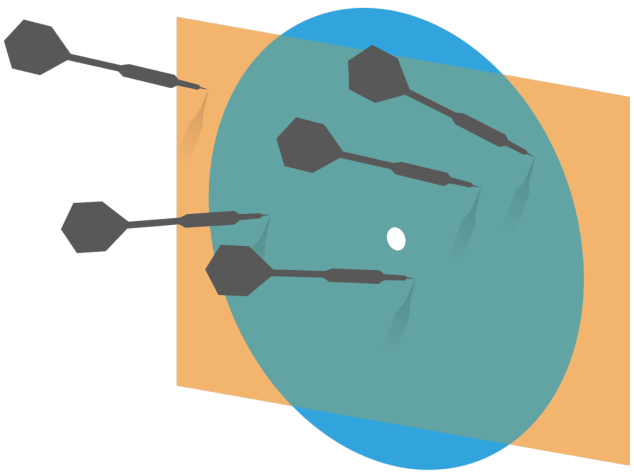
Probability and Statistics for Computer Science



Credit: wikipedia

"Probabilistic analysis is mathematical, but intuition dominates and guides the math" – Prof. Dimitri Bertsekas

Hongye Liu, Teaching Assistant Prof, CS361, UIUC, 9.3.2021

Last time

More summary statistics

Correlation coefficient

Correlation coeff. for prediction

Warm up

A game of chance

Warm up (II)

- Fill the blanks:

"I am an avid vegetarian and I enjoy eating all day long, people admire my appetite and like to watch me eat." I am a P A N D A. How many ways are there to rearrange these 5 letters? $\cancel{\frac{5!}{2!}}$ ________. If you draw 2 letters from them, how many outcomes (order matters) are there that are without "a"?

$\cancel{3 \times 3}$ 3×2

Objectives

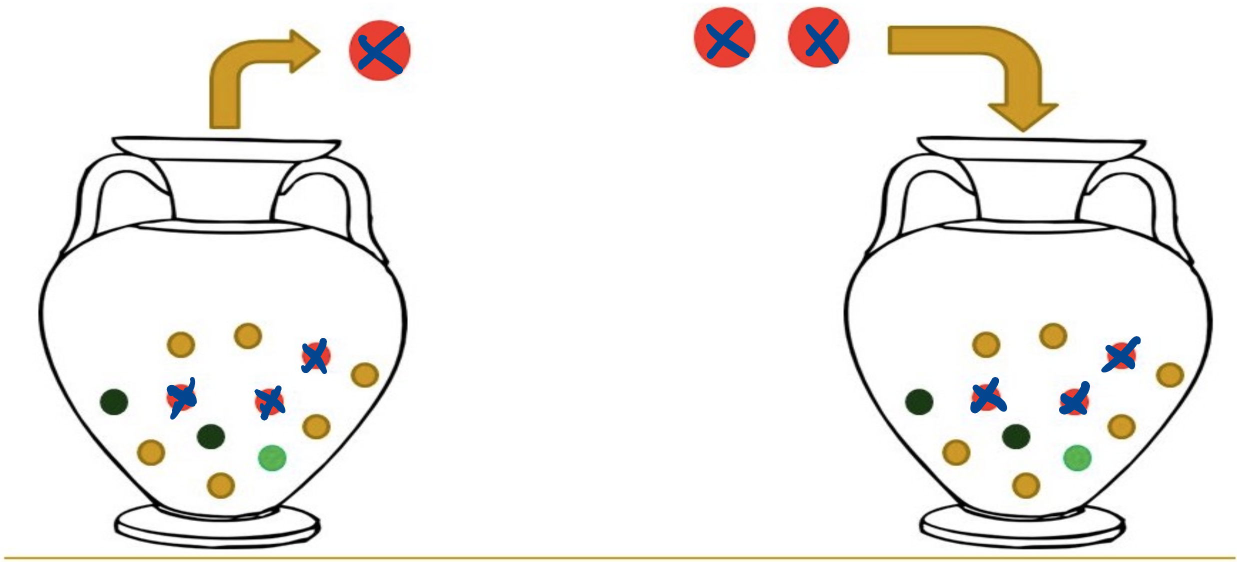
- Probability a first look
  - Outcome and Sample Space
  - Event
  - Probability

