Judgment and Decision Making

Lecture Notes

© Prof. Dr. Stephan Huber

June 15, 2023

Errata and significant changes to the initial version

- Juni 1: I revised the notes changing several small things. In particular, I changed font, I added some sections and exercises in chapter 4.

HOCHSCHULE FRESENIUS
UNIVERSITY OF APPLIED SCIENCES

This script aims to support my lecture at the HS Fresenius. It is incomplete and no substitute for taking actively part in class. Do not distribute without permission. I am thankful for comments and suggestions for improvement.

Contact

Prof. Dr. Stephan Huber
Hochschule Fresenius für Wirtschaft & Medien GmbH
Im MediaPark 4c
50670 Cologne

Office: 4b OG-1 Bü01
www.hs-fresenius.de

+49 221 973199-523
stephan.huber@hs-fresenius.de
www.t1p.de/stephanhuber

Contents

List of Exercises

List of Solutions

Preface

- Questions, comments, and suggestions are welcome.
- Material such as slides and lecture notes can be found on ILIAS.
- Workload: 62.5 h = 28 h (in-class) + 34.5 h (self-study)
- 5 Credit Points (overall) by passing a written exam with a grade better or equal to 4.0.
- Learning Outcomes:
  The students understand the process of human judgment and its implications for decision making. Areas of investigation include, but are not limited to: probability judgment, causal judgment, choice under uncertainty and moral judgment. The students know descriptive models of judgment. In particular, they understand how the discrepancies they reveal with respect to normative models can be used to build prescriptive models aiming at better judgment and decision making. Students are able to identify these discrepancies in relevant fields such as business management and to relate them to applicable models. Students are able to understand how the outcome of judgment and decision making may differ under various frameworks. They recognize the role of choice architecture in decision outcomes. They have acquired the skills to appropriately evaluate the potential and limitations of algorithm-driven decisions as well as the advantages and pitfalls of collective judgment.
- My teaching principle is **KISS** which stands for **keep it simple and straightforward**.

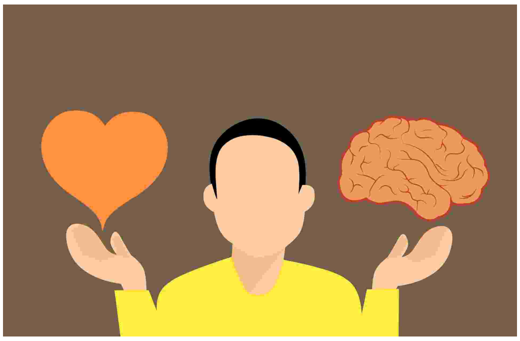
  The KISS principle states that most systems work best if they are kept simple rather than complicated; therefore, simplicity should be a key goal in design, and unnecessary complexity should be avoided.

  > However, KISS does not mean that the course is easy! If you are not able to think logically or if you are not willing to work hard, you may have problems passing the course

  Given your talent and mental capacities, I try to maximize your ability and self-confidence to solve future problems (in life and work).
- I am convinced that reading the lecture notes, preparing for class, taking actively part in class, and trying to solve the exercises without going straight to the solutions is the best method for students to
  - maximize leisure time and minimize the time needed to prepare for the exam, respectively,
  - getting long-term benefits out of the course,
  - improve grades, and
  - have more fun during lecture hours.

Chapter 1

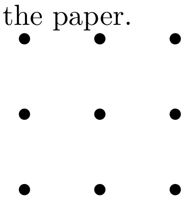
Introduction



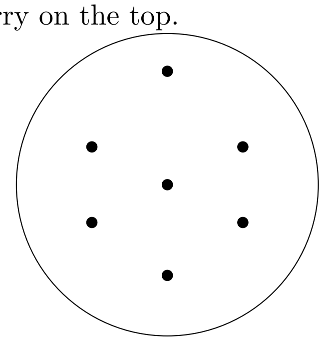
Exercise 1.1 — Puzzles (Solution → p. ??)

Try to solve the two puzzles below:

1. **The nine dots problem** Connect the dots shown below with no more than 4 straight lines without lifting your hand from the paper.



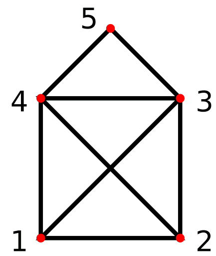
2. **The tasty cake puzzle** This picture below is a tasty cake with the nine dots representing strawberries. Cut this cake up with exactly four straight cuts so that each portion of the cake contains just one strawberry on the top.



3. Reflect on how you tried to solve the puzzles. Did you have a problem solving strategy? How did you come to the right decision? Think of restrictions you imposed on yourself which was not inherent to the problem.

Exercise 1.2 — What is the house of Santa Claus? (Solution → p. 8)

The house of Santa Claus is an old German drawing game. It goes like this: You have to draw a house in one line where you (a) must start at bottom left (point 1), (b) you are not allowed to lift your pencil while drawing and (c) it is forbidden to repeat a line. During drawing you say: "Das ist das Haus des Nikolaus".

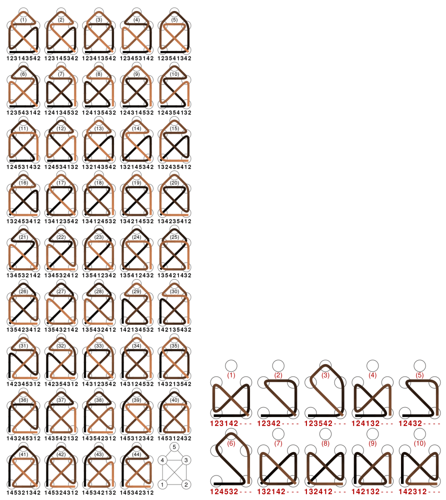


Try it and answer the following question: What is the success-rate?

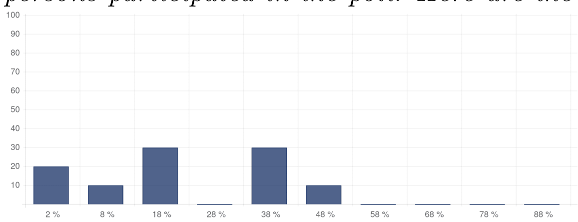
Answer the question here:

https://pingo.coactum.de/ Code: 666528

Solution to Exercise 1.2 — What is the house of Santa Claus? (Exercise → p. 8)



https://de.wikipedia.org/wiki/Haus_vom_Nikolaus
There are 44 solutions and only 10 different ways to fail. Thus, the probability to fail is about 18.5% and hence the probability to succeed is about 81.5%. In the course 10 persons participated in the poll. Here are the answers:



Nobody came close to the correct probability.

1.1 Definitions

The statement of Eilon (1969) still holds true:

> "An examination of the literature reveals the somewhat perplexing fact that most books on management and decision theory do not contain a specific definition of what is meant by a decision. One can find detailed descriptions of decision trees, discussions of game theory and analyses of various statistical treatments of payoffs matrices under conditions of uncertainty, but the definition of the decision activity itself is often taken for granted and is associated with making a choice between alternative courses of action."

1.1.1 Decision making

- The word *decision* stems from the latin verb *decidere* which can have different meanings[1] including make explicit, put an end to, bring to conclusion, settle/decide/agree (on), die, end up, fail, fall

[1] See http://www.latin-dictionary.net/search/latin/decidere

in ruin, fall/drop/hang/flow down/off/over, sink/drop, cut/notch/carve to delineate, detach, cut off/out/down, fell, flog thoroughly.

- Wikipedia (2020) defines *decision making* as follows:[2]

  > In psychology, decision-making [...] is regarded as the cognitive process resulting in the selection of a belief or a course of action among several alternative possibilities. Decision-making is the process of identifying and choosing alternatives based on the values, preferences and beliefs of the decision-maker. Every decision-making process produces a final choice, which may or may not prompt action. [...] Decision-making can be regarded as a problem-solving activity yielding a solution deemed to be optimal, or at least satisfactory. It is therefore a process which can be more or less rational or irrational and can be based on explicit or tacit knowledge and beliefs. Tacit knowledge is often used to fill the gaps in complex decision making processes.

- `www.businessdictionary.com` defines a *decision* as

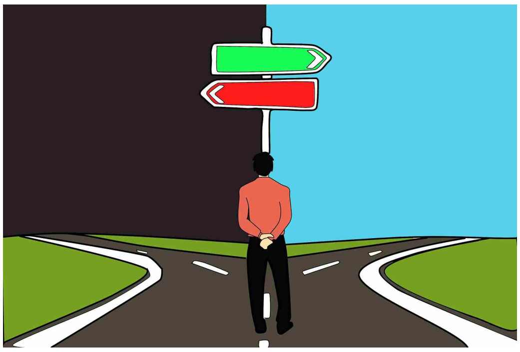
  > A choice made between alternative courses of action in a situation of uncertainty.

1.1.2 Judgment

`https://dictionary.cambridge.org` defines a *judgement* as

- the ability to form valuable opinions and make good decisions
- a decision or opinion about someone or something that you form after thinking carefully
- a decision
- an official legal decision
- a decision in a court of law
- a decision that you make, or an opinion that you have, after considering all the facts in a situation
- the ability to make good decisions or to form sensible opinions about something

1.1.3 Working definition



Let's set aside judgment and agree on a working definition for the decision that follows.

> A decision is the point at which a choice is made between alternative—and usually competing—options. As such, it may be seen as a stepping-off point—the moment at which a commitment is made to one course of action to the exclusion of others. (Fitzgerald, 2002, p. 8)

[2] *tacit knowledge* is the sort of knowledge that is hard to communicate and transfer, respectively.

1.2 Characteristics of decisions

1.2.1 Nature of decisions

All of our decisions can be roughly divided into two generic types (see Fitzgerald, 2002, p. 9f):

1. **routine decisions** Decisions that must be made at regular intervals.
2. **non-routine** Unique, random, non-recurring decision situations.

Another common method of dividing decisions into two categories is as follows:

1. **Operative Decisions:** This type of decision usually involves day-to-day business operations. There is a lot of overlap with the routine category here. Examples of this type of decision include setting production levels, deciding whether to hire or deciding whether to close a particular plant. When it comes to decisions in our daily lives, an example would be where, what, when, and what we eat for lunch.
2. **Strategic Decisions:** This type of decision is usually about company policy and direction over a long period of time. Examples of strategic decisions include entering a new market, acquiring a competitor, or withdrawing from an industry. Renting an apartment near the university or commuting to our parents is an example of a strategic decision in our personal lives.

1.2.2 Private vs. professional decision

Are the decisions we make in our daily lives different from those made by managers?

Yes and no. While private decisions affect fewer people on average, they are also about people (human resources), money (budgeting), buying and selling (marketing), how we do something (operations), or how we want to do it in the future (strategy and planning).

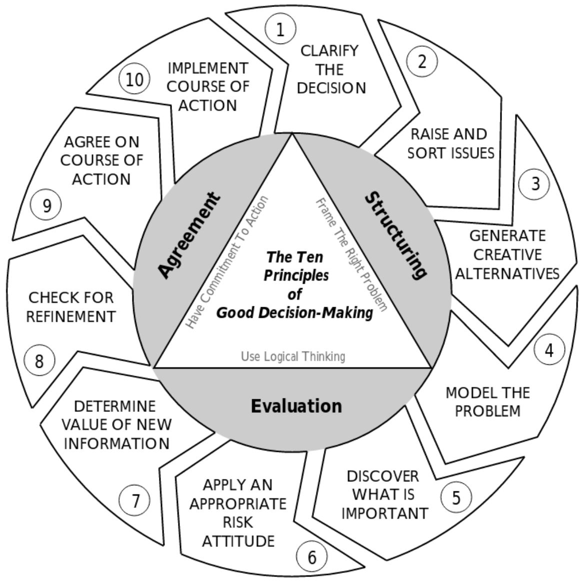
1.2.3 Scope of decisions

Some decisions are more important than others because the magnitude of the potential impact of a decision varies, i.e., the scope of a decision. For example, decisions can affect one person or millions, one pound/dollar or millions, one product/service or an entire market, one day or ten years, etc.

However, it is not entirely clear how to validate the scope. It depends a lot on the perspective of the decision maker. For example, for a small company, an investment of 10,000 euros can be a big decision, while for a multinational corporation it is a drop in the bucket. The scope for decision-making is therefore relative, not absolute. It depends entirely on the context in which the decision is made and on the characteristics of the person(s) making the decision.

1.3 Decision making strategies

1.3.1 Different schemes of a decision making process

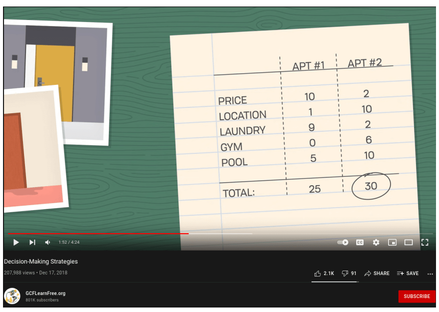


Exercise 1.3 — Pic a problem (Solution → p. ??)

- Choose a problem of your choice and try to solve the problem using the two illustrations above by making a good decision.
- Discuss in class whether these diagrams were helpful in making a wise decision or solving a problem

.

Exercise 1.4 — Decision making strategies (Solution → p. ??)



Watch https://youtu.be/pPIhAm_WGbQ. How is the nature of decisions discussed here? Does it contain a rational model of problem solving? Reflect on which ways to solve a problem and come to a decision, respectively, have been addressed.

1.3.2 The rational model

Rational decision making is a style of decision making based on objective data and a formal process of analysis. It excludes acting based on subjective feelings and intuitive approach. The model assumes that deducting decisions with full information and all alternatives allows for the creation of cognitive skills that allow the evaluation of all possible options and then the selection of the best one. (CEOpedia, 2021)

It essentially consists of a logical sequence of steps. Here is an example taken from Fitzgerald (2002, p. 13):

1. **Clearly identify the problem**. A *problem* can be defined as a perceived gap between the current reality and the desired reality.

2. **Generate potential solutions**. For routine decisions, various alternatives can be identified relatively easily using predetermined decision rules, but non-routine decisions require a creative process to find new alternatives.

3. Using appropriate analytical approaches, select a solution from the available alternatives, preferably the one with the largest expected value. In decision theory, this is called **maximizing the *expected utility* of the outcomes**.

4. **Implement the solution**. Managers often undermine implementation by not ensuring that those responsible for implementation understand and accept what they need to do, and that they have the motivation and resources needed for successful implementation.

5. **Evaluate the effectiveness of the implemented decision**

1.3.3 Rational models are not always helpful

While these rational models may be helpful, they can also be criticized in a number of ways. In particular, it is a belief that managers actually optimize their decision behavior based on rationality because they consciously select and implement the best alternatives. This is a misconception because any rational optimization is based on a number of dubious assumptions. For example, (see Fitzgerald, 2002, p. 13):

- it is hardly possible to know in advance all possible alternative solutions and to know in advance the specific outcomes that will result from each of them;
- there is actually an optimal solution, and this solution is among the alternatives identified;
- it is possible to accurately and numerically weight the various alternatives, the probabilities of their outcomes, and the relative desirability of these alternatives and outcomes;
- the decision makers always act rationally and therefore the decision making is is free of emotion, bias, and politics; and
- business decisions are driven entirely by the desire to maximize profit profits

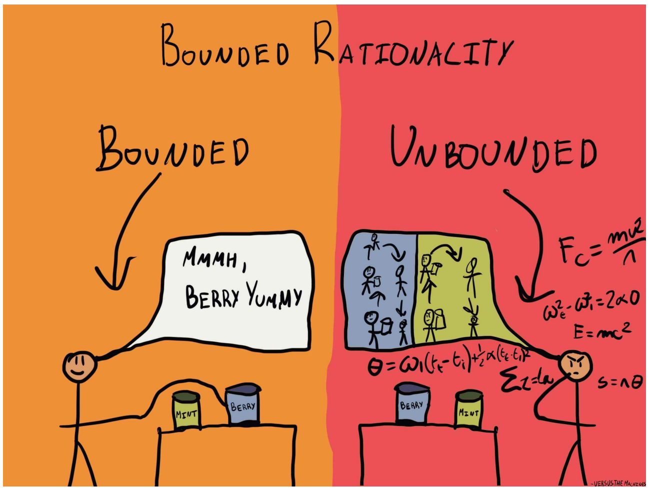
The **rational model can be considered normative** because it prescribes a strict logical sequence of steps to be followed in any decision-making situation. The rational models are based on the assumption that human behavior is logical and therefore predictable under certain circumstances. This is not necessarily what actually happens in the real world. Consider, for example, the findings in the field of behavioral economics, which clearly show that homo economicus is a (useful) concept that is weak in several aspects.

Bounded rationality

Source: Nobel Foundation archive

Herbert A. Simon (1916-2001) received the Nobel Memorial Prize in Economic Sciences in 1978 and the Turing Award[a] in 1975. He "combined different scientific disciplines and considered new factors in economic theories. Established economic theories held that enterprises and entrepreneurs all acted in completely rational ways, with the maximization of their own profit as their only goal. In contrast, Simon held that when making choices all people deviate from the strictly rational, and described companies as adaptable systems, with physical, personal, and social components. Through these perspectives, he was able to write about decision-making processes in modern society in an entirely new way"(NobelPrize.org, 2021). In particular, he proposed **bounded rationality** as an alternative basis for the mathematical and neoclassical economic modelling of decision-making, as used in economics, political science, and related disciplines.

[a] The Turing Award is an annual prize given by the Association for Computing Machinery (ACM) for contributions of lasting and major technical importance to the computer field. It is generally recognized as the highest distinction in computer science and is known as or often referred to as 'Nobel Prize of Computing'.



Source: *https://thedecisionlab.com/biases/bounded-rationality/*

Bounded rationality proposes that decision making is constrained by managers' ability to process information, i.e., the rationally is *bounded*. Managers use shortcuts and rules of thumb which are based on their prior experience with similar problems and scenarios. Given the constraints of managers in their position, they do not actually *optimize* their choice given the available information. It is more like finding a *satisfactory* solution, not necessarily the *best* of *optimal* solution.

Exercise 1.5 — Optimal vs. satisfactory solution (Solution → p. 14)

Using the following images, explain the idea of bounded rationality in the context of decision making.

Solution to Exercise 1.5 — Optimal vs. satisfactory solution (Exercise → p. 13)

a) Collecting and analyzing the available information about a product is costly. It is also difficult to analyze the importance of product features for the intended purpose.
b) Individuals often use rules of thumb to make a satisfactory decision.
c) It is difficult to understand complex situations such as the market for financial products. For some people, it is simply not possible to find the best product in these complex markets.
d) Consumers are often confronted with many variants of a product. The differences are negligible and therefore it is not worthwhile for consumers to analyze the situation in detail. Thus, they make a decision that may not be optimal, but they are satisfied with it.

1.3.4 Homo economicus

Exercise 1.6 — Are we irrational? (Solution → p. ??)

Discuss the following statement: *"Since the rationality of individuals is bounded and it is obvious that individuals do not make optimal decisions, we can say that individuals act irrationally."*

The so-called homo economicus is often modeled by the assumption of perfect rationality. Actors are assumed to always act in a way that maximizes their utility (as consumers) and profit (as producers), and to this end they are capable of arbitrarily complex reasoning. That is, it is assumed that they are always capable of thinking through all possible outcomes and choosing the course of action that leads to the best possible outcome.

Of course, the assumption of homo economicus is heroic, and I doubt that any serious economist has ever assumed that this assumption can be found one-to-one in reality. However, it is an assumption that helps to make predictions and explain the behavior of people and societies to some extent. I mean, what would happen if microeconomics went to the other extreme assuming irrationally acting individuals. The result would be, more or less, that we would not be able to make predictions and the future would be a random walk.

Nevertheless, economists and anyone who wants to make, apply, or think about economic theories should be aware of all the pitfalls that arise in our decision making and know that our ability to act rationally is limited. In real life, we often use heuristics to solve problems and make decisions.

