly choosing one student from STAT 244, $\Omega = \{\text{all students in class}\}, A = \{\text{all stat majors}\}, B = \{\text{all math majors}\}.$

Notation	Definition	Example/interpretation
$\overline{A \subset B}$ or	A is a subset of B	all stats majors in the class
$A \subseteq B$	(and possibly $A = B$)	are also math majors
$A \subsetneq B$	A is a strict subset of B	all stats majors are also
	(A = B not allowed)	math majors, but not all
		math majors are stats majors
A and B	$A \cap B = \emptyset$	there are no stats/math
are disjoint		double majors in the class



Probability Measures

Probability Measures

A *probability measure* is the function $P(\cdot)$ that specifies a probability for each subset A of Ω that satisfies the following 3 axioms.

- AXIOM 1 $P(\Omega) = 1$, where $\Omega =$ sample space.
- AXIOM 2 For any event $A \subset \Omega$, $P(A) \ge 0$.
- AXIOM 3 If A₁ and A₂ are disjoint, then

$$\mathrm{P}(A_1\cup A_2)=\mathrm{P}(A_1)+\mathrm{P}(A_2).$$

More generally, if $A_1, A_2, A_3, ...$ is a *countably* infinite collection of disjoint events, then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$$

Probability Rules Derived From Probability Axioms

1. Complement Rule: $P(A^c) = 1 - P(A)$

2. $P(\emptyset) = 0$

- 3. If $A \subset B$ then $P(A) \leq P(B)$
- 4. General Addition Rule: $P(A \cup B) = P(A) + P(B) P(A \cap B)$

Probability Rules Derived From Probability Axioms

1. Complement Rule: $P(A^c) = 1 - P(A)$

2. $P(\emptyset) = 0$

- 3. If $A \subset B$ then $P(A) \leq P(B)$
- 4. General Addition Rule: $P(A \cup B) = P(A) + P(B) P(A \cap B)$

These properties can be proved from the axioms.

Example—proof of the Complement Rule:

$$\underbrace{1 = P(\Omega)}_{\text{by Axiom 1}} = \underbrace{P(A \cup A^c) = P(A) + P(A^c)}_{\text{by Axiom 3}}$$

If a sample space Ω consists of a finite number of outcomes and each outcome is equally likely to occur, then the probability for an event $A \subset \Omega$ is

$$P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{\text{number of outcomes in } A}{\text{number of outcomes in } \Omega}.$$

If a sample space Ω consists of a finite number of outcomes and each outcome is equally likely to occur, then the probability for an event $A \subset \Omega$ is

$$P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{\text{number of outcomes in } A}{\text{number of outcomes in } \Omega}.$$

If a sample space Ω consists of a finite number of outcomes and each outcome is equally likely to occur, then the probability for an event $A \subset \Omega$ is

$$P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{\text{number of outcomes in } A}{\text{number of outcomes in } \Omega}.$$

If a sample space Ω consists of a finite number of outcomes and each outcome is equally likely to occur, then the probability for an event $A \subset \Omega$ is

$$P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{\text{number of outcomes in } A}{\text{number of outcomes in } \Omega}.$$

Caution: Outcomes Are NOT Always Equally Likely — Don't Assume!

Example. Drawing two balls at random without replacement from the box below



Which of the following is correct calculation for the probability of $A = \{\text{getting 2 Reds}\}$?

1. Sample space
$$\Omega_{=}\{0, 1, 2\}, A = \{2\} \Rightarrow P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{1}{3}$$

Caution: Outcomes Are NOT Always Equally Likely — Don't Assume!

Example. Drawing two balls at random without replacement from the box below



Which of the following is correct calculation for the probability of $A = \{ getting 2 Reds \}$?

1. Sample space $\Omega_{=}\{0, 1, 2\}$, $A = \{2\} \Rightarrow P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{1}{3}$ 2. Sample space $\Omega_{=}\{RR, RG, GR, GG\}$, $A = \{RR\}$.

$$\Rightarrow P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{1}{4}$$

Caution: Outcomes Are NOT Always Equally Likely — Don't Assume!