    Probability axioms & Properties
  - Calculating probability

Outcome

- An outcome **A** is a possible result of a random repeatable experiment

Random: uncertain, Nondeter-ministic, …



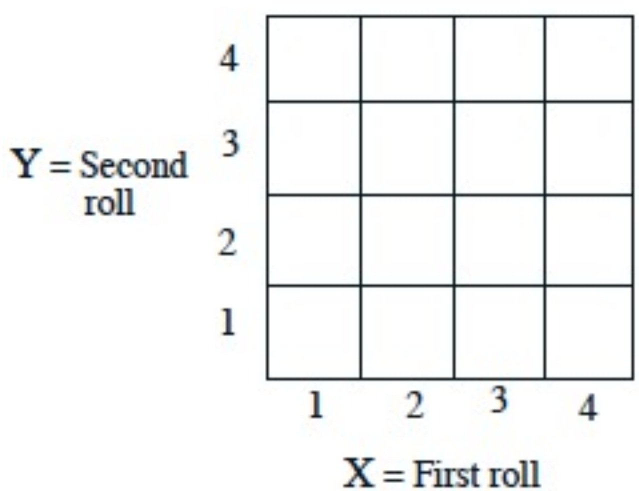
Sample space

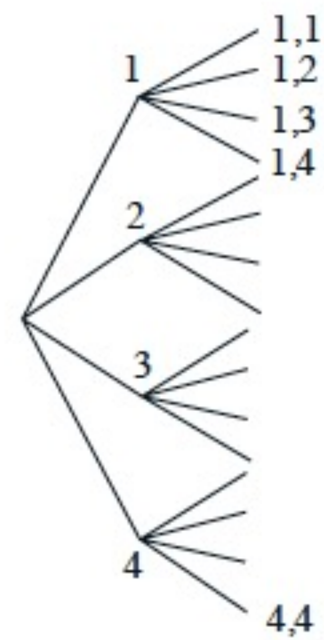
- The Sample Space, **Ω,** is the set of all possible outcomes (Mutually exclusive) associated with the experiment
- Discrete or Continuous

Sample Space example (1)

- Experiment: we roll a 4sided-die twice
- **Discrete** Sample space:

{(1,1), (1,2)….}

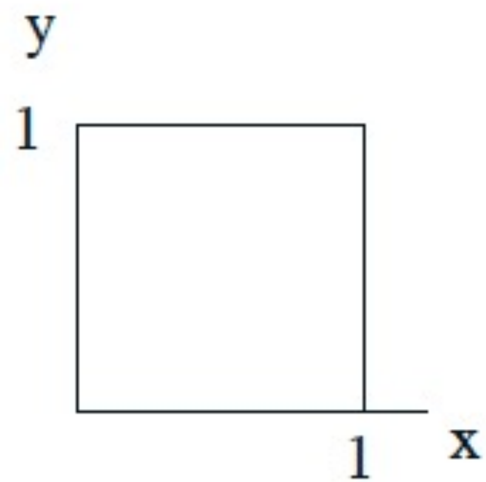




Sample Space example (2)

- Experiment: Romeo and Juliet’s date
- **Continuous** Sample space:

$$\Omega = \{(x,y) \mid 0 \le x, y \le 1\}$$

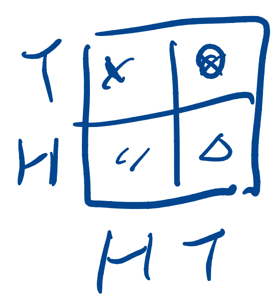


Sample Space depends on experiment (3)

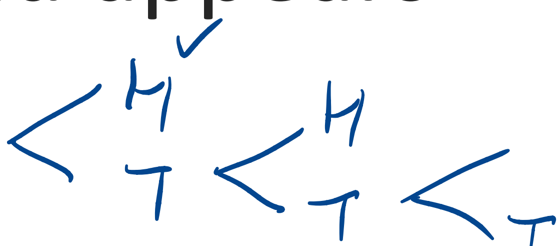
- Different coin tosses
  - Toss a fair coin

{H, T}

  - Toss a fair coin twice



  - Toss until a head appears



Sample Space depends on experiment (4)

- Drawing 2 socks one at a time from a bag containing 1 blue sock, 1 orange sock and 1 white sock **with replacement**?

- Drawing 2 socks one at a time from a bag containing 1 blue sock, 1 orange sock and 1 white sock **without replacement**?

