Introduction to Probability and Statistics

Prof. Mohammed-Slim Alouini & Prof. Hesham ElSawy
slim.alouini@kaust.edu.sa & hesham.elsawy@kfupm.edu.sa

Eng. Chaouki Bem Issaid & Eng. Lama Niyazi
chaouki.benissaid@kaust.edu.sa & lama.niyazi@kaust.edu.sa

STC Academy, Riyadh, KSA

23 June to 4 July 2019

Overview

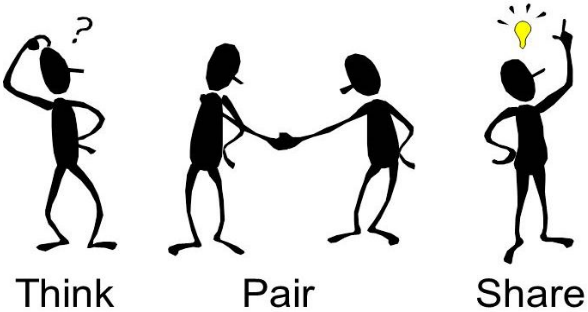
- Week 1: Basic concepts
  - D1 Basic Concepts + Combinatorics
  - D2 Discrete RVs
  - D3 Continuous RVs + Mixed RVs (Prof. Alouini)
  - D4 Simulation/Generation of RVs + Inequalities (Prof. Alouini)
  - D5 Two RVs
- Week 2: Multiple RVs & Stochastic processes
  - D1 Multiple RVs + Multivariate Gaussians (Prof. Alouini)
  - D2 Sum of RVs + Types of Convergence
  - D3 Poisson Process +Branching Process (Prof. Alouini)
  - D4 Birth-Death + Intro to Queuing (Prof. Alouini)
  - D5 Markov Chains
- Week 3: Statistics (Prof. Ombao)
  - Descriptive statistics, Frequentist Statistics, Bayesian Statistics, and Linear Regressions

Lecture Structure

- Today
  - 8:30 AM to 11:30 AM $\Longrightarrow$ Lecture
  - 1:30 AM to 4:30 AM $\Longrightarrow$ Problem solving session
- Starting tomorrow
  - 8:30 AM to 8:45 AM $\Longrightarrow$ Quiz
  - 8:45 AM to 11:30 AM $\Longrightarrow$ Lecture
  - 1:30 AM to 4:30 AM $\Longrightarrow$ Problem solving session

Lecture Structure

- All materials will be provided after 4:30 pm (mailing list)
- Follow the Think-Pair-Share (TPS) strategy (not for quizzes)



Course Objectives (Probability part)

- Understand fundamental concepts and axioms of probability
- Systematic and consistent reasoning about engineering problems that involve uncertainties
  - Random variables
  - Stochastic processes
- Formulate probabilistic models to solve engineering problems
  - Predictions
  - Decision making
  - Performance analysis

References (Probability part)

- Main
  - H. Pishro-Nik. “Introduction to probability, statistics, and random processes,” 2014, available at https://www.probabilitycourse.com/
- Other references
  - P. Z. Peebles and B. Shi, “Probability, Random Variables, and Random Signal Principles”, McGraw-Hill, New 5th Edition, 2015.
  - A. Papoulis and S. Pillai, “Probability, random variables, and stochastic processes,” Tata McGraw-Hill Education, 2002.
  - A. Leon-Garcia, “Probability and Random Processes for EE”, Addison Wesley, 2nd Edition, 1994.
  - S. Ross, “A First Course in Probability”, Prentice Hall, 5th Edition, 1998.
  - H. Kobayashi, B. Mark , and W. Turin. “Probability, random processes, and statistical analysis,” Cambridge University Press, 2011.

Lecture Objectives

- Introduction to probability
- Review on set theory
- Axioms of probability
- Events and Venn diagrams
- Counting and combinatorics

Introduction to Probability

- Is a branch of mathematics that relies on some axioms and uses mathematical argument to find consequences (theorems)

Definition: Random phenomena

Events or experiments whose outcomes we cannot predict with certainty

- Probability can
  - Illustrate relative frequencies
  - Describe beliefs

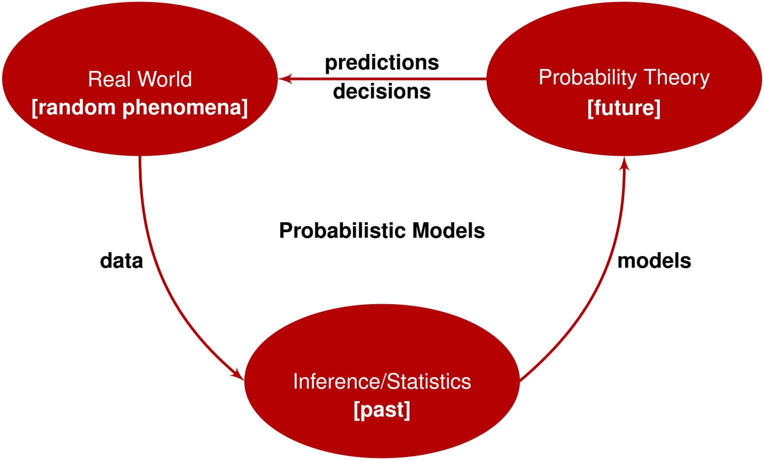
Probability as relative frequency

The number of times an event happens within a very large times of repeated trials

Definition: Personal belief

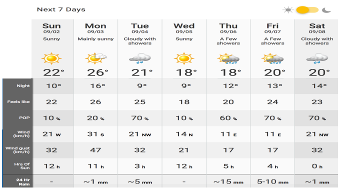
Is a quantification of our belief that an event will happen

Probability, Statistics, and Environment



Applications of Probability Theory

- Weather forecasting
- Insurance companies
- Analysis and design of systems
- Games



Next 7 Days

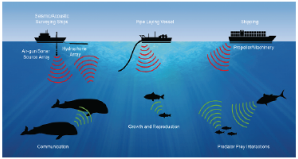
	Sun 09/02 Sunny	Mon 09/03 Mainly sunny	Tue 09/04 Cloudy with showers	Wed 09/05 Sunny	Thu 09/06 A Few showers	Fri 09/07 A Few showers	Sat 09/08 Cloudy with showers
	22°	26°	21°	18°	18°	20°	20°
Night	10°	16°	9°	9°	12°	13°	14°
Feels like	22	26	25	18	20	24	23
POP	10 %	20 %	70 %	10 %	60 %	70 %	70 %
Wind (km/h)	21 W	31 S	21 NW	14 N	11 E	11 E	21 NW
Wind gust (km/h)	32	47	32	21	17	17	32
Hrs Of Sun	12 h	11 h	3 h	12 h	5 h	4 h	0 h
24 Hr Rain	-	~1 mm	~5 mm	-	~15 mm	5-10 mm	~1 mm

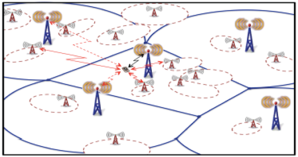
Probability and Signals

Definition: Random Signal

A time varying waveform that can only be characterized in a probabilistic manner.

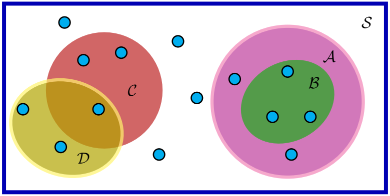
- Examples of random signals are
  - Thermal noise in the output of a receiver
  - Information bits in a communication system
  - Interference in wireless networks
  - Voltage in a solar cell or windmill
  - Sea sounds in a ship/submarine sonar system





Elements of Probabilistic Models

- A probabilistic model is defined through the tuple $(\mathcal{S}, \mathcal{F}, \mathbb{P})$
  - $\mathcal{S}$ is the outcome (sample) space
  - $\mathcal{F}$ is the set of all possible events (subsets of the outcome space)
  - $\mathbb{P}$ is the probability law
- Venn diagram representation



Set Notation

Set notation

A set is a collection of **distinct** elements. Sets are either defined via the "***tabular***" method or the "***rule***" method.

- $\mathcal{A} = \{a, b, c, d\}$
  - $a \in \mathcal{A}$
  - $e \notin \mathcal{A}$
  - $\{a, b\} \subset \mathcal{A}$
  - $\{a\} \subset \mathcal{A}$
- $\mathcal{B} = \{0, 1, 2, 3, 4, \dots\}$
  - $\{1, 3, 5, \dots\} \subset \mathcal{B}$
  - $\mathbb{N} \subset \mathcal{B} \subset \mathbb{R}$
  - $-3 \notin \mathcal{B}$
- $\mathcal{C} = \{x : \text{integers } 1 \leq x \leq 1000\}$
  - $100 \in \mathcal{C}$
  - $1.5 \notin \mathcal{C}$
  - $\mathcal{C} \subset \mathbb{N}$
- $\mathcal{D} = \{x | 0 < x \leq 10\}$
  - $1.005 \in \mathcal{D}$
  - $0 \notin \mathcal{D}$
  - $\{x | 0 < x \leq 1\} \subset \mathcal{D}$
  - $\mathcal{D} \subset \mathbb{R}$
- $\mathcal{E} = \{x | \cos(x) \geq \frac{1}{2}\}$
  - $0 \in \mathcal{C}$
  - $\frac{\pi}{2} \notin \mathcal{C}$
  - $[0, \frac{\pi}{3}] \subset \mathcal{C}$
- $\mathcal{F} = \{(x, y) : x + y \leq \frac{1}{2}\}$
  - $(0, 0) \in \mathcal{D}$
  - $(1, 0) \notin \mathcal{D}$
  - $\{(0, \frac{1}{2}), (\frac{1}{2}, 0), (\frac{1}{3}, \frac{1}{7})\} \subset \mathcal{D}$

Set Notation

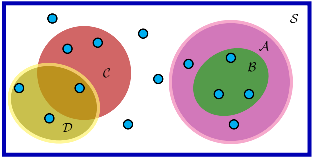
- Set of natural numbers $\mathbb{N} = \{1, 2, 3, 4, \cdots\}$
- Set of integers $\mathbb{Z} = \{\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\}$
- Set of rational numbers $\mathbb{Q} = \{\frac{a}{b} | a, b \in \mathbb{Z}, b \neq 0\}$
- Set of real numbers $\mathbb{R}$
- Set of complex numbers $\mathbb{C} = \{a + bi | a, b \in \mathbb{R}, i = \sqrt{-1}\}$
- The null (empty) set $\phi = \{\}$
- The universal set (outcome space) $\mathcal{S}$
- $\mathcal{A} = [1, 5] \Longrightarrow \mathcal{A} = \{x | 1 \leq x \leq 5\}$
- $\mathcal{A} = (1, 5) \Longrightarrow \mathcal{A} = \{x | 1 < x < 5\}$
- $\mathcal{A} = (1, 5] \Longrightarrow \mathcal{A} = \{x | 1 < x \leq 5\}$
- $\mathcal{A} = \{\clubsuit, \diamondsuit\}$ and $\mathcal{B} = \{\diamondsuit, \clubsuit\} \Longrightarrow \mathcal{A} = \mathcal{B}$