Example. Drawing two balls at random without replacement from the box below



Which of the following is correct calculation for the probability of $A = \{\text{getting 2 Reds}\}$?

1. Sample space
$$\Omega_{=}\{0, 1, 2\}, A = \{2\} \Rightarrow P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{1}{3}$$

2. Sample space $\Omega_{=}\{RR, RG, GR, GG\}, A = \{RR\}.$
 $\Rightarrow P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{1}{4}$
3.
 $\Omega = \begin{cases} (1, 2) (1, 3) (1, 4) (1, 5) \\ (2, 1) & (2, 3) (2, 4) (2, 5) \\ (3, 1) (3, 2) & (3, 4) (3, 5) \\ (4, 1) (4, 2) (4, 3) & (4, 5) \\ (5, 1) (5, 2) (5, 3) (5, 4) \end{cases}, A = \{(1, 2), (2, 1)\}, A = \{(1, 2), (2, 1)\}, B = \{(1, 2), (2, 1)$

Counting Methods

Counting Methods

To compute probabilities using the formula

$$\mathrm{P}(A) = \frac{\#(A)}{\#(\Omega)}$$

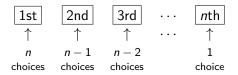
for more complex situations, systematic **counting methods** are required.

Multiplication Principle

If one experiment has m outcomes and another experiment has n outcomes, there are $m \times n$ possible outcomes for the two experiments.

Factorial

How many ways are there to arrange n distinct items in a row?



As a whole, there are

$$n \times (n-1) \times (n-2) \times \ldots \times 1 = n!$$
 ways.

n! is read "*n factorial*". Note that 0! is defined as 1.

Counting Methods — Permutation

How many ways are there to choose r items out of n distinct items, while keeping track of the order of selection is

replacement

Ex: Choose two from {Apple, Orange Pear}

 \blacktriangleright 3 \times 3 = 9 ways if drawing with replacement, and order matters \blacktriangleright 3 \times 2 = 6 ways if drawing without replacement, and order matters

Counting — Combinations

How many ways are there to choose r items out of n distinct items, regardless of order?

Counting — Combinations

How many ways are there to choose r items out of n distinct items, regardless of order?Let N be the answer. We can find N as follows.

- Consider the number of ways to choose r persons out of n, and seat them in a row of r chairs.
- We can generate all such ways by
 - first selecting r persons out of n, regardless of order. There are N ways to do so.
 - Then arranging those r persons on the r chairs. There are r! ways to do that.
- By this argument there are N × r! ways to seat n persons in a row of r chairs.
- But from the previous slide, we know that there are $\frac{n!}{(n-r)!}$ such ways. Thus

$$N \times r! = \frac{n!}{(n-r)!} \Longrightarrow N = \frac{n!}{r!(n-r)!}$$

Combinations = Binomial Coefficients

The number of ways to choose r out of n is denoted as

$$\binom{n}{r} =_n C_k = C_k^n = \frac{n!}{r! (n-r)!},$$

and the 3 notations all read as "n choose r".

They are also called the *Binomial coefficients* as they are the coefficients of the **Binomial expansion**:

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

Counting — Group Divisions (Multinomial Coefficients)

The number of ways to divide *n* items into *k* <u>labeled</u> groups of sizes n_1, n_2, \ldots, n_k respectively and $n_1 + n_2 + \cdots + n_k = n$ is

$$\binom{n}{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1! \cdot n_2! \cdot \cdots \cdot n_k!}.$$

Note that "combination" $\binom{n}{r}$ is special case of dividing *n* items into 2 groups of sizes *r* and *n* - *r* respectively.

Counting — Group Divisions — Proof

- There are $\binom{n}{n}$ ways to choose n_1 items out of n for Group #1.
- ▶ There are $\binom{n-n_1}{n_2}$ ways to choose n_2 items out of $n n_1$ for Group #2.
- ► There are $\binom{n-n_1-n_2}{n_3}$ ways to choose n_3 items out of $n n_1 n_2$ for Group #3.

There are
$$\binom{n-n_1-\dots-n_{k-1}}{n_k}$$
 ways to choose n_k items out of $n-n_1-n_2-\dots-n_{k-1}$ for Group $\#$ k.

