Online Review Course of
Undergraduate Probability and Statistics

# Review Lecture 4
# Probability, part 1

Chris A. Mack
*Adjunct Associate Professor*

Course Website: www.lithoguru.com/scientist/statistics/review.html



## Probability

- Probability Theory – a mathematical framework for reasoning about uncertainty
- Most engineering problems are solved as if they were deterministic (the same inputs always give the same output)
- Real life is messy. Two possibilities:
  – Randomness adds uncertainty to our deterministic solution; or
  – Randomness dominates the outcome



## Frequentist View

- Q: Given a coin of unknown fairness, how would you estimate the probability of getting a “head”?
- A: Flip the coin a large number of times (N). Count the number of heads (H). Then,

$$\mathbb{P}(H) \cong \frac{H}{N} \qquad \text{(weak law of large numbers)}$$

- Assumes that each trial results in independent, identically distributed outcomes
- But, what does it mean to discuss the probability of a unique event (e.g., what is the probability of ocean levels rising more than 50 cm by 2100?)



## Elements of a Probabilistic Model

- Sample Space (Ω)
  – The set of all possible outcomes of an experiment
- Probability Law (ℙ)
  – Assigns a non-negative number to each event of interest
  – Any useful probability law will follow Kolmogorov Axioms of Probability
- We’ll ignore many mathematical details



## Sample Space (Ω)

- The set of all possible outcomes of an experiment
  – experiment: the underlying process that will produce exactly one result
  – Sample space may be discrete (finite or countably infinite) or continuous
  – Outcomes must be distinct and mutually exclusive
  – Sample space must be collectively exhaustive



## Example Sample Space

- Experiment = flip a coin three times
  – Important question, does order matter?

If yes, Ω = {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}

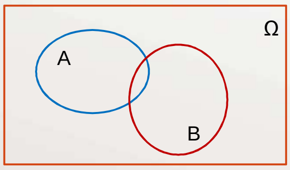
If no, Ω = {three H, two H + one T, two T + one H, three T}
or Ω = {0, 1, 2, 3} where outcome defined as number of heads

- Event – a collection of possible outcomes
  – Example events: X = HHT
    X = even number of heads
    X = two or more heads

## Set Basics

- Events are subsets of the sample space
  - We use set algebra to describe events

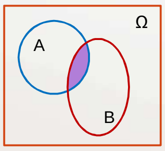



Venn Diagram





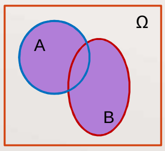
## Set Basics




Intersection:

$$A \cap B = \{x | x \in A \ and \ x \in B\}$$

such that — The set of all — is a member of



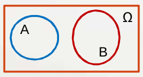

Union:

$$A \cup B = \{x | x \in A \ or \ x \in B\}$$

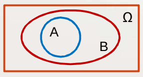




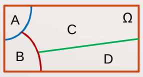
## Set Terminology




Disjoint: A and B are disjoint if they share no elements; $A \cap B = \phi$ (null set)




Subset: A is a subset of B if every element of A is found in B; $A \supset B \ if \ A \cap B = A$




Partition: A collection of disjoint sets whose union is Ω; $e.g., \ A \cup B \cup C \cup D = \Omega$


Compliment: The compliment of A is all the elements of Ω that do not belong to A; $A^c = \bar{A} = \{x \in \Omega | x \notin A\}$





## Set Algebra

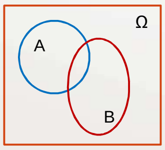
Useful Identities: $\Omega^c = \phi$ $\quad A \cup A^c = \Omega$

$(A^c)^c = A$ $\quad A \cap A^c = \phi$

DeMorgan's Laws:

$$\left(\bigcup_{i=1}^{n} E_i\right)^c = \bigcap_{i=1}^{n} E_i^{\,c}$$

$$\left(\bigcap_{i=1}^{n} E_i\right)^c = \bigcup_{i=1}^{n} E_i^{\,c}$$








## Probabilistic Law

- A probabilistic law assigns a number to each event of interest
  - ℙ(E) = probability that event E will happen
- A very common approach is to first assign probabilities to each outcome in Ω
  - For $p_i$ = probability of outcome i,

$$\mathbb{P}(E) = \sum_{all \ outcomes \ in \ E} p_i$$





## Axioms of Probability

- All valid probability laws must obey the Axioms of Probability (Kolmogorov Axioms)
  - Non-negativity: ℙ(E) ≥ 0 for all E
  - Normalization: ℙ(Ω) = 1
  - Additivity: for any sequence of disjoint events $E_i$

$$\mathbb{P}\left(\bigcup_{i=1}^{n} E_i\right) = \mathbb{P}(E_1) + \mathbb{P}(E_2) + \cdots + \mathbb{P}(E_n)$$

Recall: disjoint = mutually exclusive

## Probability Identities

- Given any probability law that obeys the probability axioms,
  - $\mathbb{P}(\emptyset) = 0$
  - $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$
  - $\mathbb{P}(E) \leq 1$
  - If $E \subset F$ then $\mathbb{P}(E) \leq \mathbb{P}(F)$
  - $\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F)$
- Can you prove these?

 13

## Applying Probability

- When mapping the real world to a probabilistic model, we have choices
  - We pick the sample space based on how we define our experiment
  - We define our probability law (constrained by the axioms of probability)
- We judge the resulting probabilistic model by its usefulness

 14

## Review #4: What have we learned?

- Explain the frequentist view of probability
- What are the two elements of a probabilistic model?
- What are the defining properties of a sample space?
- Define disjoint sets, a subset, a partition, and the compliment of a set
- What are the three probability axioms?

 15