STEVENS INSTITUTE OF TECHNOLOGY 1870

MA 540: Introduction to Probability Theory

Lecture 2: Axioms of Probability

Sample Space – Denoted Ω or *S*

- This set of all possible outcomes of an experiment is known as the **sample space** of the experiment and is denoted by S or Ω
- Examples
  - Flipping one coin – sample space consists of {H, T}
  - Flipping 3 coins – sample space consists of {(HHH), (HHT), (HTH), (THH), (HTT), (THT), (TTH), (TTT)} – curly brackets are sets, parenthesis are outcomes
  - Race with 7 people – S = {all 7! permutations of (1, 2, 3, 4, 5, 6, 7)}
  - Length of this class – $S = \{x: 0 \le x \le 2$ hours 35 minutes$\}$
    - Note this is continuous, there are uncountably many different possible outcomes

Event

- Any subset A in the sample space is called an **event.**
  - Another perspective: a set consisting of possible outcomes of the experiment
- Examples
  - **Simple**: Flipping 1 coin – 4 possible events 1. ∅, 2. {H}, 3. {T}, 4. {H, T} = S

  - **Moderate:** Flipping 3 coins – one event is A = {(HHH), (HHT), (HTH), (THH)}
    - This is the event that at least two of the coins lands heads. Alternatively, that at most one coin lands tails

  - **Hard to enumerate:** Race with 7 people – A = {all permutations starting with 5}
    - This has 6! different outcomes. The first position is fixed, the rest are note
    - This event is all outcomes where person 5 wins

  - Continuous example: Length of this class – $A = \{x: 0 \leq x \leq 2$ hours 30 minutes$\}$
    - This event is “class does not go over time” or “students don’t get mad at professor” or “professor gets good reviews” or “professor keeps job” or “professor’s children get to eat”

Question

- Consider the experiment where a die is continually rolled until a 6 is rolled or until is has been rolled 5 times.
  - Describe the sample space of this experiment?
  - What’s not in the sample space?

- Consider the events consisting of the number of dice rolls in the experiment above.
  - What outcomes are contained in the event “the die is rolled once?”
  - What outcome are contained in the event “the die is rolled twice?”
  - What outcome are contained in the event “the die is rolled five times?”

Unions

- If A is an event in S and B is an event in S, $A \cup B$ *is* the set consisting of any outcome in A, any outcome in B, or any outcome in both A and B. This is called the union of A and B
  - In the race example, the event "person 2 wins or person 3 wins" is represented by the union {Permutations of (1, 2, 3, 4, 5, 6, 7) beginning with 2} ∪ {Permutations of (1, 2, 3, 4, 5, 6, 7) beginning with 3}
  - This is a union of **disjoint** sets. There is no permutation that begins with both 2 and begins with 3
  - Unions in general do not need to be disjoint
  - For example, when flipping 3 coins, the set "the first or last coin is tails" is the union of the sets {(THH), (THT), (TTH), (TTT)} ∪{(HHT), (HTT), (THT), (TTT)}
    = {(THH), (TTH) (HHT), (HTT), (THT), (TTT)}
  - Notice some overlap between the sets in the union, but the overlapping outcomes are only listed once in the union

Intersections

- If A is an event in S and B is an event in S, $A \cap B$ *is* the set consisting of only those outcomes in both A and B. This is called the intersection of A and B. It is also written AB

- The intersection of disjoint sets is ∅
  - Think of the sets "the sum of 2 dice is 5" and "the sum of 2 dice is 6"

- The intersection of the set "at least 2 coins are heads" ∩ "at least one coin is tails" is "exactly 2 coins are heads" (Why?)

- Where the union of 2 sets is at least as large as the larger of the two sets in the union, the intersection of 2 sets is at most as large as the smaller of the 2 sets in the union

Questions

- In an experiment where we flip a coin 25 times, what is the union of the events “at least one head appears” and “at least one tail appears”?

- In an experiment where we flip a coin 25 times, what is the intersection of the events “at least one head appears” and “at least one tail appears”?

