A Brief Introduction to Probability

ECE/CS 541

Foundations

- Notion of an "outcome of an experiment"
  - Experiment characterized by something tangible, e.g.
    - Flip a coin
      - Outcomes H, T
    - Flip a coin five times
      - Outcomes are sequences, e.g. HHHHH, HHHHT, HHHTH, and so on
    - Roll die 1 and die 2
      - Outcomes : faces shown, e.g. (die1 = 2, die2 = 4), (die1=1, die2=1), (die1=6, die2=1), ...
    - Run a system for 10,000 hours
      - Outcomes : Fails, does not fail

Event Spaces

- Each outcome is from an *event space* Ω of all possible outcomes
  - Experiment, flip a coin once, Ω = {H,T}
  - Experiment, flip a coin 5 times
    - Ω = { HHHHH, HHHHT, …., TTTHT, TTTTH, TTTTT}
  - Experiment, roll die1 and die2
    - Ω = {(1,1),(1,2),(1,3), …., (6,4),(6,5),(6,6) }

Event Spaces (cont'd)

Given an event space Ω we *define* an event as a subset of Ω

Singleton event corresponds to one outcome, e.g.,

- coin face is 'H'
- Sequence of coin faces is 'HTHHT'
- die1 = 2, die2 = 4

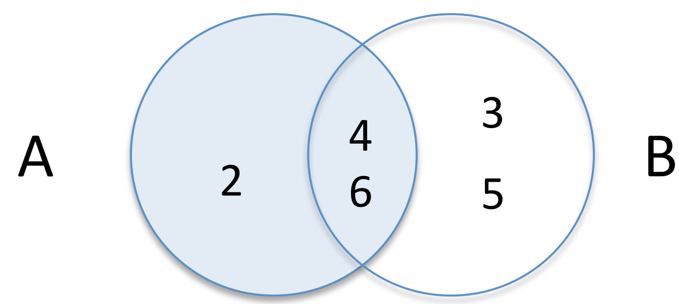
Some events refer to more interesting subsets, e.g.

- {(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)} is event that sum of die1 and die2 is 7
- Given an event and an outcome of an experiment, we can ask whether the outcome is in the event

Algebra of Events

Ways in which events combine

- Union of A and B --- set-theoretic union, describes space of outcomes that are either in A, or B, or both
  - Example: A = { die toss with even number }, B = {die toss with value at least 3}
- Venn diagram view

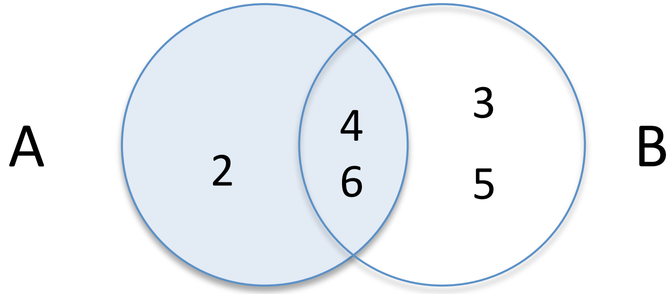


$A \cup B = \{ 2, 3, 4, 5, 6\}$

Algebra of Events

Ways in which events combine

- Intersection of A and B --- set-theoretic intersection, describes space of outcomes that are in A **and** B
- Example: A = { die toss with even number }, B = {die toss with value at least 3}
- Venn diagram view

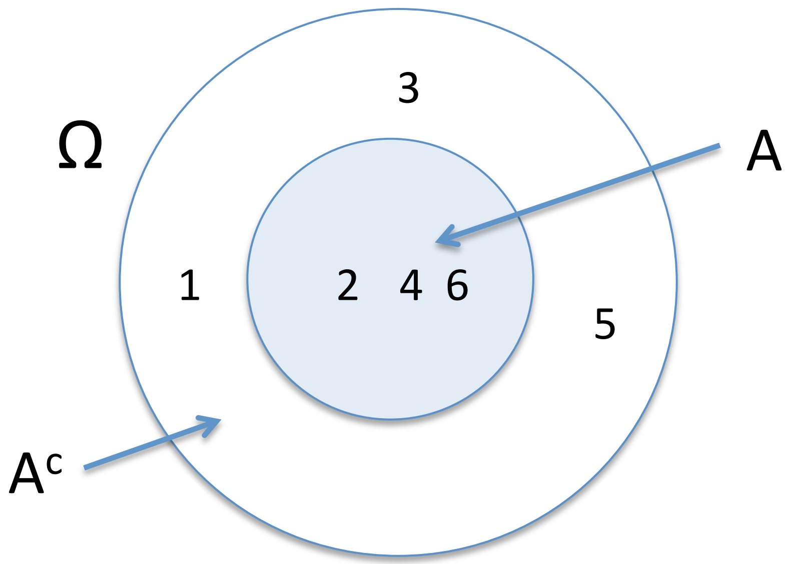


$A \cap B = \{4, 6\}$

Algebra of Events

Ways in which events combine

- Complement of A --- set theoretic "entire event space excluding A" Ω / A often denoted A^c



