ENGR 1330 Computational Thinking with Data Science

Finding Probabilities

Outline

- Concepts of sample, population, and probabilities
- Computing probability: single events, both events, at least event.

Objective

- Be able to find probabilities of events.

Sample and Population

- A **population** includes all of the elements from a set of data.

  Example: All people living in the US

- A **sample** consists one or more observations drawn from the population.

  Draw 1,000 people in all people living in the US

Probabilities

- By convention, probabilities are numbers between 0 and 1, or, equivalently, 0% and 100%; denoted by:

P(event)

- Impossible events have probability 0.
- Events that are certain have probability 1.

$$P(\textit{an event does not happen}) = 1 - P(\textit{the event happens})$$

Compute probability

When all outcomes are equally likely:

- Example: rolling an ordinary die; we can assume six faces are equally likely.

⇒Probability that the die shows an even number is:

P(shows an even number) = $\#\{2, 4, 6\}/\#\{1, 2, 3, 4, 5, 6\} = 3/6 = 0.5$

P(die shows a multiple of three) = $\#\{3, 6\}/\#\{1, 2, 3, 4, 5, 6\} = 0.333$

P(an event happens) = $\#\{outcomes\ that\ make\ the\ event\ happen\}/\#\{all\ outcomes\}$

Compute probability

When two events must both happen:

- Example: A box that contains three tickets: one red, one blue, and one green.
- Draw two tickets at random without replacement; that is, you shuffle the three tickets, draw one, shuffle the remaining two, and draw another from those two.
- What is the chance you get the green ticket first, followed by the red one?

Compute probability

Possible pairs of colors: RB, RG, BR, BG, GR, GB

P(GR) = $\#\{GR\}/\#\{RB, RG, BR, BG, GR, GB\}$ = 1/6

Other solution: green ticket picked first => 1/3
Red is next (two tickets remaining to pickup red): => ½

P(green first, then red) = $1/3 * 1/2 = 1/6$

Multiplication rule:

P(two events both happen) = P(one event happens) *
P(the other event happens, given that the first one happened)

Compute probability

When an Event an Happen in Two Different Ways :

we want the chance that one of the two tickets is green and the other red.

=> This event doesn't specify the order in which the colors must appear. So they can appear in either order.

P(one green and one red) = P(GR) + P(RG) = 1/6 + 1/6 = 1/3

Additive rule:

P(an event happens) = P(first way it can happen) +
P(second way it can happen)

Computing probability

At least one success:

We used to question about the likelihood that a particular individual in a population is selected to be in the sample.

The individual is called “success”. The problem is now finding the chance the sample contains a success.

Example: tossing a coin twice: HH, HT, TH, TT

⇒ The chance of getting at least one head in two tosses is: 3/4

Other solution:

P(at least one head in two tosses) = 1 – P(both tails) = 1 – ¼ = 3/4

P(both tails) = ½ * ½ = 1/4 Thanks to multiplication rule.

P(at least one head in 17 tosses) = ?

Computing probability

P(at least one head in 17 tosses) = $1 - (1/2)^{17}$

P(a single roll is not 6) = ?

P(a single roll is not 6) = P(1) + P(2) + P(3) + P(4) + P(5) = 5/6

P(at least one 6 in two rolls) = 1 – P(both rolls are not six) = $1 - (5/6)^2$

P(at least one 6 in 17 rolls) = 1 – P(17 rolls are not six) = $1 - (5/6)^{17}$

Compute probability

```python
import pandas as pd

numRolls = []
probabilities = []
for i in range(50):
    numRolls.append(i)
    probabilities.append(1-(5/6)**i)

rolls = {
    "NumRolls": numRolls,
    "Prob at least 6": probabilities
}
```

```python
df = pd.DataFrame(rolls)
df.plot.scatter(x="NumRolls", y="Prob at least 6")
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x17a2119
```