An introduction to mathematical models in management

Fabrizio Iozzi

September 6, 2015

Preface

First of all, I would like to state that these notes are *not* a reference manual for the topics dealt with. On the contrary, this is a very brief account of some fundamental facts with some problems and examples of applications. The reader interested in studying these subjects is strongly suggested to look at [4] for probability theory, at [2] for decision making problems and at [3] for linear programming.

The choice of the subjects and the order in which they are presented here follow the syllabus of the "Quantitative methods for management" course offered in the Graduate Program in Management at Università Bocconi and taught by the author. Many topics were presented in the class with the help of various software, mainly Microsoft Excel. However, I decided not to give any reference to any software because the software world is changing at so rapid a pace that it is almost impossible to keep up with it. Since the course was introduced in 2006, the instructors faced 4 different operating systems (Windows XP, Windows Vista, Windows 7 e Mac OSX), 3 versions of MSExcel (2003, 2007, 2010), 2 versions of an add-in [5], not to mention some interesting "features" of the combination of the above software (e.g., the different behavior the software has depending on the language of the operating system). The inclusion of references to software would have transformed this document in a neverending work in progress or a "fabbrica del Duomo." [1]

These notes are the English translation of the original Italian lecture notes. During the translation, I wish to thank Emanuele Borgonovo for his help in revisioning the original text and in rewriting parts of it. The chapter in Influence Diagrams is largely based on the corresponding article he wrote for the Encyclopedia of Medical Decision Making [1].

I would also like to thank again Francesca Beccacece and Enrico Moretto for draft editing of the Italian version and for many suggestions; Alessandra Cillo for her helpful comments and Gabriele Gurioli for providing me useful material for the expected utility section.

Finally, I apologize in advance for any typographical errors. If you find any, and I am sure you will, please list them in an email to be sent to fabrizio.iozzi@unibocconi.it.

[1] "The Duomo factory". This phrase is used to point at never ending enterprises. The "Veneranda Fabbrica Del Duomo" was estabilished in 1387 to build Milan's Duomo and for its maintenance. Construction work had lasted for centuries and the Cathedral was completed only in 1892 (!). Given its size and its rich and complex decorative and architectural apparatus, the Duomo always needs work of restoration, consolidation, etc. The Veneranda Fabbrica, therefore, has never been closed and it is still operating today.

Exercises

The exercises in these notes are a selection of exam papers and in-class assignments 2006–2009, all with well commented solutions.

Quickies

Each chapter ends with some "quickies". Quickies are questions with very brief answer, i.e. no more than one line. Quickies are neither quick questions nor they have to be answered quickly: only the answers should be concise. In general, all answers must be properly justified to get credit. For example, if the question is

> Is 10 a prime number? Why?

the right answer would be No, because 10 is the product of 2 and 5. On the contrary, the following answers are to be considered wrong for the reasons shown in parentheses:

- No, because it's even (justification is wrong–2 is even *and* is prime)
- No, because is a multiple of 5 (justification is wrong–5 is a multiple of itslef *and* is prime)
- No (justification is missing)
- No, because it can be divided by 5 and is greater than 9 (the first statement is right but the second is not)

Milano, September 6, 2015

Contents

Part I

Mathematical tools

Chapter 1

An introduction to probability

1.1 Probability in the real world

Everything in the world is uncertain, at least to a certain degree. Among sciences, some of them have been given the attribute of "exact" as they share a somewhat low degree of randomness, in the sense that in these contexts predictions are quite good most of the time. This is in contrast with the so called "social sciences", where a higher degree of uncertainty is always present and must be taken into account.

For example, consider a European entrepreneur who signed a contract to buy an asset 6 months from now for, say, 1000 US dollars. As she usually pays in euros, she many consider these two options:

- to immediately exchange some EUR into USD to get 1000 USD and to save them until the payment occurs
- to save the equivalent of 1000 USD in EUR and exchange them just before the payment

The consequences of both actions are unpredictable. If, during the 6 months, the exchange rate USD-EUR increases (i.e. the equivalent in US dollars of 1 EUR increases) it is best to wait until the payment to exchange EUR in USD, to save money. If, on the contrary, the exchange rate decreases, the best thing to do is to exchange euros for US dollars now. The effect of the same variations on the second option are opposite. Anyway, one understands there is no optimal choice. Nonetheless, if the entrepreneur wants not to be subject to risk, she could sign a contract with a bank that will exchange euros in dollars at a prefixed rate at a prefixed time in the future. Such a contract is called an *option*, and is quite useful for the entrepreneur: the bank guarantees the delivery of the foreign currency in the future with a fixed price. This way, the entrepreneur has transferred the risk to the bank and, for this reason, the bank does not sign the contract for free: the contract has a price and that's the price for the risk. The need for computing the equivalent of a certain consequence in the future has not disappeared; it has just moved from the entrepreneur to the bank.

Situations like the previous one, usually share two common problems:

- the need to determine the probability that a certain event will occur;

- the need to do some computations with the probability just found.

It is important to point out that these problems are different: the first, to find the probabilities, always comes before the second.

Classical probability

The first problem, the determination of probabilities, is an extremely difficult task. Historically, the first attempt to solve it was the "classical" theory of probability. In the classical theory of probability, if all the results of an experiment are equally likely, the probability of any event is the ratio of the number of results favorable to the event to the total number of results. For example, if the experiment is the determination of the EUR/USD exchange rate 6 months from now, and we consider just three results, "less or equal to 1.2", "greater or equal to 1.4" e "between 1.2 and 1.4", equally probable, the probability that "the rate will be greater than or equal to 1.2" is 2/3, because there are only 3 possible outcomes and only 2 out of 3 are favorable to the event.

subjective view

Following the "subjective", the determination of the probability of an event depends on the person who estimates it. For example, consider the launch of a new product and the estimation of its probability of success: a young person could judge the product too "traditional" and assess the probability as 40%. In contrast, an experienced business man could positively view the "continuity" of the product from the previous versions and estimate the same probability as 70%. Of course, both of them will base their estimates only on their respective (different) knowledge, which in turn is made up of the past experiences of the individual.

frequentist view

According to a third view, called frequentist, the determination of probabilities is an experimental process where the experiment is to be done, in the same conditions, an infinite number of times. The relative frequency of the event in the sequence of experiments will "converge" to the probability of the event. Although somehow intriguing, this view is in fact unfeasible. To repeat an experiment in the same conditions again and again is impossible; in addition, even if the relative frequency will converge, the definition gives no clue about when, i.e. there is no way to know how many experiments are to be done to get a sufficient approximation of the required probability.

As every view has its own strengths and weaknesses, we will think of the values of probabilities as being given "a priori".

The second problem, to find how to use probability in computations, is a technical one and can be solved in a formal way. What follows is a (very) brief summary of a few facts from probability theory. The reader interested in going into a greater depth into this topic could look at [4].

1.2 Formalizing probability: events and probability

Given an experiment whose outcome is uncertain, we define a "sample space".

Definition 1.2.1 (Sample space)**.** The set of all possible outcomes of an experiment is called sample space, Ω.

Elements of Ω are called *elementary events* or atomic events[1].

Definition 1.2.2 (Event)**.** A subset A of Ω is called an event.

[1] to be precise, elementary events are not elements of the sample space, but *singletons*, that is subsets of Ω with only one element

Here are three examples.

Example 1.2.1 (Tax reductions)**.** You are a general contractor and are considering the government announcement that current tax reductions could be extended for one or two years. The experiment relevant for you is the extension of the reductions so

$$\Omega = \{\text{"no extension"}, \text{"one year extension"}, \text{"two years extension"}\}$$

and the elementary events "no extension", "one year extension", "two years extension". The event corresponding to the English sentence

$$A = \{\text{"the extension has been approved"}\}$$

is

$$A = \{\text{"one year extension"}, \text{"two years extension"}\}$$

Example 1.2.2 (Share price)**.** The experiment is to observe the price of a company share at a certain time. In this case, $\Omega = \{x \in R : x > 0\}$. Elementary events are all positive real numbers. The event "the price is less than 10" corresponds to the subset $\{0 < x < 10\}$.

Example 1.2.3 (Production)**.** The experiment consists in randomly choosing a piece in a given production lot. There are two production plants, S1 e S2, that produce the same number of pieces in the same time. Each piece can be perfect or faulty. If $C1$ is a perfect piece coming from plant 1 and $D1$ is a faulty piece from the same plant, and analogously for plant 2, we have

$$\Omega = \{C1, C2, D1, D2\}.$$

The event "the piece is faulty" corresponds to $\{D1, D2\}$.

As events are subset of Ω, we recall some set theory definitions about set operations. The following examples are related to the previous ones.

Definition 1.2.3. Ω, as a subset of itself is called *certain event*; the empty set, $\emptyset$, is called *impossible event*.

Events are subsets (of Ω) and therefore usual operations on sets can be extended to events.

Definition 1.2.4. The complement of an event A is the event that occurs whenever A does not occur and is denoted by $\bar{A}$.

In example 1.2.1, if B is the event "tax reductions have been extended for one year" then $\bar{B}$ is

$$\bar{B} = \{\text{"no extension"}, \text{"two years extension"}\}.$$

Definition 1.2.5. Given two events A and B, event $A \cup B$ is called union (or Boolean[2] sum) of events A and B and occurs if A occurs or B occurs or both.. Usually, the sentence referring to the union of events is made up connecting the two sentences referring to each event with "or".

[2]from George Boole, 2 November 1815 - 8 December 1864, an English mathematician and philosopher.

In example 1.2.3, if A is the event "the piece is faulty" and B is the event "the piece comes from plant 1" then $A \cup B = \{C1, D1, D2\}$

Definition 1.2.6. Given two events A and B, event $A \cap B$ is called intersection (or Boolean product) of events A and B and occurs if both event A and B occur. Usually, the sentence referring to the union of events is made up connecting the two sentences referring to each event with "and"., the sentence referring to the union of events is made up connecting the two sentences referring to each event with "and".

In the same production example, if A is "the piece is faulty" and B is "the piece comes from plant 1" then $A \cap B = \{D1\}$

Definition 1.2.7. Two events A and B such that $A \cap B = \emptyset$ are called *mutually exclusive*. More generally, n events, $X_1, \ldots, X_n$, are said to be mutually exclusive if $X_i \cap X_k = \emptyset$ for any $i \neq k$.

One can now give the abstract definition of probability[3]:

Definition 1.2.8 (Probability measure)**.** A function P that assigns a real number $P(A)$ to any event A is a probability measure if

1. $P(A) \geq 0$ for every A;
2. $P(\Omega) = 1$
3. if A and B are events such that $A \cap B = \emptyset$ then $P(A \cup B) = P(A) + P(B)$

From the previous definition, a number of important consequences follow:

Theorem 1.2.1. *If P is a probability measure and A is an event then:*

1. *$P(\bar{A}) = 1 - P(A)$;*
2. *$P(\emptyset) = 0$,*
3. *if A and B are events then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$; this equation extends the third axiom to the case of non mutually exclusive events.*
4. *if A implies B, i.e. $A \subset B$, then $P(A) \leq P(B)$*

A not rigorous proof of these results can be found using Venn diagrams.

1.3 Conditional probability

Any assessment of a probability is made using some information. With the available data, the decision maker assigns each event a probability. If the decision maker gets more, probabilities can change.

[3]we are consciously omitting the definition of an algebra of sets, given the introductory level of these notes.

Example 1.3.1. The contractor in example 1.2.1, based on past experiences, assigns a 40% probability to the event “tax reductions will be extended”. Then, he knows that a draft of the extension is currently being written and so he increases the probability of the event to 70%. The event “a draft of the extension is being written” changed the probability of the event “tax reductions will be extended” and the contractor, given the news, is likely to change his decisions.

Example 1.3.2. The owner of the production plants in example 1.2.3 started a quality assurance process. An engineer is sent to inspect production lots. Each lot is made up of the same number of pieces from both plants. The inspector randomly extracts one piece from a lot to assess its quality. With no other information, the probability that a piece is produced in plant 1 is 50%. Further analyses show that plant 1 produces 1 faulty piece out of 10 while plant 2, being newer than plant 1, produces 1 faulty piece out of 50. The engineer further inspects the piece and finds that is it faulty. Now the probability that the piece comes from plant 1 is greater than before, because in a single lot the production is the same for both plants but plant 1 is more likely to produce faulty pieces. The occurrence of the event “the piece is faulty” changed the probability of the event “the piece comes from plant 1”.

In both examples a change of probability assessments occurred: both subjects gave an initial estimate of the probability of a certain event (prior probability) and then, taking into account some new facts, changed the initial estimate into a new one (posterior probability. The new probability is called “conditional probability”.

Prior and posterior probability

Definition 1.3.1 (Conditional probability)**.** If $P(B) > 0$ the probability of A given that B has occurred, is called “conditional probability of A given B” and is equal to

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

Using the classical probability view, the previous formula is easily explained. If there are, say, N possible outcomes of the experiment, the fact that B has occurred reduces the number of possible outcomes to $N \times P(B)$, i.e. the number of outcomes favorable to B. Of those outcomes, only $N \times P(A \cap B)$ are favorable to A and to B at the same time. The result then follows as a classical computation of probability.

Example 1.3.3. Two dice are rolled. We would like to know the probability of the event “the sum of the dice is greater than 6” (event A). If we assume the dice are fair, the number of possible outcomes is 36, while those favorable to A are 21 (see Figure 1.1). Therefore $P(A) = 21/36 \simeq 58.33\%$. Now assume we know that “the first die shows three” (event B). This fact changes our estimate on the occurrence of A into $P(A|B)$. Now the possible outcomes are just six, and among them only three are favorable to A so $P(A|B) = 50\%$ (see Figure 1.2).

The following theorem is useful in many applications.

Theorem 1.3.1. *If A and B are events, then*

$$P(B) = P(B|A)P(A) + P(B|\overline{A})P(\overline{A})$$

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

Figure 1.1: Roll of two dice: 21 outcomes with sum greater than 6 (green); 15 remaining outcomes (black). The probability that the sum is greater than 6 is 21/36.

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

Figure 1.2: Roll of two dice: if the first shows 3, sample space is shrunk to the third row. There are 3 rolls with sum greater than 6 (green) and 3 with sum less or equal to 6 (black). The probability that the sum is greater than 6 *given that the first die shows 3* is 3/6.

Substituting the total probability theorem into the conditional probability formula we get a meaningful result, the Bayes theorem

Theorem 1.3.2 (Bayes)**.** *If A and B are events and $P(B) > 0$ then* Bayes' Theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\overline{A})P(\overline{A})}$$

Bayes' theorem "swaps" the roles of A e B in the conditioning of probability. In the LHS (Right Hand Side) B is given while in the LHS (Left Hand Side) A is. The denominator is actually the probability of B, written as the sum of conditional probabilities, accordingly to the total probability theorem 1.3.1.

To understand the meaning of Bayes' theorem, we can go back to example 1.2.3. After a faulty piece is found, the engineer would like to assess the probability that is comes from plant 1. If A is the event "the piece comes from plant 1" and B is the event "the piece is faulty", the unknown probability is $P(A|B)$.

The engineer knows two facts. Firstly, the two plants produce the same number of pieces. So, $P(A) = 0.5$ and $P(\overline{A}) = 0.5$, i.e. the probability that a piece comes from plant 1 is 50% and the probability that is does *not* come from plant 1, and therefore comes from plant 2, is obviously 50%.

The other known fact is the percentage of faulty pieces produced in the two plants. Plant 1 has a 10% failure rate so the probability of a piece being faulty given that it comes from plant 1 is 10%, i.e. $P(B|A) = 0.1$. Similarly, plant 2 has a 2% failure rate so the probability that a piece is faulty given that it comes from plant 2 is 2%, $P(B|\overline{A}) = 0.02$.

Using Bayes' theorem we can get a precise value for the unknown probability:

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\overline{A})P(\overline{A})} \\ &= \frac{0.1 \times 0.5}{0.1 \times 0.5 + 0.02 \times 0.5} \\ &= \frac{0.05}{0.06} \\ &= \frac{5}{6} \simeq 83.33\% \end{aligned}$$

Given an event B, the set of all the probabilities "given B" are a probability measure themselves and therefore, if we measure mutually excluding events, their sum must be 1: $P(A|B) + P(\overline{A}|B) = 1$. The denominator is the probability the a random piece is faulty, i.e. $P(B) = 0.06$

Here is another example showing a decision making problem.

Example 1.3.4. A hauler plans to buy a new set of trucks. Experience shows that the probability that a truck would be working one year after its purchase is 70%. If F is "the truck is working after one year", then $P(F) = 0.7$ and $P(\overline{F}) = 0.3$.

Before buying the trucks, the hauler has them tested by the drivers who will be giving a positive or negative opinion about the purchase. The hauler knows that:

- when trucks have been found working after one year, 80% of the times drivers gave positive opinions
- when trucks got damaged within a year, 90% of the times drivers gave negative opinions

If $+$ means "experts give positive opinion" e $-$ means "experts give negative opinion", we can rephrase the previous facts using the formulas $P(+|F) = 0.8$ and $P(-|\overline{F}) = 0.9$. We might also write, for example, that when trucks remained working after one year, 20% of the times experts failed in predicting their status, so $P(-|F) = 0.2$. Similarly, when dealing with trucks that would eventually be found broken, 10% of the times experts gave them positive opinion, i.e. $P(+|\overline{F}) = 0.1$.

If we now ask the experts about the new trucks, we already know their answer will modify our confidence in trucks' life.

We are to compute the probability of F given one of the possible opinions. If we get a positive opinion, then

$$\begin{aligned} P(F|+) &= \frac{P(+|F)P(F)}{P(+|F)P(F) + P(+|\overline{F})P(\overline{F})} \\ &= \frac{0.8 \times 0.7}{0.8 \times 0.7 + 0.1 \times 0.3} \\ &= \frac{0.56}{0.59} = 0.9491\ldots \end{aligned}$$

and therefore $P(\overline{F}|+) = 0.0508\ldots$

If, on the contrary, the opinion will be negative, then

$$\begin{aligned} P(F|-) &= \frac{P(-|F)P(F)}{P(-|F)P(F) + P(-|\overline{F})P(\overline{F})} \\ &= \frac{0.2 \times 0.7}{0.2 \times 0.7 + 0.9 \times 0.3} \\ &= \frac{0.14}{0.41} = 0.3414\ldots \end{aligned}$$

and therefore $P(\overline{F}|-) = 0.6586\ldots$

These are not surprising results, because experts are shown to be right most of the times. Therefore, if they give a positive opinion we expect the probability of the trucks to be working after one year to increase (from 0.7 to 0.9491). On the other side, in the case of a negative opinion we expect the same probability to decrease (to 0.3414).

Numeric results depend on the experts reliability.

For example, if we assume that when trucks are working after one year, experts always give a positive opinion, we have $P(+|F) = 1$ and, consequently, $P(-|F) = 0$. If we further assume that when trucks do not last for one year, experts give a negative opinion, we have $P(-|\overline{F}) = 1$ and $P(+|\overline{F}) = 0$. If we compute the probability in this new scenario, we get:

$$\begin{aligned} P(F|+) &= \frac{P(+|F)P(F)}{P(+|F)P(F) + P(+|\overline{F})P(\overline{F})} \\ &= \frac{1 \times 0.7}{1 \times 0.7 + 0 \times 0.3} = \frac{0.7}{0.7} = 1 \end{aligned}$$

which implies $P(\overline{F}|+) = 0$, and

$$\begin{aligned} P(F|-) &= \frac{P(-|F)P(F)}{P(-|F)P(F) + P(-|\overline{F})P(\overline{F})} \\ &= \frac{0 \times 0.7}{0 \times 0.7 + 0.9 \times 0.3} = \frac{0}{0.77} = 0 \end{aligned}$$

which implies $P(\overline{F}|-) = 1$. The previous statements can be informally translated into English by saying that "experts are always right": if they give a positive opinion, the truck will work; otherwise the truck will not work. This situation is actually impossible but it serves as an extreme case that can be approached when probabilities are close to 1.

If only one of the two probabilities equals 1, e.g. $P(+|F) = 1$, while the other remains as in the original definition, $P(-|\overline{F}) = 0.9$, we get:

$$\begin{aligned} P(F|+) &= \frac{P(+|F)P(F)}{P(+|F)P(F) + P(+|\overline{F})P(\overline{F})} \\ &= \frac{1 \times 0.7}{1 \times 0.7 + 0.1 \times 0.3} \\ &= \frac{0.7}{0.73} = 0.9589\ldots \end{aligned}$$

which implies $P(\overline{F}|+) = 0.0410\ldots$, and

$$\begin{aligned} P(F|-) &= \frac{P(-|F)P(F)}{P(-|F)P(F) + P(-|\overline{F})P(\overline{F})} \\ &= \frac{0 \times 0.7}{0 \times 0.7 + 0.9 \times 0.3} = 0 \end{aligned}$$

which implies $P(\overline{F}|-) = 1$. We know that when trucks are working, experts are always right ($P(+|F) = 1$) and give a positive opinion. However, when trucks are not working ($\overline{F}$) experts give negative opinions only 90% of the times, thus they give a positive (wrong) opinion in the 10% of the time. Thus, if we get a positive opinion we are not sure the trucks are going to work or not (but we expect the probability of them working to increase). On the contrary, if we get a negative opinion we are sure the truck is not going broken within a year.

The other somewhat strange situation happens when both reliabilities equal 50%, $P(+|F) = 0.5$ e $P(-|\overline{F}) = 0.5$. We have $P(-|F) = 0.5$ and $P(+|\overline{F}) = 0.5$ and therefore:

$$\begin{aligned} P(F|+) &= \frac{P(+|F)P(F)}{P(+|F)P(F) + P(+|\overline{F})P(\overline{F})} \\ &= \frac{0.5 \times 0.7}{0.5 \times 0.7 + 0.5 \times 0.3} = \frac{0.35}{0.5} = 0.7 \end{aligned}$$

which implies $P(\overline{F}|+) = 0.3\ldots$, and

$$\begin{aligned} P(F|-) &= \frac{P(-|F)P(F)}{P(-|F)P(F) + P(-|\overline{F})P(\overline{F})} \\ &= \frac{0.5 \times 0.7}{0.5 \times 0.7 + 0.5 \times 0.3} = \frac{0.35}{0.5} = 0.7 \end{aligned}$$

which implies $P(\overline{F}|-) = 0.3$. In this scenario, prior probabilities, $P(F) = 0.7$ and $P(\overline{F}) = 0.3$, equal posterior probabilities, $P(F|+) = 0.7$ and $P(\overline{F}|+) = 0.3$. The experts' opinion does not change our confidence in the trucks being working after one year. This is easily explained if we consider that an expert with a failure rate of 50% is not an expert at all! Anyone, by tossing a coin and giving, for example, a positive opinion when it shows heads and a negative one in the other case, has the same degree of reliability of an expert with a 50% probability of failure. Thus, previous results are another extreme case, that in which the given fact conveys no information at all. Typical applications fall in between these two extreme cases.

1.4 Discrete random variables

Usually the consequences of an experiment are expressed by a number.

Example 1.4.1. In example 1.2.1, it is obvious that the contractor's profit will depend on the decision the government will make. For example, we could assume that:

1. the probability that tax reductions will not be extended, A_1, is 10%, $P(A_1) = 0.1$;
2. the probability that tax reductions will be extended by one year, A_2, is 60%, $P(A_2) = 0.6$;
3. the probability that tax reductions will be extended by two years, A_3, is 30%, $P(A_3) = 0.3$;

and that

1. if reductions are not extended, contractor's profit will increase by 100000 euros;
2. if reductions are extended by one year, contractor's profit will increase by 200000 euros;
3. if reductions are extended by two years, contractor's profit will increase by 250000 euros;

Random variable

The experiment consists in the extension of reductions. The profit increase is a number X, called *random variable*, and is bound to the result of the experiment. The values X can take are usually denoted by the corresponding lowercase letter with an index: $x_1 = 100000$, $x_2 = 200000$ e $x_3 = 250000$. Thus, to every event there is an associated probability and a value of a random variable, summarized in the following table

Events	A_1	A_2	A_3
Probability (p_1, p_2, p_3)	0.1	0.6	0.3
Values of X (x_1, x_2, x_3)	100000	200000	250000

Whenever the random variable takes only a finite number of values, as in this case in which there are only three values, the random variable is called *discrete*. The table above is an example of a *discrete probability distribution*.

1.5 Expected value of a random variable

Definition 1.5.1 (Expected value)**.** If X is a discrete random variable and p_i is the probability that it takes the value x_i, $p_i = P(X = x_i)$, the expected value of X is

$$\mathrm{E}\,(X) = \sum_{i=1}^{n} x_i \times p_i$$

Example 1.5.1. In example 1.4.1 the expected value of the random variable X is

$$\mathrm{E}\,(X) = \sum_{i=1}^{3} x_i \times p_i = 100000 \times 0.1 + 200000 \times 0.6 + 250000 \times 0.3 = 205000$$

Note that, in general, the expected value of a random variable is *not* one of the values the variable can take. One can informally explain this definition assuming that the experiment could be repeated many times in the same conditions (i.e. with the same probability distribution). In this case, when the number of experiments is large, the mean value of the whole set of experiments, i.e. the sum of the values taken by X divided by the number of experiments that have been made, should be close to the expected value of X[4].

1.6 Problems

Problem 1.6.1. An investment company has just bought 10 millions of euros of buildings in downtown Milan. According to the company CEO, the probability that the value of the buildings will increase by 30% in the next two years is 0.3, while the probability that the value will increase by 10% is 0.5. Otherwise, the value of building will remain the same. What is the expected value of the buildings after two years?