Introduction to Set Theory

- Tossing a coin once
  $\mathcal{S} = \{H, T\}$
- Tossing a coin twice
  $\mathcal{S} = \{HH, HT, TH, TT\}$
- Roll one die
  $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$
- Roll two dice
  $\mathcal{S} = \{(1,1), (1,2), \dots, (2,1), \dots, (6,6)\}$
- Scale with infinite precision
  $\mathcal{S} = \{x : 0 \leq x \leq 360\}$
- Through a dart a target with radius R
  $\mathcal{S} = \{(x, y) : x^2 + y^2 \leq R^2\}$

Set Operations

- Let $\mathcal{S}$ be a universal set
- All other sets are subsets of the universal set
- Venn diagram representation
  - $\mathcal{B}$ is a proper subset of $\mathcal{A} \implies \mathcal{B} \subset \mathcal{A}$
  - $\mathcal{A}$ and $\mathcal{C}$ are mutually exclusive sets

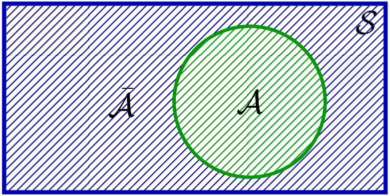


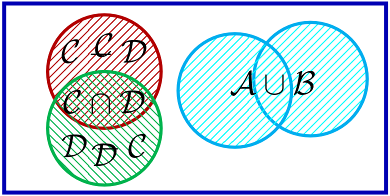
Definition : Power set

A set of n elements can have at most 2^n subsets.

- Let $\mathcal{A} = \{1, 2, 3\}$, then the power set of $\mathcal{A}$ is $\{\phi, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$

- Consider a universal set $\mathcal{S}$
  - $\mathcal{A} \subseteq \mathcal{S}$
  - $\mathcal{A}^c = \bar{\mathcal{A}} = \mathcal{S} - \mathcal{A} = \mathcal{S} \setminus \mathcal{A}$
    $x \in \bar{\mathcal{A}} \Longrightarrow x \in \mathcal{S}$ & $x \notin \mathcal{A}$
  - $\overline{(\bar{\mathcal{A}})} = (\mathcal{A}^c)^c = \mathcal{A}$
  - $\bar{\mathcal{S}} = \phi$ & $\bar{\phi} = \mathcal{S}$
- Unions and intersections
  - $x \in \mathcal{A} \cup \mathcal{B} \Longrightarrow x \in \mathcal{A}$ or $x \in \mathcal{B}$
  - $x \in \mathcal{C} \cap \mathcal{D} \Longrightarrow x \in \mathcal{C}$ & $x \in \mathcal{D}$
  - $x \in \mathcal{C} - \mathcal{D} \Longrightarrow x \in \mathcal{C}$ & $x \notin \mathcal{D}$
  - $x \in \mathcal{D} - \mathcal{C} \Longrightarrow x \in \mathcal{D}$ & $x \notin \mathcal{C}$
  - $\mathcal{C} - \mathcal{D} \neq \mathcal{D} - \mathcal{C}$
- Mutually exclusive (disjoint) sets
  - $\mathcal{A} \cap \mathcal{C} = \phi$





Set Operations

Unions and intersections rules

Let $\mathcal{A}_n$, $n = 1, 2, 3, \ldots$ be subsets of $\mathcal{S}$. Then

- $x \in \bigcup\limits_n \mathcal{A}_n \Longrightarrow x \in \mathcal{A}_n$ for some (at least one) n
- $x \in \bigcap\limits_n \mathcal{A}_n \Longrightarrow x \in \mathcal{A}_n$ for all n
- $\mathcal{A}_n \cup \mathcal{S} = \mathcal{S}$
- $\mathcal{A}_n \cap \mathcal{S} = \mathcal{A}_n$
- $\mathcal{A}_n \cup \bar{\mathcal{A}}_n = \mathcal{S}$
- $\mathcal{A}_n \cap \bar{\mathcal{A}}_n = \phi$
- $\bar{\bar{\mathcal{A}}}_n = \mathcal{A}_n$
- $\mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_n$ are partitions of $\mathcal{S}$ if
  - $\bigcup_{i=1}^n \mathcal{A}_i = \mathcal{S}$
  - $\bigcap_{i=1}^n \mathcal{A}_i = \phi$

Algebra of Sets

Algebra of Sets

- Commutative law
  - $\mathcal{A} \cup \mathcal{B} = \mathcal{B} \cup \mathcal{A}$
  - $\mathcal{A} \cap \mathcal{B} = \mathcal{B} \cap \mathcal{A}$
- Associative law
  - $(\mathcal{A} \cup \mathcal{B}) \cup \mathcal{C} = \mathcal{A} \cup (\mathcal{B} \cup \mathcal{C})$
  - $(\mathcal{A} \cap \mathcal{B}) \cap \mathcal{C} = \mathcal{A} \cap (\mathcal{B} \cap \mathcal{C})$
- Distributive law
  - $\mathcal{A} \cap (\mathcal{B} \cup \mathcal{C}) = (\mathcal{A} \cap \mathcal{B}) \cup (\mathcal{A} \cap \mathcal{C})$
  - $\mathcal{A} \cup (\mathcal{B} \cap \mathcal{C}) = (\mathcal{A} \cup \mathcal{B}) \cap (\mathcal{A} \cup \mathcal{C})$

De Morgan's Law

De Morgan's Law

- Basic form
  - $\overline{(\mathcal{A} \cup \mathcal{B})} = \bar{\mathcal{A}} \cap \bar{\mathcal{B}}$
  - $\overline{(\mathcal{A} \cap \mathcal{B})} = \bar{\mathcal{A}} \cup \bar{\mathcal{B}}$
- General form
  - $\overline{(\bigcup\limits_{n} \mathcal{A}_n)} = \bigcap\limits_{n} \bar{\mathcal{A}}_n$
  - $\overline{(\bigcap\limits_{n} \mathcal{A}_n)} = \bigcup\limits_{n} \bar{\mathcal{A}}_n$

Cartesian Product

Cartesian Product

A Cartesian product of two sets $\mathcal{A}$ and $\mathcal{B}$, denoted as $\mathcal{A} \times \mathcal{B}$, is the set containing all **ordered** pairs from $\mathcal{A}$ and $\mathcal{B}$.

$$\mathcal{A} \times \mathcal{B} = \{(x, y) | x \in \mathcal{A} \text{ and } y \in \mathcal{B}\}$$

- Example
  - Let $\mathcal{A} = \{1, 2, 3, 4, 5, 6\}$ and $\mathcal{B} = \{H, T\}$ then

$$\begin{aligned}\mathcal{A} \times \mathcal{B} = \{&(1, H), (2, H), (3, H), (4, H), (5, H), (6, H),\\ &(1, T), (2, T), (3, T), (4, T), (5, T), (6, T)\}\end{aligned}$$

  - The Euclidian plane can be constructed as $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$
  - For $\mathcal{A} \neq \mathcal{B} \Longrightarrow \mathcal{A} \times \mathcal{B} \neq \mathcal{B} \times \mathcal{A}$

Multiplication principle

Let $| \cdot |$ denote the **cardinality** (i.e., size) of a set and let $|\mathcal{A}| = N$ and $|\mathcal{B}| = M$, then $|\mathcal{A} \times \mathcal{B}| = NM$

Cardinality

Definition 1: Cardinality

The cardinality denotes the size of a set. Sets could be

- Finite $\Longrightarrow$ Discrete
- Infinite
  - Countable $\Longrightarrow$ Discrete
  - Uncountable $\Longrightarrow$ Continuous

- For finite sets, the cardinality is the number of elements in the set
- For two sets $\mathcal{A}$ and $\mathcal{B}$, the cardinality of $|\mathcal{A} \cup \mathcal{B}|$ is given by $|\mathcal{A} \cup \mathcal{B}| = |\mathcal{A}| + |\mathcal{B}| - |\mathcal{A} \cap \mathcal{B}|$
- For three sets $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$, the cardinality of $|\mathcal{A} \cup \mathcal{B} \cup \mathcal{C}|$ is given by $|\mathcal{A} \cup \mathcal{B} \cup \mathcal{C}| = |\mathcal{A}| + |\mathcal{B}| + |\mathcal{C}| - |\mathcal{A} \cap \mathcal{B}| - |\mathcal{A} \cap \mathcal{C}| - |\mathcal{B} \cap \mathcal{C}| + |\mathcal{A} \cap \mathcal{B} \cap \mathcal{C}|$

Cardinality

Definition 1: Inclusion exclusion rule

- For n finite sets, we can write

$$\left|\bigcup_{i=1}^{n} \mathcal{A}_i\right| = \sum_{i=1}^{n} |\mathcal{A}_i| - \sum_{i<j} |\mathcal{A}_i \cap \mathcal{A}_j| + \sum_{i<j<k} |\mathcal{A}_i \cap \mathcal{A}_j \cap \mathcal{A}_k|$$
$$- \cdots + (-1)^{n+1} |\mathcal{A}_i \cap \mathcal{A}_j \cdots \cap \mathcal{A}_n|$$

- Example: consider an event where
  - 10 have white shirts
  - 8 have red shirts
  - 4 have black shoes and white shirts
  - 3 have back shoes and red shirts
  - 21 have white shirts, red shirts, or back shoes
- Find the number of people with black shoes (answer is10)

Cardinality of Infinite sets

Definition : Countable Set

A set is countable if it can be put in one-to-one correspondence with the natural numbers $\mathbb{N} = \{1, 2, 3, 4, 5, \dots\}$

- Discrete sets can be countably infinite.
- Examples of countably infinite sets
  - Natural numbers
    $\{1, 2, 3, 4, 5, \dots\}$
  - Odd and even numbers

$$\begin{matrix} \{ & 1, & 3, & 5, & 7, & 9, & \dots\} \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \{ & 1, & 2, & 3, & 4, & 5, & \dots\} \end{matrix} \qquad \begin{matrix} \{ & 2, & 4, & 6, & 8, & 10, & \dots\} \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \{ & 1, & 2, & 3, & 4, & 5, & \dots\} \end{matrix}$$

  - Integers $\mathbb{Z}$

$$\begin{matrix} \{ & 0, & 1, & -1, & 2, & -2, & \dots\} \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \{ & 1, & 2, & 3, & 4, & 5, & \dots\} \end{matrix}$$

Discrete and Continuous Sets

- Examples of countably infinite sets
  - Rational numbers less than 1

$$\begin{array}{cccccccc} \{ & \frac{1}{2}, & \frac{1}{3}, & \frac{2}{3}, & \frac{1}{4}, & \frac{3}{4}, & \frac{1}{5}, & \ldots\} \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \{ & 1, & 2, & 3, & 4, & 5, & 6, & \ldots\} \end{array}$$

  - Pairs of positive integers

$$\begin{array}{cccccccc} \{ & (0,0), & (1,0), & (0,1), & (0,2), & (1,1) & (2,0), & \ldots\} \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \{ & 1, & 2, & 3, & 4, & 5, & 6, & \ldots\} \end{array}$$

- Examples of continuous (uncountable) sets
  - Real numbers $\mathbb{R}$
  - Unit interval [0,1]
  - Euclidean space (1D, 2D, ...)