Q1. Sample Space

- Drawing 2 socks one at a time from a bag containing 1 blue sock, 1 orange sock and 1 white sock **with replacement**? What is the number of unique outcomes in the sample space?

  A. 5   B. 7 $\boxed{\text{C.}}$ 9

3×3

Q2.Sample Space

- Drawing 2 socks one at a time from a bag containing 1 blue sock, 1 orange sock and 1 white sock **without replacement**? What is the number of unique outcomes in the sample space?

  A. 5 B. 6 C. 9

Sample Space in real life

- Possible outrages of a power network
- Possible mutations in a gene
- A bus' arriving time

Event

- An event **E** is a subset of the sample space Ω
- So an event is a set of outcomes that is a subset of Ω, ie.
  - Zero outcome ϕ
  - One outcome $\{A\}$
  - Several outcomes $\{A_1, A_2, A_3\}$
  - All outcomes Ω

The same experiment may have different events

- When two coins are tossed
  - Both coins come up the same?
  - At least one head comes up?

{HH
TT}

{HH, HT, TH}

Some experiment may never end

- Experiment: Tossing a coin until a head appears

  - **E**: Coin is tossed at least 3 times

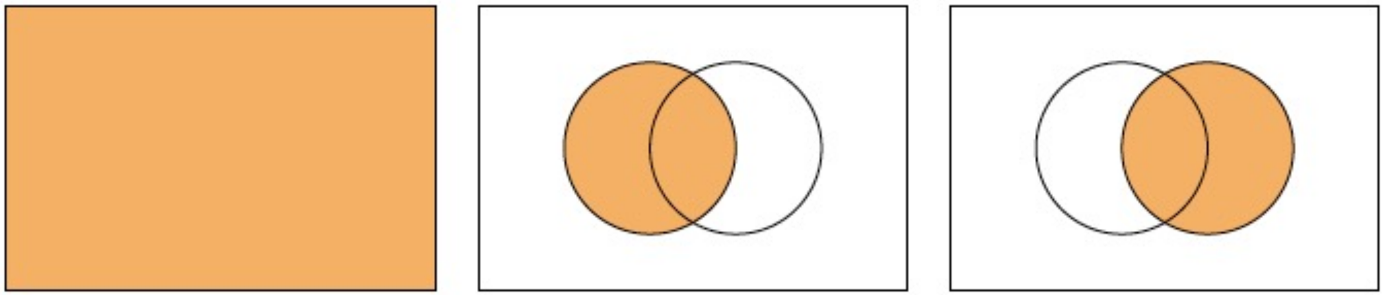
    This event includes infinite # of outcomes

TTH

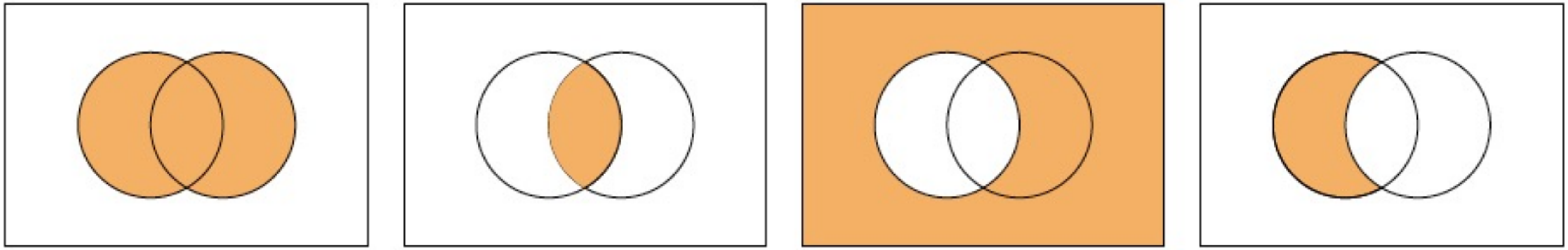
TTTH

- - - . — . . TH

Venn Diagrams of events as sets



Ω E_1 E_2



$E_1 \cup E_2$ $E_1 \cap E_2$ ${E_1}^c$ $E_1 - E_2$

Combining events

- Say we roll a six-sided die. Let

$$E_1 = \{1, 2, 5\} \; and \; E_2 = \{2, 4, 6\}$$

$\Omega = \{1, 2, 3, 4, 5, 6\}$

  - What is $E_1 \cup E_2$ $\rightarrow \{1, 2, 4, 5, 6\}$
  - What is $E_1 \cap E_2$ $\{2\}$
  - What is $E_1 - E_2$ $\{1, 5\}$
  - What is $E_1^c = \Omega - E_1$

