Part I

Motivations and Foundations

1

Quantitative Methods: Should We Bother?

If you are reading this, chances are that you are on your way to becoming a manager. Or, maybe, you are striving to become a better one. It may also be the case that the very word *manager* sounds dreadful to you and conjures up images of unjustified bonuses; yet, you might be interested in how *good* management decisions should be made or supported, in both the private and public sectors. Whatever your personal plan and taste, what makes a good manager or a good management decision? The requirements for a career in management make a quite long list, including interpersonal communication skills, intuition, human resource management, accounting, finance, operations management, and whatnot. Maybe, if you look down the list of courses offered within master's programs in the sector, you will find quantitative methods (QMs). Often, students consider this a rather boring, definitely hard, maybe moderately useful subject. I am sure that a few of my past students would agree that the greatest pleasure they got from such a course was just passing the exam and forgetting about it. More enlightened students, or just less radical ones, would probably agree that there is something useful here, but you may just pay someone else to carry out the dirty job. Indeed, they do have a point, as there are plenty of commercially available software packages implementing both standard and quite sophisticated statistical procedures. You just load data gathered somewhere and push a couple of buttons, so why should one bother learning too much about the intricacies of QMs? Not surprisingly, a fair share of business schools have followed that school of

thought, as the role of QMs and management science in their curricula has been reduced,[1] if they have not been eliminated altogether.

Even more surprisingly however, there is another bright side of the coin. The number of software packages for data analysis and decision support is increasing, and they are more and more pervasive in diverse application fields such as supply chain management, marketing, and finance. Their role is so important that even books aimed at non specialists try to illustrate the relevance of quantitative methods and analytics to a wide public; the key concept of books like *Analytics at Work* and *The Numerati* is that these tools make an excellent *competitive weapon*.[2] Indeed, if someone pays good money for expensive software tools, there must be a reason. How can we explain such a blatant contradiction in opinions about QMs? The mathematics has been there for a while, but arguably the main breakthrough has been the massive availability of data thanks to Web-based information systems. Add to that the availability of cheap computing power and better software architectures, as well as smart user interfaces. These are relatively recent developments, and it will take time to overcome the inertia, but the road is clear.

Still, one of the objections above still holds: I can just pay a specialist or, maybe, learn a few pages of a software manual, without bothering with the insides of the underlying methods. However, relying on a tool without a reasonable knowledge of its traps and hidden assumptions can be quite dangerous. The role of quantitative strategies in many financial debacles has been the subject of heated debate. Actually, the unpleasing outcome of bad surgery executed by an incompetent person with distorted incentives can hardly be blamed on the scalpel, but it is true that quantitative analysis can give a false sense of security in an uncertain world. This is why anyone involved in management needs a decent knowledge of analytics. If you are a top manager, you will not be directly involved in the work of the specialists, but you should share a common language with them and you should be knowledgeable enough to appreciate the upsides and the downsides of their work. At a lower level, if you get an esoteric error message when running a software application, you should not be utterly helpless; by the same token, if there are alternative methods to solve the same problem, you should figure out what is the best one in your case. Last but not least, a few other students of mine accepted the intellectual challenge and discovered that studying QMs can be rewarding, interesting, and professionally relevant, after all.[3]

[1] The actual term that is often used is "dumbed down," but this could sound not too politically correct to some.

[2] See Refs. [1, 7], as well as T.H. Davenport, Competing on Analytics, *Harvard Business Review*, Jan. 2006, pp. 1–9.

[3] I am quite proud to say that one of my best QMs master's students had a degree in classics, from Oxford University. I am also proud to say that a few of them changed their mind about statistics: "You know, prof, last year I got a pretty good grade in statistics, but I couldn't figure out why..."

I will spend quite a few pages trying to convince you that a good working knowledge of QMs is a useful asset for your career.

- When information is available, decisions should be based on data. True, a good manager should also rely on intuition, gut feelings, and the ability to relate to people. However, there are notable examples of managers who were considered geniuses after a lucky decision, and eventually destroyed their reputation, endangered their business, and went to jail in some remarkable cases. Without going to such extremes, even the best manager may make a wrong decision, because something absolutely unpredictable can happen. A good decision should be somewhat robust, but when things go really awry, being able to justify your move on a formal analysis of data may save your neck.

- QMs can make you a sort of universal blood donor. The mathematics behind is general enough to be applied in different settings, such as supply chain management, finance, and marketing. QMs can open many doors for you. Indeed, throughout the book I will insist on this point by alternating examples from quite different areas.

- Even if you are not a specialist, you should be able to work with consultants who have specialized quantitatively. You should be able to interact constructively with them, which means neither refusing good ideas merely because they seem complicated, nor taking for granted that sophistication always works. At the very least, you should be aware of what they are doing.

I have met some people whose idea of applying QMs is collecting data and coming up with a few summary measures, maybe some fancy plots to spice up a presentation, and that's it. In fact, QMs are much more than collecting basic descriptive statistics:

1. If QMs are to be of any utility to a manager, they should help her in *making decisions*. Unfortunately, modeling to make decisions is a rather hard topic.

2. By the same token, basic probability and statistics are not enough to meet the challenge of a complex reality. Multivariate analysis tools have been applied, but there is a gap between books covering the standard procedures and those at an advanced level.

We will try to bridge that gap, which is somewhat hard to do by just walking through a lengthy and dry list of theorems and proofs. In this chapter I will illustrate a few toy examples, that will hopefully provide you with enough motivation to proceed.

We have emphasized the role of data to make decisions. If we knew all of the relevant data in advance, then our task would be considerably simplified. Nevertheless, we show in Section 1.1 that even in such an ideal situation some

quantitative analysis may be needed. More often than not, uncertainty makes our life harder (or more interesting). In Section 1.2 we deal with different examples in which we have to make a decision under uncertainty. The standard tools that help us in such an endeavor are provided by probability and statistics, which constitute a substantial part of the book. Nevertheless, we will show that some concepts, such as probability, can be somewhat dependent on the context. Indeed, many features of real life may make a straightforward application of simple methods difficult, and we will see a few examples in Section 1.3. Finally, in Section 1.4 we will discuss how, when, and why QMs can be useful, while pointing out their limitations.

1.1 A DECISION PROBLEM WITHOUT UNCERTAINTY: PRODUCT MIX

Product mix decisions are essentially resource allocation problems. We have limited resources, such as machines, labor, and raw materials, and the problem calls for their optimal use in order to maximize profit, which is earned by producing and selling a set of items. The decision problem consists of finding the right amounts to produce for each item over a certain timespan. Profit depends on the cost of producing each item and the price at which they can be sold. Produced quantities should comply with several constraints, such as production capacity and market limitations, since we should not produce what we are not going to sell anyway.

One of the fundamental pieces of information we need is demand. The time period we work with can be a day, a week, or a month. In practice, demand varies over time and can be quite uncertain. Here we consider an idealized problem in which demand is known and constant over time. Furthermore, demand is not completely exogenous in real life, as we might influence it by pricing decisions. Price can be more or less under direct control, depending on the level of competition and the type of market we deal with; in a product mix problem we typically assume that we are price takers.

In the first example below, products are similar in the sense that they consume similar amounts of resources. In the second one, we will complicate resource consumption a bit.

1.1.1 The case of similar products

A firm[4] produces red and blue pens, whose unit production cost is 15 cents (including labor and raw material). The firm incurs a daily fixed cost, amounting to €1000, to run the plant, which can produce at most 8000 pens per day in total (i.e., including both types). Note that we are expressing the capacity

[4]This example is based on Chapter 2 of Ref. [5].

constraint in terms of the total number of pens produced, which makes sense if resource requirements are the same for both products; in the case of radically different products (say, needles and air carriers), this makes no sense, as we shall see in the next section. We are not considering changeover times to switch production between the two different items, so the above information is all we need to know from the technological perspective.

From the market perspective, we need some information about what the firm might sell and at which price. The blue pens sell for 25 cents, whereas things are a tad more complicated for the red ones. On a daily basis, the first 5000 red pens can be sold for 30 cents each, but additional ones can be sold for only 20 cents. This may sound quite odd at first, but it makes sense if we think that the same product can be sold in different markets, where competition may be different, as well as general economic conditions. Such a price discrimination can be maintained if markets are separated, i.e., if one cannot buy on the cheaper market and resell on the higher-priced market.[5] In general, there may be a complex relationship between price and demand, and in later chapters we will consider QMs to estimate and take advantage of this relationship.

The problem consists of finding how many red and how many blue pens we should produce each day. Note that we are assuming constant demand; hence, the product mix is just repeated each day. In the case of time-varying demand and changeover costs, there could be an incentive to build some inventory, which would make the problem dynamic rather than static.

1. The production manager, an ugly guy with little business background, decides to produce 5000 red and 3000 blue pens, yielding a daily profit of €50 (please, check this result). This may not sound too exciting, but at least we are in the black.

2. A brilliant consultant (who has just completed a renowned master, including accounting classes) argues that this plan does not consider how the fixed cost should be allocated between the two product types. Given the produced quantities, he maintains that €625 ($\frac{5}{8}$ of the fixed cost) should be allocated to red pens, and €375 to blue pens. Subtracting this fraction of the fixed cost from the profit contribution by blue pens, he shows that blue pens are not profitable at all, as their production implies a loss of €75 per day! Hence, the consultant concludes that the firm should just produce red pens.

What do you think about the consultant's idea? Please, *do* try finding an answer before reading further!

[5]International students may be familiar with book editions that are marked as "Not for sale in the USA."

A straightforward calculation shows that the second solution, however reasonable it might sound, implies a daily loss:

$$1000 - 5000 \times (0.30 - 0.15) - 3000 \times (0.20 - 0.15) = €100$$

It is also fairly easy to see that the simple recipe of the production manager is just based on the idea of giving priority to the item that earns the largest profit margin. Apart from that, we should realize that the fixed cost is not really affected by the decisions we are considering at this level. If the factory is kept open, the fixed cost must be paid, whatever product mix is selected. However, this does not mean that the fixed cost is irrelevant altogether. At a more strategic decision echelon, the firm could consider shutting the plant down because it is not profitable. The point is that any cost is variable, at some hierarchical level and with a suitably long time horizon.

From a formal point of view, what we have been trying to solve is a problem such as

$$\begin{aligned} \max \quad & \pi(x_r, x_b) \\ \text{s.t.} \quad & x_r + x_b \leq 8000 \\ & x_r, x_b \geq 0 \end{aligned}$$

In this mathematical statement of the problem we distinguish the following:

- Two *decision variables*, x_r and x_b, which are the amounts of red and blue pens that we produce, respectively.
- An *objective function*, $\pi(x_r, x_b)$, representing the profit we earn, depending on the selected mix, i.e., on the value assigned to the two decision variables. Our task is maximizing profit with respect to decision variables.
- A set of *constraints* on the decision variables. We should maximize profit with respect to the decision variables, *subject to* (s.t. in the model formulation) this set of constraints. The first constraint here is an inequality corresponding to the capacity limitation. Further, we have included nonnegativity requirements on sold amounts. Granted, unless you are pretty bad with marketing, you are not going to sell negative amounts, which would reduce profit. Yet, from a mathematical perspective, manufacturing negative amounts of an item could be an ingenious way to create capacity for another item, which makes little sense and must be forbidden. Constraints pinpoint a *feasible region*, i.e., a set of solutions that are acceptable, among which we should find the best one, according to our criterion.