A **heuristic** is any approach to solving a problem that uses a practical method that is not guaranteed to be optimal, perfect, or rational. However, a heuristic should–at best–be sufficient to achieve an immediate, short-term goal or approximation. Overall, people use heuristics because they either cannot act completely rationally or want to act rationally but do not have the time it would take to compute the perfect solution. Moreover, the effort is probably not worth it or simply not possible given the time constraints under which the problem must be solved.

Game theory is concerned with mathematical models of strategic interaction between rational decision makers. It is concerned with the way in which *interacting decisions* by *economic agents* produce outcomes in terms of the *preferences* (or *benefits*) of those agents, where the outcomes in question may not have been intended by any of the agents. More about this in chapter 4.

Exercise 1.7 — Lisa is pregnant (Solution → p. 59)

The following question stems from a study carried out by Kahneman and Tversky (1972): Lisa is thirty-three and is pregnant for the first time. She is worried about birth defects such as Down syndrome. Her doctor tells her that she need not worry too much because there is only a 1 in 1,000 chance that a woman of her age will have a baby with Down syndrome. Nevertheless, Lisa remains anxious about this possibility and decides to obtain a test, known as the Triple Screen, that can detect Down syndrome. The test is moderately accurate: When a baby has Down syndrome, the test delivers a positive result 86 percent of the time. There is, however, a small 'false positive' rate: 5 percent of babies produce a positive result despite not having Down syndrome. Lisa takes the Triple Screen and obtains a positive result for Down syndrome. Given this test result, **what are the chances that her baby has Down syndrome?**

a) 0-20 percent chance
b) 21-40 percent chance
c) 41-60 percent chance
d) 61-80 percent chance
e) 81-100 percent chance

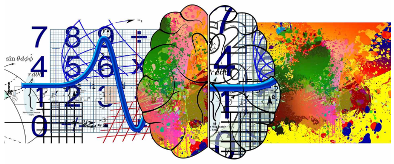
✎ Think about the question and put your answer here:
https://pingo.coactum.de/ Code: 666528

1.3.5 Review

- Decision analysis is about using information in order to come to a decision.
- A structured and rational process can help improve the chances of receiving good decision outcomes.
- As decision problems are often (too) complex to fully capture or solve rationally. Thus, a good decision analysis should try to use the available information and the existing understanding of the problem as transparent, consistent, and logical as possible.
- A complex decision problem should be simplified and hence decomposed into its basic and most important components.
- There are hundred of different *schemes* or *strategies* how to make decisions in certain circumstances. Many heuristics exist how to think, behave, and calculate to come to a wise decision.
- Mostly decisions are based on subjective expectations. These expectations are difficult to validate.
- Articulating exact expectation and preferences is a difficult task and the information that stems from these articulation is full of biases. Decision analytic tools need to take that into consideration.

Chapter 2

Rational models of decision making



2.1 Making a rational choice

Economists view the choices that people make as **rational**. A rational choice is one that compares **costs and benefits** and achieves the greatest benefit over cost for the person making the choice.

Exercise 2.1 — The businessman and the fisherman (Solution → p. ??)

A classic tale that exist in different version[a] goes like this:

> One day a fisherman was lying on a beautiful beach, with his fishing pole propped up in the sand and his solitary line cast out into the sparkling blue surf. He was enjoying the warmth of the afternoon sun and the prospect of catching a fish.
> About that time, a businessman came walking down the beach, trying to relieve some of the stress of his workday. He noticed the fisherman sitting on the beach and decided to find out why this fisherman was fishing instead of working harder to make a living for himself and his family. “You aren’t going to catch many fish that way,” said the businessman to the fisherman.
> “You should be working rather than lying on the beach!”
> The fisherman looked up at the businessman, smiled and replied, “And what will my reward be?”
> “Well, you can get bigger nets and catch more fish!” was the businessman’s answer.
> “And then what will my reward be?” asked the fisherman, still smiling. The businessman replied, “You will make money and you’ll be able to buy a boat, which will then result in larger catches of fish!”
> “And then what will my reward be?” asked the fisherman again.
> The businessman was beginning to get a little irritated with the fisherman’s questions.
> “You can buy a bigger boat, and hire some people to work for you!” he said.
> “And then what will my reward be?” repeated the fisherman.
> The businessman was getting angry. “Don’t you understand? You can build up a fleet of fishing boats, sail all over the world, and let all your employees catch fish for you!”

Once again the fisherman asked, “And then what will my reward be?”
The businessman was red with rage and shouted at the fisherman, “Don’t you understand that you can become so rich that you will never have to work for your living again! You can spend all the rest of your days sitting on this beach, looking at the sunset. You won’t have a care in the world!”
The fisherman, still smiling, looked up and said, “And what do you think I’m doing right now?”

Who is acting rationally here? The fishermen or the businessmen? What are the costs and benefits of both?

[a] This one stems from thestorytellers.com. A famous version stems from Paulo Coelho see https://paulocoelhoblog.com

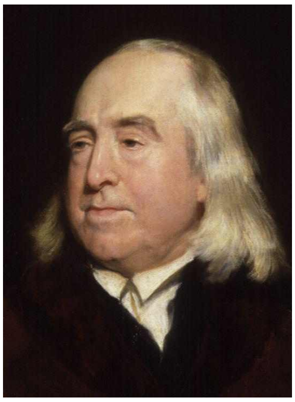
Benefit: What you gain

The utility of a thing is the gain or pleasure it brings and is determined by **preferences**. That is, what a person likes or dislikes, and the intensity of those feelings. Economists measure utility as the most a person **is willing to give up to get something**.

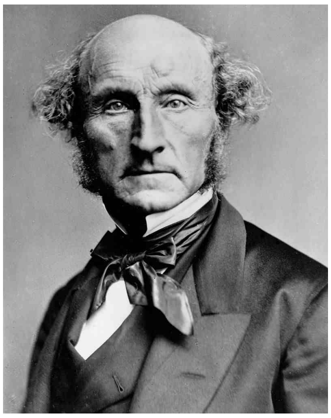
Cost: What you must give up

Thinking about a choice as a tradeoff emphasizes cost as an opportunity forgone. The **opportunity cost** of something is the highest-valued alternative that must be given up to get it.

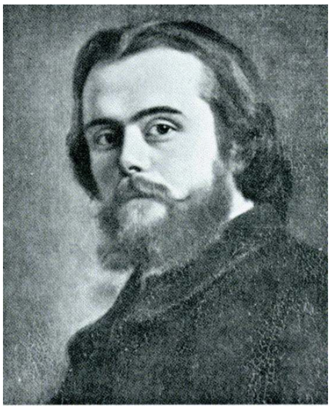
Utility as a general measure



Jeremy Bentham (1748-1832)



John Stuart Mill (1806-1873)



Léon Walras (1834-1910)

Within economics, the concept of utility is used to model worth or value. Its usage has evolved significantly over time. The term was introduced initially as a measure of pleasure or happiness within the theory of utilitarianism by moral philosophers such as Jeremy Bentham and John Stuart Mill. Later on it was popularized by Léon Walras. In Microeconomics it usually represents the satisfaction or pleasure that consumers receive for consuming a bundle of goods and services, respectively.

2.2 Three conditions of decision making

There are three general conditions that determine the design of the optimal decision making process: certainty, risk and uncertainty.

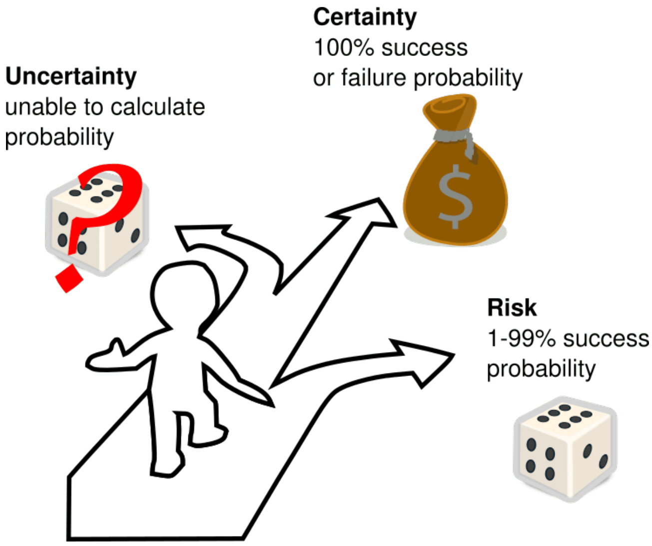


Figure 2.1: Conditions of decision making
Source: CEOpedia (2021)

certainty A condition under which taking a decision involves reasonable degree of certainty about its result, what are the opportunities and what conditions accompany this decision.

risk A condition under which taking a decision involves reasonable degree of certainty about its result, what are the opportunities and what conditions accompany this decision.

uncertainty A condition in which decision maker does not know all the choices, as well as risks associated with each of them and possible consequences.

Exercise 2.2 — Three categories (Solution → p. 18)

Read Finne (1998)[a].

a) Explain the three categories of decision making.
b) Give examples of the three categories of decision making.
c) Explain the four criteria for decision making under uncertainty.
d) Table 7 is full of errors. Can you play referee and correct them.

[a]Finne, T. (1998). The three categories of decision-making and information security. *Computers & Security*, 17(5):397–405

Solution to Exercise 2.2 — Three categories (Exercise → p. 18)

The example that is shown in Figure 7 of (Finne, 1998, p.401) contains some errors. Here is the correct table including the Hurwicz-values (we assume $\alpha = .5$):

O_{ij}	$min(\theta_1)$	$max(\theta_2)$	H_i
a_1	36	110	73
a_2	40	100	70
a_3	58	74	66
a_4	61	66	63.5

Thus, the order of preference is $a_4 \succ a_3 \succ a_1 \succ a_2$.

2.2.1 Decision making under certainty

Decision making under certainty means that we assume that we are certain about the future state of nature, i.e. we have **complete information** about the alternative actions we can take and their implied consequences. Taking decisions under certainty seem like a trivial exercise. However, there are many problems that are so complex that sometimes sophisticated mathematical techniques are needed to find the best solution.

Decision trees When we decide under certain conditions, decision trees can be a helpful tool, as explained in Finne (1998).

Payoff table In addition, a payoff table or decision matrix, as shown in Figure 2.2, is helpful because it clarifies the alternatives and outcomes:

Our decision(s) are made among a certain number of alternatives, denoted A_i. Moreover, there is more than one future *state of nature* or goal, which we denote by N_j. These future states of nature may differ in their probability of occurrence p_j. (Of course, the probability of occurrence for a state of nature is certainly equal to one).

The outcome, i.e., the payoff, depends on the alternative chosen and the future state of nature that occurs. For example, if one chooses alternative A_i and state of nature N_j occurs, the payoff is O_{ij}.

The goal, of course, is to choose the alternative A_i that gives us the most favorable outcome O_{ij}. In the following, we will briefly discuss some computational schemes that can support good decision making.

	State of Nature/Probability					
	N_1	N_2	$\dots$	N_j	$\dots$	N_n
Alternative	p_1	p_2	$\dots$	p_j	$\dots$	p_n
A_1	O_{11}	O_{12}	$\dots$	O_{1j}	$\dots$	O_{1n}
A_2	O_{21}	O_{22}	$\dots$	O_{2j}	$\dots$	O_{2n}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
A_i	O_{i1}	O_{i2}	$\dots$	O_{ij}	$\dots$	O_{in}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
A_m	O_{m1}	O_{m2}	$\dots$	O_{mj}	$\dots$	O_{mn}

Figure 2.2: Payoff matrix

For simplicity, let us assume that all objectives are independent from each other. Let's further assume that you are certain about the state of nature. Then you simply choose the alternative that gives you the best outcome for a given state of nature. This is trivial, and in most real-world scenarios it is not so simple.

Restaurant Suppose you have to decide between four different restaurants where you want to go out to eat $(a_1, \dots, a_4)$. The restaurants have different characteristics such as the quality of the food k_1, the quality of the music played k_2, the price k_3, the quality of the service k_4 and the environment k_5. The corresponding payoff table is given below. High numbers represent good quality. In other words, a_i refer to competing alternatives where to eat for dinner and k_i are certain characteristics of the respective restaurant and the numbers given in the table indicate the output.

	k_1	k_2	k_3	k_4	k_5
a_1	3	0	7	1	4
a_2	4	1	4	2	1
~~a_3~~	~~4~~	~~0~~	~~3~~	~~2~~	~~1~~
a_4	5	1	2	3	1

Figure 2.3: Example: Payoff table

Domination

Some alternative can be excluded because they are *dominated* by other alternatives. In Figure 2.3, you can see that alternative 2 is always superior to alternative 3, i.e., alternative 3 is dominated by alternative 2. Thus, it does not make sense to think about alternative 3.

Weighting

As the objectives k_j of Figure 2.3 do not represent different states of nature but represent characteristics and its corresponding utility (whatever that number may mean in particular) of one particular characteristics if we choose a respective alternative. For example, think of that you can choose one out of four restaurants, $a_1, \ldots, a_4$, in order to eat a five-course menu, $k_1, \ldots, k_5$, and the outcome represents your utility of each course, j, for each restaurant, i. We already know from the domination principle that restaurant a_3 is a bad choice.

Suppose you have a preference for the first three courses. Specifically, suppose that your preference scheme is as follows:

$$g_1 : g_2 : g_3 : g_4 : g_5 = 3 : 4 : 3 : 1 : 1$$

This means, for example, that you value the second food course four times more than the last course.

$$w_1 = 3/12; w_2 = 4/12; w_3 = 3/12; w_4 = w_5 = 1/12.$$

In order to find a decision criteria you can calculate the aggregated expected utility as follows:

$$\Phi(a_i) = \sum_{p=1}^{r} w_p \cdot u_{ip} \rightarrow max$$

The results are as follows:

	k_1	k_2	k_3	k_4	k_5	$\Phi(a_i)$
a_1	3	0	7	1	4	35/12
a_2	4	1	4	2	1	31/12
a_3	4	0	3	2	1	24/12
a_4	5	1	2	3	1	29/12

Thus, $a_1 \succ a_2 \succ a_4 \succ a_3$, i.e., you prefer alternative a_1.

Preferences and decision criteria

Instead of calculating the expected aggregated utility you can take your decision based on other preference schemes. For example, assume some courses are more important to you than others such as $k_1 \succ k_3 \succ k_2 \succ k_4 \succ k_5$. Now, if you decide based on your most important course you should choose restaurant 4 because this gives you the highest utility in that dish. If you further assume that you are not that hungry and you only like to have two dishes, then you should probably better go for restaurant 1 because the aggregated utility in the two preferred courses is 10 utility units and restaurants 2, 3 and 4 can only offer 8, 7 and 7 utility units in your two preferred courses.

Körth's Maximin-Rule According to this rule, we compare alternatives by the worst possible outcome under each alternative, and we should choose the one which maximizes the utility of the worst outcome. More concrete, the procedure consists of 4 steps:

1. Calculate the utility maximum for each column of the payoff matrix: $\overline{O}_j = \max_{i=1,\dots,m} O_{ij} \qquad \forall j$
2. Calculate for each cell the relative utility, $\frac{O_{ij}}{\overline{O}_j}$.
3. Calculate for each row the minimum:

$$\Phi(a_i) = \min_{j=1,\dots,p} \left(\frac{O_{ij}}{\overline{O}_j} \right)$$

4. Set preferences by maximizing $\Phi(a_i)$.

Exercise 2.3 — Körth: example (Solution → p. 21)

For the following payoff-matrix, calculate the order of preferences based on the Körth-rule.

O_{ij}	k_1	k_2	k_3	k_4	k_5
a_1	3	0	7	1	4
a_2	4	0	4	2	1
a_3	4	-1	3	2	1
a_4	5	1	3	3	1

Solution to Exercise 2.3 — Körth: example (Exercise → p. 21)

1. $\overline{O}_1 = 5; \overline{O}_2 = 1; \overline{O}_3 = 7; \overline{O}_4 = 3; \overline{O}_5 = 4$

2.

O_{ij}	k_1	k_2	k_3	k_4	k_5
a_1	*3/5*	*0*	*1*	*1/3*	*1*
a_2	*4/5*	*0*	*4/7*	*2/3*	*1/4*
a_3	*4/5*	*-1*	*3/7*	*2/3*	*1/4*
a_4	*1*	*1*	*3/7*	*1*	*1/4*

3.

O_{ij}	k_1	k_2	k_3	k_4	k_5	$\Phi(a_i)$
a_1	*3/5*	*0*	*1*	*1/3*	*1*	*0*
a_2	*4/5*	*0*	*4/7*	*2/3*	*1/4*	*0*
a_3	*4/5*	*-1*	*3/7*	*2/3*	*1/4*	*-1*
a_4	*1*	*1*	*3/7*	*1*	*1/4*	*1/4*

4. $a_4 \succ a_1 \sim a_2 \succ a_3$

Linear programming

Linear Programming is a common technique for decision making under certainty. It allows to express a desired benefit (such as profit) as a mathematical function of several variables. The solution is the set of values for the independent variables (decision variables) that serves to maximize the benefit or to minimize the negative outcome under consideration of certain limits, a.k.a. constraints. The method usually follows a four step procedure:

1. state the problem;
2. state the decision variables;
3. set up an objective function;
4. clarify the constraints.

Exercise 2.4 — Linear programming (Solution → p. 22)

Consider a factory producing two products, product X and product Y. The problem is this: If you can realize \$10.00 profit per unit of product X and \$14.00 per unit of product Y, what is the production level of x units of product X and y units of product Y that maximizes the profit P each day? Your production, and therefore your profit, is subject to resource limitations, or constraints. Assume in this example that you employ five workers—three machinists and two assemblers—and that each works only 40 hours a week.[a]

- Product X requires three hours of machining and one hour of assembly per unit.
- Product Y requires two hours of machining and two hours of assembly per unit.

[a]The example is taken from Morse et al. (2014, p. 134f).

Solution to Exercise 2.4 — Linear programming (Exercise → p. 22)

1. State the problem: *How many of product X and product Y to produce to maximize profit?*
2. Decision variables: *Let x = number of product X to produce per day*
Let y = number of product Y to produce per day

3. Objective function: *Maximize*

$$P = 10x + 14y$$

4. Constraints:

- *machine time=120h*
- *assembling time=80h*
- *hours needed for production of one good:*
  machine time: $x \to 3h$ *and* $y \to 2h$
  assembling time: $x \to 1h$ *and* $y \to 2h$

Thus, we get:

$$3x + 2y \leq 120 \quad \Leftrightarrow y \leq 60 - \frac{3}{2}x \quad \textit{(hours of machining time)}$$

$$x + 2y \leq 80 \quad \Leftrightarrow y \leq 40 - \frac{1}{2}x \quad \textit{(hours of assembly time)}$$

Since there are only two products, these limitations can be shown on a two-dimensional graph (Figure 2.4). Since all relationships are linear, the solution to our problem will fall at one of the corners.

