CS 112: Computer System Modeling Fundamentals

Prof. Jenn Wortman Vaughan

March 29, 2011

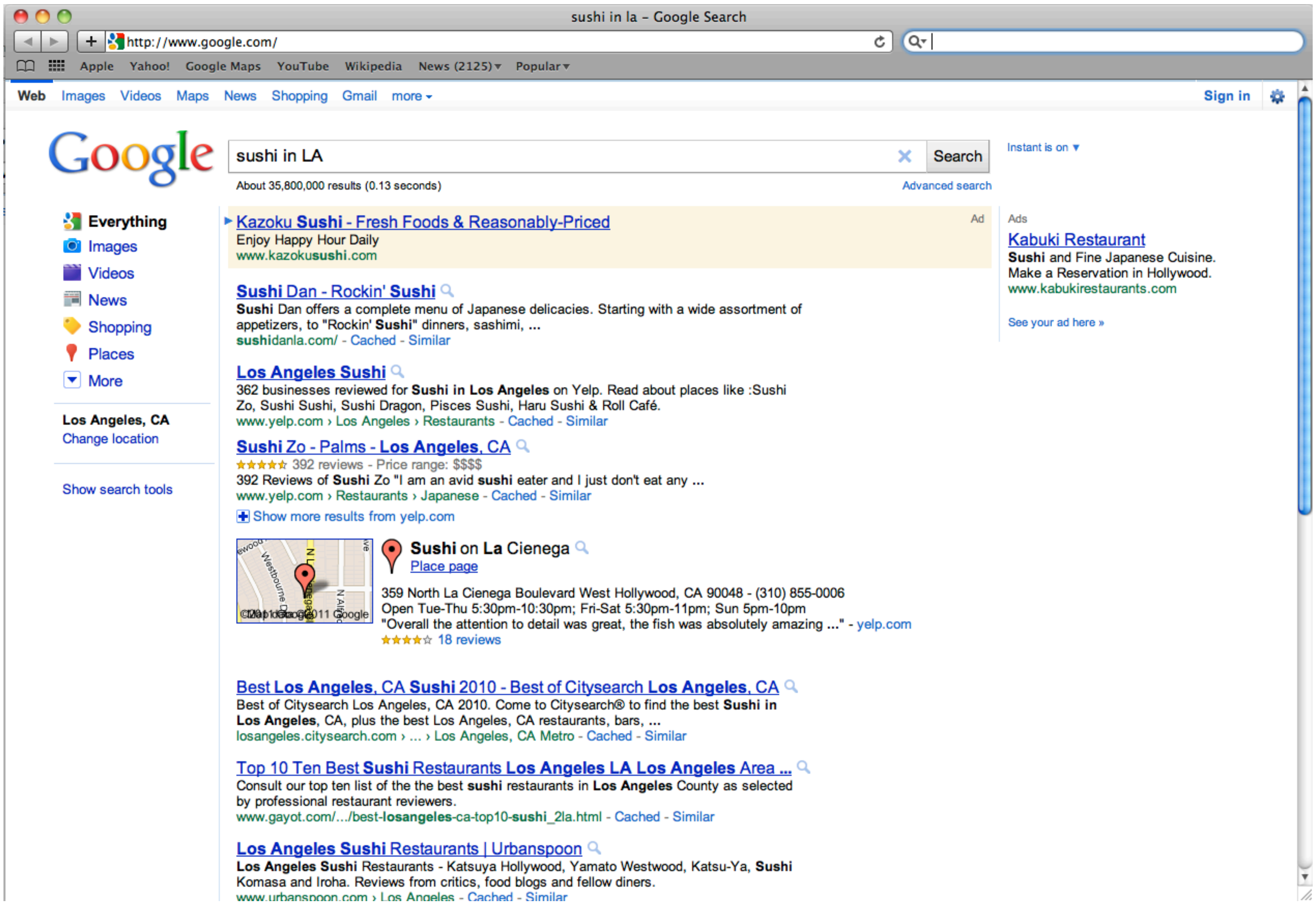
“Reasoning About Uncertainty”