  The feasible region in our case is just the shaded triangle depicted in Fig. 1.1. If you have trouble understanding how to get that figure, you might wish to refer to Section 2.3; yet, we may recall from high school

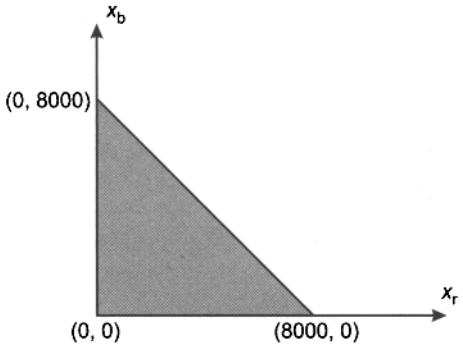

Fig. 1.1 The feasible set for the problem of red and blue pens.

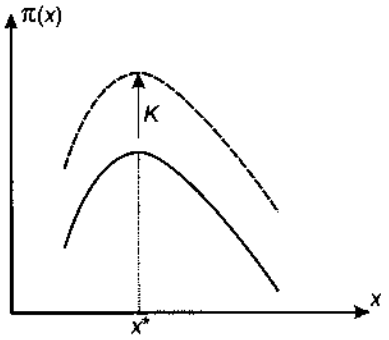

Fig. 1.2 Shifting a function up and down does not change the optimal solution.

mathematics that an equation like $ax_1 + bx_2 = c$ is the equation of a line in the plane; an inequality like $ax_1 + bx_2 \leq c$ represents one of the two half-planes separated by that line. To see which one, the easy way is checking if the origin of the plane, i.e., the point of coordinates (0, 0) satisfies the inequality, in which case it belongs to the half-plane, or not.

Intuitively, since the firm makes money by selling whatever pen it produces, the capacity constraint should be binding at the optimal solution, which means that we should look for solutions on the line segment joining points of coordinates (0, 8000) and (8000, 0). In Chapter 2 we will see how one can maximize a profit function (or minimize a cost function) in simple cases; a more thorough treatment will be given in Chapter 12. For now, we may immediately see why the fixed cost should be ignored in finding the optimal mix. Assume, for the sake of simplicity, that we have just one decision variable and consider the objective function $\pi(x)$ in Fig. 1.2. Let us denote the optimal solution of the maximization problem, $\max \pi(x)$, by x^*. We see that if the function is shifted up (or down) by a given amount K, i.e., if we solve $\max \pi(x) + K$, the optimal solution does not change. Yet, the optimal value does, and this

may make the difference between a profitable business and an unprofitable one. Whether this matters or not depends on the specific problem we are addressing.

Takeaways Even from a simple problem like this, there are some relevant lessons that deserve being pointed out:

- A simple decision problem consists of decision variables, constraints on them, and some performance measure that we want to optimize, such as minimizing cost or maximizing profit.
- Not all costs are always relevant; this may depend on the level at which we are framing the problem.
- The relationship between price and demand can be complex. In real life, data analysis can be used to quantify their link, as well as the uncertainty involved.

1.1.2 The case of heterogeneous products

We solved the previous example by a simple rule: Let us pick the most profitable item and try producing as much as we can; if we hit a market limitation, consider the next most profitable item, and go on until we run out of resource availability. However, there must be something more to it. To begin with, we had just one resource; what if there are many? Well, maybe one of them will prove to be the bottleneck and will limit overall production. But there is another issue, as we expressed the capacity constraint as the number of overall items that we could produce each day. What if each item consumes a different amount of each resource? In order to see that things may be a tad more complicated, let us consider another toy example.[6]

We are given

- Two item types (P1 and P2) that we are supposed to produce and sell
- Four resource types (machine groups A, B, C, and D) that we use to produce our end items

Note that *all* of the above resources are needed to produce an item of either type; they are not alternatives, and each part type must visit all of the machine groups in some sequence. The information we have to gather from a production engineer is the time that each piece must spend being processed on each machining center. This information is given in Table 1.1, where columns labeled $T_A, \ldots, T_D$ are the processing times (say, minutes) for each part type on each machine type. At this level, we are not really interested in the exact

[6]This example is based on Chapter 16 of Ref. [10].

Table 1.1 Data for the optimal mix problem.

Item	T_A	T_B	T_C	T_D	Cost	Price	Demand
P1	15	15	15	25	45	90	100
P2	10	35	5	15	40	100	50

sequence of machine visits; probably, some technological reason will force a sequence of operations, but we want to determine how many pieces we produce during each period. To make this point clearer, let us say that we want to find a weekly production mix. Someone else will have the task to specify what has to be processed on each machine, on each hour of each day during the week. In most problem settings there is a decision hierarchy, whereby we first specify an aggregate plan, that becomes progressively more detailed while going down the hierarchy.

From Table 1.1 we immediately see that end items differ in their resource requirements. Hence, it makes no sense to express a capacity constraint in terms of the total number of items that we can produce each week. What we need to know is how many minutes of resource availability we have each week. This depends on the work schedule, labor and machines available, etc. Each machine group may consist of many similar or identical machines; hence, we are interested in the aggregate capacity, rather than the time that each single physical machine is available. To consider a simple case, let us assume that machine availability is the same for all of the four groups: 2400 minutes. Note that this is the availability, or capacity, for *each* machine group.

Another limitation on production stems from market size. If demand is limited, there is no point in making something we can't sell (remember that, according to our assumptions, both capacity and demand are constant over time, so there is no point in building and carrying any inventory). Furthermore, we should consider the cost of producing an item and the price at which we may sell it. These market and economical data are given in the last three columns of Table 1.1. The cost given in the third column from the right refers to each single item and it may also include raw material, labor, etc. Further to that, let us say that we also incur a fixed cost of €5000 per week. We have already pointed out that this will not influence the optimal mix, but it makes the difference between being in the black or in the red. In the two last columns we see the price at which we sell each unit, which we assumed constant and independent from the number of items produced, and the weekly demand for each part type, which places an upper bound on sales.

Our task is to find the optimal production mix, i.e., a production plan maximizing profit. The task is not that difficult, as we just need two numbers. Let us denote by x_1 and x_2 the amounts of item P1 and P2 that we produce,

respectively. Yet, we must be careful to meet all of the capacity and market size constraints.

A trial-and-error approach One thing we may try is to apply the same principle of the red and blue pens: P2 looks more profitable, since its profit margin is $100 - 40 = €60$, which is larger than the $90 - 45 = €45$ of item P1. So, let us try to maximize production of item P2. From the technological data, we see immediately that the bottleneck machine group, on which P2 spends the most time, is machining center B. An upper bound on x_2 is obtained by assuming that we use all of the capacity of group B to manufacture P2:

$$35x_2 \leq 2400 \Rightarrow x_2 \leq 68.57$$

One could object that the true bound is 68, as we cannot manufacture fractional amounts of an item. Anyway, we cannot sell more than 50 pieces, so we set $x_2 = 50$, and then we maximize production of P1 using residual capacity. We should figure out which of the four capacity constraints will turn out to be binding. We can write the following set of inequalities, one per machine group, and check which one actually limits production:

$$\begin{aligned} 15x_1 + 10 \cdot 50 \leq 2400 &\Rightarrow x_1 \leq 126.67 \\ 15x_1 + 35 \cdot 50 \leq 2400 &\Rightarrow x_1 \leq 43.33 \\ 15x_1 + 5 \cdot 50 \leq 2400 &\Rightarrow x_1 \leq 143.33 \\ 25x_1 + 15 \cdot 50 \leq 2400 &\Rightarrow x_1 \leq 66 \end{aligned}$$

which yields $x_1 = 43.33$. For the sake of simplicity, let us assume that we are indeed able to make fractional amounts of items. This is somewhat true when we deal with things such as paint, and it is a sensible approximation for large numbers; rounding 1,000,000.489 up or down induces a small error. We will see in Chapter 12 why forcing integrality of decision variables may complicate things, and we should do it only when really needed. The production plan $x_1 = 43.33$, $x_2 = 50$ is feasible; unfortunately, total profit is negative:

$$45 \times 43.33 + 60 \times 50 - 5000 = -50$$

What went wrong? Maybe this is the best we can do, and we should just shut the business down, or try reducing cost, or try increasing price without reducing demand too much. Or maybe we missed something. With red and blue pens, resource consumption was the same for both items, but in our case P2 features the larger resource consumption on machine B. Maybe we should somehow consider a tradeoff between profit and resource consumption; maybe we should come up with a ratio between profit contribution and resource consumption. It is not quite clear how we should do this, since it is not true that P2 requires more time than P1 on *all* of the four resources. Nevertheless, it could well be the case that, carrying out this analysis, P1 would turn out to be more profitable. So, let us see what we get if we maximize production

of P1 first. In this case, machine group D is the bottleneck, and the same reasoning as above yields

$$25x_1 \leq 2400 \Rightarrow x_1 \leq 96$$

Now we do not reach the market bound, which is 100 for P1, but then, since we use all of the capacity of group D for item P1, we must set $x_2 = 0$. Fair enough, but profit is even worse than before: $45 \cdot 96 - 5000 = -680$.

Hopefully, the reader is starting to see that even for a toy problem such as this one, the art of quick calculations based on plausible and intuitive reasoning may fall short of our expectations. But before giving up, let us try to see if there is a way to make the problem simpler. After all, the difficulty comes mainly form capacity constraints and differentiated resource consumption. If we look a bit more carefully at Table 1.1, we see something interesting. Consider resources A and B: Are they equally important? Note that a plan that is feasible for group B must be feasible for A as well: P1 requires the same amount of time on both groups, whereas P2 has a larger requirement on B. We may conclude that group A will never be a binding resource. If we compare resource requirements for groups B and C, we immediately reach a similar conclusion. In fact, only resources B and D need to be considered.[7]

Now the perspective looks definitely better: We just need to find a solution which uses all of the resources B and D, as this will maximize production. Unless we hit a market constraint, there is no point in leaving critical resources unused. We should find two values for our two decision variables, x_1 and x_2, such that both machine groups B and D are fully utilized. This results in a system of two equations:

$$\begin{cases} 15x_1 + 35x_2 = 2400 \\ 25x_1 + 15x_2 = 2400 \end{cases} \tag{1.1}$$

We will see a bit more about solving such a system of linear equations in Chapter 3. For now, let us just say that solving this system yields the production mix $x_1 = 73.84$ and $x_2 = 36.92$, rounding numbers down to the second decimal digit; this results in a total profit of €538.46, which is positive! Intuition worked pretty well for the red and blue pens problem, but this solution is a bit harder to get by sheer intuition.

