Data Science 1
Probability
Equally Likely Approach

Edward L. Boone

Introduction

From the previous videos we understand what a Sample Space $\mathcal{S}$ is and what events are.

Now we want to know how to assign a probability to an event E, denoted $P(E)$.

One way is to use the Empirical approach which looks at the proportion of the time the event was observed.

$$\hat{P}(E) = \frac{n(E)}{n}$$

where n is the number of trials we observed and $n(E)$ is the number of those trials where E occured.

Here we use $\hat{P}$ to show that the probability is estimated from data.

However, this approach can be very time consuming and expensive as the experiment must be conducted and the outcomes must be observed and counted.

An Alternative

An alternative approach is to consider experiments where each outcome is *equally likely* and then use those to create probabilities of Events.

This works for many simple experiments.

Since the outcomes are equally likely we don't need to actually conduct the experiment.

However, we will need to be able to count (this can be more complicated than you think).

Simple Example

Many board games are played using one or more dice. If we are looking at a simple six sided die, then the sample space is:

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

If the dice is *fair* then the each face is *equally likely* to face upwards after the dice is rolled.

Since there are only six outcomes and they are equally likely each face has a probabilty of $1/6$ of facing upwards after a roll.

$$P\left(\epsdice{1}\right) = P\left(\epsdice{2}\right) = P\left(\epsdice{3}\right) = P\left(\epsdice{4}\right) = P\left(\epsdice{5}\right) = P\left(\epsdice{6}\right) = 1/6$$

Probability of Events

In order to calculate the probability of events we need to count the number of outcomes in the event as well as the number of outcomes in the sample space.

Let $|\cdot|$ denote the number of outcomes in $\cdot$.

Using this we can define the probability of and event E to be:

$$P(E) = \frac{|E|}{|\mathcal{S}|}$$

This formula works only if all the outcomes in $\mathcal{S}$ are equally likely.

This also requires us to be able to count the number of outcomes in each.

Simple Example

Suppose I am playing a simple board game with my daughter and I am in a position that if I roll a single die and get either ⚀ or ⚃ then I can stay in the game. Otherwise she wins the game. What is the probability that I stay in the game after the next roll?

Here $E = \{ \text{⚀}, \text{⚃} \}$ and $\mathcal{S} = \{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅}\}$. Hence,

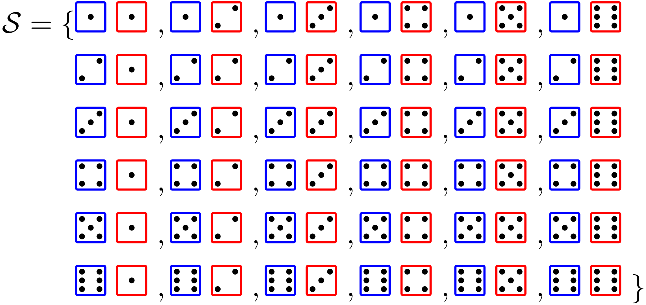
$$P(E) = \frac{|E|}{|\mathcal{S}|} = \frac{|\{ \text{⚀}, \text{⚃} \}|}{|\{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅}\}|} = \frac{2}{6} = \frac{1}{3}$$

Thus the probability that I stay in the game is $1/3$.

To use the Empirical probability approach I would have had to roll the dice many, many, many times and record the number of times that ⚀ or ⚃. In this approach we didn't even conduct the experiment once.

Two Dice

Things get considerably more complicated when playing a game with two dice. Since at each roll two faces are facing up we need to record two faces at a time. Suppose one die is blue and the other is red. Then the Sample Space is:

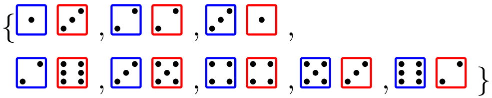


There are $|\mathcal{S}| = 36$ possible outcomes!

Two Dice

Suppose I am playing a board game with my daughter and I need a 4 or 8 to stay in the game. What is the probability that I stay in the game after the next roll?

$E =$



There are $|E| = 8$ outcomes that meet the criterion.

This gives us the probability of staying in the game as:

$$P(E) = \frac{|E|}{|S|} = \frac{8}{36} = \frac{2}{9} \approx 0.2222$$

Summary

The *Relative Frequency* approach to probability is one way to assign probabilities to events.

- Requires each outcome to be equally likely.
- Requires the ability to count all outcomes in the sample space and the event.
- Cheaper than actually running the experiment over and over.

In order to be successful at this we need to learn how to count.