Probability and Statistics

Topic 6 - Probability Rules

Aamir Alaud Din, PhD

October 16, 2023

TABLE OF CONTENTS

1. OBJECTIVES
2. THE WHY SECTION
3. PROBABILITY RULES
4. COMPUTING PROBABILITIES: EMPIRICAL METHOD
5. COMPUTING PROBABILITIES: CLASSICAL METHOD
6. SIMULATION TO GENERATE DATA BASED ON PROBABILITIES
7. SUBJECTIVE PROBABILITIES
8. SUMMARY

TABLE OF CONTENTS

1. OBJECTIVES
2. THE WHY SECTION
3. PROBABILITY RULES
4. COMPUTING PROBABILITIES: EMPIRICAL METHOD
5. COMPUTING PROBABILITIES: CLASSICAL METHOD
6. SIMULATION TO GENERATE DATA BASED ON PROBABILITIES
7. SUBJECTIVE PROBABILITIES
8. SUMMARY

OBJECTIVES

After learning this topic and studying, you should be able to:

1. Apply the rules of probabilities
2. Compute and interpret probabilities using the empirical method
3. Compute and interpret probabilities using the classical method
4. Use simulation to obtain data based on probabilities
5. Recognize and interpret subjective probabilities

TABLE OF CONTENTS

1. OBJECTIVES
2. THE WHY SECTION
3. PROBABILITY RULES
4. COMPUTING PROBABILITIES: EMPIRICAL METHOD
5. COMPUTING PROBABILITIES: CLASSICAL METHOD
6. SIMULATION TO GENERATE DATA BASED ON PROBABILITIES
7. SUBJECTIVE PROBABILITIES
8. SUMMARY

THE WHY SECTION

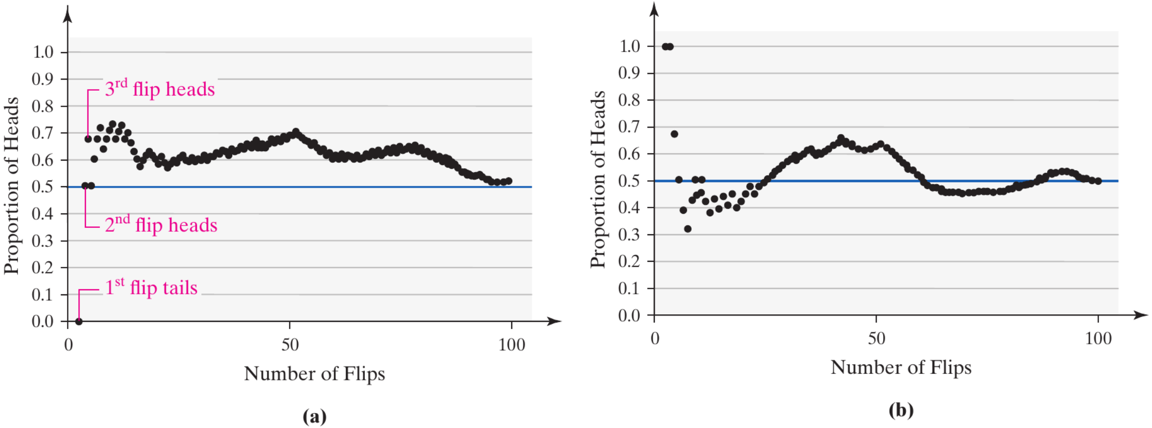
- In topic 1, we studied that the statistical parameters are reported with a certian confidence limit.
- This confidence limit is linked to probability.
- Probability is calculated using classical and empirical methods.
- The probability of an event in a set of data follows certain rules.
- Our aim is to know the rules and probability and to compute probabilities using empirical and classical methods.

TABLE OF CONTENTS

1. OBJECTIVES
2. THE WHY SECTION
3. PROBABILITY RULES
4. COMPUTING PROBABILITIES: EMPIRICAL METHOD
5. COMPUTING PROBABILITIES: CLASSICAL METHOD
6. SIMULATION TO GENERATE DATA BASED ON PROBABILITIES
7. SUBJECTIVE PROBABILITIES
8. SUMMARY

PROBABILITY RULES

- Suppose we flip a coin 100 times and compute the proportion of heads observed 0 after each toss of the coin.
- The first flip is tails, so the proportion of heads is $\frac{0}{1} = 0$; the second flip is heads, so the proportion of heads is $\frac{1}{2} = 0.5$; the third flip is heads, so the proportion of heads is $\frac{2}{3} = 0.667$, and so on.
- Plot the proportion of heads versus the number of flips and obtain the graph in Figure 1(a).
- We repeat this experiment with the results shown in Figure 1(b).
- The graphs in Figures 1(a) and (b) show that in the short term (fewer flips) the observed proportion of heads is different and unpredictable for each experiment.



