UNIVERSITY OF LINCOLN

SCHOOL OF MATHEMATICS AND PHYSICS

MTH 1005
PROBABILITY AND STATISTICS

Semester B

(7/2/2023)

Danilo Roccatano

Office: INB 3323
Email: droccatano@lincoln.ac.uk

SUMMARY OF LAST LECTURE

We have introduced the following concepts

- A short introduction to the graphical representation of data.
- Sample space, outcomes, events and probability distribution.
- Frequentist paradigm of probability

PROBABILITY

In this part we are going to learn the axioms of probability and connects them to counting problems –including combinatorial and permutations.

Learning outcomes:

- Sample spaces
- Outcomes
- Events
- Frequentist paradigm of probability
- The axioms of probability
- Calculate the probabilities of events occurring in simple sample spaces with equally likely outcomes.

BASIC DEFINITIONS

The theory of probability was developed with the aim to explain a game(gambling) governed by the law of chance.

From: Gonick & Smith The cartoon guide to Statistics

What is a probability?

Two possible definitions

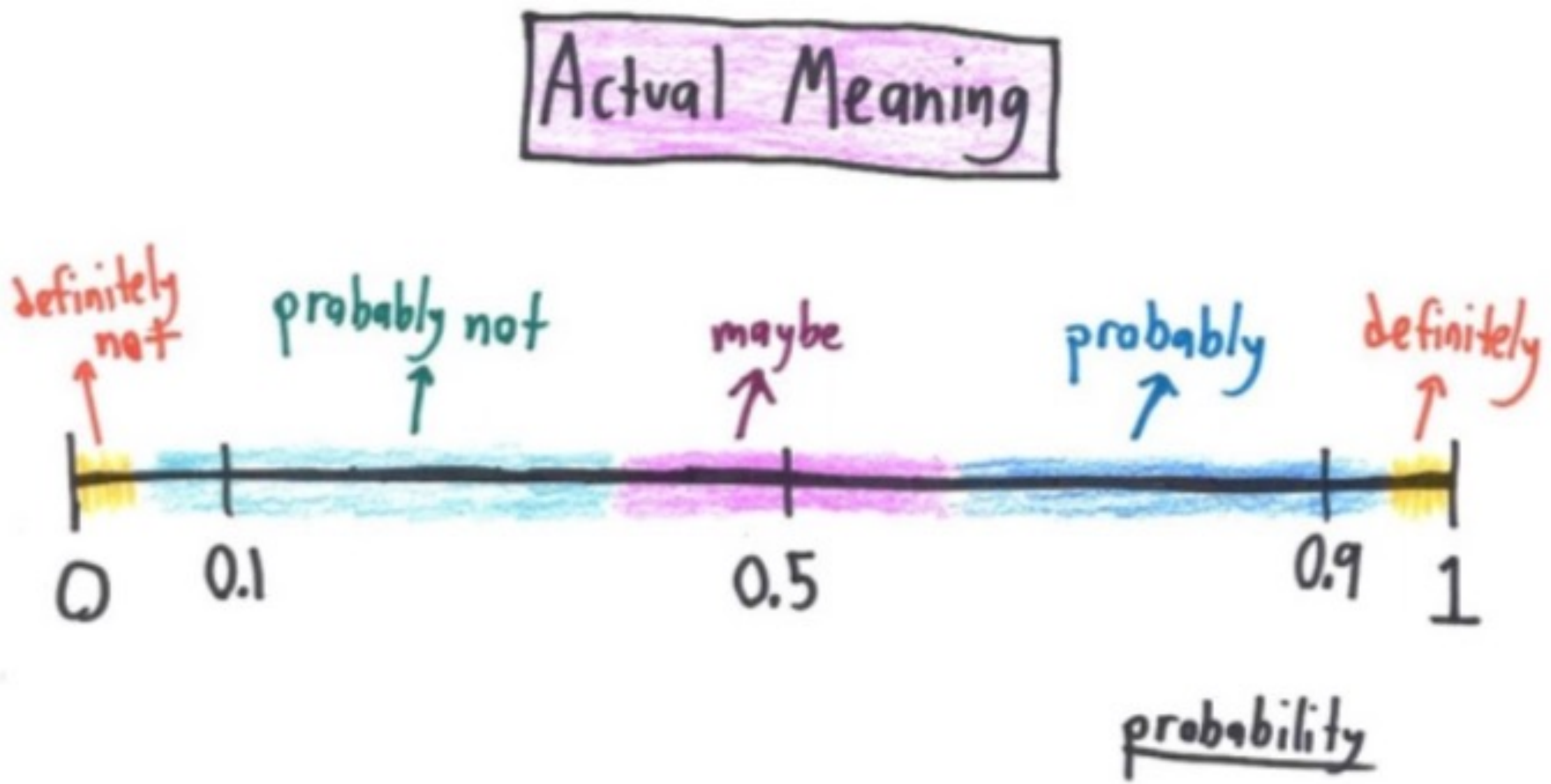
Objective (or frequentist): *the probability is obtained by repeating random experiments and counting the number of times something occurs:*

$$P(A) = p = \lim_{N \to \infty} \frac{N_A}{N}$$

Subjective (Bayesian): A probability can describe many things, but it is not possible to repeat the process many times and sample how often the event occurs. In this case, the probability is interpreted as a reasonable expectation, representing a state of knowledge or quantifying a personal belief.

