Introduction to Probability

Jamie and Amnon

September 30, 2016

Warning

WARNING: This presentation may contain equations known to the State of California to cause cancer and birth defects or other reproductive harm.

(but they are not really important)
We will start slow and formal, and then move faster and less formal.

What is Probability

- A nice mathematical framework
- Somehow related to the real world:
    - We can associate the probability function of an event with the limit of the fraction of times this event will happen when doing an infinite number of repeats.
    - When a phenomena shows results that are memoryless.

Mathematical Definitions - Sample Space

- The sample space Ω is the set of all possible outcomes we measure
  - For a coin flip, we can define:

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

  - For a balanced, 6-sided die:

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

  - For flipping 2 coins:

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

Mathematical Definitions - Event

- Events are defined as subsets of Ω:

$$E \in 2^{\Omega}$$

- Every outcome is an event
- An event can be a set of several outcomes
- Examples:
  In a die roll, an event can be even numbers

$$E = \{2, 4, 6\}$$

  In two coin flips, an event can be 2 heads:

$$E = \{HH\}$$

Mathematical Definitions - Distribution Function

Let Ω be a Sample Space
The distribution function is defined as:

$$m : \Omega \Rightarrow [0, 1]$$

Such that:

- Non negative:
$$\forall \omega \in \Omega, \quad m(\omega) \geq 0$$
- Sum is 1:
$$\sum_{\omega \in \Omega} m(\omega) = 1$$
- Example:
In a fair die roll, $m(1) = m(2) = ... = m(6) = \frac{1}{6}$

Mathematical Definitions - Probability Function

Let Ω be a Sample Space
The probability function is an extension of the distribution function.
for a given distribution function $m(\omega)$, we define a probability function:

$$P : 2^{\Omega} \Rightarrow [0, 1]$$

Such that:

$$\forall E \in 2^{\Omega}, \quad P(E) = \sum_{\omega \in E} m(\omega)$$

- Example:
  In a fair die roll:

$$P(\{1\}) = m(1) = \frac{1}{6}$$

$$P(\{2, 4, 6\}) = m(2) + m(4) + m(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

Example - two fair coin tosses

- The Sample Space is: $\Omega = \{HH, HT, TH, TT\}$
- Possible Events are:
  $E = \{HH\}$ (two heads)
  $F = \{HH, TT\}$ (two identical results)
- The Distribution Function is:
  $m(HH) = m(HT) = m(TH) = m(TT) = \frac{1}{4}$
- The Probability Function is defined for events:
  For E (2 heads), $P(E) = P(\{HH\}) = \frac{1}{4}$
  For F (two identical results), $P(F) = P(\{HH, TT\} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

Nice properties of probability functions

- Probability of the complete sample space is 1:

$$P(\Omega) = 1$$

- Probability of complement is one minus event probability:

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

- Probability of the union of disjoint events in the sum of their probabilities:

$$P(E \cup F) = P(E) + P(F) \quad \text{for } E \cap F = \emptyset$$

- HOWEVER, this does NOT always work:

$$P(E \cap F) = P(E) \cdot P(F)$$

Example - nice properties of probability functions

For two coin flips:

- The sample space $\Omega = \{HH, HT, TH, TT\}$
- Probability of the complete sample space is :

$$P(\Omega) = P(\{HH\}) + P(\{HT\}) + P(\{TH\}) + P(\{TT\}) = 1$$

- Probability of at least one tail:

Conditional Probability

- What are the probabilities of events given some event happened?
- Example: with a fair die, what is the probability of getting $\{6\}$ given that the results was even?
  (Meaning: if we throw the die a lot of times, out of the times that the number was even, what fraction of times was the number 6)

Mathematical Definition - Conditional Probability

- For an event E, the conditional probability given the event is written as:

$$P(F|E)$$

The probability that F happened given that E happened.