Problem 1.6.2. A company is planning an incentive travel to a tropical island for its best customers. According to the company board the probability that the travel will be a success is 80%. Two members of the company board are going to go to the island before approving the planned travel. When they are back, they will give an opinion about the place and its facilities. Past experience shows that when the travel was a success, the testers gave a positive opinion 90% of the times. When it was a flop, the opinion was negative 80% of the times. Now, the two members are on their way back. How can the probability of a success change?.

1.7 Quickies

Question 1.7.1. A risky investment can result in a revenue of 1, 2 or 5 millions euros. Could it be that the expected value of the investment is 6 M euros? Why?

Question 1.7.2. Let $P(A|B) = 0.4$. Is it true, in general, that $P(A|\overline{B}) = 0.6$?

[4] From a theoretical viewpoint, this statement should be made more precise and could then be rigorously justified. The reader interested in this topic can see [4].

Question 1.7.3. Let $P(A|\overline{B}) = 0.4$ Is it true, in general, that $P(\overline{A}|\overline{B}) = 0.6$?

Question 1.7.4. Let A and B be two events with $P(A) = 0.4$ and $P(B) = 0.3$. Could it be that $P(A \cup B) = 0.8$?

Question 1.7.5. Given two events A and B we know that $P(A) = 0.4$ and A implies B. Is it possible that $P(B) = 0.6$?

Part II

Decision under uncertainty

Chapter 2

Introduction

2.1 Models for Managers

One of the first decision-making problems taught often at MBAs or management specialization course to introduce decision analysis is the following (it is now a classic). You are the owner of a racing team. It is the last race of the season, and it has been a very good season for you. Your old sponsor will remain with you for the next season offering an amount of $50000, no matter what happens in the last race. However, the race is important and broadcasted on television. If you win or end the race in the first five positions, you will gain a new sponsor who is offering you $100000, besides $10000 or $5000 praise. However there are unfavorable running conditions and an engine failure is likely, based on your previous data. You estimate the damage to a total of -$30000. In addition, it would be very bad for the image of you racing team to have an engine failure in such a public race.

What shall you do, run or withdraw?

When presented to professional decision-makers, some managers answer immediately that they should run, because it is their job. Are they right or wrong? It is completely right if your objective is "to run" (It's your job, after all). Maybe a more articulare reasoning would be that if you continue running and races don't go well, then you won't be able to run anymore, because you'll have no funds.

Now, is there a solution to this problem?

There is a set of problem which are much more complicated than this one. For example, you are planning a space mission for NASA and you wish insights on its risk and safety. Or you manage a vary large facilty and have to order the right amount of spare parts. Or you are a human resource manager and need to establish the proper set of incentives to your working team. Or you are a policy maker and wish to make decision on the policy to adopt for fighting climating change.

It is clear that solving it requires the right mix of managerial intuition, competence (i.e., knowledge of the specific aspects of the problem) and ability. However, these managerial skills are undoubtedly enriched by the utilization of a decision-support model. In fact, managers and decision-makers benefit from the utilization of decision-support models in virtually every field.

One of the learning objectives of this course is to help you in understanding

Figure 2.1: The process of decision making

what is a model, so that to lern how to use it. In other words, we cannot expect too much out of a model, but we cannot even expect too little.

2.2 The decision process

The modern decision-making process is iterative, as described in Figure .

The manager faces a problem. Then, she can rely on her intuition and come to a decision. This is probably the way you can follow in day-to-day practice. However, there is a set of problems in which the decision-maker feels the need to combine all her available information into a model. She then performs several analyses of the model and uses the results to make a (possibly) better informed decision.

The next quote is attributed to Frederick W. Smith, former CEO of the Federal Express Corporation: ”By modeling various alternatives for future system design, Federal Express has, in effect, made its mistakes on paper. Computer modeling works; it allows us to examine many different alternatives and it forces the examination of the entire problem”.

Now, we come to the question of what is a mathematical model for a manger. From a scientific viewpoint, a managerial model is not different to any scientific model. In this respect, a scientific model is an abstract representation of a real situation. Through abstraction we identify the essential elements of the problem and, when possible, associate these elements with mathematical laws. If the set of equations we come up with is solvable, then the solution we find becomes an important part of the information based on which we can make our decision.

The risk associated with the abstraction process is to overlook important elements of the problem. The ability of including the essential elements of the problem without overlooking marks the difference between a good model and a bad one. The abstraction process is in fact the actual creation of the model. It consists of the “translation” of real things (persons, firms, rules, etc.) into abstract mathematical objects (variables, functions, etc.). Every mathematical

Abstraction process

object is bound to the real thing it represents and this binding is crucial: it can provide useful insights to look into the results and to reduce the complexity of the problem. The more faithful the "translation", the more the model is close to reality, the closer the conclusions we derive from it are to the real problem. However, as with other types of translation, a "perfect translation", i.e. a perfect mathematical copy of the real world, does not exist. What we get is a set of several models, each of them being imperfect/limited when considered from some perspective.

The models described in these notes are optimization models. They are called *decision making problems*. The mathematical solution to these problem helps the decision-maker to figure out the course of action that allows her to come closer to reaching her goals.

Optimization models

Decision making problems

In the next section, we deal with the fundamental tools developed for dealing with decision-making under uncertainty.

Chapter 3

Influence diagrams[1]

3.1 Introduction

Influence diagrams (IDs) are graphical tools for the representation and solution of decision-making problems. By representation, one means the identification of the decision-making problem elements. In particular, an influence diagram reveals the probabilistic dependences among the uncertain quantities and the state of information at each decision stage. By solution, one means the determination of the preferred alternative (best strategy selection) given the state of information. Influence diagrams grant decision-makers the possibility of representing complex decision-making problems in a thorough albeit compact fashion. It is this strength over other representation techniques that have made the use of influence diagrams widespread.

An ID is a acyclic graph consisting of *nodes* and *arcs*. Nodes are of three types: decision, chance and value. Value nodes end the diagram. Decision nodes represent a decisions, namely the selection among alternatives operated by the decision-maker. The result of a decision-node is under the control of the decision-maker. Chance nodes (also called event nodes) represent the outcome of a statistical experiment (a random event) whose outcome cannot be controlled by the decision-maker: it is uncertain. The graphical representation of these nodes is displayed in Table 3.1. Nodes

We must emphasize here that IDs are a compact representation of the decision problem. Thus, nodes "hide" their content: the available alternatives in decision nodes or the possible outcomes of chance nodes are hidden below each node.

Example 3.1.1. A transportation company must move some goods using an airplane or a ship. The airplane is cheaper and faster than the ship. However, a just started volcano eruption is jeopardizing air transport because volcanic ashes could ground the airplane for days, forcing the delivery of goods to be delayed. In this case, both the cost of the stocking of the goods at the airport and the cost for the delay, could be greater than that of a ship transport, that is not affected by the eruption.

In this example we have 3 nodes: one to represent the uncertainty about the volcano; another to represent the decision that must be made ("ship or

[1] This chapter is largely based on [1].

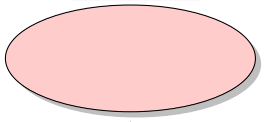
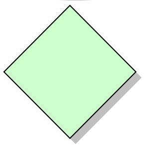
shape	node type
	Decision
	Chance
	Value

Table 3.1: Node types and corresponding shapes

airplane") and the third representing the final value of the operation. See Figure 3.1

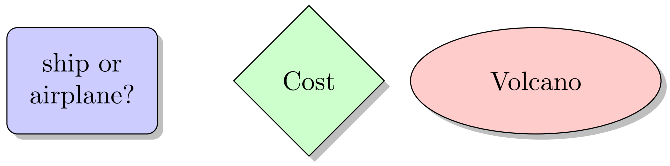


Figure 3.1: Nodes in example 3.1.1

Example 3.1.2. A physician must select the treatment for a patient. The first stage of the treatment foresees to choose between cures A or B. The two cures have a different efficacy, with their overall effect strongly dependent on the patient's response. After one week the physician re-evaluates the patient's conditions. Depending on the evaluation results, the physician has to decide between continuing with cure A, switching to B or resort to a third cure, C. The problem contains two (sequential) decisions.

The complete ID for this example is shown in Figure 3.2.

Decision nodes display the decisions to be taken at different stages of the decision analysis problem at hand. A variable contained in a decision node is under the control of the decision-maker, who selects the alternative that maximizes the decision-maker's preferences. In Figure 1, decision node "Cure A or B" represents the first selection between cures A and B, the node "Cure A, B or C" represents the selection between A, B and C. The second selection is made after re-evaluation of the patient's conditions.

Chance nodes represent variables or events whose knowledge or realization is out of the control of the decision-maker. Chance nodes are sometimes referred to as uncertainty or event nodes. Each chance node contains all possible realizations of the corresponding uncertain variable. Realizations are called outcomes. In Figure 3.2, the chance node "Patient Condition" represents the

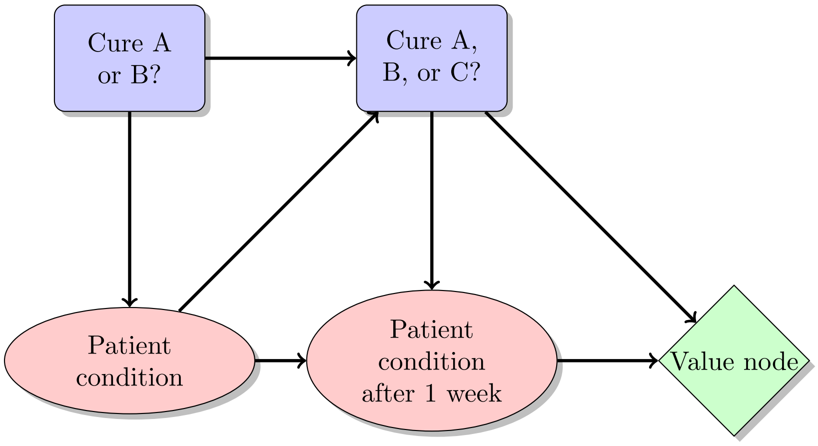


Figure 3.2: Influence diagram of example 3.1.2

conditions of the patient after selection of A or B. If the analyst/decision-maker considers that three possible states, namely “fully recovered”, “partially recovered” and “worsened”, are possible, then the chance node will have three outcomes. The decision-maker’s state of belief of the likelihood of the outcomes is characterized by a corresponding conditional probability distribution.

Value nodes contain the decision-maker’s utility for each consequence. A consequence is the end state of the world resulting as a combination of selected alternatives and outcomes of uncertain events. Utility is a quantification of preference and must be assessed consistently with the axioms of Decision Theory. Value nodes are occasionally referred to as utility nodes.

Arrows joining nodes in an influence diagram are called arcs. Arcs are grouped into the categories of informational, and conditional and functional.

Informational arcs are arcs ending in a decision node. If the informational arc stems from a chance node, then it indicates that the decision-maker is aware of the outcome of the chance node at the moment the decision is made. If an arc connects two decision nodes, then the decision-maker is aware of the previously selected alternatives. Arcs connecting decision nodes are also called no-forgetting arcs. Informational arcs imply time precedence. For this reason, it is not possible to reverse the direction of an informational arc.

Arrows ending into chance nodes are conditional arcs. Conditional arcs indicate the presence of a possible probabilistic dependence among the distribution of the random variables contained in the two chance nodes that the arcs link. We recall that probabilistic dependence is a weaker relationship than causal dependence. Let X and Y be the two random variables represented by the two chance nodes. Two cases are possible: X is probabilistically dependent on Y or it is not. If there might be a probabilistic dependence, this is displayed by the presence of an arc. The direction of the arrow shows the state of information of the decision-maker, i.e., whether the decision-maker is willing/capable of expressing the probabilities as $P(X|Y)$ or as $P(Y|X)$. A conditional arc does not necessarily correspond to time precedence. In fact, it is always possible to reverse a conditional arc using Bayes’ theorem, provided that this operation is performed in consistent fashion. The strong assertion is the lack of a conditional arc between two nodes: the random variables are independent.

Finally, functional arcs are arcs that enter the final node of the diagram, namely, arcs entering the value node.

The influence diagram in Figure 3.2 contains six arcs. The arc between decision nodes "Cure A or B" and "Cure A, B or C" is an informational arc. It denotes the fact that the physician, at the moment of the second decision, is aware of whether cure A or B has been previously selected. The arc from chance node "Patient Conditions" and decision node "Cure A, B or C" is also an informational arc, denoting the fact that the physician is informed of the patient's conditions after cure A or B has been adopted. The arc between the chance nodes "Patient Conditions" and "Patient Conditions after 1 week" is a conditional arc representing the fact that the decision-maker considers the outcomes of the second chance node, i.e., the conditions of the patient after the revision of the treatment has been undertaken, dependent on the conditions of the patient after the first cure A or B has been selected.

Arcs

In an ID nodes are joined by *arcs*. Arcs are the logical links between nodes. Arcs are always *oriented*, i.e. they go from a parent node to a child node, following the arrow on the arc. An arc can be:

- *informational*, if it ends in a decision node. The presence of an informational arc means that the entity represented in the parent is known at the time the decision is made;
- *conditional*, if it ends in an event node. The presence of a conditional arc means that the child event depend on the entity represented in the parent.
- *functional*, if it ends in a result node.

In example 3.1.1 there are two functional arcs. The cost is directly influenced by the outcome of the volcano node and, independently, by the decision that has been made. The new ID is drawn in Figure 3.3

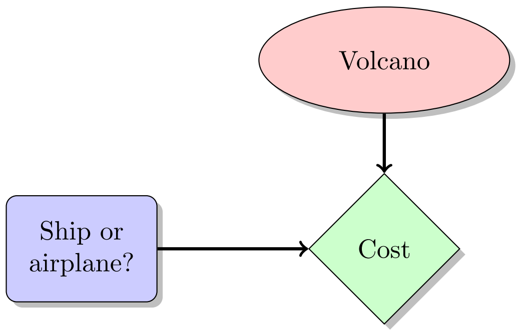


Figure 3.3: Influence diagram for example 3.1.1

For now, the ID in Figure 3.3 has no quantitative information. In fact, as we have seen, it refers to the first levels of IDs, namely graphical representation.

To solve the decision-making problem, we need to complete the diagram with the necessary numbers. These are:

- the numerical values of the outcomes in the end node (utilities, payoffs, costs) and
- the conditional probability tables or probability distributions, to be associated with each node.

Let us start with the a probability distribution of the chance node. This distribution is assigned by the decision maker (subjective view) based on her degree-of-belief about the outcomes of the events. For example, the decision maker thinks that volcanic ashes may ground airplanes with a 60% probability. Thus, the "Cost" node is associated with a conditional probability table with the cost for each transpotation means and each possible event. If airplane cost is 5000 euros normally and 12000 euros if the ashes ground planes and ship cost is 8000 euros, the new ID is that of Figure 3.4.

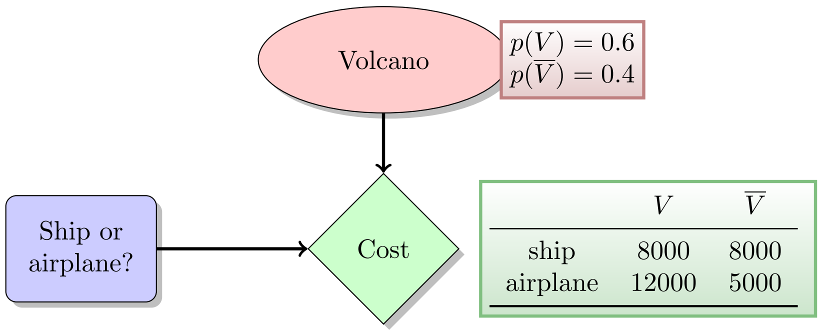


Figure 3.4: Influence diagram of example 3.1.1

3.2 Information

In example 3.1.1 the decision affects the cost of the transportation. The decision maker is in the worst position, being unable to control the uncertainty of the problem.

Let us assume that, before making any decision, the company takes a look at the predictions about the volcanic activity in the near future. We now must add another event node in the diagram, that of "Predictions": the decision maker takes note of the predictions and then decides. Thus, an informational arc must go from the "Prediction" node to the "Ship or airplane?" node. From a mathematical point of view, this changes the probabilities in the decision process. If P is, say, the event "Predictions say the volcano is going to erupt", the relevant probability will be $p(V|P)$. Predictions changed the information of the company and the probabilities must change accordingly.

Predictions are in turn affected by the volcanic activity: geologists study the behavior of the volcano and then give their view about the future, Therefore, there is a conditional arc going from the "Volcano" node to the "Predictions" node (see Figure 3.5).

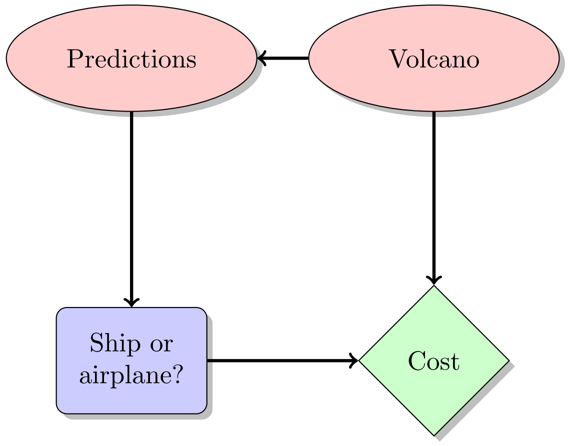


Figure 3.5: ID of example 3.1.1 with predictions (imperfect information).

Imperfect information

The ID in Figure 3.5 shows an *imperfect information*. Information (i.e. knowledge of the predictions) is imperfect because predictions do not rule out any of the possible outcomes of the future, but can decrease the randomness for the decision maker. A synonym for imperfect information is *sample information*, because in many cases predictions are based on sample observations. For example, to get a prediction about the customer satisfaction for a new product, a company can test the product with a small sample of people. The answer the company gets is not that of the whole market but is a hopefully good prediction of it.

If the company could wait for the volcano to erupt or not, the decision will be the best possible one. The corresponding ID would be that of Figure 3.6. In this case, the arc from the volcano to the decision is an informational one because in this scenario the company knows the outcome of the random experiment (the eruption of the volcano) in advance. In this case we have *perfect* information.

Perfect information

With perfect information, the decision process does not contain any randomness anymore and it reduces to an optimization problem. Even if the perfect information scenario does not actually happen, it is nonetheless interesting as an extreme case: the maximum advantage we can get from a sample information can not exceed what could be gained in the perfect information framework. We will return on this in chapter 4.

3.3 Iterated decisions

As a slight variation on the above situation, let us consider the position of a decision maker who has no available information about the volcano, but knows it can be obtained, at a certain cost. Under these circumstances, the decision maker has to decide whether or not to wait for the predictions. Given all the other data, this decision is obviously based on the reliability of the prediction. The more the one who makes the prediction is reliable, the more it is convenient to ask for predictions. Obviously, if the reward we get from the prediction is less than what we pay for it, it makes no sense to go that way. This extreme

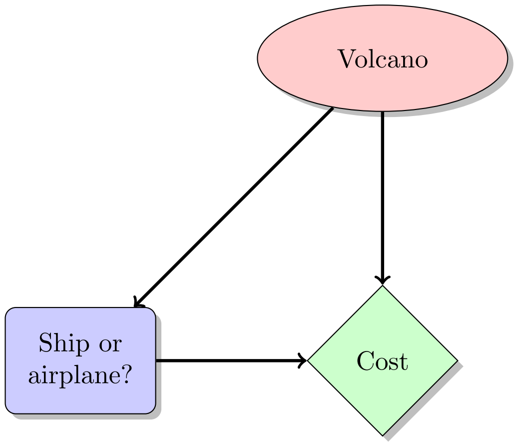


Figure 3.6: ID of example 3.1.1 with perfect information

case is that of perfect information, because no predictor could be better than the one who knows the future in advance. On the contrary, if the predictor is completely unreliable, it is useless to ask for predictions, and to pay for them. Actually, we are often between these two extreme cases.

The ID representing this case is that of Figure 3.7. The informational arc between the decision to ask for predictions and the decision on the means of transport brings an important consequence: the "Ship or airplane?" node is influenced by its parent, because the decision on the means of transport *follows* that on the predictions. Given the cost of predictions, the value of the "Cost" node depends also on the "Wait for predictions" node. Thus, we must add an informational arc between the two nodes (see Figure 3.7).

3.4 Additional information on ID

One distinguishes three levels of influence diagrams: graphical, functional and numerical. The graphical level displays nodes and arcs, evidencing probabilistic dependence and the flow of information, i.e., the information available before each decision (Figure 3.2). At the functional level one introduces the outcomes, the conditional distributions and the alternatives of each chance and decision node, respectively. The numerical level is the level at which the values of the conditional probabilities and utilities are inserted. The insertion of the numerical values is necessary for the solution of the decision-making problem.

Significant research on the solution of influence diagrams has then been undertaken in the fields of computer science and operations research, and numerous algorithms have been developed improving the efficiency of the original ones. The availability of algorithms has made available commercial software that allows decision-makers to implement and solve decision analysis problems through their representation in the form of influence diagrams. In the original work of Ronald A. Howard and James E. Matheson the solution of influence diagrams is envisioned in a two step approach, through conversion of the influence diagram into the corresponding decision tree. A few years later, Ross

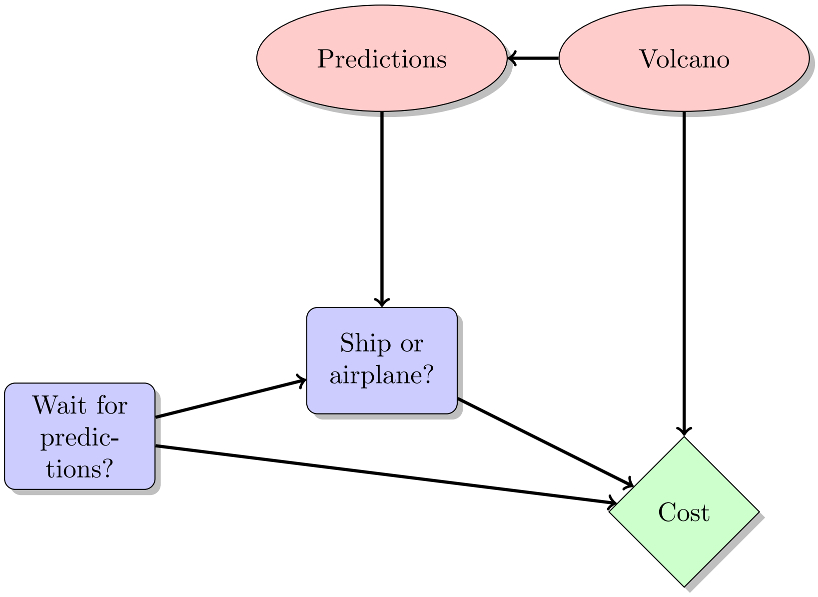


Figure 3.7: Influence diagram of example 3.1.1 with the decision on waiting for predictions

D. Shachter proposes the first algorithm for direct solution of influence diagrams. The direct solution algorithm proceeds through arc reversals and node elimination. These operations follow strict rules that allow not to distort the calculation of the expected utilities and the flow of information. Such operations are called value-preserving. The four main types can be listed as follows, according with the taxonomy of Joseph A. Tatman and Ross D. Shachter: arc reversal, chance node removal through summation, chance node removal by conditional expectation and decision-node removal by maximization. The procedure foresees first the removal of barren nodes, followed by the iterative application of the four operations, until the best strategy is identified.

Influence diagrams are often utilized in conjunction with decision trees. Some commercial software allows users to first structure the model in the form of an influence diagram and then to obtain the corresponding decision trees. Decision trees allow to display the combinations of choices and outcomes that lead to each consequence, thus providing a detailed description of the decision-making problem. However, their size increases exponentially with the number of nodes. Not all influence diagrams can be directly converted into a decision tree and to one influence diagram there can correspond more than one decision tree. The conditions that assure the possibility of transforming an influence diagram into a decision tree are the single decision maker condition and the no forgetting condition. The reader interested in these topics is referred to [1].

3.5 Problems

Problem 3.5.1. For the next Fall–Winter season, a fashion company is setting up the price list within the end of August. The main competitor will publish its list price at the end of July. After both lists are published large-scale retailers' orders will determine the final profit.

Draw the ID for this scenario.

3.6 Quickies

Question 3.6.1. How many types of node can appear in an ID?

Question 3.6.2. Can an informational arc end in an event node?

Question 3.6.3. Can a conditional arc end in an event node?

Question 3.6.4. An arc starts from an event node A and ends into an event node B. What can be said about the probabilities in node B?

Question 3.6.5. You must decide whether or not to buy shares. The market is highly unpredictable. Draw an ID about this scenario.

Chapter 4

Decision trees

4.1 Introduction

A *decision tree* (DT) is one of the graphical tools for the representation and solution of decision-analysis problems. Their are well known for their immediacy, namely, their capability of displaying in a straightforward way all the elements of a decision-problem. However, as they display all possible combinations of decisions and events present in a decision-problem, their size grows exponentially with the number of element in the diagram.

From a mathematical viewpoint, a DT is a graph with 3 types of nodes: Nodes

- decision nodes
- event nodes
- result nodes

connected by branches. Branches emanating from decision nodes represent the alternatives available to the decision maker. Branches originating in an event node represent the possible outcomes of a random experiment. No branch starts from any result node: result nodes are "end" nodes.