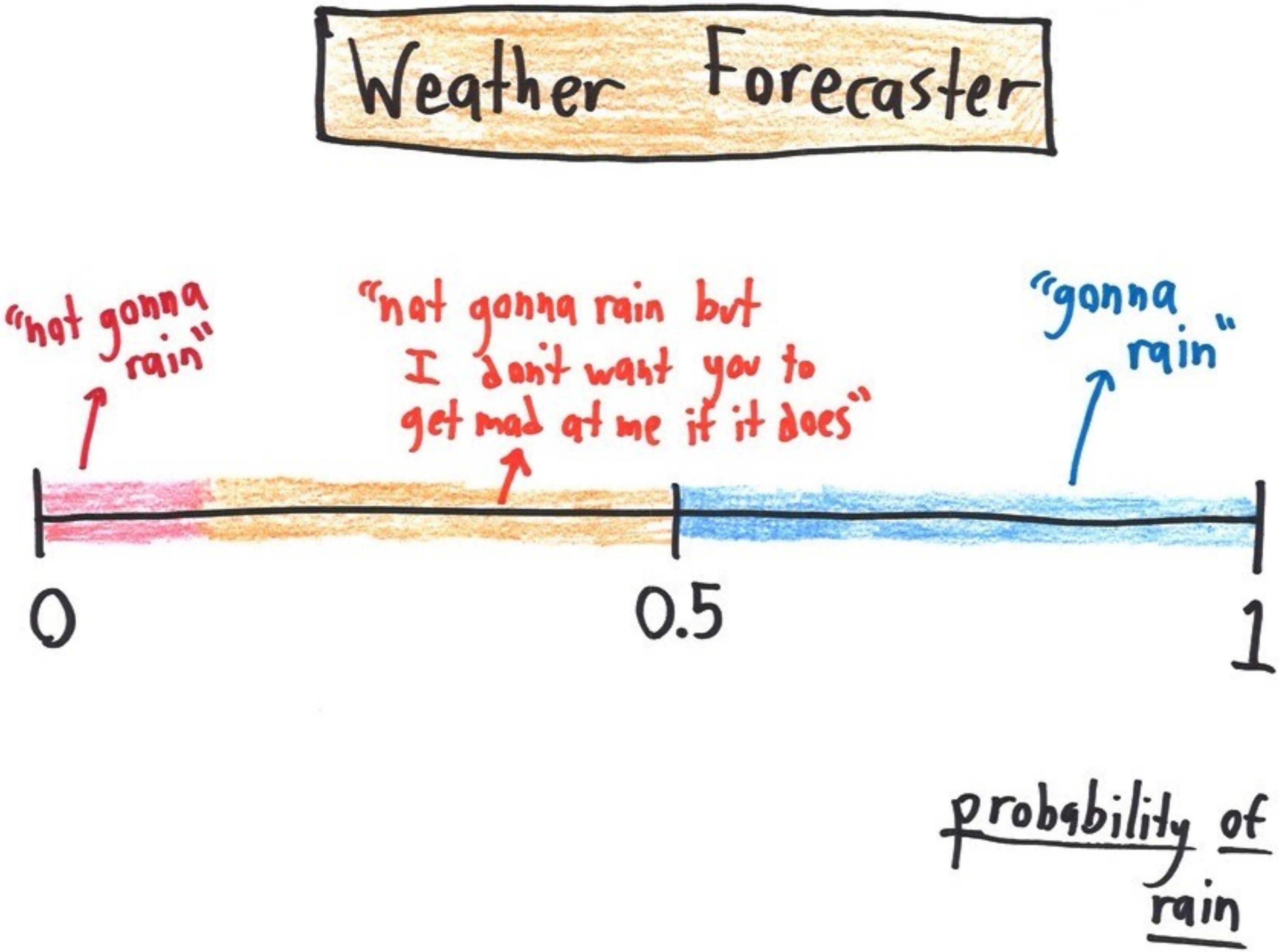
Bayesian methods are widely accepted and used in artificial intelligence in the field of machine learning, computer vision.

What is a probability?