$\{3, 4, 6\}$

Frequency Interpretation of Probability

- Given an experiment with an outcome **A**, we can calculate the probability of **A** by repeating the experiment over and over

$$P(A) = \lim_{N->\infty} \frac{number\ of\ time\ A\ occurs}{N}$$

$P(\Omega) = 1$

- So,

$$0 \leq P(A) \leq 1$$

$$\sum_{A_i \in \Omega} P(A_i) = 1$$

Axiomatic Definition of Probability

- A probability function is any function P that maps sets to real number and satisfies the following **three** axioms:

  1) Probability of any event E is non-negative

$$P(E) \geq 0$$

  2) Every experiment has an outcome

$$P(\Omega) = 1$$

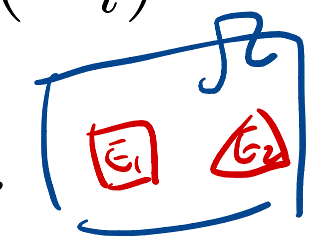
Axiomatic Definition of Probability

3) The probability of disjoint events is additive

$$P(E_1 \cup E_2 \ \cup \ldots \cup \ E_N) = \sum_{i=1}^{N} P(E_i)$$

Mutually exclusive

$$if \ E_i \cap E_j = \emptyset \ for \ all \ i \neq j$$



$P(E_1 = \{1, 2\} \cup E_2 = \{3, 4\}) = P(E_1 \cup E_2)$

$= P(E_1) + P(E_2)$

Q3. Disjoint/Mutual Exclusive

- Toss a coin 3 times

The event "exactly 2 heads appears" and "exactly 2 tails appears" are disjoint.

