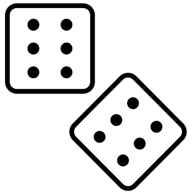
Welcome to Advanced Data Analysis (PHYS 605)

Prof. Claudia Gomes da Rocha

claudia.gomesdarocha@ucalgary.ca

Department of Physics and Astronomy
Faculty of Science, University of Calgary

Probability



Chance? Likelihood? What comes to your mind when referring to probability?

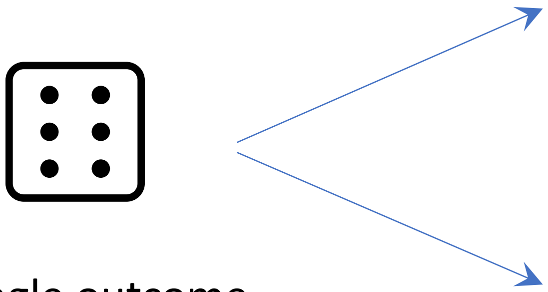
The probability of something occurring is the quantification of the chance of observing a particular outcome given an event.

If we consider a sample set $\Omega_X = \{x_i\}$ that contains all possible elementary events x_i, then we can define the probability of obtaining a certain outcome, or event, as $P(x_i)$.

An event may be the result of a single experiment, or one single data point collected by an unrepeatable experiment. Many data points can be collected from an ensemble of events.

Quantifying the probability of a repeatable experiment can be used to make predictions of the outcomes of future experiments. We cannot predict the outcome of a given experiment with certainty; however, <u>we can assign a level of confidence to our predictions</u>.

Probability

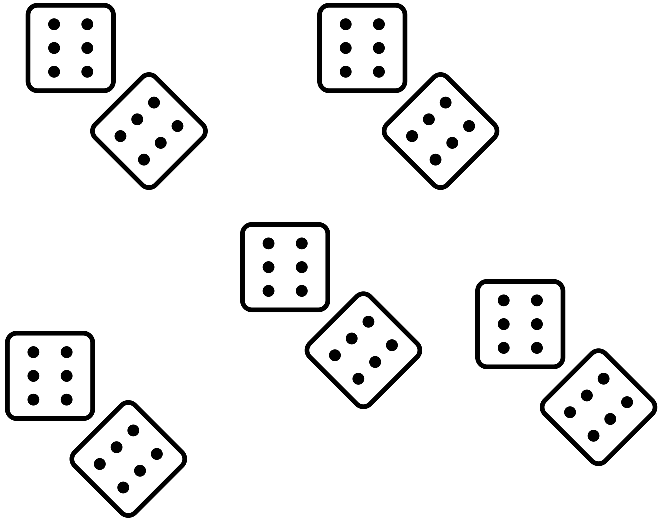


Agree with predictions

Disagree with predictions

A single outcome from any random process…

However, it is very difficult to get any sense of the prediction quality from only one test.



Repeated “trials”: in physics, we often use the concept of an “ensemble” containing some N number of identically prepared systems.

In practice, a set of multiple trials is often obtained by repeated trials of a single system. The trials are expected to be independent. For that, it may require waiting between trials for some characteristic correlation time in order to let the system “forget” the previous trial.

Or try to prepare as many identical systems as possible and perform independent measurements at them.

This discussion often invokes the ‘**ergodic hypothesis**’, but without careful thought about its validity.

Do you remember what the ergodic hypothesis is? If not, RECAP!

Assigning probabilities to possible outcomes is one of the key tasks of the **theory of probability**.

1. Classical method based on the repetition of the experiment.

2. Method based on empirical knowledge of the experiment.

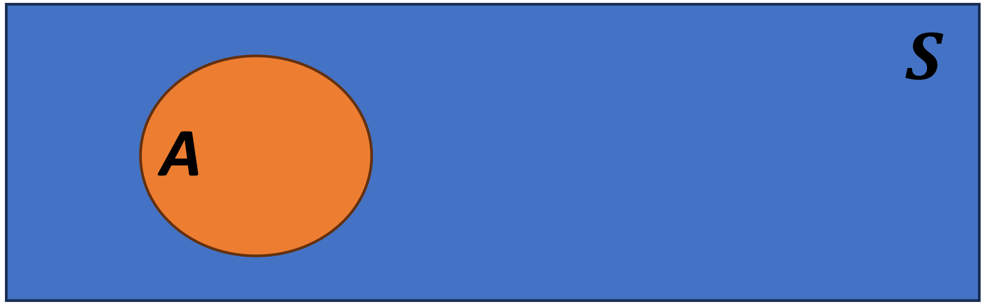
<u>There is an element of subjectivity that enters the analysis and the interpretation of results when dealing with probabilities.</u>

It is therefore the task of the researcher or analyst to keep track of any assumptions made and to account for them in the interpretation of the results (and to communicate those clearly and transparently to others).