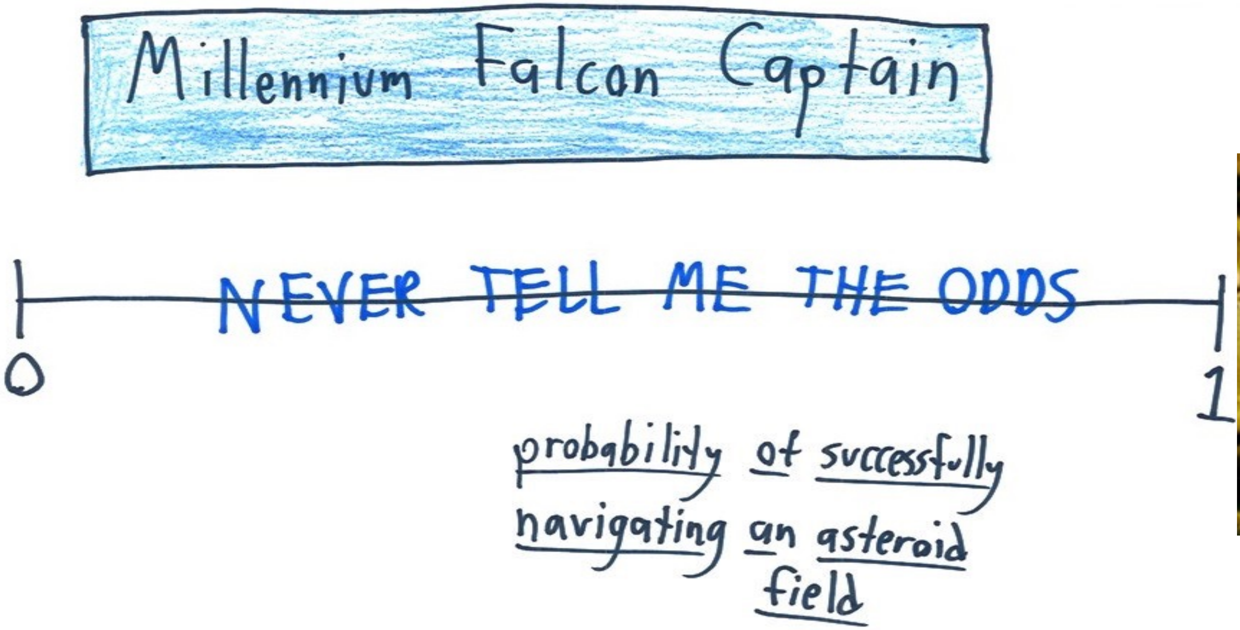
https://mathwithbaddrawings.com/2015/09/23/what-does-probability-mean-in-your-profession/

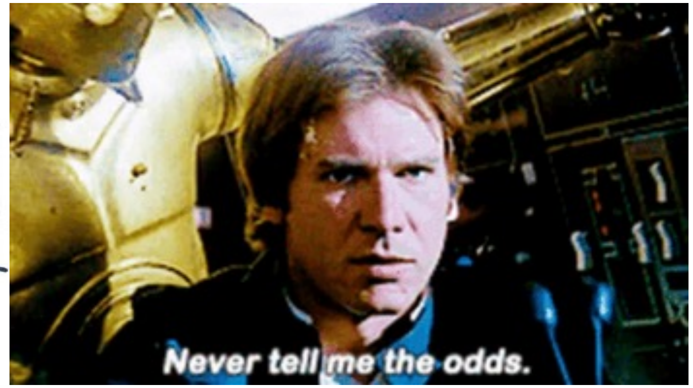
What is a probability?



https://mathwithbaddrawings.com/2015/09/23/what-does-probability-mean-in-your-profession/

What is a probability?





https://giphy.com/gifs/millennium-9mKMpXfvrwSl2

BASIC DEFINITIONS

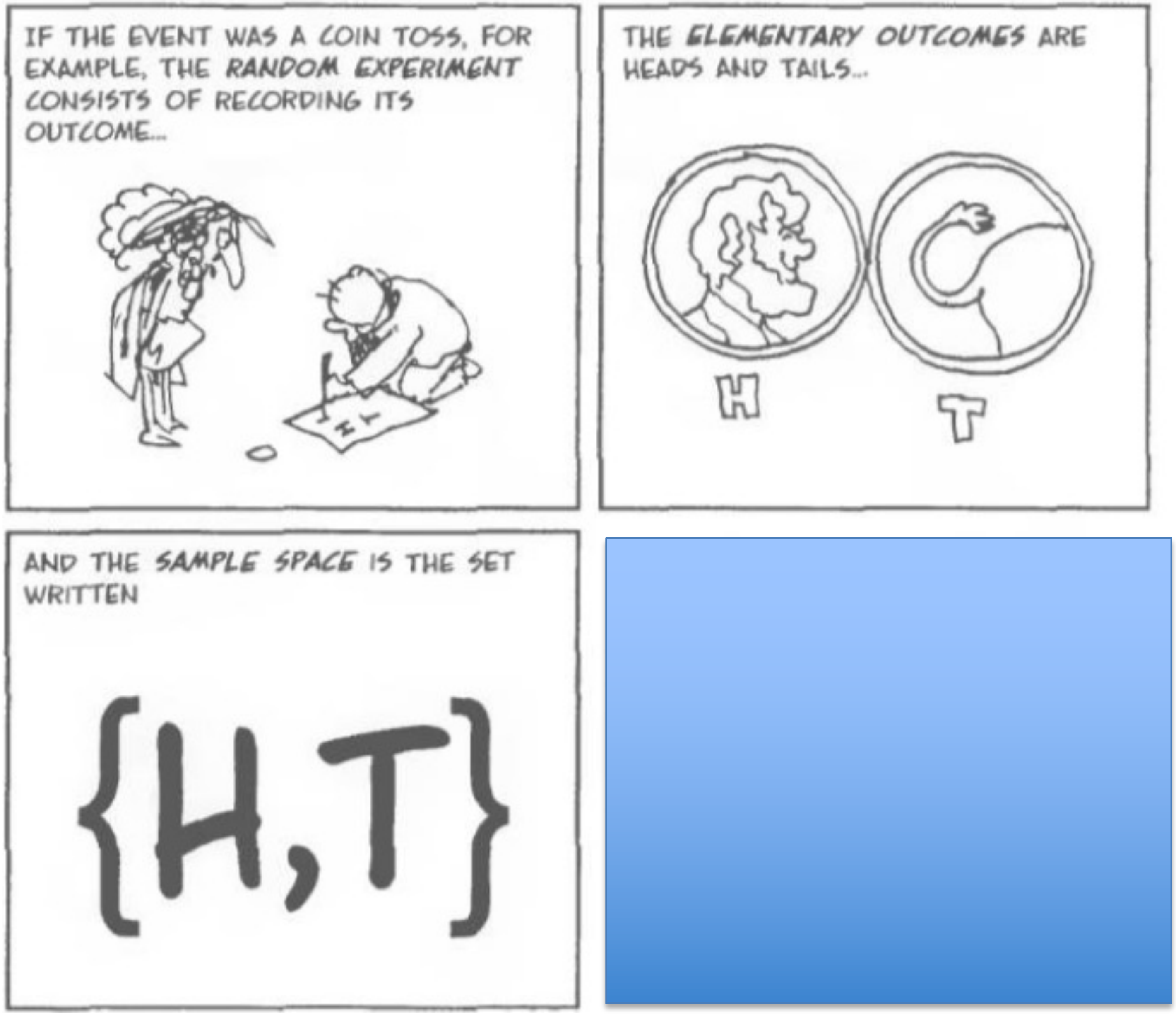
As a scientist, we call the game of the gambler a **random experiment**: in general, is a process of observing the outcome of a chance event (ex., the rolling of a dice, the radioactive decay, the fluctuation of the stock market).