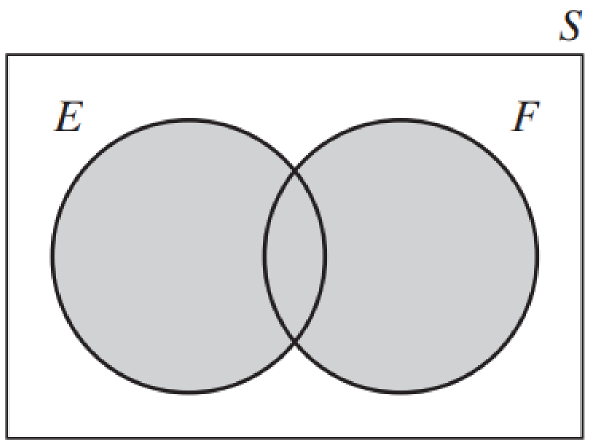
Multiple unions/intersections, complements, subset

- Extension beyond 2 sets:
  - $\cup_{i=1}^{\infty} A_i$ is the union of all events A_i, and represents the set of outcomes which occur in *any* set A_i
  - $\cap_{i=1}^{\infty} A_i$ is the intersection of all events A_i, and represents the set of outcomes which occur in *every* set A_i
  - $\cup_{i=1}^{\infty}(0, \frac{1}{i})$ - the union of intervals running from 0 to 1/i = (0 , 1)
  - $\cup_{i=1}^{\infty}(0, \frac{1}{i})$ - the intersection of intervals running from 0 to 1/i = Ø

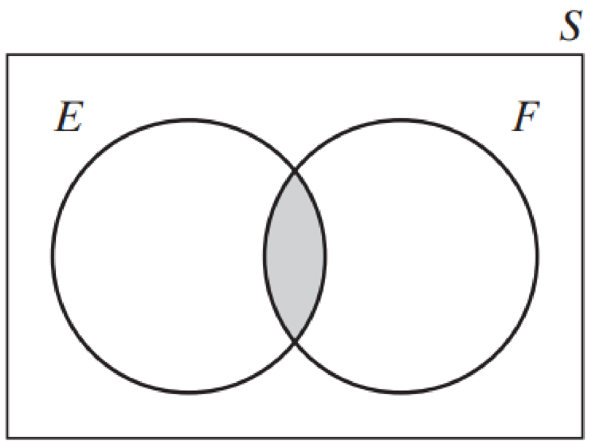
- Complement of *A*, denoted A^c, comprises all outcomes in *S* but not in A
  - By definition, $A \cap A^c = \emptyset$
  - Simplest case: 1 coin flip – if *A* = {H}, A^c = {T}

- If all outcomes in *A* are also in *B*, we say that *A* is a subset of *B*, or *A* is contained in *B* or $A \subseteq B$

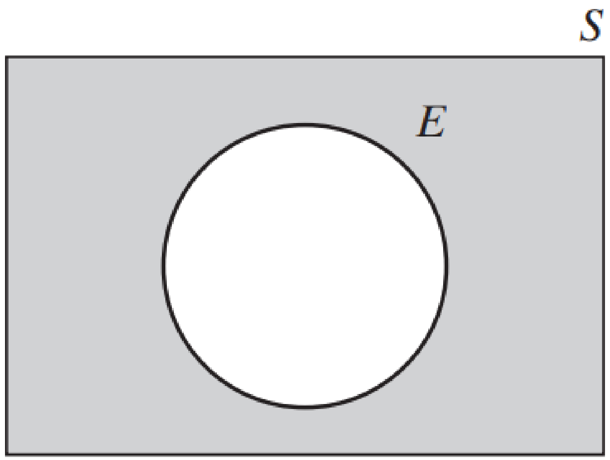
Venn Diagrams



(a) Shaded region: $E \cup F$.



(b) Shaded region: EF.



(c) Shaded region: E^c.

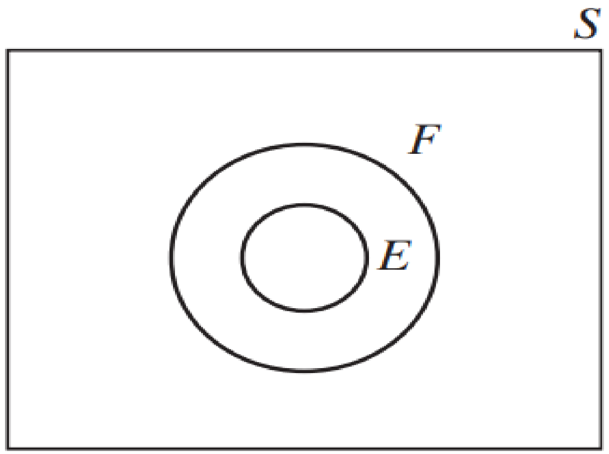
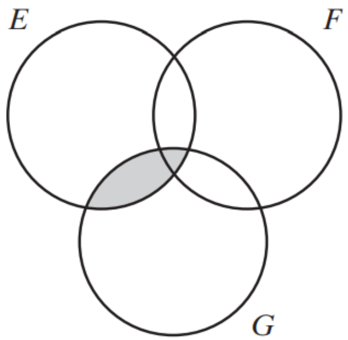


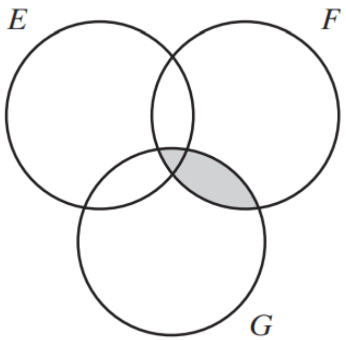
Figure 2.2 $E \subset F$.

