CENG 222
Probability and Statistics

Introduction and Overview

Nesli Erdoğmuş & Burak Galip Aslan

Contents

- Making decisions under uncertainty
- Overview of the lecture content
- A simple exercise

Making decisions under uncertainty

Decision: the process of making **choices** by identifying a decision, gathering information, and assessing alternative **resolutions**.

Making decisions under uncertainty

- **Decision:** the process of making **choices** by identifying a decision, gathering information, and assessing alternative **resolutions**.

- This course is about uncertainty, measuring and quantifying uncertainty, and making decisions under uncertainty. Loosely speaking, by uncertainty we mean the condition when results, outcomes, the nearest and remote future are **not** completely determined; their development depends on a number of factors and just on a pure **chance**.

Uncertainty

Uncertainty

- Why did not they know it in advance, when the event was scheduled?

Uncertainty

- Why did not they know it in advance, when the event was scheduled?

- To support these words, a meteorologist predicts, say, a 60% chance of rain. Why cannot she let us know exactly whether it will rain or not, so we'll know whether or not to take our umbrellas? Yes, because of uncertainty. Because she cannot always know the situation with future precipitation for sure.

Uncertainty

⚽ **Sparta Prag - Galatasaray**

Galatasaray eşleşmenin ilk ayağında zorlanmasına rağmen 3-2 kazanarak tur için avantaj elde etti. Sarı-kırmızılı ekip o maçın ardından oynadığı Ankaragücü deplasmanında da rakipler için korkutucu bir görüntü sergiledi. Kırmızı kart cezalısı Nelsson ve Avrupa listesine yazılamayan Köhn'den yararlanamayacak olsalar da sahaya çıkacak 11 tur için yeterli olacaktır. Sparta Prag'da sol stoper Krejci, sol kanat bek Rynes ve sol açık Birmancevic kart cezaları sebebiyle sahaya çıkamayacak. Aynı hatta oynayan ve fark yaratan bir sistemin önemli dişlileri olan bu oyunculardan yararlanamayacak olmak Çek ekibini mutlaka zorlayacaktır. Galatasaray'ın bu karşılaşmayı da kazanacağına inanıyorum. `KG Var' alternatif seçenek olarak akıllarda yer almalı.

Making decisions under uncertainty

- This book is about measuring and dealing with uncertainty and randomness. Through basic theory and numerous examples, it teaches:

Making decisions under uncertainty

- This book is about measuring and dealing with uncertainty and randomness. Through basic theory and numerous examples, it teaches:
  - how to evaluate probabilities, or chances of different results (when the exact result is uncertain)

Making decisions under uncertainty

- This book is about measuring and dealing with uncertainty and randomness. Through basic theory and numerous examples, it teaches:
  - how to evaluate probabilities, or chances of different results (when the exact result is uncertain)
  - how to select a suitable model for a phenomenon containing uncertainty and use it in subsequent decision making

Making decisions under uncertainty

- This book is about measuring and dealing with uncertainty and randomness. Through basic theory and numerous examples, it teaches:
  - how to evaluate probabilities, or chances of different results (when the exact result is uncertain)
  - how to select a suitable model for a phenomenon containing uncertainty and use it in subsequent decision making
  - how to evaluate performance characteristics and other important parameters for new devices and servers

Making decisions under uncertainty

- This book is about measuring and dealing with uncertainty and randomness. Through basic theory and numerous examples, it teaches:
  - how to evaluate probabilities, or chances of different results (when the exact result is uncertain)
  - how to select a suitable model for a phenomenon containing uncertainty and use it in subsequent decision making
  - how to evaluate performance characteristics and other important parameters for new devices and servers
  - how to make optimal decisions under uncertainty.

Overview of the texbook

- **Chapter 2** introduces a language that we'll use to describe and quantify uncertainty. It is a language of **Probability**.
- When outcomes are uncertain, one can identify more likely and less likely ones and assign, respectively, high and low probabilities to them.
- Probabilities are numbers between 0 and 1, with 0 being assigned to an impossible event and 1 being the probability of an event that occurs for sure.

Overview of the texbook

- Using the introduced language, we shall discuss random variables as quantities that depend on chance.
- They assume different values with different probabilities. Due to uncertainty, an exact value of a random variable cannot be computed before this variable is actually observed or measured.
- Then, the best way to describe its behavior is to list all its possible values along with the corresponding probabilities.

Overview of the texbook

- Using the introduced language, we shall Such a collection of probabilities is called a **distribution**.
- Amazingly, many different phenomena of seemingly unrelated nature can be described by the same distribution or by the same **family of distributions**. This allows a rather general approach to the entire class of situations involving uncertainty.

Overview of the texbook

- As an application, it will be possible to compute probabilities of interest, once a suitable family of distributions is found. **Chapters 3 and 4** introduce families of distributions that are most commonly used in computer science and other fields.

Overview of the texbook

- In modern practice, however, one often deals with rather complicated random phenomena where computation of probabilities and other quantities of interest is far from being straightforward.
- In such situations, we will make use of **Monte Carlo methods**. Instead of direct computation, we shall learn methods of simulation or generation of random variables (**Chapter 5**).

Overview of the texbook

- If we are able to write a computer code for simulation of a certain phenomenon, we can immediately put it in a loop and simulate such a phenomenon thousands or millions of times and simply count how many times our event of interest occurred.
- This is how we shall distinguish more likely and less likely events. We can then **estimate** probability of an event by computing a proportion of simulations that led to the occurrence of this event.