Set theory

Sample space S: set of all possible outcomes of an experiment. For example, the sample space of the roll of a die is $S = \{1,2,3,4,5,6\}$.

Event A: is a subset of S so $A \subset S$ and it can represent a number of possible outcomes for the experiment. For example, an event *A* called "even number" in the die roll experiment can be represented as $A = \{2,4,6\}$. If the outcome of an experiment is included in *A*, then event *A* has occurred.



There is always one subset called *empty* set that contains <u>no outcomes</u>, i.e., $A = \emptyset$ ("impossible event").

On the other hand, a subset can contain *all possible outcomes*, so that the event subset equals the sample space, i.e., $A \subseteq S$.

Set theory

A given set with *n* elements $A = \{a_1, a_2, \dots, a_n\}$, we can use the mathematical notation to indicate if an element $a_i \in A$ (is an element of *A*) or $a_i \notin A$ (is not an element of *A*).

A set with *n* elements has a total of 2^n subsets. For instance, the sample space of a coin toss is $S = \{heads, tails\}$ and that can have 4 subsets $\{\emptyset\}, \{heads\}, \{tails\}, \{heads, tails\}$.

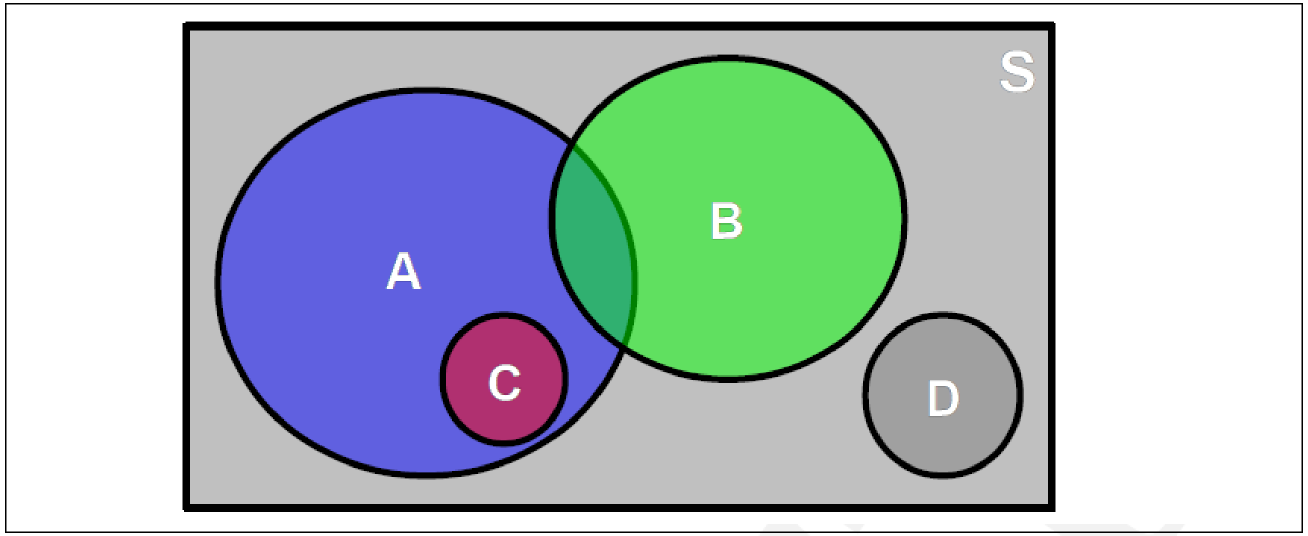


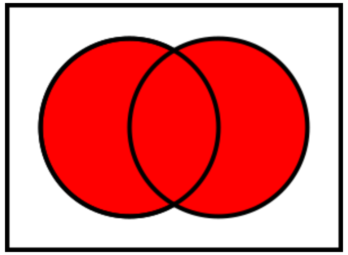
Figure 4.1: Venn diagram with subset $C \subset A$ (or superset $A \supset C$) and intersection (product) $A \cap B$.