Laws

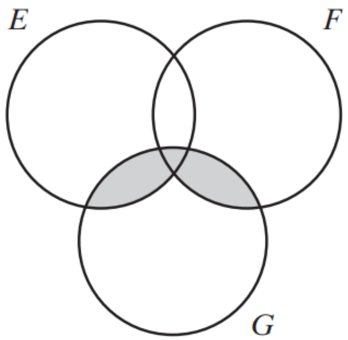
- Commutativity: $E \cup F = F \cup E$ $E \cap F = F \cap E$
- Associativity: $(E \cup F) \cup G = E \cup (F \cup G)$ $(E \cap F) \cap G = E \cap (F \cap G)$
- Distributivity: $(E \cup F) \cap G = (E \cap G) \cup (F \cap G)$ $(E \cap F) \cup G = (E \cup G) \cap (F \cup G)$



(a) Shaded region: EG.



(b) Shaded region: FG.



(c) Shaded region: $(E \cup F)G$.

Questions

- What is the complement of A^c?
- If $A \subseteq B$,
  - What is the union of A and B?
  - What is the intersection of A and B?
- What's the complement of $(A \cup A^c) \cap B$?

De Morgan’s laws

- $(\bigcup_{i=1}^{n} A_i)^c = \bigcap_{i=1}^{n} A_i^c$
  - The complement of the union of sets is composed of all those elements which are in the complement of every set
  - To be in the union, an element must only appear in at least one A_i. To be in the intersection, it must this not appear in all
- $(\bigcap_{i}^{n} A_i)^C = \bigcup_{i=1}^{n} A_i^c$
  - The complement of the intersection of sets $\{A_i\}$ is the union of the each of the complements
  - To be in the intersection, an element must be in all sets $\{A_i\}$. To be in the union of complements, an element must only not appear in at least one A_i
- Proof on board

Probability as relative frequency

- Repeat an experiment under identical conditions n times. For event A, let n(A) be the number of times event A occurs
- We want: P(A) = $\lim_{n \to \infty} \frac{n(A)}{n}$
  - The limiting proportion of times event A occurs
- This is intuitive, but there are some embedded assumptions that the limit exists and is unique
- For this reason, we use a different definition of probability and build from there

Axioms of probability

1. 0 ≤ P(A) ≤ 1
   - "Probabilities take on values between 0 and 1"

2. P(S) = 1
   - "With probability 1, the outcome will be a point in the sample space"

3. Letting A_1, A_2, A_3... be mutually exclusive events (that is $A_i \cap A_j$ = Ø if i ≠ j), then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

   - "For any sequence of mutually exclusive events, the probability of at least one of these events occurring is the sum of the individual probabilities"

Consequences of axioms of probability

- If A_1 = *S*, $A_i: i > 1 = \emptyset$, then the events are mutually exclusive. Using axioms 2 and 3, that means $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$
  = P(S) + $\sum_{i=2}^{\infty} P(A_i)$ = 1
  Hence P(∅) = 0

- Consequence: Letting $A_1, A_2, \ldots, A_n$ be mutually exclusive events, then

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

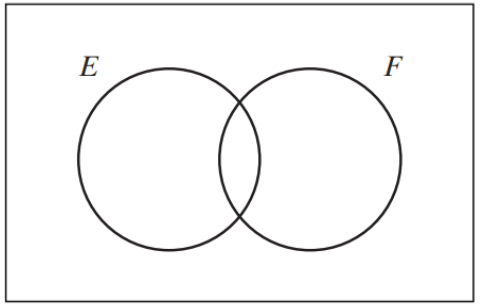
simply by letting all $A_i: \mathrm{i} > \mathrm{n} = \emptyset$