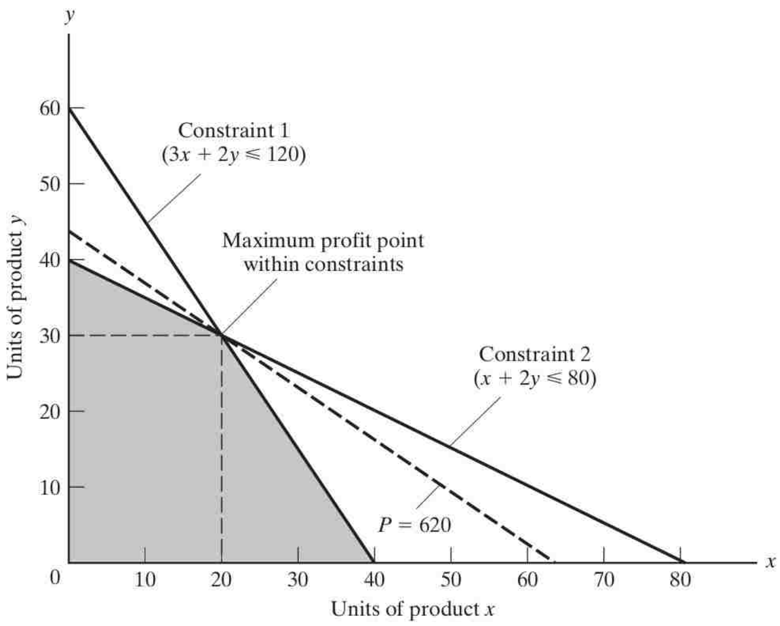


Figure 2.4: Linear program example: constraints and solution

To draw the isoprofit function in a plot with the good y on the y-axis and good x on the x-axis, we can re-arrange the objective function to get

$$y = \frac{1}{14}P - \frac{10}{14}x$$

To illustrate the function let us consider some arbitrarily chosen levels of profit in Figure 2.5:

- *$350 by selling 35 units of X or 25 units of Y*
- *$700 by selling 70 units of X or 50 units of Y*
- *$620 by selling 62 units of X or 44.3 units of Y.*

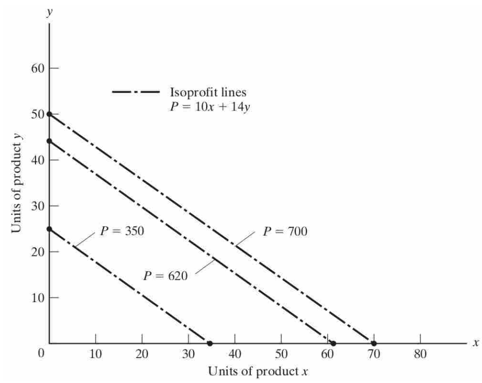


Figure 2.5: Linear program example: isoprofit lines

To find the solution, begin at some feasible solution (satisfying the given constraints) such as $(x, y) = (0, 0)$*, and proceed in the direction of steepest ascent of the profit function (in this case, by increasing production of Y at $14.00 profit per unit) until some constraint is reached. Since assembly hours are limited to 80, no more than 80/2, or 40, units of Y can be made, earning* $40 \cdot \$14.00$*, or $560 profit. Then proceed along the steepest allowable ascent from there (along the assembly constraint line) until another constraint (machining hours) is reached. At that point, (x,y) = (20,30) and profit P = (20 * +10.00) + (30 * +14.00), or $620. Since there is no remaining edge along which profit increases, this is the optimum solution.*

2.2.2 Decision making under uncertainty

When it comes to decision-making under uncertainty you need to have a criterion that helps you to come to a *rational* decision. The choice of criteria is a matter of taste and should at best fit with overall (business) strategy.

Laplace criterion

In case of different possible states of nature and no information on their probability of occurencies, the Laplace criterion simply assigns equal probabilities to the possible pay offs for each action and then selecting that alternative which corresponds to the maximum *expected* pay off. An example is given in Finne (1998).

Maximax criterion

If you like to *go for cup*, i.e., like to have the best out of the best possible outcome without taking into consideration the fact that this may potentially also result in the worst possible scenario, then, you can choose the alternative that gives the maximum possible output.

Minimax criterion

The Minimax (or maximin) criterion is a conservative criterion because it is based on making the best out of the worst possible conditions. Again, examples on how to calculate it are given in Finne (1998).

Savage Minimax criterion

This criterion aims to minimize the worst-case regret to perform as closely as possible to the optimal course. Since the minimax criterion applied here is to the regret (difference or ratio of the payoffs) rather than to the payoff itself, it is not as pessimistic as the ordinary minimax approach. For more details, please read Finne (1998).

Hurwicz criterion

The Hurwicz criterion allows the decision maker to calculates a weighted average between the best and worst possible payoff for each decision alternative. Then, the alternative with the maximum weighted average is chosen.

For each decision alternative, the weight α is used to compute Hurwicz the value:

$$H_i = \alpha \cdot \overline{O}_i + (1-\alpha) \cdot \underline{O}_i$$

where $\overline{O}_i = \max_{j=1,\dots,p} O_{ij} \quad \forall i$ and $\underline{O}_i = \min_{j=1,\dots,p} O_{ij} \quad \forall i$, i.e., the respective maximum and minimum Output for each alternative, i. The example that is shown in Figure 7 of (Finne, 1998, p.401) contains some errors. Here is the *correct table* including the Hurwicz-values (we assume a $\alpha = .5$):

O_{ij}	$min(\theta_1)$	$max(\theta_2)$	H_i
a_1	36	110	73
a_2	40	100	70
a_3	58	74	66
a_4	61	66	63.5

Thus, the order of preference is $a_4 \succ a_3 \succ a_1 \succ a_2$.

2.2.3 Decision making under risk

When some information is given about the probability of occurrence of states of nature, we speak of *decision-making under risk*. The most straight forward technique to make a decision here is to maximize the expected outcome for each alternative given the probability of occurrence, p_j.

However, the expected utility hypothesis states that the subjective value associated with an individual's gamble is the statistical expectation of that individual's valuations of the outcomes of that gamble, where these valuations may differ from the Euro value of those outcomes. Thus, you should better look on the utility of a respective outcome rather than on the outcome itself because the utility and outcome do not have to be linked in a linear way. The St. Petersburg Paradox by Daniel Bernoulli in 1738 is considered the beginnings of the hypothesis.

2.2.4 St. Petersburg Paradox

Ininite St. Petersburg lotteries

Suppose a casino offers a game of chance for a single player, where a fair coin is tossed at each stage. At the beginning, the player stakes \$1. Every time a head appears, the stake is doubled. The game continues until the first tails appears, at which point the player receives $\$2^{k-1}$, where k is the number of tosses (number of heads) plus one (for the final tails). For instance, if tails appears on the first toss, the player wins \$0. If tails appears on the second toss, the player wins \$2. If tails appears on the third toss, the player wins \$4, and so on. The extensive form is given in figure 2.6.

Given the rules of the game, what would be a fair price for the player to pay the casino in order to enter the game?

To answer this question, one needs to consider the expected payout: The player has a 1/2 probability of winning \$1, a 1/4 probability of winning \$2, a 1/8 probability of winning \$4, and so on. Thus, the overall expected value can be calculated as follows:

$$E = \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 4 + \frac{1}{16} \cdot 8 + \dots$$

This can be simplified as:

$$E = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots = +\infty.$$

That means the expected win for playing this game is an infinite amount of money. Based on the expected value, a risk-neutral individual should be willing to play the game at any price if given the opportunity. The willingness to pay of most people who have given the opportunity to play the game deviates dramatically from the objectively calculable expected payout of the lottery. This describes the apparent paradox.

In the context of the St. Petersburg Paradox, it becomes evident that relying solely on expected values is inadequate for certain games and for making well-informed decisions. Expected utility, on the other hand, has been the prevailing concept used to reconcile actual behavior with the notion of rationality thus far.

Finite St. Petersburg lotteries

Let us assume that at the beginning, the casino and the player agrees upon how many times the coin will be tossed. So we have a finite number I of lotteries with $1 \leq I \leq \infty$.

To calculate the expected value of the game, the probability $p(i)$ of throwing any number i of consecutive head is crucial. This probability is given by

$$p(i) = \underbrace{\frac{1}{2} \cdot \frac{1}{2} \cdot \dots \cdot \frac{1}{2}}_{i \text{ Faktoren}} = \frac{1}{2^i}$$

The payoff $W(I)$ is, if head appears I-times in a row by

$$W(I) = 2^{I-1}$$

The expected payoff $E(W(I))$ if the coin is flipped I times is then given by

$$E(W(I)) = \sum_{i=1}^{I} p(i) \cdot W(i) = \sum_{i=1}^{I} \frac{1}{2^i} \cdot 2^{i-1} = \sum_{i=1}^{I} \frac{1}{2} = \frac{I}{2}$$

Thus, the expected payoff grows proportionally with the maximum number of rolls. This is because at any point in the game, the option to keep playing has a positive value no matter how many times

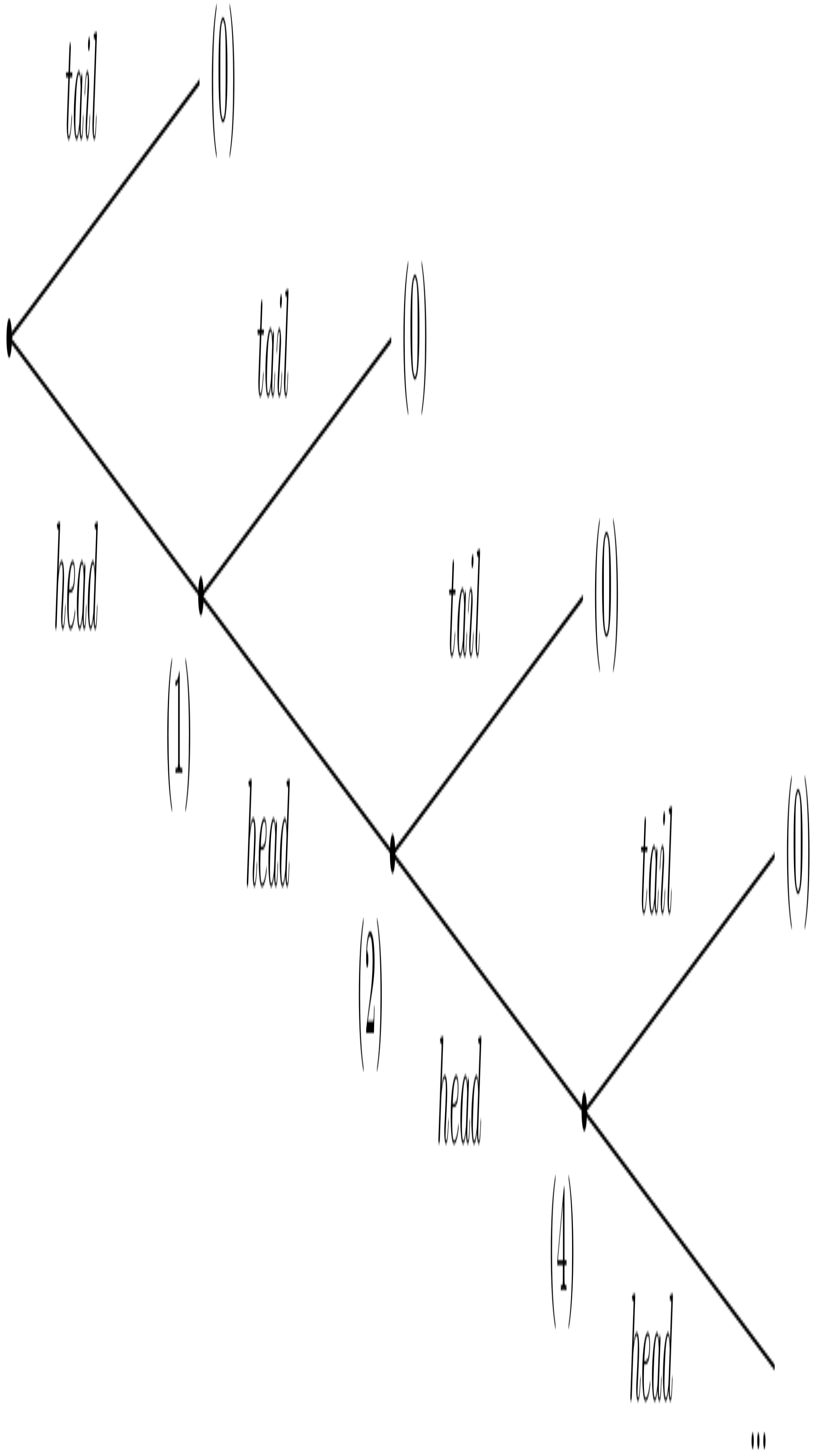


Figure 2.6: Extensive form of the St. Petersburg paradox

head has appeared before. Thus, the expected value of the game is infinitely high for an unlimited number of tosses but not so for a limited number of tosses. Even with a very limited maximum number of tosses of, for example, $I = 100$, only a few players would be willing to pay \$50 for participation. The relatively high probability to leave the game with no or very low winnings leads in general to a subjective rather low evaluation that is below the expected value.

In the real world, we understand that money is limited and the casino offering this game also operates within a limited budget. Let's assume, for example, that the casino's maximum budget is \$20,000,000. As a result, the game must conclude after 25 coin tosses because $2^{25} = 33,554,432$ would exceed the casino's financial capacity. Consequently, the expected value of the game in this scenario would be significantly reduced to just \$12.50. Interestingly, if you were to ask people, most would still be willing to pay less than \$12.50 to participate. How can we explain this? Well, it is not the expected outcome that matters but the utility that stems from the outcome.

The impact of output on utility matters

Daniel Bernoulli (1700 - 1782) worked on the paradox while being a professor in St. Petersburg. His solution builds on the conceptual separation of the expected payoff and its utility. He describes the basis of the paradox as follows:

> "Until now scientists have usually rested their hypothesis on the assumption that all gains must be evaluated exclusively in terms of themselves, i.e., on the basis of their intrinsic qualities, and that these gains will always produce a utility directly proportionate to the gain.'' (Bernoulli, 1954, p. 27)

The relationship between gain and utility, however, is not simply directly proportional but rather more complex. Therefore, it is important to evaluate the game based on expected utility rather than just the expected payoff.

$$E(u(W(I))) = \sum_{i=1}^{I} p(i) \cdot u(W(i)) = \sum_{i=1}^{I} \frac{1}{2^i} \cdot u\left(2^{i-1}\right)$$

Daniel Bernoulli himself proposed the following logarithmic utility function:

$$u(W) = a \cdot \ln(W),$$

where a is a positive constant. Using this function in the expected utility, we get

$$E(u(W(I))) = \sum_{i=1}^{I} \frac{1}{2^i} \cdot a \cdot \ln\left(2^{i-1}\right) = a \cdot \sum_{i=1}^{I} \frac{i-1}{2^i} \ln 2 = a \cdot \ln 2 \cdot \sum_{i=1}^{I} \frac{i-1}{2^i}.$$

The infinite series, $\sum_{i=1}^{I} \frac{i-1}{2^i}$, converges to 1 ($\lim_{I\to\infty} \sum_{i=1}^{I} \frac{i-1}{2^i} = 1$). Thus, given an ex ante unbounded number of throws, the expected utility of the game is given by

$$E(u(W(\infty))) = a \cdot \ln 2.$$

In experiments in which people were offered this game, their willingness to pay was roughly between 2 and 3 Euro. Thus, the suggests logarithmic utility function seems to be a pretty realistic specification. The main reason is mathematically that the increasing expected payoff has decreasing marginal utility and hence the utility function reflects the risk aversion of many people.

Exercise 2.5 — Rationality and risk (Solution → p. ??)

There are 90 balls in an box. It is known that 30 of them are red, the remaining 60 are blue or green. An individual can choose between the following lotteries:

	Payoff	Probability
Lottery 1	100 Euro if a red ball is drawn 0 Euro else	$p = \frac{1}{3}$
Lottery 2	100 Euro if a blue ball is drawn 0 Euro else	$0 \leq p \leq \frac{2}{3}$

()

In a second variant it has the choice between the following lotteries:

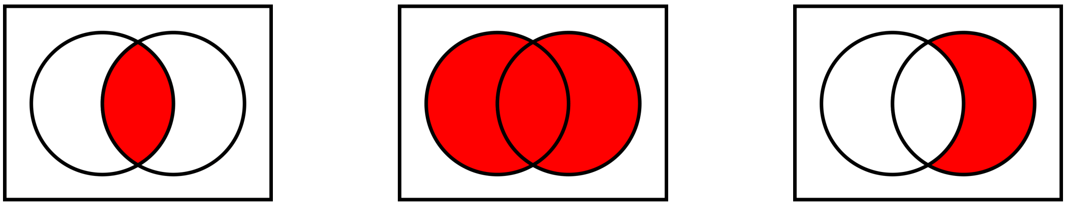
	Payoff	Probability
Lottery 3	100 Euro if a red or green ball is drawn or 0 Euro else	$\frac{1}{3} \leq p \leq 1$
Lottery 4	100 Euro if a blue or green ball is drawn or 0 Euro else	$p = \frac{2}{3}$

a) Which of the lotteries does the individual choose on the basis of expected values (risk neutral)?

b) Which of the lotteries does the individual choose on the basis of expected utility if the utility of a payoff of x is given by $u(x) = x^2$?

c) Empirical studies, e.g. Camerer and Weber (1992), show, however, that most individuals will usually choose lotteries 1 and 4. will. Discuss: Is this consistent with rational behavior?

2.3 Bayes' theorem



(a) Intersection: $A \cap B$ (b) Union: $A \cup B$ (c) Relative complement: $\neg A \cup B$

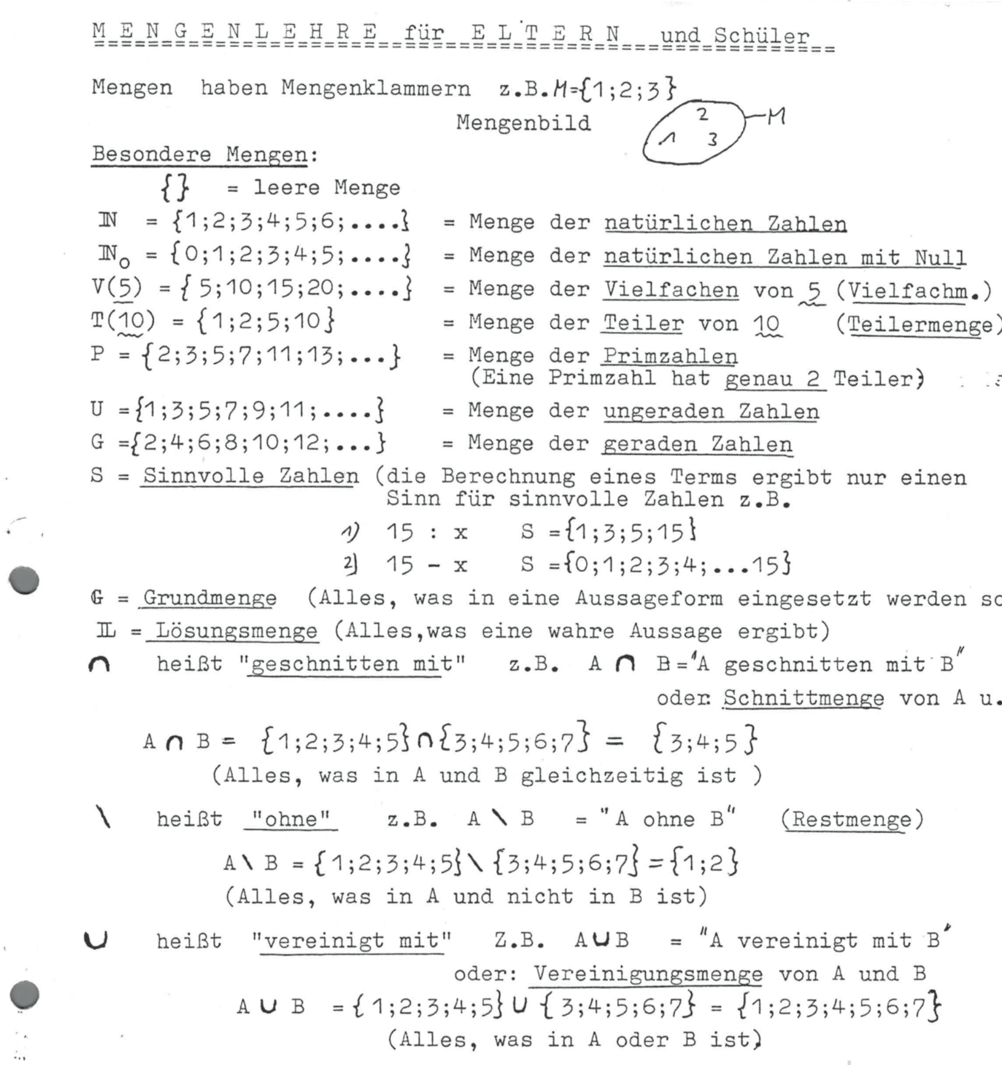
Figure 2.7: Basic Operations in Venn Diagrams

Learning objectives

- Understand and use the terminology of probability.
- Determine whether two events are mutually exclusive and whether two events are independent.
- Calculate probabilities using the Addition Rules and Multiplication Rules.
- Calculate with conditional probabilities using Bayes Theorem.
- Construct and interpret Venn Diagrams.
- Construct and interpret Tree Diagrams.