CS 112: ~~Computer System Modeling Fundamentals~~

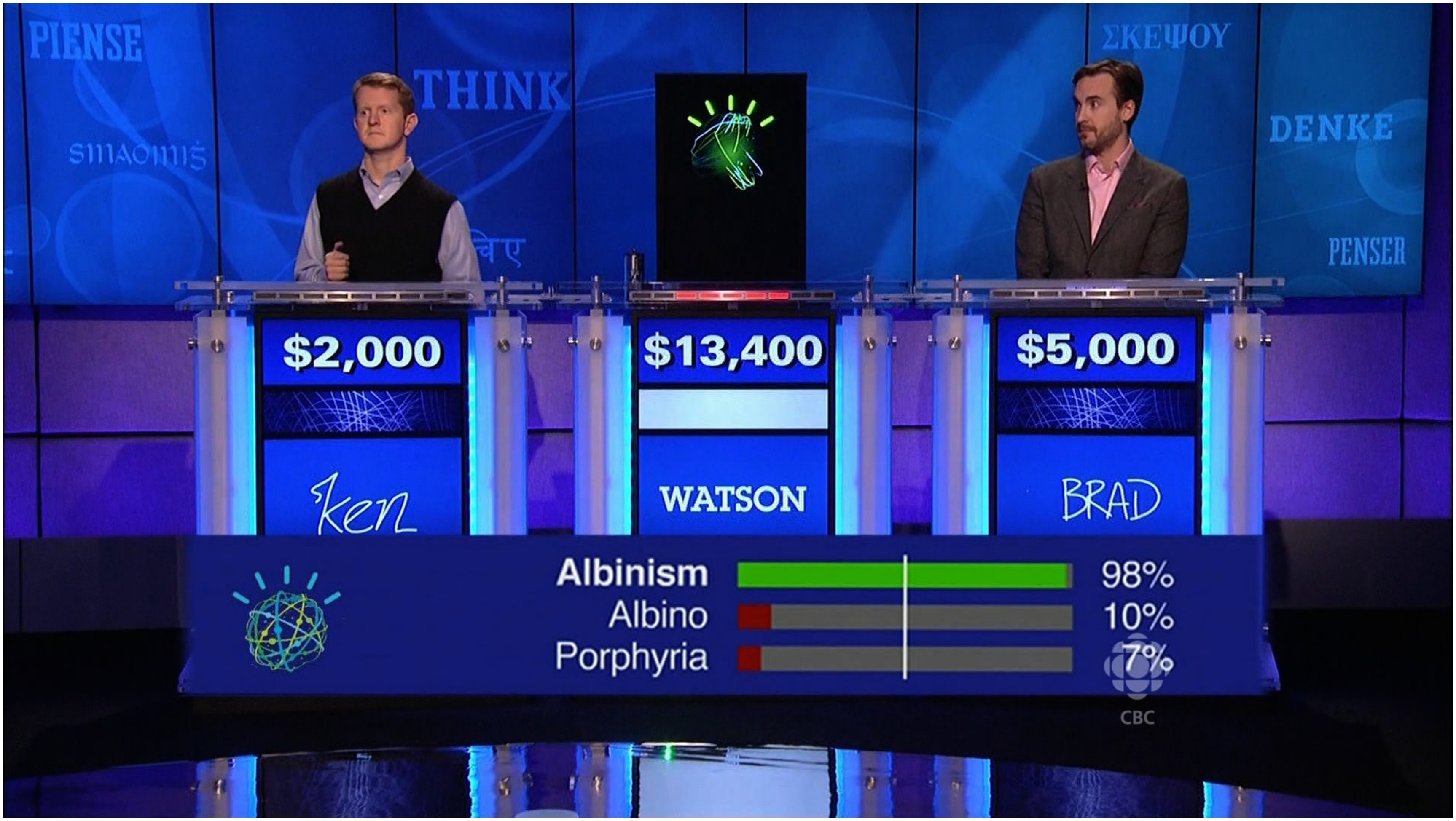
Prof. Jenn Wortman Vaughan

March 29, 2011

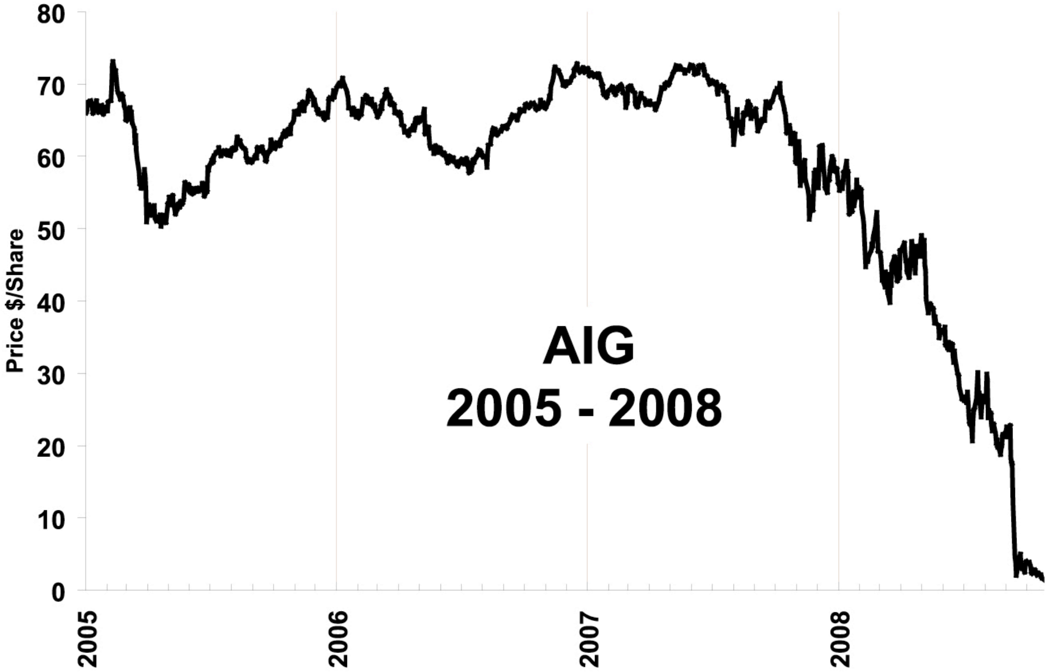
Uncertainty in Computer Science



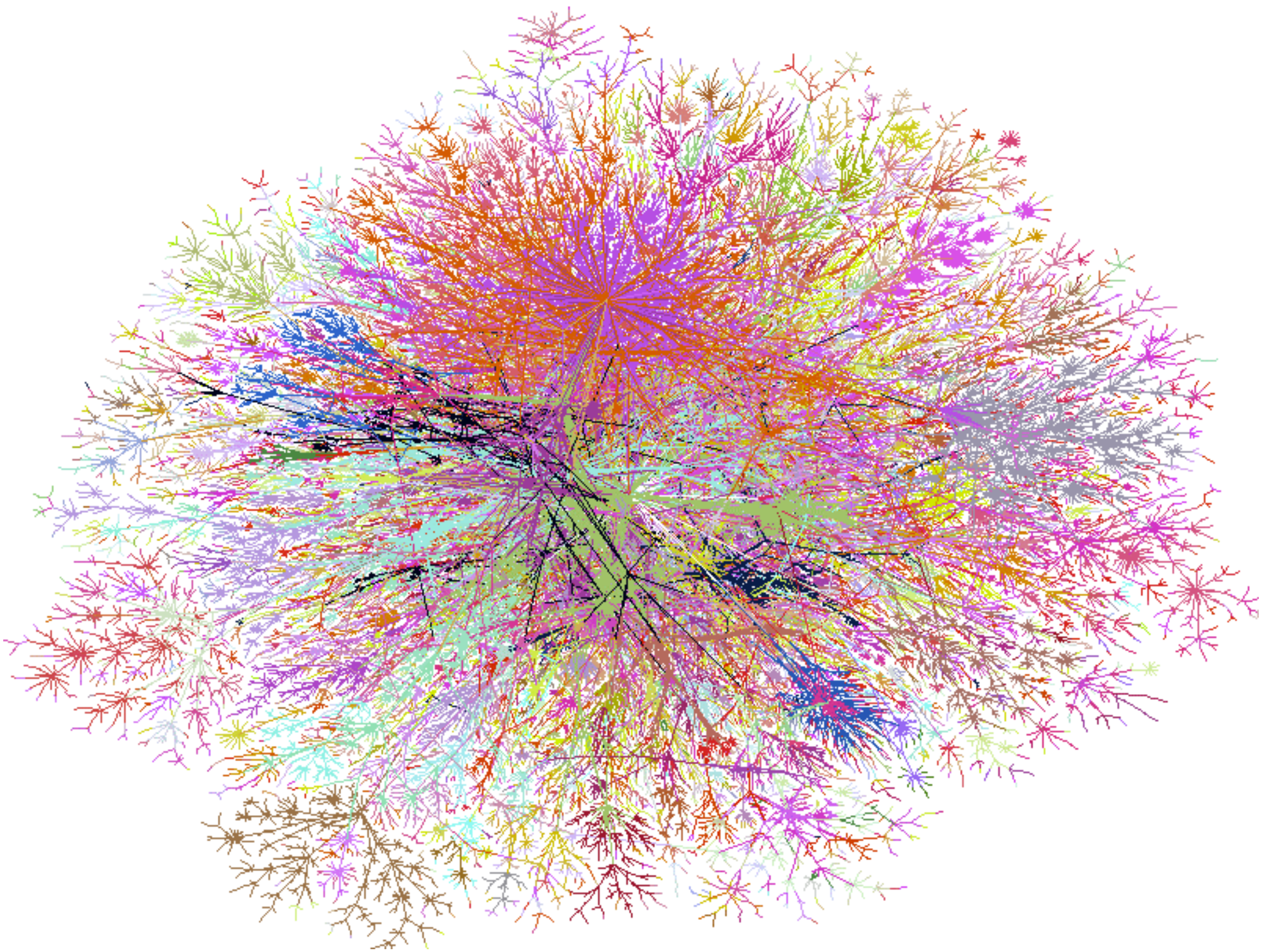
Uncertainty in Computer Science



Uncertainty in Computer Science



Uncertainty in Computer Science



Uncertainty in Computer Science

To: Jenn Wortman Vaughan
From: Jeff Vaughan
Subject: Plans for tonight ✓

To: Jenn Wortman Vaughan
From: Jens Palsberg
Subject: Meeting ✓

To: Jenn Wortman Vaughan
From: Bob Smith
Subject: V14GR4 4 U ✗

To: Jenn Wortman Vaughan
From: NIPS Committee
Subject: Paper decision ?

Uncertainty in Computer Science

Uncertainty is everywhere in computer science.

In this course, you will develop foundational, mathematical reasoning skills that will help you deal with it.

Course Overview

Part I: Probability Theory

- Sample space and events
- Probability laws and their basic properties
- Conditional probability and independence
- Bayes' rule
- Counting problems
- Random variables
- Expectation, mean, and variance
- Covariance and correlation
- Limit theorems

Part II: On to something new!

Markov Chains: How can we reason about random processes that evolve over time?

Part II: On to something new!

Markov Chains: How can we reason about random processes that evolve over time?

- Example: Random surfer used in Google's PageRank

Part II: On to something new!

Markov Chains: How can we reason about random processes that evolve over time?

- Example: Random surfer used in Google's PageRank

Statistical Inference: How can we estimate properties of random variables or model parameters from observations?

Part II: On to something new!

Markov Chains: How can we reason about random processes that evolve over time?

- Example: Random surfer used in Google's PageRank

Statistical Inference: How can we estimate properties of random variables or model parameters from observations?

- Example: Estimating parameters of a spam filter

Part II: On to something new!

Markov Chains: How can we reason about random processes that evolve over time?