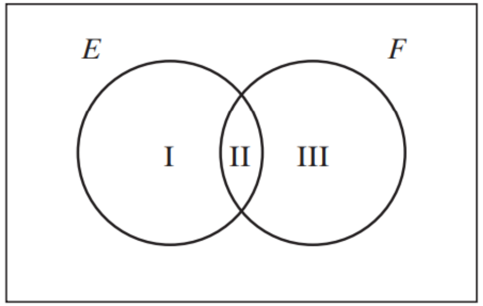
- As A ∪ A^c = *S*, we can use the above with axiom 2 to deduce that P(A) = 1- P(A^c)

More consequences

- If $A \subseteq B$, $P(A) \leq P(B)$
  - $B = A \cup (A^c \cap B)$ - Draw it to check. A and $A^c \cap B$ are mutually exclusive since any event in $A^c \cap B$ is also in A^c, so $P(B) = P(A) + P(A^c \cap B)$. Therefore, using axiom 1, $P(A) \leq P(B)$.

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - Disjointify: $P(A \cup B) = P(A \cup (A^c \cap B)) = P(A) + P(A^c \cap B)$
  - Similarly: $P(B) = P(A^c \cap B) + P(A \cap B)$, therefore $P(A^c \cap B) = P(B) - P(A \cap B)$
  - Plugging The second result into the first, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - This proof works, but it's much easier to see in a Venn diagram





Linda the bank teller example

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Which is more probable?

- Linda is a bank teller.
- Linda is a bank teller and is active in the feminist movement.

Linda the bank teller example

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Which is more probable?

- Linda is a bank teller.
- Linda is a bank teller and is active in the feminist movement.

- Must be the first one. Let A be the event "Linda is a bank teller". Let B be the event "Linda is is active in the feminist movement."
- The first proposition is $P(A)$. The second proposition is $P(A \cap B)$
- $(A \cap B) \subseteq A$ therefore $P(A \cap B) \leq P(A)$

Inclusion – Exclusion identity

- What if we want to find the probability of the union of more than 2 sets?
- General formula called the inclusion-exclusion identity

$$P(\cup_{i=1}^{n} E_i) = \sum_{r=1}^{n} (-1)^{r+1} \sum_{i_1 < \cdots < i_r} P(E_{i_1} \cdots E_{i_r})$$

- Note if all sets are mutually exclusive, all terms where r > 1 are necessarily = 0
- We can use this for upper and lower bounds on probabilities:

$$P(\cup_{i=1}^{n} E_i) \leq \sum_{i=1}^{n} P(E_i)$$

$$P(\cup_{i=1}^{n} E_i) \geq \sum_{i=1}^{n} P(E_i) - \sum_{j<i} P(E_i E_j)$$

Sample spaces with equally likely outcomes

- If an experiment has a finite sample space, e.g. $S = \{1, 2, 3, ..., N\}$, **where each outcome is equally likely to occur**, that is P({1}) =P({2}) = P({3}) = … = P({N}):
  - By axioms 2 and 3, P({1}) =P({2}) = P({3}) = … = P({N}) = 1/N

- In this case, the probability of an event is simply the proportion of outcomes in the event

- P(A) = $\frac{\text{Number of outcomes in } A}{\text{Number of outcomes in } S}$

- We saw this last week in an example where we wanted to know the probability of getting k heads in n fair coin flips.

Example - committee

- A committee of 5 is to be selected from a group of 6 men and 9 women. If the selection is made randomly, what is the probability that the committee consists of 3 men and 2 women?

Example - committee

- A committee of 5 is to be selected from a group of 6 men and 9 women. If the selection is made randomly, what is the probability that the committee consists of 3 men and 2 women?

- Numerator: number of ways to choose 3 men out of 6 men multiplied by the number of ways to choose 2 women out of 9 women. In all cases order doesn't matter

- Denominator: Number of ways to choose 5 people out of 15 people

$$\frac{\binom{6}{3}\binom{9}{2}}{\binom{15}{5}}$$

Example - birthday problem

- What is the probability that in a room of n people, at least 2 people share a birthday? For simplicity, assume no leap day birthdays
- Hint: Easier to calculate probability all birthdays are different, then simply take the complement.