A. True

B. False

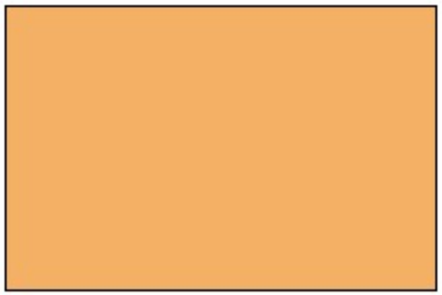
HTH HHT

THT TTH

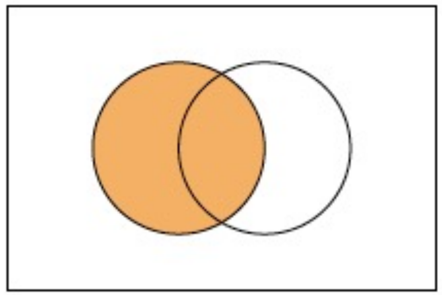
THH

~~HHTT~~ HTT

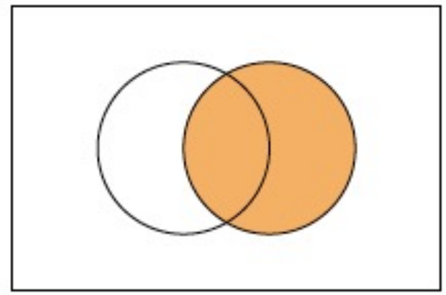
Venn Diagrams of events as sets



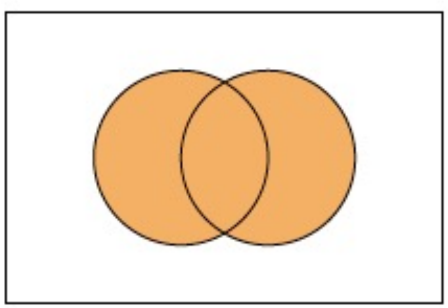
Ω



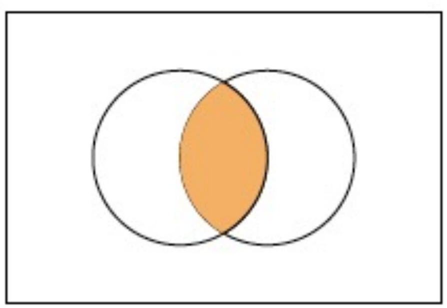
E_1



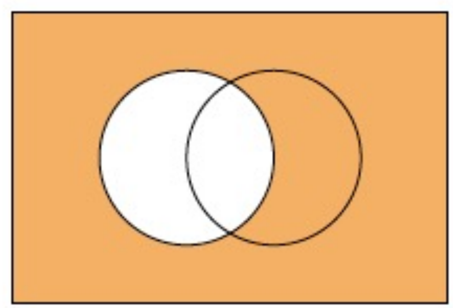
E_2



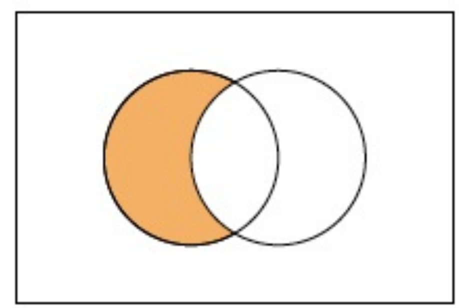
$E_1 \cup E_2$



$E_1 \cap E_2$



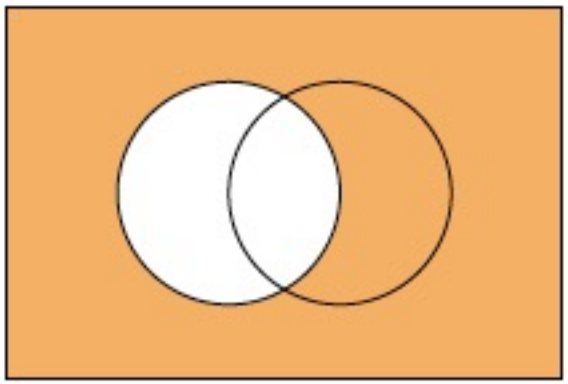
${E_1}^c$



$E_1 - E_2$

Properties of probability

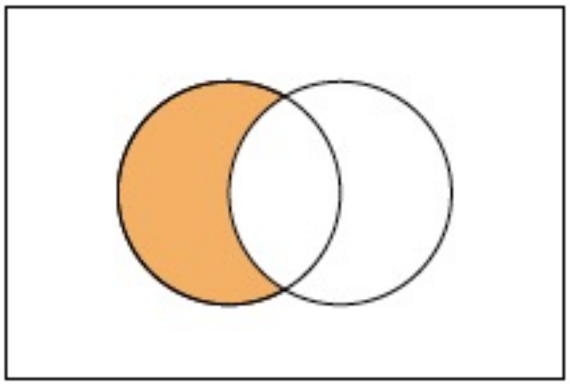
- The complement



$$P(E^c) = 1 - P(E)$$

- The difference

$$P(E_1 - E_2) = P(E_1) - P(E_1 \cap E_2)$$



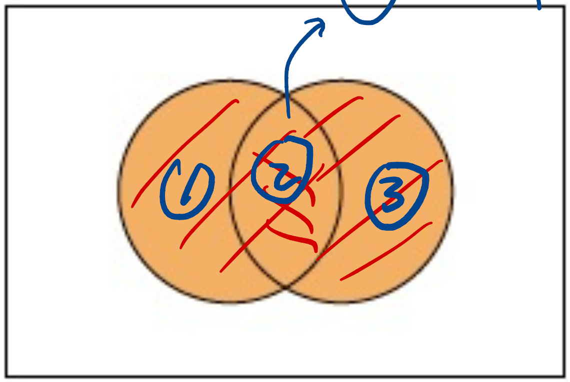
$p(E_1 - E_2) = p(E_1) - p(E_1 \cap E_2)$

Properties of probability

- The union

$$P(E_1 \cup E_2) =$$
$$P(E_1) + P(E_2)$$
$$- P(E_1 \cap E_2)$$

② is $P(E_1 \cap E_2)$



$P(E_1) = ① + ②$ $\quad$ $P(E_2) = ② + ③$

- The union of multiple E

$P(E_1 \cup E_2) = ① + ② + ③$

$$P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) + P(E_3)$$
$$- P(E_1 \cap E_2) - P(E_2 \cap E_3) - P(E_3 \cap E_1)$$
$$+ P(E_1 \cap E_2 \cap E_3)$$

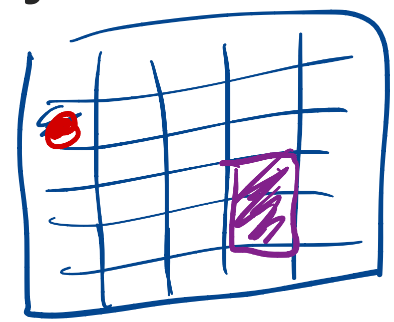
The Calculation of Probability

- Discrete countable finite event
- Discrete countable infinite event
- Continuous event