Cardinality: Useful results

Theorems

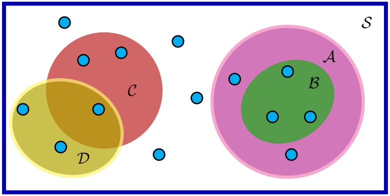
- Any subset of countable set is countable
- Any superset of uncountable set is uncountable
- Let $\mathcal{A}_1, \mathcal{A}_2, \cdots$ be a list of countable sets then $\bigcup_i \mathcal{A}_i$ is countable
- If $\mathcal{A}$ and $\mathcal{B}$ are countable, then $\mathcal{A} \times \mathcal{B}$ is also countable

Progress...

- Last section
  - Set Operations
  - De Morgan’s Law
  - Cartesian product
  - Set cardinality
- Next section
  - Random Experiments
  - Axioms of Probability

Random Experiment

- Elements of a random experiment
  - Sample space with elements that are
    - Mutually exclusive
    - Collectively exhaustive
  - Events
    - Subsets of the sample space
  - Probability law
    - Assigns non-negative number for each event



Random Experiment

Definitions

- Trial, realization, or instantiation: performance of a random experiment
- Outcome: A result of a random experiment.
- Sample Space: The set of all possible outcomes.
- Event: A subset of the sample space

- Tossing a coin once
  $\mathcal{S} = \{H, T\}$
- Tossing a coin twice
  $\mathcal{S} = \{HH, HT, TH, TT\}$
- Roll one die
  $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$
- Roll two dice
  $\mathcal{S} = \{(1, 1), (1, 2), \dots, (2, 1), \dots, (6, 6)\}$
- Scale with infinite precision
  $\mathcal{S} = \{x : 0 \leq x \leq 360\}$
- Through a dart a target with radius R
  $\mathcal{S} = \{(x, y) : x^2 + y^2 \leq R^2\}$

Events

- Consider the experiment of rolling a die
- The outcome space is $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$
- Events are defined in terms of subsets of the outcome space.
- The event of getting an odd number $E_1 = \{1, 3, 5\}$ occurs if the outcome of the experiment is 1, 3, or 5.
- The event of getting a number greater than 3 is $E_2 = \{4, 5, 6\}$ occurs if the outcome is 4, 5, or 6.
- the event $E_3 = E_1 \cup E_2 = \{1, 3, 4, 5, 6\}$ occurs if E_1 **or** E_2 occur. That is, the outcome is 1, 3, 4, 5, or 6.
- The event $E_4 = E_1 \cap E_2 = \{5\}$ occurs if the E_1 **and** E_2 occurs. That is, the outcome is 5.

Definition : Union and intersection of events

Consider the events $E_1, E_2, \cdots, E_n$, then

- The event $\bigcup_{i=1}^{n} E_i$ occurs if **at least** one of the events E_i, $1 \leq i \leq n$, occurs.
- The event $\bigcap_{i=1}^{n} E_i$ occurs if **all of** of the events E_i, $1 \leq i \leq n$, occur.

Probability

- We assign a probability measure between 0 and 1 to events.
- The probability describes the likelihood that an event to occur.

Axioms of Probability

- Axiom 1: (non-negativity)
  $\mathbb{P}(\mathcal{A}) \geq 0$
- Axiom 2: (normalization)
  $\mathbb{P}(\mathcal{S}) = 1$
- Axiom 3: (Countable additivity)
  if $\bigcap\limits_{n} \mathcal{A}_n = \phi \Longrightarrow \mathbb{P}(\bigcup\limits_{n} \mathcal{A}_n) = \sum\limits_{n} \mathbb{P}(\mathcal{A}_n)$

- Some consequences of probability axioms
  - $\mathbb{P}(\mathcal{A}) \leq 1$
  - $\mathbb{P}(\phi) = 0$
  - if $\mathcal{B} \subset \mathcal{A} \Longrightarrow \mathbb{P}(\mathcal{B}) \leq \mathbb{P}(\mathcal{A})$
  - $\mathbb{P}(\mathcal{A} \cup \mathcal{B}) = \mathbb{P}(\mathcal{A}) + \mathbb{P}(\mathcal{B}) - \mathbb{P}(\mathcal{A} \cap \mathcal{B})$

Consequences of Probability Axioms

- $\mathbb{P}(\mathcal{A}) \leq 1$

$$
\begin{array}{rccc}
\overset{(a2)}{\Longrightarrow} & \mathbb{P}(\mathcal{S}) & = & 1 \\
 & \mathbb{P}(\mathcal{A} \cup \bar{\mathcal{A}}) & = & 1 \\
\overset{(a3)}{\Longrightarrow} & \mathbb{P}(\mathcal{A}) + \mathbb{P}(\bar{\mathcal{A}}) & = & 1 \\
 & \mathbb{P}(\mathcal{A}) & = & 1 - \mathbb{P}(\bar{\mathcal{A}}) \\
\overset{(a1)}{\Longrightarrow} & \mathbb{P}(\mathcal{A}) & \leq & 1
\end{array}
$$

- $\mathbb{P}(\phi) = 0$

$$
\begin{array}{rccc}
\overset{(a2)}{\Longrightarrow} & \mathbb{P}(\mathcal{S}) & = & 1 \\
 & \mathbb{P}(\mathcal{S} \cup \bar{\mathcal{S}}) & = & 1 \\
 & \mathbb{P}(\mathcal{S} \cup \phi) & = & 1 \\
\overset{(a3)}{\Longrightarrow} & \mathbb{P}(\mathcal{S}) + \mathbb{P}(\phi) & = & 1 \\
\overset{(a2)}{\Longrightarrow} & 1 + \mathbb{P}(\phi) & = & 1 \\
\Longrightarrow & \mathbb{P}(\phi) & = & 0
\end{array}
$$

Consequences of Probability Axioms

- if $\mathcal{B} \subset \mathcal{A} \Longrightarrow \mathbb{P}(\mathcal{B}) \leq \mathbb{P}(\mathcal{A})$

$$\begin{array}{rcl} \mathbb{P}(\mathcal{A}) & = & \mathbb{P}(\mathcal{B} \cup \{\mathcal{A} - \mathcal{B}\}) \\ \mathbb{P}(\mathcal{A}) & = & \mathbb{P}(\mathcal{B}) + \mathbb{P}(\mathcal{A} - \mathcal{B}) \\ \mathbb{P}(\mathcal{A}) - \mathbb{P}(\mathcal{A} - \mathcal{B}) & = & \mathbb{P}(\mathcal{B}) \\ \overset{(a1)}{\Longrightarrow} \quad \mathbb{P}(\mathcal{A}) & \geq & \mathbb{P}(\mathcal{B}) \end{array}$$

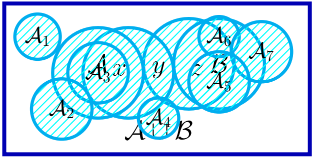
Consequences of probability axioms

- $\mathbb{P}(\mathcal{A} \cup \mathcal{B}) = \mathbb{P}(\mathcal{A}) + \mathbb{P}(\mathcal{B}) - \mathbb{P}(\mathcal{A} \cap \mathcal{B})$
  Let $x = \mathcal{A} - \mathcal{B}$, $z = \mathcal{B} - \mathcal{A}$, and $y = \mathcal{A} \cap \mathcal{B}$, then

$$\begin{aligned}
\mathbb{P}(\mathcal{A} \cup \mathcal{B}) &= \mathbb{P}(x \cup y \cup z) \\
\mathbb{P}(\mathcal{A} \cup \mathcal{B}) &= \mathbb{P}(x) + \mathbb{P}(y) + \mathbb{P}(z) \\
\mathbb{P}(\mathcal{A} \cup \mathcal{B}) &= \mathbb{P}(x) + \mathbb{P}(y) + \mathbb{P}(z) + \mathbb{P}(y) - \mathbb{P}(y) \\
\mathbb{P}(\mathcal{A} \cup \mathcal{B}) &= \mathbb{P}(x \cup y) + \mathbb{P}(z \cup y) - \mathbb{P}(y) \\
\mathbb{P}(\mathcal{A} \cup \mathcal{B}) &= \mathbb{P}(\mathcal{A}) + \mathbb{P}(\mathcal{B}) - \mathbb{P}(\mathcal{A} \cap \mathcal{B})
\end{aligned}$$

  if $\mathcal{A} \cap \mathcal{B} = \phi \Longrightarrow \mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\phi) = 0$
  $\Longrightarrow \mathbb{P}(\mathcal{A} \cup \mathcal{B}) = \mathbb{P}(\mathcal{A}) + \mathbb{P}(\mathcal{B})$ (Axiom 3)

- Union bound:
  $\mathbb{P}(\bigcup_n \mathcal{A}_n) \leq \sum_n \mathbb{P}(\mathcal{A}_n)$



Example

Remember

- $\mathbb{P}(\mathcal{S}) = 1$
- $\mathbb{P}(\mathcal{A}) + \mathbb{P}(\bar{\mathcal{A}}) = 1$
- $\mathbb{P}(\mathcal{A} \cup \mathcal{B}) = \mathbb{P}(\mathcal{A}) + \mathbb{P}(\mathcal{B}) - \mathbb{P}(\mathcal{A} \cap \mathcal{B})$