In total,

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3} \times \cdots \times \binom{n-n_1-\dots-n_{k-1}}{n_k}$$

$$= \frac{n!}{n_1!(n-n_1)!} \cdot \frac{(n-n_1)!}{n_2!(n-n_1-n_2)!} \cdot \frac{(n-n_1-n_2)!}{n_3!(n-n_1-n_2-n_3)!} \cdots \frac{(n-n_1-\dots-n_{k-1})!}{n_k! \, 0!}$$

$$= \frac{n!}{n_1! \cdot n_2! \cdots \cdot n_k!} = \binom{n}{n_1, n_2, \dots, n_k}$$

Counting — Group Divisions — Proof

- There are $\binom{n}{n}$ ways to choose n_1 items out of n for Group #1.
- ▶ There are $\binom{n-n_1}{n_2}$ ways to choose n_2 items out of $n n_1$ for Group #2.
- ► There are $\binom{n-n_1-n_2}{n_3}$ ways to choose n_3 items out of $n n_1 n_2$ for Group #3.

There are
$$\binom{n-n_1-\dots-n_{k-1}}{n_k}$$
 ways to choose n_k items out of $n-n_1-n_2-\dots-n_{k-1}$ for Group $\#$ k.

In total,

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3} \times \cdots \times \binom{n-n_1-\dots-n_{k-1}}{n_k}$$

$$= \frac{n!}{n_1!(n-n_1)!} \cdot \frac{(n-n_1)!}{n_2!(n-n_1-n_2)!} \cdot \frac{(n-n_1-n_2)!}{n_3!(n-n_1-n_2-n_3)!} \cdots \frac{(n-n_1-\dots-n_{k-1})!}{n_k!0!}$$

$$= \frac{n!}{n_1! \cdot n_2! \cdots \cdot n_k!} = \binom{n}{n_1, n_2, \dots, n_k}$$

Counting — Group Divisions — Proof

- There are $\binom{n}{n}$ ways to choose n_1 items out of n for Group #1.
- ▶ There are $\binom{n-n_1}{n_2}$ ways to choose n_2 items out of $n n_1$ for Group #2.
- ► There are $\binom{n-n_1-n_2}{n_3}$ ways to choose n_3 items out of $n n_1 n_2$ for Group #3.

There are
$$\binom{n-n_1-\cdots-n_{k-1}}{n_k}$$
 ways to choose n_k items out of $n-n_1-n_2-\cdots-n_{k-1}$ for Group $\#$ k.

In total,

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3} \times \cdots \times \binom{n-n_1-\dots-n_{k-1}}{n_k}$$

$$= \frac{n!}{n_1!(n-n_1)!} \cdot \frac{(n-n_1)!}{n_2!(n-n_1-n_2)!} \cdot \frac{(n-n_1-n_2)!}{n_3!(n-n_1-n_2-n_3)!} \cdots \frac{(n-n_1-\dots-n_{k-1})!}{n_k!0!}$$

$$= \frac{n!}{n_1!\cdot n_2!\cdots n_k!} = \binom{n}{n_1, n_2, \dots, n_k}$$

Counting Examples

Count how many ways each is possible:

1. Split 10 students into two teams, with 4 students and 6 students:

$$\binom{10}{4} = \frac{10!}{4!\,6!} = 210$$

2. Split 10 students into Section 1 and Section 2, each with 5 students:

$$\binom{10}{5} = \frac{10!}{5! \, 5!} = 252$$

3. Split 10 students into two teams, each with 5 students:

$$egin{pmatrix} 10 \ 5 \end{pmatrix} \cdot rac{1}{2} = 126$$

4. Split 15 students into 3 teams, each with 5 students:

$$\binom{15}{5,5,5} \cdot \frac{1}{3} = 126126.$$

Example: Finding Probabilities by Counting

20 cards are drawn randomly from a standard 52 card deck. What is the probability that exactly half of the 20 cards are Red?

Ans:

Example: Finding Probabilities by Counting

If you arrange 4 pennies and 4 dimes into a random order, what is the probability that they alternate?

Ans: