# STA257H1 F - Probability and Statistics I - Fall 2014

Instructor: Olga Chilina

Office: SS6002

Office hours:

Mondays, 2:00 - 3:00 pm

Wednesdays, 5:00 - 6:00 pm

Email: olgac@utstat.toronto.edu

Webpage:
http://www.utstat.toronto.edu/~olgac/sta257_2014/

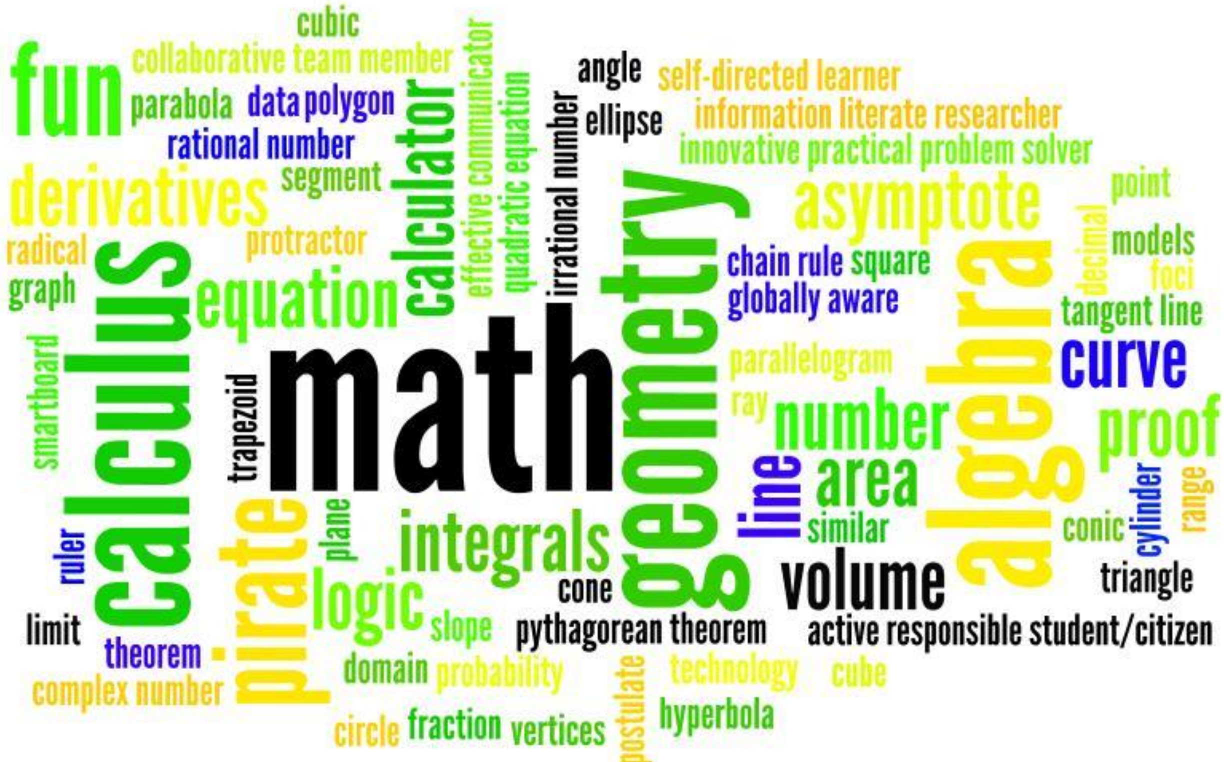
Lecture time and location:

L0101: Mon 3:00 - 5:00 pm, Wed 3:00 - 4:00 pm in MC102

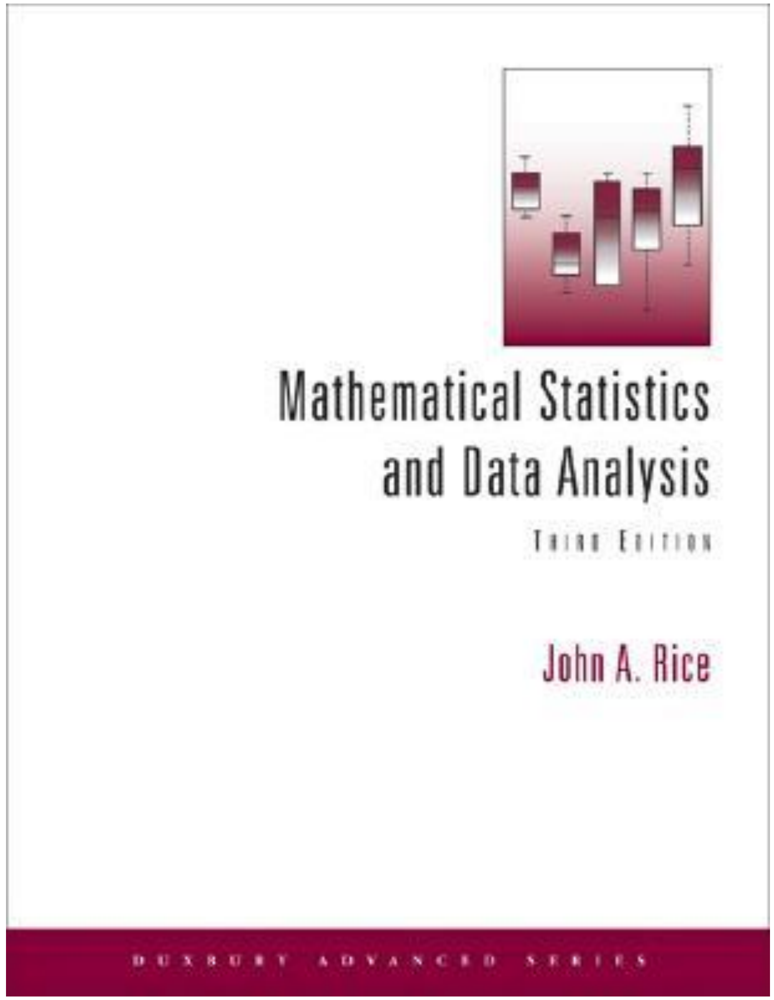
L5101: Wed 7:00 - 10:00 pm in SS2102

**Necessary Background**

You must be taking, or have taken and passed, a 2nd year calculus course as well as 2nd year linear algebra. Students not having this background should consider other alternative courses. Please do not attempt this course without the necessary background.



# Textbook




MATHEMATICAL STATISTICS AND DATA ANALYSIS, 3rd Edition, by John Rice

# Tutorials

Tutorials begin the week of Sept 15.

Tutorials meet every Wednesday:

L0101: 4:10 - 5:00 pm
L5101: 6:10 - 7:00 pm

Tutorial rooms will be posted on the course web site and blackboard prior to Sept 15.

Assignments will be posted on the course web site, consisting of suggested exercises, mostly from the textbook. Do not hand them in. Bring your solutions to tutorials, along with your questions about these exercises or the related theory and concepts.

Expect a quiz on the material as well. The first quiz is on Sept. 17!

# Evaluation

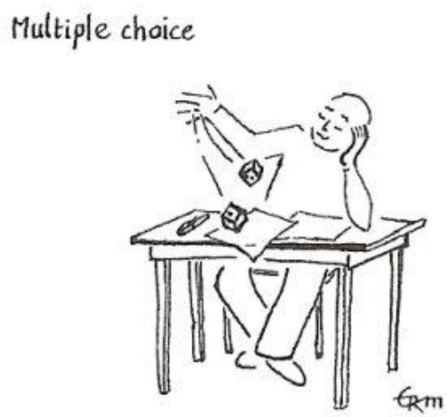

Tutorial weekly quizzes: 15%

Two-hour midterm test: 35%

Three-hour final exam: 50%

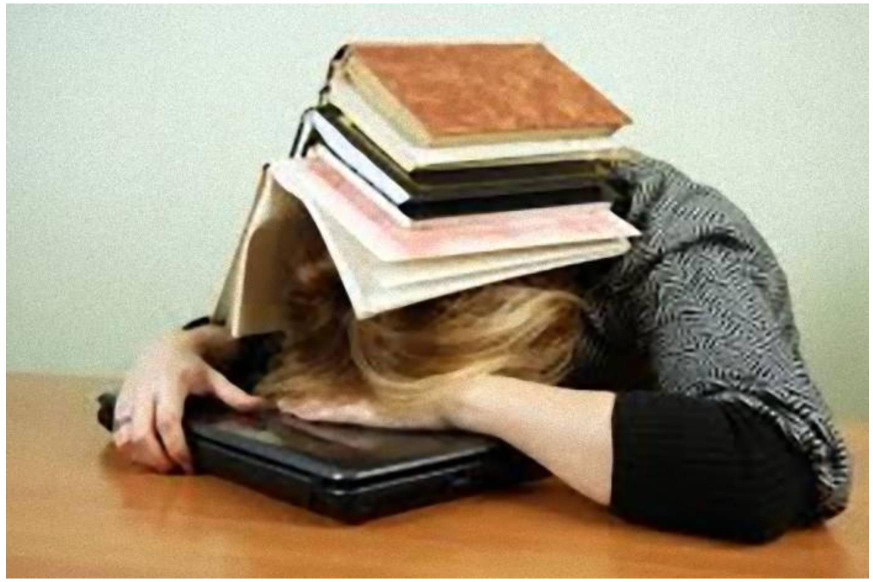
# Quizzes

- Quizzes will be given in tutorial.
- A typical quiz will be a multiple choice question, you get either 1 mark (for the correct answer) or 0.3 (for attendance).




- Your TA will record your mark for each quiz. So be sure to attend the correct tutorial, and to know your TA's name.
- ***If you miss a tutorial/quiz due to illness, late enrolment, etc., please discuss this matter with your TA, and not your lecturer.*** If ill, bring some proof.

## Midterm Test/Final Exam



The midterm test is on

L0101: Oct 20, 3:00 - 5:00 pm

L5101: Oct 22, 7:00 - 9:00 pm

**Programmable calculators are not permitted on test and exam**.

A one-sided 8-1/2"x 11" aid sheet, hand-written, is allowed on the test (two-sided on final exam).

**You must bring your student identification to term tests as well as the final exam.** The day and time for the final exam will be announced later.

# Missed Midterm Test

- There are **no make-up tests**.

- Should you miss the term test due to illness, you must submit to your lecturer, within one week, completed by yourself and your doctor, **the 'U of T Student Medical Certificate',** obtainable from your college registrar, the Office of the Faculty Registrar (SS1006), the Stats Dept. office, or the Koffler health service.

- The test's weight will then be shifted to the final exam.

- **If proper documentation is not received, your test mark will be zero.** ☹

# Additional help


- For continued discussion and questions outside of class, try posting on the Piazza discussion forums. The instructor and TAs will be monitoring them regularly.

- You can visit instructor (SS6002) and TAs (SS1091) during their office hours.

- There is a drop-in Statistics Aid Centre in New College: Wetmore Hall 68A. See http://www.utstat.toronto.edu/wordpress/?page id=154 for the schedule.

- E-mail should only be used for emergencies or personal matters.

# Tentative Lecture Outline

**Week 1**: Introduction to probability: set notations, Venn diagrams, probability models. Basic combinatorics.

**Week 2:** Rules of probability. Conditional probability. Law of total probability. Bayes' rule. Independence.

**Week 3:** Random variables. Discrete case: Bernoulli, Binomial, Geometric, negative Binomial, Hypergeometric, and Poisson distributions. Distribution function.

**Week 4:** Continuous case: Uniform, Exponential, Gamma, Beta, and Normal distributions. Density function. Poisson Processes.

**Week 5**: Expectation. Moments. Variance. Functions of random variables. Indicator functions and random variables.

**Week 6**: General Normal distribution. Chi-square distribution. Short review for the midterm.

**Week 7: TERM TEST on weeks 1- 6 material**

**Week 8:** Conditional probability on a joint distribution. Marginal density function. More on independence of random variables.

**Week 9**: Conditional densities. Covariance. Correlation. Conditional expectation. Markov's and Chebyshev's inequalities. Laws of large numbers.

**Week 10:** Convolution. Cauchy distribution. Jacobian, change of variables for two dimensional case. F and t distributions.

**Week 11**: Order statistics. Probability generating functions. Moment generating functions.

**Week 12**: Convergence in Distribution. Continuity theorem for mgf's. Central Limit Theorem. Different types of convergence.

Note: This corresponds to parts of Chapters 1-6 of the textbook and some additional material not found in the text.

## Let's get started!!!

Lecture 1

# Probability

H T

Example: When you toss a coin, there are only two possible outcomes, heads and tails.

$$P(H) = \frac{1}{2} = P(\frac{1}{2})$$



What if we toss a coin two times?

HH HT TH TT

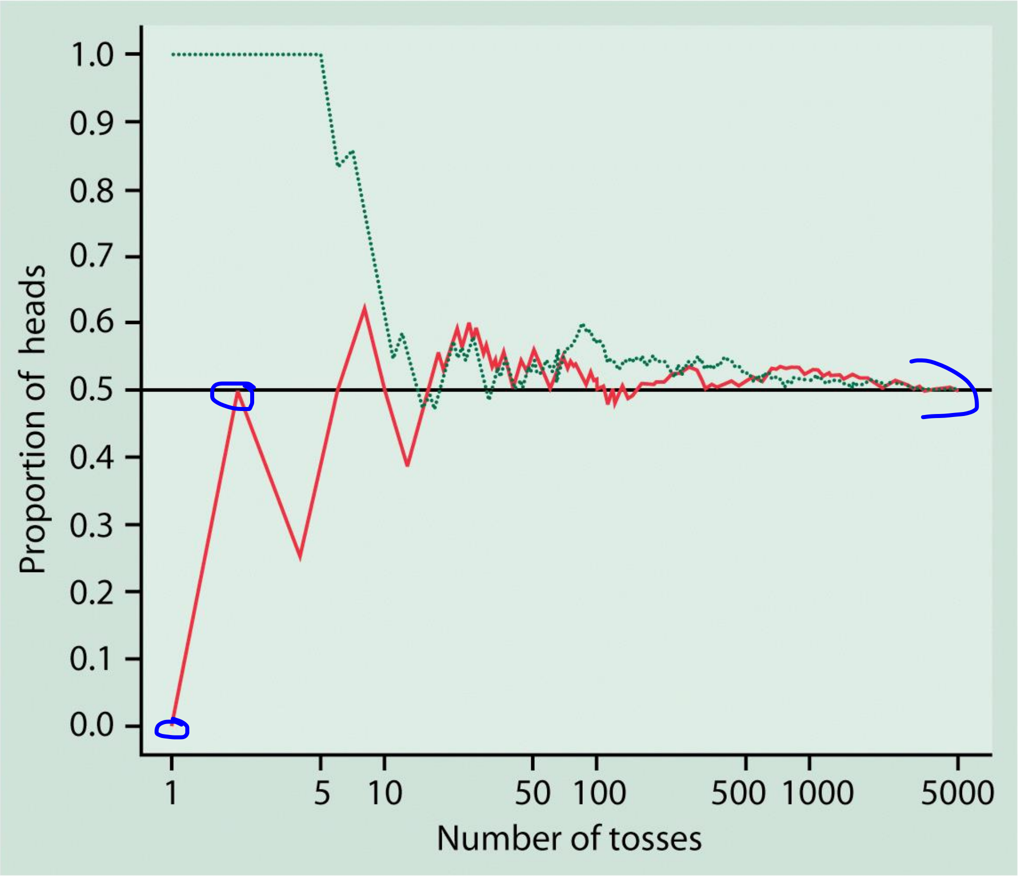
$$P(HH) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(H)P(H)$$

$$P(\text{exactly one } H)$$

$$= P(HT) + P(TH) = \frac{1}{2}$$

Figure below shows the results of tossing a coin 5000 times twice.

For each number of tosses from 1 to 5000, we have plotted the proportion of those tosses that gave a head.




Trial A (solid line) begins tail, head, tail, tail. Trial B starts with five straight heads, so the proportion of heads is 1 until the sixth toss.

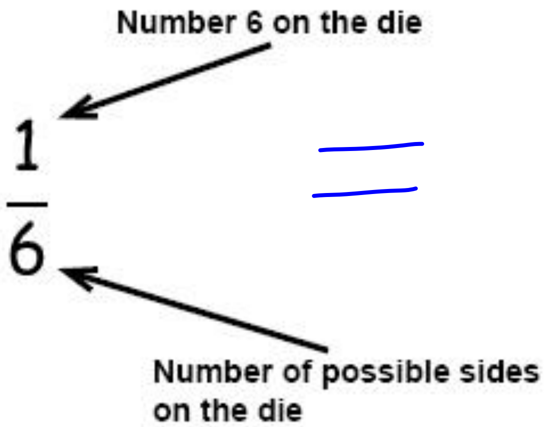
<u>Caution</u>: Probability describes only what happens in the long run

(so **probability of a head is 0.5**).

# Randomness

Definition: We call a phenomenon **random** if individual outcomes are uncertain but there is nonetheless a regular distribution of outcomes in a large number of repetitions.

The **probability** of any outcome of a random phenomenon is the proportion of times the outcome would occur in a very long series of repetitions.




$$= \frac{1000}{6000}$$

History of many tosses of a coin:

- The French naturalist Count Buffon (1707-1788) tossed a coin 4,040 times.

  Result:

  2,048 heads, or proportion 2,048/4,040 = 0.5069 for heads.

- Around 1900, the English statistician Karl Pearson heroically tossed a coin 24,000 times.

  Result:

  12,012 heads, a proportion of 0.5005.

- While imprisoned by the Germans during World War II, the South African statistician John Kerrich tossed a coin 10,000 times.

  Result:

  5,067 heads, proportion of heads = 0.5067.

## Set Notation

We shall use capital letters *A*, *B*, *C*,... to denote sets of elements.

<u>Example</u>: $A = \{1, 2, 3\}$

Let's denote by Ω the set of all elements under consideration, i.e. the set of all possible outcomes of a random phenomenon that we are interested in at the moment.

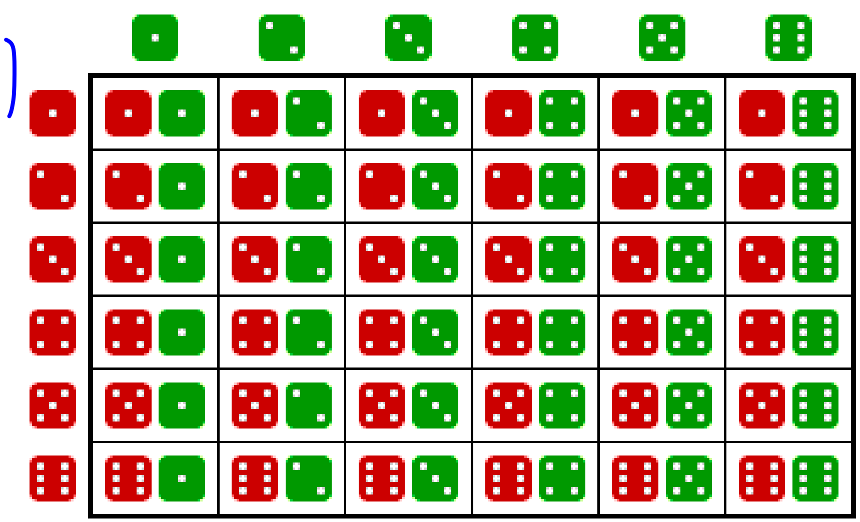
<u>Examples</u>:

(1) Toss a coin (we call a coin **fair** if the probability of heads is 0.5)

$\Omega = \{H, T\}$      $|\Omega| = 2$

(2) Roll two dice

$$\Omega = \{(1,1), (1,2), \ldots, (1,6), \ldots, (6,1), \ldots, (6,6)\}$$



$|\Omega| = 36$

Definition: We call $A$ a **subset** of $B$, $A \subseteq B$, if every element in $A$ is also in $B$, i.e.

$$\text{any } x \in A \Rightarrow x \in B$$

Example: $A = \{1, 2, 3\}$, $B = \{1, 2, 3, 4\}$

$$A \subseteq B \qquad A \subset B$$

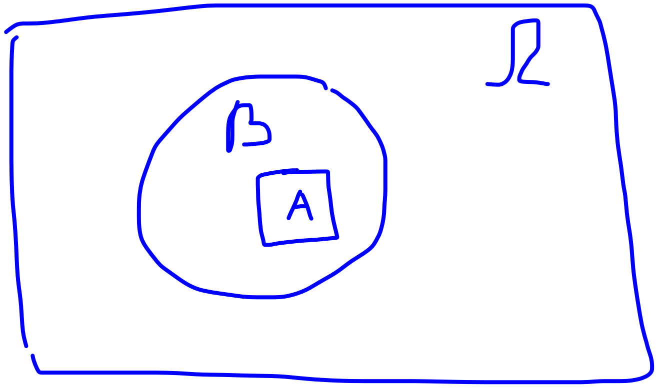
We say, $A$ is a **proper subset** of $B$, $A \subset B$, if $A$ is a subset of $B$, and there is at least one element in $B$ which is not in $A$.

$$\text{there is } y \in B \text{ s.t. } y \notin A$$

Definition: The **null**, or **empty set**, $\emptyset$, is the set consisting of no points. Thus, $\emptyset$ is a subset of every set.

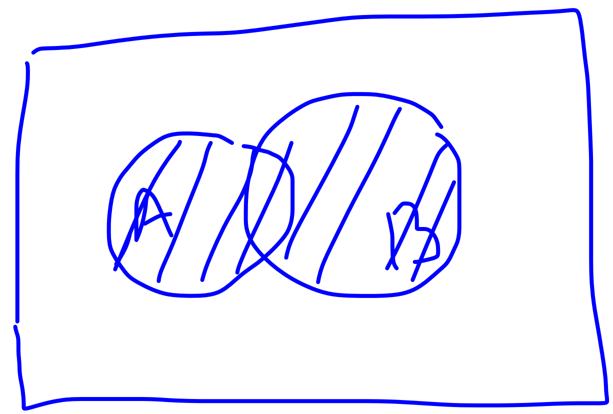
$$\emptyset \subseteq A, \text{ for any } A$$