If this seems too hard, please have a reality check. We had to solve just a toy problem, ignoring all of the complications that make real life so fun:

- We had to deal with just two end items (they may easily be thousands).
- Demand was known with certainty (you wish).

[7] A whole managerial philosophy, called the *theory of constraints*, has been born, based on the principle that you may simplify a problem and better focus your effort by concentrating on bottlenecks, i.e., the factors that really limit performance.

- All of the relevant data were constant over time (same as above).
- We did not consider interactions between demands for different end items (if a customer wants both items P1 and P2, and we have not enough of one of them, we might well lose the whole order).
- We did not consider availability of raw materials (one of the most amusing moments you might experience in life is when you cannot finish the assembly of a \$100,000 item because you miss a little screw worth a few cents).
- We did not consider changeover times between different item types (on very old press lines in the automotive industry, setting up production for another model required 11 hours).
- We did not consider detailed execution and timing.
- We did not consider substitution between raw materials; in some blending processes (food and oil), there are some degrees of freedom making the choice even more complicated (we cover blending problems in Section 12.2.3).
- We did not include integrality constraints on the decision variables, which would probably make our approach unsuitable (we will see how to cope with this complication in Section 12.6.2).

If we realize the true complexity of a real-life problem, it is no surprise that sometimes even getting a *feasible* solution (let alone an optimal one) may be difficult without some quantitative support. Hence, we need a more systematic approach.

A model-based approach In the case of red and blue pens, we hinted at the possibility of building a mathematical representation of a decision problem. Maybe, this can be helpful in a complex setting. To begin with, we want to maximize profit. Formally, this means that we want to maximize a function such as

$$45x_1 + 60x_2$$

We have already remarked that fixed costs do not change where the optimal solution is, so subtracting €5,000 is inconsequential. From the work we have carried out before, we see that capacity constraints can be represented as a set of inequalities:

$$\begin{aligned} 15x_1 + 10x_2 &\leq 2400 \\ 15x_1 + 35x_2 &\leq 2400 \\ 15x_1 + 5x_2 &\leq 2400 \\ 25x_1 + 15x_2 &\leq 2400 \end{aligned}$$

If we also include nonnegativity of decision variables and market bounds, we end up with the following mathematical problem:

$$\begin{aligned} \max \quad & 45x_1 + 60x_2 \\ \text{s.t.} \quad & 15x_1 + 10x_2 \leq 2400 \\ & 15x_1 + 35x_2 \leq 2400 \\ & 15x_1 + 5x_2 \leq 2400 \\ & 25x_1 + 15x_2 \leq 2400 \\ & 0 \leq x_1 \leq 100 \\ & 0 \leq x_2 \leq 50 \end{aligned} \tag{1.2}$$

This is an example of a linear programming problem, where *linear* is due to the fact that decision variables occur linearly: you do not see products such as $x_1 \cdot x_2$, powers such as x_1^2, or other weird functions such as $\sin x_2$. Real-life problems may involve thousands of decision variables, but they can be solved by many computer packages implementing a solution strategy called the *simplex method*, and (guess what?) using this magic you get the optimal solution above. By the way, good software will also spot and get rid of irrelevant constraints to speed up the solution process.

More on this in Chapter 12, but in this simple case we may visualize things graphically in order to better understand why the first simple-minded approach failed.

A graphical solution As with red and blue pens, we are dealing here with a bidimensional problem. Each (linear) inequality corresponds to a half-plane. Since we must satisfy a set of such constraints, the set of feasible solutions is the intersection of half-planes, and is illustrated in Fig. 1.3. The shaded figure is a polyhedron, resulting from the intersection of the relevant constraints: these are the capacity constraints for groups B and D, and the market bound for item P2.

The parallel lines shown in the figure are the level curves of the profit function. For instance, to visualize all of the product mixes yielding a profit contribution of €2000 (neglecting the fixed cost), we should draw the line corresponding to the linear equation

$$45x_1 + 60x_2 = 2000$$

Changing the desired value of profit contribution, we draw a set of parallel lines; three of them are displayed in Fig. 1.3. It is also easy to see that profit increases by moving in the northeast direction, i.e., by increasing production of both part types.

There is an infinite set of feasible mixes (barring integrality requirements on decision variables), but we see that a only a very few of them are relevant: those corresponding to the vertices (or extreme points) of the polyhedron, i.e., points $M_0, M_1, \ldots, M_4$. Point M_0, the origin of the axes, corresponds to

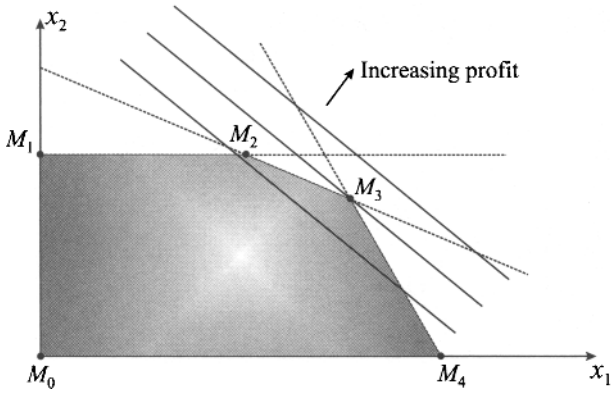

Fig. 1.3 Graphical solution of the optimal mix problem.

making nothing and is not quite interesting. Point M_1 corresponds to making 50 items of P2, and none of P1; in fact, during our first attempt, we moved to that point first, and then to point M_2, with coordinates (43.33, 50), which was our first mix. Point M_4, with coordinates (96, 0), represents the second tentative mix we came up with. We see that the second solution was worse than the first one by checking the level curves of profit. Since level curves are parallel lines, and we should move along the direction of increasing profit, we see that the optimal solution must be a feasible point that "touches" the level curve with the highest profit. This happens at point M_3, which in fact corresponds to the optimal mix.

The slope of the level curves depends on the profit margin of each item. For instance, if we increase the profit margin of P1, the lines rotate clockwise; if profit margin of P1 is increased enough, the optimal mix turns out to be point M_4. In general, changing the economics of the problem will result in different optimal mixes, as expected, but they will always be extreme points of the feasible set, and there are not so many of them. Whatever profit margins are, only points M_2, M_3, and M_4 can be candidate optimal solutions. If level curves happen to be parallel to an edge of the feasible set, we have an infinite number of optimal solutions, but we may just consider one corresponding to a vertex. In fact, the standard approach to solving a linear programming model, via the simplex method, exploits this property to find an optimal solution with stunning efficiency even for large-scale problems involving thousands of variables and constraints.

Incidentally, if we insist on producing integer amounts, we should only consider points with integer coordinates within the polyhedron. We may draw this feasible set as a grid of discrete points. Doing so, the optimal mix turns out to be $x_1 = 73$, $x_2 = 37$, with total profit 505. Understandably, profit is reduced by adding a further constraint on production volume. It is tempting to conclude that we may easily get this solution by solving the previous prob-

lem and then rounding the solution to the closest integer point on the grid. Unfortunately, this is not always the case, and quite sophisticated methods are needed to solve problems with integer decision variables efficiently.

Takeaways

- Intuition may fail when tackling problems with many constrained decision variables.
- Mathematics may yield an optimal solution for the *model*. Because modeling calls for simplification, this need not be the best solution of our *problem*, but it may be a good starting point.
- Sophisticated software packages are available to tackle mathematical model formulations. Hence, we need to concentrate on modeling rather than on complicated solution procedures. Indeed, in Chapter 12 we focus on models for decision making, while just giving a glimpse of the computational solution procedures. This is extended in Chapter 13 to cope with uncertainty.
- Nevertheless, a suitable background in calculus and algebra is needed to gain a proper understanding of the involved approaches; this is the subject of Chapters 2 and 3.

1.2 THE ROLE OF UNCERTAINTY

We often have to make decisions here and now, without complete knowledge about problem data or the occurrence of future events. In distribution logistics, significantly uncertain demand must be faced; in finance, several sources of risks affect the return of an investment portfolio. In all of these settings, the future effect of actions is not known for sure. Uncertainty can take several forms. In the simplest case, we may be able to gather past information and use that to generate a set of plausible future scenarios. This is where the standard tools of probability and statistics come into play. They will be the subject of Part II of the book, and are typically considered the core of any course on QMs. To get gradually acquainted with them, let us consider a few toy problems.

1.2.1 A problem in supply chain management

In the product mix problem, we assumed perfect knowledge of future demand, but, unfortunately, exact demand forecasts are a bit of scarce commodity in the true world. Indeed, the standard trouble in supply chain management is purchasing an item for which demand information is quite uncertain. If we order too much, one or more of the following scenarios might occur:

- Finance will suffer, as money is tied up in inventories.
- Items may become obsolete because of fads or product innovation, and money will be lost in inventory writeoffs.
- Perishable items may run out of their shelf life before being sold, and money will be lost again.

On the other hand, if we do not order enough items, we may not be able to meet customer demand and revenue will suffer (as well as our career; life is hard, isn't it?).

To take our first baby steps, let us consider a relatively simple version of the problem. We are in charge of purchasing an item with a very limited shelf life. Both purchased quantities and demand are given as small integer numbers, which makes sense for a niche product. Items are purchased for delivery at the beginning of each week, and any unsold item is scrapped at the end of the same week; hence, each time we face a brand-new problem, in the sense that nothing is left in inventory from the previous time periods. Demand for the next week is not known, but we do have some information about past demand. The following list shows demand for the past 20 weeks:

$$\begin{array}{cccccccccc} 3, & 1, & 3, & 2, & 3, & 2, & 2, & 4, & 5, & 2 \\ 1, & 2, & 2, & 3, & 5, & 2, & 3, & 2, & 4, & 1 \end{array} \tag{1.3}$$

The big question is: How many items should we order right now?

When asked this question, most students suggest considering the *average* demand, which is easily calculated as

$$\overline{D} = \frac{3+1+3+2+\cdots+4+1}{20} = \frac{52}{20} = 2.6$$

Not too difficult, even though this result may leave us a bit uncertain, as we cannot really order fractional amounts of items. Yet, it seems that a reasonable choice could be between 2 and 3.