The **elementary outcomes** are all possible results of the random experiment.

The **sample space** is the set or collection of all the elementary outcome.

SAMPLE SPACES

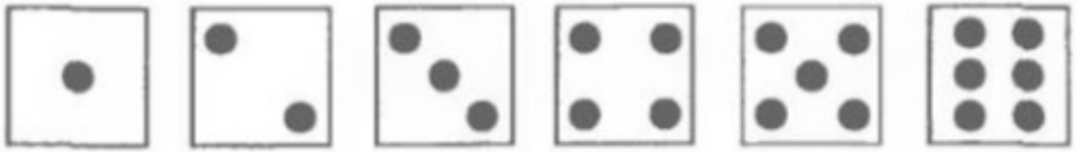
Example *What is the sample space of tossing a coin?*



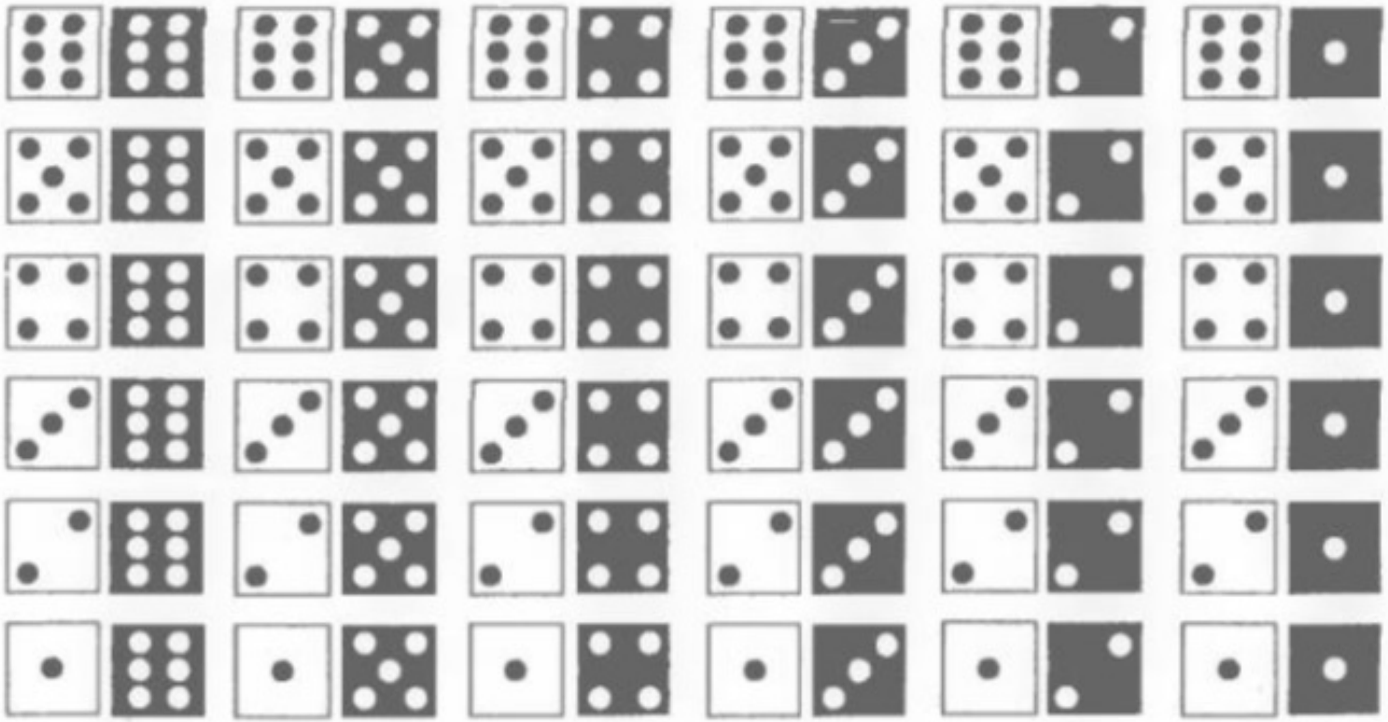
From: Gonick & Smith The cartoon guide to Statistics