Every node has an associated value, according to the rules that also apply to ID nodes. Each that will be explained below. A decision-node corresponds to a selection, namely, a max operation. The decision-maker selects the alternative that maximizes her utility. A chance node is characterized by an expectation operation, namely, we will need to compute its expected value for solving the diagram. Thus, chance nodes require that each branch is associated with a probability to be assigned by the decision-maker.

Nodes in a DT are represented by different shapes, according to Table 4.1.

A DT must be read from left to right and from top to bottom.

Example 4.1.1. A consumer electronic company must decide whether to develop a new product or not. In the next quarter, with no new product, revenues are expected to be 350000 euros, with production costs at 50000 euros. If the new product is developed, production costs raise to 120000 euros and revenues depend on the success of the new product. If the product is successful, which is 70% likely, revenues are expected to be 500000 euros. If the product is not successful revenues will be 300000 euros.

shape	node type
rectangle	Decision
circle	Event
triangle	Result

Table 4.1: DT: node types and their graphical representation

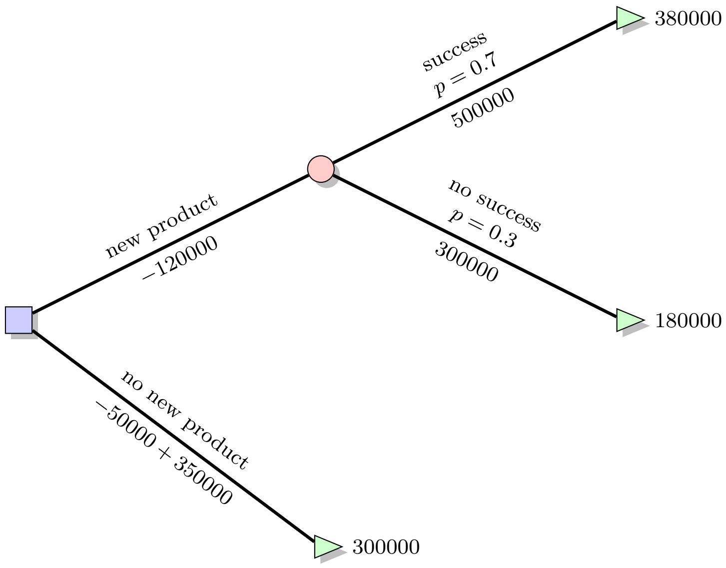


Figure 4.1: DT of example 4.1.1.

There are 5 nodes in this example: one is the decision node ("develop?"), one is the event node ("success?"), and the remaining three are the result of the decisions in the three cases (no new product, new successful product, new unsuccessful product). The event node is linked to the decision about the development. In Figure 4.1 we put all the relevant data. To compute the value of result nodes, values on the arcs have been added. For example, 380000 is the sum of the "new product" arc and the "success" arc.

To evaluate a DT we must compute the value of all intermediate nodes in the tree and then compute the value of the root node. This is done by "folding back" (or "roling back") the tree. The algorithm, in fact, starts from result nodes and ends at the root node. The value of a node is given by the following rules:

- the value of an event node is the expected value of its outcomes

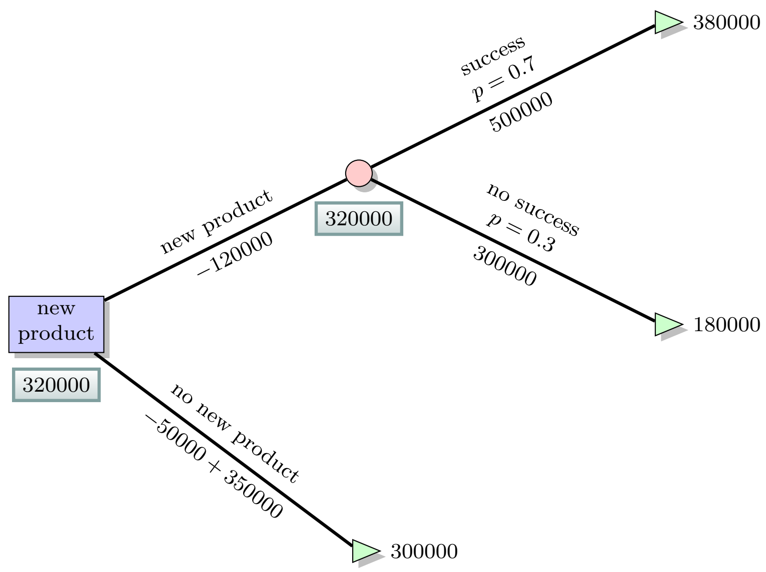


Figure 4.2: DT of example 4.1.1 with numeric values.

- the value of a decision node is the maximum value of the nodes of its options

In example 4.1.1 there's only one event node with two outcomes: success, worth 380000 with probability 0.7, and no success, worth 180000 with probability 0.3. Thus, the event node is worth

$$380000 \times 0.7 + 180000 \times 0.3 = 320000$$

Now we can compute the value of the decision (root) node. The decision has two options: "new product", worth 320000, and "no new product", worth 300000. Therefore, the decision node is worth 320000 and the decision to be made is to develop the "new product" (see Figure 4.2).

In a tree with decision nodes and result nodes only, the decision is simple: choose the maximum among all the alternatives. In this case, one is dealing with a problem with decision under "certainty". One simply chooses the preferred alternative. A typical case is the choice of the taste of an icecream or, when you go shopping, the purchase of, say, a new purse. It might require a while to think of the icecream-taste or of the type of purse to buy, but it is only a question of preferences. Provided that you have the necessary amount of money, there is no doubt that the icecream will be of the taste you ordered.

Conversely, if there are event nodes, you are in the presence of uncertainty. To evaluate the problem, the theory states that the decision maker has to convert the uncertainty associated with the event represented by the node into a number. This is done by using the expected value of the alternatives. Thus, the decision-maker is selecting the decision that maximizes her "expected" payoff (or, in a more general sense, utility.)

Here, it is important to recall that the expected value is not necessarily a value of one of the alternatives, i.e. the event node value is generally not one of the possible outcome of the experiment. In the example above, there is no way to get a profit of 320000 euros: either we get 380000 or 180000 euros. A possible meaning for the expected value is this: if we were to repeat the decision many times in the same conditions on the average we would get 320000 per decision.

4.2 Sensitivity analysis

A mathematical model can be generally seen as a series of mathematical objects (equations, inequalities, etc.) that processes a set of variables (called exogenous) to produce an output. The output is called endogenous variable or decision-support criterion, depending on the applications. When the exogenous variables are assigned numerical values, they become "parameters" of the model. The value of the model output changes if we change the values of the parameters.

The above-considerations hold for generic models and decision-support tools. IDs and DTs are part of this family and make no exception. In example 4.1.1 the values for cost, revenues and the probabilities are numbers given by the decision maker. Using these numbers, a well defined algorithm produces the final result.

However, in real world situations, it is very unlikely that decision-makers can assign a certain value to all exogenous variables. Most of the values that appear in a mathematical model are known within a given tolerance. For example, in assessing the revenues of the new product the amount of 500000 euros could be the base case of estimates that range from 100000 to 750000. How do we deal with this variability?

The answer is to perform *sensitivity analysis*. Sensitivity analysis tools are a set of methods that have been developed in the management sciences and in the scientific literature to help decision-makers making the most out of their mathematical models.

The sensitivity analysis methods we are going to examine in this course respond to some of the questions one can ask to the model. They have, therefore, the purpose of providing an introduction to sensitivity analysis. Many recent and much more sophisticated methods are available.

The first method we examine is called "one way" sensitivity analysis. It consists of inspecting the output of the model as one-factor-at-a-time is varied within its variation range. The remaining factors are held fixed at their base case value.

This type of analysis points out the boundaries within which the solution is still the best and, if the solution changes, the way it changes.

We proceed this way. First, the estimate of the revenues for the new product becomes a variable r. If r is greater than 500000 euros then it will be optimal to develop the new product. However, if r is small enough the optimal choice would be not to develop the product because revenues will be less than the increased cost (120000 euros). The expected value as a function of r is

$$(r - 120000) \times 0.7 + 180000 \times 0.3 = 0.7r - 30000$$

To develop the new product we need this quantity to be greater than that on

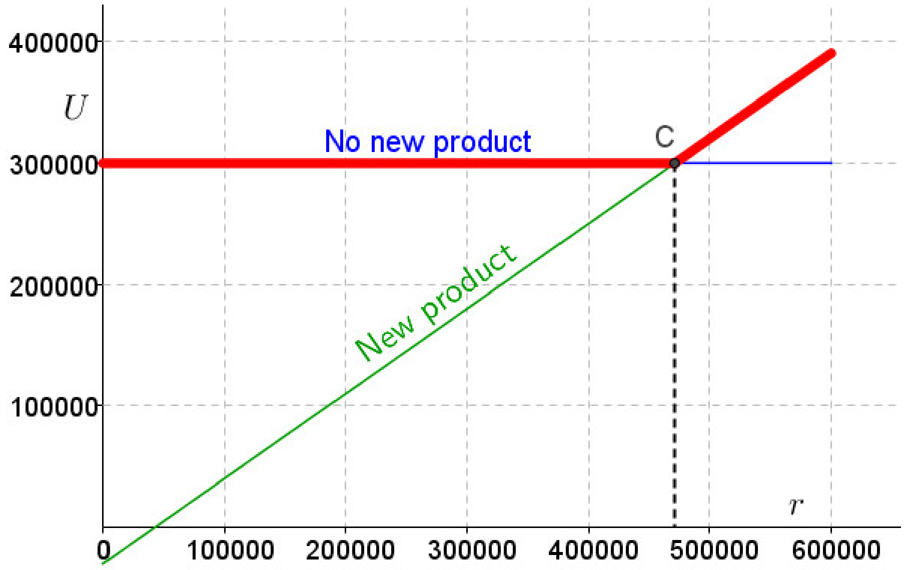


Figure 4.3: The red line represents, in the interval $0 \leq r \leq 600000$, the decision value with respect to the revenue r.

the other branch of the tree, which is 300000, i.e.

$$0.7r - 30000 > 300000.$$

We finally get $r > 471428$ euros. Thus, even if with predicted revenues of 500000 euros it is optimal to go for the new product, a decrease of less than 30000 euros changes the optimal choice to keeping the status quo. The plot in figure 4.3 shows the problem.

In a similar line of thought, we could consider the probability of success as a variable (leaving other parameters untouched). In the original form of the problem, the probability is 70%. We expect the new product to be optimal if that number is close enough to 1. If the probability of success is small the expected values of revenues will be small too and they will be less then the cost of the development. If p is the probability of success then the value of the event node is

$$380000p + 180000(1 - p) = 180000 + 200000p.$$

To be greater than the other alternative, we should have

$$180000 + 200000p > 300000$$

i,e, $p > 0.6$. If the probability of success exceeds 60% it is better to develop the new product; otherwise, one the company should keep on with the current production. Also in this case the plot in figure 4.4 shows the problem.

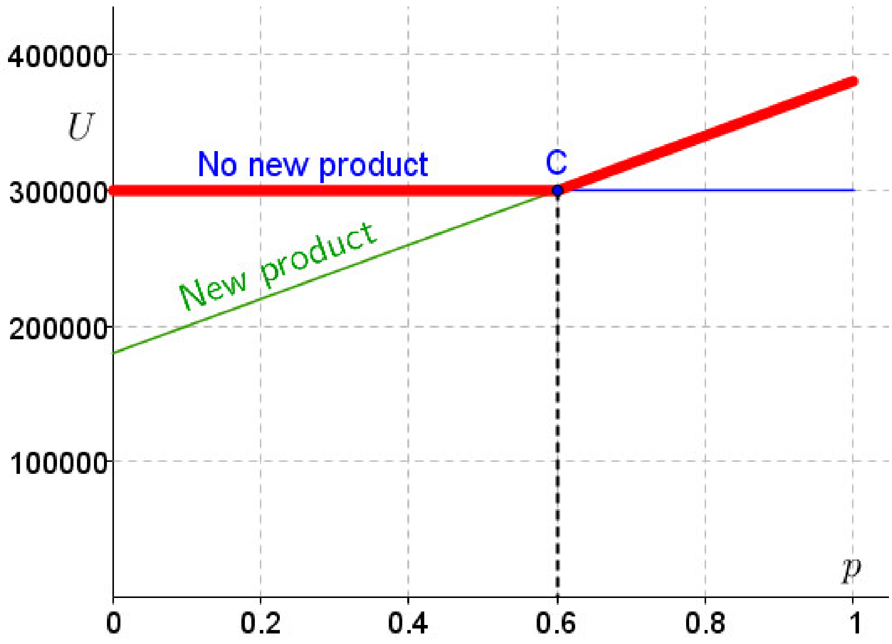


Figure 4.4: The red function represents, in the interval $0 \leq p \leq 1$, the decision value with respect to the probability p.

To summarize

> With the data in the original problem, it is optimal to develop a new product with an expected profit of 320000 euros. In addition:
>
> - if, keeping the other parameters unchanged, the probability of success changes but remains over 60%, the optimal solution is the same as before;
> - if, keeping the other parameters unchanged, estimated revenues change but remain over 471428 euros, the optimal solution is the same as before;
> - in all other cases, a more detailed analysis is needed.

To go on with the analysis, we should consider the variations of the two parameters at the same time. A further insight could be the extension of this analysis to all the parameters in the model. These tasks are computationally complex and can be done only with the help of a computer and a dedicated software.

4.3 Information

An integral concept of any decision analysis is the value of information. The decision-maker might (and should) ask the question of what elements in the decision-making problem are worth further investigation and/or if there is some essential information she should get for making a "better-informed" decision.

Maybe the arrival of new evidence can make her change her mind about the problem. However, information has a cost and she might be willing to pay a maximum amount in order to collect the new evidence. How much should she afford?

The answer to this question is obtained by appraising the value of information. We are going to discuss this concept, as usual, by means of an example. However, before starting, we state a short premise. Information is obtained by consulting a source. This source can be perfectly reliable or imperfect. In the first case, one talks about the expected value of perfect information (EVPI), in the second by the expected value of sample information (EVSI). As we are to see, EVPI can be seen as a special case of EVSI. Also, EVPI should always be greater or equal to EVSI: we are willing to spend more for perfect than for imperfect information.

Example 4.3.1. Let us assume that the decision-maker wishes to get additional information on the acceptance of the new product with a market research test. Let S be the event "the product is successful" and T be the event "the research test was positive". Let us assume that each time the product was successful, the test gave a positive answer with probability 80% while when the product was not successful, the test gave a negative answer with probability 90%, that is

$$P(T|S) = 0.8 \qquad P(\overline{T}|\overline{S}) = 0.9$$

and consequently

$$P(\overline{T}|S) = 0.2 \qquad P(T|\overline{S}) = 0.1;$$

Moreover, we know that

$$P(S) = 0.7 \qquad P(\overline{S}) = 0.3.$$

The decision tree in the new configuration is that of Figure 4.5.

To compute the values on the tree we must compute the relevant probabilities. From theorem 1.3.1 we have

$$p(T) = p(T|S)p(S) + p(T|\overline{S})p(\overline{S}) = 0.8 \times 0.7 + 0.1 \times 0.3 = 0.59$$

and

$$p(\overline{T}) = p(\overline{T}|S)p(S) + p(\overline{T}|\overline{S})p(\overline{S}) = 0.2 \times 0.7 + 0.9 \times 0.3 = 0.41$$

$(p(T) + p(\overline{T}) = 1)$. From theorem 1.3.2 we have

$$p(S|T) = \frac{p(T|S)p(S)}{p(T)} = \frac{0.8 \times 0.7}{0.59} = \frac{56}{59} = 0.9491\ldots$$

and

$$p(S|\overline{T}) = \frac{p(\overline{T}|S)p(S)}{p(\overline{T})} = \frac{0.2 \times 0.7}{0.41} = \frac{14}{41} = 0.3415\ldots;$$

we can now compute

$$p(\overline{S}|T) = 0.0509\ldots \qquad p(\overline{S}|\overline{T}) = 0.6585\ldots$$

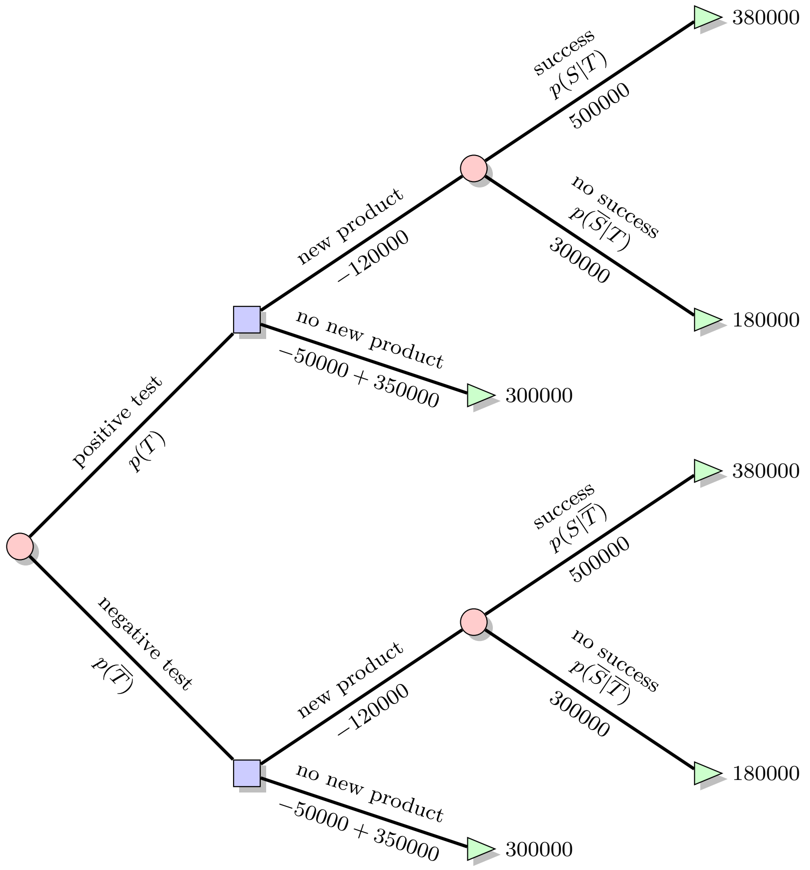


Figure 4.5: DT of example 4.3.1.

All the probabilities are known and we can finally compute the value of the tree as in Figure 4.6.

The new value for the root node is 341193.8 euros, greater than that of 4.2. In fact, in the new scenario the test research gave information about the new product and this information increased the value of the entire tree. The increase is measured by the difference between the root nodes: $341193.8 - 320000 = 21193.8$. This is the *expected value of sample information*.

Expected value of sample information

The information is partial, because the test is reliable but could give incorrect answers($P(T|S)$ and $P(\overline{T}|\overline{S})$ are both less than 1). The perfect information scenario would be that in which the decision about the production logically follows the completely predictable outcome. It is as if a magician could give a perfect view of the future and, based on the magician information, the decision maker solves the problem. The perfect information DT is shown in Figure 4.7.

The perfect information value of the root node is 356000 euros and to get it we need perfect information. Thus, this is the maximum we can get for this problem. The difference between the root node in the original problem, $356000 - 320000 = 36000$ is the *expected value of perfect information*.

Expected value of perfect information

If information is free, it is obvious that we should include it in the tree

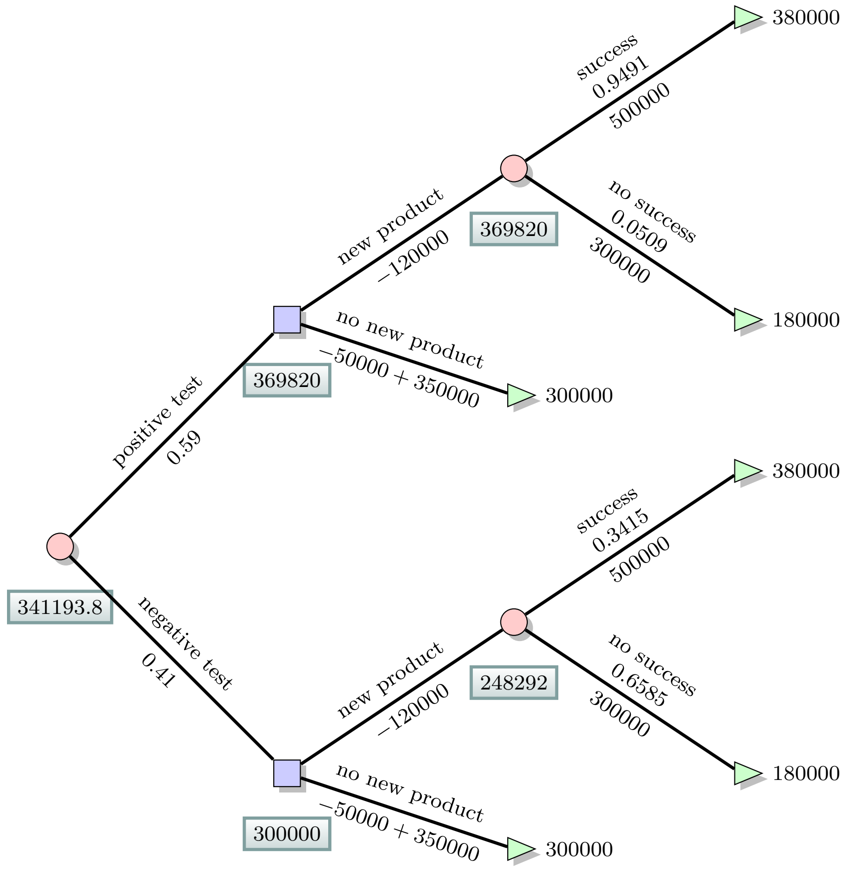


Figure 4.6: AD of example 4.3.1 with sample information.

and increase the value of the root node. If information is available at a cost, it makes sense to compare its cost with the benefit we get from it. If, for example, the cost of the market research in example 4.3.1 is 15000 euros, then we should do the research because there is a 21193.8 euros increase in profits, so the balance is positive. If there are more sources of information, each of them should be compared with its cost and we expect that the more expensive they are the more they are reliable, as the give a greater profit.

Even if the cost of information is not known, the decision maker can establish an upper bound for that cost. As the value of perfect information is 36000 euros, any information source that is more expensive than that should not be considered, because the cost would be greater than the benefit.

4.4 Problems

Problem 4.4.1. A pharmaceutical company has a patent for a molecule that could be effective in treating a certain disease. The cost of research for that molecule is 10 millions euros. The probability that the molecule is actually effective is only 10% but if it is so, revenues will grow to 200 millions euros.

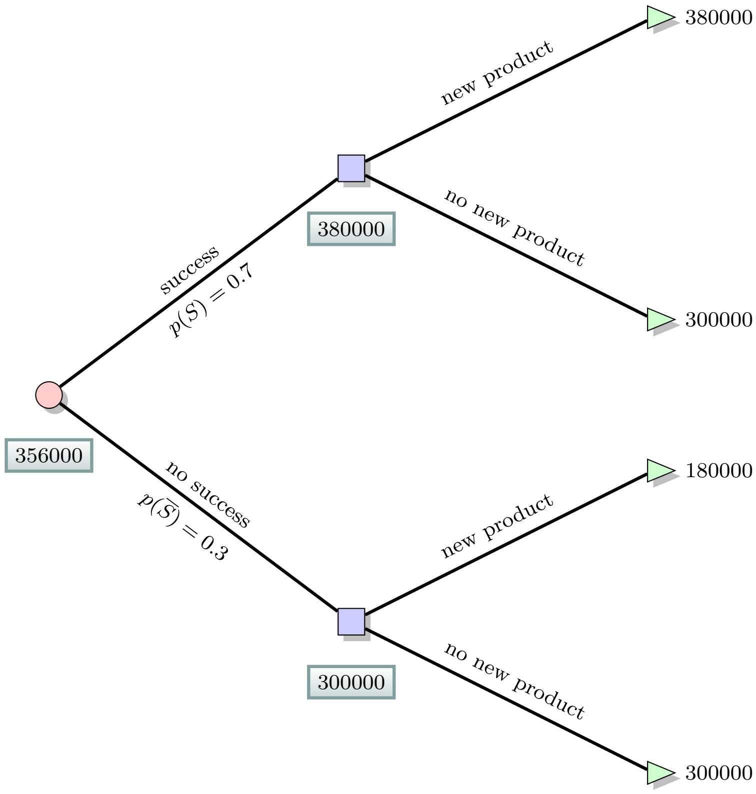


Figure 4.7: DT for example 4.3.1 with perfect information.

The other option is not to do any research with no cost and no revenues.

1. Draw a decision tree and find the best decision;
2. compute the value of perfect information;
3. before making a decision the company can ask an expert about the healing properties of the molecule. The expert can give a positive or negative opinion. If the molecule is successful, the probability that the expert will give a positive opinion is 95%, while if the molecule is not effective, the probability of a negative opinion is 85%. Draw a new decision tree including the expert opinion and the value of sample information. If the expertise costs 7 millions euros, should we ask for it?

Problem 4.4.2. A company is trying to decide whether to bid for a certain contract or not. They estimate that merely preparing the bid will cost 10000 euros. If their company bid then they estimate that there is a 50% chance that their bid will be put on the "short-list", otherwise their bid will be rejected. Once "short-listed" the company will have to supply further detailed information (entailing costs estimated at 5000 euros). After this stage their bid will either be accepted or rejected.