- Example: Random surfer used in Google's PageRank

Statistical Inference: How can we estimate properties of random variables or model parameters from observations?

- Example: Estimating parameters of a spam filter
- Might cover some basic ideas from machine learning if we have time…

Course Logistics

Course Staff

Instructor: Jenn Wortman Vaughan (jenn@cs.ucla.edu)

Office Hours: Wednesday, 1–3pm, Boelter 4532H

No office hours this week!

TA: Ethan Schreiber (ethan@cs.ucla.edu)

Office Hours: Monday, 11:30–1:30, Boelter 2432

Discussions: Friday, 4–5:50pm, in this room

Grader: Brian Geffon (briangeffon@gmail.com)

Textbook and Material

Required textbook:

Introduction to Probability (2nd Edition)

by Dimitri P. Bertsekas and John N. Tsitsiklis

Other materials, including reading assignments, problem sets, and any lecture slides, will be posted at:

http://www.cs.ucla.edu/~jenn/courses/S11.html

Breakdown of Grades

Five Homework Assignments (30%)

- Mostly pencil-and-paper, a little programming

In-class Quizzes (10%)

- Given at the start of class, lowest quiz dropped

Midterm (30%)

- Tuesday, May 3, one sheet of hand-written notes allowed

Final Exam (30%)

- Thursday, June 9, cumulative, same rules as midterm

Homework Policies

Grading: On each assignment, a subset of problems will be graded. Your grade will be based on these problems only.

Homework Policies

Grading: On each assignment, a subset of problems will be graded. Your grade will be based on these problems only.

Late Assignments: No late assignments will be accepted.

Homework Policies

Grading: On each assignment, a subset of problems will be graded. Your grade will be based on these problems only.

Late Assignments: No late assignments will be accepted.

Regrades: Requests for regrades must be submitted in writing. Requests will not be accepted until 48 hours after assignments are returned, and all requests must be submitted within one week.

Homework Policies

Grading: On each assignment, a subset of problems will be graded. Your grade will be based on these problems only.

Late Assignments: No late assignments will be accepted.

Regrades: Requests for regrades must be submitted in writing. Requests will not be accepted until 48 hours after assignments are returned, and all requests must be submitted within one week.

Academic Honesty: Collaboration is encouraged, but you must follow the academic honesty policy!

Academic Honesty

- Each student must write down his or her own solutions independently in his or her own words.
- Each student must submit a list of anyone with whom the assignment was discussed.
- All sources (internet included) must be properly credited.
- Solution sets from this course or any other course cannot be used under any circumstances.

What is probability?

Interpretation: Relative Frequency

Example: If we say a well-manufactured coin lands on heads with probability 0.5, we might mean that if we flip the coin a large (infinite) number of times, the coin should land on heads about half the time.

Interpretation: Relative Frequency

Example: If we say a well-manufactured coin lands on heads with probability 0.5, we might mean that if we flip the coin a large (infinite) number of times, the coin should land on heads about half the time.

What is the probability that a card randomly selected from a deck is a spade?

Interpretation: Relative Frequency

Example: If we say a well-manufactured coin lands on heads with probability 0.5, we might mean that if we flip the coin a large (infinite) number of times, the coin should land on heads about half the time.

What is the probability that a card randomly selected from a deck is a spade?

What is the probability that a spaceship will remain operational during a mission to Mars? Hmmm….

Interpretation: Subjective Belief

Example: A detective believes that a particular suspect committed the crime with probability 0.7

Interpretation: Subjective Belief

Example: A detective believes that a particular suspect committed the crime with probability 0.7

Can be useful when we need a principled, systematic way to make choices in the presence of uncertainty

Interpretation: Subjective Belief

Example: A detective believes that a particular suspect committed the crime with probability 0.7

Can be useful when we need a principled, systematic way to make choices in the presence of uncertainty

Related to gambling odds…

Basic Definitions

- An experiment is an activity (real or conceptual) that results in exactly one of several possible outcomes.