BASIC DEFINITIONS

THE SAMPLE SPACE OF THE THROW OF A *SINGLE DIE* IS A LITTLE BIGGER.



AND FOR A *PAIR* OF DICE, THE SAMPLE SPACE LOOKS LIKE THIS (WE MAKE ONE DIE WHITE AND ONE BLACK TO TELL THEM APART):



SAMPLE SPACES

Note that sample spaces are not unique, sometimes there will be more than one way to break down and list the possible outcomes.

Example: An 'experiment' consists of flipping two coins? Give two possible sample spaces.

Define a single coin flip

$$F = \{H, T\}.$$

Then we have

$$S = \{F \times F\} = \{(H, H), (H, T), (T, H), (T, T)\}$$

SAMPLE SPACES

Alternatively, we could have three outcomes - 0 heads come up between the two flips, 1 head or 2 heads

$$S = \{0, 1, 2\}$$

Now our element {0} is identical to {T , T } previously.
But

$$\{1\} = \{(H, T), (T, H)\}$$

$$\{2\} = \{(H, H)\}$$

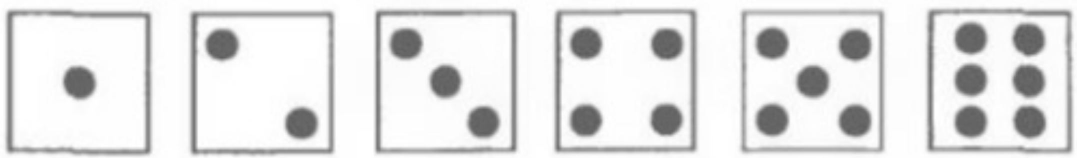
OUTCOMES

Definition: *Outcomes are the individual elementary elements of the sample space.*
(If the sample space is finite, or countably in finite, it will correspond to a discrete random variable.)

Example

A die is rolled once. We let X be the outcome of the experiment. The sample space for the experiment is the 6-element set

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



and each outcome i , for *i= 1…6*, corresponds to the number of dots on the face that faces up.

EVENTS

Definition: an event is any subset, E , of the sample space, S. $E \subset S$.

(Recall that a set E is a subset of set S , if all of the members of E are also contained in S)

Example what is the event, E , corresponding to a die coming up with an even number?

E = {2, 4, 6}

EVENTS

Example: *Consider a road race with the runners labelled 1 to 4. What is the event A that runner 1 won the race?*

Set builder notation

$$S = \{\text{all orderings of } (1,2,3,4)\} = \{(x_1, x_2, x_3, x_4 : x_i = 1, 2, 3, 4, x_i \neq x_j)\}$$

$$A = 1234, 1243, 1324, 1342, 1423, 1432 = \text{all ordering of } (1234) \text{ beginning with } 1$$
$$= \{(x_1, x_2, x_3, x_4 : x_1 = 1, x_i = 2, 3, 4, x_i \neq x_j)\}$$

note that $A \subset S$

What is the event that runner 2 lost the race?

$$B = 1342, 1432, 3142, 3412, 4132, 4312 = \text{all ordering of } (1234) \text{ ending with } 2$$
$$= \{(x_1, x_2, x_3, x_4 : x_4 = 2, x_i = 1\ \ 3, 4, x_i \neq x_j)\}$$

note again that $B \subset S$

SET THEORY AND PROBABILITY

Our events are be considered as sets and therefore we can apply results from set theory.

"OR" or Unions ∪

The event E_A or E_B has occured is denoted $E_A \cup E_B$, called the union of E_A and E_B, and is composed of the outcomes that are in either E_A or E_B.

"AND" or Intersections ∩

The event that E_A and E_B has occured is denoted $E_A \cap E_B$, called the intersection of E_A and E_B, and is composed of the outcomes that are in both E_A and E_B.

SET THEORY AND PROBABILITY

Example - Four runners are racing.
What is the event, , that runner 1 was first and runner 2 was last?