- The weather forecast states the following
  - The probability that it rains today is 60%
  - The probability that it rains tomorrow is 50%
  - The probability it does not rain either today or tomorrow is 30%
- Find
  - The probability that it rains today or tomorrow
  - The probability that it will rain today and tomorrow
  - The probability that it will rain today but not tomorrow
  - The probability that it will either rain today or tomorrow but not both

Discrete Probability Models

- The outcome space is countable and can be listed as $\mathcal{S} = \{s_1, s_2, s_3, \cdots\}$.
- Using axiom 3, the probability of any event $\mathcal{A} \subset \mathcal{S}$ can be obtained as

$$\mathbb{P}(\mathcal{A}) = \mathbb{P}\left(\bigcup_{s_i \in \mathcal{A}} \{s_i\}\right) = \sum_{s_i \in \mathcal{A}} \mathbb{P}(s_i)$$

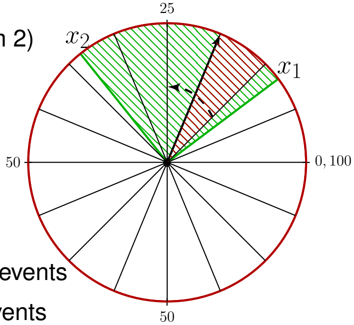
- Example: a player consider a game where you win (k-2)$ with probability $\frac{1}{2^k}$, where $k \in \mathbb{N}$. Find
  - The probability to win more than 3$
  - The probability to lose 1$
- In case of finite probability space with equally likely outcomes, we have

$$\mathbb{P}(\mathcal{A}) = \frac{|\mathcal{A}|}{|\mathcal{S}|}$$

- Example: find the probability of getting an even number when rolling a fair die

Continuous Probability Models

- The outcome space is uncountable and cannot be listed, and hence, summations should be substituted by interval integrals.
- Example: rotating a fair wheel
- $\mathcal{S} = \{0 \leq x \leq 100\}$
- $\mathbb{P}(x_1 < x < x_2) = \frac{x_2 - x_1}{100}$ (satisfies Axiom 1)
- Setting $x_2 = 100$ and $x_1 = 0$ (satisfies Axiom 2)
- Break the scale into N equal segments
- Let $x_n = \frac{(n)100}{N}$, where $n = 1, 2, 3, \ldots, N$
- Define $A_n = \{x_{n-1} < x < x_n\}$
  Then $\mathbb{P}(A_n) = \frac{1}{N}$
  $$\mathbb{P}(\bigcup_{n=1}^{N} A_n) = \sum_{n=1}^{N} \mathbb{P}(A_n)$$
- Events with zero probabilities vs impossible events
- Events with probability one vs must occur events



Solving Probability Problems

Steps to Solve Probability Problems

1. Specify the sample space
2. Identify the event of interest
3. Specify the probability law
4. Calculate the probability

Solving Probability Problems

Coin tossing until first head

- What is:
  - The probability that the first H comes after even number of coin tossing
  - The probability that the first H comes after 5 coin tossing

- Outcome space: $\mathcal{S} = \{n \in \mathbb{N}\}$
- Probability law: $\mathbb{P}(n) = \frac{1}{2^n}$
- Events of interest are
  - $n \in \{2, 4, 6, \ldots\}$
  - $n < 5$

Solving Probability Problems

Sum of two dice

- What is:
  - The probability that the output is equal to 7
  - The probability that the output is equal to 1
  - The probability that the output is greater than 4

- Outcome space: $\mathcal{S} = \{2, 3, 4, \ldots, 12\}$

$$\begin{array}{cccccc}
(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\
(2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6)
\end{array}$$

- Probability law
  $\mathbb{P}(2) = \mathbb{P}(12) = 1 \times \frac{1}{36}$
  $\mathbb{P}(3) = \mathbb{P}(11) = 2 \times \frac{1}{36} = \frac{1}{18}$
  $\mathbb{P}(4) = \mathbb{P}(10) = 3 \times \frac{1}{36} = \frac{1}{12}$
- Events of interest are
  - $\mathbb{P}(7) = \frac{1}{6}$, $\quad \mathbb{P}(1) = 0$, $\quad$ and $\quad \mathbb{P}(n > 4) = \frac{15}{18}$

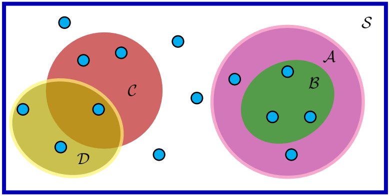
Progress...

- Last section
  - Random experiment
  - Axioms of probability
  - Discrete probability models
  - Continuous probability models
- Next section
  - Joint probability
  - Conditional probability
  - Total probability

Types of Probabilities

Marginal, joint, and conditional probabilities

- Marginal probability: $\mathbb{P}(\mathcal{A})$ ✓
- Joint probability: $\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\mathcal{A}) + \mathbb{P}(\mathcal{B}) - \mathbb{P}(\mathcal{A} \cup \mathcal{B})$
- Conditional Probability: $\mathbb{P}(\mathcal{A}|\mathcal{B})$?



Conditional Probability

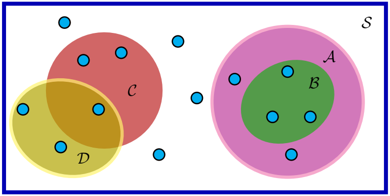
Conditional Probability

Provides a systematic way to incorporate new acquired knowledge into a probabilistic model.

Notation

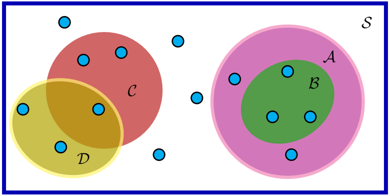
The probability that event $\mathcal{A}$ ***occurs*** given that event $\mathcal{B}$ ***occured***

$$\mathbb{P}\{\mathcal{A}|\mathcal{B}\}$$



Example

- Consider an outcome space $\mathcal{S}$ with n equiprobable elements $\mathcal{S} = \{x_1, x_2, \ldots, x_n\}$
- $\mathbb{P}(\{x_i\}) = \frac{1}{n}$
- $\mathbb{P}\{\mathcal{A}\} = \frac{5}{14}$ $\quad \mathbb{P}\{\mathcal{A}|\mathcal{B}\} = 1$ $\quad \mathbb{P}\{\mathcal{A}|\mathcal{C}\} = 0$
- $\mathbb{P}\{\mathcal{B}\} = \frac{3}{14}$ $\quad \mathbb{P}\{\mathcal{B}|\mathcal{A}\} = \frac{3}{5}$ $\quad \mathbb{P}\{\mathcal{B}|\mathcal{D}\} = 0$
- $\mathbb{P}\{\mathcal{C}\} = \frac{3}{14}$ $\quad \mathbb{P}\{\mathcal{C}|\mathcal{D}\} = \frac{1}{3}$ $\quad \mathbb{P}\{\mathcal{C}|\mathcal{A}\} = 0$



Conditional Probability Law

Conditional Probability Law

- The conditional probability is defined as
  - Assuming that $\mathbb{P}(\mathcal{B}) > 0$

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

  - Assuming that $\mathbb{P}(\mathcal{A}) > 0$

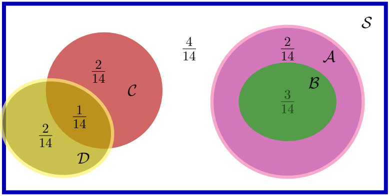
$$P(\mathcal{B}|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{B})}{\mathbb{P}(\mathcal{A})}$$

Example

Conditional Probability Law

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

- Consider an outcome space $\mathcal{S}$ with the shown probability law
- $\mathbb{P}\{\mathcal{A}\} = \frac{5}{14}$ $\quad \mathbb{P}\{\mathcal{A}|\mathcal{B}\} = \frac{\frac{3}{14}}{\frac{3}{14}} = 1$ $\quad \mathbb{P}\{\mathcal{A}|\mathcal{C}\} = \frac{0}{\frac{3}{14}} = 0$
- $\mathbb{P}\{\mathcal{B}\} = \frac{3}{14}$ $\quad \mathbb{P}\{\mathcal{B}|\mathcal{A}\} = \frac{\frac{3}{14}}{\frac{5}{14}} = \frac{3}{5}$ $\quad \mathbb{P}\{\mathcal{B}|\mathcal{D}\} = \frac{0}{\frac{3}{14}}$
- $\mathbb{P}\{\mathcal{C}\} = \frac{3}{14}$ $\quad \mathbb{P}\{\mathcal{C}|\mathcal{D}\} = \frac{\frac{1}{14}}{\frac{3}{14}} = \frac{1}{3}$ $\quad \mathbb{P}\{\mathcal{C}|\mathcal{A}\} = \frac{0}{\frac{5}{14}} = 0$



Axioms of Probability

Axioms of Probability

Conditional probability satisfies the axioms of probability

- Axiom 1: (non-negativity)
  $\mathbb{P}(\mathcal{A}|\mathcal{B}) \geq 0$
- Axiom 2: (normalization)
  $\mathbb{P}(\mathcal{S}|\mathcal{B}) = 1$
  $\mathbb{P}(\mathcal{B}|\mathcal{B}) = 1$
- Axiom 3: (additivity) can be generalized to countable additivity
  if $\{\mathcal{A} \cap \mathcal{B}\} \cap \{\mathcal{C} \cap \mathcal{B}\} = \phi \Longrightarrow \mathbb{P}(\mathcal{A} \cup \mathcal{C}|\mathcal{B}) = \mathbb{P}(\mathcal{A}|\mathcal{B}) + \mathbb{P}(\mathcal{C}|\mathcal{B})$

- Proof of axiom 3: consider the two disjoint sets $\mathcal{A}$ and $\mathcal{C}$, then

$$\begin{aligned}
\mathbb{P}(\mathcal{A} \cup \mathcal{C}|\mathcal{B}) &= \frac{\{\mathbb{P}(\mathcal{A} \cup \mathcal{C}\} \cap \mathcal{B})}{\mathbb{P}(\mathcal{B})} = \frac{\mathbb{P}(\{\mathcal{A} \cap \mathcal{B}\} \cup \{\mathcal{C} \cap \mathcal{B}\})}{\mathbb{P}(\mathcal{B})} \\
&= \frac{\mathbb{P}(\{\mathcal{A} \cap \mathcal{B}\}) + \mathbb{P}(\{\mathcal{C} \cap \mathcal{B})\}}{\mathbb{P}(\mathcal{B})} = \frac{\mathbb{P}(\{\mathcal{A} \cap \mathcal{B}\})}{\mathbb{P}(\mathcal{B})} + \frac{\mathbb{P}(\{\mathcal{C} \cap \mathcal{B})\}}{\mathbb{P}(\mathcal{B})} \\
&= \mathbb{P}(\mathcal{A}|\mathcal{B}) + \mathbb{P}(\mathcal{C}|\mathcal{B})
\end{aligned}$$