Proportion of Heads
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
3rd flip heads
2nd flip heads
1st flip tails
0
50
100
Number of Flips
(a)
Proportion of Heads
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
50
100
Number of Flips
(b)

PROBABILITY RULES

- As the number of flips increases, however, both graphs tend toward a proportion of 0.5.
- This is the basic premise of probability.
- Probability is the measure of the likelihood of a random phenomenon or chance behavior occurring.
- It deals with experiments that yield random short-term results or outcomes yet reveal long-term predictability.
- The long-term proportion in which a certain outcome is observed is the probability of that outcome.
- So we say that the probability of observing a head is $\frac{1}{2}$ or 50% or 0.5 because, as we flip the coin more times, the proportion of heads tends toward $\frac{1}{2}$.

- This phenomenon, which is illustrated in Figure above, is referred to as the Law of Large Numbers.

The Law of Large Numbers

As the number of repetitions of a probability experiment increases, the proportion with which a certain outcome is observed gets closer to the probability of the outcome.

- We now introduce some terminology that we will need to study probability.

Homework 3

Write a python program to draw the plot like the one shown in figure above. Write the code for drawing an ace from the deck of 52 playing cards.

PROBABILITY RULES

- We now introduce some terminology that we will need to study probability.

Sample Space

The sample space, S, of a probability experiment is the collection of all possible outcomes.

- We now introduce some terminology that we will need to study probability.

Sample Space

An event is any collection of outcomes from a probability experiment. An event consists of one outcome or more than one outcome. We will denote events with one outcome, sometimes called simple events, e_i. In general, events are denoted using capital letters such as E.

PROBABILITY RULES

- In the following probability rules, the notation $P(E)$ means "the probability that event E occurs."

Rules of Probabilit

1. The probability of any event E, $P(E)$, must be greater than or equal to 0 and less than or equal to 1. That is, $0 \leq P(E) \leq 1$.
2. The sum of the probabilities of all outcomes must equal 1. That is, if the complete sample space $S = \{e_1, e_2, \ldots, e_n\}$, then

$$P(e_1) + P(e_2) + \cdots + P(e_n) = 1$$

- If an event is impossible, the probability of the event is 0.

PROBABILITY RULES

- If an event is a certainty, the probability of the event is 1.
- The closer a probability is to 1, the more likely the event will occur.
- The closer a probability is to 0, the less likely the event will occur.
- Be careful of this interpretation. An event with a probability of 0.75 does not have to occur 75 times out of 100.
- Rather, we expect the number of occurrences to be close to 75 in 100 trials.
- The more repetitions of the probability experiment, the closer the proportion with which the event occurs will be to 0.75 (the Law of Large Numbers).
- One goal of this course is to learn how probabilities can be used to identify unusual events.

Unusual Event

An unusual event is an event that has a low probability of occurring.

- Typically, an event with a probability less than 0.05 (or 5%) is considered unusual, but this cutoff point is not set in stone.
- The researcher and the context of the problem determine the probability that separates unusual events from not so unusual events.
- For example, suppose that the probability of being wrongly convicted of a capital crime punishable by death is 3%.
- Even though 3% is below our 5% cutoff point, this probability is too high in light of the consequences (death for the wrongly convicted), so the event is not unusual (unlikely) enough.

PROBABILITY RULES

- We would want this probability to be much closer to zero.
- Now suppose that you are planning a picnic on a day having a 3
- In this context, you would consider "rain" an unusual (unlikely) event and proceed with the picnic plans.
- The point is this: Selecting a probability that separates unusual events from not so unusual events is subjective and depends on the situation.
- Statisticians typically use cutoff points of 0.01, 0.05, and 0.10.
- Next, we introduce three methods for determining the probability of an event: the empirical method, the classical method, and the subjective method.

PROBABILITY RULES

EXAMPLE 1

A probability experiment consists of rolling a single fair die.

1. Identify the outcomes of the probability experiment.
2. Determine the sample space.
3. Define the event $E =$ "roll an even number."

EXAMPLE 2

The color of a plain M&M milk chocolate candy can be brown, yellow, red, blue, orange, or green. Suppose a candy is randomly selected from a bag. Table 1 shows each color and the probability of drawing that color. Show that the data represents a probability model.

Table 1

Color	Probability
Brown	0.13
Yellow	0.14
Red	0.13
Blue	0.24
Orange	0.20
Green	0.16

TABLE OF CONTENTS

1. OBJECTIVES
2. THE WHY SECTION
3. PROBABILITY RULES
4. COMPUTING PROBABILITIES: EMPIRICAL METHOD
5. COMPUTING PROBABILITIES: CLASSICAL METHOD
6. SIMULATION TO GENERATE DATA BASED ON PROBABILITIES
7. SUBJECTIVE PROBABILITIES
8. SUMMARY

COMPUTING PROBABILITIES: EMPIRICAL METHOD

- Probabilities deal with the likelihood that a particular outcome will be observed.
- For this reason, we begin our discussion of determining probabilities using the idea of relative frequency.
- Probabilities computed in this manner rely on empirical evidence, that is, evidence based on the outcomes of a probability experiment.