Basic Definitions

- An experiment is an activity (real or conceptual) that results in exactly one of several possible outcomes.
- The set of all possible outcomes of an experiment is called the sample space, typically denoted Ω.

Basic Definitions

- An experiment is an activity (real or conceptual) that results in exactly one of several possible outcomes.
- The set of all possible outcomes of an experiment is called the sample space, typically denoted Ω.
- Examples:
  - Flipping a single coin.

Basic Definitions

- An experiment is an activity (real or conceptual) that results in exactly one of several possible outcomes.
- The set of all possible outcomes of an experiment is called the sample space, typically denoted Ω.
- Examples:
  - Flipping a single coin.
  - Flipping two coins in a row.

Basic Definitions

- An experiment is an activity (real or conceptual) that results in exactly one of several possible outcomes.
- The set of all possible outcomes of an experiment is called the sample space, typically denoted Ω.
- Examples:
    - Flipping a single coin.
    - Flipping two coins in a row. (Multiple alternatives..)

Basic Definitions

- An experiment is an activity (real or conceptual) that results in exactly one of several possible outcomes.
- The set of all possible outcomes of an experiment is called the sample space, typically denoted Ω.
- Examples:
    - Flipping a single coin.
    - Flipping two coins in a row. (Multiple alternatives..)
    - Pulling a random card from a deck.

Basic Definitions

- An experiment is an activity (real or conceptual) that results in exactly one of several possible outcomes.
- The set of all possible outcomes of an experiment is called the sample space, typically denoted Ω.
- Examples:
  - Flipping a single coin.
  - Flipping two coins in a row. (Multiple alternatives..)
  - Pulling a random card from a deck.
  - Running a computer program.

Samples Spaces

There can be more than one right way to define the sample space for a particular experiment.

Always has to satisfy two properties:

Samples Spaces

There can be more than one right way to define the sample space for a particular experiment.

Always has to satisfy two properties:

- Elements must be mutually exclusive

  (only one can be the real outcome)

Samples Spaces

There can be more than one right way to define the sample space for a particular experiment.

Always has to satisfy two properties:

- Elements must be mutually exclusive
  (only one can be the real outcome)
- Elements must be exhaustive
  (we always obtain an outcome from the sample space)

Events

- An event is a collection of outcomes – that is, a subset of the sample space.

Events

- An event is a collection of outcomes – that is, a subset of the sample space.

- If the experiment is a toss of a die, events could include:

  "the number is odd"

  "the number is less than 3"

  "the number is 2"

Events

- An event is a collection of outcomes – that is, a subset of the sample space.
- If the experiment is a toss of a die, events could include:

  "the number is odd"

  "the number is less than 3"

  "the number is 2"

- Each of these can be represented as a set of atomic events – the basic elements of the sample space.

Events

Since events are sets, we can apply set operations to them…

$A = \{1,3,5\}$ “the number is odd”

$B = \{1,2,3\}$ “the number is ≤ 3”

Events

Since events are sets, we can apply set operations to them…

$A = \{1,3,5\}$ “the number is odd”

$B = \{1,2,3\}$ “the number is ≤ 3”

$A \cup B$ “the number is odd OR the number is ≤ 3”

Events

Since events are sets, we can apply set operations to them…

$A = \{1,3,5\}$ “the number is odd”

$B = \{1,2,3\}$ “the number is ≤ 3”

$A \cup B$ “the number is odd OR the number is ≤ 3”

$A \cap B$ “the number is odd AND the number is ≤ 3”

Probability Axioms

We can encode our knowledge or beliefs about the likelihood of particular sets of outcomes using a probability law, which assigns a probability P(A) to each event A

Probability Axioms

We can encode our knowledge or beliefs about the likelihood of particular sets of outcomes using a probability law, which assigns a probability P(A) to each event A

Probability Axioms:

Probability Axioms

We can encode our knowledge or beliefs about the likelihood of particular sets of outcomes using a probability law, which assigns a probability P(A) to each event A

Probability Axioms:

- Nonnegativity: $P(A) \geq 0$ for every event A

Probability Axioms

We can encode our knowledge or beliefs about the likelihood of particular sets of outcomes using a probability law, which assigns a probability P(A) to each event A