## Venn Diagrams




The **union** of *A* and *B*, $A \cup B$, is the set of all elements in *A* or in *B*, or both, i.e.

$$A \cup B = \{x \in \Omega : x \in A \text{ or } x \in B\}$$




<u>Example</u>:

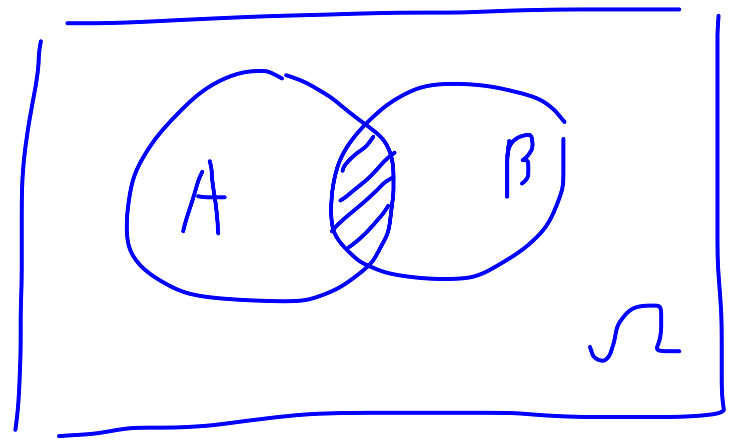
$$A = \{1, 2, 3\}$$

$$B = \{2, 3, 7, 9\}$$

$$A \cup B = \{1, 2, 3, 7, 9\}$$

The **intersection** of $A$ and $B$, $A \cap B$, is the set of all elements in both $A$ and $B$, i.e.

$$A \cap B = \{x \in \Omega : x \in A \text{ and } x \in B\}$$




Example: $A = \{1, 2, 3\}$

$B = \{2, 3, 7, 9\}$

$A \cap B = \{2, 3\}$

If $A \subset \Omega$, then the **complement** of $A$, $\bar{A}$ or $A^c$, is the set of elements that are in $\boldsymbol{\Omega}$ but not in $A$, i.e.

$$A^c = \{x \in \Omega : x \notin A\}$$


Example: $\Omega = \{1, 2, 3, \ldots, 10\}$

$A = \{1, 2, 3\}$

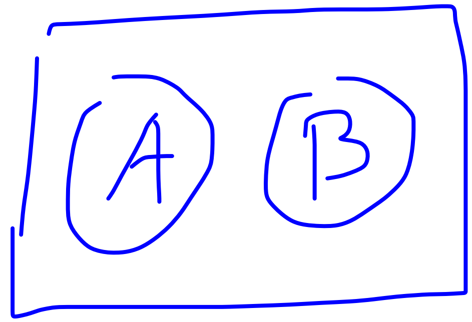
$A^c = \{4, 5, 6, \ldots, 10\}$

$A \cup A^c = \Omega$

$A \cap A^c = \emptyset$

$(A^c)^c = A$

Two sets, $A$ and $B$, are **disjoint**, or **mutually exclusive**, if they don't have any elements in common, i.e. $A \cap B = \emptyset$




Example: $\Omega = \{1, \dots, 10\}$

$A = \{1, 2, 3\}$

$B = \{5, 7\}$

$A \cap B = \emptyset$

Ex: Roll a die

$A = \{\text{odd \#}\} = \{1, 3, 5\}$

$B = \{\text{even \#}\} = \{2, 4, 6\}$

$A \cap B = \emptyset$

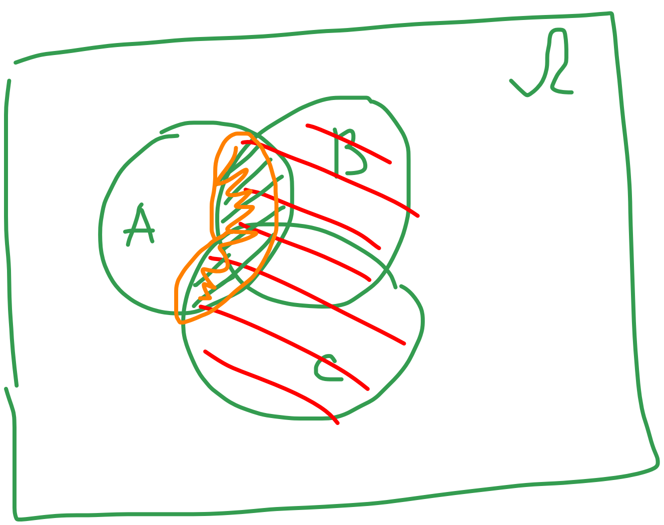
## Distributive Laws for Sets

$a(b+c) = ab + ac$

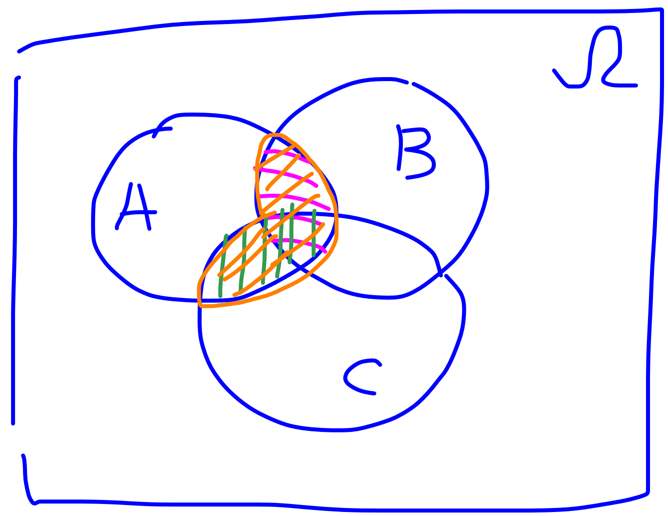
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \rightarrow \text{exe}$$