- It can be shown that:

$$P(F|E) = \frac{P(F \cap E)}{P(E)}$$

(IMPORTANT FORMULA)

Examples - Conditional Probability

For a 6-sided fair die:

- The probability we get a 6 given the number was even: $P(\{6\}|\{2,4,6\}) = \frac{P(\{6\}\cap\{2,4,6\})}{P(\{2,4,6\})} = \frac{P(\{6\})}{P(\{2,4,6\})} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$
- The probability we get a 6 given the number was odd: $P(\{6\}|\{1,3,5\}) = \frac{P(\{6\}\cap\{1,3,5\})}{P(\{1,3,5\})} = \frac{P(\emptyset}{P(\{1,3,5\})} = 0$

Examples - Conditional Probability

For 2 coin flips:

- The probability we get 2 similar results given the first was head:

$$P(\{HH, TT\}|\{HH, HT\}) = \frac{P(\{HH, TT\} \cap \{HH, HT\})}{P(\{HH, HT\})} =$$

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

Complicated Example - Conditional Probability

A rare and lethal disease with a prevalence of 0.1% of the population. Luckily, we have a test with 99% accuracy (so if you have the disease, 99% of the times the test will return positive, and if you are healthy, 99% of the times the test will return negative

Unfortunately, you tested positive for the disease. What is the probability you actually have it?

Complicated Example - Conditional Probability

Denote H for healthy, S for sick, P for test positive and N for test negative.
We have:

- $\Omega = \{HP, HN, SP, SN\}$
- $P(\{SP, SN\}) = 0.001$
- $P(\{SP\}|\{SP, SN\}) = 0.99$
- $P(\{HN\}|\{HP, HN\}) = 0.99$

What is $P(\{SP, SN\}|\{SP, HP\})$?

Complicated Example - Conditional Probability

$$P(\{HN\}|\{HP, HN\}) = 0.99$$

$$\Downarrow$$

$$P(\{HN\}) = 0.99 \cdot P(\{HP, HN\}) = 0.99 \cdot (1 - 0.001) \approx 0.99$$

$$\Downarrow$$

$$P(\{HP\}) \approx 1 - 0.99 = 0.01$$

and

$$P(\{SP\}|\{SP, SN\}) = 0.99$$

$$\Downarrow$$

$$P(\{SP\}) = 0.99 \cdot P(\{SP, SN\}) = 0.99 \cdot 0.001 \approx 0.001$$

therefore:

$$P(\{SP, SN\}|\{SP, HP\}) = \frac{(\{SP\}}{P(\{SP, HP\}} = \frac{0.001}{0.001 + 0.01} = 0.1$$

So we still have hope (only 10% of being sick)

Random Variables

- A random variable is "a quantity whose value depends in some clearly-defined way on a set of possible random events".
- Usually denoted by capital letters (i.e. X,Y)
- For every repeat/world manifestation, the random variable gets a value
- Can have several random variables. The value each random variable gets are from the same repeat.

Example - Random Variables

for throwing 2 dice

- W is defined as the sum of the 2 dice
- X is defined as the value of the first die
- Y is defined as the value of the second die
- Z is defined as 1 if the 2 dice are equal, 0 if they are different

For a given repeat, say we got the values 4,5 in our 2 dice. Then we have:

- $W = 9$
- $X = 4$
- $Y = 5$
- $Z = 0$

Mathematical Definition - Independence

Two random variables are said to be independent if the value we get for one do not change the probability function of the other:
Definition:
X,Y are independent if:

$$P(Y|X) = P(Y) \quad \forall x \in X$$

Example - Independence

from the previous example (for throwing 2 dice)

- W is defined as the sum of the 2 dice
- X is defined as the value of the first die
- Y is defined as the value of the second die
- Z is defined as 1 if the 2 dice are equal, 0 if they are different

We have

- W and X are dependent (i.e. if the sum is 12, both dice must be 6)
- W and Y are dependent (i.e. if the sum is 12, both dice must be 6)
- X and Y are independent (the value of one die will not affect the other)
- are W and Z dependent?

Example - Independence

The indepedence of two random variables also depends on the sample space they are in.

For example (WITH BACTERIA!): Choose two people (A and B) from the world and look at the level of Streptococcus in their saliva (denote S_A and S_B).