Counting to determine probability of countable finite event

- From the last axiom, the probability of event **E** is the sum of probabilities of the disjoint outcomes

$$P(E) = \sum_{A_i \in E}^{\mathbf{N}} P(A_i)$$



- If the outcomes have equal probability,

$E : 3$ or 6 or 1 $\quad |E| = 3 \quad |\Omega| = 10$

$$\frac{2x}{30x}$$

$$P(E) = \frac{number\ of\ outcomes\ in\ E}{total\ number\ of\ outcomes\ in\ \Omega} = \frac{3}{10}$$

$\Omega : 1, 2, \ldots 10$

Probability using counting: (1)

- Tossing a fair coin twice:
  - Prob. that it appears the same?

$$\Omega = \{HH, HT, TH, TT\}$$

$$E = \{HH\} + \{TT\} = \{HH, TT\}$$

$$P(E) = \frac{|E|}{|\Omega|} = \frac{1}{2}$$

  - Prob. that at least one head appears?

0.75

HH
HT
TH

$$P(E) = \frac{3}{4}$$

Probability using counting: (2)

- 4 rolls of a 5-sided die:

  E: they all give different numbers

- Number of outcomes that make the event happen: $|E|$ $|\Omega|$

  $5 \times 4 \times 3 \times 2 =$

- Number of outcomes in the sample space

  $|\Omega| = 5^4$

- Probability: $P(E) = \frac{|E|}{|\Omega|} =$

Probability using counting: (2)

- What about N-1 rolls of a N-sided die?

  E: they all give different numbers

  - Number of outcomes that make the event happen: $N \times N-1 \times \cdots \times 2$

  - Number of outcomes in the sample space

    N^{N-1}

  - Probability: $\frac{N!}{N^{N-1}}$

Probability by reasoning with the complement property

- If $P(E^c)$ is easier to calculate

$$P(E) = 1 - P(E^c)$$

Probability by reasoning with the complement property

- A person is taking a test with **N** true or false questions, and the chance he/she answers any question right is 50%, what's probability the person answers **at least** one question right?

E^c = none is right

$|E^c| = 1^N$ $|\Omega| = 2^N$

$(0.5)^N$

$P(E) = 1 - P(E^c) = 1 - 0.5^N$

Probability by reasoning with the union property

- If E is either E1 or E2

$$P(E) = P(E_1 \cup E_2) =$$

$$P(E_1) + P(E_2) - P(E_1 \cap E_2)$$

Probability by reasoning with the properties (2)

- A person may ride a bike on any day of the year equally. What's the probability that he/she rides on a Sunday or on 15^{th} of a month?

1 day in 2021 that is Sun. & 15th

E_1: Sun. E_2: 15th $E_1 \cap E_2$

$$P(E_1 \cup E_2) = \frac{52}{365} + \frac{12}{365} - \frac{1}{365}$$

Counting may not work

- This is one important reason to use the method of reasoning with properties

What if the event has infinite outcomes

- Tossing a fair coin until head appears
  - Coin is tossed at least 3 times

    This event includes infinite # of outcomes.

    And the outcomes don't have equal probability.

TTH, TTTH, TTTTH….

Assignments

- Work on Module Week2,
- Quiz1,
- HW1
- HW2

$$\frac{\overbrace{1 \times 1 \times \cdots}^{N}}{\underbrace{2 \times 2 \cdots 2}_{N}} = \frac{1}{2^N} = \left(\frac{1}{2}\right)^N$$

Commutative

$$A \cap B = B \cap A$$

$$A \cup B = B \cup A$$

Associative

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

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

Distributive

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

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

Idempotent

$A \cap A = A$

$A \cup A = A$

Identity

$A \cup \phi = A$

$A \cap \phi = \phi$

$A \cup \Omega = \Omega$

$A \cap \Omega = A$

Complement

$A \cup A^c = \Omega$

$A \cap A^c = \phi$

$\Omega^c = \phi$

$\phi^c = \Omega$

De Morgan's

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

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

Additional References

- Charles M. Grinstead and J. Laurie Snell "Introduction to Probability”
- Morris H. Degroot and Mark J. Schervish "Probability and Statistics”

See you next time

See You!