Exercise 2.6 — School Stuff (Solution → p. ??)

The next chapter will deal with stochastics. In Germany, stochastics is taught in the Gymnasium. If you are not from Germany, it was probably also part of your school experience. When I moved to Cologne in 2020, I found the following page that I had received from my high school math teacher. It was September 1993 and I was a desperate fifth grader in my fourth week. Perhaps you'd like to share your experiences with stochastics? In particular, please let me know if you are familiar with the content of Figure 2.7.



MENGENLEHRE für ELTERN und Schüler

Mengen haben Mengenklammern z.B. $M=\{1;2;3\}$

Mengenbild

Besondere Mengen:

$\{\}$ = leere Menge

$\mathbb{N} = \{1;2;3;4;5;6;....\}$ = Menge der natürlichen Zahlen

$\mathbb{N}_0 = \{0;1;2;3;4;5;....\}$ = Menge der natürlichen Zahlen mit Null

$V(5) = \{5;10;15;20;....\}$ = Menge der Vielfachen von 5 (Vielfachm.)

$T(10) = \{1;2;5;10\}$ = Menge der Teiler von 10 (Teilermenge)

$P = \{2;3;5;7;11;13;...\}$ = Menge der Primzahlen (Eine Primzahl hat genau 2 Teiler)

$U = \{1;3;5;7;9;11;....\}$ = Menge der ungeraden Zahlen

$G = \{2;4;6;8;10;12;...\}$ = Menge der geraden Zahlen

S = Sinnvolle Zahlen (die Berechnung eines Terms ergibt nur einen Sinn für sinnvolle Zahlen z.B.

1) $15 : x$ $S = \{1;3;5;15\}$

2) $15 - x$ $S = \{0;1;2;3;4;...15\}$

G = Grundmenge (Alles, was in eine Aussageform eingesetzt werden so

$\mathbb{L}$ = Lösungsmenge (Alles,was eine wahre Aussage ergibt)

$\cap$ heißt "geschnitten mit" z.B. $A \cap B$ = "A geschnitten mit B" oder Schnittmenge von A u.

$A \cap B = \{1;2;3;4;5\} \cap \{3;4;5;6;7\} = \{3;4;5\}$

(Alles, was in A und B gleichzeitig ist)

$\setminus$ heißt "ohne" z.B. $A \setminus B$ = "A ohne B" (Restmenge)

$A \setminus B = \{1;2;3;4;5\} \setminus \{3;4;5;6;7\} = \{1;2\}$

(Alles, was in A und nicht in B ist)

$\cup$ heißt "vereinigt mit" Z.B. $A \cup B$ = "A vereinigt mit B" oder: Vereinigungsmenge von A und B

$A \cup B = \{1;2;3;4;5\} \cup \{3;4;5;6;7\} = \{1;2;3;4;5;6;7\}$

(Alles, was in A oder B ist)

2.3.1 Terminology: $P(A)$, $P(A|B)$, Ω, $\cap$, $\neg$, ...

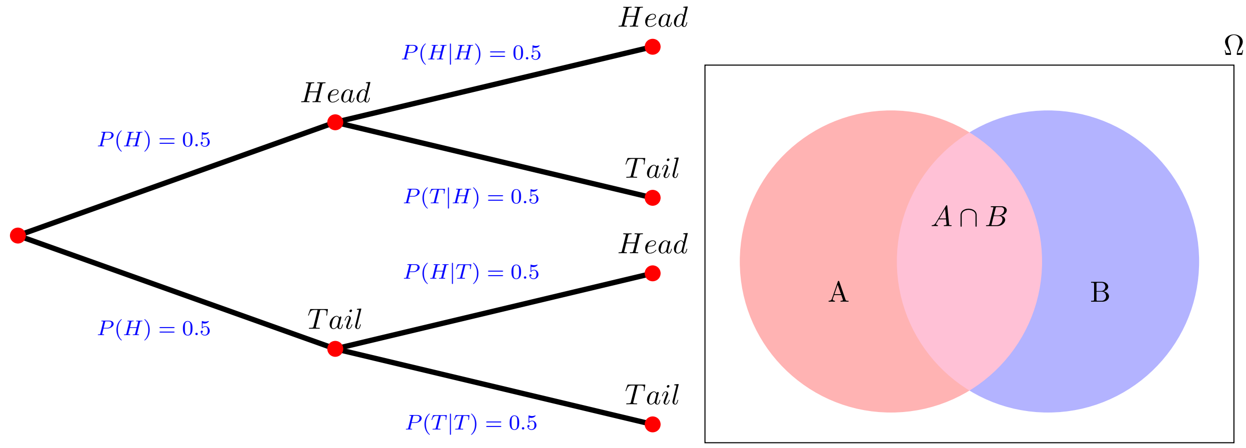
Sample space

A result of an **experiment** is called an **outcome**. An experiment is a planned operation carried out under controlled conditions. Flipping a fair coin twice is an example of an experiment. The **sample space** of an experiment is the set of all possible outcomes. The Greek letter Ω is often used to denote the sample space. For example, if you flip a fair coin, $\Omega = \{H, T\}$ where the outcomes heads and tails are denoted with H and T, respectively.

Exercise 2.7 — Sample space (Solution → p. 30)

Find the sample space for the following experiments:

a) One coin is tossed.
b) Two coins are tossed once.
c) Two dices are tossed once.
d) Picking two marbles, one at a time, from a bag that contains many blue, B, and red marbles, R.



(a) Tree diagramm (b) Venn diagramm

Figure 2.8: Diagrams to represent a sample space

Solution to Exercise 2.7 — Sample space (Exercise → p. 29)

a) $\Omega = \{head, tail\}$
b) $\Omega = \{(head, head), (tail, tail), (head, tail), (tail, head)\}$
c) Overall, 36 different outcomes: $\{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), \ldots, (6,6)\}$
d) $\Omega = \{(B,B), (B,R), (R,B), (R,R)\}$.

Overall, there are three ways to represent a sample space:

1. to list the possible outcomes (see Exercise 2.7),
2. to create a tree diagram (see Figure 2.8a), or
3. to create a Venn diagram (see Figure 2.8b).

Probability

Probability is a measure that is associated with how certain we are of outcomes of a particular experiment or activity. The probability of an event A, written $P(A)$, is defined as

$$P(A) = \frac{\text{Number of outcomes favorable to the occurrence of } A}{\text{Total number of equally likely outcomes}} = \frac{n(A)}{n(\Omega)}$$

For example, A dice has 6 sides with 6 different numbers on it. In particular, the set of **elements** of a dice is $M = \{1, 2, 3, 4, 5, 6\}$ or $M = \{$⚀, ⚁, ⚂, ⚃, ⚄, ⚅$\}$. Thus, the probability to receive a ⚅ is $1/6$ because we look for one wanted outcome in six possible outcomes.

Exercise 2.8 — Probability (Solution → p. 30)

When a fair dice is thrown, what is the probability of getting
a) the number 5,
b) a number that is a multiple of 3,
c) a number that is greater than 6,
d) a positive number that is less than 7.

Solution to Exercise 2.8 — Probability (Exercise → p. 30)

A fair dice is an unbiased dice where each of the six numbers is equally likely to turn up. The sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$.

a) Let A be the event of getting the number 5, $A = \{5\}$. Then, $P(A) = \frac{1}{6}$.
b) Let B be the event of getting a multiple of 3, $B = \{3, 6\}$. Then, $P(B) = \frac{1}{3}$.
c) Let C be the event of getting a number greater than 6, $C = 7, 8, \ldots$. Then, $P(C) = 0$ as there is no number greater than 6 in the sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$. A probability of 0 means the event will never occur.
d) Let D be the event of getting a number less than 7, $D = \{1, 2, 3, 4, 5, 6\}$. Then, $P = 1$ as the event will always occur.

The complement of an event ($\neg - \textit{Event}$)

The complement of event A is denoted with a $\neg A$ or sometimes with a superscript 'c' like A^c. It consists of all outcomes that are *not* in A. Thus, it should be clear that $P(A) + P(\neg A) \overset{!}{=} 1$. For example, let the sample space be $\Omega = \{1, 2, 3, 4, 5, 6\}$ and let $A = \{1, 2, 3, 4\}$. Then, $\neg A = \{5, 6\}$; $P(A) = \frac{4}{6}$; $P(\neg A) = \frac{2}{6}$; and $P(A) + P(\neg A) = \frac{4}{6} + \frac{2}{6} = 1$.

Independent events (AND-events)

Two events are independent when the outcome of the first event does not influence the outcome of the second event. For example, if you throw a dice and a coin, the number on the dice does not affect whether the result you get on the coin. More formally, two events are independent if the following are true:

$$\begin{aligned} P(A|B) &= P(A) \\ P(B|A) &= P(B) \\ P(A \cap B) &= P(A)P(B) \end{aligned}$$

To calculate the probability of two independent events (X and Y) happen, the probability of the first event, $P(X)$, has to be multiplied with the probability of the second event, $P(Y)$:

$$P(X \text{ and } Y) = P(X \cap Y) = P(X) \cdot P(Y),$$

where $\cap$ stands for 'and'. For example, let A and B be $\{1, 2, 3, 4, 5\}$ and $\{4, 5, 6, 7, 8\}$, respectively. Then $A \cap B = \{4, 5\}$.

Exercise 2.9 — Three dices (Solution → p. 31)

Suppose you have three dice. Calculate the probability of getting three ⚄.

Solution to Exercise 2.9 — Three dices (Exercise → p. 31)

The probability of getting a ⚄ on one dice is 1/6. The probability of getting three ⚄ is:

$$P(\text{⚄} \cap \text{⚄} \cap \text{⚄}) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{216}$$

Dependent events (|-Events)

Events are dependent when one event affects the outcome of the other. If A and B are dependent events then the probability of both occurring is the product of the probability of A and the probability of A *after* B has occurred:

$$\begin{aligned} & P(A \cap B) = P(A) \cdot P(B|A), \\ \Leftrightarrow\ & P(B|A) = \frac{P(A \cap B)}{P(A)}. \end{aligned}$$

where '$|A$' stands for 'after A has occurred', or 'given A has occurred'. In other words, $P(B|A)$ is the probability of B given A. This equation is also known as the **Multiplication Rule**.

The conditional probability of A given B is written $P(A|B)$. $P(A|B)$ is the probability that event A will occur given that the event B has already occurred. A conditional reduces the sample space. We calculate the probability of A from the reduced sample space B. The formula to calculate $P(A|B)$ is

$$P(A|B) = \frac{P\left(A \cap B\right)}{P\left(B\right)}$$

where $P(B)$ is greater than zero. This formula is also known as **Bayes' Theorem**, which is a simple mathematical formula used for calculating conditional probabilities, states that

$$P(A)P(B|A) = P(B)P(A|B)$$

This is true since $P(A \cap B) = P(B \cap A)$ and due to the fact that $P(A \cap B) = P(B \mid A)P(A)$, we can write Bayes' Theorem as

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}.$$

The box below summarizes the important facts w.r.t. Bayes' Theorem.

For example, suppose we toss a fair, six-sided die. The sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$. Let A be face is 2 or 3 and let B be face is even (2, 4, 6). To calculate $P(A|B)$, we count the number of outcomes 2 or 3 in the sample space $B = \{2, 4, 6\}$. Then we divide that by the number of outcomes B (rather than Ω).

We get the same result by using the formula. Remember that Ω has six outcomes.

$$P(A \mid B) = \frac{P(B \cap A)}{P(B)} = \frac{\frac{\text{number of outcomes that are 2 or 3 and even}}{6}}{\frac{\text{number of outcomes that are even}}{6}} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}$$

Bayes' Theorem
The theorem states that

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

if $P(B) \neq 0$ and A and B are events. It simply uses the following logical facts:

$$P(B \cap A) = P(A \cap B),$$

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \text{ and}$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)},$$

or, to put it in one line:

$$P(A \cap B) = P(B \cap A) = P(A \mid B)P(B) = P(B \mid A)P(A).$$

Sometimes, it is helpful to re-write the Theorem as follows:

$$P(A) = P(A|B)P(B) + P(A|\neg B)P(\neg B), \text{ and}$$

$$P(B) = P(B|A)P(A) + P(B|\neg A)P(\neg A),$$

Exercise 2.10 — Purse (Solution → p. 33)

A purse contains four €5 bills, five €10 bills and three €20 bills. Two bills are selected randomly without the first selection being replaced. Find the probability that two €5 bills are selected.

Solution to Exercise 2.10 — Purse (Exercise → p. 33)

There are four €5 bills. There are a total of twelve bills. The probability to select at first a €5 bill then is $P(€5) = \frac{4}{12}$*. As the the result of the first draw affects the probability of the second draw, we have to consider that there are only three €5 bills left and there are a total of eleven bills left. Thus,*

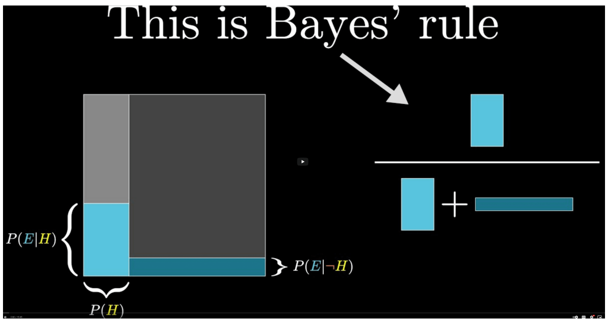
$$P(€5|€5) = \frac{3}{11}$$

and

$$P(€5 \cap €5) = P(€5) \cdot P(€5|€5) = \frac{4}{12} \cdot \frac{3}{11} = \frac{1}{11}.$$

The probability of drawing a €5 bill and then another €5 bill is $\frac{1}{11}$.

2.3.2 Bayes' theorem and the case of false positive



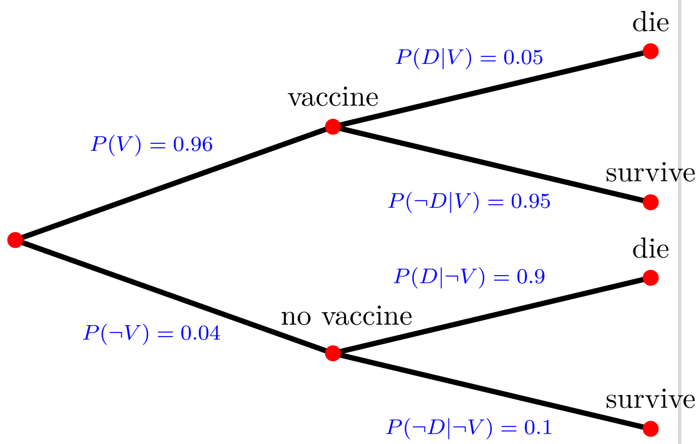
Watch the two videos linked here:

- *Bayes theorem* https://youtu.be/HZGCoVF3YvM and
- *The quick proof of Bayes' theorem* https://youtu.be/U_85TaXbeIo

Also, consider this interactive tool: https://www.skobelevs.ie/BayesTheorem/

Exercise 2.11 — To be vaccinated or not to be (Solution → p. 34)

The *Tree Diargram* shows probabilities of people to have a vaccine for some disease. Moreover, it shows the conditional probabilities of people to die given the fact they were vaccinated or not. D denotes the event of *die* and $\neg D$ denotes *not die*, i.e., survive; V denotes the event of *vaccinated* and $\neg V$ *not vaccinated.*



a) Calculate the overall probability to die, $P(D)$
b) Calculate the probability that a person that has died was vaccinated, $P(V|D)$.

Disclaimer: The case presented here is fictitious. The data given here are purely fictitious and serve only to practice the method.

Solution to Exercise 2.11 — To be vaccinated or not to be (Exercise → p. 33)

a)

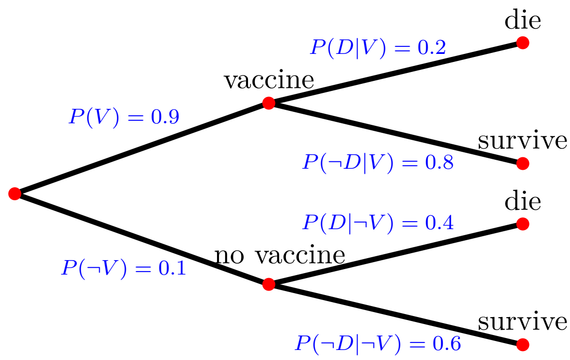
$$\begin{aligned} P(D) &= P(V) \cdot P(D|V) + (P(\neg V) \cdot P(D|\neg V)) \\ &= .96 \cdot .05 + .04 \cdot .09 \\ &= 0.048 + 0.036 \\ &= 0.084 \end{aligned}$$

b)

$$P(V \mid D) = \frac{P(D \mid V)P(V)}{P(D)} = \frac{.05 \cdot .96}{.084} \approx .5714285$$

Exercise 2.12 — To die or not to die (Solution → p. 34)

You read on *Facebook* that in the year 2021 about over 80% of people that died were vaccinated. You are shocked by this high probability that a dead person was vaccinated, $P(V|D)$. You decide to check this fact. Reading the study to which the Facebook post is referring, you find out that the study only refers to people above the age of 90. Moreover, you find the following *Tree Diargram*. It allows checking the fact as it describes the vaccination rates and the conditional probabilities of people to die given the fact they were vaccinated or not. In particular, D denotes the event of *die* and $\neg D$ denotes *not die*, i.e., survive; V denotes the event of *vaccinated* and $\neg V$ *not vaccinated*.



a) Calculate the overall probability to die, $P(D)$
b) Calculate the probability that a person that has died was vaccinated, $P(V|D)$.
c) Your calculations shows that the fact used in the statement on *Facebook* is indeed true. Discuss whether this number should have an impact to get vaccinated or not.

Disclaimer: The case presented here is fictitious. The data given here are purely fictitious and serve only to practice the method.

Solution to Exercise 2.12 — To die or not to die (Exercise → p. 34)

a)

$$P(D) = 0.9 \cdot 0.2 + 0.1 \cdot 0.4 = 0.22$$

b)

$$P(V \mid D) = \frac{P(D \mid V)P(V)}{P(D)} = \frac{0.2 \cdot 0.9}{0.22} = \frac{0.036}{0.22} \approx 0.8181$$

c)

Exercise 2.13 — Corona False Positive (Solution → p. 35)

Suppose that Corona infects one out of every 1000 people in a population and that the test for it comes back positive in 99% of all cases if a person has Corona. Moreover, the test also produces some false positive, that is about 2% of uninfected patients also tested positive.
Now, assume you are tested positive and you want to know the chances of having the disease. Then, we have two events to work with:
A: you have Corona
B: your test indicates that you have Corona
and we know that

$$\begin{aligned} P(A) &= .001 && \rightarrow \text{one out of 1000 has Corona} \\ P(B|A) &= .99 && \rightarrow \text{probability of a positive test, given infection} \\ P(B|\neg A) &= .02 && \rightarrow \text{probability of a false positive, given no infection} \end{aligned}$$

As you don't like to go into quarantine, you are interested in the probability of having the disease given a positive test, that is $P(A|B)$?

Disclaimer: The case presented here is fictitious. The data given here are purely fictitious and serve only to practice the method.