Solving Probability Problems

Rolling of two dice

- Given that event $\mathcal{B}$, defined as $\min(x, y) = 2$, has occurred
- Find:
  - $\mathbb{P}(\max(x, y) = 1)$
  - $\mathbb{P}(\max(x, y) = 4)$
  - $\mathbb{P}(\max(x, y) > 4)$

- Outcome space is

$$
\begin{array}{cccccc}
(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\
(2,1) & {\color{red}(2,2)} & {\color{red}(2,3)} & {\color{red}(2,4)} & {\color{red}(2,5)} & {\color{red}(2,6)} \\
(3,1) & {\color{red}(3,2)} & (3,3) & (3,4) & (3,5) & (3,6) \\
(4,1) & {\color{red}(4,2)} & (4,3) & (4,4) & (4,5) & (4,6) \\
(5,1) & {\color{red}(5,2)} & (5,3) & (5,4) & (5,5) & (5,6) \\
(6,1) & {\color{red}(6,2)} & (6,3) & (6,4) & (6,5) & (6,6)
\end{array}
$$

Example

Rolling of two dice

- Given that event $\mathcal{B} \to \min(x, y) = 2$ occurred, what is:
  - $\mathbb{P}(\max(x, y) = 1) = \frac{0}{\frac{9}{36}}$
  - $\mathbb{P}(\max(x, y) = 4)$
  - $\mathbb{P}(\max(x, y) > 4)$

- Outcome space is

(1,1)	$(1, 2)$	$(1, 3)$	$(1, 4)$	$(1, 5)$	$(1, 6)$
$(2, 1)$	$(2, 2)$	$(2, 3)$	$(2, 4)$	$(2, 5)$	$(2, 6)$
$(3, 1)$	$(3, 2)$	$(3, 3)$	$(3, 4)$	$(3, 5)$	$(3, 6)$
$(4, 1)$	$(4, 2)$	$(4, 3)$	$(4, 4)$	$(4, 5)$	$(4, 6)$
$(5, 1)$	$(5, 2)$	$(5, 3)$	$(5, 4)$	$(5, 5)$	$(5, 6)$
$(6, 1)$	$(6, 2)$	$(6, 3)$	$(6, 4)$	$(6, 5)$	$(6, 6)$

Example

Rolling of two dice

- Given that event $\mathcal{B} \to \min(x,y) = 2$ occurred, what is:
  - $\mathbb{P}(\max(x,y) = 1) = 0$
  - $\mathbb{P}(\max(x,y) = 4) = \frac{\frac{2}{36}}{\frac{9}{36}}$
  - $\mathbb{P}(\max(x,y) > 4)$

- Outcome space is

$(1,1)$	$(1,2)$	$(1,3)$	**(1,4)**	$(1,5)$	$(1,6)$
$(2,1)$	$(2,2)$	$(2,3)$	**(2,4)**	$(2,5)$	$(2,6)$
$(3,1)$	$(3,2)$	$(3,3)$	**(3,4)**	$(3,5)$	$(3,6)$
(4,1)	**(4,2)**	**(4,3)**	**(4,4)**	$(4,5)$	$(4,6)$
$(5,1)$	$(5,2)$	$(5,3)$	$(5,4)$	$(5,5)$	$(5,6)$
$(6,1)$	$(6,2)$	$(6,3)$	$(6,4)$	$(6,5)$	$(6,6)$

Example

Rolling of two dice

- Given that event $\mathcal{B} \rightarrow \min(x, y) = 2$ occurred, what is:
  - $\mathbb{P}(\max(x, y) = 1) = 0$
  - $\mathbb{P}(\max(x, y) = 4) = \frac{2}{9}$
  - $\mathbb{P}(\max(x, y) > 4) = \frac{\frac{4}{36}}{\frac{9}{36}} = \frac{4}{9}$

- Outcome space is

$$
\begin{array}{cccccc}
(1,1) & (1,2) & (1,3) & (1,4) & \textbf{(1,5)} & \textbf{(1,6)} \\
(2,1) & (2,2) & (2,3) & (2,4) & \textbf{(2,5)} & \textbf{(2,6)} \\
(3,1) & (3,2) & (3,3) & (3,4) & \textbf{(3,5)} & \textbf{(3,6)} \\
(4,1) & (4,2) & (4,3) & (4,4) & \textbf{(4,5)} & \textbf{(4,6)} \\
\textbf{(5,1)} & \textbf{(5,2)} & \textbf{(5,3)} & \textbf{(5,4)} & \textbf{(5,5)} & \textbf{(5,6)} \\
\textbf{(6,1)} & \textbf{(6,2)} & \textbf{(6,3)} & \textbf{(6,4)} & \textbf{(6,5)} & \textbf{(6,6)}
\end{array}
$$

Conditional Probabilities

Conditional Probabilities

- Conditional probabilities are used to incorporate new information to a probabilistic model and update the probability law
- Construct multistage model of a probabilistic experiment (chain rule)

Example

Example

Consider a region where airplanes fly with probability 5%. An airplane radar system, with the aspects listed below, is deployed in that region.

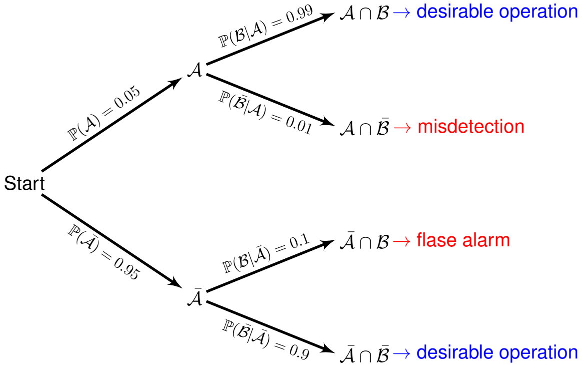
- Radar Aspects
  - Detection probability 99%
  - False alarm probability 10%
- Find:
  - The probability that an airplane flies over the region and the radar screen shows it
  - The probability that the radar screen shows an airplane
  - The probability that there is an airplane given that the radar screen shows an airplane
  - The probability of faulty radar operation (i.e., error)

Solution

- We first define the following two events
  - Event $\mathcal{A}$: an airplane flies over the region
  - Event $\mathcal{B}$: the radar screen shows an object
- This will lead to the following four outcomes
  1. An airplane flies over the region and the radar screen shows it $\mathcal{A} \cap \mathcal{B}$ (desirable)
  2. An airplane flies over the region and the radar screen shows nothing $\mathcal{A} \cap \bar{\mathcal{B}}$ (undesirable [misdetection])
  3. No airplane flies over the region and the radar screen shows an object $\bar{\mathcal{A}} \cap \mathcal{B}$ (undesirable [false alarm])
  4. No airplane flies over the region and the radar screen shows nothing $\bar{\mathcal{A}} \cap \bar{\mathcal{B}}$ (desirable)

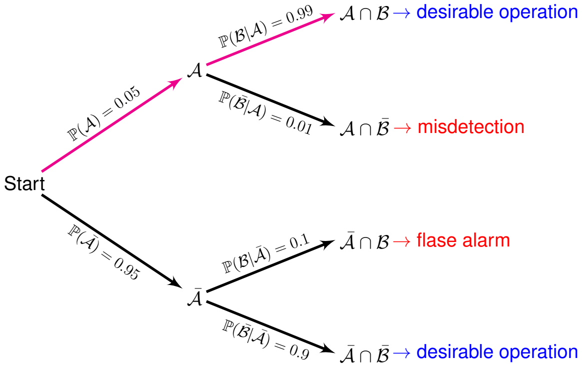
Solution

- Use the conditional probability law to build the following tree diagram



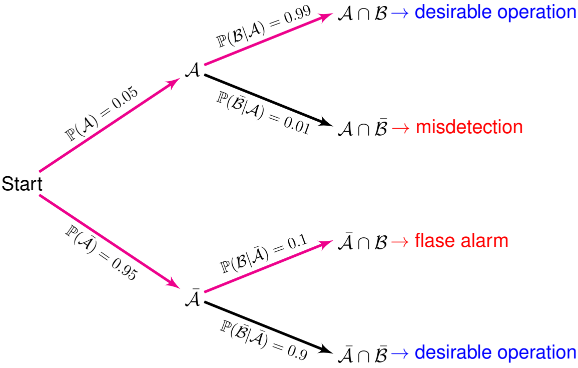
Solution

- The probability that an airplane flies over the region (i.e., event $\mathcal{A}$) and the radar screen shows it (i.e.,event $\mathcal{B}$) is
  $\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B}|\mathcal{A}) = 0.05 \times 0.99 = 0.0495$



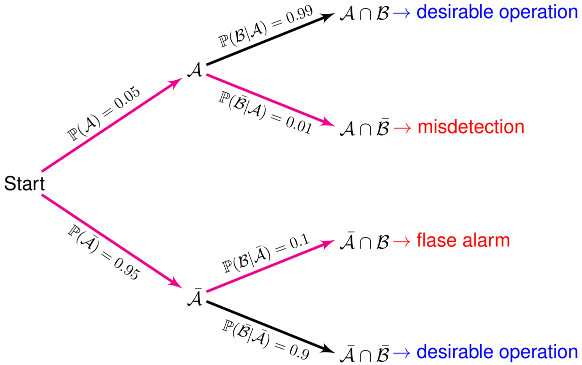
Solution

- The probability that the radar screen shows an object (i.e., event $\mathcal{B}$) is $\mathbb{P}(\mathcal{B}) = \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B}|\mathcal{A}) + \mathbb{P}(\bar{\mathcal{A}})\mathbb{P}(\mathcal{B}|\bar{\mathcal{A}}) = 0.05 \times 0.99 + 0.95 \times 0.1 = 0.1445$



Solution

- The probability that there is an airplane (i.e., event $\mathcal{A}$) given that the radar screen shows an airplane (i.e., event $\mathcal{B}$) is $\mathbb{P}(\mathcal{A}|\mathcal{B}) = \frac{\mathbb{P}(\mathcal{A}\cap\mathcal{B})}{\mathbb{P}(\mathcal{B})} = \frac{0.0495}{0.1445} = 0.342$
- The probability of faulty radar operation $\mathbb{P}(\text{error}) = \mathbb{P}(\mathcal{A})\mathbb{P}(\bar{\mathcal{B}}|\mathcal{A}) + \mathbb{P}(\bar{\mathcal{A}})\mathbb{P}(\mathcal{B}|\bar{\mathcal{A}}) = 0.05\times0.01 + 0.95\times0.1 = 0.0955$