Boolean Algebra

Algebra of sets is Boolean algebra

- Ω plays the role of 1, $\varnothing$ plays role of 0
- Set union plays the role of logical OR
- Set intersection plays the role of logical AND
- Set complement plays the role of NOT

$\Omega \cap A = A$, $\Omega \cup A = \Omega$, $\varnothing \cap A = \varnothing$, $\varnothing \cup A = A$

$A \cup A^c = \Omega$, $A \cap A^c = \varnothing$

- Commutative laws $A \cup B = B \cup A$, $A \cap B = B \cap A$
- Distributive laws

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

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

- DeMorgan’s Law

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

Now some “probability”

We aim to give weights to events

- N.B. For a given event A and outcome E, we can say that E is in A or not
  - Example: A = { sum of two die is 7 }, E = (die1=2,die2=6)
- We have a notion then that if we repeated an experiment many times, some fraction of those outcomes would be in A, denote P(A)
  - This the ‘frequency’ interpretation of probability

Axioms of Probability

Function P: $2^{\Omega} \longrightarrow [0,1]$ is a *probability measure* if

- If $A_1, A_2, \ldots, A_n$ are disjoint (i.e., $A_j \cap A_k = \emptyset$ for $j \neq k$) then $P(A_j \cup A_k) = P(A_j) + P(A_k)$

- $P(A) \geq 0$ for all events $A \subseteq \Omega$

- $P(\Omega) = 1$

That's all!

Consequences of Axioms

You can prove the following from these axioms

1. $P(\varnothing) = 0$

2. $P(A) + P(A^c) = 1$

Consequences of Axioms

You can prove the following from these axioms

$A \cup \emptyset = A$

1. $P(\emptyset) = 0$

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

$\Rightarrow P(\emptyset) = 0$

2. $P(A) + P(A^c) = 1$

$A \cup A^c = \Omega$

$P(A \cup A^c) = P(\Omega)$

$P(A) + P(A^c) = P(\Omega)$

$= 1$

Consequences of Axioms

You can prove the following from these axioms

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

From point of view of Venn diagram think of P() as “area”

Consequences of Axioms

You can prove the following from these axioms

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

From point of view of Venn diagram think of P() as "area"

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

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

$P(A \cap B) + P(A \cap B^c) + P(A \cap B) + P(B \cap A^c) - P(A \cap B)$

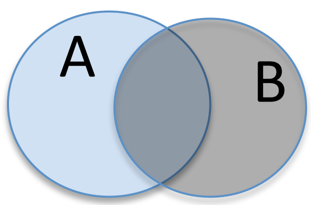
$= P(A \cup B)$

Consequences of Axioms

You can prove the following from these axioms

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

From point of view of Venn diagram think of P() as "area"

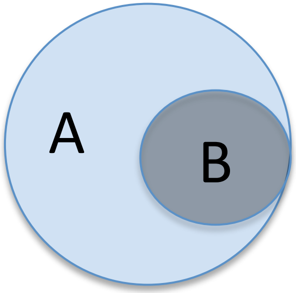


Consequences of Axioms

You can prove the following from these axioms

4. If $B \subset A$, then $P(B) \leq P(A)$, furthermore

$$P(A/B) = P(A) - P(B)$$



Consequences of Axioms

You can prove the following from these axioms

5. $P(A_1 \cup A_2 \cup \ldots A_n) \leq P(A_1) + P(A_2) + \ldots + P(A_n)$

Consequences of Axioms

You can prove the following from these axioms

5. $P(A_1 \cup A_2 \cup \ldots A_n) \leq P(A_1) + P(A_2) + \ldots + P(A_n)$

Proof by Induction

$n=2$

$P(A_1 \cup A_2) \leq P(A_1) + P(A_2)$

$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A \cap B)$

Case proven for $n=2$

Consider any $n \geq 2$

Define $B = A_2 \cup A_3 \cup \cdots A_k$

$P(A_1 \cup A_2 \cup \cdots \cup A_k) = P(A_1 \cup B)$ by induction hypothesis

$P(B) = P(A_2 \cup \cdots A_k) \leq P(A_2) + \cdots P(A_k)$

$\leq P(A_1) + P(B)$

$\leq P(A_1) + P(A_2) + \cdots + P(A_k)$

Proven

Conditional Probability

Denote the probability of A given B by $P(A \mid B)$

Frequency interpretation

- n experiments, $c(\omega,n)$ number whose outcome is ω
- $c(B,n)$ number whose outcomes are in B
- Frequency of $\omega \in B$ relative to experiments with outcomes in B