Solution to Exercise 2.13 — Corona False Positive (Exercise → p. 35)

In order to come to an answer, lets draw a table of the probabilities that may be of interest:

	A	$\neg A$	*sum*
B	$P(A \cap B)$	$P(\neg A \cap B)$	$P(B)$
$\neg B$	$P(A \cap \neg B)$	$P(\neg A \cap \neg B)$	$P(\neg B)$
	$P(A)$	$P(\neg A)$	*1*

Please note, the symbol "¬" simply abbreviates "NOT" and the symbol "∩" stands for "AND". The probability of you having both the disease and a positive test, $P(A \cap B)$, is easy to calculate:

$$\underbrace{P(A \cap B) = P(B|A)P(A)}_{\text{a.k.a. multiplication rule}} = .99 \cdot .001 = .00099$$

Also, it is straight forward to calculate the probability of having both, no infection and a positive test:

$$P(\neg A \cap B) = P(B|\neg A)P(\neg A) = .02 \cdot .999 = .01998$$

Knowing that, it is clear that the overall probability of being diagnosed with Corona is

$$P(B) = .00099 + .0199 = .02097$$

That means, out of 1000 people about 21 are on average diagnosed with Corona while only one person actually is infected with Corona. Thus, your probability of having Corona once your test came out to be positive, $P(A|B)$, is approximately

$$P(A|B) \approx \frac{1}{21} = 0,047619048$$

and more precisely

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{.00099}{.02097} = 0,0472103.$$

In other words, with a probability of more than 95%, you may go into quarantine without infection:

$$P(\neg A|B) = \frac{P(\neg A \cap B)}{P(B)} = \frac{.01998}{.02097} = 0,9527897 \quad (= 1 - P(A|B))$$

*Given the accuracy of the test, this number appears to be rather high to many people. The high test accuracy of 99% and the rather low number of 2% false positives, however, is misleading. This is sometimes called the **false positive paradox**. The source of the fact that many people think $P(A|B)$ is much lower is that they don't consider the impact of the low probability of having the disease, $P(A)$, on $P(A|B)$, $P(B|A)$, and $P(B|\neg A)$ respectively (also, see the Linda case). Moreover, many people don't understand the false positive rate correctly.*
To summarize, we know

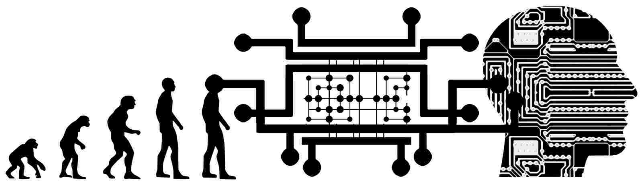
	A	$\neg A$	$\sum$
B	*.00099*	*.01998*	*.02097*
$\neg B$	$P(A \cap \neg B)$	$P(\neg A \cap \neg B)$	$P(\neg B)$
	.001	$P(\neg A)$	*1*

The four remaining unknowns can be calculated by subtracting in the columns and adding across the crows, so that the final table is:

	A	$\neg A$	$\sum$
B	*.00099*	*.01998*	*.02097*
$\neg B$	*.00001*	*.97902*	*.97903*
	.001	*.999*	*1*

Chapter 3

Automated decision making



3.1 Data

Data are characteristics or information, usually numerical, that are collected through observation. In a more technical sense, data is a set of values of qualitative or quantitative variables about one or more persons or objects [. . .]

Although the terms 'data' and 'information' are often used interchangeably, these terms have distinct meanings. In some popular publications, data is sometimes said to be transformed into information when it is viewed in context or in post-analysis. In academic treatments of the subject, however, data are simply units of information. [. . .]

Data is measured, collected and reported, and analyzed, whereupon it can be visualized using graphs, images or other analysis tools. Data as a general concept refers to the fact that some existing information or knowledge is represented or coded in some form suitable for better usage or processing. Raw data ('unprocessed data') is a collection of numbers or characters before it has been 'cleaned' and corrected by researchers. Raw data needs to be corrected to remove outliers or obvious instrument or data entry errors [. . .]'. Data processing commonly occurs by stages, and the 'processed data' from one stage may be considered the 'raw data' of the next stage. Field data is raw data that is collected in an uncontrolled 'in situ' environment. Experimental data is data that is generated within the context of a scientific investigation by observation and recording. **(Wikipedia)**

3.2 Data science

Data science is an inter-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from many structural and unstructured data. Data science is related to data mining and big data.

Data science is a "concept to unify statistics, data analysis, machine learning and their related methods" in order to "understand and analyze actual phenomena" with data. It employs techniques and theories drawn from many fields within the context of mathematics, statistics, computer science, and information science.

3.3 Types of data

Structured: Data is stored, processed, and manipulated in a traditional relational database management system (e.g., temperature)
Un-structured: Data that is commonly generated from human activities and does not fit into a structured database format (e.g., text)
Semi-structured: Data does not fit into a structured database system, but is nonetheless structured by tags that are useful for creating a form of order and hierarchy in the data
Big data *see below*
Dark data *see below*

Structured vs. semi-structured data Business Intelligency requires analysts to deal with both structured and semi-structured data. The term semi-structured data is used for all data that does not fit neatly into relational or flat files, which is called structured data. We use the term semi-structured (rather than the more common unstructured) to recognize that most data has some structure to it. A survey indicated that 60% of CIOs and CTOs consider semi-structured data as critical for improving operations and creating new business opportunities (Blumberg and Atre, 2003).

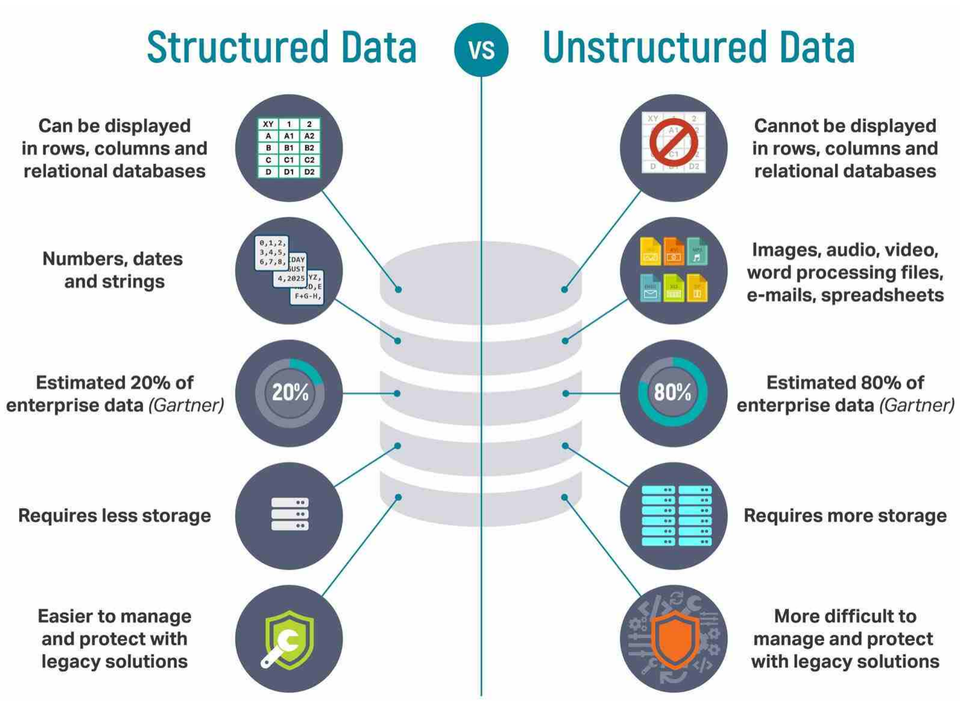


Figure 3.1: Structured vs. semi-structured data

Executive at Fortune 500 telecommunciations provider

"We have between 50,000 and 100,000 conversations with our customers daily, and I don't know what was discussed. I can see only the end point – for example, they changed their calling plan. I'm blind to the content of the conversations." (see Blumberg and Atre, 2003)

Semi-structured data: Examples

For example, e-mail is divided into messages and messages are accumulated into file folders.

Business processes, Chats, E-mails, Graphics, Image files, Letters, Marketing material, Memos, Movies, News items, Phone, conversations, Presentations, Reports, Research, Spreadsheet files, User group files, Video files, Web pages, White papers, Word processing text

Exercise 3.1 — Data Mining (Solution → p. ??)

- Watch: `https://youtu.be/EH3bp5335IU`
- Read the Wikipedia page of 'Data Mining'

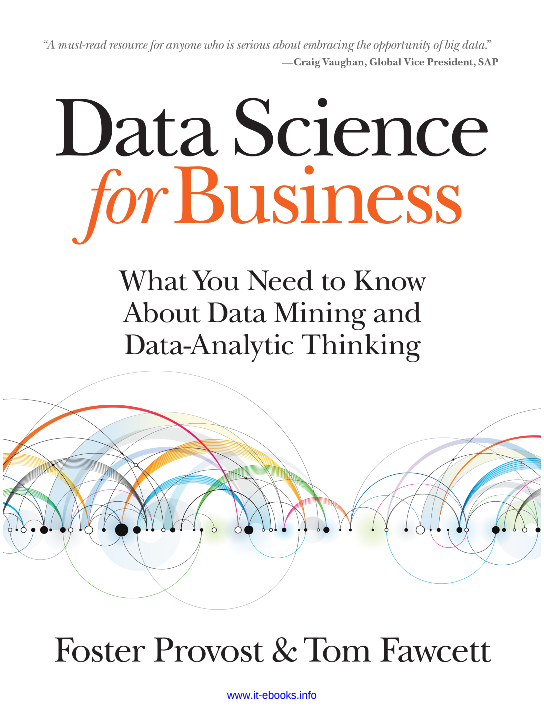


Figure 3.2: Data Science for Business by Provost and Fawcett (2013)

A data scientist view on Provost and Fawcett (2013, ch.1+2)

In the following, you see a wordcloud where the size of the word represents its frequency in the text, e.g., data was counted about 600 times, mining about 140 times and science about 100 times. It is a nice tool to analyse unstructured data such as the text of two chapters. Unstructured data does not have a pre-defined data model or is not organized in a pre-defined manner.

Figure 3.3: Wordcloud on Provost and Fawcett (2013, ch.1+2)
Source: own calculations

3.4 Big data

Big data is a field that treats ways to analyze, systematically extract information from, or otherwise deal with data sets that are too large or complex to be dealt with by traditional data-processing application software. [...] Big data challenges include capturing data, data storage, data analysis, search, sharing, transfer, visualization, querying, updating, information privacy and data source. Big data was originally associated with three key concepts: volume, variety, and velocity. When we handle big data, we may not sample but simply observe and track what happens. Therefore, big data often includes data with sizes that exceed the capacity of traditional software to process within an acceptable time and value.

Current usage of the term big data tends to refer to the use of predictive analytics, user behavior analytics, or certain other advanced data analytics methods that extract value from data, and seldom to a particular size of data set. [...] Analysis of data sets can find new correlations to "spot business trends, prevent diseases, combat crime and so on." Scientists, business executives, practitioners of medicine, advertising and governments alike regularly meet difficulties with large data-sets in areas including Internet searches, fintech, urban informatics, and business informatics. Scientists encounter limitations in e-Science work, including meteorology, genomics, connectomics, complex physics simulations, biology and environmental research. **(Wikipedia)**

Big data characteristics

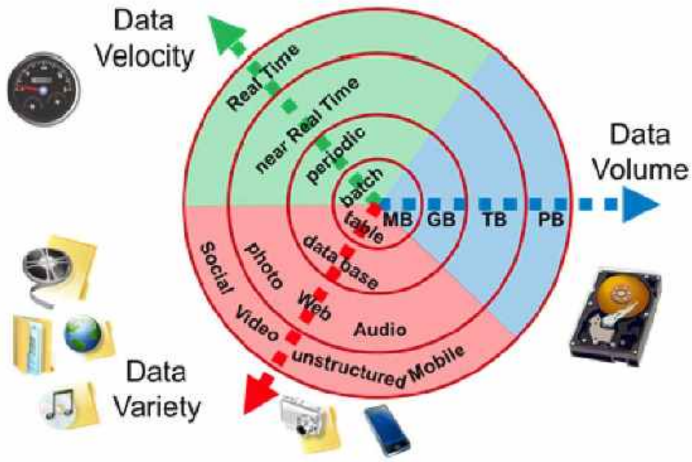


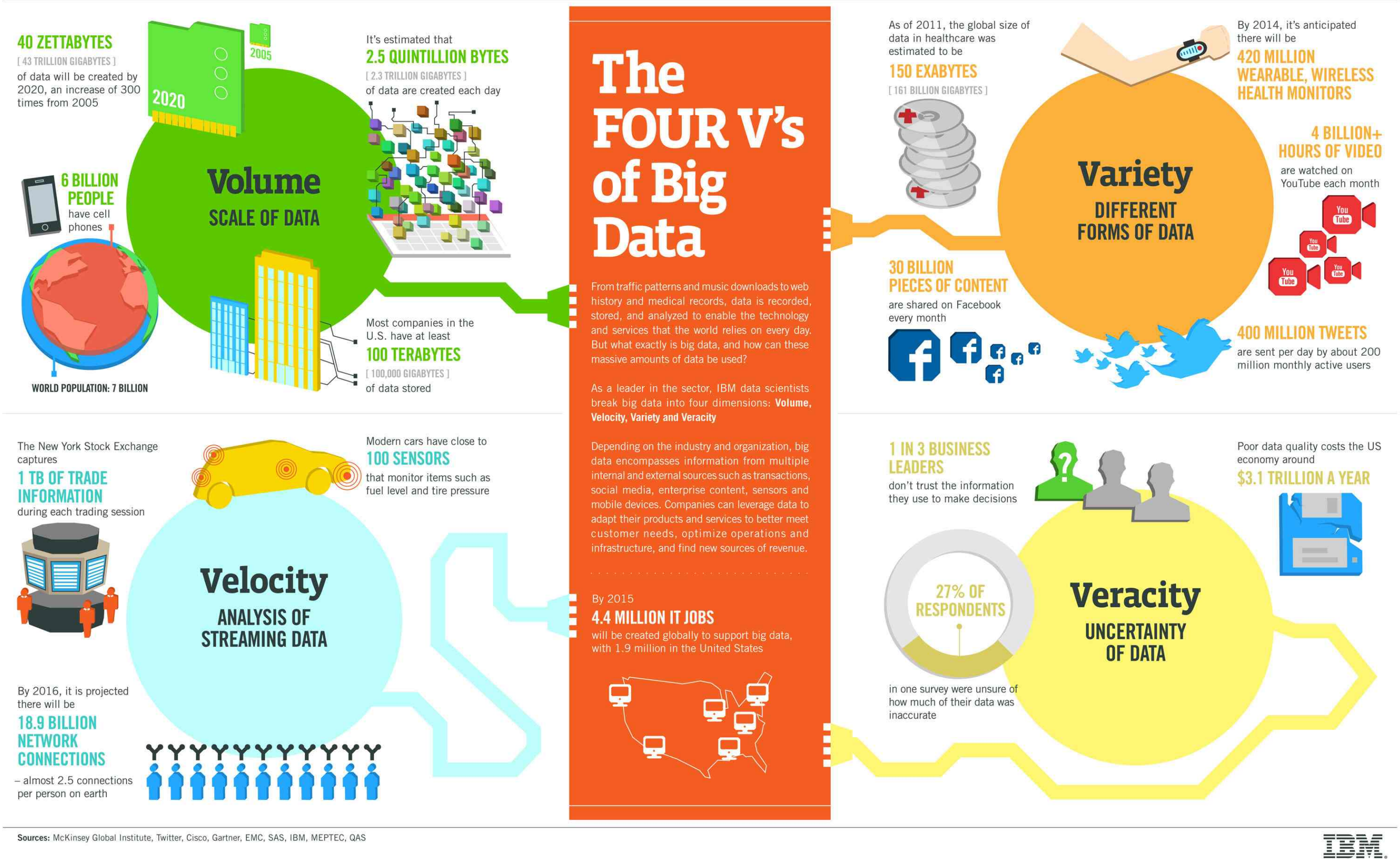
Figure 3.4: Big data characteristics

Other mentioned data characteristics are:
Veracity, Exhaustive, Fine-grained and uniquely lexical, Relational, Extensional, Scalability, Value, Variability

Volume: The amount of data matters. With big data, you'll have to process high volumes of low-density, unstructured data. This can be data of unknown value, such as Twitter data feeds, clickstreams on a webpage or a mobile app, or sensor-enabled equipment.

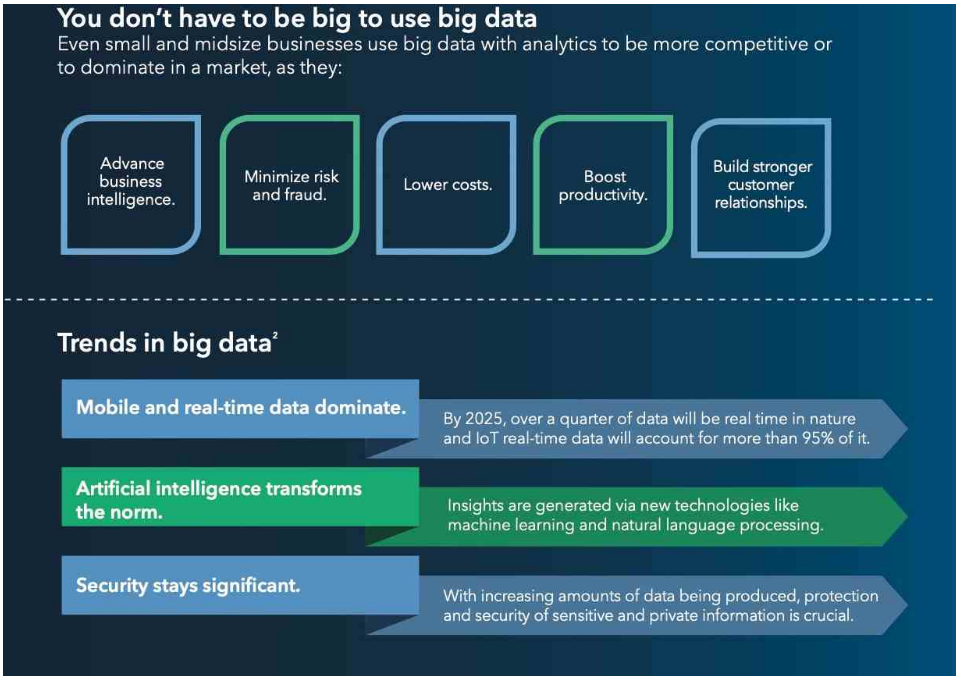
Velocity: The speed at which the data is generated and processed to meet the demands and challenges. Normally, the highest velocity of data streams directly into memory versus being written to disk. Some internet-enabled smart products operate in real time or near real time and will require real-time evaluation and action

Variety: Variety refers to the many types of data that are available. Traditional data types were structured and fit neatly in a relational database. With the rise of big data, data comes in new unstructured data types. Unstructured and semistructured data types, such as text, audio, and video, require additional preprocessing to derive meaning and support metadata



The FOUR V's of Big Data
From traffic patterns and music downloads to web history and medical records, data is recorded, stored, and analyzed to enable the technology and services that the world relies on every day. But what exactly is big data, and how can these massive amounts of data be used?
As a leader in the sector, IBM data scientists break big data into four dimensions: Volume, Velocity, Variety and Veracity
Depending on the industry and organization, big data encompasses information from multiple internal and external sources such as transactions, social media, enterprise content, sensors and mobile devices. Companies can leverage data to adapt their products and services to better meet customer needs, optimize operations and infrastructure, and find new sources of revenue.
By 2015
4.4 MILLION IT JOBS
will be created globally to support big data, with 1.9 million in the United States
Volume
SCALE OF DATA
40 ZETTABYTES
[43 TRILLION GIGABYTES]
of data will be created by 2020, an increase of 300 times from 2005
2020
2005
6 BILLION PEOPLE
have cell phones
WORLD POPULATION: 7 BILLION
It's estimated that
2.5 QUINTILLION BYTES
[2.3 TRILLION GIGABYTES]
of data are created each day
Most companies in the U.S. have at least
100 TERABYTES
[100,000 GIGABYTES]
of data stored
Velocity
ANALYSIS OF STREAMING DATA
The New York Stock Exchange captures
1 TB OF TRADE INFORMATION
during each trading session
Modern cars have close to
100 SENSORS
that monitor items such as fuel level and tire pressure
By 2016, it is projected there will be
18.9 BILLION NETWORK CONNECTIONS
– almost 2.5 connections per person on earth
Variety
DIFFERENT FORMS OF DATA
As of 2011, the global size of data in healthcare was estimated to be
150 EXABYTES
[161 BILLION GIGABYTES]
30 BILLION PIECES OF CONTENT
are shared on Facebook every month
By 2014, it's anticipated there will be
420 MILLION WEARABLE, WIRELESS HEALTH MONITORS
4 BILLION+ HOURS OF VIDEO
are watched on YouTube each month
400 MILLION TWEETS
are sent per day by about 200 million monthly active users
Veracity
UNCERTAINTY OF DATA
1 IN 3 BUSINESS LEADERS
don't trust the information they use to make decisions
27% OF RESPONDENTS
in one survey were unsure of how much of their data was inaccurate
Poor data quality costs the US economy around
$3.1 TRILLION A YEAR
Sources: McKinsey Global Institute, Twitter, Cisco, Gartner, EMC, SAS, IBM, MEPTEC, QAS
IBM

Big Data: What It Is and Why It Matters
What is big data?1
Big data sets are too large and complex to be processed by traditional methods.
Consider that in a single minute there are:
277,777 Instagram stories posted
511,200 tweets sent
4,500,000 YouTube videos watched
4,497,420 Google searches conducted
18,100,000 texts sent
188,000,000 emails sent
The 3 V's of big data – Plus 2
These are the defining properties or dimensions of big data.
Volume the amount of data
Velocity the speed of data
Variety the types of data
Variability the variety of and changes in data flows
Veracity the quality of data
How do organizations optimize the value of big data?
Regardless of location, size, sources, owners or users, these steps can unleash value from an organization's complex data landscape (data fabric).
Set a big data strategy.
Identify big data sources.
Access, manage and store big data.
Analyze big data.
Make data-driven decisions.