The company estimate that the labor and material costs associated with the contract are 127000 euros. They are considering three possible bid prices, namely 155000, 170000 and 190000 euros. They estimate that the probability of these bids being accepted (once they have been short-listed) is 0.90, 0.75 and 0.35 respectively.

What should the company do and what is the expected monetary value of your suggested course of action?

Problem 4.4.3. A call center employee spends part of his work time betting on the Internet, instead of contacting potential customers. This way the company looses 500 euros per week. The call center manager has the right to fire him without justifications. In this case, the company will get back the lost revenue.

Another opinion is to wait for the employee to be caught in the act. As this is a violation of the job contract, besides being fired, the employee must refund (and the company gains) 400 euros.

Server logs show that the employee bets for 60% of his work time so the probability of finding him playing is 0.6.

- Draw an ID for the problem.
- Draw the corresponding DT.
- What should the manager do? To wait o fire the employee immediately?

A software allows to log all the Internet connections going through the call center. According to the documentation, the software could provide a solid evidence against the employee.

- Draw a decision tree and find the maximum price to pay for such a software.

Further investigations show that, for privacy concerns, the above software is illegal. The manager, therefore, could interview some of the employee's colleagues to get an idea of his future behavior. However, time lost in interview is a loss (150 euros) and the colleagues opinions are, obviously, biased. If the employee is actually betting, they will tell the truth only 50% of the times, while if he does not bet they will be truthful 95% of the times.

- Should the manager interview the employee's colleagues?

4.5 Quickies

Question 4.5.1. An event node has two alternatives, worth 10 and 20 (with given probabilities). Can the node be worth 30?

Question 4.5.2. An event node has two branches: 10 with probability p and 20 with probability $1-p$. How much is p if the value of the node is 16?

Question 4.5.3. Considering the DT in Figure 4.8 can we point out the optimal decision?

Question 4.5.4. Considering the DT in Figure 4.9 can we point out the optimal decision?

Question 4.5.5. Considering the DT in Figure 4.10 can we point out the optimal decision?

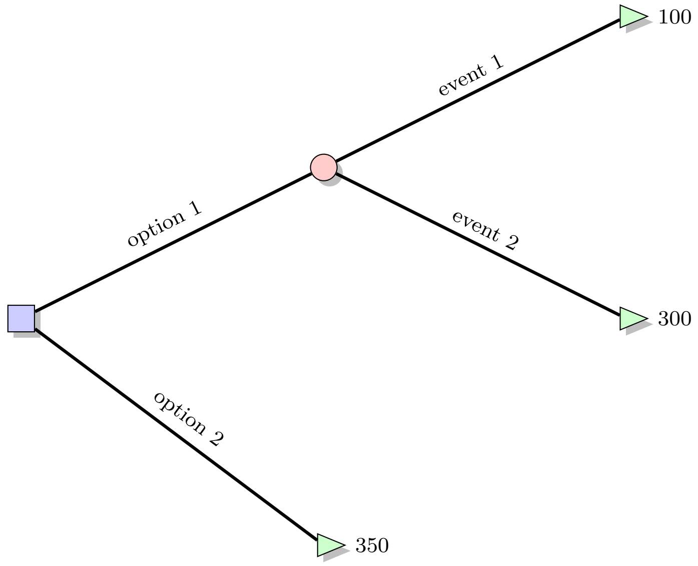


Figure 4.8: DT for quicky 4.5.3.

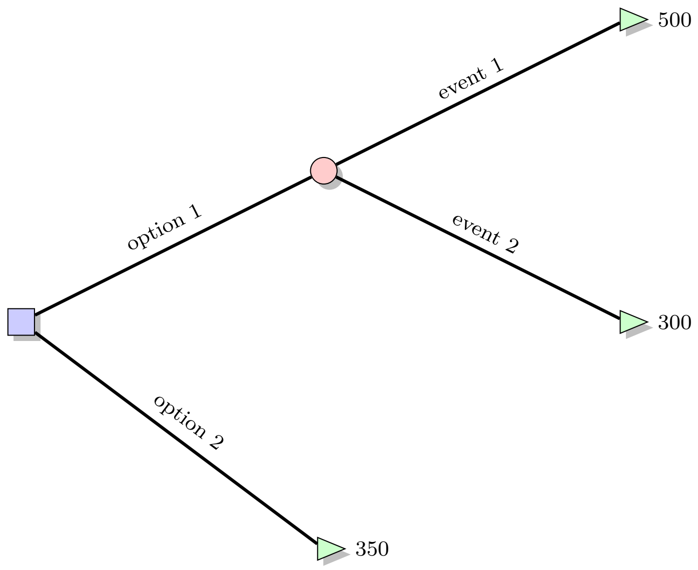


Figure 4.9: DT for quicky 4.5.4.

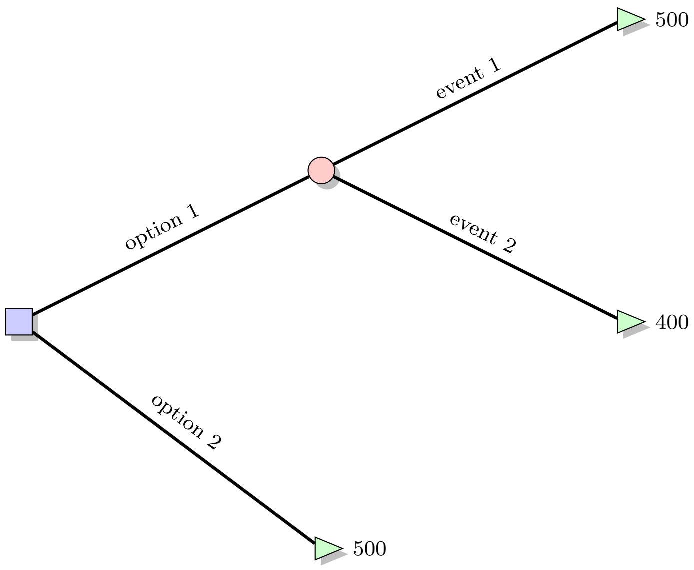


Figure 4.10: DT for quicky 4.5.5.

Chapter 5

Utility

5.1 Risk attitude

Let us consider the following example.

Example 5.1.1. A person has the opportunity to change his job. If he changes his job there is a 50% chance that the new job will be highly rewarding and his annual income will increase by 30000 euros; otherwise, he will get the same salary he gets now. If he does not change his job, his salary will increase by 10000 euros.

This situation is represented in Figure 5.1.

If we compute the value of the root node using the expected value we find the optimal choice is to change job. However, a real person facing this opportunity might take a different approach. To analyze all possible attitudes, let the certain salary increase (the one that he gets if he does not change) be

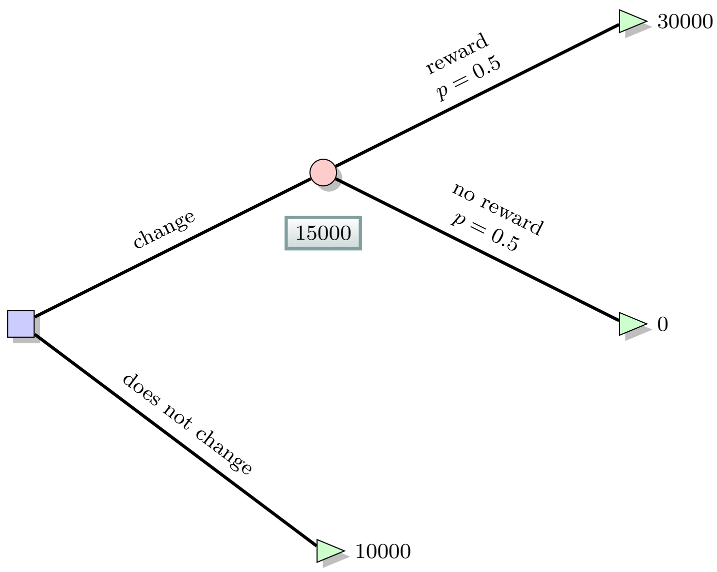


Figure 5.1: DT for example 5.1.1.

x. To make a further abstraction, let us assume this is a lottery (i.e. a random variable) where the prize is 30000 euros and the probability of winning is 50%. For what value of x the decision maker is indifferent between x and the lottery?

The answer depends on the decision maker. Let us assume there is a person, A, for whom $x = 12000$. A judges equally to get 12000 euros or to participate in the lottery. A knows that the expected value of the lottery is 15000 euros, but is willing to waive $15000 - 12000 = 3000$ euros to get 12000 euros *for sure*. Using a proper mathematical terminology, 12000 euros are the *certainty equivalent* of the "lottery". As the lottery is, in fact, a random variable X, the certainty equivalent is usually denoted by a function of X, $z(X)$.

Certainty equivalent

risk aversion

A is *risk averse*. Facing an uncertain alternative with expected value 15000 euros, A would rather give up and "pay" 3000 euros to change the uncertain option with the certain amount. 12000 euros is the minimum price for which A would sell the ticket of the lottery. The difference between the expected value of the lottery and its certainty equivalent is the *risk premium*. If the risk premium is positive, the decision maker is risk averse.

Risk premium

Another person, B could think that $x = 20000$. This person "loves" uncertainty: to have him sell the lottery ticket we need 5000 euro more than the lottery expected value. B is said to be "risk seeking". For a risk seeking person, the risk premium is negative.

Risk seeking

Risk neutrality

A third person, C, for whom $x = 15000$ would be called *risk neutral*. C's risk premium is 0. C is a decision maker who completely agrees with the expected value assessments.

So, from a risk attitude viewpoint, the analysis in the previous chapters was done as if we were risk neutral people. Psychological and economical studies show that many people are risk averse, especially in business matters. Changing the point of view could possibly change the optimal decisions. In the previous example, A would rather get 12000 euro and stay with the company instead of changing his job. In the 20th century, von Neumann and Morgenstern proposed the expected utility theory.

5.2 Expected utility

In the expected utility theory the subject must be able to order a set of random variables according to his preferences. We will assume these random variables take values in an interval $[a, b] \subset \mathbb{R}$. If the subject follows some basic requirements, known as *von Neumann and Morgenstern axioms* (vNM), it can be shown that it exists a positive function $u : [a, b] \to \mathbb{R}$ such that a random variable X is preferred to Y if and only if $\mathrm{E}(u(X)) > \mathrm{E}(u(Y))$. The function u is called *utility function*, and depends on the person making the judgments. The number $\mathrm{E}(u(X))$ is called *expected utility of X*.

Everyone who follows vNM axioms has an associated utility function that describes his preferences: faced with two alternatives, *the decision maker chooses the option with the greatest expected utility*. The principle guiding the choice is that of *maximization of expected utility criterion* (MEU). Under very unrestrictive hypotheses, u is a strictly increasing continuous function in $[a, b]$ and thus has inverse. The inverse function, u^{-1}, is strictly increasing and therefore

$$\mathrm{E}(u(X)) > \mathrm{E}(u(Y)) \quad \text{if and only if} \quad u^{-1}(\mathrm{E}(u(X))) > u^{-1}(\mathrm{E}(u(Y)))$$

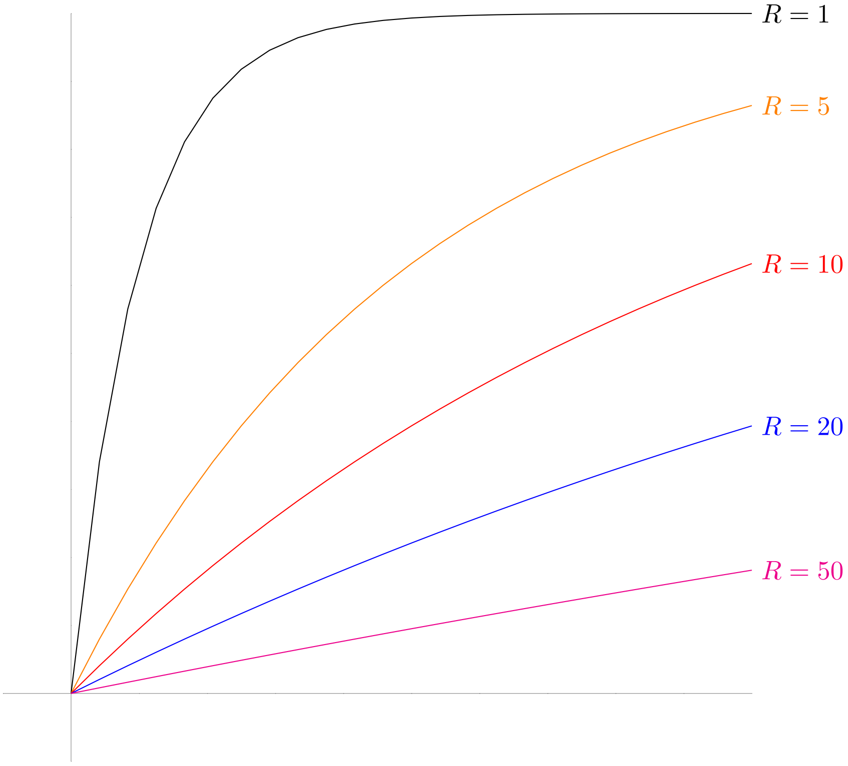


Figure 5.2: Exponential Utility with different values of R.

As for any random variable X the certainty equivalent $z(X)$ is $z(X) = u^{-1}(\mathrm{E}\,(u(X)))$, we can rewrite the previous equation in this way:

$$\mathrm{E}\,(u(X)) > \mathrm{E}\,(u(Y)) \quad \text{if and only if} \quad z(X) > z(Y). \tag{5.1}$$

The MEU criterion is equivalent to maximizing the certainty equivalent (MCE). If the subject is risk neutral we have $\mathrm{E}\,(X) = z(X)$, and the above equation becomes

$$\mathrm{E}\,(u(X)) > \mathrm{E}\,(u(Y)) \quad \text{if and only if} \quad \mathrm{E}\,(X) > \mathrm{E}\,(Y)$$

which is the maximum expected value criterion (the one we used in the previous chapters). When the subject is risk averse one must use 5.1 and is bound to find the function u corresponding to that subject. Expected utility theory shows that *a subject is risk averse if and only if its utility function is concave.* As a model for utility functions, the family of functions

$$u(x) = 1 - e^{-x/R}$$

has been widely accepted. It is called "exponential utility". The parameter $R > 0$ is called *risk tolerance*. The graph of $u(x)$ for some values of R is shown in Figure 5.2.

The larger R, the less concave is the curve, the closer it is to a straight line.

Exponential utility functions family has these features:

- to find one member in the family only one number is needed, R;
- a quantity called *Arrow-Pratt measure of absolute risk aversion* defined by $-u''/u'$ is a constant. This roughly means that the subject behavior against risk does not depend on the amount of the lottery.
- exponential utility is defined over $[0, +\infty)$ with range $[0, 1)$. As it is, it is not immediately fit for any problem. If w is the worst outcome of the problem we should modify u to get $u(w) = 0$. Geometrically, we need to horizontally displace the graph in such a way that the intersection with the x axis is $x = w$ and not $x = 0$. To get this, we can put $u(x) = 1 - e^{-(x-w)/R}$.
- once $u(x)$ is known, to compare certainty equivalents we need to get $z(x)$ and thus to invert u. This is easily done with the exponential utility function because
$$u^{-1}(x) = -R \ln(1 - x)$$

Going back to the first example, if the person has an exponential utility function with $R = 15$ his utility function is

$$u(x) = 1 - e^{-x/15}$$

and the expected utility of random variable X in thousands of euros is

$$0.5u(30) + 0.5u(0) = 0.5 \left(1 - e^{-30/15}\right) = 0.432332$$

with a certainty equivalent of

$$z(X) = u^{-1}(0.432332) = -15 \ln(1 - 0.432332) = 8.493287$$

smaller than 10, the (certain) increase without changing job. So this person would rather not to change his job. Note that if R increases (i.e. the risk tolerance increases), the risk aversion decreases and the optimal decision could change. If $R = 21$ the certainty equivalent would be 10.04466, pointing to the change job alternative.

To summarize, the introduction of a utility function changes the value of nodes in a DT and could possibly lead to different decisions. To help the decision process, many software give the user the option to use utility function specifying its parameters and computing the new values.

5.3 Problems

Problem 5.3.1. Consider example 4.1.1 with all the amounts *expressed in thousands of euros*. With the help of a computer, point to the optimal decision for a subject with exponential utility function depending on the value for the risk tolerance.

5.4 Quickies

Question 5.4.1. A lottery ticket wins 100 with probability 30%. If for a person A the certainty equivalent of this lottery is 20, how can we describe A's attitude toward risk?

Question 5.4.2. A risk averse person buys a lottery ticket that wins an amount x with probability 1%. The ticket costs 1 euro. If the person acted rationally, could x be any number?

Question 5.4.3. Given a utility function $u = \sqrt{x}$ and a choice between two alternatives 10 e 20 with probabilities 0.3 and 0.7, compute the certainty equivalent of the choice.

Question 5.4.4. An insurance broker has to make his last call for today to try to sell a life insurance plan. He has two potential customers, both the same age and in good health. Their preferences are well modeled by an exponential utility function. They only differ for the value of R. Who should the broker call?

Question 5.4.5. Another way to see the preference for 10000 euros instead of the job change in example 5.1.1 is to think of a person who assigns a different value for the probability of the "lottery" (here represented by the reward in the new job). How shall one change probabilities of the two alternatives to decide for the current job?

Part III

Linear optimization

Chapter 6

Linear programming

6.1 Some definitions

Linear programming (LP) is a mathematical method originally developed to allocate limited resources among competing activities in an optimal way.

Limited resources

LP considers problems with limited resources. The word "resource" should have a broad meaning, referring not only to raw materials, etc., but also to labor hours, people, money, space, etc. Being limited is a common feature of many LP models (e.g. a machine can work for no more than 8 hours per day).

Competing activities

Activities that need resources are typically competing, in the sense that they use the same resources (e.g., if the machine is making a certain product, it can not make a different product at the same time).

Optimal allocation

Allocation of resources is the act of programming activities, i.e. the construction of a plan that says what is done and how (e.g., the machine will work on product A for 5 hours and then on product B for 3 three hours). Allocation should be optimal, i.e. it should maximize or minimize a certain function, called objective function.

Linearity

"Linear" means that all mathematical objects in this part of these notes are linear expressions, i.e. first degree equation and inequalities in the decision variables. This in turn means that all the quantities are additive, that is the quantity corresponding to the sum of two levels of a resource is the sum of the level of the two resources. For example, we will assume that if a machine requires 2 hours to process a unit of a certain product, then it will need $2k$ hours to process k units of the same product. It's easy to see that, generally speaking, this is not an easy requirement to meet. When this happens, the model must be written using non linear equations and inequalities. We will not deal with such kind of problems.

In the following paragraphs we will build a LP model by defining its three main parts: the decision variables, a system of constraints and an objective function. We stress the importance of this approach: one must always define these objects before even trying to look at a solution.

6.2 A LP problem

Example 6.2.1. A chemical company produces 4 types of reagents, R_1, R_2, R_3 and R_4, with 3 machines, M_1, M_2 and M_3. Table 6.1 shows the time (in

	R_1	R_2	R_3	R_4
M_1	3	4	7	1
M_2	5	7	1	1
M_3	1	1	2	1

Table 6.1: Time consumption for each machine for 1000 liters of reagent

	revenue (euros)
R_1	6.2
R_2	3.2
R_3	18
R_4	2

Table 6.2: Revenues for each liter of reagent sold

hours) needed by each of the three machines to process 1000 liters of reagent. For example, to get 1000 liters of R_1, M_2 mus work for 5 hours.

The company owns 3 machines of type M_1, 4 of type M_2 and one of type M_3. Each machine can work for no more than 100 hours per week.

Every week, the company must produce at least 10000 liters of each of the reagents R_2, R_3 and R_4.

The revenues from the selling of 1 liter of reagent is shown in Table 6.2 and the company is sure to sell the whole weekly production.

The company would like to plan the weekly production of the reagents in such a way as to maximize its revenues.

Decision variables

The starting point in the resolution of a LP problem is the definition of the *decision variables*. These variables measure the resources to be allocated to the activities in the problem. The number of decision variables depend on the problem.

In example 6.2.1 what the decision maker is to decide is the production plan, the quantities to be determined are the liters of each reagent to produce. Thus, there are 4 decision variables: R_1, R_2, R_3 and R_4. However, before going on, we should point out that within a given problem, each decision variable should satisfy a certain number of preliminary constraints.

As R_1, R_2, R_3 and R_4 represent quantities of a certain physical asset they are to be non negative numbers. This kind of constraint is often present in LP problems. Note that the zero value is usually acceptable as it represent the missing of that resource from the plan.

What are the units of measure of the decision variables? Sometimes the choice is obvious but in particular cases, as that of the example 6.2.1, the answer might not be unique. In fact, we could measure reagents in liters or, to follow Table 6.1, in thousands of liters.

Should we consider only integer numbers or also floating point numbers for the variables? If we use thousands of liters as units, we are forced to use floating point variables, because it should be possible to produce, say, "3.1415 thousands of liters", approximately 3141 liters and a half. The same can be said even if we take liters as units because, ideally, the quantities measured by the variables are indefinitely divisible. However, in some case, as we shall see in section 8.2), this condition is not met. When, for example, decision variables represent the number of people employed in a given activity, it is obvious that these numbers should be integers and can not be floating point numbers.

To make a choice, we state that R_1, R_2, R_3 and R_4 will be the number of *thousands of liters of reagent* and therefore will take non negative real values,

$$R_i \in \mathbb{R}, R_i \geq 0, i = 1, 2, 3, 4.$$

Constraints

As we deal with limited resources, decision variables are to be constrained. We already know that they can not be negative. However, any LP problem will usually contain other restrictions that, mathematically, will be expressed as equations and inequalities in the decision variables.

In the statement of the problem it is said that "Every week, the company must produce at least 10000 liters of each of the reagents R_2, R_3 and R_4. " The mathematical translation is made up of three inequalities:

$$\begin{aligned} R_2 &\geq 10 \\ R_3 &\geq 10 \\ R_4 &\geq 10 \end{aligned}$$

as variables represent thousands of liters.

Other constraints come from the limited number of machine hours. Machine M_1 is available for no more than 300 hours. Every thousand of liters of reagent 1 uses M_1 for 3 hours; thus, if the production plan states we are to produce R_1 thousands liters of reagent 1, we will use machine 1 for $3R_1$ hours. Similarly, for reagent 2 we will use machine M_1 for $4R_2$ hours, for reagent 3 we will use machine 1 for $7R_3$ hours and for R_4 hours for reagent 4. Summing up, for a production plan of R_1, R_2, R_3 and R_4 thousands liters, machine M_1 us used for $3R_1 + 4R_2 + 7R_3 + R_4$ hours. This sum should be less or equal to 300 hours (because we have 3 machines of type M_1 and each of them is available for 100 hours). In mathematical terms

$$3R_1 + 4R_2 + 7R_3 + R_4 \leq 300.$$

On the same line of thought, we get two similar inequalities for the other two machines:

$$\begin{aligned} 5R_1 + 7R_2 + R_3 + R_4 &\leq 400 \\ R_1 + R_2 + 2R_3 + R_4 &\leq 100. \end{aligned}$$

The decision variables must therefore satisfy a solution of the following system of inequalities:

$$\begin{cases} 3R_1 + 4R_2 + 7R_3 + R_4 \leq 300 \\ 5R_1 + 7R_2 + R_3 + R_4 \leq 400 \\ R_1 + R_2 + 2R_3 + R_4 \leq 100 \\ R_2 \geq 10 \\ R_3 \geq 10 \\ R_4 \geq 10 \\ R_i \geq 0, i = 1, 2, 3, 4. \end{cases}$$

Feasible solutions

These inequalities are the problem *constraints*. The set of all solutions of these system of inequalities is called the *set, or region, of the feasible solutions*. Each feasible solution is a feasible decision, which *satisfy every constraint*. Sometimes the feasible set is empty: this means that the constraints are too restrictive and the problem has no solution. For example, the following constraints system

$$\begin{cases} R_1 + R_2 \leq 50 \\ R_1 + 2R_2 \geq 120 \\ R_i \geq 0, i = 1, 2. \end{cases}$$

has a null feasible set. In fact, the second equation can be written as $R_1 + R_2 + R_2 \geq 120$. The first equation implies that the sum of the first two terms is less than or equal to 50, which in turn implies that R_2 is greater than 70, which is impossible, because the two variables are non negative and each of them must be less or equal to 50. This could be too simple an example but with typical LP problems with many variables and many constraints, it is very difficult to spot a null feasible set and this can be done only with the help of a computer.

If there are solutions, i.e. if the feasible set is not empty, then the optimal solution must be searched in this set.

Objective function

The last step in the formulation of a LP problem is the provision of a way to choose the optimal solution among all the possible solutions. The choice is made by either maximizing or minimizing a certain function, called *objective function*. In the current example, we want revenues to be at their maximum. As we plan the production of R_1, R_2, R_3 and R_4 thousands liters of each of the 4 reagents, and given the data in Table 6.2, total revenues amount to

$$6200R_1 + 3200R_2 + 18000R_3 + 2000R_4$$

and we require this quantity to be the maximum possible.