<u>Proof</u>:




=




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

Method 2: Take $x \in A \cap (B \cup C)$. Show $x \in (A \cap B) \cup (A \cap C)$

## De Morgan's Laws

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

$$(A \cap B)^c = A^c \cup B^c \longrightarrow \text{exe}$$

Proof: Show any $x \in (A \cup B)^c \overset{?}{\iff} x \in A^c \cap B^c$

$x \in (A \cup B)^c \iff x \notin A \cup B \iff x \notin A$ and $x \notin B$

$\iff x \in A^c$ and $x \in B^c \iff x \in A^c \cap B^c$

Note: For any collection of sets $A_1, A_2, A_3, \ldots$ in $\Omega$

$$\boxed{\left(\bigcup_{n=1}^{\infty} A_n\right)^c = \bigcap_{n=1}^{\infty} (A_n)^c}$$

$$\left(\bigcap A_n\right)^c = \bigcup A_n^c$$

**Finite, Countable Infinite, and Uncountable**

- A set A is **finite** if it contains a finite number of elements
- A set A is **countable infinite** if it can be put into a one-to-one correspondence with the set of positive integers, ℕ.

Example:

$A = \{1, 2, 3\} \rightarrow$ finite

$\mathbb{Q} = \{\text{rationals}\} \rightarrow$ countable

$\mathbb{Z} = \{\text{integers}\}$

$\mathbb{I} = \mathbb{Q}^c = \{\text{irrationals}\} \rightarrow$ uncountable

$\mathbb{R} = \mathbb{Q} \cup \mathbb{Q}^c$

$(a, b)$

interval

uncountable

## Probability Model

A description of a random phenomenon in the language of mathematics is called a **probability model**.

For example, to build a probability model for a coin tossing example we have to

- List all possible outcomes
- Assign probabilities to those outcomes

$X = \#$ of heads

$$X: \quad 0 \quad 1$$

$$P: \quad \frac{1}{2} \quad \frac{1}{2}$$

Sometimes we need a more complex definition of probability models.

Definition: An **experiment** is the process by which an observation is made, and it can result in one or more outcomes that we shall call **events**.

Example: We toss a die. Consider the following events:

A = {we observe an odd number} $= \{1, 3, 5\}$

B = {we observe a number less than 4} $= \{1, 2, 3\}$

C = {we observe either 4 or 6} $= \{4, 6\}$

$E_1$ = {we observe 1}, $E_2$ = {we observe 2}, ... , $E_6$ = {we observe 6}

$A = E_1 \cup E_3 \cup E_5$

An event *A*, which can be decomposed into other events, is called a **compound event**. An event that cannot be decomposed is called a **simple event**.

Each simple event corresponds to one and only one element, called a **sample point**.

$\{1\}, \{2\}, \ldots, \{6\}$ are sample points

Definition: The set consisting of all possible sample points is called the **sample space**, **Ω**, associated with an experiment.

# <u>Ingredients for Probability Model</u>:

## - Discrete Sample Space

- $\Omega$ is a discrete sample space if it contains either a finite or a countable number of distinct sample points.
- For a discrete sample space it suffices to assign probabilities to each sample point.
- There are experiments for which the sample space is not countable and hence is not discrete. For example, the experiment that consists of measuring the blood pressure of patients with heart disease.

## - $\boldsymbol{\sigma}$-algebra

A **$\boldsymbol{\sigma}$-algebra**, $\mathcal{F}$, is a collection of subsets of $\Omega$ satisfying the following properties:

- $\mathcal{F}$ contains $\emptyset$ and $\Omega$
- $\mathcal{F}$ is closed under taking complement, i.e. $A \in \mathcal{F} \Longrightarrow A^c \in \mathcal{F}$
- $\mathcal{F}$ is closed under taking countable union, i.e. $A_1, A_2, A_3, \ldots \in \mathcal{F} \Longrightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

Claim: $\mathcal{F}$ is closed under taking countable intersection.

Proof: $A_1, A_2, A_3, \ldots \in \mathcal{F} \overset{?}{\Longrightarrow} \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$

by def. $\Rightarrow A_1^c, A_2^c, A_3^c, \ldots \in \mathcal{F}$

$$\Longrightarrow \bigcup_{i=1}^{\infty} A_i^c \in \mathcal{F} \Longrightarrow \left(\bigcup A_i^c\right)^c \in \mathcal{F}$$

$$\left(\bigcup A_i^c\right)^c \underset{\text{De Morgan}}{=} \bigcap \left(A_i^c\right)^c = \bigcap A_i \in \mathcal{F}$$

$\blacksquare$

## - Probability Measure

Let Ω be a sample space associated with an experiment.

To every event $A \in \mathcal{F}$, we assign a number $P(A)$, called the **probability measure** of $A$, mapping $\mathcal{F} \to [0,1]$, so that

- $0 \leq P(A) \leq 1$
- $P(\Omega) = 1$   $\Omega$ - sure event
- If $A_1, A_2, A_3, \ldots \in \mathcal{F}$, $A_i \cap A_j = \emptyset, i \neq j$, then
$$P(A_1 \cup A_2 \cup A_3 \cup \ldots) = P(A_1) + P(A_2) + P(A_3) + \cdots$$
$$= \sum_{i=1}^{\infty} P(A_i) \longrightarrow \text{countable additivity}$$

Finite case: $P(A_1 \cup \ldots \cup A_n) = P(A_1) + \cdots + P(A_n)$

Example: We toss two balanced coins.