$P(A|B)$ is the frequency of $\omega \in A \cap B$ relative to the frequency of B

Conditional Probability

Denote the probability of A given B by P(A | B)

$$P(A|B) = \sum_{\omega \in A \cap B} \lim_{n \to \infty} \frac{c(\omega, n)}{c(B, n)}$$

$$= \sum_{\omega \in A \cap B} \frac{\lim_{n \to \infty} \frac{c(\omega, n)}{n}}{\lim_{n \to \infty} \frac{c(B, n)}{n}}$$

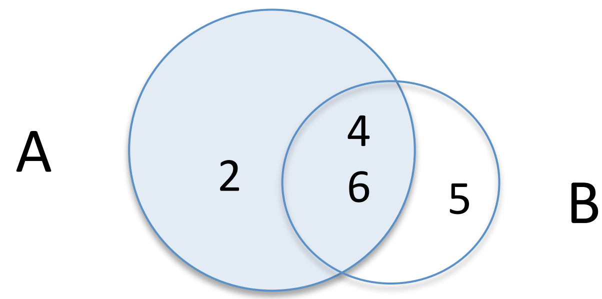
$$= \frac{\lim_{n \to \infty} \frac{c(A \cap B, n)}{n}}{\lim_{n \to \infty} \frac{c(B, n)}{n}} = \frac{P(A \cap B)}{P(B)}$$

Conditional Probability

Example

- A is die toss shows even number
    - But we know also the die toss is 4 or larger (event B)

Venn diagram view



Frequency of {2,4,6} in space of {4,5,6} is same as frequency of {4,6} in {4,5,6} : (P(4)+P(6)) / (P(4) + P(5) + P(6))

In general $P(A \mid B) = P(A \cap B) / P(B)$

Independent Events

Events A and B are said to be independent if

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

Whether or not this is true depends on definitions of A and B

Example: 3 consecutive coin tosses

A = {at least 1 heads}, B = {exactly 1 heads}

A = { HHH, HHT, HTH, THH, HTT, THT, TTH }, P(A) = 7/8

B = {HTT,THT,TTH}, P(B) = 3/8

$A \cap B$ = {HTT,THT,TTH}, P(B) = 3/8

$$P(A \cap B) = 3/8 \quad \neq \quad (7/8)*(3/8)$$

Independent Events

Events A and B are said to be independent if

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

Whether or not this is true depends on definitions of A and B

Example: 3 consecutive coin tosses

A = {first two tosses have exactly one H}, B = {last toss is H}

A = { HTH, HTT, THH, THT }, $P(A) = 4/8 = ½$

B = { HHH, HTH, THH, TTH }, $P(B) = 4/8 = ½$

$A \cap B$ = {HTH,THH} = $P(A \cap B) = 2/8 = ¼$

$¼ = P(A \cap B) = ½ * ½ = P(A)*P(B)$

So A and B are independent

Independent Events and Conditional Probability

An intuitive relationship:

if A is independent of B, then P(A) does not depend on B:

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

Conditional Probability

Denote the probability of A given B by P(A | B)

P(A|B) is the frequency of $\omega \in A \cap B$ relative to the frequency of B

$$P(A|B) = \sum_{\omega \in A \cap B} \lim_{n\to\infty} c(\omega,n)/c(B,n)$$

$$= \lim_{n\to\infty} \sum_{\omega \in A \cap B} c(\omega,n)/c(B,n)$$

$$= \lim_{n\to\infty} \sum_{\omega \in A \cap B} (c(\omega,n)/n) / (c(B,n)/n)$$

$$= (\sum_{\omega \in A \cap B} \lim_{n\to\infty} c(\omega,n)/n) / (\lim_{n\to\infty} c(B,n)/n)$$

$$= (\lim_{n\to\infty} c(A \cap B,n)/n) / (\lim_{n\to\infty} c(B,n)/n)$$

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