Approximating Probabilities Using the Empirical Approach

The probability of an event E is approximately the number of times event E is observed divided by the number of repetitions of the experiment.

$$P(E) \approx \text{relative frequency of } E = \frac{\text{frequency of } E}{\text{number of trials of experiment}}$$

COMPUTING PROBABILITIES: EMPIRICAL METHOD

- When we find probabilities using the empirical approach, the result is approximate because different trials of the experiment lead to different outcomes and, therefore, different estimates of $P(E)$.
- For example, probability of finding the event "head" in 20 flips in two different experiments will not be exactly the same.
- Surveys are probability experiments. Why?
- Each time a survey is conducted, a different random sample of individuals is selected.
- Therefore, the results of a survey are likely to be different each time the survey is conducted because different people are included.

COMPUTING PROBABILITIES: EMPIRICAL METHOD

EXAMPLE 3

An insurance agent currently insures 182 teenage drivers (ages 16 to 19). Last year, 24 of the teenagers had to file a claim on their auto policy. Find the probability that a teenager will file a claim in a given year.

EXAMPLE 4

The data in Table 2 represent the results of a survey in which 200 people were asked their means of travel to work.

1. Use the survey data to build a probability model for means of travel to work.
2. Estimate the probability that a randomly selected individual carpools to work. Interpret this result.
3. Would it be unusual to randomly select an individual who walks to work?

COMPUTING PROBABILITIES: EMPIRICAL METHOD

Table 2

Means of Travel	Frequency
Drive alone	153
Carpool	22
Public transportation	10
Walk	5
Other means	3
Work at home	7

Table 3

Means of Travel	Probability
Drive alone	0.765
Carpool	0.11
Public transportation	0.05
Walk	0.025
Other means	0.015
Work at home	0.035

TABLE OF CONTENTS

1. OBJECTIVES
2. THE WHY SECTION
3. PROBABILITY RULES
4. COMPUTING PROBABILITIES: EMPIRICAL METHOD
5. COMPUTING PROBABILITIES: CLASSICAL METHOD
6. SIMULATION TO GENERATE DATA BASED ON PROBABILITIES
7. SUBJECTIVE PROBABILITIES
8. SUMMARY

COMPUTING PROBABILITIES: CLASSICAL METHOD

- The empirical method gives an approximate probability of an event by conducting a probability experiment.
- The classical method of computing probabilities does not require that a probability experiment actually be performed.
- Rather, it relies on counting techniques.
- The classical method of computing probabilities requires equally likely outcomes.
- An experiment has equally likely outcomes when each outcome has the same probability of occurring.
- For example, when a fair die is thrown once, each of the six outcomes in the sample space, $\{1, 2, 3, 4, 5, 6\}$, has an equal chance of occurring whereas this is not the case with a loaded die.

COMPUTING PROBABILITIES: CLASSICAL METHOD

Computing Probability Using the Classical Method

If an experiment has n equally likely outcomes and if the number of ways that an event E can occur is m, then the probability of E, $P(E)$, is

$$P(E) = \frac{\text{number of ways that E can occur}}{\text{number of possible outcomes}} = \frac{m}{n}$$

So, if S is the sample space of this experiment,

$$P(E) = \frac{N(E)}{N(S)}$$

where $N(E)$ is the number of outcomes in E, and $N(S)$ is the number of outcomes in the sample space.

COMPUTING PROBABILITIES: CLASSICAL METHOD

EXAMPLE 5

A pair of fair dice is rolled. Fair die are die where each outcome is equally likely.

1. Compute the probability of rolling a seven.
2. Compute the probability of rolling "snake eyes"; that is, compute the probability of rolling a two.
3. Comment on the likelihood of rolling a seven versus rolling a two.

- In simple random sampling, each individual has the same chance of being selected.
- Therefore, we can use the classical method to compute the probability of obtaining a specific sample.

COMPUTING PROBABILITIES: CLASSICAL METHOD

EXAMPLE 6

Sophia has three tickets to a concert, but Yolanda, Michael, Kevin, and Marissa all want to go to the concert with her. To be fair, Sophia randomly selects the two people who can go with her.

1. Determine the sample space of the experiment. In other words, list all possible simple random samples of size $n = 2$.
2. Compute the probability of the event "Michael and Kevin attend the concert."
3. Compute the probability of the event "Marissa attends the concert."
4. Interpret the probability in part (c).