Sets can be represented in a Venn diagram by circles drawn inside a rectangle representing the universal set.

Union (OR)

The union of two sets contains all elements that are present in the first set, in the second set, or in both sets. The sum or union of two sets *A* and *B*

$$A + B \quad \leftrightarrow \quad A \cup B$$

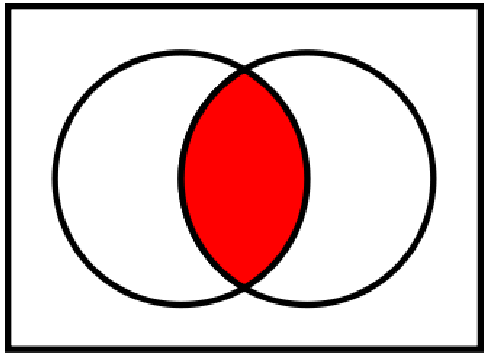


is a set containing elements of both sets. This operation is commutative and associative:

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

For example, if A = {1, 3, 5, 7} and B = {1, 2, 4, 6, 7} then A ∪ B = {1, 2, 3, 4, 5, 6, 7}.

Intersection (AND)



The product or intersection of two sets A and B

$$AB \quad \leftrightarrow \quad A \cap B$$

is a set consisting of all elements that are common to both sets. It is the set containing all elements of A that also (AND) belong to B; equivalently, all elements of B that also (AND) belong to A. This operation is commutative, associative, and distributive:

$$AB = BA \qquad (AB)C = A(BC) \qquad A(B \cup C) = AB \cup AC$$

For example, if A = {1, 2, 3} and B = {2, 3, 4} then A ∩ B = {2, 3}.

Two sets A and B are *mutually exclusive* if they have NO common elements, i.e., $AB = \{\emptyset\}$.
Several sets $A_1, A_2, \ldots$ are mutually exclusive if each pair of sets is mutually exclusive, i.e.,

$$A_iA_j = \{\emptyset\} \qquad \forall\, i \neq j$$

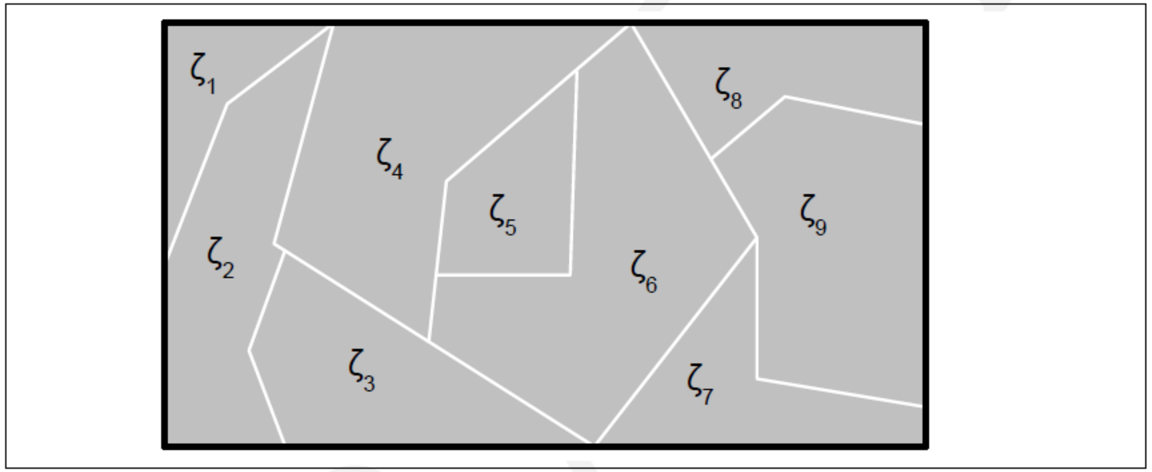
The union and intersection operations can be extended to more than two events!