Other students suggest that we should stock the *most likely* value of demand. To see what this means exactly, it would be nice to see some more structure in the demand history, maybe by counting the frequency at which each value has occurred in the past. If we sort demand data, we get the following picture:

$$\underbrace{1,\ 1,\ 1,}_{3\text{ times}}\ \underbrace{2,\ 2,\ 2,\ 2,\ 2,\ 2,\ 2,\ 2,}_{8\text{ times}}\ \underbrace{3,\ 3,\ 3,\ 3,\ 3,}_{5\text{ times}}\ \underbrace{4,\ 4,}_{2\text{ times}}\ \underbrace{5,\ 5}_{2\text{ times}}$$

These numbers provide us with the frequencies at which each value occurred in the observed timespan. If we divide each frequency by the number of observations, we get *relative frequencies*. For instance, the relative frequency

Table 1.2 Frequencies (F), relative frequencies (F_{rel}), and cumulated (relative) frequencies (F_{cum}) for demand data.

Value	F	F_{rel}	F_{cum}
1	3	0.15	0.15
2	8	0.40	0.55
3	5	0.25	0.80
4	2	0.10	0.90
5	2	0.10	1.00

of the value 2 is $\frac{8}{20} = 0.4$ or, in percentage terms, 40%. We may also calculate average demand by using relative frequencies:

$$\begin{aligned}
\overline{D} &= \frac{3 \times 1 + 8 \times 2 + 5 \times 3 + 2 \times 4 + 2 \times 5}{20} \\
&= \frac{3}{20} \times 1 + \frac{8}{20} \times 2 + \frac{5}{20} \times 3 + \frac{2}{20} \times 4 + \frac{2}{20} \times 5 \\
&= 0.15 \times 1 + 0.40 \times 2 + 0.25 \times 3 + 0.10 \times 4 + 0.10 \times 5 \\
&= 2.6
\end{aligned}$$

Not surprisingly, we get the same average as above. We see that average demand is a weighted average of observed values, where weights correspond to relative frequencies. If we believe that the future will reflect the past, relative frequencies provide us with useful information about the likelihood of each demand value in the future.

Frequencies and relative frequencies are tabulated in columns 2 and 3 of Table 1.2. Be sure to note that relative frequencies cannot be negative and add up to 1, or 100%. Frequencies and relative frequencies may also be visualized using a histogram, as shown in Fig. 1.4. The observed values are reported on the horizontal axis (abscissa); the vertical axis (ordinate) may represent frequencies (a) of relative frequencies (b). The two plots are qualitatively the same, as relative frequencies are just obtained by normalizing frequencies with respect to the number of observations. After a quick glance at the graphical representation of relative frequencies, the intuitive idea of a "likelihood measure" of each demand value comes to mind rather naturally. Indeed, it is possible to interpret relative frequencies as *probabilities*. However, some caution should be exercised and we will see in Chapters 5 and 14 that probability is not such a trivial concept, as there are alternative interpretations. Still, this intuitive interpretation may be useful in many practical cases.

Looking at Table 1.2, we see that the most likely value (or the most frequent value in the past, to be honest with ourselves) is 2, which is not too different from the average value. In descriptive statistics, the most likely value

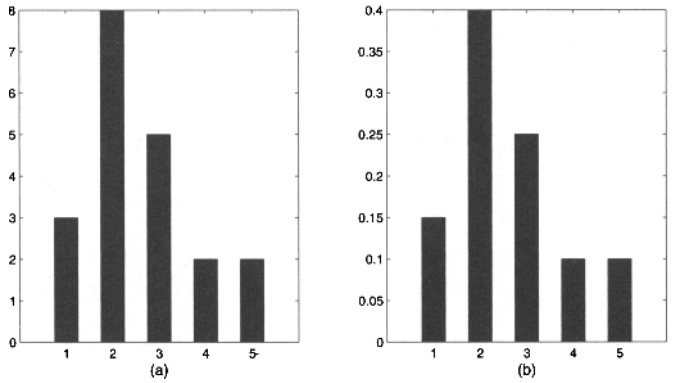

Fig. 1.4 Histograms visualizing frequencies and relative frequencies for demand data.

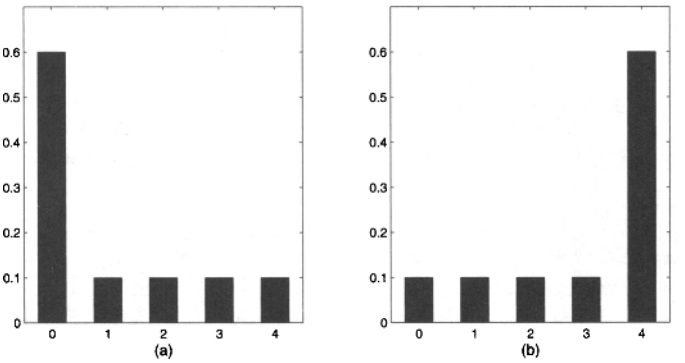

Fig. 1.5 Two skewed distributions.

is called *mode*. Since we get similar solutions by considering either mean or mode, we could be tricked into believing that we will always make a sensible choice by relying on them. Before we get so overconfident, let us consider the histograms of relative frequencies in Fig. 1.5. In histogram (a), we see that the most likely value is zero, but would we really stock nothing? Probably not. The two histograms in Fig. 1.5 are two examples of asymmetric cases. They are "skewed" into opposite directions, and we probably need a way to characterize skewness. We will deal with this and other summary measures in Chapter 4, but it is already clear that mean and mode do not tell the whole story and they are not always sufficient to come up with a solution for a decision problem. Lack of symmetry is likely to affect our stocking decisions, but there is still another essential point that we are missing: dispersion. Consider the two histograms in Fig. 1.6. Histogram (a) looks more concentrated, which arguably suggests less uncertainty about future demand with respect to histogram (b). We need some ways to measure dispersion as well, and to

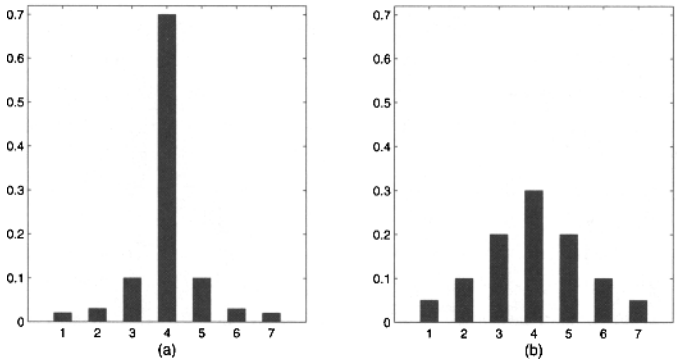

Fig. 1.6 The role of dispersion.

figure out how it can affect our choice. Indeed, we need some ways to characterize uncertainty, and this motivates the study of descriptive statistics (to be carried out in Chapter 4). This is fine, but it is utterly useless, unless we find a way to use that information to come up with a decision. It is important to realize how many points we are missing, if we just consider relative frequencies.

The role of economics. If we have a stockout, i.e., we run out of stock and do not meet the whole customer demand, how much money do we lose? And what if we have an overage, i.e., we stock too much and have to scrap perished or obsolete items? To see the point, consider the following problem. We have to decide how many T-shirts to make (or buy) for an upcoming major sport event. Producing and distributing a T-shirt costs €5; each T-shirt sells for €20, but unsold items at the end of the event must be sold at a markdown price, resulting in a loss.[8] Let us assume that the discount on sales after the event is 80%, so that the markdown price is €4. A credible forecast, based on similar events, suggests that the expected value of sales is 12,000 pieces. We will clarify what we mean by *expected value* exactly, but you may think of it as the "best forecast" given our knowledge. However, demand is quite uncertain. A consultant, considering demand uncertainty and the risk of unsold items, suggests to keep on the safe side and produce just 10,000 pieces. Is this a good idea?

Please! Wait and think about the question before going on.

[8]The example may not sound too significant, but this kind of situation is typical of fashion items and, given the pace of technological innovation, presents some features shared by huge markets such as consumer electronics.

When we sell a T-shirt, our profit is €15; if we have to mark down, we lose only €1. Given that, most people would probably suggest a more aggressive strategy and buy a bit more than the expected value. Indeed, most fashion stores mark prices down at some time, which means that they tend to overstock. Would you change your idea if profit margin were €2 and the cost of an unsold item were €5? Economics must play a role here, as well as dispersion. Without any information about uncertainty, we cannot specify how much above or below the expected value of demand we should place our order. A plain point forecast, i.e., a single number, is not enough for robust decision making, a point that we will stress again in Chapters 10 and 11, when dealing with regression and time series models for forecasting.

Predictable vs. unpredictable variability. Consider once again the demand data in (1.3), but this time imagine that the time series, in chronological order, is

$$1,\ 1,\ 1,\ 2,\ 2,\ 2,\ 2,\ 2,\ 2,\ 2,\ 2,\ 3,\ 3,\ 3,\ 3,\ 3,\ 4,\ 4,\ 5,\ 5$$

Mean, mode, etc., are not affected by this reshuffling of data, but should we neglect the clear pattern that we see? There is a trend in demand, which is not captured by simple summary measures. And what should we do with a demand pattern such as the following one?

$$1,\ 2,\ 3,\ 4,\ 3,\ 2,\ 1,\ 2,\ 3,\ 4,\ 3,\ 2,\ 1,\ 2,\ldots$$

In this case, we notice a seasonal pattern, with regular up- and downswings in demand. Trend and seasonality contribute to demand variability, but we should set predictable and unpredictable components of variability apart. In chapter 11 we describe some simple methods for doing so.

The role of time and intertemporal dependence. The previous point shows that time does play a role, when we can identify partially predictable patterns such as trend and seasonality. Time may also play a role when our assumptions about ordering and shelf life are less restrictive. Assume that the shelf life is longer than the time between the orders we issue to suppliers. In making our decision, we should also consider the inventory level, and this would make the problem dynamic rather than static. A safe guess is that this is no simplification.

An even subtler point must be considered in order to properly represent unpredictable variability. I will illustrate it with a real-life story. A few years ago in Turin, where I live, there was a period of intense rain followed by an impressive flood. A weird thing with such an event is that there is way too much water in the streets, but you do not get any from your water tap at home. In that case, the high level of the main river in

the city prevented the pumping stations from working. This problem, as I recall, was solved quickly, but the immediate consequence was a race to buy any bottle of mineral water around (with plenty of amicable exchange of ideas between tactful customers at retail stores). Now, if you were the demand manager for a company selling mineral water, would you interpret that spike in demand as a signal of an increasing market share? Well, not really, I guess. On the contrary, you could expect a period of low demand, when households deplete their unusually high inventories. More generally, if consumption of an item is relatively steady over time, a spike in demand (maybe due to a trade promotion) is likely to be followed by a period of low demand.[9] In order to take such issues into account, we need statistical tools to investigate correlation, which is the subject of Chapter 8. Correlation is useful in many settings where we ask questions about random variables rather than random events, such as

- If demand for an item has been larger than usual today, can we say something about demand tomorrow? Will it be larger or smaller than usual?
- If the return from a financial investment has been good, does this tell us something about the return of the same investment in the future, or maybe about the return from other investments?