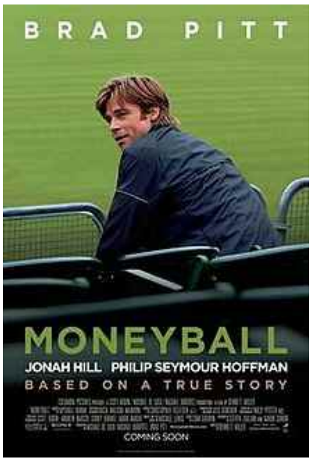
3.5 Dark data

Dark data is data that is collected through various computer network operations but is not used in any way to gain insight or make decisions. An organization's ability to collect data may exceed its capacity with which to analyze the data. In some cases, the company is not even aware that the data is being collected. Approximately 90 percent of the data generated by sensors and analog-to-digital converters is never used.

My first PC

Commodore C64, price about 750 Euro, CPU $<$ 1 MHz, RAM 64 KByte, i.e., 0,064 MB or 0,000064 GB, Graphics 320 x 200 with 16 colors, no HD but 5,25" floppy disks of 166KByte and a max of 144 files.
My main PC today has 4100Mhz at 8 cores and 16 threats with 32GB of RAM and my video card can do 3840 x 2160 at four screens, 1 NVmE of 1 TB, 2 SSD of 500GB each, 2 HD of 4TB each and 3TB in a cloud. My first external HD in 2002 had 160GB and cost 170 Euro.

3.6 Data analysis...



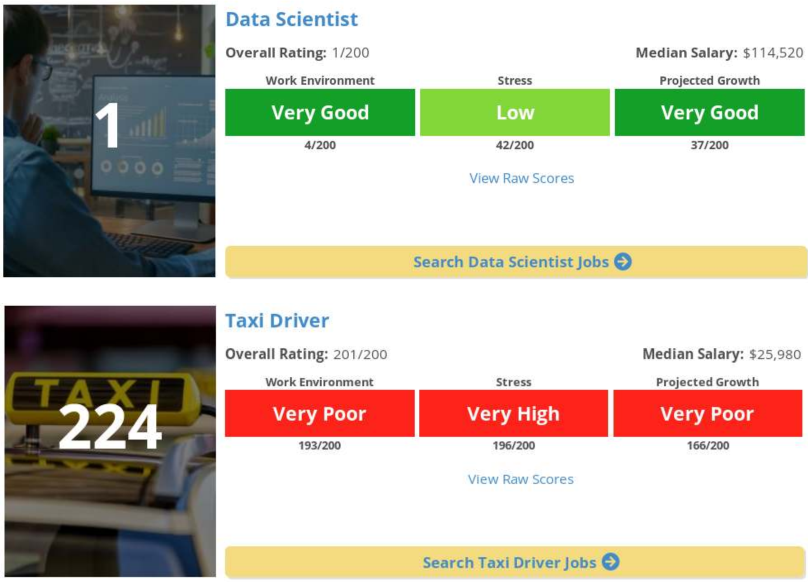
... is a process of inspecting, cleansing, transforming and modeling data with the goal of discovering useful information, informing conclusion and supporting decision-making. Data analysis has multiple facets and approaches, encompassing diverse techniques under a variety of names, and is used in different business, science, and social science domains. In today's business world, data analysis plays a role in making decisions more scientific and helping businesses operate more effectively.

- Please watch: https://youtu.be/J36ZfXBsGjs
- Sport Economics[a] is a well-accepted discipline, see: https://journals.sagepub.com/home/jse

[a] It covers both the ways in which economists can study the distinctive institutions of sports, and the ways in which sports can allow economists to research many topics, including discrimination and antitrust law.

3.7 Case study: Data analysis for NYC taxi drivers

Best and worst job: What do they have in common?

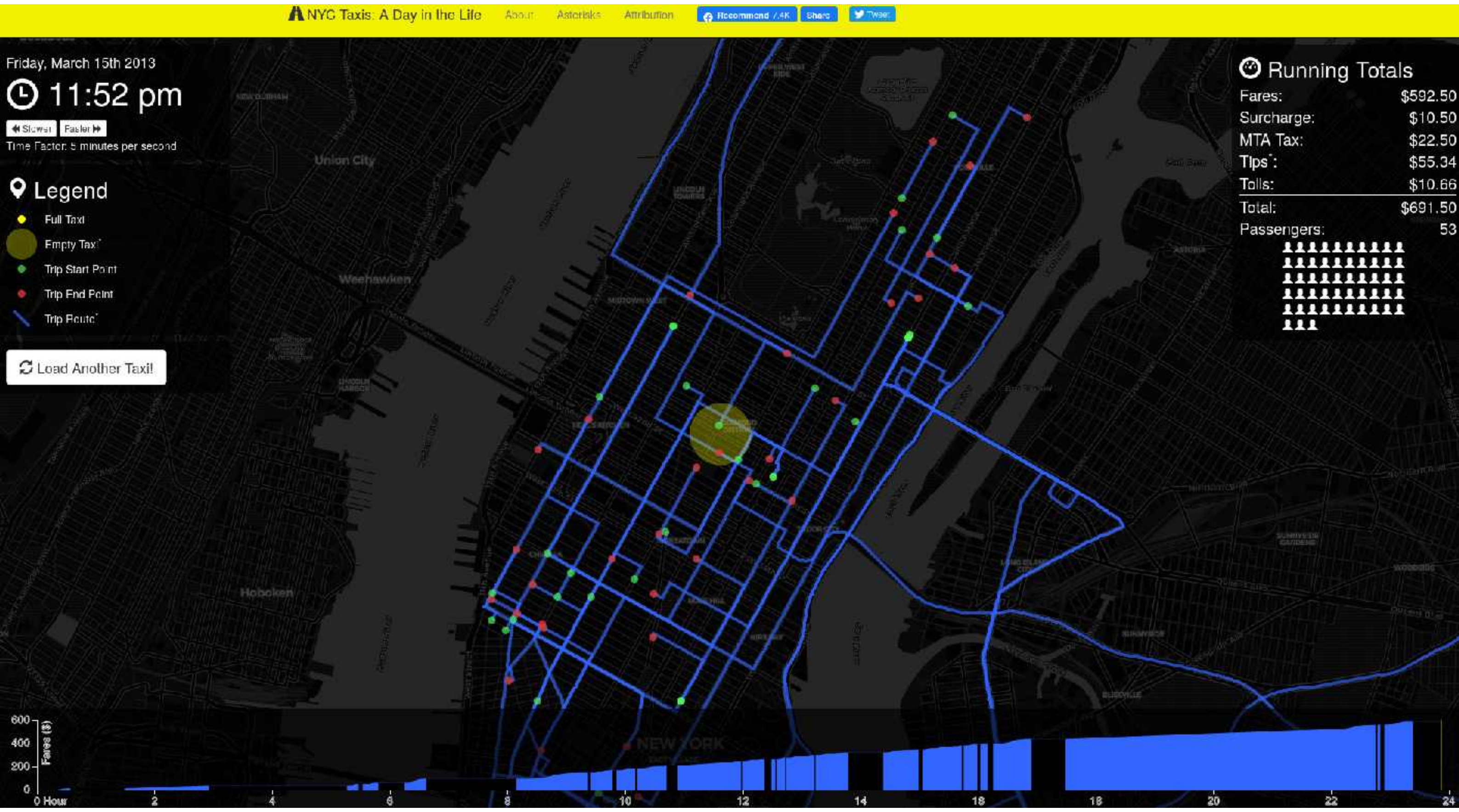


University Professor is ranked third with the best work environment and less stress, but only a median salary of $76,000.
Source: https://www.careercast.com/jobs-rated/2019-jobs-rated-report

NYC taxis: A Day in the life

Visit: http://chriswhong.github.io/nyctaxi/

This visualization displays the data for one random NYC yellow taxi on a single day in 2013. See where it operated, how much money it made, and how busy it was over 24 hours.



Discuss: What makes taxi drivers happy?

Maybe a driver wants

- more profit (fare-costs+tips),
- more tips,
- less time being 'empty' and driving around searching,
- to know fuel saving-routes,
- to work in areas with less stressful traffic (How to avoid Manhattan?),
- to start at the optimal place and time,
- to know when and where to rest without loosing much, etc.

3.8 Machine learning, AI, automated decision-making

While **machine learning** (ML) is based on the idea that machines should be able to learn and adapt through experience, artificial intelligence (AI) refers to a broader idea where machines can execute tasks *smartly*. AI applies ML techniques to solve actual problems and to automate decision making.

3.9 Machine learning. . .

. . . is the study of computer algorithms that improve automatically through experience. In particular, machine learning is a form of artificial intelligence (AI) as it provides machines and systems to automatically learn and improve from experience. Machine learning algorithms build a mathematical model based on sample data, known as 'training data', in order to make predictions or decisions without being explicitly programmed to do so.

. . . for making predictions: if you want a model to determine future trends; machine learning algorithms are the best bet. This falls under the paradigm of supervised learning. It is called supervised because you already have the data based on which you can train your machines (for example, a fraud detection model can be trained using a historical record of fraudulent purchases).

. . . for pattern discovery: If you don't have the parameters based on which you can make predictions, then you need to find out the hidden patterns within the dataset to be able to make meaningful predictions. This is nothing but the unsupervised model as you don't have any predefined labels for grouping. The most common algorithm used for pattern discovery is Clustering.

3.10 What is automated decision-making?

Automated decision-making is the process of making a decision by automated means without any human involvement. These decisions can be based on factual data, as well as on digitally created profiles or inferred data. Examples of this include:

- an online decision to award a loan; and
- an aptitude test used for recruitment which uses pre-programmed algorithms and criteria.

Automated decision-making often involves **profiling**, but it does not have to.

> Demetis and Lee (2018): "Another well-known example comes from Amazon. The vast majority of prices are defined by algorithms in so far as Amazon vendors "use algorithmic pricing to ensure that they can automatically change their product prices based on a competitor"' [39], with the result that vendors are being forced to engage in this practice for fear of losing out to the competition. Meanwhile, the algorithmic interactions between vendors carry the possibility of developing unpredictable consequences. Such algorithmic pricing on Amazon can be found in the example of the book entitled The Making of a Fly by evolutionary biologist Peter Lawrence. This book came to be priced at $23,698,655.93 (plus $3.99 shipping) as two sellers were using algorithms to adjust the price of the book in response to one another. It took 10 days for humans to notice and intervene to bring back the prices to normal levels [43]; ironically, "normal levels" merely indicated a temporary human decision that would allow the continuation of algorithmic pricing."

3.11 What is profiling?

Profiling analyzes aspects of an individual's personality, behavior, interests and habits to make predictions or decisions about them. In particular, profiling' means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyze or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location or movements.

Watch: https://youtu.be/7-MNbzv8lAA

You are carrying out profiling if you:

- collect and analyze personal data on a large scale, using algorithms, AI or machine-learning;
- identify associations to build links between different behaviors and attributes;
- create profiles that you apply to individuals; or
- predict individuals' behavior based on their assigned profiles.

Organizations obtain personal information about individuals from a variety of different sources. Internet searches, buying habits, lifestyle and behavior data gathered from mobile phones, social networks, video surveillance systems and the Internet of Things are examples of the types of data organizations might collect.

They analyze this information to classify people into different groups or sectors. This analysis identifies correlations between different behaviors and characteristics to create profiles for individuals. This profile will be new personal data about that individual.

Organizations use profiling to

- find something out about individuals' preferences;
- predict their behavior; and/or
- make decisions about them.

Profiling can use **algorithms**. An algorithm is a sequence of instructions or set of rules designed to complete a task or solve a problem. Profiling uses algorithms to find correlations between separate datasets. These algorithms can then be used to make a wide range of decisions, for example to predict behavior or to control access to a service. Artificial intelligence (AI) systems and machine learning are increasingly used to create and apply algorithms.

Although many people think of marketing as being the most common reason for profiling, this is not the only application.

What are the benefits of profiling and automated decision-making? Profiling and automated decision making can be very useful for organisations and also benefit individuals in many sectors, including healthcare, education, financial services and marketing. They can lead to quicker and more consistent decisions, particularly in cases where a very large volume of data needs to be analysed and decisions made very quickly.

Examples Profiling is used in some medical treatments, by applying machine learning to predict patients' health or the likelihood of a treatment being successful for a particular patient based on certain group characteristics.

Less obvious forms of profiling involve drawing inferences from apparently unrelated aspects of individuals' behavior.

Using social media posts to analyze the personalities of car drivers by using an algorithm to analyze words and phrases which suggest 'safe' and 'unsafe' driving in order to assign a risk level to an individual and set their insurance premium accordingly.

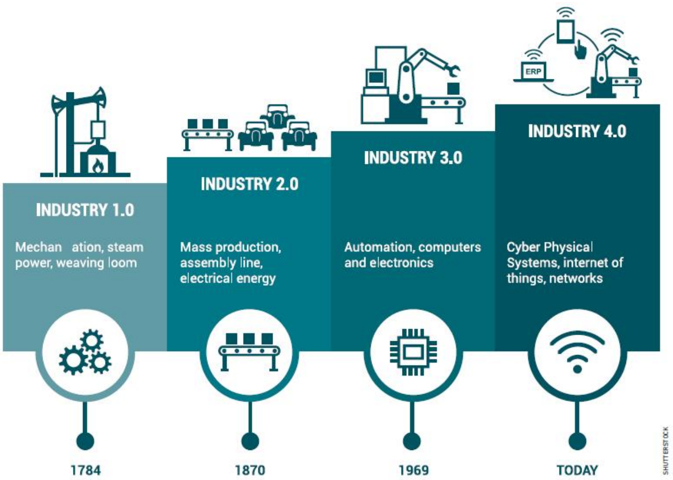
What are the risks? Although these techniques can be useful, there are potential risks:

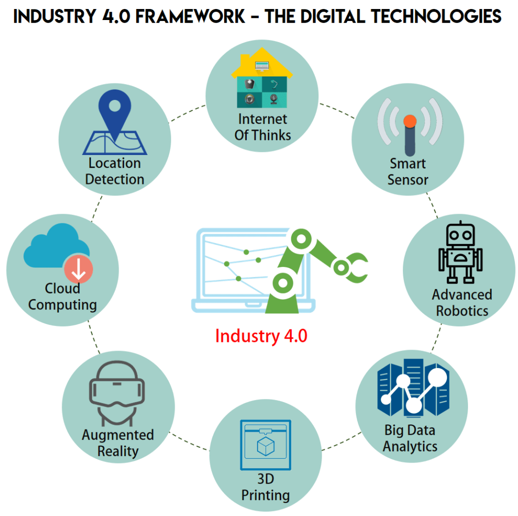
- Profiling is often invisible to individuals.
- People might not expect their personal information to be used in this way.
- People might not understand how the process works or how it can affect them.
- The decisions taken may lead to significant adverse effects for some people.

Just because analysis of the data finds a correlation doesn't mean that this is significant. As the process can only make an assumption about someone's behaviour or characteristics, there will always be a margin of error and a balancing exercise is needed to weigh up the risks of using the results.

3.12 Industry 4.0

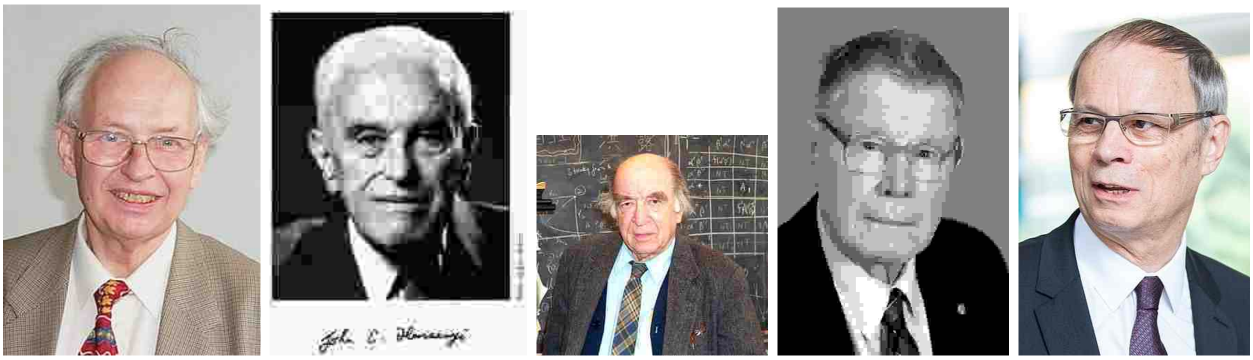
Smart industry or *INDUSTRIE 4.0* refers to the technological evolution from embedded systems to cyber-physical system. 'INDUSTRIE 4.0 represents the coming fourth industrial revolution on the way to an Internet of Things, Data and Services. Decentralized intelligence helps create intelligent object networking and independent process management, with the interaction of the real and virtual worlds representing a crucial new aspect of the manufacturing and production process.