Table 4

Yolanda, Michael	Yolanda, Kevin
Yolanda, Marissa	Michael, Kevin
Michael, Marissa	Kevin, Marissa

COMPUTING PROBABILITIES: CLASSICAL METHOD

EXAMPLE 7

Suppose that a survey asked 500 families with three children to disclose the gender of their children and found that 180 of the families had two boys and one girl.

1. Estimate the probability of having two boys and one girl in a three-child family using the empirical method.
2. Compute and interpret the probability of having two boys and one girl in a three-child family using the classical method, assuming boys and girls are equally likely.

- In comparing the results of Examples 7(a) and 7(b), notice that the two probabilities are slightly different.
- Empirical probabilities and classical probabilities often differ in value, but, as the number of repetitions of a probability experiment increases, the empirical probability should get closer to the classical probability.

COMPUTING PROBABILITIES: CLASSICAL METHOD

- However, it is possible that the two probabilities differ because having a boy or having a girl are not equally likely events.
- Maybe the probability of having a boy is 0.505 and the probability of having a girl is 0.495.
- Consider the Vital Statistics data in the Statistical Abstract of the United States.
- In 2012, for example, 2,022,000 boys and 1,931,000 girls were born.
- Based on empirical evidence, the probability of a boy is approximately

$$\frac{2,022,000}{2,022,000 + 1,931,000} = 0.512$$

- According to classical method, the result will be 0.5 because having boy or girl is equally likely.

TABLE OF CONTENTS

1. OBJECTIVES
2. THE WHY SECTION
3. PROBABILITY RULES
4. COMPUTING PROBABILITIES: EMPIRICAL METHOD
5. COMPUTING PROBABILITIES: CLASSICAL METHOD
6. SIMULATION TO GENERATE DATA BASED ON PROBABILITIES
7. SUBJECTIVE PROBABILITIES
8. SUMMARY

SIMULATION TO GENERATE DATA BASED ON PROBABILITIES

- Look at the functions in random package of python.
- Use the module to solve the following example.

EXAMPLE 8

1. Simulate the experiment of sampling 100 three-child families to estimate the probability that a three-child family has two boys.
2. Simulate the experiment of sampling 1000 three-child families to estimate the probability that a three-child family has two boys.
3. Compare result obtained with classical method.

TABLE OF CONTENTS

1. OBJECTIVES
2. THE WHY SECTION
3. PROBABILITY RULES
4. COMPUTING PROBABILITIES: EMPIRICAL METHOD
5. COMPUTING PROBABILITIES: CLASSICAL METHOD
6. SIMULATION TO GENERATE DATA BASED ON PROBABILITIES
7. SUBJECTIVE PROBABILITIES
8. SUMMARY

SUBJECTIVE PROBABILITIES

Subjective Probability

A subjective probability of an outcome is a probability obtained on the basis of personal judgment.

NOTE

The subtopic is neither technical and nor difficult, so, you are advised to study the subtopic yourself.

TABLE OF CONTENTS

1. OBJECTIVES
2. THE WHY SECTION
3. PROBABILITY RULES
4. COMPUTING PROBABILITIES: EMPIRICAL METHOD
5. COMPUTING PROBABILITIES: CLASSICAL METHOD
6. SIMULATION TO GENERATE DATA BASED ON PROBABILITIES
7. SUBJECTIVE PROBABILITIES
8. SUMMARY

SUMMARY

- The probability is the chance of occurrence of an event.
- As the number of repetitions of a probability experiment increases, the proportion with which a certain outcome is observed gets closer to the probability of the outcome.
- The collection of all possible outcomes is the sample space.
- An event is any collection of outcomes from a probability experiment.
- An event consists of one outcome or more than one outcome.
- The probability rules are:
  1. The probability of an event lies between 0 and 1.
  2. Sum of probabilities of all the events is 1.

SUMMARY

- The probability of an impossible event is 0.
- The probability of an event with certainty is 1.
- An unusual event is an event that has a low probability of occurring.
- The empirical probability of an event E is approximately the number of times event E is observed divided by the number of repetitions of the experiment.
- Empirical probabilities are approximate as they are based on data which changes by changing the samples.
- Surveys are probability experiments because each time a survey is conducted, a different random sample of individuals is selected.

SUMMARY

- If an experiment has n equally likely outcomes and if the number of ways that an event E can occur is m, then the probability of E, $P(E)$, is

$$P(E) = \frac{\text{number of ways that } E \text{ can occur}}{\text{number of possible outcomes}} = \frac{m}{n}$$

- The results of empirical and classical probability are slightly different due to their approach of computing probabilities.
- Simulation can be used to generate data of probability experiments.
- Simulation generated probability experiments may give reliable results comparable with classical probabilities.
- A subjective probability of an outcome is a probability obtained on the basis of personal judgment.


Thank You!