The role of alternative items and competitors. We have just considered one item, disregarding possible interactions with other ones. In practice, items may interact in many ways:

- *Shelf space.* Limited shelf space at a retail store must be allocated to different items; then, stocking decisions are not independent.
- *Substitute products.* A stockout on one item may be almost irrelevant, if the customer switches easily to a substitute item that we sell.
- *Complementary products.* Stocking out on one item may have a detrimental effect on other items, too; imagine a customer order consisting of several lines, related to different items; the customer could cancel the order if some order line is not satisfied.
- *New products.* If the assortment is changed by the introduction of a new item, sales of older items are likely to be affected by cannibalization.

[9]Incidentally, in such a case, a trade promotion might have the only effect of making the life of logistic managers a misery, without contributing much to the bottom line. In fact, some retail stores adopt a policy of *everyday low prices*.

By the same token, we should not disregard the role of competitors. If a competitor is about to launch a new product, we should plan in advance a suitable reaction. In such a case, just looking at past sales will be as safe and smart as driving a car by just looking into the rear mirror. Furthermore, pricing is likely to affect sales along multiple dimensions: the price of the item, the price of related items, and the price asked by competitors. We should investigate, among other things, the relationship between price and demand. One way of doing that involves regression models, which are the subject of Chapters 10 and 16.

The role of sampling uncertainty. When we evaluate summary measures such as the mean of a variable, we look at a limited set of past data. But how reliable is that information? Intuitively, the more data we have, the better. However, looking too far into the past is dangerous, as we might take into account information that is hardly relevant for new market conditions. Would you use information about stock returns before World War II to manage a pension fund? In Chapter 9 we deal with inferential statistics, which may help us in assessing the degree of confidence we can have in our estimates.

Observables vs. unobservables. Finally, available data may not be what is actually relevant and needed. We have taken for granted that demand data were at our disposal. Unfortunately, in many practical settings we do not really observe demand, but *sales*. If there is a stockout on an item at our retail store, will the customer inform anyone that her demand was unmet? Not necessarily; maybe she will just go and buy at another retail store. If we are lucky, she will just settle for a product substitute. But even in this case, what we gather is sales data by scanning bar codes at the cash desk or point of sale. Clearly, this can result in an underestimation of actual demand. In other cases, data are available, but they are not gathered because of a wrong business process. Imagine a business to business setting, where a potential customer calls about immediate product availability. When the desired items are not in stock, she tries with another supplier. If the business process is such that only agreed-on orders are entered into the information system, disregarding lost sales, we are underestimating demand again. Often we have to settle for a proxy of what we cannot observe directly, and this might affect decisions.

After this long list of complicating factors,[10] you may feel a little overwhelmed, but please don't: There is a long array of powerful quantitative methods that we may integrate with good old common sense in order to tackle challenging

[10]If you hate supply chain management, and you are fond of finance, do not fret. The good news is that you will see plenty of examples related to finance in what follows. The bad news is that the list of complicating factors is overwhelming in that setting as well.

problems. Anyway, if you want to take a short route, you could always try to do just a little better than your competitor. Say that he is able to meet demand completely in 80% of the weeks. The probability of not having a stockout is one of the many ways in which one can measure the *service level*. By choosing a stocking level S, you are setting the probability of satisfying all customers. For instance, with the data of Table 1.2, if you choose $S = 1$, demand will be met only when it is 1, which happens with probability 0.15. If $S = 2$, you meet demand when it is 1 or 2; summing the respective relative frequencies, we see that this happens with probability $0.15 + 0.40 = 0.55$. The pattern is clear: The service level for increasing stocking levels is obtained by summing probabilities. This leads us to the concept of *cumulative relative frequencies*, which are displayed in the last column of Table 1.2. If you want to do a little better than the competitor, setting a 85% service level, you should stock four items, implying a 90% service level.

This last observation leads us to concepts such as cumulative distributions and quantiles. They are fundamental in many areas, such as inferential statistics and risk management, and will be among the most important topics in Chapters 6 and 7.

1.2.2 Squeezing information out of known facts

As we have remarked in the previous section, we often have to cope with the impossibility of gathering in advance all the information we need to make a decision, partly because of uncertainty about future events, and partly because some variable cannot be observed directly. However, this is no good reason not to do our best to exploit *partial* information. In customer relationship management, we are not able to read a customer's mind; nevertheless, we may observe her behavioral patterns and try to infer if she will be loyal and hopefully profitable, or not. The analytical tools of probability and statistics are commonly used for this kind of analysis.

A typical and quite familiar domain in which we have to settle for an observable proxy of an unobservable variable is medicine. Hence, to get a feeling for the involved issues, let us consider a medical problem. Quite often we undergo an exam to determine whether we are suffering from a certain illness. Thankfully, it is a rare occurrence that our bodies must be ripped open in order to check the state of the matter. Clinical tests are an indirect way to draw conclusions about something that cannot, or should not, be inspected too directly. Now, suppose that we have the following information:

- A kind of illness affects 0.2% of a population.
- A test can reveal the illness, *if a person is ill*, with a probability of 99.9%; i.e., if you are ill, there is a small probability (0.1%) that the exam will fail to detect your true state.
- The probability of a false positive is 1%; in other words, if one is perfectly healthy, the test may wrongly report the illness with probability 1%.

The issues of how this information is obtained and how reliable it is pertain to the domain of inferential statistics; for now, we will take someone's word for it. We see that there are two kinds of potential errors:

- We may fail to detect illness in an ill subject (missed positive).
- We may see a problem where there is none (false positive).

We would like to know whether a person is ill, but the only thing we can observe is the result of the test. If the test is positive, we would like to draw the conclusion that the patient is ill in order to start an adequate treatment. However, we cannot be that sure. Hence, the question is: If the test is positive for a person, what is the probability that he is actually ill?

When I ask students this question, I make it somewhat easier and just ask them to tell in which of the following intervals the required probability lies: $(0, 25)\%$, $(25, 50)\%$, $(50, 75)\%$, or $(75, 100)\%$?

Please! Pause for a while before reading further, and try giving your answer.

If you have selected an interval with high probability, you are in good company. From my experience, the most voted interval is $(50, 75)\%$. I guess the reason is that the professor maliciously insists on pointing out that the test is reliable, in the sense that if you are ill, the test will tell you that with high probability. Some students take me seriously, and vote $(75, 100)\%$. Most students probably temper that with a bit of skepticism and choose $(50, 75)\%$, based on some *good old intuition*. Very few students choose the correct answer, i.e., $(0, 25)\%$.

Let us try to get the answer by a somewhat crude line of reasoning.

1. Say that our population consists of 10,000 persons. How many should we expect to be ill? Taking 0.2% of 10,000 yields 20 persons; 99.9% of them will be reported ill, which is 19.98 (almost 20).

2. How many false positives do we expect to get? There are (on the average) 9980 healthy persons out of the total of 10,000; 1% of them will be incorrectly reported ill, i.e., 99.80 (which is almost 100, i.e., 1% of the whole population).

3. Then, on the average we will get

$$19.98 + 99.8 = 119.78$$

positives, but only 19.98 of them are in fact ill. So, the correct answer to our question is

$$\frac{19.98}{119.78} = 16.68\%$$

This is certainly much less than 50%. The problem is that the fraction of ill people is (luckily) low, in fact much lower than the number of false positives. Indeed, we could get close to the right answer simply by noting that the fraction of false positives is about 1% and the fraction of correct positives is about 0.2%; if we take the ratio

$$\frac{0.2}{0.2+1} = \frac{1}{6} = 0.1667 \approx 16.68\%$$

we see that it was *bad* intuition leading us into the trap, as a closer look into the data is enough to show that false positives have a large impact here.

Now, on the basis of this result, you should sit a while and broaden your view by asking a few questions:

1. If this is a cheap test, should we use a more expensive but more reliable one?
2. Should we use this test for mass screening?
3. What happens in the case of a false positive?
4. What are the social and psychological costs of a false positive?

Of course, there are no general answers, as they depend on the impact of this illness, its mortality rate, etc. Moreover, spending resources for this illness may subtract funds from programs aimed at preventing and curing other types of illnesses. The last question, in particular, shows that in a decision problem there are issues that are quite hard to quantify.[11] So, quantitative methods will definitely *not* provide us with all of the answers. Still, they may be quite useful in providing us with information that is fundamental for a correct analysis, information that we are going to miss if we rely on *bad* intuition.

Generalizing a bit, we note that a priori, if you pick a person at random the probability that he is ill is 0.2%. If you know that the test has been positive, that probability is updated to a larger value. The new probability of the event of interest (he is ill) is *conditional* on the occurrence of another event (the test is positive). We will learn about conditional probabilities in Chapter 5, where we develop a tool to tackle such problems more systematically, namely, Bayes' theorem. Bayes' theorem is the foundation of Bayesian statistics, which is increasingly relevant in practical settings in which either most past data are not relevant or you have none, and you have to rely on your subjective assessment (see Chapter 14).

[11] I recall the story of that guy who was diagnosed in a terminal state a while ago. He started spending all of his money to live what precious time was left to him in the most entertaining way. The good news was that the doctors were wrong; you imagine the bad news.

1.2.3 Variations on coin flipping

In this section, we consider the mother of all random experiments: coin flipping. This is a good way to get acquainted with probabilities, which are not necessarily relative frequencies, as well as with investments under uncertainty. We illustrate plain coin flipping first, and then some more interesting variations on the theme.

Plain coin flipping Assuming that we flip a fair coin, what is the probability of its landing head? Probably, you will not flip the coin one million times to conclude that this probability is 0.5. The reasoning you are following is, in fact, based on some form of symmetry that is exploited whenever you think about gambling by simple mechanisms such as throwing dice, picking cards at random, or spinning a roulette. Hopefully, you start to see that there may be different concepts of probability, a topic that we will discuss a bit in Section 5.1. The very nature of probabilities has been the subject of heated debate, but let us leave such philosophical issues aside and assume that you play a simple lottery. A fair coin is flipped: You win \$10 when it lands head, and you lose \$5 when it lands tail. If the game is repeated a large number of times, how much money should you win with each flip, on the average? Looking back at how we computed average demand in Section 1.2.1, you might suggest the idea of multiplying each possible payoff outcome by its probability:

$$\mathrm{E}[P] = 0.5 \times 10 + 0.5 \times (-5) = 2.5$$

We are using a new notation here, as $\mathrm{E}[\cdot]$ denotes the *expected value* of a random variable. Indeed, we are not taking averages based on observed data, looking backward; rather, we are stating something about the future, looking forward.

It would be nice to play that game a large number of times without having to pay for the privilege of flipping the coin. Still, a basic law of economics says that there is no free lunch; hence, someone will ask you for some money to play the game. How much money would you be willing to pay exactly? Suppose that someone says that, in order to play a fair game, the price is \$2.5; note that, if you pay that amount, the expected overall profit is

$$-2.5 + 0.5 \times 10 + 0.5 \times (-5) = 0$$

which does sound fair. Would you accept? Maybe yes, as the worst that can happen is losing \$7.5, which will not change your life. But let's scale the game up a million times: You may win \$10,000,000, you may lose \$5,000,000. Would you still be willing to pay \$2,500,000 for playing the game? I guess that the answer is no; probably, you would not play the game even for free.