Define E_1 the event that the first runner won the race $E_1 = \{1234, 1243, 1324, 1342, 1423, 1432\}$
and E_2 the event that the second runner lost the race $E_2 = \{1342, 1432, 3142, 3412, 4132, 4312\}$

A is given by the intersection of E_A and E_B, so outcomes that are in E_1 *and* E_2

$$\begin{aligned} E_1 \cap E_2 &= \{1234, 1243, 1324, 1342, 1423, 1432\} \cap \{1342, 1432, 3142, 3412, 4132, 4312\} \\ &= \{1342, 1432\} \end{aligned}$$

PROBABILITY THEORY: A BIT OF HISTORICAL BACKGROUND

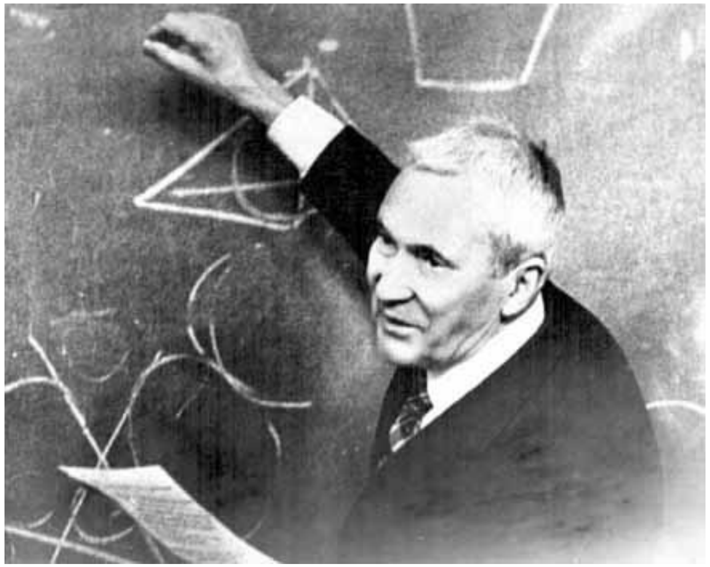
Christiaan Huygens
(1629-1695)

Rev. Thomas Bayes
(1701-1761)

Pierre-Simon Laplace
(1749-1827)

Andrey Markov
(1856-1922)

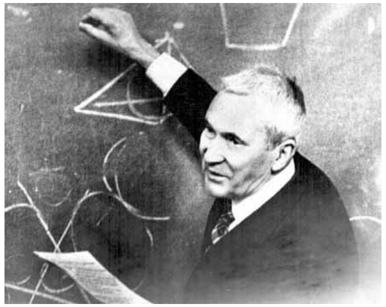
Modern probability theory is based on an axiomatic description of the properties of probability formulated by the Russian mathematician A. N. Kolmogorov (1903-1987).



http://enacademic.com/dic.nsf/enwiki/54484

AXIOMS OF PROBABILITY

We suppose that for each event of an experiment with sample space there is a number , which is in accord with the following three axioms:



Axiom 1 - The probability that an event will come from the sample space is unity

$$P(S) = 1$$

Axiom 2

$$P(A) \leq 1 \text{ for all } A \subset S$$

Axiom 3

$$P(A \cup B) = P(A) + P(B) \text{ if } A \cap B = \emptyset$$

We call a function P(A) the probability of event if these axioms hold.

AXIOMS OF PROBABILITY

Note that we are using '*intuition*', ' *physical insight*', or some other magic to decide the probabilities of the individual outcomes!

There is no general way to do it for a given problem.

We use every piece of information available to us and then compare it to reality if we aim to represent the real world with the probability distribution.

SAMPLE SPACES WITH EQUALLY LIKELY OUTCOMES

Often, we can reasonably assume that each outcome of an experiment is equally likely - **rolling a fair die**, for instance.

Or mathematically if the sample space S is a finite set
S = {1, 2, . . . , N }

it is often natural to assume that

$$P(\{1\}) = P(\{2\}) = \cdots = P(\{N\}) = p$$

SAMPLE SPACES WITH EQUALLY LIKELY OUTCOMES

Now it follows from axioms 2 and 3 that

$$1 = P(S) = P(\{1\}) + P(\{2\}) + \cdots + P(\{N\}) = Np$$

It then follows from axiom 3 that for any event $E \subset S$ its probability will be given by the number of elementary outcomes in E

$$P(E) = \frac{Number\ of\ outcomes\ in\ E}{N}$$

This means that to determine probabilities we need to count the number of ways in which a given event can occur.

SAMPLE SPACES WITH EQUALLY LIKELY OUTCOMES

In the more general case, we need to assign specific probabilities to the elementary outcomes, $P(\omega)$, so that

$$1 = P(S) = P(\{1\}) + P(\{2\}) + \cdots + P(\{N\})$$

in order to satisfy axiom 2.

It then follows from axiom 3 that for any event $E \subset S$ its probability will be given by

$$P(E) = \sum_{\omega \in E} P(\omega)$$

This means that to determine probabilities we need to count the number of ways in which a given event can occur.

SAMPLE SPACES WITH EQUALLY LIKELY OUTCOMES

Example. An experiment consists of throwing a coin twice. As we saw we can assign the sample space in two ways:

either

$$S_1 = \{HH, HT, TH, TT\}$$

with

$$P(HH) = \frac{1}{4}, P(HT) = \frac{1}{4}, P(TH) = \frac{1}{4}, P(TT) = \frac{1}{4}$$

or

$$S_2 = \{0, 1, 2\}.$$

In the later case, we take the number of times a head appears as our outcomes.

