Basics of Probability

Lecture 7-8

Introduction

Probability statements are everywhere around us. Examples of probability statements include

1. There is a 60 percent chance of its raining today.
2. The chance of me winning the lottery is one in 80 million.
3. There is a 50-50 chance of observing a head when a fair coin is tossed.

Just what is meant by chance in these statements? Chance is a measure of uncertainty, and we call this measure probability. In this lecture we will study this concept of probability.

Probability: a step by step definition (I)

Randomness suggests unpredictability. A simple example of *randomness*: when a coin is tossed, the outcome is uncertain.

The outcome could be either an observed head (H) or an observed tail (T). Because the outcome of the toss cannot be predicted for sure, we say that it displays randomness.

Probability: a step by step definition (II)

A random experiment is an experiment in which the outcome on each trial is uncertain and distinct.

Examples of random experiments are rolling a die, selecting items at random from a manufacturing process to examine for defects, selection of numbers by a lottery machine, etc.

Probability: a step by step definition (III)

When we toss a coin, we have two possible outcomes, summarized by $\{H, T\}$.

Such a list is called a **Sample Space.** Generally the sample space is indicated with the Greek letter $\mathbf{\Lambda}$

Example : A fair regular six-sided die is rolled. List the sample space for this random experiment. Solution: Let $\mathbf{\Lambda}$ represent the sample space. Then $\mathbf{\Lambda} = \{1, 2, 3, 4, 5, 6\}$.

Probability: a step by step definition (IV)

We may only be interested in part of the sample space. For example we may be concerned with even numbers $A = \{2, 4, 6\}$. This is a *subset* of the sample space. Such subsets are called **events**

If the outcomes in a sample space are equally likely to occur, then the **classical probability of an event A** is defined to be

$$P(A) = \frac{\text{Number of simple events in A}}{\text{Total number of simple events in the sample space}}$$

Probability: Some simple examples

When a child is born, the gender of the child is either a boy (B) or a girl (G), summarized by $\mathbf{\Lambda} = \{B, G\}$. If we consider a two-child family, the possibilities can be summarized by $\mathbf{\Lambda} = \{BB, BG, GB, GG\}$.

If a two-child family is selected at random, what is the probability of there being two boys in the family?

the event of two boys occurs once, and there are four simple events in the sample space, thus

$$P(BB) = \frac{1}{4} = 0.25$$

Set theory

Basic notions

Mathematically speaking events are sets.

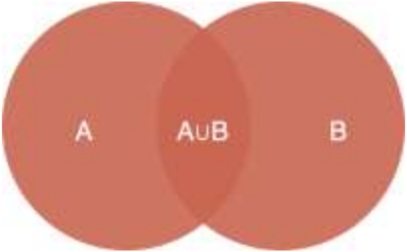
Definition: A set is a finite or infinite collection of objects

Sets can be **finite** (contains a finite number of objects) or **infinite** (contains an infinite number of objects)

The **cardinality** of a given set is the number of objects that belong to such set. I.e. if $E = \{1, 2, 3\}$, the cardinality of E, $\# E = 3$,

Recap of Set theory

- **Union** $(A \cup B)$ given two events A, B, everything that is in *either* A, B *or in both*.

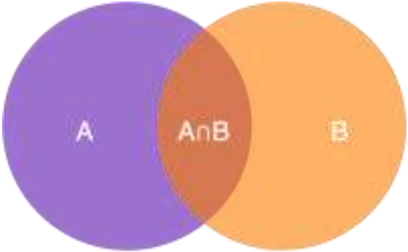


Example A = "the die returns an even number", B = "the die returns a 5"

$\Rightarrow A \cup B = \{2, 4, 5, 6\}$

Recap of Set theory

- **Intersection** $(A \cap B)$ given two events A, B, everything that is in *both* A and B.