Chain rule

Chain rule

- $\mathbb{P}(\mathcal{A}|\mathcal{B}) = \frac{\mathbb{P}(\mathcal{A}\cap\mathcal{B})}{\mathbb{P}(\mathcal{B})}$
- $\mathbb{P}(\mathcal{A}\cap\mathcal{B}) = \mathbb{P}(\mathcal{B})\mathbb{P}(\mathcal{A}|\mathcal{B})$
- $\mathbb{P}(\mathcal{A}\cap\mathcal{B}) = \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B}|\mathcal{A})$
- $\mathbb{P}(\mathcal{A}\cap\mathcal{B}\cap\mathcal{C}) = \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B}|\mathcal{A})\mathbb{P}(\mathcal{C}|\mathcal{A}\cap\mathcal{B})$
- General form

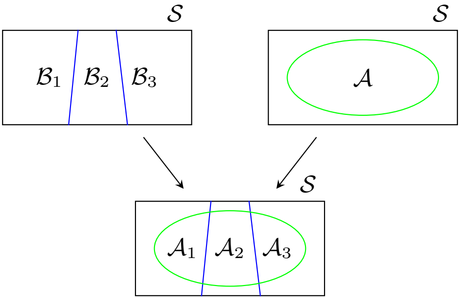
$$\mathbb{P}(\bigcap_{i=1}^{n}\mathcal{A}_i) = \mathbb{P}(\mathcal{A}_1)\prod_{i=2}^{n}\mathbb{P}(\mathcal{A}_i|\mathcal{A}_1\cap\mathcal{A}_1\cap\cdots\cap\mathcal{A}_{i-1})$$

Law of Total Probability

Law of total probability

Consider that $\mathcal{S}$ is divided into N subregions denoted as $\mathcal{B}_n$, $n = 1, 2, \ldots, N$ such that $\mathcal{B}_n \cap \mathcal{B}_m = \phi$ for $n \neq m$ and $\bigcup_{n=1}^{N} \mathcal{B}_n = \mathcal{S}$.
Then, the probability of an event $\mathcal{A}$ can be written as

$$\mathbb{P}(\mathcal{A}) = \sum_{n=1}^{N} \mathbb{P}(\mathcal{B}_n)\mathbb{P}(\mathcal{A}|\mathcal{B}_n)$$



Progress...

- Last section
  - Joint probability
  - Conditional probability
  - Total probability
- Next section
  - Bayes' theorem
  - Independent events

Bayes' Theorem

Bayes' Theorem

- Let $\mathcal{B}_n$, $n = 1, 2, 3, \ldots$ be a set of disjoint event such that $\bigcup_n \mathcal{B}_n = \mathcal{S}$ and $\mathbb{P}(\mathcal{B}_n) > 0 \; \forall n$. Also, let $\mathcal{A} \subset \mathcal{S}$ be an event such that $\mathbb{P}(\mathcal{A}) > 0$, then

$$P(\mathcal{A}|\mathcal{B}_n) = \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{B}_n)}{\mathbb{P}(\mathcal{B}_n)} \quad \Longrightarrow \quad \mathbb{P}(\mathcal{A} \cap \mathcal{B}_n) = P(\mathcal{A}|\mathcal{B}_n)\mathbb{P}(\mathcal{B}_n)$$

- It is also true that

$$P(\mathcal{B}_n|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{B}_n \cap \mathcal{A})}{\mathbb{P}(\mathcal{A})} \quad \Longrightarrow \quad P(\mathcal{B}_n|\mathcal{A}) = \frac{P(\mathcal{A}|\mathcal{B}_n)\mathbb{P}(\mathcal{B}_n)}{\mathbb{P}(\mathcal{A})}$$

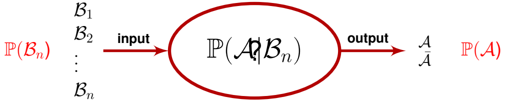
- Using the law of total probability $\mathbb{P}(\mathcal{A}) = \sum_n \mathbb{P}(\mathcal{A} \cap \mathcal{B}_n)$

$$\mathbb{P}(\mathcal{B}_n|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A}|\mathcal{B}_n)\mathbb{P}(\mathcal{B}_n)}{\mathbb{P}(\mathcal{B}_1)\mathbb{P}(\mathcal{A}|\mathcal{B}_1) + \mathbb{P}(\mathcal{B}_2)\mathbb{P}(\mathcal{A}|\mathcal{B}_2) + \cdots + \mathbb{P}(\mathcal{B}_n)\mathbb{P}(\mathcal{A}|\mathcal{B}_n)}$$

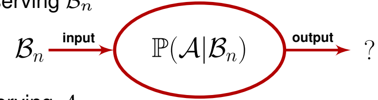
Bayes' Theorem

$$\mathbb{P}(\mathcal{A}|\mathcal{B}_n) = \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{B}_n)}{\mathbb{P}(\mathcal{B}_n)}, \quad \mathbb{P}(\mathcal{A}|\mathcal{B}_n) = \frac{\mathbb{P}(\mathcal{B}_n|\mathcal{A})\mathbb{P}(\mathcal{A})}{\mathbb{P}(\mathcal{B}_n)}, \quad \mathbb{P}(\mathcal{B}_n|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A}|\mathcal{B}_n)\mathbb{P}(\mathcal{B}_n)}{\mathbb{P}(\mathcal{A})}$$

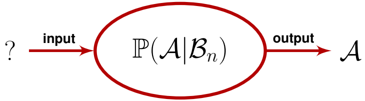
- Model (i.e., system behavior)



- Prediction while observing $\mathcal{B}_n$



- Inference while observing $\mathcal{A}$



The likelihood that $\mathcal{B}_n$ occurred is $\mathbb{P}(\mathcal{B}_n|\mathcal{A})$

Bayes' Theorem

Bayes' Theorem [Inference]

$$\mathbb{P}(\mathcal{B}_n|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A}|\mathcal{B}_n)\mathbb{P}(\mathcal{B}_n)}{\mathbb{P}(\mathcal{A}\cap\mathcal{B}_1) + \mathbb{P}(\mathcal{A}\cap\mathcal{B}_2) + \cdots + \mathbb{P}(\mathcal{A}\cap\mathcal{B}_n)}$$

$$\mathbb{P}(\mathcal{B}_n|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A}|\mathcal{B}_n)\mathbb{P}(\mathcal{B}_n)}{\mathbb{P}(\mathcal{B}_1)\mathbb{P}(\mathcal{A}|\mathcal{B}_1) + \mathbb{P}(\mathcal{B}_2)\mathbb{P}(\mathcal{A}|\mathcal{B}_2) + \cdots + \mathbb{P}(\mathcal{B}_n)\mathbb{P}(\mathcal{A}|\mathcal{B}_n)}$$

- Inference $\Longrightarrow$ updating beliefs about $\mathcal{B}_n$

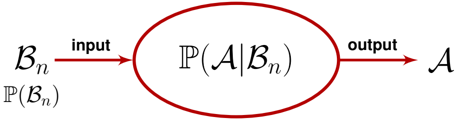
$$? \xrightarrow{\textbf{input}} \mathbb{P}(\mathcal{A}|\mathcal{B}_n) \xrightarrow{\textbf{output}} \mathcal{A}$$

The likelihood that $\mathcal{B}_1$ occurred is $\mathbb{P}(\mathcal{B}_1|\mathcal{A})$
$\mathcal{B}_2$ occurred is $\mathbb{P}(\mathcal{B}_2|\mathcal{A})$

Termenologies

Bayes' Theorem

- $\mathbb{P}(\mathcal{B}_n)\ \forall n$ are denoted as ***priori probabilities***
- $\mathbb{P}(\mathcal{A}|\mathcal{B}_n)\ \forall n$ are denoted as ***transition probabilities***
- $\mathbb{P}(\mathcal{B}_n|\mathcal{A})\ \forall n$ are denoted as ***posteriori probabilities***

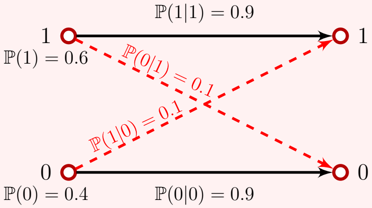


The likelihood that the input is $\mathcal{B}_1$ is $\mathbb{P}(\mathcal{B}_1|\mathcal{A})$
$\mathcal{B}_2$ is $\mathbb{P}(\mathcal{B}_2|\mathcal{A})$

Example

Example

Consider a communication system that transmits either 1 or 0. Transmitted bits go through a ***binary symmetric channel*** with the following state digram



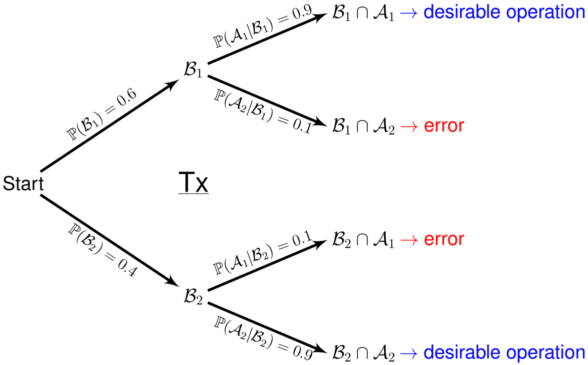
- Find the probability of error for the depicted communication system

Solution

- We first define the following four events
  - Event $\mathcal{B}_1$: probability of sending 1
  - Event $\mathcal{B}_2$: probability of sending 0
  - Event $\mathcal{A}_1$: probability of receiving 1
  - Event $\mathcal{A}_2$: probability of receiving 0
- This will lead to the following four outcomes
  1. $\mathcal{B}_1 \cap \mathcal{A}_1$ (desirable)
  2. $\mathcal{B}_2 \cap \mathcal{A}_2$ (desirable)
  3. $\mathcal{B}_1 \cap \mathcal{A}_2$ (undesirable [error])
  4. $\mathcal{B}_2 \cap \mathcal{A}_1$ (undesirable [error])