Let $A$ = {exactly one head} $= \{TH, HT\}$

$B$ = {at least one head} $= \{TH, HT, HH\}$

Find $P(A), P(B), P(A \cap B), P(A \cup B)$, and $P(\bar{A} \cup B)$.

Solution: $\Omega = \{HH, TH, HT, TT\} = \{\omega_1, \omega_2, \omega_3, \omega_4\}$

$P(A) = \frac{1}{2}$ $\quad A = \{\omega_2, \omega_3\}$ $\quad P(\omega_i) = \frac{1}{4}$

$P(B) = P(A) + P(HH) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$

$B = A \cup \{HH\}$

$$\boxed{\begin{aligned} P(A) &= P(\omega_2) + P(\omega_3) \\ &= \frac{1}{4} + \frac{1}{4} \\ &= \frac{2}{4} = \frac{1}{2} \end{aligned}}$$

$P(A \cap B) = P(A) = \frac{1}{2}$

$P(A \cup B) = P(B) = \frac{3}{4}$

$P(A^c \cup B) = P(\Omega) = 1$

Proposition: $P(\emptyset) = 0$

Proof:

For any $A$, $A = A \cup \emptyset$, $A \cap \emptyset = \emptyset$

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

$$\Rightarrow P(\emptyset) = 0$$

**Three Ingredients**: $(\Omega, \mathcal{F}, P)$

A *probability space* consists of three elements:

- set $\Omega$
- σ-algebra $\mathcal{F}$
- probability measure $P$

## Calculating Probabilities

Suppose $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$. For any $A$,

$$\boxed{P(A) = \sum_{\omega_i \epsilon A} P(\omega_i)}$$

If the outcomes are equally likely, then

$$P(\omega_1) = \cdots = P(\omega_n) = \frac{1}{n}$$

$$P(A) = \frac{\#\ of\ outcomes\ for\ which\ A\ occurs}{n} = \frac{|A|}{n}$$

where $|A|$ is the number of elements in $A$.

<u>Example</u>: Roll a die and let $A$ = {even number} $= \{2, 4, 6\}$

$$P(A) = \frac{|A|}{n} = \frac{3}{6} = \frac{1}{2}$$

$n = 6$, $\Omega = \{1, 2, \dots, 6\}$

# Counting Methods

## - Multiplication Principle

Suppose we want to make a series of decisions. There are $c_1$ choices for decision 1 and for each of these there are $c_2$ choices for decision 2, etc. Then the number of ways the series of decisions can be made is

$$c_1 \cdot c_2 \cdot c_3 \cdots$$

Example: You have 2 pairs of jeans, 3 shirts and 2 pairs of shoes. In how many ways you can choose an outfit?

$$2 \cdot 3 \cdot 2 = 12$$

Quiz

Example: How many 3 letter words can be composed from the English Alphabet if

(a) There is no limitation
(b) The words have 3 different letters

(a) $\underline{\quad}\ \underline{\quad}\ \underline{\quad}$
$26 \cdot 26 \cdot 26 = 26^3$

(b) $\underline{\quad}\ \underline{\quad}\ \underline{\quad} = 26 \cdot 25 \cdot 24$
$26 \quad 25 \quad 24$

Example: How many 3 digit numbers are there?

$\underline{\quad}\ \underline{\quad}\ \underline{\quad}$
$9 \quad 10 \quad 10 = 900$

all digits different: $\underline{\quad}\ \underline{\quad}\ \underline{\quad}$
$9 \quad 9 \quad 8 = 648$

## - Combinations

The number of **combinations** of *n* objects taken *k* at a time is the number of subsets, each of size *k*, that can be formed from the *n* objects.

$\{1, 2, 3, 4\}$

$n = 4$

$k = 2$

Notation: $C_k^n$ or $\binom{n}{k}$ ($n$ choose $k$)

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

$n! = n(n-1) \cdot \ldots \cdot 2 \cdot 1$

$5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$

$\{1,2\}, \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}, \{3,4\}$

Facts:

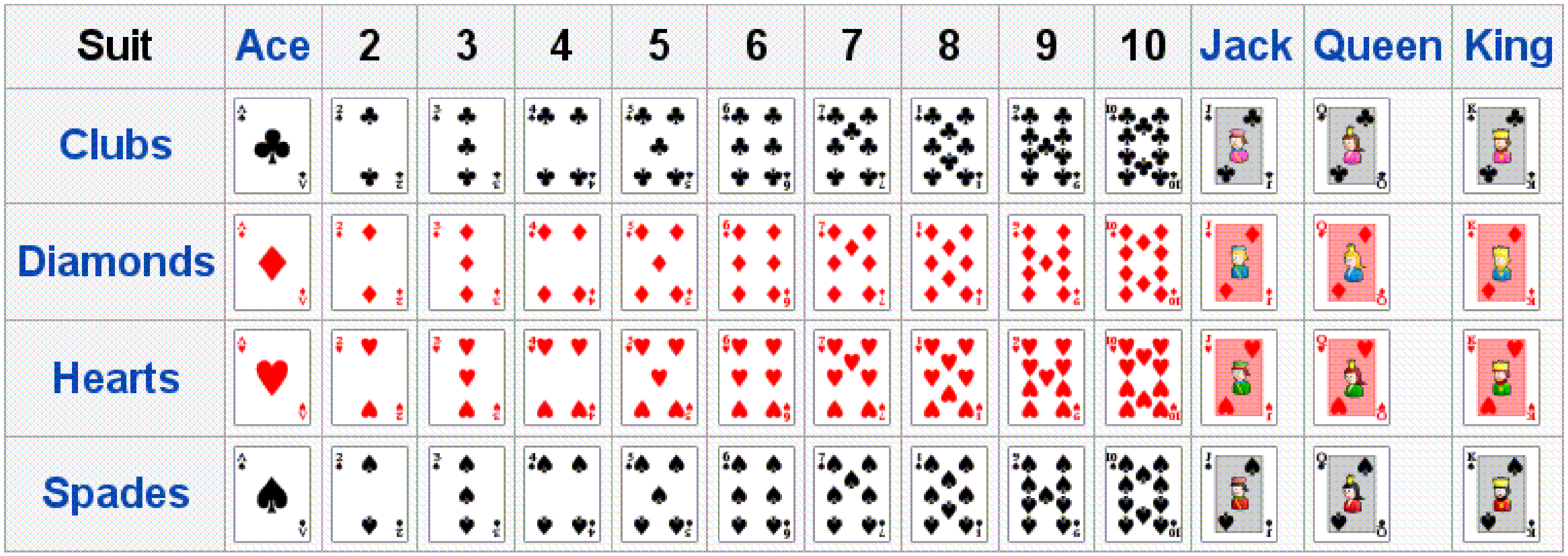
- $\binom{n}{0} = 1 = \binom{n}{n}$
- $\binom{n}{1} = n = \binom{n}{n-1}$
- $\binom{n}{k} = \binom{n}{n-k}$