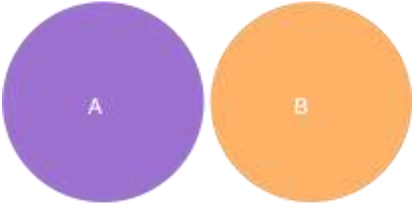
Example A = "the die returns an even number", B = "the die returns a number smaller than 5"

$\Rightarrow A \cap B = \{2, 4\}$

Recap of Set theory

Empty intersection

- **Intersection** $(A \cap B)$ given two events A, B, everything that is in *both* A and B.

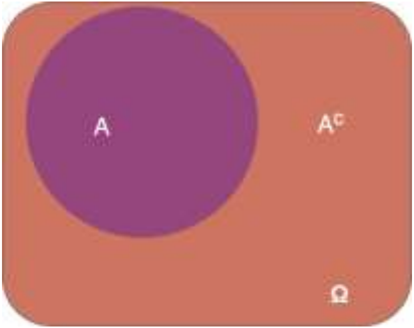


Example A = "the die returns an even number", B = "the die returns a 5"

$\Rightarrow A \cap B = \emptyset$ A and B are **disjoint**

Recap of Set theory

- **Complement** (A^c or $\bar{A}$) everything that is not in A.



Example A = "the die returns an even number", A^c = "the die returns an odd number"

Probability

A formal definition of probability is hard to be given. Practically speaking, the **probability** is a **set function** on which the following holds.

Probability Axioms...

- $0 \leq P(A) \leq 1$
- $P(\mathbf{\Lambda}) = 1$
- $P(\emptyset) = 0$

and some trivial consequences

- $P(A^c) = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

if A, B are disjoint then $P(A \cup B) = P(A) + P(B)$

The basic ingredients - an evergreen Example

Random phenomenon: throw of a die

- **Sample Space**: all the possible outcomes
  - $\boldsymbol{\Lambda} = \{1, 2, 3, 4, 5, 6\}$
- **Event**: “the die returns an even number”
  - $E = \{2, 4, 6\}$
- **Probability**:
  - $P(E) = \frac{}{\# \boldsymbol{\Lambda}} = \frac{1}{2}$

Set Theory is important

For computing probabilities

Exercise "students" In a sample of 100 college students, 60 said that they own a car, 30 said that they own a stereo, and 10 said that they own both a car and a stereo.

1. Compute probabilities for these events, and depict this information on a Venn diagram.

Let C be the event that a student owns a car, and let D be the event that a student owns a stereo. Thus $P(C) = 0.6$, $P(D) = 0.3$, and $P(C \cap D) = 0.1$

2. Compute the probability that a student has car but hasn't a stereo

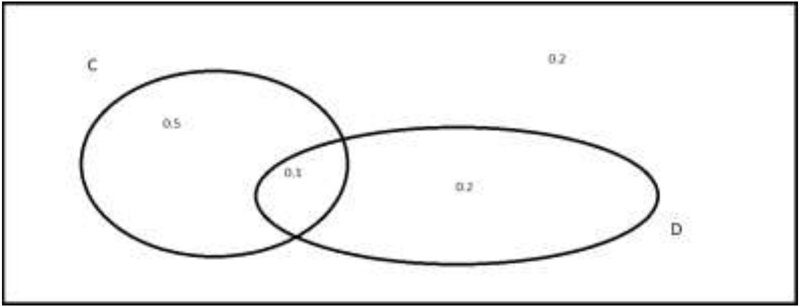
$$P(\text{"Only a Car"}) = P(C) - P(C \cap D) = 0.6 - 0.1 = 0.5$$

Exercise "students"

Venn Diagram

3. Compute The probability that a student has either a car or a Stereo

$$P(\text{"A Car OR A Stereo"}) = P(C \cup D) = P(C) + P(D) - P(C \cap D) =$$
$$= 0.6 + 0.3 - 0.1$$



Exercises

- Two coins are tossed. Note: Each coin has two possible outcomes H (Heads) and T (Tails).

1. Get the sample space.
2. Find the probability that two heads are obtained.

.

Exercises

- Which of the following is an impossible event?

1. Choosing an odd number from 1 to 10.
2. Getting an even number after rolling a single 6-sided die.
3. Choosing a white marble from a jar of 25 green marbles.
4. None of the above.