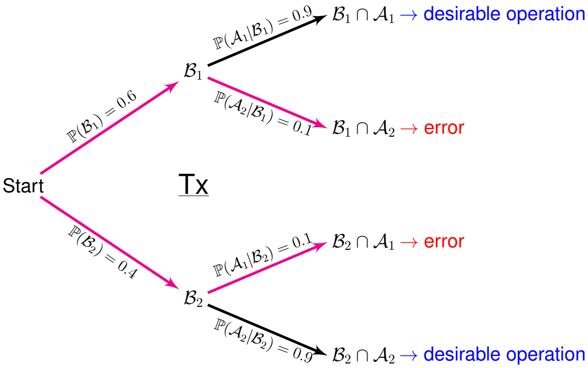
Solution

- Applying the multi-stage transmission model



Solution

- Applying the cain rule
- $\mathbb{P}(\text{error}) = \mathbb{P}(\mathcal{B}_1)\mathbb{P}(\mathcal{A}_2|\mathcal{B}_1) + \mathbb{P}(\mathcal{B}_2)\mathbb{P}(\mathcal{A}_1|\mathcal{B}_2) = 0.01$



Independent Events

Statistical independence

- Two events $\mathcal{A}$ and $\mathcal{B}$ are statistically independent if and only if (iff):

$$\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B})$$

- If $\mathcal{A}$ and $\mathcal{B}$ are statistically independent, then

$$\mathbb{P}(\mathcal{A}|\mathcal{B}) = \mathbb{P}(\mathcal{A}) \quad \text{and} \quad \mathbb{P}(\mathcal{B}|\mathcal{A}) = \mathbb{P}(\mathcal{B})$$

- Mutually exclusive events are not statistically independent
- Let $\mathcal{A}$ and $\mathcal{B}$ be two mutually exclusive event such that $\mathbb{P}(\mathcal{A}) > 0$ and $\mathbb{P}(\mathcal{B}) > 0$.

$$\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = 0 \neq \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B})$$

- Any two sets that have $\mathcal{A} \cap \mathcal{B} = \phi$ cannot be statistically independent

Independent Events

Example

- Let $\mathcal{S} = \{1, 2, 3, 4, 5, 6, 7, 8\}$ be a set of equiprobable elements.
- Consider the two events $\mathcal{A} = \{1, 3, 5, 7\}$ and $\mathcal{B} = \{7, 8\}$
- Are $\mathcal{A}$ and $\mathcal{B}$ statistically independent?

Solution 1

- The marginal probability of $\mathcal{A}$ is $\mathbb{P}(\mathcal{A}) = \frac{1}{2}$
- The marginal probability of $\mathcal{B}$ is $\mathbb{P}(\mathcal{B}) = \frac{1}{4}$
- The joint probability of $\mathcal{A}$ and $\mathcal{B}$ is $\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \frac{1}{8}$
- Since $\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B})$, then $\mathcal{A}$ and $\mathcal{B}$ are statistically independent

Independent Events

Example

- Let $\mathcal{S} = \{1, 2, 3, 4, 5, 6, 7, 8\}$ be a set of equiprobable elements.
- Consider the two events $\mathcal{A} = \{1, 3, 5, 7\}$ and $\mathcal{B} = \{7, 8\}$
- Are $\mathcal{A}$ and $\mathcal{B}$ statistically independent?

Solution 2

- The marginal probability of $\mathcal{B}$ is $\mathbb{P}(\mathcal{B}) = \frac{1}{4}$
- Given that $\mathcal{A}$ occurred, the conditional probability that $\mathcal{B}$ occurs is $\mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{1}{4}$
- Since $\mathbb{P}(\mathcal{B}|\mathcal{A}) = \mathbb{P}(\mathcal{B})$, then $\mathcal{A}$ and $\mathcal{B}$ are statistically independent

Independent Events

Statistical independence

If $\mathcal{A}$ and $\mathcal{B}$ are statistically independent, then

- $\bar{\mathcal{A}}$ and $\mathcal{B}$ are independent
- $\mathcal{A}$ and $\bar{\mathcal{B}}$ are independent
- $\bar{\mathcal{A}}$ and $\bar{\mathcal{B}}$ are independent

Union of independent events

Let $\mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_n$ be a set of independent events, then using De Moegan's law we have

$$\mathcal{A}_1 \cup \mathcal{A}_2 \cup \cdots \cup \mathcal{A}_n = \overline{\left(\bar{\mathcal{A}}_1 \cap \bar{\mathcal{A}}_2 \cap \cdots \cap \bar{\mathcal{A}}_n\right)}$$

Then we have

$$\begin{aligned}
\mathbb{P}\left(\mathcal{A}_1 \cup \mathcal{A}_2 \cup \cdots \cup \mathcal{A}_n\right) &= 1 - \mathbb{P}\left(\bar{\mathcal{A}}_1 \cap \bar{\mathcal{A}}_2 \cap \cdots \cap \bar{\mathcal{A}}_n\right) \\
&= 1 - \mathbb{P}\left(\bar{\mathcal{A}}_1\right)\mathbb{P}\left(\bar{\mathcal{A}}_2\right)\cdots\mathbb{P}\left(\bar{\mathcal{A}}_n\right) \\
&= 1 - \left(1 - \mathbb{P}\left(\mathcal{A}_1\right)\right)\left(1 - \mathbb{P}\left(\mathcal{A}_2\right)\right)\cdots\left(1 - \mathbb{P}\left(\mathcal{A}_n\right)\right)
\end{aligned}$$

Independent Events

Statistical independence

- Events $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ are statistically independent if and only if (iff):

$$\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B})$$
$$\mathbb{P}(\mathcal{A} \cap \mathcal{C}) = \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{C})$$
$$\mathbb{P}(\mathcal{B} \cap \mathcal{C}) = \mathbb{P}(\mathcal{B})\mathbb{P}(\mathcal{C})$$
$$\mathbb{P}(\mathcal{A} \cap \mathcal{B} \cap \mathcal{C}) = \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B})\mathbb{P}(\mathcal{C})$$

Independent Events

Example

- Let $\mathcal{S} = \{1, 2, 3, 4, 5, 6, 7, 8\}$ be a set of equiprobable elements.
- Consider the events $\mathcal{A} = \{1, 3, 5, 7\}$, $\mathcal{B} = \{7, 8\}$, $\mathcal{C} = \{1, 2, 3, 5, 7\}$, and $\mathcal{D} = \{3, 5, 7\}$
- Check the statistical independence of $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$

Solution

- The marginal probabilities of the events are $\mathbb{P}(\mathcal{B}) = \frac{1}{2}$, $\mathbb{P}(\mathcal{B}) = \frac{1}{4}$, $\mathbb{P}(\mathcal{C}) = \frac{5}{8}$, and $\mathbb{P}(\mathcal{D}) = \frac{3}{8}$
- $\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \frac{1}{8} = \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B})$ ✓
- $\mathbb{P}(\mathcal{A} \cap \mathcal{C}) = \frac{1}{2} \neq \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{C})$ ✗
- $\mathbb{P}(\mathcal{A} \cap \mathcal{D}) = \frac{3}{8} \neq \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{D})$ ✗
- $\mathbb{P}(\mathcal{B} \cap \mathcal{C}) = \frac{1}{8} \neq \mathbb{P}(\mathcal{B})\mathbb{P}(\mathcal{C})$ ✗
- $\mathbb{P}(\mathcal{B} \cap \mathcal{D}) = \frac{1}{8} \neq \mathbb{P}(\mathcal{B})\mathbb{P}(\mathcal{D})$ ✗. . . and so on

Independent Events

Statistical independence

- Generally, events $\mathcal{A}_n$, $n = 1, 2, \ldots, n$ are statistically independent if and only if (iff):

$$
\begin{aligned}
\mathbb{P}(\mathcal{A}_n \cap \mathcal{A}_m) &= \mathbb{P}(\mathcal{A}_n)\mathbb{P}(\mathcal{A}_m) & \forall n \neq m \\
\mathbb{P}(\mathcal{A}_n \cap \mathcal{A}_m \cap \mathcal{A}_l) &= \mathbb{P}(\mathcal{A}_n)\mathbb{P}(\mathcal{A}_n)\mathbb{P}(\mathcal{A}_k) & \forall n \neq m \neq l \\
&\cdots \\
\mathbb{P}(\mathcal{A}_1 \cap \mathcal{A}_2 \cap \cdots \cap \mathcal{A}_N) &= \mathbb{P}(\mathcal{A}_1)\mathbb{P}(\mathcal{A}_2)\mathbb{P}\cdots(\mathcal{A}_N)
\end{aligned}
$$

Conditional Independence

Conditional Independence

Two events $\mathcal{A}$ and $\mathcal{B}$ are conditionally independent given an event $\mathcal{C}$ with $\mathbb{P}(\mathcal{C}) > 0$ if

$$\mathbb{P}(\mathcal{A} \cap \mathcal{B}|\mathcal{C}) = \mathbb{P}(\mathcal{A}|\mathcal{C})\mathbb{P}(\mathcal{B}|\mathcal{C})$$

This implies that

$$\mathbb{P}(\mathcal{A}|\mathcal{B},\mathcal{C}) = \mathbb{P}(\mathcal{A}|\mathcal{C})$$

$$\begin{aligned}
\mathbb{P}(\mathcal{A}|\mathcal{B},\mathcal{C}) &= \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{B}|\mathcal{C})}{\mathbb{P}(\mathcal{B}|\mathcal{C})} \\
&= \frac{\mathbb{P}(\mathcal{A}|\mathcal{C})\mathbb{P}(\mathcal{B}|\mathcal{C})}{\mathbb{P}(\mathcal{B}|\mathcal{C})} \\
&= \mathbb{P}(\mathcal{A}|\mathcal{C})
\end{aligned}$$

Example

- A box contains two coins: a regular coin and one fake two-headed coin (P(H)=1). I choose a coin at random and toss it twice. Define the following events.
  - $\mathcal{A}$= First coin toss results in an H.
  - $\mathcal{B}$= Second coin toss results in an H.
  - $\mathcal{C}$= Coin 1 (regular) has been selected.
- Find
  - $\mathbb{P}(\mathcal{A}|\mathcal{C})$
  - $\mathbb{P}(\mathcal{B}|\mathcal{C})$
  - $\mathbb{P}(\mathcal{A} \cap \mathcal{B}|\mathcal{C})$
  - $\mathbb{P}(\mathcal{A})$ and $\mathbb{P}(\mathcal{B})$
  - $\mathbb{P}(\mathcal{A} \cap \mathcal{B})$

Progress...