All the previous requirements is summarized as follows:

$$\begin{array}{rll} \max & 6200R_1 + 3200R_2 + 18000R_3 + 2000R_4 & \\ \text{s.t.} & 3R_1 + 4R_2 + 7R_3 + R_4 \leq 300 & (\text{machine } M_1) \\ & 5R_1 + 7R_2 + R_3 + R_4 \leq 400 & (\text{machine } M_2) \\ & R_1 + R_2 + 2R_3 + R_4 \leq 100 & (\text{machine } M_3) \\ & R_2 \geq 10 & (\text{min. production } R_2) \\ & R_3 \geq 10 & (\text{min. production } R_3) \\ & R_4 \geq 10 & (\text{min. production } R_4) \\ & R_i \geq 0, i = 1, 2, 3, 4. & (\text{non neg. var.}) \end{array}$$

	value
Reagent R1	0
Reagent R2	10
Reagent R3	35.714
Reagent R4	10
Objective	694857

Table 6.3: Solution of example 6.2.1: final values of decision variables and objective

	Value	Status	Slack
Machine 1	300	Binding	0
Machine 2	115.714	Not binding	284.285
Machine 3	91.429	Not binding	8.571
Minimum R2	10	Binding	0
Minimum R3	35.714	Not binding	25.714
Minimum R4	10	Binding	0

Table 6.4: Solution of example 6.2.1: constraints values

where "s.t." stands for *subject to*. All the objects in the previous summary is indeed linear. When an LP problem is in the form just outlined, i.e. it has:

- an objective function to be *maximized*
- all the decision variables bound to be ≥ 0
- all the constraint in the $\leq$ form with variables on the left and resources (numbers) on the right

it is said to be in *standard form*.

Solution

To get the solution to problem in example 6.2.1 we need a computer. The final output of the software is usually similar to Tables 6.3 and 6.4.

The optimal production plan is to produce 10000 l of reagent 2, 35714 l of reagent 3, 10000 l of reagent 4 and not to produce reagent 1 at all. Note that:

- although the revenues per unit for reagent 1 is greater than that of reagent 2 and 4, reagent 1 does not appear in the optimal solution
- while type 1 machines are used exhaustively, type 2 and type 3 machines are used only partially. Type 2 machines are used for 115 hours out of 400 available hours; type 3 machines are used for 91 hours out of 100 available hours; remaining hours are shown in the 4th column in Table 6.4
- minimum production constraints on reagents 2, 3 and 4 are met, with a surplus for constraint 3. Surplus is shown in the 4th column in Table 6.4

Binding and not binding constraints

In an optimal solution, when the value of a constraint is equal to its right hand side, we say that the constraint is *binding*. When the values are different, we say that the constraint is *not binding*. If a constraint is binding, the corresponding resource is exhausted: in the optimal plan machine 1 has no hours left. On the contrary, a not binding constraint points out a resource that is not exhausted and what is left is reported in the "slack" column. For minimum production constraints, we could make a similar, opposite, observation.

Another way to look at the same topic is to think of changing the left hand sides of the constraints. In binding constraints a variation on the left hand side in the same direction of the constraint (i.e. an increase for a greater than or equal constraint, a decrease for a less than nor equal constraint) *forces* a change in the solution. If machine 1 available hours decrease by one, going to 299, the optimal solution must change because it is not a feasible solution anymore! Similarly, if, say, the reagent 2 minimum production level increases to 11 thousands, the previous optimal solution is again not feasible. On the contrary, for not binding constraints there can be "room for change". For example, if a type 2 machine breaks down, making the available hours decrease to 300, the optimal solution will not change. And an increase in the minimum production of reagent 3 to 20 thousands would not force to change the plans, as the optimal solution is already at 35714 liters .

6.3 Graphical solution

The graphical method for LP can be used only for problems with two decision variables. We illustrate it with the following example.

Example 6.3.1. A chemical firm produces two types of textile dyes, C_1 ed C_2. Each dye is made dispersing the pigment in a liquid. Each week 1000 mg of pigment and 40 hours of labor are available. Total weekly production must not exceed 700 l of dye. The quantity of C_1 dye can exceed the quantity of C_2 dye by no more than 350 l. Each l of C_1 dye requires 2 mg of pigment and 3 minutes of labor. Each l of C_2 dye requires 1 mg of pigment and 4 minutes of labor. Each l of C_1 dye makes 8 euros in revenues; each l of C_2 makes 5 euros.

The managing board, given that C_1 is more profitable than C_2, suggested to use all the resources to produce the maximum of C_1 and then to use the remaining resources to produce C_2. Thus, the firm produced 450 l of C_1, for a revenues of $450 \times 8 = 3600$ euros, and 100 l of C_2, for an additional $100 \times 5 = 500$ euros in revenues. The total revenues are 4100 euros. Is this solution optimal?

We have 2 decision variables: x, the planned production of C_1, and y, the planned production for C_2, both expressed in l. The objective function is the total revenues, $8x + 5y$. Moreover:

1. the total used pigment must not exceed 1000 mg: $2x + y \leq 1000$;
2. labor hours must be less than or equal to 40: $3x + 4y \leq 2400$;
3. total production must not exceed 700 l: $x + y \leq 700$
4. C_1 production must not exceed C_2 by more than 350 l: $x \leq y + 350$.

The LP problems is the following:

$$
\begin{array}{rll}
\max & 8x + 5y & \\
\text{s.t.} & 2x + y \leq 1000 & \text{(pigment)} \\
& 3x + 4y \leq 2400 & \text{(labor hours)} \\
& x + y \leq 700 & \text{(total prod.)} \\
& x - y \leq 350 & \text{(mix prod.)} \\
& x, y \geq 0 &
\end{array}
$$

Inequalities in the Cartesian plane

Analytical geometry is a tool that links algebraic objects, equations and inequalities, with geometric ones: each point in the plane is linked to a couple of real numbers and an equation in the two variables, x and y, links, through its solutions, a geometrical object in the plane. If, for example, we *briefly* say that the equation $2x + y - 1000 = 0$ represent a straight line in the plane, we actually mean that each solution of the equation $2x + y - 1000 = 0$ is a couple of numbers x and y, and this couple is linked to a point in the plane. Therefore, the set of (all) the solutions of the equation is linked to the set of the corresponding points. These point are displaced on a straight line. It is then obvious that if a point does not lie on a straight line, its coordinates are not a solution of the equation and vice versa.

The set of solutions of the inequality $2x + y \leq 1000$ is made up with the solutions of the equation $2x + y = 1000$ and the solutions of the inequality $2x + y < 1000$. The former are the point on the straight line. The latter are to be found among the point that do not belong to the line, that is among the point for which $2x + y \neq 1000$. Geometrically, this is the union of the two half planes created by the line $2x + y = 1000$. the origin, $(0, 0)$, is a solution of the inequality ($2 \times 0 + 0 = 0 < 1000$). For any other point $P = (\alpha, \beta)$ lying in the same half plane as the origin, we could think of drawing a line from the origin to P not intersecting the straight line $2x + y = 1000$. If $2\alpha + \beta > 1000$ then this would mean that we can change the value of $2x + y$ from a value less than 1000 to a value greater than 1000. Continuity of the expression $2x + y$ and of the line OP imply that somewhere between the origin and P there should be a point in which $2x + y = 1000$. But this can not happen, because O and P are on the same side of the straight line. So the value of $2x + y$ in P, and that of any point on the same side of P, is to be of the same sign of that in O, i.e. less than 1000. The straight line parts the plane into two half-planes: in one region $2x + y > 1000$ while in the other $2x + y < 1000$. As we know that in the origin $2x + y < 1000$, we can say that the graphical solution of $2x + y < 1000$ is the half-plane with the origin. The solution to the original problem, $2x + y \leq 1000$, is the union of the half-plane and the straight line.

The same can be said about non negativity constraints: requiring that $x > 0$ means we are considering only the points to the right of the y-axis and similarly for $y > 0$. Both constraints bound the feasible solution to the first quadrant.

The feasible region will be solution of the system of all the inequalities of the system.

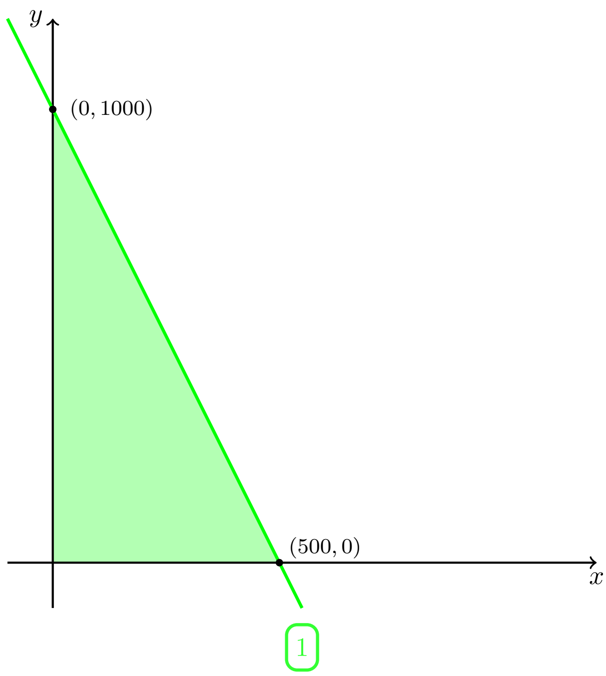


Figure 6.1: Feasible region in ex. 6.3.1 with the constraint $2x + y \leq 1000$

Graphical solution

Let us draw the feasible region step by step, taking one constraint at a time:

1. the pigment used in the process is to be less than 1000 mg: $2x+y \leq 1000$. We draw the straight line and shadow the feasible region. See Figure 6.1.

2. labor hours have to be less than or equal to 40: $3x + 4y \leq 2400$. Again, we draw the straight line $3x+4y = 2400$, and we note that the half-plane we are interested in is that with the origin. Thus, the feasible region is given by this half-plane intersected with the previous intermediate result. See Figure 6.2.

3. whole production must not exceed 700 l: $x + y \leq 700$. See Figure 6.3.

4. C_1 production must not exceed that of C_2 by more than 350 l: $x \leq y+350$. See Figure 6.4.

The resulting area is the feasible set. Every point in this region represents a production plan that satisfies the constraints.

The target is to choose among all the feasible solution to get the one that maximizes the objective function, $8x + 5y$. Let us take the set of parallel straight lines of equation $8x + 5y = z$, where z is a real number. Each value of z determines a unique straight line and therefore a value for the objective. We would like to know:

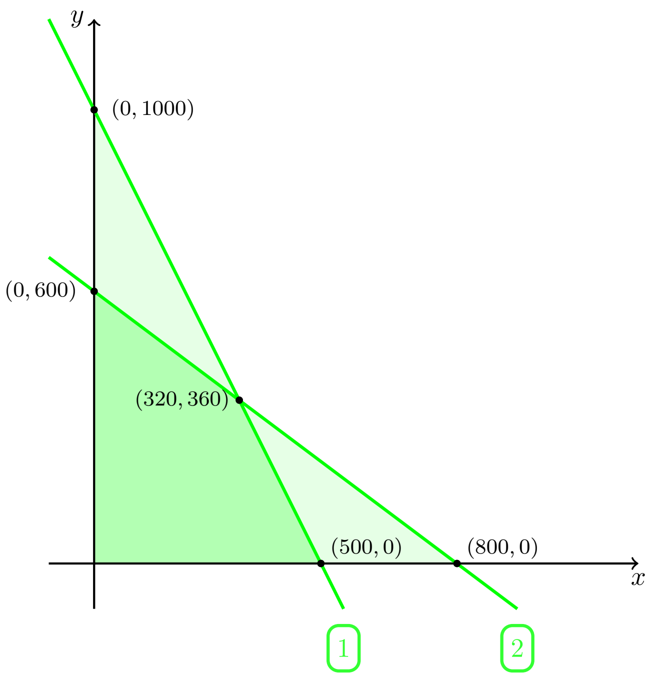


Figure 6.2: Feasible region in ex. 6.3.1 with $2x + y \leq 1000$ and $3x + 4y \leq 2400$

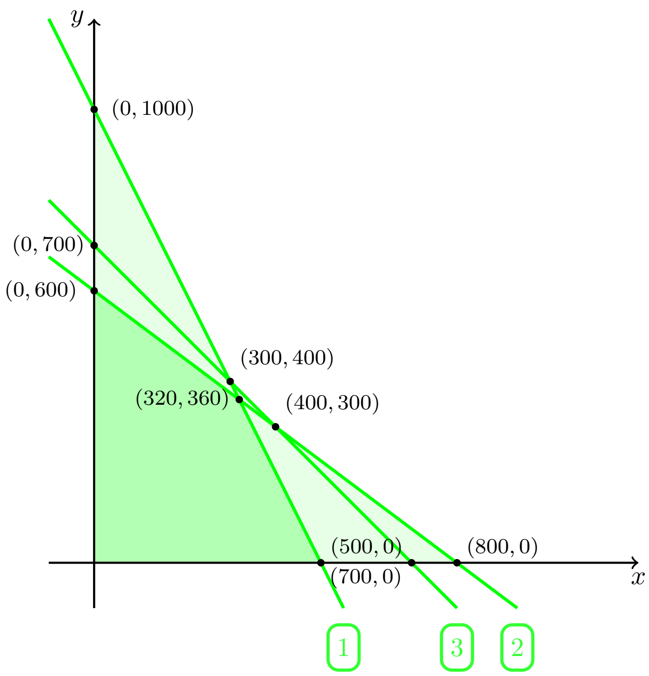


Figure 6.3: Feasible region in ex. 6.3.1 with $2x + y \leq 1000$, $3x + 4y \leq 2400$ and $x + y \leq 700$

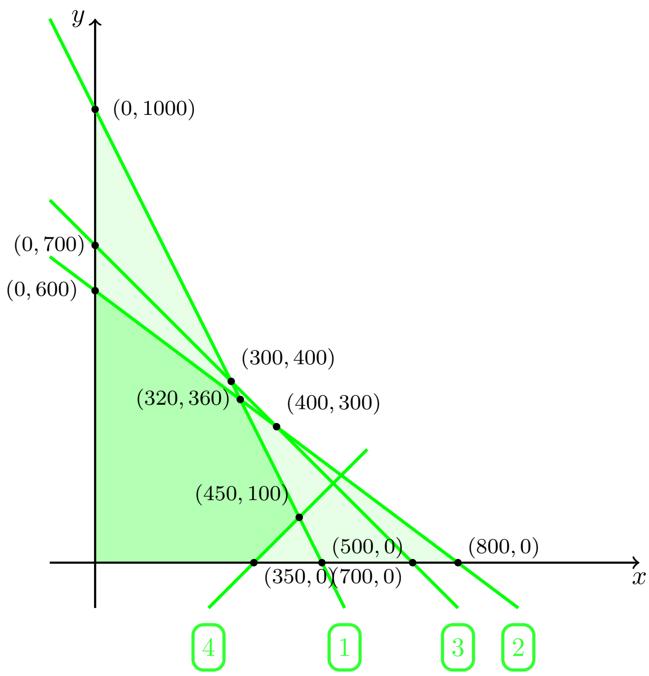


Figure 6.4: Feasible region in ex. 6.3.1 with all the constraints.

- if at least one point in the feasible region is a point of at least one of the parallel lines
- among the points that satisfy the previous requirement, which is that (or those) whose objective value is the highest possible.

For every non empty feasible region, the answer to the first question is always yes, because the parallel lines span the entire plane. To answer the second question, we show in Figure 6.5 3 straight lines corresponding to the values $z = 100, 1000, 2000$.

As the line moves to the right, z increases. Thus, every feasible solution that lies on $8x+5y = 1000$ is better than any that lies on $8x+5y = 100$. As we increase z, i.e. we move the line to the right, the line will eventually have no intersection with the feasible region. The borderline point is $(320, 360)$: the line that passes through this point is that with $z = 8 \times 320 + 5 \times 360 = 4360$. This is the maximum value of the objective that satisfies the constraints, for moving the line further to the right we get no feasible solution. The optimal solution to the problem is produce 320 l of C_1 and 360 of C_2. The corresponding revenues are 4360 euros. See Figure 6.6.

The optimal solution is a vertex at the intersection of the straight lines corresponding to constraints 1 and 2. This means that when choosing the optimal solution, the resources represented in these constraints are completely allocated. In fact, with $x = 320$ and $y = 360$ the pigment used in the process is $2 \times 320 + 360 = 1000$, i.e. all the available pigment. The same can be said for labor hours. Constraints 1 and 2 are *binding*. In contrast, the optimal solution

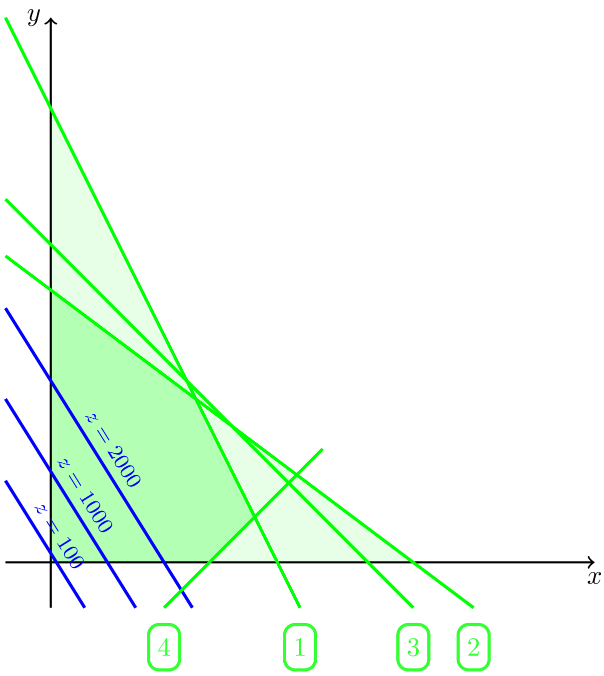


Figure 6.5: Feasible region and objective function for some values of z.

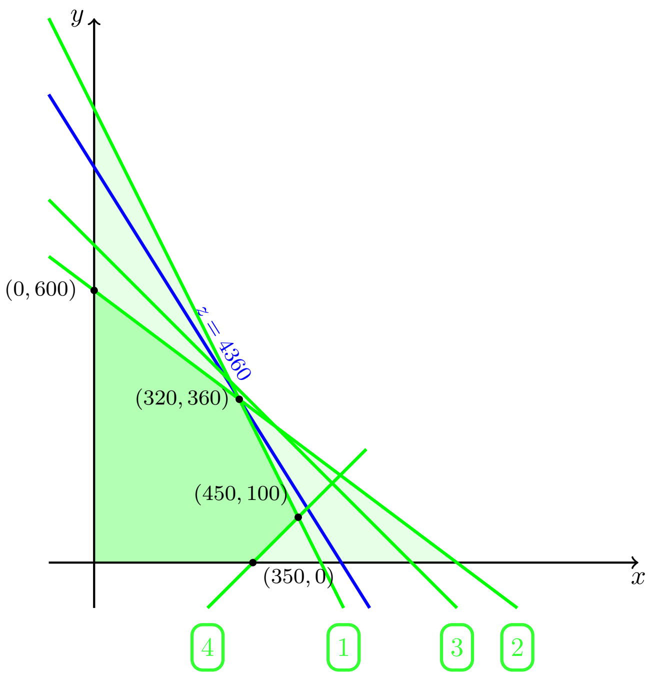


Figure 6.6: Graphical solution of ex. 6.3.1.

Vertices	Values of z
$(0, 0)$	0
$(350, 0)$	2800
$(450, 100)$	4100
$(320, 360)$	4360
$(0, 600)$	3000

Table 6.5: Vertices of the feasible region and values of the objective in ex. 6.3.1.

does not lie on the other two straight lines, corresponding to constraints 3 and 4. These constraints are *not binding* and the corresponding resources are not completely allocated.

By the word "resource" we do not necessarily mean a material resource. Constraint 4, for example, defines a feasible mix of production, not a quantity of something. The fact that a constraint is not binding means that its right hand side (the binding value of the resource) can be changed without altering the optimal solution.

It can be shown that the optimal solution in a LP problem is always in a vertex of the feasible region. If the optimal vertices are more than one, this means that all the points in the segment joining the two vertices are optimal solutions as well and thus that there are infinite optimal solutions (see section 7.5).

The preceding result suggests a quick method for finding the optimal solution. As the optimum must be on a vertex of the feasible region, it is enough to check the value of the objective in all the vertices. The vertex with the highest value is the optimal solution. In the current example, vertices and corresponding values of the objective are shown in Table 6.5. It is easy to see that the maximum is reached in $(320, 360)$.

6.4 Problems

Problem 6.4.1. A company produces two types of pipes: high temperature resistant (A) and ultra high temperature resistant (AA). For every ton of AA pipes, the company get 39 euros of revenuesm while each ton of A pipes generates a revenues of 31 euros. The production of a ton of AA type requires 7 hours, while for a ton of A type only 5 hours are needed. The machinery can provide no more than 56000 work hours.

The company would like to produce at least 10000 tons of pipes, no matter what type. The company would like to maximize revenues meeting the constraints. Formulate a LP model and solve it graphically.

Problem 6.4.2. A farm buys feed for its animals from two suppliers, 1 and 2. Feeds are mixed together before being given to animals. Each feed contains ingredients which are essential to the life of the animals and to their growth, such as vitamins, minerals, etc.

One kg of feed from supplier 1 contains 5 g of ingredient A, 4 g of ingredient B and 0.5 g of ingredient C. One kg of feed from supplier 2 contains 10 g of A, 3 g of B and does not contain C.

Supplier 1 cost is 0.2 euro per kg; supplier 2 cost is 0.3 euro per kg.

The monthly need for each animal is 90 g of A, 48 g of B and 1.5 g of C.

Write a LP model and determine the composition of the mix that, while meeting the animal needs, is the cheapest. Use the graphical method to solve the problem.

6.5 Quickies

Question 6.5.1. What is the "slack" in a constraint in a standard LP model?

Question 6.5.2. What does "linear" menas in LP?

Question 6.5.3. What is a binding constraint?

Question 6.5.4. Are two LP problems that differ only in the order of the constraints equal?

Question 6.5.5. What is the feasible set in a LP problem?

Chapter 7

Sensitivity analysis

7.1 General framework

A LP model contains an objective function and a number of equalities and inequalities. All of them are defined through their coefficients. These coefficients represent the parameters of the problem, such as prices, revenues, costs, quantities, etc. that are assumed to be valid at the moment in which the decision maker will start the execution of the production plan. As we already wrote in Chapter 4, in doing this a couple of issues should be noted.

The first is that all the coefficients in a LP problem are typically the result of more or less reliable estimates of the quantities in place. These are usually known with a certain degreee of confidence. Thus we might wish to assess what happens when they vary in a given range.

The second is that the solution of a LP problem is something that will be implemented in the future and is therefore subject to another degree of uncertainty. For example, a LP problem might include the price of a certain item. Based on this particular value an optimal solution has been found. Now, the decision maker runs the risk that, when the production is completed and the product is about to be sold, market conditions have changed and the price is either too high or too low to be reasonable.

As in Chapter 4, to overcome both issues we must see how the optimal solution changes as single coefficients change. This analysis is called *sensitivity analysis*. We show how it works on the problem already solved graphically.

7.2 Objective's coefficients

In example 6.3.1, the objective function coefficients of choice variables x and y in represent revenues coming from the sale of 1 l of C_1 and 1 l of C_2. Let us call the two coefficients (marginal revenues) a and b. These values (8 and 5) are defined at the moment the model is built and could change when dyes are actually sold. A change in the coefficient of the objective function causes a change in the slope of the parallel lines and we must see if these changes could affect the optimal solution or not.

Assume, for simplicity, that there could be a variation of a while b remains fixed. We are therefore studying the behavior of the lines $ax + 5y = z$. These lines are called iso-revenues lines. The slope of the lines is $-a/5$ and as far as this is between that of constraints 1 and 2, the optimal solution is the same (the optimum is the vertex at the intersection of constraints 1 and 2). The slope of constraint 1 is -2 while that of constraint 2 is $-3/4$. Thus, $(320, 360)$ is still the optimal solution if a is such that

$$-2 < -\frac{a}{5} < -\frac{3}{4}.$$

Multiplying by -5 we get

$$3.75 < a < 10.$$

As a is 8, we can state that

> if a increases by a maximum of 2 $(= 10 - 8)$ or decreases by a maximum of 4.25 $(= 8 - 3.75)$ the optimal solution remains unchanged.

The difference between the maximum value of a for which the solution does not change and the original value of a is called *allowable increase*. Similarly, the difference between the minimum value of a for which the solution does not change and the original value of a is called *allowable decrease* . The overall variation interval for the coefficient is called allowable range

allowable increase and decrease

If a changes in the allowable range, the optimal solution does not change but the optimum value does change. For example, if a becomes 9, the optimal solution is still $(320, 360)$ but the value of the objective is now $9 \times 320 + 5 \times 360 = 4680$. This change, however, is of minor importance because the good news is that the already in progress plan is still the optimal one. The greater the allowable increases and decreases the less the decision maker is worried about his plan.

An analogous computation can be done for b. If we have $8x + by = z$, with $b > 0$, the slope of the objective is $-8/b$ and the previous requirements are:

$$-2 < -\frac{8}{b} < -\frac{3}{4}.$$

Multiplying by $-b$ we get

$$4 < b < \frac{32}{3}.$$

Summing up we could say that

> if y coefficient in the objective increases by a maximum of $17/3 = 5.66\ldots$ $(= 32/3 - 5)$ or decreases by a maximum of 1 $(= 5 - 4)$ the optimal solution remains unchanged.

As in the previous case, even if the optimal solution does not change, the value of the objective function at the optimum does change. When dealing with more than two variables these calculations are done by a computer whose output is a table with maximum allowable increase and decrease for any objective coefficient.