What you have seen is an example of how risk aversion affects our behavior. This is a fundamental ingredient of any decision under uncertainty, including investments in financial assets or new products. Decision making under

uncertainty and risk aversion are the subject of Chapter 13. For now, let us consider a simple investment decision that looks quite similar to coin flipping.

Fancy coin flipping Suppose that you are the lucky manager of a movie company, which has just signed a contract with a new and promising director.[12] The contract provides that you may produce zero, one, or two movies with this director, during the next 2 years. To clarify, we could produce a movie now, investing some money immediately and hopefully collecting some payoff in, say, one year. At the end of the first year, we could produce a new movie,[13] collecting revenues at the end of the second year. One decision we have to make is whether we should produce those movies or not. In pondering the decision, note that any advance money you have paid to the good director is gone; in other words, it is a *sunk* cost. If, on second thought, you feel that producing a movie with him will turn out a disaster, you should not suffer from the sunk cost syndrome, i.e., the tendency to go on at any cost since you have already spent money on your endeavor.[14] On the contrary, you should only consider the money you are about to spend now to produce the movie, and the prospects for future revenue.

To spice things up, let us say that there is another dimension, as we may select one out of two production and marketing plans, a conservative and an aggressive one. We may invest more money in well-known actors and/or special effects, as well as in marketing the movie around the world.

1. Production/marketing plan A costs 2500 whatever monetary unit you like. This is the money you have to spend for the privilege of flipping the coin. We have some uncertainty about the success of our pet group. If the movie is successful, revenue after one year will be 4400 the same whatever; if it's a flop, you make nothing. To consider a case as close as possible to coin flipping, and to simplify calculations, let us assume that the probability of a success is 50%. Of course, coin flipping is a rather crude model of uncertainty for such a situation, and you should really ponder how to come up with a set of sensible scenarios along with their probabilities. By the way, the astute reader will suspect that here we should work with yet another, more subjective concept of probability. The symmetry of coin flipping or dice throwing does not apply and if this is an unknown director, looking at the past relative frequencies case might just make no sense. This happens whenever you deal with a brand new product.[15]

[12] This section is based on an example described in HBS business case [8].

[13] Well, maybe it will not really be a *new* movie; sequels are all the rage, aren't they?

[14] While we are talking about movies, imagine that you have paid a ticket for a movie that you find disgusting and/or insufferably boring after 20 minutes. Should you stay seated there for two more painful hours, just because you paid for the ticket?

[15] As a case in point, when the first IBM computer, a piece of hardware as large as a big room, was working (and heating whoever was around it), an IBM boss said that it was

2. Production/marketing plan B is more aggressive and expensive, as it costs 4000, but it increases revenue by 50% in the case of a hit, without changing probabilities (hence, we have two possible outcomes: 6600 or 0, with 50% probability). Arguably, a different marketing plan could also change success/flop probabilities, but for the sake of simplicity we will neglect this.

Both plans look like a coin flipping experiment, but in this case we should consider time or, to be more precise, the time value of money. Let us digress a bit in order to clarify this issue.

A little digression: time value of money Would you prefer to receive \$100 right now or \$100 in one year? Well, easy answer, and you surely concur that \$100 in one year is not the same as \$100 today. But what if you have to choose between \$100 right now and \$106 in one year? In order to make the two amounts comparable, the standard approach is to take money in the future and discount it back to now. To do so, we apply a *discount rate*, which is related to an interest rate.

Say that you may invest money at an annual risk-free interest rate of 5%; if you invest \$100 now at that rate, in one year you will own

$$\$100 \times (1 + 0.05) = \$105$$

The general rule is that to evaluate money in one year, you multiply money now by $(1+r)$, where r is the prevailing interest rate. If you leave that money invested for 2 years, and interest on interest is earned, the money in 2 years will be

$$\$100 \times (1 + 0.05) \times (1 + 0.05) = \$100 \times (1 + 0.05)^2 = \$110.25$$

Note the effect of compounding; without that, in 2 years you would get only \$110, i.e., the initial capital plus twice the interest payment. Now, if you see things the other way around, how much are \$106 in one year worth now? A little thought should suggest the following rule: Discount money in the future using the interest rate, and compare it with money now:

$$\frac{\$106}{1.05} = \$100.95 > \$100$$

which suggests the opportunity of waiting one year to grab almost one extra dollar.

This is what cash flow discounting is all about. We will see a bit more about financial mathematics later.[16] In practice, the real issue is estimating the

a very nice piece of equipment and that probably they were going to sell 10 in the entire world. Hardly a good prophecy, and the same difficulty in forecasting sales applied when the first cell phones were being produced.

[16] See Examples 2.7 and 2.33.

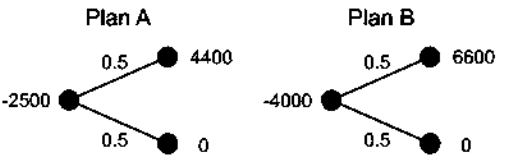

Fig. 1.7 Graphical representation of alternative plans for producing and marketing a movie.

uncertain cash flows and choosing a suitable discount rate. This may be difficult, as the discount rate should take investment risk into account. For now, let us assume that we discount cash flows with an annual discount rate of 10%.

Back to fancy coin flipping One sensible starting point in deciding whether we should trust the good director with our money is to find out how much we expect to make using each alternative marketing plan to produce one movie. Let us consider plan A. We pay 2500 for sure now, and there is a 50–50 chance of making 4400 or 0 in one year. Taking discounting into account, the money we expect to make is

$$-2500+\frac{0.5\times 4400+0.5\times 0}{1.10}=-2500+\frac{2200}{1.10}=-2500+2000=-500 \quad (1.4)$$

This −500 is the (expected) *net present value* (NPV) of our investment. A basic rule of investment analysis says that we should invest in projects with positive NPV. Unfortunately, the idea of adopting marketing plan A to produce a movie now does not seem quite promising. What about using the bolder plan B? Well, repeating the analysis yields the following NPV:

$$-4000+\frac{0.5\times 6600+0.5\times 0}{1.10}=-1000$$

an even bleaker perspective. A useful way of visualizing lotteries is illustrated in Fig. 1.7, which depicts a *scenario tree*. The leftmost node corresponds to "now," i.e., the current state of the world, where we have to decide whether we should spend 2500 or 4400 to start one of the marketing plans. The two nodes on the right correspond to possible states of the world in the future. In our case, we have just two possible scenarios, success or flop, which are associated to given payoffs and probabilities. Clearly, this a two-state scenario tree is a very crude representation of uncertainty; in practice, scenario trees may be definitely richer.

It seems there is no hope, but wait: We just considered one movie! What if we produce two over the next 2 years? If one game of coin flipping is a bad lottery, can we make it any better by just repeating it? To illustrate the point with plain coin flipping, what is the probability of getting two heads in a row, or any other possible outcome? The possible head–tail sequences when you

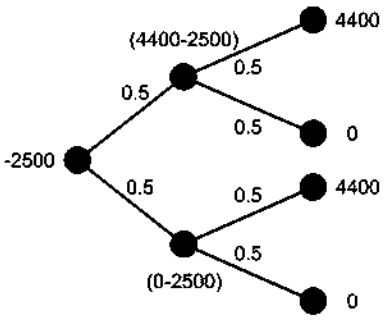

Fig. 1.8 Repeating the fancy coin flipping experiment twice.

flip a coin twice are

$$HH, \qquad HT, \qquad TH, \qquad HH$$

The usual symmetry argument suggests that none of these four outcomes is more likely than the others; hence, the probability of each sequence should just be $\frac{1}{4} = 0.25$ (note once again that probabilities add up to 1). A deeper reasoning should reveal that, unless the coin has some memory, the result of the second flip has nothing to do with the result of the first one; the two events are *independent*. We will learn in Chapter 5 that the joint probability of two independent events is just the product of their probabilities. For instance, we obtain the following equation, which confirms the previous result:

$$\mathrm{P}\{HH\} = \mathrm{P}\{H\} \times \mathrm{P}\{H\} = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

Now, what should we conclude if we consider the movie production problem as no more than a fancy variation on coin flipping? The situation is depicted in Fig. 1.8, where we produce two movies using plan A (incidentally, we ignore inflation and assume that future production and marketing costs will be the same as today). We have four equally likely scenarios: success–success, success–flop, flop–success, flop–flop. Each scenario has probability 0.25. We should discount cash flows occurring in 1 and 2 years; we should also consider the *net* cash flow in one year, resulting from collecting revenues (if any) and investing in the new movie. For the success–success scenario, the NPV is as follows:

$$-2500 + \frac{4400 - 2500}{1.10} + \frac{4400}{(1.10)^2} \approx 2863.64$$

Please carry out the calculation for the other three scenarios, and verify that the expected NPV is

$$\begin{aligned} &0.25 \times 2863.64 + 0.25 \times (-772.73) + 0.25 \times (-1136.36) \\ &\quad +0.25 \times (-4772.73) = -954.55 \end{aligned}$$

The result is negative again, but this is hardly a surprise; if flipping a coin once has a negative expected profit, repeating that gamble twice will not produce a positive result, because we are replicating the same experiment. In fact, a much smarter way to obtain the result is by noting that we are just carrying out the experiment with plan A twice in a row, and we may use the result of Eq. (1.4). We have just to discount the expected NPV for the second movie back from the end of year 1 to now:

$$-50 + \frac{(-50)}{1.10} = -954.55$$

By the same token, if we try mixing plans A and B in some sequence, we will hardly get any better, *if* our problem is just a fancy variation of coin flipping.

But is the movie production case really like coin flipping? In Section 1.2.2 we learned about *conditional* probabilities: If the test is positive, this does change your probability of being ill. If the first movie is a success (or a flop), can we say that producing the next one will just be another coin flipping experiment? Probably not, as the result of the first trial tells us something about market reaction to our product. In fact, we need an assessment of *conditional* probabilities, i.e., the probability that the second movie is a success (or a flop), conditional on the fact that the first one has been a success (or a flop). Reasonably, if the first movie is a success, the probability that the second one is a success as well is greater than 0.5. By the same token, if we have a flop in the first trial, we raise our expectation for another flop with the second movie. Coins are supposed to have no memory, but the tendency to stack movie sequels one after another suggests that moviegoers are not memoryless coins. Estimating those conditional probabilities may be quite hard, but let us consider a very extreme and overly simplified case, just to illustrate the point. Assume that if the first attempt is a success, the second one will certainly be as well (i.e., the conditional probability of a second hit after the first one is 1); on the other hand, assume that after a flop, we will have another flop for sure. In such a case, it is easy to see that if the first movie is a success, we should certainly produce the second one; on the other hand, if the first movie is a flop, we will just cut our losses and forget about it. This is what managerial flexibility is all about; we do not plan everything in advance, but we adapt our decisions when we gather additional information and revise our beliefs.