- Last section
  - Bayes' theorem
  - Independent events
- Next section
  - Combined experiments
  - Counting
  - Bernoulli trials

Combined Experiments

- Develop a systematic procedure to study uncertain situations that involve several experiments
  - Flipping a coin N times $\{H\ H\ H\ T\ T\ H\ H\ T\ \cdots\ H\}$
  - Rolling a die N times $\{1\ 2\ 3\ 5\ 4\ 2\ 1 \cdots 5\}$
  - Select several cards from a deck of cards with/without replacement $\{J\heartsuit\ 10\clubsuit\ 5\diamondsuit\ 2\spadesuit\ 7\spadesuit\ K\diamondsuit\ \cdots\ Q\heartsuit\}$
  - Select several balls from urns with/without replacement $\{\bullet\bullet\bullet\bullet\bullet\bullet\bullet\ \cdots\ \bullet\}$
- Remember [rule to think about uncertain situation]:

Steps to Solve Probability Problems

1. Specify the sample space
2. Identify the event of interest
3. Specify the probability law
4. Calculate the probability

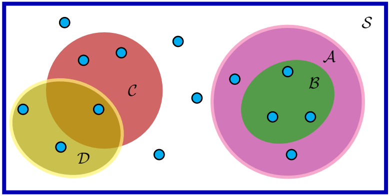
Combined Experiments

- For uniform probability laws

$$\mathbb{P}(\mathcal{A}) = \frac{|\mathcal{A}|}{|\mathcal{S}|}$$

$$\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \frac{|\mathcal{A} \cap \mathcal{B}|}{|\mathcal{S}|}$$

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



Combined Experiments

- The outcome space in combined experiments is determined by the Cartesian product
$$\mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{S}_3 \cdots \times \mathcal{S}_n$$
- The size of the combined outcome space is determined via the multiplicative rule
- Examples
  - Flipping a coin twice
    $\mathcal{S} = \{HH, HT, TH, TT\}$
  - Rolling a die twice
    $\mathcal{S} = \{(1,1),(1,2),(1,3),\ldots,(2,1),\ldots,(6,6)\}$
- The outcome space in combined experiments may be vast
- Counting techniques help identifying the cardinality of the outcome space and the subset constituting the events of interest

Combined Experiments

Terminologies

- Sampling or draw: choosing an element of a set
- With and without replacement: do we allow repetition of previously selected items or not
- Ordered and unordered: does order matter or not

- We are interested in the following four cases
    1. Ordered sampling with replacement
    2. Ordered sampling without replacement
    3. Unordered sampling without replacement
    4. Unordered sampling with replacement

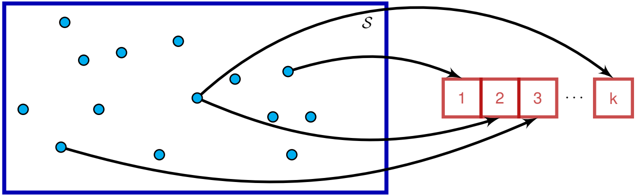
Ordered sampling with replacement

- Consider a outcome space with n elements
- We select k samples at random and put them in an ordered list and allow repetition
- The number of ways the k element list can be constructed is given by

$$\underbrace{n \times n \times n \times n \cdots \times n}_{k \text{ times}} = n^k$$

- If we select k samples from an outcome space of size $|\mathcal{S}_1| = n$ and t samples from an outcome space $|\mathcal{S}_2| = m$, then number of ways the list can be done is

$$\underbrace{n \times n \cdots \times n}_{k \text{ times}} \times \underbrace{m \times m \cdots \times m}_{t \text{ times}} = n^k m^t$$



Ordered sampling with replacement

- Examples
  - Car plates constructed from ***3 letters*** and ***4 digits***
    $26 \times 26 \times 26 \times 10 \times 10 \times 10 \times 10$
  - Uniforms given that we are selecting ***5 ties***, ***3 pants***, ***3 shirts***, ***2 jackets***
    $5 \times 3 \times 3 \times 2$

Ordered sampling without replacement

Permutations

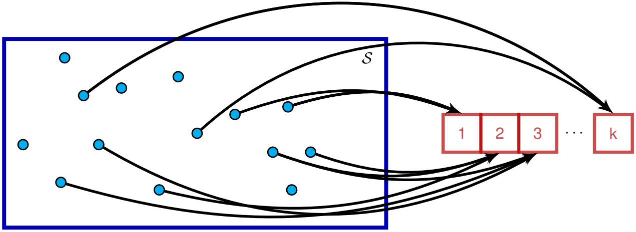
- Describes the all possible ways in which a set or number of things can be ordered or arranged
- Order matters and repetition is not allowed
- For a set of n elements, the number of possible permutations of these elements is given by

$$P_n^n = n! = n \times (n-1) \times (n-2) \times (n-3) \times \cdots \times 3 \times 2 \times 1$$

- If only k elements of a set of n elements will be selected at random and ordered, then the number of possible permutations of these k-elements is given by

$$P_k^n = \frac{n!}{(n-k)!} = n \times (n-1) \times (n-2) \times (n-3) \times \cdots \times (n-k+1)$$

Ordered sampling without replacement



- k students are attending a course, what is the probability that at least any two of them have the same birthday?
- The size of the birthday outcome space is $n = 365$
- Let the event $\mathcal{A}$ be the event that at least any two of them have the same birthday
- $\bar{\mathcal{A}}$ the event that none of them have the same birthday

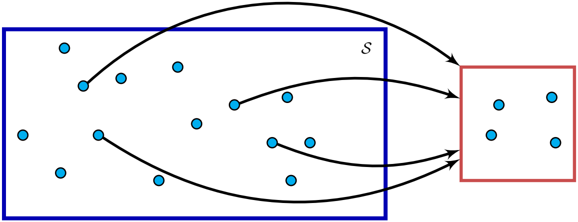
$$P(\mathcal{A}) = 1 - P(\bar{\mathcal{A}}) = 1 - \frac{|\bar{\mathcal{A}}|}{|\mathcal{S}|} = 1 - \frac{P_k^n}{n^k}$$

Unordered sampling without replacement

Combinations

- Describes the all possible ways in which a subset of k-elements can be formed from a set of n-elements.
- Ordering does not mater and replacement is not allowed.
- The number ways a subset of k-elements can be formed from a set of n-elements is given by

$$C_k^n = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$



Unordered sampling without replacement

- Example: choosing 3 cards from a deck of 52 cards, what is the probability of having at least one ace?
- Order does not matter.
- The outcome space is of size $\binom{52}{3}$
- The number of combinations without any aces is $\binom{48}{3}$
- Let $\mathcal{A}$ be the event of having at least one ace in the chosen three cards, then

$$P(\mathcal{A}) = 1 - P(\bar{\mathcal{A}}) = 1 - \frac{|\bar{\mathcal{A}}|}{|\mathcal{S}|} = 1 - \frac{\binom{48}{3}}{\binom{52}{3}}$$

Unordered sampling without replacement

Useful identity

- Recall that for a set of n elements, the number of all subsets of n is given by the power set of size

$$\#\text{of all combinations} = 2^n$$

- The number of all subsets can be interpreted at the number of all possible combinations of all sizes $k \in \{0, 1, 2, \cdots, n\}$
- Hence, we have the following identity

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

Permutations & Combinations

Permutations & Combinations for Subsets

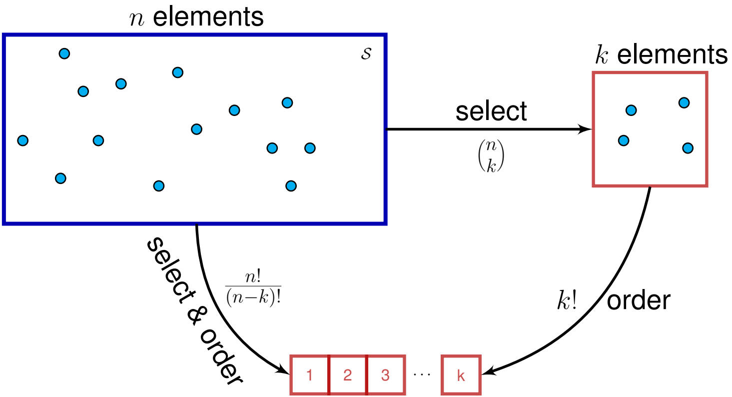
- Randomly select k elements from an outcome space with n elements

$$\#\text{of all combinations of size } k = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

- Randomly select ***<u>and order</u>*** k elements from an outcome space with n elements

$$\#\text{of all permutations of size } k = \frac{n!}{(n-k)!} = P_k^n; \quad k = 1, 2, \ldots, n$$

Permutations & Combinations



Bernoulli Trials

Example

- Consider a coin flipping experiment where the probability of getting a head is p and the probability of getting a tail is $1-p$
- Find the probability of having exactly k heads out of N coin flipping experiments.

Solution 1

- Consider that $k=3$ $N=10$
- $\mathbb{P}\{HTTHHTTTTT\} = p^3 \times (1-p)^7$
- $\mathbb{P}\{HTTHTTTTHT\} = p^3 \times (1-p)^7$
- $\mathbb{P}\{TTTHTHTHT\} = p^3 \times (1-p)^7$
- Generally, the probability of having a pattern with k heads and $N-k$ tails is $p^k \times (1-p)^{N-k}$
- The number of patterns that satisfy such property is $\binom{N}{k}$
- Hence, the probability of having any patternsthat satisfies k heads and $N-k$ tails is

$$\binom{N}{k} p^k (1-p)^{N-k}$$

Binomial Probability

Binomial Probability

- Let p be the probability of success and $1 - p$ be the probability of failure
- The probability of getting k successes out of N trials is

$$\mathbb{P}(\#\text{successes} = k) = \binom{N}{k} p^k (1-p)^{N-k}; \quad k = 0, 1, 2, \ldots, N$$

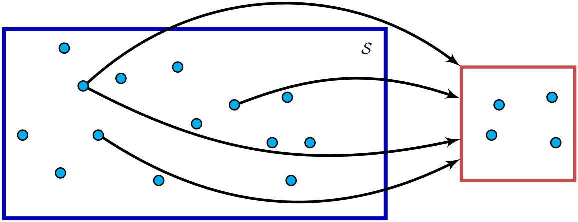
$$\sum_{k=0}^{N} \binom{N}{k} p^k (1-p)^{N-k} = ?1$$

Unordered sampling with replacement

Combinations

- Describes the all possible ways in which a subset of k-elements can be formed from a set of n-elements.
- Ordering does not mater and replacement is allowed.
- The number ways a subset of k-elements can be formed from a set of n-elements with replacement is given by

$$\binom{n+k-1}{k} = \binom{n+k-1}{n-1}$$



Questions?