Chapter 4

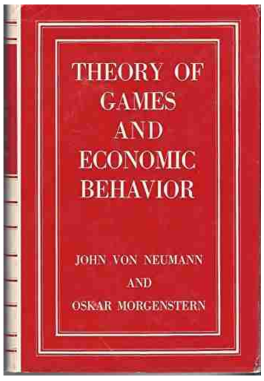
Game theory



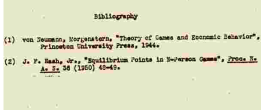
Reinhard Selten, John Harsanyi, Leonid Hurwicz, Thomas Schelling, Jean Tirole

4.1 Introduction

Game theory is the study of mathematical models of **strategic interaction among rational decision-makers**. It has applications in all fields of social science, as well as in logic, systems science and computer science. In the 21st century, game theory applies to a wide range of behavioral relations, and is now an umbrella term for the science of logical decision making in humans, animals, and computers.
Modern game theory began with the idea of mixed-strategy equilibria in two-person zero-sum games and its proof by *John von Neumann*. His work and in particular his jointly written with *Oskar Morgenstern* from 1944 *Theory of Games and Economic Behavior* which considered cooperative games of several players was the beginning of modern game theory. The second edition of this book provided an axiomatic theory of expected utility, which allowed mathematical statisticians and economists to treat decision-making under uncertainty.
Game theory has been widely recognized as an important tool in many fields. As of 2014, with the Nobel Memorial Prize in Economic Sciences going to game theorist *Jean Tirole*, **eleven game theorists have won the economics Nobel Prize** including *Reinhard Selten* from Germany together with *John Harsanyi* and *John Nash* in 1994.



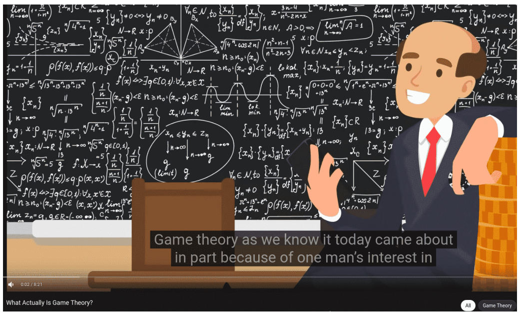
The reference list of John Nash's doctoral thesis:



Bibliography

(1) von Neumann, Morgenstern, "Theory of Games and Economic Behavior", Princeton University Press, 1944.

(2) J. F. Nash, Jr., "Equilibrium Points in N-Person Games", Proc. N. A. S. 36 (1950) 48-49.



https://youtu.be/YueJukoFBMU — Game Theory Explained in One Minute

- Game Theory is the analysis of strategic interactions between two or more players.
- It analyzes behavior when some **players** make **strategic decisions** – their actions affect others **payoff** and, in turn, others' actions affect their **payoffs**.
- It can be applied to real market situations as well as to many different interactions between people.
- It is based on the assumption that players act rationally and maximize their own benefit, and enables economic research through experiments in a lab or from the field.

4.2 Representation of a game: the normal form

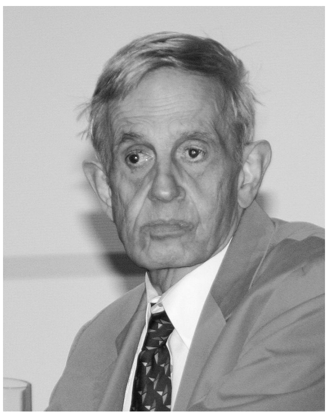
The matrix provided is a normal-form representation of a game in which players move simultaneously (or at least do not observe the other player's move before making their own) and receive the payoffs as specified for the combinations of actions played.

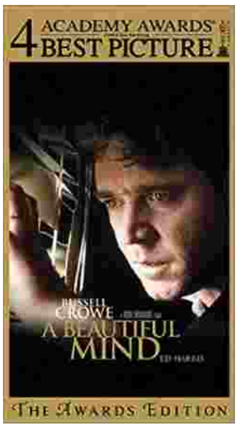
		Person B	
		work	shirk
Person A	work	10 ; 10	5 ; 11
	shirk	11 ; 5	6 ; 6

Here, shirking from work is the **dominant strategy** for both players.

4.3 Nash equilibrium

4.3.1 John Forbes Nash Jr. (1928–2015)





https://youtu.be/jE24tzle1m8 — Dr. John Nash on his life before and after the Nobel Prize
https://youtu.be/ZoZe2tPd3Hk — What is Nash Equilibrium?

https://youtu.be/2d_dtTZQyUM — Nash Equilibrium (taken from *A Beautiful Mind*)

https://library.princeton.edu/special-collections/sites/default/files/Non-Cooperative_Games_Nash.pdf

4.3.2 Identifying a Nash equilibrium

To find a Nash equilibrium in a normal form game, we can look for the best responses for both players in a game. We do so by putting a star to the payoff attained by the best response of a player for all the strategies of the other player. For example, we put a star in the top left corner next to the 4 because S1 is the best response by player B to the action S1 of player A.

Notice that the bottom right corner box has a particular feature: it shows that the strategies of played by all the (two) players and resulting in that outcome are best responses to the others' players best responses. That defines a **Nash equilibrium**.

		Person B		
		S1	S2	S3
Person A	S1	0 ; 4*	4* ; 0	5 ; 3
	S2	4* ; 0	0 ; 4*	5 ; 3
	S3	3 ; 5	3 ; 5	6* ; 6*

The Nash equilibrium is a concept of game theory where the optimal outcome of a game is one where no player has an incentive to deviate from his chosen strategy after considering an opponent's choice.

4.3.3 The prisoners dilema

The Prisoner's Dilemma is the most well-known example of game theory. It shows why two completely rational individuals might not cooperate, even if it appears that it is in their best interests to do so.

https://youtu.be/9uDUv1TpGxI?t=158 — What Actually Is Game Theory?

Consider the example of two criminals arrested for a crime. Prosecutors have no hard evidence to convict them. However, to gain a confession, officials remove the prisoners from their solitary cells and question each one in separate chambers. Neither prisoner has the means to communicate with each other. The criminals are now confronted by the officials with four possible scenarios.

1. If both confess, they will each receive a eight-year prison sentence.
2. If Prisoner 1 confesses, but Prisoner 2 does not (he aims to *cooperate* with Prisoner 1), Prisoner 1 will get free and Prisoner 2 will get twenty years.
3. If Prisoner 2 confesses, but Prisoner 1 does not (he aims to *cooperate* with Prisoner 1), Prisoner 1 will get twenty years, and Prisoner 2 will get free.
4. If neither confesses, each will serve two years in prison.

The prisoner's dilemma

		Person B	
		Confess	Cooperate
Person A	Confess	8 years ; 8 years	0 years ; 20 years
	Cooperate	20 years ; 0 years	2 years ; 2 years

Let us now look how individuals would rationally think what to do:

- If A assumes that B confesses, A would also confess.
- If A assumes that B cooperates, A would confess.

As the same logic applies for B, we can say that the strategy of choice is to confess while the most favorable strategy for both would be to cooperate. The game theoretical equilibrium (both confess) of this game can be called a **Nash equilibrium** because it suggests that both players will make the move that is best for them individually even if it is worse for them collectively.

Exercise 4.1 — Chicken Game (Solution → p. ??)

a) Explain what is understood as a chicken game in game theory.
b) Explain what is understood as a lemon market in game theory.

4.4 Structure of games

4.4.1 Elements of games

- Number of players
- Number of strategies and alternative actions, repsectively
- Payoff functions
- State of information (who knows what)
- Timing of actions and information

4.4.2 Classes of games

- cooperative vs. non-cooperative
- Static vs. dynamic
- one-shot vs. repeated
- non-zero-sum vs. zero-sum
- perfect information vs. non-perfect information
- symmetric information vs. asymmetric information
- deterministic vs. non-deterministic payoffs (random)

4.5 Representation of a game: the extensive-form

An representation allows the explicit representation of a number of key aspects, like the sequencing of players' possible moves, their choices at every decision point, the (possibly imperfect) information each player has about the other player's moves when they make a decision, and their payoffs for all possible game outcomes. Extensive-form games also allow for the representation of incomplete information in the form of chance events modeled as *moves by nature*.

Exercise 4.2 — Matching pennies (random and simultaneous version) (Solution → p. 53)

Write down the following game in the normal and the extensive form:
Matching pennies (random and simultaneous version) is a game with two players (1, 2). Both players flip simultaneously a penny high. Each penny falls down and shows either head up or tail up. If the two pennies match (either both heads up of both tails up) player 2 wins and player 1 must pay him a Euro. If the two pennies do not match player 1 wins and player 2 must pay him a Euro.
Describe the elements of the game and the class of this game.

Exercise 4.3 — Matching pennies (random version) (Solution → p. 54)

Write down the following game in the extensive form:
Matching pennies (random and simultaneous version) is a game with two players (1, 2). **Player 1 starts by flipping a fair penny high, catches it and then turns it over into the other hand so that the result is hidden to the other player. Then, player 2 flips the coin.** If the two pennies match (either both heads up of both tails up) player 2 wins and player 1 must pay him a Euro. If the two pennies do not match player 1 wins and player 2 must pay him a Euro.

Exercise 4.4 — Matching pennies (strategic version) (Solution → p. 54)

Write down the following game in the extensive form and discuss the strategies of both:
Matching pennies (strategic version) is a game with two players (1, 2). Player 1 starts **and decides whether to put a coin with either head up or tail up onto a table. Player 2 can see the decision of player 1. Then, player 2 decides whether to put a coin with head or table on the table.** If the two pennies match (either both heads up of both tails up) player 2 wins and player 1 must pay him a Euro. If the two pennies do not match player 1 wins and player 2 must pay him a Euro.

Solution to Exercise 4.2 — Matching pennies (random and simultaneous version)
(Exercise → p. 52)

It is a game that belongs to the following classes:

- *non-cooperative,*
- *static,*
- *one-shot,*
- *zero-sum,*
- *perfect information*
- *symmetric information*
- *non-deterministic payoffs*

Elements of the game:

- *Number of players: 2*
- *Number of strategies: No strategies as whether head or tail shows up is random*
- *Payoff functions:*

$$\text{Player 1} = \begin{cases} 1, & \textit{if } (T,T) \text{ or } (H,H) \\ -1, & \textit{otherwise} \end{cases}$$

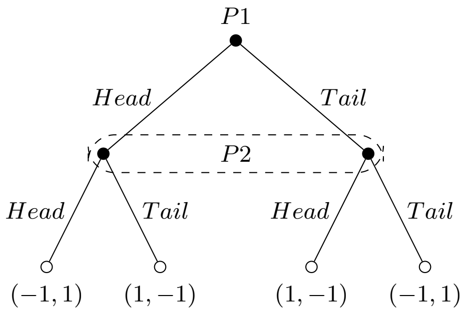
$$\text{Player 2} = \begin{cases} 1, & \textit{if } (H,T) \text{ or } (T,H) \\ -1, & \textit{otherwise} \end{cases}$$

- *State of information: Everybody knows the rules and is perfectly informed*
- *Timing of actions and information: both throw the coin at the same time and see the result at the same time*

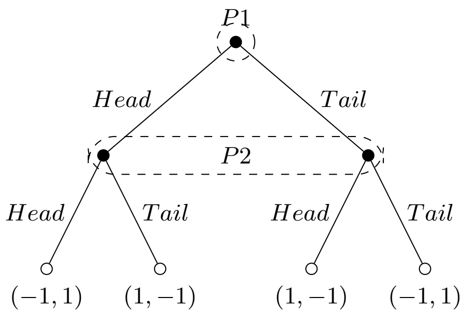
		Person 2	
		Head	*Tail*
Person 1	*Head*	*-1 ; 1*	*1 ; -1*
	Tail	*1 ; -1*	*-1 ; 1*

Solution to Exercise 4.3 — Matching pennies (random version) (Exercise → p. 53)

As player 2 has no idea what player 1 has chosen, he cannot come up with any strategy that increases his winning rate.

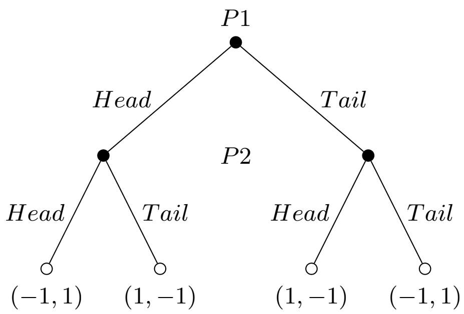


The dashed circled line indicates that player 2 actually is not informed about whether P1 decided head or tail. As we have now introduced how to graphically show that some players have a restricted information set, we can draw the extensive form also for the random and simultaneous version of the matching pennies game. Please note that the dashed circle around player 1 is redundant an hence it is a convention not to draw it sometimes.



Solution to Exercise 4.4 — Matching pennies (strategic version) (Exercise → p. 53)

As player 2 has complete information about the decision of player 1, he cannot always come up with the choice that makes him win. That is, if player 1 chooses head(/tail) player one will also choose head(/tail).



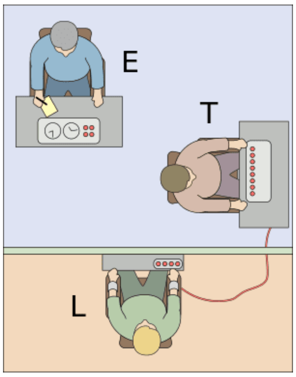
4.6 Behavioral economics

4.6.1 Behavioral experiments

Behavioral experiments are important in economics because they enable validating economic theories in a controlled environment. For example, experiment can help

- to select the right theories for the respective question,
- to test the robustness of models w.r.t. certain circumstances, or
- to calibrate of parameters such as risk aversion.

Moreover, experiments can provide new empirical insights on which new theories can be grounded.



Milgrim Experiment (starting 1961)

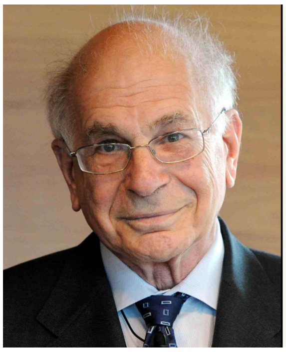
To see how behavioral experiments actually works, or better worked because of ethnical reasons, let us look at the old, controversial and infamous **Milgrim Experiment**

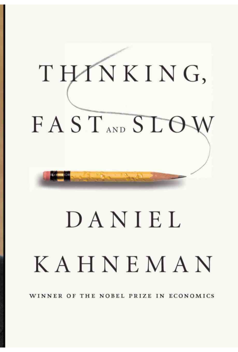
▶ `https://youtu.be/3YOox59J0Bk` — The Milgram Experiment - Shock Study on Obedience Conclusions

▶ `https://youtu.be/Kzd6Ew3TraA` (short) or `https://youtu.be/rdrKCilEhC0` (long) — Milgrim Experiment (original recordings)

4.6.2 Behavioral economic experiments

Tversky and Kahneman's judgment under uncertainty





▶ `https://youtu.be/3IjIVD-KYF4` — An Introduction to Tversky and Kahneman's Work

- Behavioral Economics studies the effects of psychological, cognitive, emotional, cultural and social factors on the decisions of individuals and institutions and how those decisions vary from those implied by classical economic theory.
- It is primarily concerned with the bounds of rationality of economic agents. Behavioral models typically integrate insights from psychology, neuroscience and microeconomic theory. The study of behavioral economics includes how market decisions are made and the mechanisms that drive public choice.
- Behavioral Economic studies have proven that many people **do not act as assumed by theories in microeconomics**. In most of these theories it is claimed that the individuals act rational. That means like a *homo economicus*. This economic man is the portrayal of humans as agents who are consistently rational, narrowly self-interested, and who pursue their subjectively-defined ends optimally.

4.7 St. Peterburg revisited

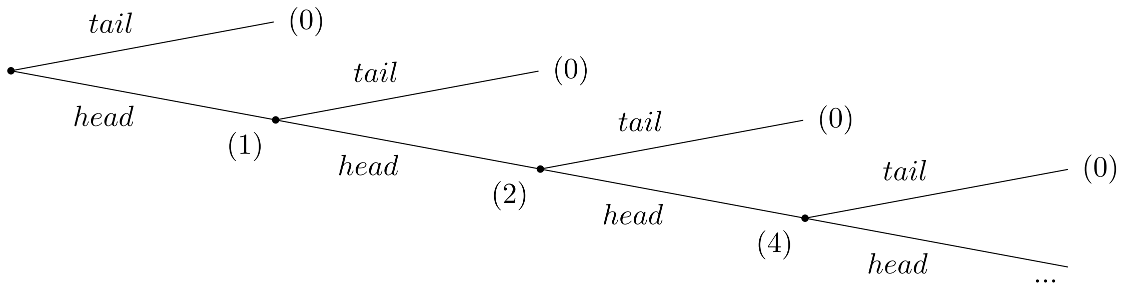
Rationality of actors against the background of well-specified preferences is a central assumption in all areas of economics. In the St.-Petersburg-Paradoxon discussed in **??** it becomes very clear that the orientation towards expected values is not adequate for games, but that expected utility must be the relevant concept if one wants to reconcile actual behavior with the concept of rationality.

In the following, we will write down the St. Peterburg game explained on page **??** in the extensive-form[1]. Moreover, we will state the payoff function of the St. Petersburg game and also state the expected payoff if the game is played I-times. From experiments we know that the willingness to pay for

[1]See p. 52 if you don't know what the extensive form is.

participating does not increase with the number of rounds the game is played. How could we fix this paradox?

St. Peterburg game in extensive form The extensive form is given below:



First, the expected value of the game must now be calculated. For this the probability $p(i)$ of throwing any number i of consecutive heads is crucial. This probability is given by

$$p(i) = \underbrace{\frac{1}{2} \cdot \frac{1}{2} \cdot \ldots \cdot \frac{1}{2}}_{i \text{ Faktoren}} = \frac{1}{2^i}$$

The payoff $W(i)$ is, if head appears i-times in a row by

$$W(i) = 2^{i-1}$$

The expected payoff $E(W(I))$ if the coin is flipped I times is then given by

$$E(W(I)) = \sum_{i=1}^{I} p(i) \cdot W(i) = \sum_{i=1}^{I} \frac{1}{2^i} \cdot 2^{i-1} = \sum_{i=1}^{I} \frac{1}{2} = \frac{I}{2}$$

Thus, the expected payoff grows proportionally with the maximum number of rolls. This is because at any point in the game, the option to keep playing has a positive value no matter how many times head has appeared before. Thus, the expected value of the game is infinitely high for an unlimited number of rolls. Even with a very limited maximum number of rolls of, for example, $I = 100$, only a few players would be willing to pay 50 Euro for participation. The relatively high probability to leave the game with no or very low winnings leads in general to a subjective rather low evaluation that is below the expected value.

The basic idea of *Daniel Bernoulli*, which enables a solution of this paradox lies in the conceptual separation of the expected payoff and its utility. He describes the basis of the paradox as follows:

> "Until now scientists have usually rested their hypothesis on the assump- tion that all gains must be evaluated exclusively in terms of themselves, i.e., on the basis of their intrinsic qualities, and that these gains will always produce a utility directly proportionate to the gain." (Bernoulli, 1954, p. 27)

Thus, we should evaluate the game in categories of expected utility:

$$E(u(W(I))) = \sum_{i=1}^{I} p(i) \cdot u(W(i)) = \sum_{i=1}^{I} \frac{1}{2^i} \cdot u\left(2^{i-1}\right)$$

Daniel Bernoulli himself proposed the following logarithmic utility function:

$$u(W) = a \cdot \ln(W),$$

where a is a positive constant. Using this function in the expected utility, we get

$$E(u(W(I))) = \sum_{i=1}^{I} \frac{1}{2^i} \cdot a \cdot \ln\left(2^{i-1}\right) = a \cdot \sum_{i=1}^{I} \frac{i-1}{2^i} \ln 2 = a \cdot \ln 2 \cdot \sum_{i=1}^{I} \frac{i-1}{2^i}.$$

The infinite series, $\sum_{i=1}^{I} \frac{i-1}{2^i}$, converges to 1 ($\lim_{I\to\infty} \sum_{i=1}^{I} \frac{i-1}{2^i} = 1$). Thus, given an ex ante unbounded number of throws, the expected utility of the game is given by

$$E(u(W(\infty))) = a \cdot \ln 2.$$

In experiments in which people were offered this game, their willingness to pay was roughly between 2 and 3 Euro. Thus, the suggests logarithmic utility function seems to be a pretty realistic specification. The main reason is mathematically that the increasing expected payoff has decreasing marginal utility and hence the utility function reflects the risk aversion of many people.