Partition

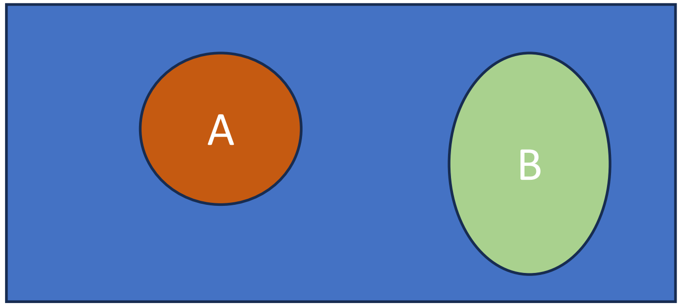
A number of events or a collection of subsets represented by ζ_i is said to be a *partition* of the sample space if they satisfy the two properties:

$$\zeta_1 \cup \cdots \cup \zeta_n = \bigcup_{i=1}^{n} \zeta_i = S$$

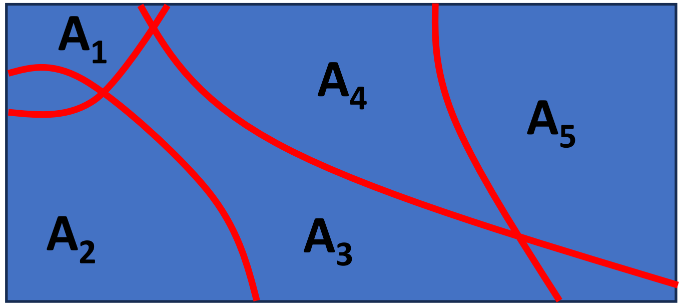
$$\zeta_i \zeta_j = \{\emptyset\} \quad \text{for} \quad i \neq j$$



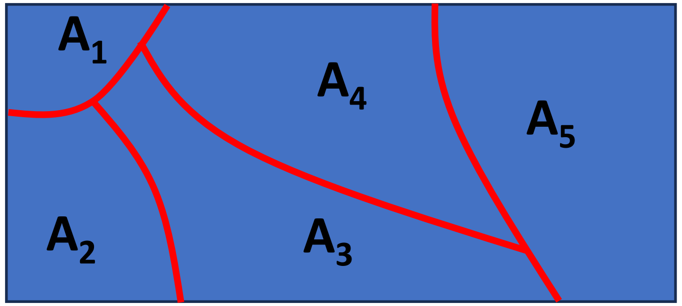
MUTUALLY EXCLUSIVE



COLLECTIVELY EXHAUSTIVE



BOTH: MUTUALLY EXCLUSIVE SETS OF EVENTS PARTION THE SAMPLE SPACE WITHOUT INTERSECTIONS!



Example:

The set {1, 2, 3} has these five partitions (one partition per item):

{ {1}, {2}, {3} }
{ {1, 2}, {3} }
{ {1, 3}, {2} }
{ {1}, {2, 3} }
{ {1, 2, 3} }

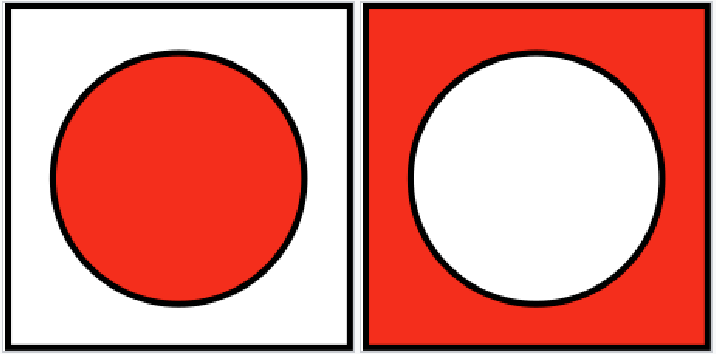
The following are not partitions of {1, 2, 3}:

{ {}, {1, 3}, {2} } is not a partition (of any set) because one of its elements is the empty set.
{ {1, 2}, {2, 3} } is not a partition (of any set) because the element 2 is contained in more than one block.
{ {1}, {2} } is not a partition of {1, 2, 3} because none of its blocks contains 3; however, it is a partition of {1, 2}.

Complement

A complement $\bar{A}$ of set A consists of all elements of S that are NOT in A:

$$A \cup \bar{A} = S \qquad A\bar{A} = \{\emptyset\} \qquad \bar{\bar{A}} = A \qquad \bar{S} = \{\emptyset\}$$



If A is the area colored red in this image… … then the complement of A is everything else.

Probability and set theory

The concept of probability embodies the notion of randomness, i.e., the fact that one cannot predict with complete certainty the outcome, or value, of a particular measurement. Probability can be formally defined in the context of **set theory**. Indeed, measurement outcomes can be viewed as members of a set.