Probability Axioms:

- Nonnegativity: $P(A) \geq 0$ for every event A
- Additivity: If A and B are disjoint events, then

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

Probability Axioms

We can encode our knowledge or beliefs about the likelihood of particular sets of outcomes using a probability law, which assigns a probability $P(A)$ to each event A

Probability Axioms:

- Nonnegativity: $P(A) \geq 0$ for every event A
- Additivity: If A and B are disjoint events, then

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

- Normalization: $P(\Omega) = 1$

Probability Axioms

Probability Axioms:

- Nonnegativity: $P(A) \geq 0$ for every event A
- Additivity: If A and B are disjoint events, then

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

- Normalization: $P(\Omega) = 1$

Other properties can be derived from these axioms…

Probability Axioms

Probability Axioms:

- Nonnegativity: $P(A) \geq 0$ for every event A
- Additivity: If A and B are disjoint events, then

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

- Normalization: $P(\Omega) = 1$

Other properties can be derived from these axioms…

- Suppose that A, B, and C are all disjoint events. What is $P(A \cup B \cup C)$?

Probability Axioms

Probability Axioms:

- Nonnegativity: $P(A) \geq 0$ for every event A
- Additivity: If A and B are disjoint events, then

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

- Normalization: $P(\Omega) = 1$

Other properties can be derived from these axioms…

- Suppose that A, B, and C are all disjoint events. What is $P(A \cup B \cup C)$?
- What is $P(\varnothing)$?

Probability Axioms

Probability Axioms:

- Nonnegativity: $P(A) \geq 0$ for every event A
- Additivity: If A and B are disjoint events, then

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

- Normalization: $P(\Omega) = 1$

Other properties can be derived from these axioms…

- Suppose that A, B, and C are all disjoint events. What is $P(A \cup B \cup C)$?
- What is $P(\varnothing)$?
- What is $P(A^c)$?

Other Properties of Probability Laws

For any events A and B,

- If $A \subset B$ then $P(A) \leq P(B)$

Other Properties of Probability Laws

For any events A and B,

- If $A \subset B$ then $P(A) \leq P(B)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Other Properties of Probability Laws

For any events A and B,

- If $A \subset B$ then $P(A) \leq P(B)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A \cup B) \leq P(A) + P(B)$

Other Properties of Probability Laws

For any events A and B,

- If $A \subset B$ then $P(A) \leq P(B)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A \cup B) \leq P(A) + P(B)$

Challenge: Try to prove these properties on your own using the axioms of probability and the rules of set theory!

Conditional Probability

- Suppose we have a fair die. If we are told that the number rolled is odd, what is the probability that it is 3?

Conditional Probability

- Suppose we have a fair die. If we are told that the number rolled is odd, what is the probability that it is 3?

$$P(\text{number is 3} \mid \text{number is odd})$$

Conditional Probability

- Suppose we have a fair die. If we are told that the number rolled is odd, what is the probability that it is 3?

  P(number is 3 | number is odd)

- When all outcomes are equally likely, we can define

  $P(A \mid B) = \# \text{ elements in } A \cap B \,/\, \# \text{ elements in } B$

Conditional Probability

- Suppose we have a fair die. If we are told that the number rolled is odd, what is the probability that it is 3?

$$P(\text{number is } 3 \mid \text{number is odd})$$

- When all outcomes are equally likely, we can define

$$P(A \mid B) = \#\text{ elements in } A \cap B \,/\, \#\text{ elements in } B$$

- Generalizing this idea, we can define

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

Conditional Probability

- Suppose we have a fair die. If we are told that the number rolled is odd, what is the probability that it is 3?

$$P(\text{number is 3} \mid \text{number is odd})$$

- When all outcomes are equally likely, we can define

$$P(A \mid B) = \#\text{ elements in } A \cap B \,/\, \#\text{ elements in } B$$

- Generalizing this idea, we can define

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

- Can show that conditional probabilities satisfy our 3 axioms with B playing the role of Ω