Exercise 4.5 — Rationality and risk (Solution → p. ??)

There are 90 balls in an box. It is known that 30 of them are red, the remaining 60 are blue or green. An individual can choose between the following lotteries:

Payoff	probability	
Lottery 1	100 Euro if a red ball is drawn 0 Euro else	$p = \frac{1}{3}$
Lottery 2	100 Euro if a blue ball is drawn 0 Euro else	$0 \leq p \leq \frac{2}{3}$

In a second variant it has the choice between the following lotteries:

	Payoff	probability
Lottery 3	100 Euro if a red or green ball is drawn 0 Euro else	$\frac{1}{3} \leq p \leq 1$
Lottery 4	100 Euro if a blue or green ball is drawn 0 Euro else	$p = \frac{2}{3}$

a) Which of the lotteries does the individual choose on the basis of expected values (risk neutral)?
b) Which of the lotteries does the individual choose on the basis of expected utility if the utility of a payoff of x is given by $u(x) = x^2$?
c) Empirical studies, e.g. Camerer and Weber (1992), show, however, that most individuals will usually choose lotteries 1 and 4. will. Discuss: Is this consistent with rational behavior?

Bibliography

Bernoulli, D. (1954). Exposition of a new theory on the measurement. *Econometrica*, 22(1):23–36. [Cited on pages 27 and 56.]

Blumberg, R. and Atre, S. (2003). The problem with unstructured data. *DM Review*, 13(42-49):62. [Cited on page 38.]

Camerer, C. and Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, 5(4):325–370. [Cited on pages 28 and 57.]

CEOpedia (2021). Rational decision making. Accessed March 11, 2023. [Cited on pages 11 and 18.]

Demetis, D. and Lee, A. S. (2018). When humans using the IT artifact becomes IT using the human artifact. *Journal of the Association for Information Systems*, 19(10):5. [Cited on page 46.]

Eilon, S. (1969). What is a decision? *Management Science*, 16(4):B172–B189. [Cited on page 8.]

Finne, T. (1998). The three categories of decision-making and information security. *Computers & Security*, 17(5):397–405. [Cited on pages 18, 19, and 24.]

Fitzgerald, S. P. (2002). *Decision Making*. Capstone Publishing. [Cited on pages 9, 10, 11, and 12.]

Kahneman, D. and Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3(3):430–454. [Cited on page 15.]

Morse, L. C., Babcock, D. L., and Murthy, M. (2014). *Managing Engineering and Technology*. Pearson, sixth edition edition. [Cited on page 22.]

NobelPrize.org (2021). Herbert A. Simon: Facts. Accessed March 11, 2023. [Cited on page 13.]

Provost, F. and Fawcett, T. (2013). *Data Science for Business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media. [Cited on pages 39 and 40.]

Wikipedia (2020). Wikipedia, the free encyclopedia. `https://en.wikipedia.org/`. [Online; accessed 3-March-2020]. [Cited on page 9.]

Appendix A

Solutions to exercises

Solution to Exercise 1.7 — Lisa is pregnant (Exercise → p. 15)

At the beginning of the Inferential Statistics course at summer 2020, I also asked student this question. Here are the results of the 16 students who participated:

Option	*Frequency of answers*	*Percentage*
a)	*3*	*19%*
b)	*2*	*13%*
c)	*2*	*13%*
d)	*3*	*19%*
e)	*6*	*38%*

How did they reach their answers? Like most people, they decided that Lisa has a substantial chance of having a baby with Down syndrome. The test gets it right 86 percent of the time, right? That sounds rather reliable, doesn't it? Well it does, but we should not rely on our feelings here and better do the math because the correct result would be that there is ***just a 1.7 percent chance of the baby having a Down syndrome****.*

Now, let us proof the result that there is ***just a 1.7 percent chance of the baby having a Down syndrome*** *using Bayesian Arithmetic.*

Watch the two videos linked here:

- *Bayes theorem* https://youtu.be/HZGCoVF3YvM *and*
- *The quick proof of Bayes' theorem* https://youtu.be/U_85TaXbeIo

Also, consider this interactive tool: https://www.skobelevs.ie/BayesTheorem/

Solution: *Let A be the event of the Baby has a Down syndrom and B the test is positive. Then,*

$$
\begin{aligned}
P(A) &= 0.001 \\
P(B \mid A) &= 0.86 \\
P(B \mid \neg A) &= 0.05 \\
P(B) &= \frac{999 \cdot 0.05}{1000} + \frac{1 \cdot 0.86}{1000} = \frac{50.81}{1000} = 0.05081 \\
P(A \mid B) &= \frac{P(B \mid A)P(A)}{P(B)} = \frac{0.86 \cdot 0.001}{0.05081} = 0.016925802
\end{aligned}
$$

Exercise A.1 — How good should a test be? (Solution → p. 60)

Suppose that a disease infects one out of every 100 people in a population and that the test for it comes back positive in 99% of all cases if a person has the disease. Moreover, the test also produces some false positive, that is 1% of positive tests belong to uninfected patients.

a) Calculate the probability of not having the disease given a positive test.

b) Suppose you have two options to improve the test: (1) 99,2% of all test come back positive if a person has the disease, or (2) only 0.9% of uninfected persons are tested positive. Which option would you prefer?

c) Calculate the false positive rate that assures that the probability of not having the disease given a positive test is below 10%.

Solution to Exercise A.1 — How good should a test be? (Exercise → p. 59)

a) *We want to know $P(\neg A|B)$ and we have*

$$\begin{aligned} P(A) &= .01 \text{ and hence } P(\neg A) = .99 \\ P(B|A) &= .99 \\ P(B|\neg A) &= .01 \end{aligned}$$

Using Bayes Theorem we know that

$$\begin{aligned} P(\neg A|B) &= \frac{P(B|\neg A)P(\neg A)}{P(B\cap A)+P(B\cap\neg A)} \\ &= \frac{P(B|\neg A)P(\neg A)}{P(B|A)P(A)+P(B|\neg A)P(\neg A)} \\ &= \frac{.01\cdot .99}{.99\cdot .01+.01\cdot .99} = \frac{.0099}{.0099+.0099} = \frac{.0099}{.0198} = .5 \end{aligned}$$

Thus, the probability table looks like that:

	A	*¬A*	
B	*.0099*	*.0099*	*.0198*
¬B	*.001*	*.9801*	*.9802*
	.01	*.99*	*1*

b) *(1) ⇒ $P(B|A) = .992$ and we want to know $P(\neg A|B)$. Using Bayes Theorem we know that*

$$\begin{aligned} P(\neg A|B) &= \frac{P(B|\neg A)P(\neg A)}{P(B\cap A)+P(B\cap\neg A)} \\ &= \frac{P(B|\neg A)P(\neg A)}{P(B|A)P(A)+P(B|\neg A)P(\neg A)} \\ &= \frac{.01\cdot .99}{.992\cdot .01+.01\cdot .99} = \frac{.0099}{.01982} \approx .4994 \end{aligned}$$

Let us see if other things have changed, too:

	A	*¬A*	
B	*.00992*	*.0099*	*.01982*
¬B	*.00008*	*.9801*	*.98018*
	.01	*.99*	*1*

(2) ⇒ $P(B|\neg A) = .009$ and we want to know $P(\neg A|B)$. Using Bayes Theorem we know that

$$\begin{aligned} P(\neg A|B) &= \frac{P(B|\neg A)P(\neg A)}{P(B|A)P(A)+P(B|\neg A)P(\neg A)} \\ &= \frac{.009\cdot .99}{.99\cdot .01+.009\cdot .99} = \frac{.00891}{.0099+.00891} = \frac{.00891}{.01881} \approx .4737 \end{aligned}$$

Let us see if other things have changed, too:

	A	*¬A*	
B	*.0099*	*.00891*	*.01881*
¬B	*.0001*	*.98109*	*.98119*
	.01	*.99*	*1*

c) From the question, we know

$$
\begin{aligned}
P(A) &= .01 \\
P(\neg A) &= .99 \\
P(B|A) &= .99
\end{aligned}
$$

and using probability theory's multiplication rule, we also know

$$P(A \cap B) = P(B|A)P(A) = .99 \cdot .01 = .0099.$$

Now, we search for the largest possible false positive rate, i.e.,

$$P(B|\neg A) =?$$

which yields a probability of not having the disease given a positive test of below 10%, i.e.,

$$P(\neg A|B) < .1$$

Using Bayes Theorem, we can state that (for more explanations to that see box below)

$$
\begin{aligned}
P(\neg A|B) &= \frac{P(\neg A \cap B)}{P(B)} \\
&= \frac{P(\neg A \cap B)}{P(A \cap B) + P(\neg A \cap B)} \\
&= \frac{P(B|\neg A)P(\neg A)}{P(B|A)P(A) + P(B|\neg A)P(\neg A)} \\
&= \frac{P(B|\neg A) \cdot .99}{.0099 + (P(B|\neg A) \cdot .99)} < .1
\end{aligned}
$$

This is an equation with only one unknown. Thus, we can solve it:

$$
\begin{aligned}
&\frac{P(B|\neg A) \cdot .99}{.0099 + (P(B|\neg A) \cdot .99)} < .1 \\
&\Leftrightarrow P(B|\neg A) \cdot .99 < .1 \cdot [.0099 + (P(B|\neg A) \cdot .99)] \\
&\Leftrightarrow \underbrace{P(B|\neg A)}_{x} \cdot .99 < .1 \cdot [.0099 + (\underbrace{P(B|\neg A)}_{x} \cdot .99)] \\
&\Leftrightarrow x \cdot .99 < .1 \cdot [.0099 + x \cdot .99)] \\
&\Leftrightarrow .99x < .00099 + .099x \\
&\Leftrightarrow .99x - .099x < .00099 \\
&\Leftrightarrow x(.99 - .099) < .00099 \\
&\Leftrightarrow x < \frac{.00099}{.891} \\
&\Leftrightarrow x < 0,001111111 = \frac{1}{900} \\
&\Leftrightarrow P(B|\neg A) < \frac{1}{900}
\end{aligned}
$$

Appendix B

Additional exercises

Judgement and Decision Making
—
Additional Exercises

© Prof. Dr. Stephan Huber*

HS Fresenius, Cologne, November 18, 2021

1. (5 points) Given the following payoff matrix, calculate the best decision based on the Körth's rule.

O_{ij}	k_1	k_2
a_1	1	4
a_2	5	2
a_3	3	3

2. (5 points) Given the following utility matrix. What would be the weights for the characteristics k_j that would make the decider to be indifferent—with respect to the aggregated expected utility—between all three alternatives a_i.

O_{ij}	k_1	k_2
a_1	1	4
a_2	5	2
a_3	3	3

3. Suppose you want launch a product on the market. The following table gives you the expected market shares, x, and the and their respective probabilities of occurrence:

Market share in %	10	12	15	18	20
Probability of occurrence	.1	.2	.3	.2	.2

   (a) (3 points) Calculate the expected market share.

   (b) (3 points) Given the utility function, $u(x) = \frac{x^2}{5}$, calculate your expected utility.

   (c) (3 points) Assume that launching the product costs €50,000 and that your utility function with respect to money, y, is $u(y) = \frac{y}{1{,}000}$. Would you decide to launch the product if you are risk-neutral?

   (d) (3 points) Assume you can launch the product also for €40,000 which would change your success on the market as follows:

Market share in %	10	12	15	18	20
Probability of occurrence	.2	.2	.2	.2	.2

   Is this *cheap way* to launch the product preferable to the more expensive one?

4. The entrepreneur 'A' classifies an investment, which generates revenue of €2.000 with probability of 25%, a revenue of €400 with a probability of 50%, and no revenue with a probability of 25%. In

*stephan.huber@hs-fresenius.de

the same action, his competitor 'B' assumes to have a revenue of €2.000 with a probability of 40%, a revenue of €400 with a probability of 40%, and with a probability of 20% no revenue.

(a) (2 points) Calculate the expected revenue of A.

(b) (2 points) Calculate the expected revenue of B.

(c) (4 points) Suppose A is risk-neutral, which statements are correct:

- (A) A invests as long as the costs of the investment do exceed his expected revenue.
- (B) A invests as long as the costs of the investment do not exceed his expected revenue.
- (C) A is indifferent whether to invest or not if the costs of the investment are equal to his expected revenue.

(d) (4 points) Suppose B is a risk-taker, which statements are correct:

- (A) B invests even if the costs of the investment slightly exceed his expected revenue.
- (B) B invests when the costs of the investment do not exceed his expected revenue.
- (C) B is indifferent whether to invest or not if the costs of the investment are equal to his expected revenue.

5. Given is the following payoff matrix:

	z_1	z_2	z_3	z_4	z_5
a_1	1	1	1	1	1
a_2	0	0	2	1	1
a_3	1	1	0	0	0

(a) (2 points) Whats the optimal choice of action given the Maximin-strategy (not Körth's maximin-rule), i.e., you try to have the *best worst-case.*

(b) (2 points) Whats the optimal choice of action given the Maximax-strategy, i.e., you try to have the *best best-case.*

(c) (2 points) Whats the optimal choice of action given the Hurwicz-criterion and a $\alpha = 0$.

(d) (2 points) Whats the optimal choice of action given the Hurwicz-criterion and a $\alpha = 1$.

(e) (2 points) Whats the optimal choice of action given the Hurwicz-criterion and a $\alpha = .6$.

(f) (2 points) Whats the optimal choice of action given the Laplace-criterion.

6. (12 points) **Mark the correct answer(s):**
Watch the TEDx-Talk of Gerd Gigerenzer *How do smart people make smart decision?*':
`https://youtu.be/-Lg7G8TMe_A`. In his talk he mentioned:

- (A) Albert Einstein
- (B) Benjamin Franklin
- (C) Enrique Iglesias
- (D) Harry Markowitz
- (E) Paul Krugman
- (F) Expected utility theory
- (G) The best decision under risk is not the best decision under uncertainty

(H) Heuristics are indispensable for good decisions under uncertainty.

(I) Heuristics can do better than even optimization strategies in a world of risk, not in a world of uncertainty.

(J) Complex problems do not require complex solutions.

(K) The miracle of the Hudson River

(L) The miracle of Joey Ramone

(M) More information, time, and computation is always better.

7. **[Decision Making Under Uncertainty]** Given the following payoff table.

	z_1	z_2	z_3	z_4	z_5
a_1	3	3	3	3	3
a_2	5	4	3	2	1
a_3	2	2	6	6	7
a_4	1	5	1	5	6
a_5	2	2	2	1	3

(a) (3 points) State which alternatives can be excluded because they are dominated by other alternatives. (Hint: exclude dominated alternative in the calculations of the following questions.)

(b) (6 points) Calculate the preferred alternative, a_i, given the *Laplace criterion.*

(c) (3 points) What is the best alternative given the *Maximax criterion.*

(d) (3 points) What is the best alternative given the *Minimax criterion.*

8. **[Decision Making Under Certainty]** Given the following payoff table.

	z_1	z_2	z_3	z_4	z_5
a_1	1	2	3	4	3
a_2	4	3	2	1	4
a_3	4	5	0	5	6
a_4	1	5	1	5	6
a_5	2	2	2	1	3
a_6	3	4	0	5	3

(a) (4 points) State which alternatives can be excluded because they are dominated by other alternatives.

(b) (5 points) Suppose your preference scheme is as follows:

$$g_1 = \frac{1}{2}; \quad g_2 = 1; \quad g_3 = 2; \quad g_4 = 1; \quad g_5 = 1.$$

Find the order of preference based on the aggregated expected utility.

9. **[Decision Making Under Uncertainty]** Given the following payoff table.

	z_1	z_2	z_3	z_4
a_1	1	2	3	9
a_2	5	6	7	8
a_3	1	7	1	5
a_4	2	2	2	2
a_5	4	4	4	7

(a) (2 points) State which alternatives can be excluded because they are dominated by other alternatives. (Hint: exclude dominated alternatives in the calculations of the following questions.)

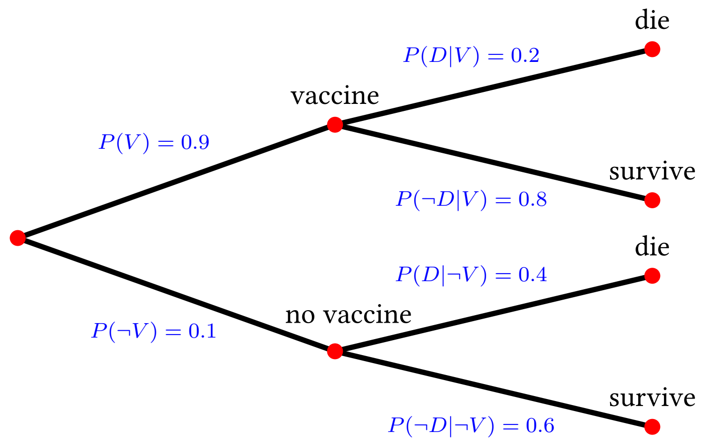
(b) (2 points) Using the *Laplace criterion*, calculate the preferred alternative(s), a_i.

(c) (2 points) Using the *Maximax criterion*, calculate the best alternative(s), a_i.

(d) (2 points) What is the best alternative given the *Minimax criterion*.

10. **[Bayes Theorem]**

You read on *Facebook* that in the year 2021 about over 80% of people that died were vaccinated. You are shocked by this high probability that a dead person was vaccinated, $P(V|D)$. You decide to check this fact. Reading the study on which the Facebook post is referring to, you find out that the study only refers to people above the age of 90. Moreover, you find the following *Tree Diargram*. It allows to check the fact as it describes the vaccination rates and the conditional probabilities of people to die given the fact they were vaccinated or not. In particular, D denotes the event of *die* and $\neg D$ denotes *not die*, i.e., survive; V denotes the event of *vaccinated* and $\neg V$ *not vaccinated*.



(a) (2 points) Calculate the overall probability to die, $P(D)$

(b) (4 points) Calculate the probability that a person that has died was vaccinated, $P(V|D)$.

(c) (2 points) Your calculations shows that the fact used in the statement on *Facebook* is indeed true. Discuss whether this number should have an impact to get vaccinated or not.

11. (10 points) **[Bayes Theorem: COVID-17 and Death]**

Suppose you analyze the causes of death of people over the course of a year in times of COVID-17 a.k.a. *the Corona disease*. Let's assume that the base rate of someone dying before Corona, $P(Die)$, was 1%, and that the base rate for having COVID-17, $P(Corona)$, is 5%. Moreover, suppose you know that from all people, 4% of those that die have COVID-17, i.e., $P(Corona|Die) = 0.04$.

(a) Calculate the probability that a person will die from COVID-17, i.e., $P(Die|Corona)$.

(b) Calculate the probability that a person is infected with COVID-17 and dies in the given period, i.e., $P(Die \cap Corona)$.

12. (12 points) **[Bayes Theorem: Elderly Fall and Death]**

    Suppose you analyze the causes of death of elderly persons (over 80 years of age) over the course of a year. Let's assume that the base rate of someone elderly dying, $P(Die)$, is 10%, and that the base rate for elderly people falling, $P(Fall)$, is 5%. Moreover, you know that from all elderly people, 7% of those that die had a fall, i.e., $P(Fall|Die) = 0.07$.

    (a) Calculate the probability that an elderly person will die from a fall, i.e., $P(Die|Fall)$.

    (b) Calculate the probability that an elderly person falls and dies in the given period, i.e., $P(Die \cap Fall)$.

    (c) Calculate the probability that an elderly person falls and doesn't in the given period, i.e., $P(Die \cap \neg Fall)$.

    (d) Calculate the probability that an elderly person does neither fall nor die in the given period, i.e., $P(\neg Die \cap \neg Fall)$.

References