One sensible idea is to be conservative at the first trial and using plan A for the first movie; if we are lucky, and discover that the director can deliver a successful movie, we will be more aggressive and take plan B next year. This strategy is depicted in Fig. 1.9. In the figure, we do not clearly distinguish nodes at which we make a decision and nodes at which we observe results, but we will see a clearer representation of a dynamic decision-making process under uncertainty in Chapter 13, where we deal with decision trees. If we

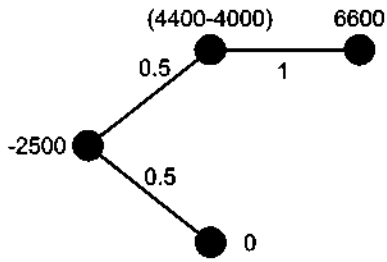

Fig. 1.9 Fancy coin flipping with memory.

calculate the expected NPV, we obtain

$$-2500 + \frac{0.5 \times (4400 - 4000) + 0.5 \times 0}{1.10} + \frac{0.5 \times 1 \times 6600}{(1.10)^2} = 409.09$$

Please beware of a common mistake. The last cash flow in the expression, 6600, occurs with probability 0.5, not 0.25. In order to evaluate the probability of each node in the scenario tree we must take the product of the conditional probabilities along the path leading from the root of the scenario tree (the "now" node) to that node. In our case, $0.5 \times 1 = 0.5$. This positive result shows that maybe our investment in the director was not so bad. Clearly, we should assess how much this conclusion depends on the probabilities that we have assumed. This process of checking the influence of uncertain data on the decision suggested by a model is called *sensitivity analysis*.

If you want to generalize a bit, a situation like this is common in risky research and development (R&D) projects. You could try a risky venture that, if successful, will pave the way to huge revenues by expanding the business. When analyzing investments, one should account for managerial flexibility and the possibility of learning over time by gathering new information. This approach leads to so-called real options;[17] in the real options literature, our example is known as a *growth* option.

Coin flipping in the short or long run Before leaving coin flipping aside, it may be instructive to consider in more detail the idea of repeating bets. Suppose that you can play the plain coin-flipping game, where you may either win 10 million or lose 5 million of whatever monetary unit you like. The coin is fair and memoryless, and you can play the game for free, so that the expected payoff is 2.5 million.

1. Would you play the game once?

2. Would you play the same game 1000 times in a row?

[17]They are called "real," as opposed to financial options based on financial assets such as stock shares. We will consider options in a few examples in the following chapters.

The answer to the first question might be no, as the chance of being broke is far from negligible, even though the expected payoff is rich. But if you may repeat the game several times, you have the possibility of recovering your losses. In the long run, the bet is quite profitable indeed, but this is no solace if you can get out of the business after a losing streak of bad outcomes, without the time to recover.[18] Again, risk aversion does play an important role and considering long-run, or expected, profits may not be always appropriate. In other words, the plain expected value of the payoff fails to take into account risk in the short term. In the movie production case, we have taken for granted that a proper discount rate can take risk aversion into account, but the issue is far from trivial and is the subject of quite some controversy in corporate finance. We investigate the role of risk in decision making under uncertainty in Chapter 13.

1.3 ENDOGENOUS VS. EXOGENOUS UNCERTAINTY: ARE WE ALONE?

A colleague of mine draws the line between engineers (like myself) and economists as follows:

- An engineer believes that fellow human beings are a little dumb, but otherwise they are generally good chaps. All you have to do is to find a clever solution for them, and they will live happily ever after.
- An economist believes that fellow human beings are not dumb at all. The problem is that you cannot turn your back on them, unless you have a taste for an instant stab.

In fact, the need for analyzing the interactions between noncooperative decision makers was the driving force for the development of game theory. We will consider related issues in Section 14.3, where we investigate the role of misaligned incentives when multiple stakeholders are involved in a decision-making process. The interaction of multiple actors may also change the nature of the uncertainty involved in decision making. Simple probability models assume that uncertainty is known and fixed; if we have enough information, we can make a good decision. However, it might be the case that the very uncertainty is influenced by decisions.

Example 1.1 To illustrate, let us go once again back to the inventory management problem of Section 1.2.1. There, we used historical data to characterize demand uncertainty and planned on using that information to come up with a decision. One tough question is: Does our decision affect the demand

[18] As the economist John Maynard Keynes aptly put it, in the long run we are all dead.

distribution? To get the point, imagine that you stock very few items. In the long run, you will end up losing customers, not just orders. What's even worse, maybe this will also affect the demand for other items. □

In practice, there are many cases in which demand is affected by stocking decisions. At retail stores, the shelf space allocated to a product may have a remarkable effect on demand. More generally, uncertainty need not be purely *exogenous*, i.e., given by outside factors that are not going to change (at least in the short term); uncertainty may be partially *endogenous*, i.e., influenced by our actions. Representing the dependence between our decisions and uncertainty may be very difficult, but at least one should be aware of the issue.

More generally, we should consider that management decisions are made in a social context, where other people could

- React to our decisions
- Use the same information we have
- Observe our behavior to gather information

Example 1.2 To see another example in a supply chain management context, suppose that we are selling a perishable item. We start selling it in the morning, and we know that at the end of the day we will scrap what's left. One could consider a dynamic pricing strategy, whereby price is marked down at the end of the day; after all, it seems better to recover some money even if this implies selling below cost. That's fine, but we are not considering the possibility of *strategic* behavior on the part of customers. Knowing that price will be reduced, they could wait and buy at dusk, just before closing time.[19] The point is that there is no separation between the set of customers who buy during the first part of the day and customers who buy later. We may apply price discrimination, i.e., asking different prices for the same product on different markets, provided they are well separated, possibly in space. The above revenue management policy is a sort of failed price discrimination strategy in time, rather than in space.

There could also be other reasons not to mark down some products. Some fresh and quickly perishable produce is not really sold for profit at retail stores. The real intent is to promote a positive image of "freshness" that has a positive impact on sales of other, more profitable items. The message would be lost by selling almost perished items in the evening. □

Social issues are even more fundamental in finance. A whole discipline, behavioral finance, was born to investigate the interplay between psychology and

[19]This is what happened at the fish counter at a big retail store in front of my home. Very few people bought fish all day long, but queues materialized after 6 p.m., not to mention a terrible smell. As you may imagine, this brilliant revenue management policy disappeared rather quickly. Much to the chagrin of my nostrils, the smell didn't follow.

finance. When you buy a good, you know its selling price and you may plan accordingly. However, financial markets are based on a competitive auction mechanism, so that prices depend on investors' behavior, on their beliefs (or lack of information), and on their beliefs about other investors' beliefs. The resulting pattern may be quite complex, and indeed it leads to bubbles and crashes. Everything is made even more complicated by deliberate trading strategies.

Example 1.3 (Short selling) A *short sale* is a trading strategy whereby you sell an asset, say, a stock share, that you do not own. To do so, you have to borrow that asset from someone else. It is a strategy that makes sense if you expect the asset price to drop. To see this, assume that the current price is $S_0 = \$50$ at time $t = 0$. If you sell the asset short, you borrow the asset and sell it for \$50. Clearly, you will have to give the asset back to its legitimate owner at some later time $t = T$, which means that you will have to buy the asset at a price S_T in the future. If you are right and the asset price falls, e.g., to $S_T = \$40$, your reward will be \$10 per share.

At times of financial turmoil, wealthy speculators could do a lot of short selling, which may itself depress prices, resulting in a self-fulfilling prophecy. Selling an asset short in large volumes can lead to lower prices, even if the economic and business fundamentals of the firm are good. Uninformed traders may just follow the lead, just because they see a drop in price, even if they do not know why this is happening. A perverse feedback mechanism can be the result, and indeed short selling is sometimes prohibited, when markets are under severe pressure. The role of short sellers (the "shorts") is actually debated, but the point here is that the interaction between actors can result in weird patterns.

A similar speculative practice is *predatory trading*. Suppose that you own a large amount of an asset and you know that Mr. X has to sell a large amount of the same asset, because he is running short of liquidity. You could sell the asset before Mr. X, which will result in a significant price drop, as the large trading volume will affect prices. Mr. X will have to sell, anyway, which will depress prices further. Then, you may buy the asset back at a lower price, getting back to your previous portfolio, plus some extra cash. ∎

Empirical research is carried out to verify if certain speculative practices such as predatory trading are actually carried out at a significant level. These very empirical research lines do use a lot of statistical methods, and have a practical impact as market regulation should be devised in such a way to avoid excessive distortions. Empirical finance is beyond the scope of this book, but in the next section we analyze in more detail a somewhat stylized, yet very significant, example to see the interplay between organized markets and uncertainty.

1.3.1 The effect of organized markets: pricing a forward contract

An asset (e.g., a stock share) is sold now at a spot price $S_0 = \$50$, where $t = 0$ is current time. The spot price is the prevailing price at which the asset is exchanged at any moment. The spot price in one year is, of course, uncertain, but say that the expected price in one year is $\mathrm{E}[S_1] = \$60$. A type of contract that is commonly traded on many assets is a forward contract.[20] This contract is signed between two parties who agree on exchanging the asset at some future time (say, one year) for a price that is agreed on *now*. This fixed price, that we denote by F_0, is the *forward price* at time $t = 0$. In one year, one of the two parties will sell the item to the other one at price F_0, no matter what spot price S_1 will prevail in one year. The party agreeing to buy is said to hold the *long position*; the party agreeing to sell is said to hold the *short position*. Depending on the relationship between the agreed-on forward price F_0 and the future spot price S_1, one of the two parties will turn out to be happy, at the expense of the counterpart. If it turns out that $S_1 > F_0$, the long position will benefit since she can buy at a price lower than the current spot price. The short position will benefit if $S_1 < F_0$, since she can sell at a higher price. A priori, there are different reasons for trading such contracts. One reason is hedging risks away: By fixing the price for the future, long and short parties eliminate risk ex ante. Ex post, one of them will regret the hedging decision, but risk management strategies should be evaluated *before* uncertainty is resolved. Another reason could be speculation of traders having firm beliefs about future spot prices and willing to bet on their forecasts.

Now the problem is: What should F_0 be? Actually, it is the balance between demand and offer that determines prices; still, there are many reasons why one should be interested in finding a sensible estimate of a fair price. One possible guess is that a fair price should be $\mathrm{E}[S_1] = \$60$. Note that money and the underlying asset are exchanged in the future, so we do not need to discount cash flows, as the forward price F_0 will be paid in one year. We have seen before that using expected values to value assets or lotteries ignores risk aversion. Taking this into account, one could argue that the price should be a bit less than \$60. But how much less? There are many actors in the markets; whose risk aversion matters most? It seems that there is no hope to find a sensible forward price! By the way, each actor might have a different view of the future, as well as different information, but to keep it simple, let us assume that everyone agrees on the above expectation about S_1.