Sensitivity analysis on objective coefficients provides the decision maker with the range within which each coefficient could change without changing the optimal solution.

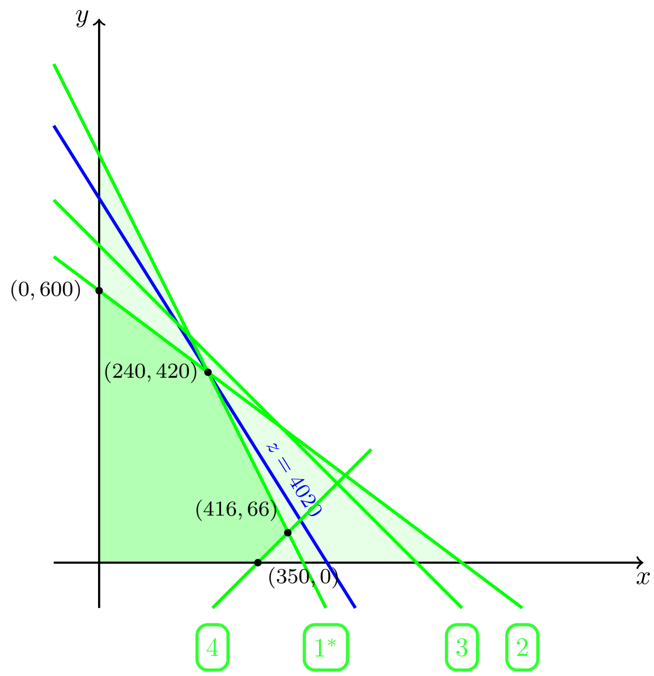


Figure 7.1: Graphical solution to example 6.3.1 with $2C_1 + C_2 \leq 900$.

7.3 Resources level and shadow prices

A different managerial question that can be addressed through sensitivity analysis concerns available resources, i.e. the left hand sides of constraints inequalities. We note that by changing the levels of the resources we could, in principle, change our production plan. Actually, we are changing the shape of the feasible region and so the question is whether it is worth going to the market to get additional resources. This then, boils down to determining how much a unit of additional resource is worth to us.

Let us start with constraint 1: $2x + y \leq 1000$. If we decrease 1000 to, say, 900 the corresponding straight line moves left and the maximum moves left as well (see Figure 7.1 where the new constraint is named 1^* and is parallel to the old).

The new optimal point is $(240, 420)$ with an objective maximum value of 4020 euros. The decrease in the maximum is easily explained taking into account the new resource level: as now there is less pigment available, there will be less dye and then a decrease in revenues.

The maximum has changed and the optimal solution has changed as well. What is the same as in the original solution is the fact that the optimal point is still at the intersection of constraint 1 and constraint 2 (even if the former has moved from the original position). This means that binding constraints are the same as in the original optimum. Note that the change in level of the binding resources can (not necessarily must) force a change in the optimal solution, while a change in not binding constraints does not change the optimum.

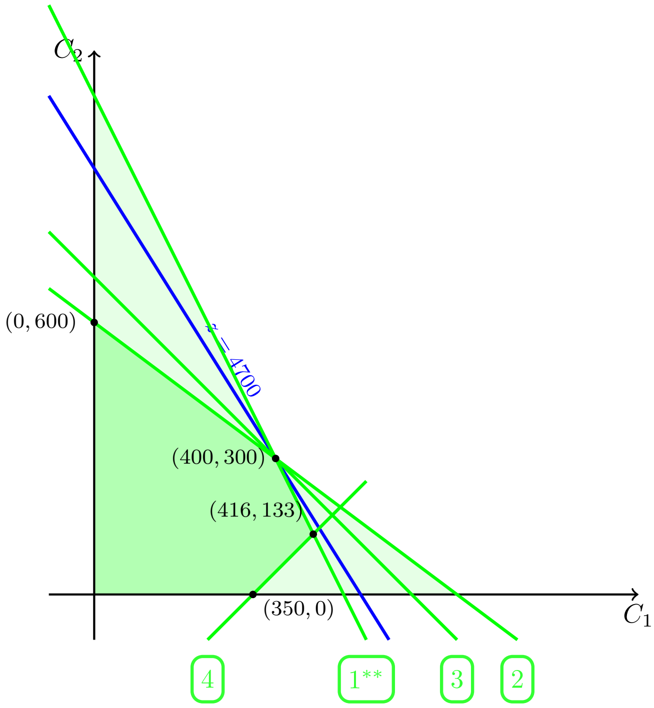


Figure 7.2: Graphical solution of example 6.3.1 with $2C_1 + C_2 \leq 1100$.

If we further reduce the available pigment, the optimal point moves towards $(0, 600)$. That is an "extreme point": a further reduction in the resource level moves the point of maximum (or minimum) from the intersection of constraints 1 and 2 to the intersection of constraint 1 and the y-coordinate axes (which is a non-negativity constraint). As the set of binding constraints has changed, the previous analysis is not valid anymore.

Let us now go back to the original problem, and move to a different situation, namely the increasing of the available pigment. In this case, constraint 1 moves to the right. If the increase is 100 mg the optimal solution moves from $(320, 360)$ to the intersection of constraints 1, 2 and 3, at $(400, 300)$ (see Figure 7.2).

Further increments in constraint 1 level move the optimal solution on the intersection between constraint 2 and 3 and again the set of binding constraints changes. Summarizing

> if the left hand side of constraint 1 increases by no more than 100 or decreases by no more than 400 the optimal solution remains on the intersection of constraints 1 and 2.

Shadow prices

A 100 units increase in the right hand side of the pigment constraint has caused the maximum revenues to increase by 340 euros. On average, this means 3.4 euros more per unit of pigment. As the model is linear, this average is actually equal to the "marginal" increase. In the differential calculus jargon, if b is the right hand side and f is the objective, we would write

$$\frac{\Delta f}{\Delta b} = 3.4$$

when Δb is within the bounds found in the previous section. The value of $\Delta f/\Delta b$ is called "shadow price" of the pigment. Shadow prices corresponding to different constraints are usually denoted by λ_1, λ_2, etc.

To explain the meaning of shadow prices, consider a decision maker who wonders what happens when he buys one more unit of a certain resource whose market price is c. Having more resources, the value of the objective function at the maximum increases by the shadow price, λ. To purchase the new resources is profitable only if $\lambda > c$. Thus, λ is indeed a sort of price, even if it not a real one, as it is computed by the model. Its value depends on the coefficient appearing in the model. This is the reason why this price is labeled "shadow".

If we proceed in the same way for constraint 2 we discover that a labor hour has a shadow price of 0.4 euros. If the market cost of a labor hour were less than 0.4 euros, say 0.3 euros, it would be profitable to buy extra hours to increase the profit.

Constraints 3 and 4 are not binding and therefore increasing the values of the corresponding resources does not cause the objective to change. Their resources are already more than what is needed to reach the maximum and therefore the shadow price is zero.

There is a simple relation between shadow prices and binding constraints:

> In an inequality constraint, the product of each slack and the corresponding shadow price is always 0. Therefore, if there is a slack (and the constraint is not binding), then the shadow price is 0; if, on the contrary, the shadow price is different from 0, then the slack is zero (and the constraint is binding).

This principle is called "complementary slackness".

It should be noted that the aforementioned product could be zero also when a shadow price vanishes *and* the corresponding constraint is binding. However, this happens only in some particular cases, see section 7.5. Shadow prices are defined similarly when resources are decreased. From a mathematical point of view, shadow prices are both left and right derivatives.

The sign of a shadow price is related to the type of constraint. In a less than (¡) constraint, increasing the resource level *extends* the feasible set and the objective can change or remain the same as before (because the previous optimal point is still feasible). In a maximization problem objective can increase ($\lambda > 0$) or not ($\lambda = 0$); in a minimization problem, objective can decrease ($\lambda < 0$) or not ($\lambda = 0$).

Conversely, when the constraint is a greater than (¿) an increase in the resource level *shrinks* the feasible set and the original optimal solution could be not feasible anymore. In a maximum problem, objective can decrease ($\lambda < 0$) or not ($\lambda = 0$); in a minimum problem objective can increase ($\lambda > 0$) or not (see Table 7.1).

When the LP problem has more than two variables the solution is usually obtained by a computer program which outputs a standard report with the

	ob. max.	ob. min.
$\leq$	+	-
$\geq$	-	+

Table 7.1: Shadow price sign and type of constraint and objective function.

Name	Final Value	Reduced Cost	Objective Coefficient	Allowable Increase	Allowable Decrease
Reagent R1	0.000	-1514	6200	1514.285714	1E+30
Reagent R2	10.000	0	3200	7085.714286	1E+30
Reagent R3	35.714	0	18000	1E+30	3533.33332
Reagent R4	10.000	0	2000	571.428571	1E+30

Table 7.2: Sensitivity report (1) for ex. 6.2.1: decision variables and objective function

Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
Machine 1	300.000	2.571	300	30	180
Machine 2	115.714	0.000	400	1E+30	284.2857143
Machine 3	91.429	0.000	100	1E+30	8.571428571
Minimum R2	10.000	-7.086	10	44.22222222	10
Minimum R3	35.714	0.000	10	25.71428571	1E+30
Minimum R4	10.000	-0.571	10	12	10

Table 7.3: Sensitivity report (2) for ex. 6.2.1: constraints

most relevant facts. For example 6.2.1 Microsoft Excel Solver produced two tables, 7.2 and 7.3.

7.4 Reduced cost

In report 7.2 there is a “reduced cost” column. To explain the meaning of this column, let us go back to the original optimal solution. The optimal plan is to produce 10000 l of reagent 2, 35714 of reagent 3, 10000 l of reagent 4 e not to produce reagent 1. Reagent 1 is missing from the optimal plan because, in comparison with other reagents that use the same resources, its unit revenue is too low to be profitable. If, starting from the optimal plan, we were somehow forced to produce reagent 1, e.g. for marketing reasons, the objective will decrease.

How much will we loose? In example 6.2.1 the unit revenue from reagent 1 is 6200 euros (a unit means a thousand liters). If we were to produce this amount of reagent 1 there will be no problem with minimum constraints for reagents 2, 3 and 4, because they would be automatically satisfied.

Machine 2 and machine 3 type constraints are not binding and therefore have slacks greater than 0. As one unit of reagent 1 uses 5 hours of machine 2 and 1 hour of machine 3 and corresponding slacks are of 284 hours and 8 hours, there is no problem here in making the new production.

Machine 1 constraint, however, is binding. To produce one unit of 1 we need 3 hours which must be taken off the optimal solution, reducing the quantity of the other reagents. There are 297 hours left for reagents 2, 3 and 4. In the new plan we get 6200 euros from the production of reagent 1 but we pay for the reduction. As machine 1 shadow price is 2571.429 euros, a 3 hours reduction costs $3 \times 2571.429 = 7714.287$ euro. The final balance is

$$+6200 - 7714.287 = -1514.287$$

euros. This is what is written in the "reduced cost" on the reagent 1 row. The reduced cost is the balance of moving the corresponding decision variable from zero to 1. If the target is to maximize the objective, this balance is negative (because if moves from the optimal solution).

Looking at the same problem from another point of view, we could say that the reduced cost is the amount by which we should increase the profit of the a decision variable to make it appear in the optimal solution. In this example, if the unit revenue were equal to or greater than 7714.287, the production of reagent 1 would increase the value of the objective and therefore reagent 1 would appear in the optimal solution.

Whenever a decision variable is missing from the optimal plan, because it is not profitable, the reduced cost quantifies this statement by telling *how much* it is not profitable.

7.5 Special cases

we will now examine some special LP problems.

Degeneracy

In the LP problem

$$\begin{array}{rl} \max & 3x + 2y \\ \text{s.t.} & x + y \leq 100 \\ & x \leq 50 \\ & y \leq 50 \\ & x, y \geq 0 \end{array}$$

the optimal solution is $x^* = 50$, $y^* = 50$ with a maximum objective value equal to 250. The first constraint is binding but if the resource level increases by one unit, to 101, the optimal solution does not change, because of the other two constraints. The shadow price of the resource in constraint 1 is 0, even if the constraint is binding.

If the resource level decreases by one unit, to 99, the optimal solution must change and necessarily it would be $x = 50$, $y = 49$, with a maximum at 248 (x is more profitable than y and then is better to decrease y than x). The shadow price is then 2.

Thus, in this case the shadow price is not uniquely defined but depends on the way we change the resource level. The problem lies in the fact that the first inequality is the sum of the other two inequalities. In mathematical terms, the three inequalities are not linearly independent. This is called *degeneracy*. For this example, this problem is easy to spot, but in more general models, with hundredths of variables and inequalities, this can be done only by the computer

Degeneracy

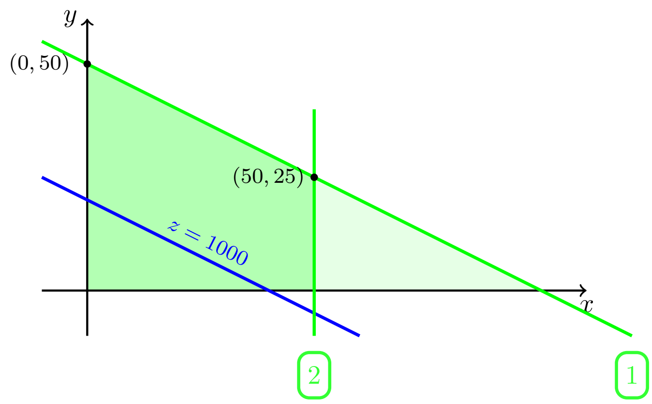


Figure 7.3: Feasible region and objective function for $z = 1000$.

and the optimization software typically points out these problems in the final report.

Infinite solutions

Let us solve the following LP problem graphically.

$$
\begin{aligned}
\max \quad & 25x + 50y \\
\text{s.t.} \quad & x + 2y \leq 100 \\
& x \leq 50 \\
& x, y \geq 0
\end{aligned}
$$

We obtain the result in Figure 7.3.

As the iso-objective line and that of constraint 1 are parallel, when they intersect they end up having an entire segment in common, namely the segment joining $(0, 50)$ and $(50, 25)$. Optimal solutions are therefore infinite, every point in the segment.

If we perform a sensitivity analysis on a problem like this one we would get strange results. A basic computer software, for example, could find only $(0, 50)$ and ignore all other solutions. Again, what is easy to spot in a two variable problem, could be invisible in a real world application and here the role of software is crucial: a good software for optimization is one that spots problems in the system of constraints and warns the user about possible oddities.

7.6 Problems

Problem 7.6.1. The Cookie Store sells three types of biscuit box: chocolate chip, pecan chip and twists. All biscuits are made with chocolate, nuts and sugar. The store has 312 pounds of sugar, 125 pounds of chocolate and 50 pounds of nuts. For marketing reasons, at least 5 chocolate boxes are to be produced. The ingredients of each type of biscuit box is shown in Table 7.4. Revenues for each type of box is shown in Table 7.5.

ingredient	choc. chip	pecan chip	twists
sugar	21	15	9
chocolate	10	5	0
nuts	1	2	0

Table 7.4: Ingredients in biscuits

box type	revenue
chocolate chip	20
pecan chip	25
twists	17

Table 7.5: Unit revenue for each type of biscuit box

Write a LP model to find the optimal production plan to maximize revenues. If some biscuits are left after each box is packed, they will be sold at the same price at a coffee shop nearby. Conduct a sensitivity analysis and answer the following questions:

1. How many biscuits it is optimal to cook? What is the total revenue?
2. Which are the not binding constraints?
3. How many pound of nuts are needed in the optimal solution?
4. Will you suggest to the Store manager to produce more chocolate biscuits? Why?
5. What should the pecan chips price be to make it profitable to produce them?
6. If we had 7 more pounds of sugar, what would the optimal total profit be?
7. What will happen to the optimal solution if the price of a chocolate box increases to 24? Why? What would then be the total revenue?

Problem 7.6.2. A farmer must determine how many acres of corn and wheat to plant. An acre of wheat yields 5 tons of wheat and requires 6 hours of labor per week. An acre of corn yields 4 tons of corn and requires 10 hours of labor per week. All wheat can be sold at 30 euros a ton, and all corn can be sold at 50 euros a ton. 45 acres of land and 350 hours per week of labor are available. Government regulations require that no more than 140 tons of wheat and no more than 120 tons of corn be produced during the current week. Formulate a LP to tell the farmer how to maximize the total revenue from wheat and corn. Then:

1. Perform a sensitivity analysis
2. If the farmer had only 40 acres, how much would the revenue be?
3. If the wheat price dropped to 26, what would the optimal solution be?

4. Determine the allowable increase and decrease ranges within which the optimal solution does not change. If only 130 tons of wheat could be sold, would the answer to the original problem change?

Problem 7.6.3. A bank must decide the composition of the portfolio of the bonds to sell. Three bonds can be sold at unit prices equal to 92 (bond 1), 95 (bond 2) and 96 (bond 3), in dollars. The bank target is to maximize revenue from the selling of the bonds, fulfilling the following constraints:

- at least 70% of the bond should be of type 1 or 2;
- bonds of type 1 must not be more than three times bonds of type 2;
- the portfolio should have no more than 100 bonds.

With these data:

1. formulate a LP problem and solve it, determining the composition of the optimal portfolio, its revenue and the set of binding constraints.
2. perform a sensitivity analysis
3. if the bond 2 price raised to 96, would the portfolio change? How much would the revenue change?
4. if the bond 3 price raised to 100, would the portfolio change? How much would the revenue change?
5. if the bank issued 19 more bonds, how much would the revenue increase? Which type of bonds would be issued? How much should the bond 1 price increase for the bond to be issued by the bank?

Problem 7.6.4. Orange Krush sells oranges crates and orange juice in bottles. For each kg of oranges used to get juice, the company gets a revenue of 1.5 euro for a cost of 1.05 euro. For each kg of oranges used for crates, the company gets a revenue of 0.50 euro for a cost of 0.20 euro. 100000 kg of oranges are available, the 40% of which must be used to produce juice. In addition, at least 20000 kg of oranges must be sold in crates. The problem is to:

1. find the optimal sell plan, maximizing the margin and its value.
2. find the set of binding constraints.
3. find the allowable profit increase and decrease for each kg of oranges to keep the optimal solution unchanged.
4. is the minimum crates production constraint binding, given that its shadow price is negative? Why is it so?

7.7 Quickies

Question 7.7.1. In a production problem, a certain raw material must be less than or equal to 100. The optimal profit is 100. The shadow price for that material is 2. What happens if we buy one more unit of the raw material, provided this change is within the allowable range?

Question 7.7.2. In a production problem, a certain raw material must be less than or equal to 100. The optimal profit is 100. The shadow price for that material is 2. Will you suggest the manager to buy an additional unit of raw material on the market at the price of 3?

Question 7.7.3. In an objective function the coefficient of a decision variable is 5 and its allowable change is $(-2, +3)$. If the coefficient increases by 1, does the objective change?

Question 7.7.4. What is the reduced cost of a variable in a LP problem?

Question 7.7.5. If a constraint is not binding, can the shadow price be different from 0?

Chapter 8

LP Examples

This is a brief review of some examples of application of linear programming. I focused on the construction of the model and not on its solution which, as I said in the preface, is thought to be done with the help of a computer.

8.1 Minimum problems

Example 8.1.1. A fashion company has recently started a new top brand. The marketing campaign is largely based on advertising on a satellite channel. Advertisements will be broadcast on sports channels (for men's clothing) and movies channels (for women's clothing). Each minute of advertisement on the sport channel costs 100000 euros and each minute on the movie channel costs 50000 euros. Each ad on the sport channel is watched by 2 millions females and 12 millions males. Each commercial on the movie channel is watched by 7 millions females and 2 millions males.

The company has a target to reach 28 millions women and 24 millions men. How many commercials should be broadcast on each channel to reach the target at the minimum cost? In formulating the problem take into account that the previous numbers are average estimates and the commercial can last for whatever amount of time, even less than a minute.

Decision variables are easy to determine: let x_1 be the number of commercials on the movie channel and x_2 the number of commercials on the sport channel. Obviously, $x_1, x_2 \geq 0$ and the last information about the commercial implies that x_1 e x_2 are real numbers. The total cost is $50x_1 + 100x_2$ and must be minimized. With x_1 ads broadcast on movie channels the company reaches $7x_1$ millions women and with x_2 ads on sports channel the company reaches $2x_2$ millions women for a total of $7x_1 + 2x_2$ millions women. This must be greater than 28 millions, i.e.

$$7x_1 + 2x_2 \geq 28$$

In the same way, for men we have

$$2x_1 + 12x_2 \geq 24$$

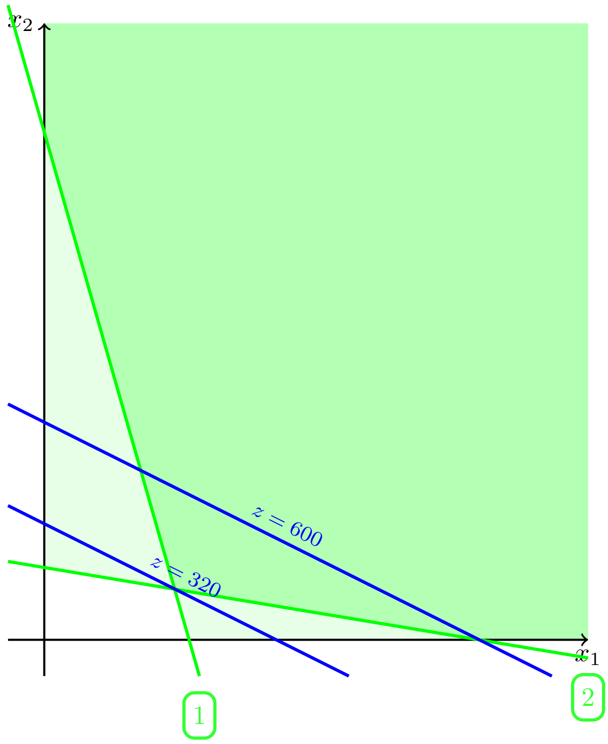


Figure 8.1: Example 8.1.1: feasible set and objective function for $z = 600$ and $z = 320$.

The LP problem is

$$\begin{array}{rll} \min & 50x_1 + 100x_2 & \\ \text{s.t.} & 7x_1 + 2x_2 \geq 28 & \text{women} \\ & 2x_1 + 12x_2 \geq 24 & \text{men} \\ & x_1, x_2 \geq 0 & \end{array}$$

This is an example of a minimum cost problem. A plan must be devised fulfilling a set of “greater than” constraints and minimizing a cost function. Solving the problem with the graphical method we get Figure 8.1.

In this type of problems, the feasible set is usually unbounded. The reason is that there is always a solution that satisfies all the constraints. The question is to find the solution the has the minimum cost. In this case, the optimal solution is $(3.6, 1.4)$ with the objective value equal to 320.

The so called “diet problems” belong to this category as well. In a diet problem, we consider a number of foods, say meat, milk, fruit, etc. and their constituent elements, say vitamins, proteins, sugars, fats, minerals, etc.. Each organism needs a minimum amount of the elements per day and to get them eats various food. A good diet is one who provides the organism with the right quantity of vitamins, etc. Among all possible good diets, the best one is the less expensive. Frequently, the composition of a diet is given as a percentage of the total food, as in the following example.

oil	1	2	3	4
% di A	20	40	70	80
% di B	20	10	60	70

Table 8.1: Substances A e B in basic oils 1, 2, 3 and 4

oil	1	2	3	4
Unit cost	1.5	2.2	4.3	4

Table 8.2: Unit cost of basic oils

Example 8.1.2. To produce burning oil, a factory mixes 4 basic oils. Each basic oil contains 2 substances, A and B that must be in the final mix as respectively its 30% and its 40%. The composition in basic oils is shown in Table 8.1.

Unit cost for basic oils are shown in Table 8.2.

Determine the less expensive mix that satisfies the constraints on A and B.

Let $x_1, \ldots, x_4$ the percentages of basic oils in the mix. If x_1 is the percentage of basic oil 1 in the mix and oil 1 contains 20% of A, the contribution of oil 1 to the quantity of A is $0.2x_1$ and similarly for other basic oils and for substance B. The constraints on substances are therefore

$$\begin{aligned} 0.2x_1 + 0.4x_2 + 0.7x_3 + 0.8x_4 \geq 0.3 \\ 0.2x_1 + 0.1x_2 + 0.6x_3 + 0.7x_4 \geq 0.4 \end{aligned}$$

As $x_1, \ldots, x_4$ are percentage of the mix, we must add a constraint asking their sum is 1:

$$x_1 + x_2 + x_3 + x_4 = 1$$

The LP problem is:

$$\begin{array}{rll} \min & 1.5x_1 + 2.2x_2 + 4.3x_3 + 4x_4 & \\ \text{s.t.} & 0.2x_1 + 0.4x_2 + 0.7x_3 + 0.8x_4 \geq 0.3 & \text{(Substance A)} \\ & 0.2x_1 + 0.1x_2 + 0.6x_3 + 0.7x_4 \geq 0.4 & \text{(Substance B)} \\ & x_1 + x_2 + x_3 + x_4 = 1 & \text{(Percentages)} \\ & x_i \geq 0, i = 1, 2, 3, 4. & \end{array}$$

and its solution is $x_1 = 0.6$, $x_2 = 0$, $x_3 = 0$ e $x_4 = 0.4$.