- There are 4 parents, 3 students and 6 teachers in a room. If a person is selected at random, what is the probability that it is a teacher or a student?

1. 4/13
2. 7/13
3. 9/13
4. None of the above.

Random variables

From the sample space to random variables

So far we introduced the sample space Λ and we computed probabilities associated to subset of the sample space.

P = "Primary School", S="Secondary School" M="Master"

Ω

For example we learnt how to compute $P("P")$ or $P("M")$ and so on…

Random variables

Coneptually speaking

Now our probabilistic experiment is to choose a student at random and record the number of years schooling (Y)

1. $P \rightarrow 5$
2. $S \rightarrow 10$
3. $M \rightarrow 15$

Random variables

Black Box

You can think Y as an abstract box that takes as input a student and produces a number y which is the number of years schooling of that particular student

$$\text{Student} \rightarrow \boxed{Y} \rightarrow y \text{ (years of schooling)}$$

just like a standard mathematical function

$$x \rightarrow \boxed{f} \rightarrow f(x) \text{ (years of schooling)}$$

Notation

Y : Random Variable y : Numerical Value

Random variable

Going back to the years of schooling

The random variable (r.v.) Y **randmoly picks up** the student and gives back in output his/her years of schooling. Thus there is a probability law associated to its values y (this i generally called **probability distribution**)

$$P(x) = \begin{cases} 6 & \text{if} \quad x = 5 \\ 2/10 & \text{if} \quad x = 10 \\ 2/10 & \text{if} \quad x = 15 \end{cases}$$

Random variable

Little bit more formally

Typically we are not interested in the single outcome itself or in the events but in a *function* of them.

A **random variable** E is any function from the sample space to the real numbers.

In formulas:

$$E : \mathbf{\Lambda} \to \mathrm{A}$$

N:B: As the random variable is defined on **the sample space** we can associate a probability to the values that the random variable assumes

Random Variables

Some additional examples

- toss a coin three times and **count** the number of tails
- roll two dice and **sum** the values of the faces

NB A random variable is a *number*: we can do all sorts of operations with it!

Random variables: discrete and continous

- A **discrete** random variable is a variable that represents numbers found by counting. For example: number of marbles in a jar, number of students present or number of heads when tossing two coins.
- When we have to use intervals for our random variable or all values in an interval are possible, we call it a continuous random variable. Thus, continuous random variables are random variables that are found from measuring - like the height of a group of people or distance traveled while grocery shopping or student test scores. In this case, E is continuous because E represents an infinite number of values on the number line.

Distribution of a random variable

an example of how to derive it

- Toss a coin three times. X is the random variable representing the *number of Tails*

m	$P(m)$	x
HH H	1/8	0
THH	1/8	1
HTH	1/8	1
HHT	1/8	1
TTH	1/8	2
THT	1/8	2
HTT	1/8	2
TTT	1/8	3

x	$p_x = P(E = x)$
0	$1/8 \times 1 = 1/8$
1	$1/8 \times 3 = 3/8$
2	$1/8 \times 3 = 3/8$
3	$1/8 \times 1 = 1/8$

The distribution of a random variable p_x is just a convenient way of summarizing single outcomes probabilities.

Constructing the probability distribution

- Λ represents the sets of all the possible outcomes after tossing two coins
- E Represents the number of tails after tossing 2 coins
- $x = \{0, 1, 2, 3\}$

x	$p_x = P(E = x)$
0	$1/8\times1 = 1/8$
1	$1/8\times3 = 3/8$
2	$1/8\times3 = 3/8$
3	$1/8\times1 = 1/8$

$$P(x) = \begin{cases} 1 & \text{if } x = 0 \\ 3/8 & \text{if } x = 1 \\ 3/8 & \text{if } x = 2 \\ 1/8 & \text{if } x = 3 \end{cases}$$

Home Task

- Two dice are rolled:

1. Construct the sample space. How many outcomes are there?
2. Find the probability of rolling a sum of 7.
3. Find the probability of getting a total of at least 10.
4. Find the probability of getting a odd number as the sum.