SAMPLE SPACES WITH NOT EQUALLY LIKELY OUTCOMES

In the case of the probabilities of the individual outcomes are ***not* equal** –

there are twice as many ways of getting the outcome one head ({H,T},{T,H})

as there are of getting zero ({T,T}

or two heads ({H,H}).

So, it makes sense to have

$$P(0) = \frac{1}{4}, P(1) = \frac{1}{2}, P(2) = \frac{1}{4}$$

OTHER EXAMPLES

A six sides die is rolled. In this case, the sample space is given by S={1,2,3,4,5,6}

For a fair die, each outcome can be argued to be equally likely. We can deduce the probability of the outcomes as

$$1 = P(S) = P(\{1\}) + P(\{2\}) + \cdots + P(\{6\}) = 6p$$

which implies that $p = \frac{1}{6}$, as you'd expect.

Define two events,

- A, "the roll is less than 5", so $A = \{1, 2, 3, 4\}$
- B, "the roll is more than 2", so $B = \{3, 4, 5, 6\}$

The number of elements in *A* is *n(A) =n(B)=4*

So $P(A) = n(A)p = 4\frac{1}{6}$ *and*

$$P(B) = n(B)p = 4\frac{1}{6}$$

ANOTHER USEFUL EXAMPLE

What is the probability of $A \cup B$?

In words what is the probability that either

"the dice roll is less than 5

OR

“the die roll is more than 2”

ANOTHER USEFUL EXAMPLE

The answer is not $n(A) + n(B)$ because the two sets have an intersection

$$A \cap B = \{1,2,3,4\} \cap \{3,4,5,6\} = \{3,4\}$$

One of the basic rules of sets is we don't need to keep duplicate members - so

$$A \cup B = \{1,2,3,4\} \cup \{3,4,5,6\} = \{1,2,3,4,3,5,6\} = \{1,2,3,4,5,6\} = S$$

We see in this case that

$$n(A \cup B) = n(A) + n(B) - n(A \cap B)$$

which for probabilities proportional to number of elements also gives

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

SUMMARY

- The probability of any event can be decided once we have the *probability function*, $P(\omega)$, for our experiment.

- The probability function is the function $P(\omega)$, where the ω are the elementary outcomes of the experiment.
- It has domain S.
- It obeys the axioms of probability.
- The actual probabilities are assigned based on some model; they are outside the idealised mathematical model.
- We will demonstrate that the interpretation of $P(E)$ as the relative frequency of the event E when the experiment is repeated many times is consistent with the axioms.
- Belief based interpretations can also satisfy these requirements.

NOTE

Notations are not totally standard in probability theory. I'll use what I think is the closest to standard and should allow you to read most works on probability.

For example, in some textbook, you can find that the sample space is called Ω, the individual outcomes ω_i and the probability function $m(\omega)$.

OTHER USEFUL RELATIONSHIPS AND DEFINITIONS

Definition 1: Complement of an event E (with respect to the sample set S) is

$$P(\bar{E}) = 1 - P(E)$$

Definition 2: E_1 and E_2 are mutually exclusive if

$$P(E_1 \cap E_2) = P\{0\} = 0$$

Definition 3: A set of events $E_1, E_2, \ldots E_n$ of some experiment are said to be *exhaustive* if

$$E_1 \cup E_2 \cup E_3 \cup \cdots E_n = S$$

Definition 4: If E_1 and E_2 are events of the same experiment

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

PRINCIPLES OF COUNTING - COMBINATORIALS AND PERMUTATIONS

Suppose two experiments are carried out

- experiment 1 can result in any of m outcomes
- for each outcome of experiment 1 experiment 2 can have *n* outcomes
- together there are *mn* outcomes.

We can write all possible outcomes as an ordered pair, *(m, n)* .
The *mn* possible outcomes can be tabulated

(1,1)	(1,2)	$\cdots$	(1,n)
(2,1)	(2,2)	$\cdots$	(2,n)
$\vdots$	$\vdots$	$\vdots$	$\vdots$
(m,1)	(m,2)	$\cdots$	(m,n)

Example: 2 balls are randomly drawn from a bowl that contains 6 black and 5 white balls.

- what is the chance that we draw one black and one white ball?

We want to apply

$$P(E) = \frac{\text{Number of outcomes in } E}{N}$$

First calculate the total number of possible outcomes …

First calculate N the total number of possible outcomes:

- ***experiment 1***, the first ball can be picked in *11* different ways
  $m = 11$.

- ***experiment 2***, for each of the results of experiment 1, we can pick the next ball in one of 10 ways
  $n = 10$.