8.2 Integer programming

Decision variables are not always real numbers. If, e.g., a variable represents the number of employees in a company, non integer numbers have no meaning at all. The request that variables are only integers radically changes the problem into a new one and its solutions.

When one or more than one variable must take only integer values, we speak of an Integer Programming problem (IP). To each IP problem, one can always

point	obj.
$(0,0)$	0
$(0,1)$	5
$(0,2)$	10
$(1,0)$	1
$(1,1)$	6
$(2,0)$	2
$(2,1)$	7

Table 8.3: Objective values at the feasible points in ex. 8.2.1.

associate a LP problem, that obtained by removing the integer constraint. This is the "relaxed" problem of an IP.

The first approach to an IP is to solve the relaxed LP problem and remove the non integer parts of the solution. To do this, there are usually two techniques:

- truncation, the removal of the digits after the decimal point, e.g. truncating $e = 2.718\ldots$ we get 2;
- rounding, the removal of the digits after the decimal point with, possibly, the substitution of the last one if the subsequent is greater than 4, e.g. rounding $e = 2.718\ldots$ we get 3.

The following example shows that both techniques produce wrong solutions.

Example 8.2.1. Consider the problem

$$\begin{array}{rl} \max & x + 5y \\ \text{s.t.} & x \leq 2 \\ & x + 10y \leq 20 \\ & x, y \geq 0 \text{ integers} \end{array}$$

As $x \leq 2$ and x must be integer and non negative, x could be $x = 0, 1, 2$. Therefore, if $x = 0$, then y can be $y = 0, 1, 2$ (or the inequality $x + 10y \leq 20$ is not true); if $x = 1$ or $x = 2$, y could be $y = 0, 1$ (for the same reason). The only feasible solutions are: $(0,0)$, $(0,1)$, $(0,2)$, $(1,0)$, $(1,1)$, $(2,0)$, $(2,1)$. The value of the objective function in these points are shown in Table 8.3. The optimal solution is $(0,2)$.

The relaxed problem is:

$$\begin{array}{rl} \max & x + 5y \\ \text{s.t.} & x \leq 2 \\ & x + 10y \leq 20 \\ & x, y \geq 0 \end{array}$$

and its graphical solution is shown in Figure 8.2.

From the figure we see that the relaxed solution is $(2, 1.8)$ with objective equal to 11. The rounded solution is $(2,2)$, which is not feasible (the second constraint is false). The truncated solution is $(2,1)$, feasible but not optimal: objective is 7. Finally, note that the optimal solution, $(0,2)$, is "far" from that of the relaxed problem.

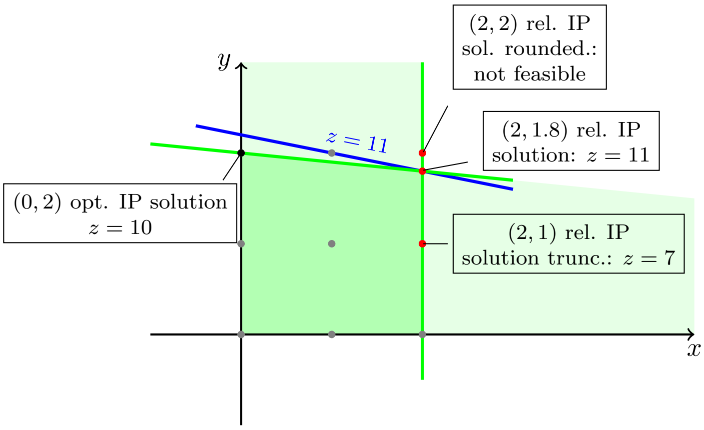


Figure 8.2: Graphical solution of ex. 8.2.1. Solutions of relaxed problem are in red; IP solution in black

fund	cost	revenue
1	5000	16000
2	7000	22000
3	4000	12000
4	3000	8000

Table 8.4: Costs and revenues of funds in ex. 8.3.1

In real IP problems, with hundredths of constraints and variables, the graphical approach is impossible. In these problems, the integer coordinates points grow exponentially with the number of decision variables and to look at the objective value at each of them would require an infinite time[1]. To solve IP problems, many very clever algorithms have been invented and are now implemented in many commercial and open source software.

8.3 Binary variables

IP problems contain an important subset, those in which variables can take only two values, 0 and 1. These variables are called *binary variables*. Their use is shown in the following example.

Example 8.3.1. 14000 euros are available to invest in one or more funds chosen in a group of 4. Funds cost and revenues are shown in Table 8.4. It is not possible to invest in more than one share of the same fund (e.g., one can not buy two shares of fund 1). Find the mix that maximizes the revenues.

[1]this is not a joke. Let us take the $[0, 2]$ interval as the range of some decision variables. With one variable, the are only 3 points, 0, 1 e 2. With two variables, there are 9 points, on a square from $(0, 0)$ to $(2, 2)$. With three variables we have a cube with 27 points, from $(0, 0, 0)$ to $(2, 2, 2)$. With 60 variables, we would have 3^{60} integer points, approximately 10^{41}. The most performing computer now can do 10^{15} operations per second(source: `http://www.top500.org/`. To look at each integer coordinates point, such a computer would need $10^{41}/10^{15} = 10^{26}$ seconds, i.e. 10^{19} years, one billion times the age of the universe.

time slot	minimum policemen
0-4	40
4-8	35
8-12	60
12-16	55
16-20	50
20-24	50

Table 8.5: Minimum number of policemen for each time slot

For each fund we must decide whether to buy it or not. Let x_1 be a variable that takes the value 1 if we buy fund 1 or 0 if we do not buy it. Total revenue is given by

$$16x_1 + 22x_2 + 12x_3 + 8x_4$$

(in thousands of euros) and must be maximized. The only constraint is that of budget. Only 14000 euros are available so

$$5x_1 + 7x_2 + 4x_3 + 3x_4 \leq 14$$

(again in thousands of euros). The IP problem is

$$\begin{array}{rl} \max & 16x_1 + 22x_2 + 12x_3 + 8x_4 \\ \text{s.t.} & 5x_1 + 7x_2 + 4x_3 + 3x_4 \leq 14 \\ & x_1, x_2, x_3, x_4 \text{ binary} \end{array}$$

The optimal solution is $x_1 = 0$ and $x_2 = x_3 = x_4 = 1$, i.e. buy one share of funds 2, 3, and 4 with a revenue of 42000 euros. Note that fund 1, which does not appear in the optimal solution, is the one with the highest return, 220%, while we have 214% for fund 2, 200% for fund 3 and 166% for fund 4.

8.4 Assignment problems

Another example of application of LP are assignment problems, as the following example.

Example 8.4.1. a police district has a number of agents providing security 24 hours a day. Policemen work on 4-hours shift or 8-hours shift. Each day the are 6 slots of 4 hours and in each slot a minimum number of policemen must be on duty, according to Table 8.5.

Costs for each shift are shown in Tables 8.6 and 8.7.

The same structure repeats everyday. Find the number of policemen to be assigned to each shift to minimize costs.

We have integer decision variables, corresponding to the number of agents working on a given shift. We have two types of shift, 4 hours and 8 hours, and for each shift there are 6 possible times to start. Let x_0 be the number of policemen with a 4-hours shift starting at midnight, x_4 the number of policemen with a 4-hours shift starting at 4, x_8 the number of policemen with a 4-hours shift starting at 8 AM, and so on till x_{20}. In the same way, let y_0 be the number of policemen with a 8-hours shift starting at midnight, etc. till y_{20}.

time slot	cost per policeman
0-4	6
4-8	6
8-12	4
12-16	4
16-20	4
20-24	5

Table 8.6: Policeman cost for 4-hours shift

time slot	Policeman cost
0-8	15
4-12	15
8-16	8
12-20	8
16-24	8
20-4	11

Table 8.7: Policeman cost for 8-hours shift

Policemen on duty between 8 AM and noon? Those who start at 8 AM, x_8 e y_8, and those who started at 4 AM with a 8-hours shift, i.e. y_4. Summing up, policemen on duty from 8 to noon are $x_8 + y_8 + y_4$ and this number should be greater than or equal to 60.

For the other time slots we make the same calculations (note that policemen starting at 8 PM on a 8-hours shift are on duty until 4 AM of the next day).

The objective is to find the minimum of

$$6x_0 + 6x_4 + 4x_8 + 4x_{12} + 4x_{16} + 5x_{20} + 15y_0 + 15y_4 + 8y_8 + 8y_{12} + 8y_{16} + 11y_{20}$$

and the IP problem is

$$\begin{array}{lll}
\min & 6x_0 + 6x_4 + 4x_8 + 4x_{12} + 4x_{16} + 5x_{20}+ & \\
 & 15y_0 + 15y_4 + 8y_8 + 8y_{12} + 8y_{16} + 11y_{20} & \\
\text{s.t.} & x_0 + y_0 + y_{20} \geq 40 & \text{(shift 0-4)} \\
 & x_4 + y_0 + y_4 \geq 35 & \text{(shift 4-8)} \\
 & x_8 + y_4 + y_8 \geq 60 & \text{(shift 8-12)} \\
 & x_{12} + y_8 + y_{12} \geq 55 & \text{(shift 12-16)} \\
 & x_{16} + y_{12} + y_{16} \geq 50 & \text{(shift 16-20)} \\
 & x_{20} + y_{16} + y_{20} \geq 50 & \text{(shift 20-24)} \\
 & x_i, y_i \geq 0, i = 0, 4, 8, 12, 16, 20, \text{ integers} &
\end{array}$$

The solution is $x_0 = 40$, $x_4 = 35$, $x_8 = 5$, $y_8 = 55$, $y_{16} = 50$ and all other variables equal to 0. The minimum cost is 1310.

8.5 Problems

Problem 8.5.1. A pharmaceutical company has two main stores, A and B, with respectively 100 and 250 boxes of a certain drug. The stores supply three

	from A	from B
to C	5	3
to D	10	4
to E	12	9

Table 8.8: Transportation costs in ex. 8.5.1.

wholesalers, C, D and E, who requested, respectively, at least 50, 120, and 180 boxes of the drug. Transportation costs per box are shown in Table 8.8. Find the optimal plan.

Problem 8.5.2. A food company makes two types of marmalade, with oranges or with lemons. Fruits are bought at 0.3 euros per kg. Each kg is processed for an hours before being available for packaging.

Each kg of fruit provides 3 hg of lemon marmalade and 4 hg of orange marmalade. Lemon marmalade is sold at 0.7 euro per hg and that of oranges is sold at 0.6 euro per hg.

Applying a special industrial process, starting from "base" marmalades, one can get "flavored marmalades" ("cinnamon orange" and "lemon and lime"). Lemon and lime is sold at 1.8 euro per hg and cinnamon orange at 1.4 euro per hg. In the special process, no product is lost but there is an additional cost (0.4 euro per hg) and more time is needed (3 hours for lemon marmalades and 2 hours for orange).

Each year, the company has 6000 labor hours and can buy a maximum of 4000 kg of fruit. Formulate a LP problem to maximize profits.

8.6 Quickies

Question 8.6.1. What is a binary variable?

Question 8.6.2. What is a IP problem?

Question 8.6.3. A IP problem has been solved and the optimal value of the objective function is 2.78. Is this possible?

Question 8.6.4. In a LP problem, the feasible set is unbounded. Might the problem admit just one solution or the number of solutions is necessarily infinite?

Question 8.6.5. The solution to an IP problem is obtained by solving the corresponding LP problem and then rounding the solution. True or false?

Solutions to exercises

Probability

1.6.1 The value of the buildings in millions of euros is a random variable X whose values are $x_1 = 13$, $x_2 = 11$ and $x_3 = 10$ with probabilities $p_1 = 0.3$, $p_2 = 0.5$ and $p_3 = 0.2$. The expected value of x is

$$\mathrm{E}\,(X) = 13 \times 0.3 + 11 \times 0.5 + 10 \times 0.2 = 11.4$$

1.6.2 Let S be "the travel is successful" and "+" be "the members' opinion is positive". Before the arrival of the two, we have $P(S) = 0.8$ and therefore $P(\overline{S}) = 0.2$. Moreover, we know that $P(+|S) = 0.9$ and $P(\overline{+}|\overline{S}) = 0.8$. From these equations we compute $P(\overline{+}|S) = 0.1$ and $P(+|\overline{S}) = 0.2$.

As the members' opinion is still unknown, we have to compute both $P(S|+)$ and $P(S|\overline{+})$. Using Bayes' theorem we have

$$\begin{aligned} P(S|+) &= \frac{P(+|S)P(S)}{P(+|S)P(S) + P(+|\overline{S})P(\overline{S})} \\ &= \frac{0.9 \times 0.8}{0.9 \times 0.8 + 0.2 \times 0.2} = \frac{0.72}{0.76} \\ &= \frac{18}{19} \simeq 0.9473 \end{aligned}$$

and

$$\begin{aligned} P(S|\overline{+}) &= \frac{P(\overline{+}|S)P(S)}{P(\overline{+}|S)P(S) + P(\overline{+}|\overline{S})P(\overline{S})} \\ &= \frac{0.1 \times 0.8}{0.1 \times 0.8 + 0.8 \times 0.2} = \frac{0.08}{0.24} \\ &= \frac{1}{3} \simeq 0.3333 \end{aligned}$$

Decision problems

3.5.1 There are 4 nodes: one decision node (the Price List), two event nodes (the Competitor price list and the orders)and one result node (the final Profit). The profit depends on the other 3 nodes. The company price list depends on the competitor price list, because the decision is made when the competitor list is known. Orders depend on both price lists. The ID is displayed in Figure 1.

4.4.1

1. The DT is shown in Figure 2. The optimal decision is to do the research. The expected profit is 10 millions euros.
2. to compute the value of perfect information, look at the DT in Figure 3. The root node value is 19 millions euros and so the value of perfect information is 9 millions euros.
3. from the DT in Figure 4 we know that the value of sample information is 6.7 millions euros. So we should not ask the opinion to the expert .

4.4.2 The DT is shown in Figure 5. Note that:

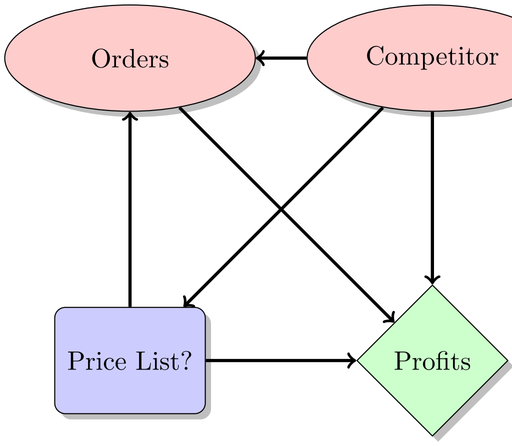


Figure 1: Influence diagram for problem 3.5.1

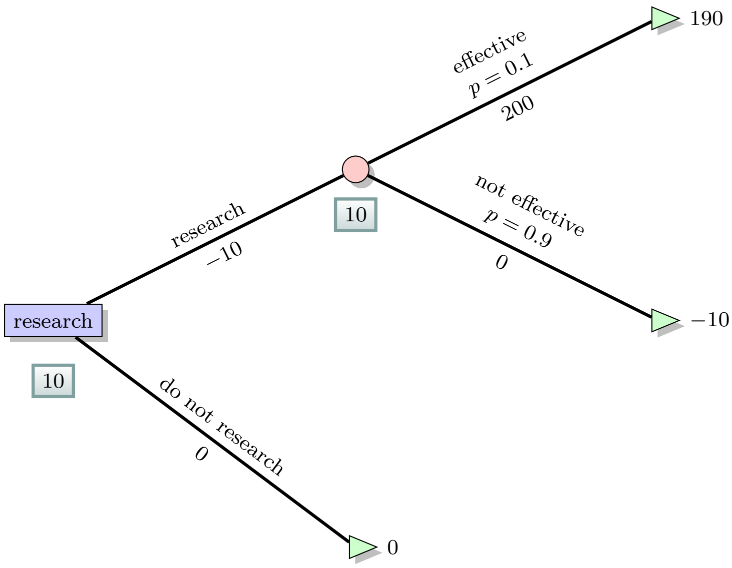


Figure 2: DT for problem 4.4.1.

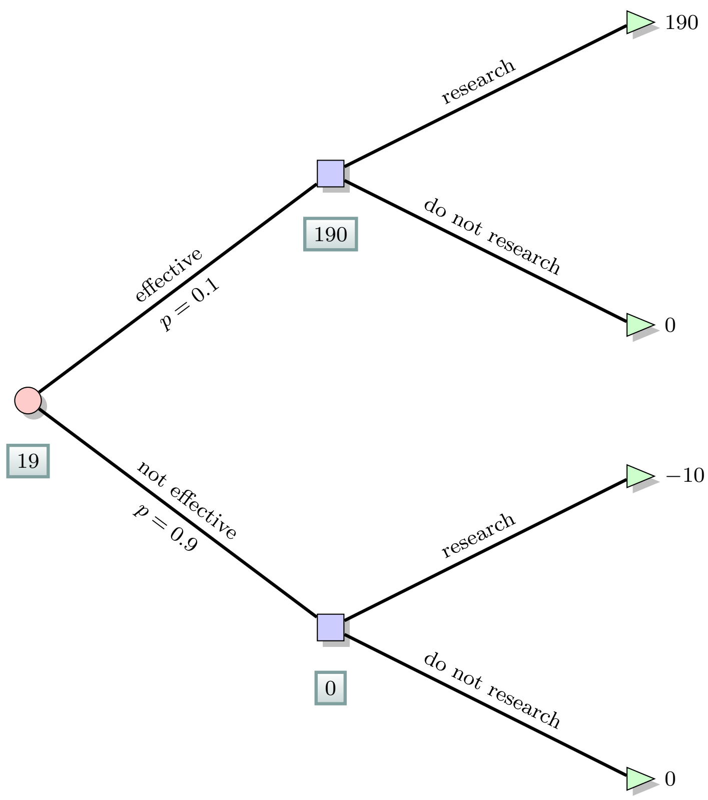


Figure 3: DT for problem 4.4.1 with perfect information.

- the company may withdraw even when it has been shortlisted
- the material cost is to be taken into account only when the bid is accepted
- acceptance percentage decreases if the bid increases

4.4.3 The ID is shown in Figure 6 and the corresponding DT is that in Figure 7.

The use of software transform the scenario in one of perfect information. The new DT is shown in Figure 8.

The expected value of perfect information is $240 - 40 = 200$ and this is the maximum price to pay for the software.

Interviews are imperfect information. Let S be the event "employee bets"; we have $P(S) = 0.6$ and $P(\overline{S}) = 0.4$. Let C be the event "colleagues say he bets". We have $P(C|S) = 0.5$ and $P(\overline{C}|\overline{S}) = 0.95$. As sources of information, the employee's colleagues are very reliable when he does not bet and completely unreliable when he bets. Applying Bayes' theorem (see 1.3.2) we compute the needed conditional probabilities, i.e.

$$P(S|C) = \frac{P(C|S)P(S)}{P(C|S)P(S) + P(C|\overline{S})P(\overline{S})} = \frac{0.5 \times 0.6}{0.5 \times 0.6 + 0.05 \times 0.4} = 0.9375$$

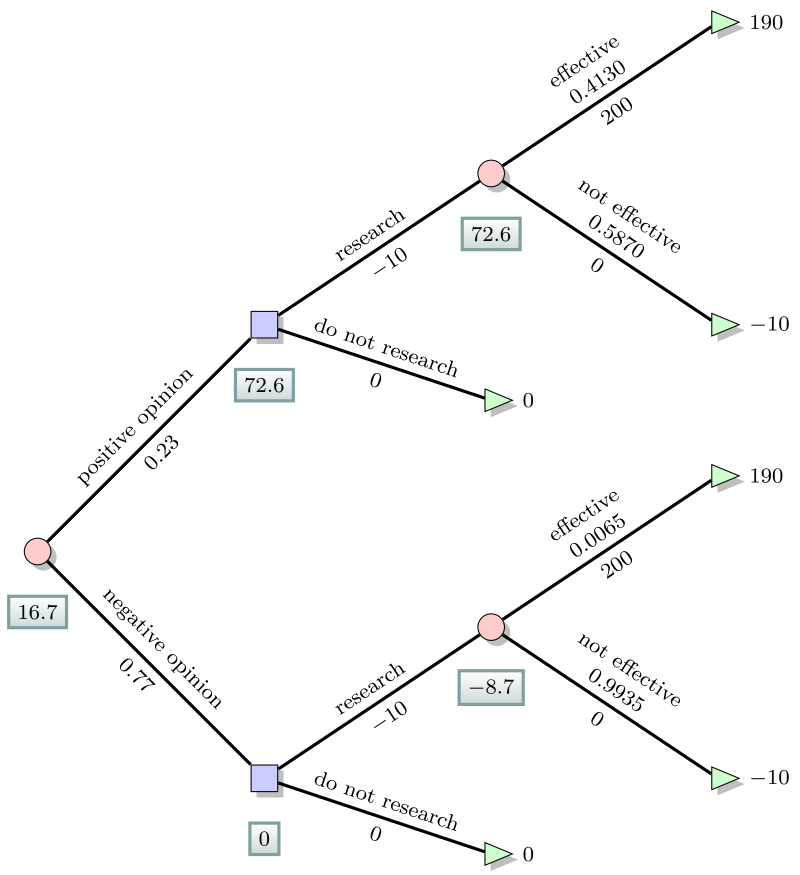


Figure 4: DT for problem 4.4.1 with sample information.

and

$$P(\overline{S}|\overline{C}) = \frac{P(\overline{C}|\overline{S})P(\overline{S})}{P(\overline{C}|\overline{S})P(\overline{S}) + P(\overline{C}|S)P(S)} = \frac{0.95 \times 0.4}{0.95 \times 0.4 + 0.5 \times 0.6} \simeq 0.5588.$$

The decision tree is shown in Figure 9.

The root node is worth 110 euros. Thus, the EVSI is $110 - 40 = 70$ euros. As the cost of interviewing the other employees is 150 euros, the manager should not interview his colleagues.

5.3.1 Let $u(x) = 1 - e^{-x/R}$ be the utility function. The inverse function is $u^{-1}(x) = -R\ln(1-x)$. The "new product" option has certainty equivalent

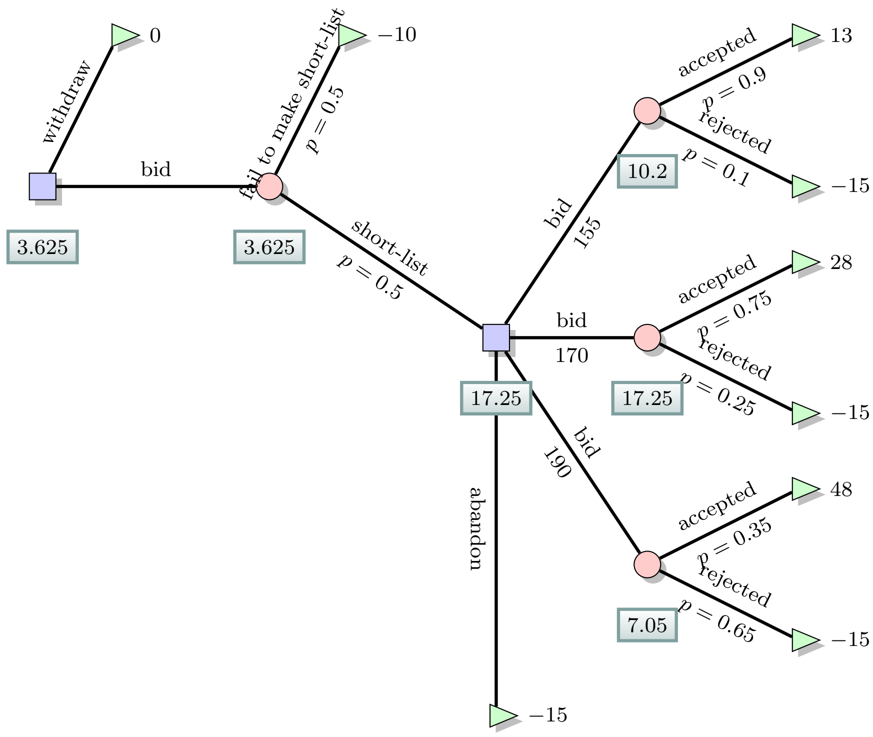


Figure 5: DT of problem 4.4.2.