A subset *A* of S, noted $A \subseteq S$, may be empty or it may contain one, few, or all elements of S. It is then possible to assign the probability $p(A)$ of some outcome *A* in terms of the outcome of multiple trials. For *N* trials, we expect that the number of cases n_A will be roughly proportional to *N*

$$n_A \approx p(A)N$$

This suggests an empirical definition of probability as

$$p(A) \equiv \lim_{N\to\infty} \frac{n_A}{N}$$

Probability and set theory

$$f_A = \frac{n_A}{N} \approx p(A)$$

The fraction f_A can be seen as a relative frequency associated with the phenomenon of interest. This "frequency" can provide an indication of the likelihood, that is, the **probability** of an event to occur. This is also known as the "frequentist interpretation" of the notion of probability.

Note that $\langle f_A \rangle$ will converge to $p(A)$ for large *N*, but the result of any given sequence of trials may fluctuate around the average.

The probability of an event *A* ranges from zero (an outcome that cannot happen) to unity (an outcome that can always happen):

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

Probability and set theory

Kolmogorov Axioms:

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

$$p(S) = 1$$

If A and B are two mutually exclusive sets, i.e., $AB = \{\emptyset\}$, then $p(A \cup B) = p(A) + p(B)$.

PS: If $AB \neq \{\emptyset\}$, then the third axiom can be generalized to $p(A \cup B) = p(A) + p(B) - p(A \cap B)$ to account that outcomes in the overlap region should be counted only once.

The axioms above set some “ground rules” of probability theory, but they provide NO guidance on how to assign probabilities.

How to assign probabilities??

Probability and set theory

But note!

Given two independent events x_i and x_j, where $i \neq j$, the probability of x_i occurring is independent of the probability of x_j happening, and the probability for one OR the other event to occur is given by $P\left(x_i \; or \; x_j\right) = P\left(x_i \; \cup \; x_j\right) = P(x_i) + P\left(x_j\right)$

Given two independent events x_i and x_j, the probability of x_i occurring is independent of the probability of x_j happening, and the probability for both events to occur is given by $P\left(x_i \; AND \; x_j\right) = P\left(x_i \; \cap \; x_j\right) = P(x_i)P\left(x_j\right)$.

The assignment of probabilities requires interpretations; one we already cited is the *frequentist (classical) interpretation* which is based on the repetition of experiments a large number of times under the same conditions.

This interpretation should be considered if you do have access to repeated experiments!

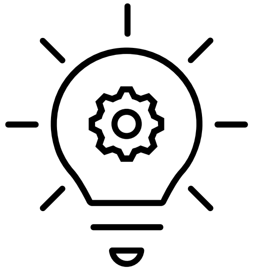
But what if you are dealing with phenomena that you do not have access to repeated experiments, or they consist of events that occur just so rarely?

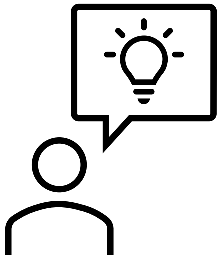
Plausibility? Risk assessment? Truthfulness tests or test of hypotheses?

The other interpretation that can be used to assign probabilities is based on prior knowledge of the phenomena or experiment but without the requirement of a large number of repetitions. This is referred to as the *Bayesian interpretation*.

The probability assigned to an event represents the *degree of belief* that an event will occur in a given try of the experiment, bringing an element of subjectivity to the study.

Bayesian probabilities are assigned based on a quantitative or practical understanding of the nature of the experiment and in accordance with the Kolmogorov axioms.





Conditional probability

Describes the occurrence of an event A, knowing that another event B has also occurred (by assumption, presumption, assertion or evidence). This can also be understood as the fraction of probability B that intersects with A:

Probability of A and B

$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$

Probability of A given B

Probability of B

Important probability rules

$$p(A) = 1 - p(\bar{A})$$

$$p(A \cap B) = p(B|A)\, p(A)$$

If A and B are independent: $p(A \cap B) = p(B)\, p(A)$
If A and B are mutually exclusive: $p(A \cap B) = 0$

The concept of conditional events and independent events determines whether or not one of the events has an effect on the **probability** of the other event.

Conditional probability

For instance, a team might have a probability of 30% of winning their local soccer tournament in the final game.

- Conditional Scenario: if it rains the team's chances may change (for the better or possibly for the worse). The probability of winning is affected by the weather - conditional.

- Independent Scenario: if the game is played in an enclosed stadium, in such a case the weather may have no effect on the team's chances - independent.

Example

$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$

Calculate the probability of obtaining 8 as the sum of two rolls of a die, given that the first roll was a 3.

IMPORTANT: there is only 1 die that is thrown twice.