- If B is chosen independently of A, S_A and S_B are independent.
- If B is always chosen as the same individual as A, S_A and S_B are dependent
- However, if we're only looking at a time series of one individual (so we randomly choose A and B from samples of a time series of one individual), S_A and S_B are independent again.

Independence and Probability

An important property of independent random variables is:

$$P(X \text{ and } Y) = P(X) \cdot P(Y) \quad \text{for X,Y independent}$$

(The probability of $X = x$ and $Y = y$ is the product of the probabilities of $P(X = x)$ and $P(Y = y)$)
This can be easily derived since we know:

$$P(X|Y) = \frac{P(X \text{ and } Y)}{P(Y)}$$

and

$$P(X|Y) = P(X)$$

Quick note: we can define also an independence for a group of random variables, and in theory it's not enough that only all pairs in the group are independent (but usually it's close enough)

Mathematical Definition - Expectation

If you are a (smart) gambler in a casino, you would like to play games that will let you win money on the long run.
The average winning over a large number of repeats is:

$$\sum_{\text{result} \in \text{game}} P(\text{result}) \cdot \text{Gain}(\text{result})$$

Since we have the relation between probability and real life, we define:
For a numeric random variable X, the Expectation of X is:

$$E(X) = \sum_{x \in X} P(X) \cdot X$$

Expectation - Properties

- Additivity : $E(X+Y) = E(X) + E(Y)$
- Linearity : $E(c \cdot X) = c \cdot E(X)$ for any number c

Example:
If you play a game where we toss a fair die and then:

- if the result is even you get 1$ if it's odd you lose 2$
- if the result is '1', you gain 5$ otherwise you don't get anything

What is your avergae gain/loss over a lot of games?

Mathematical Definition - Variance and Standard Deviation

Another thing that may interest us is how much the results vary from the mean.
A useful (but not only) way to quantify it is using Variance:

$$V(X) = \sum_{x \in X} P(x) \cdot (x - E(X))^2 = E[(X - E(X)^2]$$

and we define the standard deviation as the square root of the variance:

$$std(X) = \sqrt{V(X)}$$

Example - Variance and Standard Deviation

- If we flip a coin, X is 2 if it is head, 0 if tail:
  - $E(X) = 0.5 \cdot 2 + 0.5 \cdot (0) = 1$
  - $V(X) = 0.5 \cdot (2-1)^2 + 0.5 \cdot (0-1)^2 = 1$
  - $std(X) = \sqrt{V(X)} = \sqrt{1} = 1$
  - and a lot of time we're interested in the relative level of fluctuations $\frac{std(X)}{E(X)} = \frac{1}{1} = 1$
- If we flip a coin again, but now X is 20 if it is head, 0 if tail:
  - $E(X) = 0.5 \cdot 20 + 0.5 \cdot (0) = 10$
  - $V(X) = 0.5 \cdot (20-10)^2 + 0.5 \cdot (0-10)^2 = 100$
  - $std(X) = \sqrt{V(X)} = \sqrt{1} = 10$
  - relative level of fluctuations $\frac{std(X)}{E(X)} = \frac{10}{10} = 1$

Distributions

A lot of random variables have similar probability structures for their events (and therefore enable similar mathematical analysis and have similar properties).
We can group them into families called Distributions
Examples:

- Tossing a fair coin
- Throwing a fair 6 sided die, is the result even or odd
- The sex of a random person in the world
- Throwing a fair 6 sided die, is the result '6'

Binomial Distribution

The random variable examples from the previous all have a Binary distribution. We have two possible outcomes (denote them A and B)
The probability of getting A is denoted by $P(A)$
The probability of getting B is $1 - P(A)$
In a Binomial distribution, we repeat a binary trial n times and count how many A we get.

- See ipython notebook!

Multinomial Distribution

- See ipython notebook
- also relevant to REAL MICROBIOME experiments!

Normal Distribution and Central limit theorem

The sum of a lot of INDEPENDENT random variables always goes to the same distribution
It is called the Normal Distribution
Happens a lot in nature.
See IPython notebook

The End