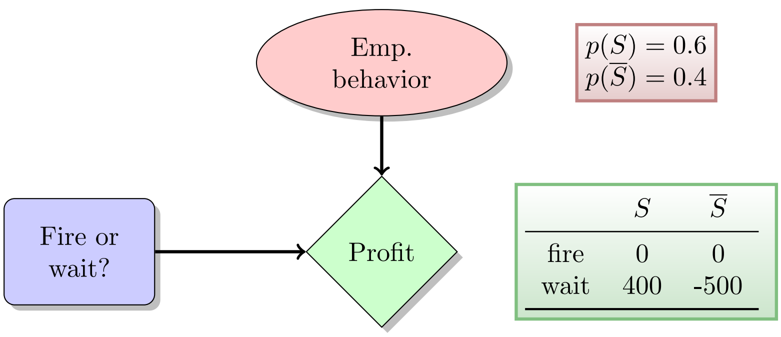


Figure 6: ID for problem 4.4.3. S is the event "employee bets".

given by

$$\begin{aligned}
z &= u^{-1}(0.7u(380) + 0.3u(180)) \\
&= -R\ln\left(1 - 0.7(1 - e^{-380/R}) - 0.3(1 - e^{-180/R})\right) \\
&= -R\ln\left(0.7e^{-380/R} + 0.3e^{-180/R}\right) \\
&= -R\ln\left[e^{-380/R}(0.7 + 0.3e^{200})\right] \\
&= -R\ln e^{-380/R} - R\ln\left(0.7 + 0.3e^{200/R}\right) \\
&= 380 - R\ln\left(0.7 + 0.3e^{200/R}\right)
\end{aligned}$$

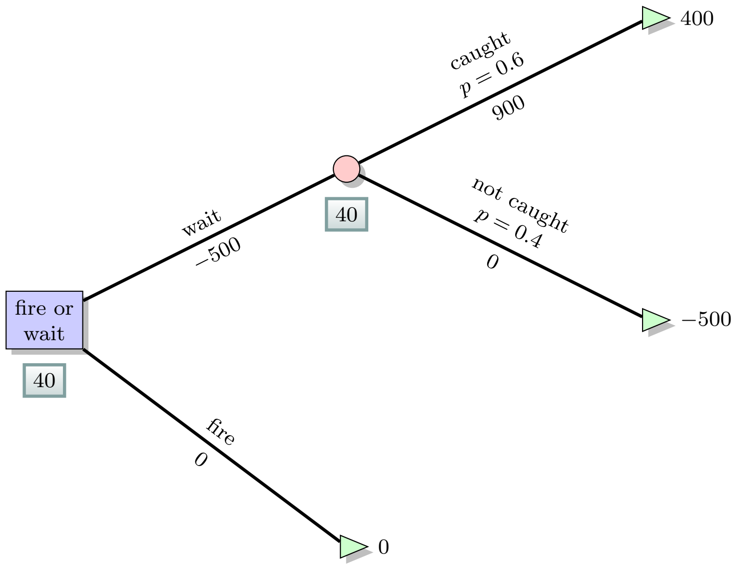


Figure 7: DT for problem 4.4.3.

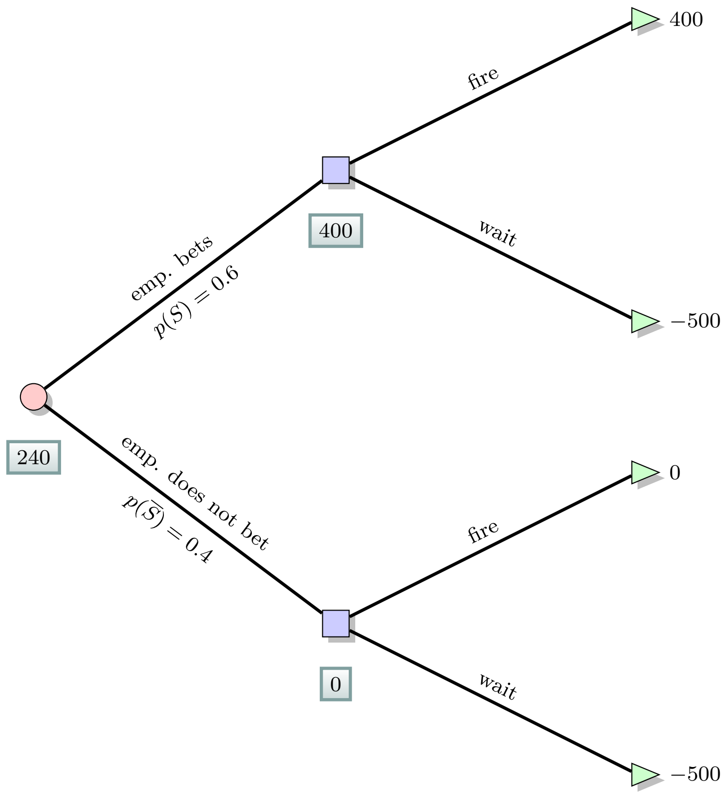


Figure 8: DT of problem 4.4.3 with perfect information.

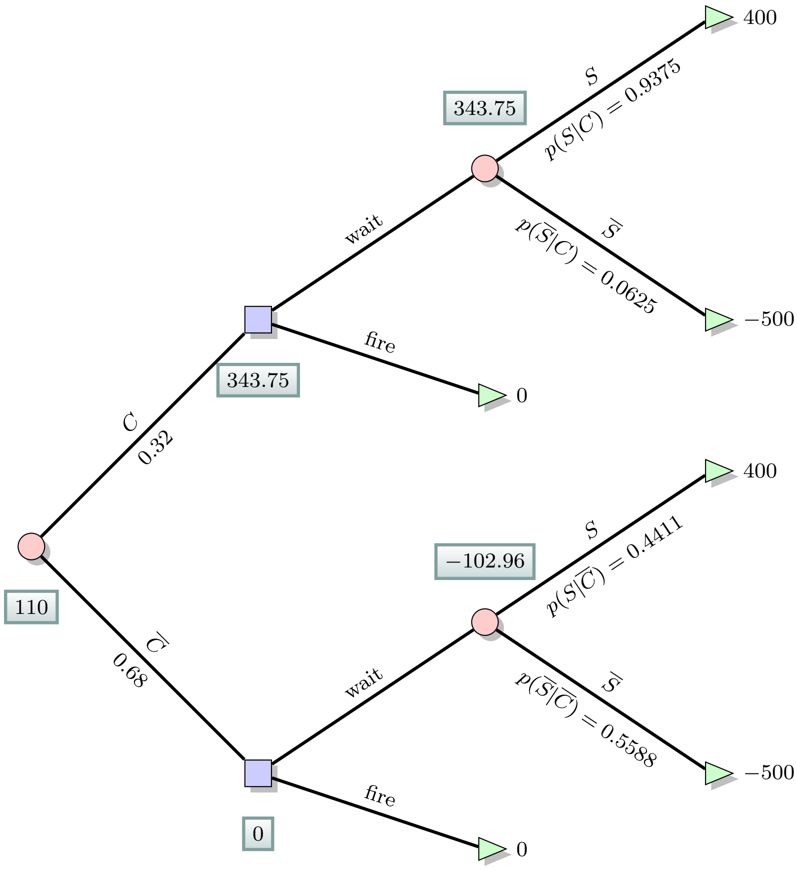


Figure 9: DT for problem 4.4.3 with sample information.

To compare this with the "certain" option, worth 300, we should compare $g(R) = R \ln\left(0.7 + 0.3e^{200/R}\right)$ with 80. Using a computer, we list some values of g (see Table 1). Table 1 shows that 80 corresponds to a value of R^* between 200 and 250. A numerical comparison shows that $R^* \approx 229$.. Thus

- for a subject with $R < 229$ (more risk averse), the optimal decision is not to develop a new product
- for a subject with $R > 229$ (less risk averse), the optimal decision is to develop a new product
- for a subject with $R \approx 229$ both decisions are equivalent.

R	$g(R)$
10	187.960272
50	141.893784
100	107.045861
150	91.309894
200	83.147044
250	78.275729
300	75.068868
350	72.807196
400	71.130045

Table 1: Values of $g(R) = R \ln\left(0.7 + 0.3e^{200/R}\right)$.

Linear programming

6.4.1 Let x and y be the tons of production of pipes of type A and AA. The problem can be formulated as follows:

$$\begin{array}{rll} \max & 31x + 39y & \\ \text{s.t.} & x + y \geq 10000 & \text{(min. total prod.)} \\ & 5x + 7y \leq 56000 & \text{(max work hours)} \\ & x, y \geq 0 & \text{(non negative variables)} \end{array}$$

The graphical solution is shown in Figure 10

the objevtive is 334000 in $(7000, 3000)$, 310000 in $(10000, 0)$ and 347200 in $(11200, 0)$. The maximum revenue is reached when the production is 11200 tons of type A pipes and no type AA pipes.

6.4.2 Let x and y be the kg to buy each month from supplier 1 and 2. The LP model is :

$$\begin{array}{rll} \min & 20x + 30y & \\ \text{s.t.} & 5x + 10y \geq 90 & \text{(minimum di A)} \\ & 4x + 3y \geq 48 & \text{(minimum di B)} \\ & 0.5x \geq 1.5 & \text{(minimum di C)} \\ & x, y \geq 0 & \text{(non negative var.)} \end{array}$$

The graphical solution is shown in Figure 11

The objective is 420 in $(3, 12)$, 312 in $(8.4, 4.8)$ e 360 in $(18, 0)$. The minimum is reached when buying 8.4 kg from supplier 1 and 4.8 kg from supplier 2.

7.6.1 There are 3 decision variables, the number of boxes to produce. As left biscuits are being sold at the same price, we should take into account non integer solutions as well.

1. The optimal solution is to produce 5 chocolate chip boxes, 23 twists boxes and no pecan chip box. Total revenue is 491.
2. Not binding constraints are those of chocolate and nuts, because the optimal solution does not exhaust them.

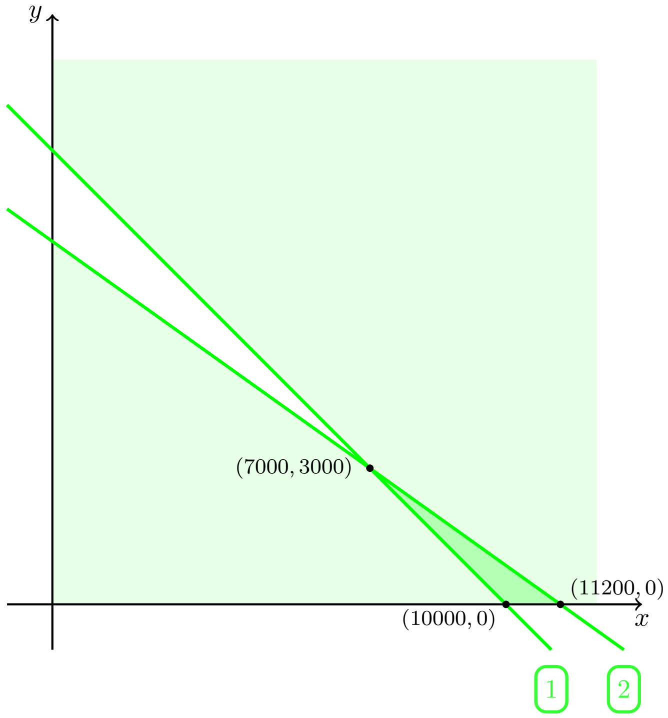


Figure 10: Graphical solution to problem 6.4.1.

3. In the optimal solution 5 pounds of chocolate are used (45 are left).

4. The Store should not modify the optimal plan, because it is, obviously, optimal.

5. Pecan chips are missing from the optimal solution because their revenue is too low. To make them appear in the optimal solution, their unit revenue should increase by (minus) the reduced cost 3.33 to reach 28.33.

6. Sugar is a binding resource and therefore with more sugar available the optimal solution could change. Allowable increase for sugar is $+\infty$ and 7 is thus within this bound. As sugar shadow price is 1.89 the revenue will increase by $7 \times 1.89 = 13.22$.

7. Allowable increase for chocolate is 19.66, thus the proposed increase (from 20 to 24) leaves the optimal plan unchanged. Revenues will by $5 \times 4 = 20$ to 511.

7.6.2 1. Let x_1 be the number of acres planted with wheat and x_2 those planted with corn. The revenue for each acre is the the revenue per ton

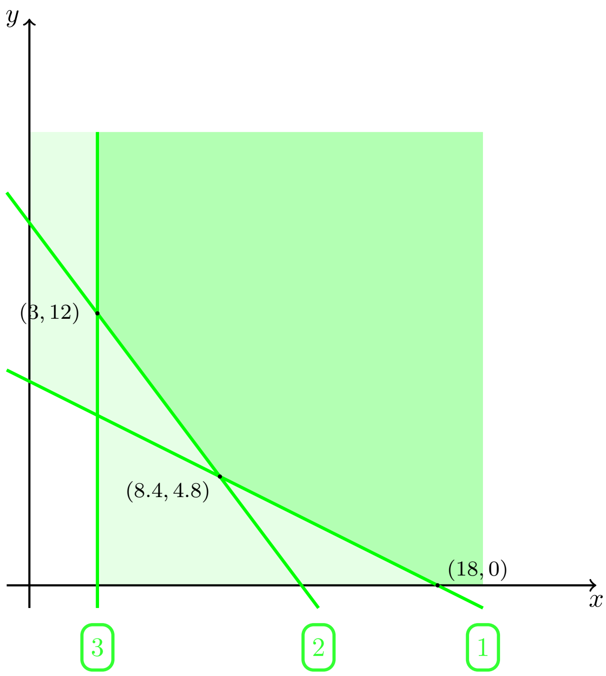


Figure 11: Graphical solution of problem 6.4.2.

times the tons per acre. The LP model is

$$
\begin{array}{rll}
\max & 150x_1 + 200x_2 & \\
\text{s.t.} & x_1 + x_2 \leq 45 & \text{(total area)} \\
& 5x_1 \leq 140 & \text{(max sell wheat)} \\
& 4x_2 \leq 120 & \text{(max sell corn)} \\
& 6x_1 + 10x_2 \leq 350 & \text{(max labor hours)} \\
& x_1, x_2 \geq 0 &
\end{array}
$$

and its optimal solution is $x_1 = 25$ and $x_2 = 20$, with a revenue of 7750. Wheat and corn constraints are not binding, with slacks 15 t of wheat and 40 t of corn. Area and labor constraints are binding.

2. x_1 coefficient range of allowable change is $(-30, +50)$; that of the coefficient of x_2 is $(-50, +50)$. Area shadow price is 75, with allowable increase of 1.2 and allowable decrease of 6.66; labor hours shadow price is 12.5, with allowable increase of 40 and allowable decrease of 12.

3. If the farmer had only 40 acres there would be a decrease of 5, which is allowable. Revenue will decrease by 75 (reduced cost) times 5 (change), i.e. 375, to 7375.

4. If the wheat price dropped to 26 the revenue per unit would be 130, 20 less than before. This decrement is allowable and the revenue would decrease by 20 times 25 (the number of acres planted with wheat), $25 \times 20 = 500$, to 7250.

5. As the allowable decrease is 15, moving from 140 to 130 will not change the solution.

Note that the problem could have been solved by the graphical method.

7.6.3 1. The optimal portfolio is made up of 70 bonds of type 2 and 30 of type 3, with a revenue of 9530. The binding constraints are the first and the third.

2. If the price of bond 2 raised to 96, an increment of 1, allowable by the sensitivity analysis, the optimal solution would not change and the revenue will increase by one for each bond 2, i.e. by 70 dollars.

3. The increase of bond 3 price to 100 is allowable and then the solution will not change and the profit will increase by 120 dollars.

4. The issue of 10 more bonds is allowable and will not change the optimal solution. The bonds will be shared proportionally to bond 2 and bond 3 types, according to the current optimal plan (i.e. 7 of type 2 and 3 of type 3). Bond 1 has a reduced cost of -3 and therefore its price must increase to 95 to make it enter the optimal solution.

7.6.4 1. The optimal plan is to produce 20000 kg of orange crates and 80000 kg of orange juice. The maximum profit is 42000.

2. Total oranges and minimum crates constraints are binding.

3. Orange crates: unit profit can decrease by any amount or increase by less than 0.15; orange juice unit profit can increase by any amount or decrease by less than 0.15.

4. Yes. Let C be the optimal quantity of orange to be sold in crates, $C \geq 20000$. If $C > 20000$ (not binding constraint), then 20000 could be increased without changing the optimal solution and the optimal profit. A negative shadow price points out that by increasing 20000 by one we get a reduction in the profit. Thus, a negative shadow price is not compatible with $C > 20000$ and therefore $C = 20000$, i.e. the constraint is binding.

8.5.1 The total quantity available at the stores, $350 = 100 + 250$, is equal to the quantity demanded by the wholesalers, so that the plan is surely feasible (the feasible set is not empty). Boxes can not be further divided so variables must be integer.

With 2 stores and 3 wholesalers, there are 6 (2×3) possible lines of transport: AC, AD, AE, BC, BD, BE. we must decide how many boxes to move on each line. Let $x_{AC}, \ldots, x_{BE}$ the number of boxes to be moved.

C gets x_{AC} and x_{BC} and therefore $x_{AC} + x_{BC} \geq 50$. The same can be said for the other two wholesalers. For the stores, we must ask that the quantity that goes out of the store is not greater than the store initial quantity. For store A we have $x_{AC} + x_{AD} + x_{AE} \leq 100$, and similarly for B.

The IP problems is:

$$\begin{array}{lll} \min & 5x_{AC} + 10x_{AD} + 12x_{AE} + 3x_{BC} + 4x_{BD} + 9x_{BE} & \\ \text{s.t.} & x_{AC} + x_{AD} + x_{AE} \leq 100 & \text{(A)} \\ & x_{BC} + x_{BD} + x_{BE} \leq 250 & \text{(B)} \\ & x_{AC} + x_{BC} \geq 50 & \text{(C)} \\ & x_{AD} + x_{BD} \geq 120 & \text{(D)} \\ & x_{AE} + x_{BE} \geq 180 & \text{(E)} \\ & x_{AC}, x_{AD}, x_{AE}, x_{BC}, x_{BD}, x_{BE}, \text{ integers} & \end{array}$$

and its solution is $x_{AC} = 50$, $x_{AD} = 0$, $x_{AE} = 50$, $x_{BC} = 0$, $x_{BD} = 120$ e $x_{BE} = 130$.

8.5.2 Decision variables are the quantities of lemon and orange marmalades, base and special: let x_1 the production (in hg) of lemon marmalade, x_2 the production of lemon and lime marmalade, x_3 the production of orange marmalade and x_4 the production of cinnamon orange marmalade.

There is an additional decision variable, the total fruit to be bought, x_5, which is not given. For this quantity there is only an upper bound. To buy the maximum of fruit is not necessarily the best thing to do as there is another constraint (labor hours) and it can be that the optimal plan will not use all the available fruit. Unused fruit will therefore be a cost unbalanced by revenues. Variables are real numbers.

The objective is the profit. Revenues are given by $0.7x_1 + 0.6x_3 + 1.8x_2 + 1.4x_4$. Costs are $0.3x_5$ for buying fruit and special process $0.4x_2 + 0.4x_4$. The objective function is

$$\begin{array}{l} 0.7x_1 + 0.6x_3 + 1.8x_2 + 1.4x_4 - 0.3x_5 - (0.4x_2 + 0.4x_4) = \\ 0.7x_1 + 1.4x_2 + 0.6x_3 + x_4 - 0.3x_5 \end{array}$$

to be maximized.

Fruit constraint is $x_5 \leq 4000$.

Labor hours are 6000. Those used in the process are $x_5 + 3x_2 + 2x_4$, and therefore $x_5 + 3x_2 + 2x_4 \leq 6000$.

These constraints are not enough to set up the model. In fact, for each kg of fruit (i.e. of x_5) we get 3 hg of lemon marmalade (x_1) and starting from the latter we get lime and lemon marmalade (i.e. x_3). Therefore, the total production of lemon marmalade (base and special) must sum up to the total lemon available, $x_1 + x_3 = 3x_5$. Doing the same for oranges we get $x_2 + x_4 = 4x_5$.

The LP problem is:

$$\begin{array}{lll} \min & 0.7x_1 + 1.4x_2 + 0.6x_3 + x_4 - 0.3x_5 & \\ \text{s.t.} & x_5 \leq 4000 & \text{(max fruit)} \\ & x_5 + 3x_2 + 2x_4 \leq 6000 & \text{(max labor hours)} \\ & x_1 + x_3 - 3x_5 = 0 & \text{(tot lemon)} \\ & x_2 + x_4 - 4x_5 = 0 & \text{(tot oranges)} \\ & x_1, \ldots, x_5 \geq 0 & \end{array}$$

and its solution is $x_1 = 11333.333$, $x_2 = 666.667$, $x_3 = 16000$, $x_4 = 0$, $x_5 = 4000$. The profit is 17266.667.

Answer to quick questions

Probability

1.7.1 No, because the expected value of a random variable always lies between the minimum and the maximum of the values the random variable can take.

1.7.2 No, because B and $\overline{B}$ are two different events and the conditional probability distribution they generate are different.

1.7.3 Yes, because A and $\overline{A}$ are mutually exclusive events and $A \cup \overline{A} = \Omega$ and therefore their probabilities, given $\overline{B}$ must sum up to 1.

1.7.4 No. From the equation

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.7 - P(A \cap B)$$

we see that $P(A \cup B)$ is less or equal to 0.7.

1.7.5 Yes. If A implies B then $P(A) \leq P(B)$ and then $P(B)$ can be any number greater or equal to 0.4, e.g. 0.6.

Decision problems

3.6.1 3: decision, event and result node.

3.6.2 No, informational arcs always end in a decision node.

3.6.3 Yes, conditional arcs end in an event node or in a result node.

3.6.4 As A is connected to B by a conditional arc the decision maker assigns the probabilities in B after A has occurred. Thus, probabilities in B are conditional on possible outcomes in A.

3.6.5 The ID is drawn in Figure 12 and it is almost identical to that at the beginning of this chapter:

4.5.1 No, because the value of an event node is the expected value of the alternatives and as such is bound to be between the minimum and the maximum of them.

4.5.2 $p = 0.4$, because $10 \times 0.4 + 20 \times 0.6 = 16$.

4.5.3 Yes, the optimal choice is 2 because the expected value of the event node will always be less than the other option.

4.5.4 No, because the value of option 2 lies between the minimum and the maximum of the event node.

4.5.5 Yes, because the value of option 2 is always greater than or equal to the maximum of the value of the event node.

5.4.1 As A's certainty equivalent is less than the expected value of the lottery (which is 30), A is risk averse.

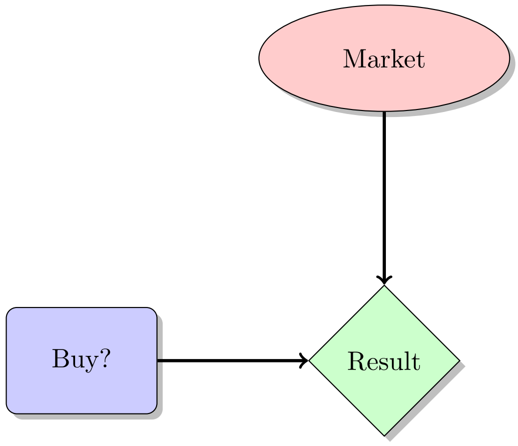


Figure 12: ID of the Buy/Not buy shares problem

5.4.2 The expected value of the lottery is $E = 0.01x$; as the person is risk averse, the certainty equivalent z should be less than E, $z < E$. Given that the person bought the ticket, its price should be less than the certainty equivalent z, i.e. $1 < z$. Therefore, $1 < E$, i.e. $1 < 0.01x$, or $x > 100$

5.4.3 The certainty equivalent is

$$z(X) = u^{-1}(p(u(x_1)) + (1-p)u(x_2)) = \left(0.3\sqrt{10} + 0.7\sqrt{20}\right)^2 = 16.6396$$

5.4.4 To sign an insurance contract means to exchange a "certainty equivalent" (the premium for risk) for an uncertain thing (being alive at a certain future time). Both customers have the same life expectancy so the only difference is in the way they see at the potential risk of dying, which is identical for the two people. The one who is more risk averse (the one with the lesser R) is more willing to pay for the insurance and therefore should be the one to call.

5.4.5 As the expected value of the change is $30000 \times p + 0 \times (1-p) = 30000p$, the current job would be preferred if $30000p < 10000$, i.e. $p < 1/3$.

Linear programming

6.5.1 The slack is the quantity of the resource that remains unused in the optimal solution.

6.5.2 Linear means that every mathematical object in a LP problem is a linear (i.e. first degree) equation or inequality in the decision variables.

6.5.3 A constraint is binding if in the optimal solution the corresponding resource is exhausted or, equivalently, that the equality relation is true.

6.5.4 Yes, because the order of the constraints does not matter (it does not change the feasible set)

6.5.5 The feasible set is the set whose elements are the solution of the constraints system. It is the set of values that solve the constraint system.

7.7.1 The optimal solution changes and the profit increases by 2.

7.7.2 No. With the additional unit, the profit will increase by 2 and with a cost of three the balance will be negative (loss of 1).

7.7.3 Yes. The optimal solution does not change but the value of the objective does change.

7.7.4 The reduced cost is the minimum amount by which a coefficient of the variable in the objective function must change to make the variable appear in the optimal solution, or to make it disappear if it already appears.

7.7.5 No. If there are resources left, the shadow price necessarily vanishes.

8.6.1 It is an integer variable that takes only two values, 1 and 0.

8.6.2 It is a problem in which variables can take only integer values.

8.6.3 Yes, because the integer constraint is a variable constraint and does not interfere with the coefficients of the objective function that can take non integer values as well.

8.6.4 The problem can admit just one solution, because even if the feasible set in unbounded, the minimum or the maximum of the objective function can be on the boundary of the set.

8.6.5 False. Rounding (and truncating) are not valid methods to get the solution of an IP problem.

Bibliography

[1] Emanuele Borgonovo. Influence diagrams. In *Encyclopedia of Medical Decision Making*. SAGE, 2008.

[2] Robert T. Clemen and Terence Reilly. *Making Hard Decisions*. Brooks Cole, 2005.

[3] George B. Dantzig and Mukund N. Thapa. *Linear Programming*. Springer, 1997.

[4] Geoffrey R. Grimmett and David R. Stirzaker. *Probability and Random Processes*. Oxford University Press, 2001.

[5] Mike Middleton. Treeplan. `http://www.treeplan.com`.

Index