Overview of the texbook

- In **Chapter 8**, we turn to Statistical Inference. While in Probability, we usually deal with more or less clearly described situations (models), in Statistics, all the analysis is based on **collected** and **observed data**.
- Given the data, a suitable model (say, a family of distributions) is fitted, its parameters are **estimated**, and conclusions are drawn concerning the entire totality of observed and unobserved subjects of interest that should follow the same model.

Probability example

A typical Probability problem sounds like this:

- A folder contains 50 executable files. When a computer virus or a hacker attacks the system, each file is affected with probability 0.2. Compute the probability that during a virus attack, more than 15 files get affected.

Probability example

A typical Probability problem sounds like this:

- A folder contains 50 executable files. When a computer virus or a hacker attacks the system, each file is affected with probability 0.2. Compute the probability that during a virus attack, more than 15 files get affected.
- Notice that the situation is rather clearly described, in terms of the total number of files and the chance of affecting each file. The only uncertain quantity is the number of affected files, which cannot be predicted for sure.

Statistics example

A typical Statistics problem sounds like this:

- A folder contains 50 executable files. When a computer virus or a hacker attacks the system, each file is affected with probability p. It has been observed that during a virus attack, 15 files got affected. Estimate p. Is there a strong indication that p is greater than 0.2?

Statistics example

A typical Statistics problem sounds like this:

- A folder contains 50 executable files. When a computer virus or a hacker attacks the system, each file is affected with probability p. It has been observed that during a virus attack, 15 files got affected. Estimate p. Is there a strong indication that p is greater than 0.2?
- This is a practical situation. A user only knows the objectively observed data: the number of files in the folder and the number of files that got affected.

Statistics example

A typical Statistics problem sounds like this:

- A folder contains 50 executable files. When a computer virus or a hacker attacks the system, each file is affected with probability p. It has been observed that during a virus attack, 15 files got affected. Estimate p. Is there a strong indication that p is greater than 0.2?
- Based on that, he needs to estimate p, the proportion of all the files, including the ones in his system and any similar systems.

Statistics example

- One may provide a point estimator of p, a real number, or may opt to construct confidence interval of “most probable” values of p.
- Similarly, a meteorologist may predict, say, a temperature of 70 degrees of F, which, realistically, does not exclude a possibility of 69 or 72 degrees, or she may give us an interval by promising, say, between 68 and 72 degrees.

Exercise 1.4

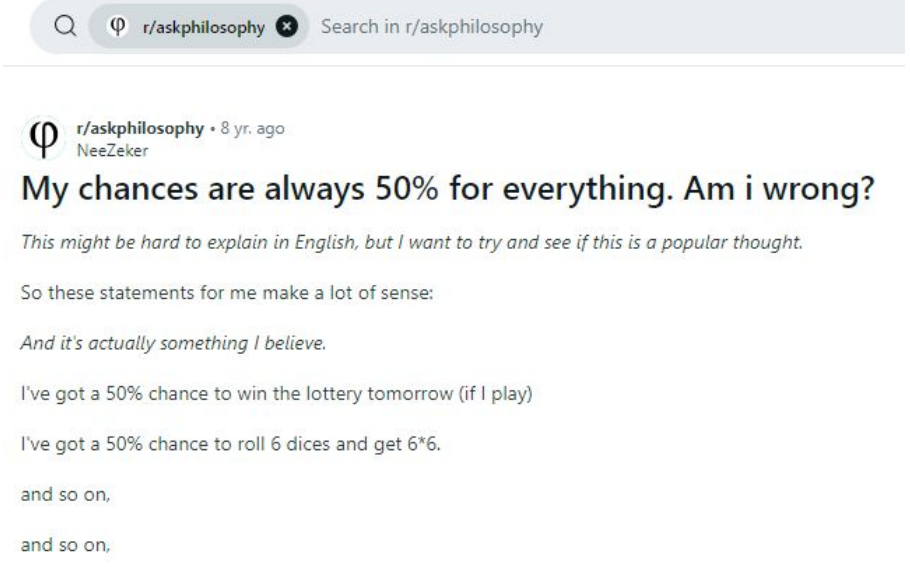
In a famous joke, a rather lazy student tosses a coin in order to decide what to do next:

- If it turns up heads, play a computer game.
- If tails, watch a video.
- If it stands on its edge, do the homework.
- If it hangs in the air, study for an exam.

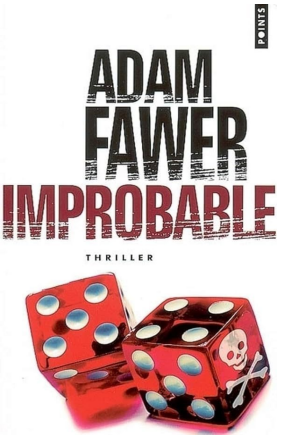
a. Which events should be assigned probability 0, probability 1, and some probability strictly between 0 and 1?

b. What probability between 0 and 1 would you assign to the event “watch a video”, and how does it help you to define “a fair coin”?

Does this student sound familiar to you?



Have you read this book?



International Thriller Writers Award recipients[11] [hide]

Year	Category	Author	Title	Publisher
2006	Best First Novel	Adam Fawer	*Improbable*	William Morrow
	Best Novel	Christopher Reich	*The Patriots Club*	Delacorte Press
	Best Paperback Original	R. Cameron Cooke	*Pride Runs Deep*	Jove
	Best Screenplay		*Caché*	Michael Haneke
	Thrillermaster	Clive Cussler		

Impossible vs. Implausible vs. Improbable

- Impossible is, it can't happen.
- Improbable is that is not likely to happen.
- Implausible is that it is not logical that such a thing would happen.

References

WIKIPEDIA

Google images