Example - birthday problem

- What is the probability that in a room of n people, at least 2 people share a birthday? For simplicity, assume no leap day birthdays
- Hint: Easier to calculate probability all birthdays are different, then simply take the complement.
- Lets start simply at n = 2. 2 slots. First person has 365 days to choose from. Second person has 364 days to choose from. Assuming each day has equal probability of being chosen, the probability of nobody sharing a birthday is $\frac{365}{365} * \frac{364}{365} = \frac{364}{365}$. Therefore, by complements, the probability of anyone sharing a birthday is $\frac{1}{365}$

Example - birthday problem

- What is the probability that in a room of n people, at least 2 people share a birthday? For simplicity, assume no leap day birthdays
- Hint: Easier to calculate probability all birthdays are different, then simply take the complement.
- Lets start simply at n = 2. 2 slots. First person has 365 days to choose from. Second person has 364 days to choose from. Assuming each day has equal probability of being chosen, the probability of nobody sharing a birthday is $\frac{365}{365} * \frac{364}{365} = \frac{364}{365}$. Therefore, by complements, the probability of anyone sharing a birthday is $\frac{1}{365}$
- For n < 366, we have n slots, where the probability of not matching any previous birthday is $\frac{365 - i + 1}{365}$.
  - Therefore, the probability of no birthdays matching is $\frac{365!}{365^n * (365-n)!}$. The probability of any 2 people sharing a birthday is 1 – that value
  - For n = 23, the probability of 2 people sharing a birthday is about .5073

Example - cards

- A poker hand consists of 5 cards. If the cards have distinct consecutive values and are not all of the same suit, we say that the hand is a straight. For instance, a hand consisting of the five of spades, six of spades, seven of spades, eight of spades, and nine of hearts is a straight. What is the probability that one is dealt a straight?

Sequences and limits

- Increasing sequence of events:
  - $A_1 \subseteq A_2 \subseteq A_3 \subseteq A_4 \subseteq \ldots A_n \subseteq A_{n+1} \subseteq \ldots..$
  - Example: all values x such that 1/n < x <1
  - Can define a limit = $\lim_{n\to\infty} A_i = \cup_{i=1}^{n} A_i$
- Decreasing sequence of events:
  - $A_1 \supseteq A_2 \supseteq A_3 \supseteq A_4 \supseteq \ldots A_n \supseteq A_{n+1} \supseteq \ldots..$
  - Example: 1/n > x > 0
  - Can define a limit = $\lim_{n\to\infty} A_i = \cap_{i=1}^{n} A_i$

- Proposition 6.1: For an increasing or decreasing sequence of events $\lim_{n\to\infty} P(A_i) = \mathrm{P}(\lim_{n\to\infty} A_i)$
  - The left side is the limit of a number. The right side is the probability of the limiting value of a sequence of sets

Using limits to deal with infinity

Suppose that we possess an infinitely large urn and an infinite collection of balls labeled ball number 1, number 2, number 3, and so on. Consider an experiment performed as follows: At 1 minute to 12 P.M., balls numbered 1 through 10 are placed in the urn and ball number 10 is withdrawn. (Assume that the withdrawal takes no time.) At $\frac{1}{2}$ minute to 12 P.M., balls numbered 11 through 20 are placed in the urn and ball number 20 is withdrawn. At $\frac{1}{4}$ minute to 12 P.M., balls numbered 21 through 30 are placed in the urn and ball number 30 is withdrawn. At $\frac{1}{8}$ minute to 12 P.M., and so on. The question of interest is, How many balls are in the urn at 12 P.M.?

Using limits to deal with infinity

However, let us now change the experiment and suppose that at 1 minute to 12 P.M., balls numbered 1 through 10 are placed in the urn and ball number 1 is withdrawn; at $\frac{1}{2}$ minute to 12 P.M., balls numbered 11 through 20 are placed in the urn and ball number 2 is withdrawn; at $\frac{1}{4}$ minute to 12 P.M., balls numbered 21 through 30 are placed in the urn and ball number 3 is withdrawn; at $\frac{1}{8}$ minute to 12 P.M., balls numbered 31 through 40 are placed in the urn and ball number 4 is withdrawn, and so on. For this new experiment, how many balls are in the urn at 12 P.M.?

Using limits to deal with infinity

Let us now suppose that whenever a ball is to be withdrawn, that ball is randomly selected from among those present. That is, suppose that at 1 minute to 12 P.M. balls numbered 1 through 10 are placed in the urn and a ball is randomly selected and withdrawn, and so on. In this case, how many balls are in the urn at 12 P.M.?

Another interpretation of probability

- Another interpretation of probability is as a “degree of belief”
- Under this interpretation, the axioms of probability should still hold
  - Some people’s beliefs may not, for example, sum to 1, but this is incoherent with probability
  - This is the reality of dealing with people. In economics, people exhibit irrationality incompatible with the logic of “homo economicus”
  - This is not actually a problem with probability, but with colloquial reasoning

Homework 2

- Due Wednesday 9/27
- Problems from textbook Chapter 2:
  - 36
  - 41
- Theoretical problems from textbook Chapter 2:
  - 2
  - 13
  - 20