order does not matter

$\binom{4}{2} = \frac{4!}{2!\,2!} = \frac{4 \cdot 3}{2} = 6$

<u>Example</u>: Two cards are drawn from a standard 52-card playing deck.

**Example set of 52 poker playing cards**




In how many different ways can you do that? $\Omega = \{\text{two cards}\}$

$$|\Omega| = \binom{52}{2} = \frac{52!}{2!\,50!} = \frac{52\cdot 51}{2!} = 26\cdot 51 = 1326$$

In how many ways can you pick up an ace and a face card?

$$|A| = |\{\text{ace \& face}\}| = 4\cdot 12 = 48$$

$$P(A) = \frac{48}{1326}$$

## - Permutations

An ordered arrangement of *n* distinct objects is called a **permutation**.

The number of permutations of *k* subjects selected from *n* distinct objects is

$$n(n-1)(n-2)\cdots(n-k+1)$$

$$\frac{n!}{}$$

Notation: $P_k^n = n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}$ $> \quad C_k^n = \frac{n!}{k!(n-k)!}$

$\Omega = \{1, 2, 3\}$

Example: $n = 3$, $k = 2$

$n = 3, \; k = 2$

$P_2^3 = \frac{3!}{1!} = 6$

$C_2^3 = \frac{3!}{2!1!} = 3$

combinations $\leftarrow$ $\{1,2\}, \{1,3\}, \{2,3\}$

Permutations $\leftarrow$ $\underbrace{(1\,2)\;(2\,1)\;(1\,3)\;(3\,1)\;(2\,3)\;(3\,2)}$

Note: The number of ordered arrangements of *k* subjects selected with replacement from *n* objects is $n^k$

with replacement: $+\;(1\,1)\;(2\,2)\;(3\,3)$

$9 = 3^2 = n^k$

**Binomial Theorem**

For any numbers *a*, *b*, and any positive integer *n*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

$\binom{n}{k}$ is called a **binomial coefficient**.

$$(a+b)^2 = 1 \cdot b^2 + 2 \cdot ab + 1 \cdot a^2$$

$n = 2$

$$\binom{n}{k} = \binom{2}{k}$$

$k = 0 \quad \binom{2}{0} = 1$

$k = 1 \quad \binom{2}{1} = 2$

$k = 2 \quad \binom{2}{2} = 1$

Fact: The number of subsets of a set of size $n$ is $2^n$.

Proof:

Binomial Thm: $$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \quad (*)$$

Ex: $\{1,2,3\}$, $n=3$

The subsets are:

$\{1\}, \{2\}, \{3\}$

$\{1,2\}, \{1,3\}, \{2,3\}$

$\{1,2,3\}, \emptyset$

8 of them

$8 = 2^3$

\# of all subsets =

= \# of subsets of size 0

+ \# of subsets of size 1

+ \# of subsets of size 2

⋮

+ \# of subsets of size n

$$= \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n}$$

$$= \sum_{k=0}^{n} \binom{n}{k} = \sum_{k=0}^{n} \binom{n}{k} 1^k 1^{n-k} \overset{(*)}{=} (1+1)^n$$

$$= 2^n \quad \blacksquare$$

## Multinomial Coefficients

The number of ways to partition $n$ distinct objects into $k$ distinct groups containing $n_1, n_2, \dots, n_k$ objects respectively, where each object appears in exactly one group and $\sum_{i=1}^{k} n_i = n$, is

$$\boxed{\binom{n}{n_1\ n_2\ \dots\ n_k} = \frac{n!}{n_1!\,n_2!\cdots n_k!}}$$

$$(a_1 + a_2 + \cdots + a_k)^n = \sum_{n_i} \underbrace{\binom{n}{n_1\ n_2\ \dots\ n_k}} a_1^{n_1} a_2^{n_2} \cdots a_k^{n_k}$$

<u>Example</u>: A committee of 10 members is to be divided into three subcommittees of size 4, 3, and 3. In how many ways can it be done?

$n = 10, \quad k = 3$

$n_1 = 4, \quad n_2 = 3, \quad n_3 = 3$

$$\sum_{i=1}^{3} n_i = 4 + 3 + 3 = 10 = n$$

$$\binom{n}{n_1\ n_2\ n_3} = \binom{10}{4\ 3\ 3} = \frac{10!}{4!\,3!\,3!} = 4200$$