- so total number of outcomes possible is ***mn = 110***

Shall we write it in set builder notation

$$B = \{(x_1, x_2): x_1 = 1,2, \dots 11, x_2 = 1,2, \dots 10, x_2 \neq x_1\}$$

and so if we were to list all the ordered pairs in a table they would be 11*10

$$N = n(B) = |B| = 11 * 10 = 110$$

PRINCIPLES OF COUNTING - COMBINATORIALS AND PERMUTATIONS

Now we have the denominator we need to calculate the number of ways of getting one black and one white ball. Either

- the first ball is black, which can occur in 6 ways (*m*), for each way there are 5 ways the second ball can be white (*n*)

- the first ball is white (5 ways, *m*) and the second ball is black (6 ways, *n*)

Either of the two ways (or, union of the two events) leads to an event with one black and one white ball.

Therefore, by the axiom 3, the total number of ways to get one black and one white ball is

$$5*6+6*5=60.$$

Shall we write it in set builder notation.

We number the balls so that 1 to 6 are white and 7 to 11 are black.

The first ball is x_1 and second x_2

$$T = \{(x_1, x_2): x_1 = 1,2,3,4,5,6, x_2 = 7,8,9,10,11\}$$

$$U = \{(x_1, x_2): x_1 = 7,8,9,10,11, x_2 = 1,2,3,4,5,6\}$$

Therefore

$$P(E) = \frac{\text{Number of outcomes in } E}{N} = \frac{5*6+6*5}{110} = \frac{6}{11}$$

FACTORIALS

If there are more than two experiments to be performed, then the above rule generalises to $n_1 \cdot n_2 \cdot \cdots \cdot n_r$ *possible outcomes of r experiments.*

So continuing our previous example, how many ways are there to select all the balls from the bowl?

In this case there are $11 \cdot 10 \cdot 9 \cdot \ldots \cdot 2 \cdot 1 = 11!$

Lets calculate this just to see how big these numbers get!

i	i!
1	1
2	2
3	6
4	24
5	120
6	720
7	5040
8	40320
9	362880
10	3628800
11	39916800

FACTORIALS : STIRLING APPROXIMATION

A useful approximation to the factorial is the Stirling approximation

$$\ln n! = n ln\, n - n + O(\ln n)$$

The big *O* notation indicate the error (difference with the correct value) that can be quantify with a constant $\frac{1}{2}\ln(2\pi n)$, yielding to the more precise formula

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

N	N!	Stirling approximation	Error %
1	1	1.00	0.227445%
2	2	2.00	0.032602%
3	6	6.00	0.009986%
4	24	24.00	0.004266%
5	120	120.00	0.002198%
6	720	720.01	0.001276%
7	5040	5040.04	0.000805%
8	40320	40320.22	0.000540%
9	362880	362881.38	0.000380%
10	3628800	3628810.05	0.000277%
11	39916800	39916883.11	0.000208%
12	479001600	479002368.48	0.000160%
13	6227020800	6227028659.89	0.000126%
14	87178291200	87178379323.32	0.000101%

PERMUTATIONS

Simple Definition: A full set of permutations is all the ways of arranging some distinguishable objects. Each permutation swaps between two of these ways of arranging them.

Rigorous Definition: A permutation is a one-to-one mapping of a set onto itself whilst changing the ordering of the elements.

For instance if A = {a, b, c} a possible permutation would be (using Cauchy's two-line notation)

$$\sigma = \begin{pmatrix} a & b & c \\ c & b & a \end{pmatrix}$$

where the permutation sends a to c , b to c, and c to a .

PERMUTATIONS

The number of permutations of n objects can be found from the counting rules to be **n!** , in much the same way as for the ball selection example above.

EXAMPLE

Find all the permutations of A = {a, b, c}

$$B = \{(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1)\}$$

and

$$n(B) = 3!$$

COMBINATIONS

Definition: We define the number of combinations of *r* objects taken from a set of size *n* , $\binom{n}{r}$, for $r \leq n$, by

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

$\binom{n}{r}$ is the number of combinations of *n* objects taken *r* at a time, also referred to as **the binomial coefficient.**

The binomial theorem states that

$$(x+y)^n = \sum_{r=0}^{n} \binom{n}{r} x^r y^{n-r}$$

We'll encounter it again when we will talk about the binomial and Poisson distributions.

COMBINATIONS

EXAMPLE

We need to select student representatives for the school - a committee of 5 people is selected from 8 physicists and 30 mathematicians

a) If we decide we need to have 2 physicists and 3 mathematicians on the committee, how many possible sets of representatives are there?

we have two 'experiments' here

- First, we choose 2 physicists out of 8 - i.e., $\binom{8}{2}$ = 28
- secondly for each choice of the physicists we can select 3 mathematicians out of 30 - i.e. . $\binom{30}{3}$=4060

So, in total 28*4060 = 113680 possible committees of this composition...!

COMBINATIONS

b) what is the probability of getting a committee with 4 mathematicians and 1 physicist?

In this case, we need to consider all the students (30+8=38)

The total number of ways of selecting the committee is $\binom{38}{5}$ = 501942

the number of ways of getting 4 mathematicians and 1 physicist is basically the same as part a) above

$$\binom{30}{4} \cdot \binom{8}{1}$$

So, the probability is

$$P = \frac{\binom{30}{4} \cdot \binom{8}{1}}{\binom{38}{5}} = 0.44$$