To get some clue, let us assume that the forward price is set as the expected future (spot) price, $F_0 = \mathrm{E}[S_1] = \$60$. Also assume that we may lend or borrow money at an annual risk-free rate of 10%. In this setting, you could adopt the following trading strategy: Borrow \$50 now, buy the asset, and

[20] In practice, futures contracts are more commonly traded. They differ from forward contracts with respect to some institutional features that enhance their liquidity, such as standardization of contracts and measures to avoid defaults by one of the two parties.

enter a short position in the forward (i.e., you agree to sell the asset at the forward price in the future, no matter what the spot price will be). This trade is costless at time $t = 0$, since cash flows cancel each other and (in principle) no money is needed to enter a forward contract. In one year, you will pay back \$55 to the guy who loaned you that money but, whatever the spot price, you will be able to sell the asset for \$60, earning \$5 *for sure*. If this were the case, we would have found the perfect money-making machine. Please note that you make money *without taking chances* and *without any need for initial cash*, as the initial net cash flow is zero. If you enter a long position at $F_0 = \$60$ and the spot price turns out to be $S_1 = \$100$, you will make \$40 by using the forward contract to buy the asset at \$60 and selling it immediately on the spot market. This is much better than the \$5 you could make by the riskless strategy, but in doing so, you are taking chances, as the spot price could be well below the forward price and you would lose money. A riskless trading strategy that does *not* require any money is an example of an *arbitrage* opportunity.

Could we scale the investment up and make an arbitrary amount of money out of nothing?

Now, remember that you are not alone. If such an arbitrage opportunity were available, many other investors would follow your pattern and buy the underlying asset; this would increase the spot price (not to mention the fact that many people would borrow money, which would affect the interest rate). What could happen is not quite clear, as it depends on complex market dynamics, but we can conclude that those three numbers (spot price \$50, forward price \$60, risk-free interest rate 10%) are not consistent; they cannot be equilibrium prices and rates.

So, we know that \$60 is too high a forward price. It is easy to see that any price higher than \$55 leads to a similar arbitrage opportunity. On the other hand, assume that F_0 is smaller than \$55, say, \$52. In this case, you could reverse the trade by selling the asset short (see Example 1.3 on short selling). You borrow the asset, sell it at the current spot price $S_0 = 50$, invest the proceeds at the risk-free rate, and enter a long position in the forward contract. This long position comes in handy, as you will have to give the asset back to the lender in one year. You will do so at the given forward price, \$52, after collecting the money you lent plus interest, which is now \$55; then, you made \$3 for free. A similar strategy can be applied for any forward price less than \$55.

The conclusion is that the only arbitrage free forward price is \$55. Generalizing a bit, what we found is that, in order to rule out arbitrage opportunities, the forward price should be

$$F_0 = S_0(1 + R_f)$$

which is a somewhat stunning conclusion: uncertainty does not play any role, as it is the risk-free interest rate R_f that determines the forward price. This

conclusion should be taken with care, as it is at odds, e.g., with what we read on newspapers about speculation with futures on oil prices. In fact, the reasoning assumes somewhat idealized financial markets, and it can be more or less justified, depending on the actual asset we are dealing with. While a financial asset can be practically stored at no cost, oil cannot; by the same token, selling oil short is certainly not that easy. Pricing forward and futures contracts is affected by many issues such as transaction costs, differences between lending and borrowing rates, etc. Still, there is an important lesson here:

Uncertainty is important in decision making, but collective behavior may have some surprising effects on it.

1.4 QUANTITATIVE MODELS AND METHODS

Hopefully, the examples in the previous sections have shown the relevance, as well as the limitations, of quantitative analysis for real-life business decisions. In the following chapters we cover a wide variety of tools, which could be somewhat confusing for the reader. Hence, it is a good idea to set a conceptual framework to classify and interpret the different approaches, which should not be regarded as a disordered array of technicalities. Applying a quantitative analysis typically requires two steps:

- Building a quantitative *model* of a system, or part of it
- Solving it by some suitable *method*

To see the difference, consider the linear programming model to find the optimal mix in Section 1.1.2. After you have written the model down, a numerical solution procedure should be applied to come up with the answer we need. Lightning fast software packages to solve linear programming are widely available; hence, emphasis should be placed on model building, rather than model solving. In fact, data collection, a correct assessment of objectives and constraints, and the ability to interpret the solution are definitely more critical to success than plain number crunching.

Indeed, the title of the book is arguably somewhat misleading, even though quantitative methods is the standard name of courses covering these subjects. Nevertheless, it is often essential to have at least a rough idea of the inner working of solution methods, because

1. The way the model is formulated may have an impact on available methods, as well as their efficiency.
2. The solution method may be valid only if some assumptions about the data hold; ignoring these assumptions may result in gross mistakes if a method is applied to the wrong problem.

Another important healthy principle that we should keep in mind is[21]

All models are wrong, but some are useful.

Indeed, any model is a *simplified* representation of reality, and it need not be quantitative. Qualitative models are very useful in business process reengineering (BPR), and are often based on some graphical formalism to clarify relationships between actors, chain of events by activity diagrams, etc. Quantitative models, on the contrary, are based on *numerical* information.

1.4.1 Descriptive vs. prescriptive models

A quantitative model can be

- *Descriptive*, if its purpose is to shed some light on the relationships between two (or more) variables of interest (e.g., do sales depend significantly on advertisement expenditure?) or to predict system performance as a function of some design variable (e.g., the average waiting time in the queue, as a function of the number of tellers in a bank).
- *Prescriptive*, when the aim is (more ambitiously) to find a solution, subject to economical or technological constraints, so that costs are minimized or profits are maximized (see, e.g., the optimal mix problem).

Typical examples of descriptive models that we cover in the book are

- Simulation models for performance evaluation (Section 9.7)
- Linear regression models (Chapters 10 and 16)
- Time series models for forecasting (Chapter 11)

All of these models are used to generate information that helps in coming up with a decision, but they are not aimed at generating the decision directly. Examples of prescriptive models whose output is the decision itself are

- The economic order quantity model (Section 2.1)
- Linear programming models (Chapter 12)

It should be clear that prescriptive models are more ambitious and, in principle, they could even be used to automate the decision process. This applies to strictly technical problems, especially when the time to make a decision is quite limited. In most business settings, however, prescriptive models should be regarded as decision *supports*, and we must be aware of their limitations:

[21] The quote is generally attributed to the statistician George E.P. Box, 1979.

- The output is affected by data uncertainty (which is not only of a statistical nature; demand is uncertain because we cannot forecast the future exactly, whereas some cost are hard to quantify–how much is the inventory holding cost for an item?). Sensitivity analysis should be an integral part of a quantitative analysis.

- Some tradeoffs between conflicting objectives are difficult to assess quantitatively, possibly requiring the interaction between multiple stakeholders.

- More often than not, we have to approximate parts of a problem to make it tractable, and expert judgment is needed to evaluate the impact on these simplifications on the viability of the model and of the solution that we obtain from solving it. This process is called *model validation.*

Because of these limitations, it is sometimes argued that quantitative analysis should be confined to academia, and that common sense and a lot of practical experience are what is really needed. It is certainly true that human knowledge is an extremely valuable asset; what is not true is that it is incompatible with quantitative analysis. Furthermore, there is another more and more important factor: time. The rate of change in business conditions is faster and faster. For instance, when you change a product line, you have to redesign the whole supply chain to support its production. There is simply no time to do that manually, and some automatic support is needed. Intuition and experience are essential, but they are not enough.

1.5 QUANTITATIVE ANALYSIS AND PROBLEM SOLVING

Even if the problem is too complex to rely on the decision proposed by the solution of a model, we should not underestimate the value of model building per se. The model building process itself is a valuable activity as it requires the following ingredients:

- *Gathering data.* Quite often, complex organizations do not pursue a disciplined approach to data management. Important data are missed, some are duplicated with possible inconsistencies, some are not shared between different offices, and errors are not discovered because no one really uses that information. The need to collect data to build a model may force an improvement of the related processes. Furthermore, model building can help in transforming a huge amount of useless data, into useful and shared *information.*

- *Structuring the problem.* This requires sharing information and understanding the multiple dimensions of a problem, as well as the points of view of other stakeholders. This is important in large organizations,

where conflicting views are the rule rather than the exception, and entities within a firm pursue hidden agendas without agreeing on a shared understanding. It is always important to keep in mind that any quantitative analysis is doomed to failure if some relevant actor is not involved and motivated.

If model building and model solving result in a solution to the business problem, the solution itself must be implemented, monitored, and adapted when required by new circumstances. Fostering the culture and the discipline needed to monitor a solution and to assess the improvement in key performance indices is again valuable per se.

We close this section by remarking that quantitative analysis can play an important role in the following circumstances:

- When an objective support for decisions is needed. Rationalizing the analysis may ease potential conflicts between stakeholders.
- When the decision process leading to a recommendation must be explicitly documented. Because of uncertainty, even the best decision may result in a bad outcome. If you can back your decision with serious analysis, chances are that you will be able to save your neck.
- When the relationship between variables is too complex to be analyzed intuitively; in such a case, intuition can lead to wrong decisions because we are not able to fully grasp the impact of decisions.
- When the number of decision variables is too large to be managed even by the best human experts.
- When there are many difficult constraints and even finding a *feasible* (let alone optimal) solution is very hard. One example is train timetabling, which is a daunting task because of the number of shared resources (trains, crews, rails) and the constraints on their use, such as rules constraining how personnel shifts are scheduled.
- When tradeoffs between conflicting criteria must be assessed objectively (e.g., customer service vs. inventory holding cost).

Problems

1.1 Consider again the growth option problem of Section 1.2.2. We want to check the impact of less extreme assumptions about the conditional probabilities for the second movie, but we are unsure which values we should use. So, we assume that the conditional probability of a second success, after a first success, is $0.5 + \alpha$, for some unknown value of α; this is also the probability of a second flop, after a first flop. The analysis in Section 1.2.2 shows that

if $\alpha = 0.5$, we should produce the first movie. What is the limit value of α below which we should change our mind?

1.2 Consider the optimal mix model of Section 1.1.2. How could we extend the model to cope with

- Third-party suppliers offering items at given cost?
- The possibility of overtime work at some resource centers?

For further reading

- There are many books covering quantitative methods for business applications. Most of them deal only with one side of the coin, either probability/statistics or decision models. A book illustrating many different kinds of applications is Ref. [6], where many kinds of problems and methods are presented, even though skipping many details and applicability conditions. Another book aimed at a relatively broad treatment of quantitative models and methods is Ref. [2], which covers less material than the previous reference, but at a deeper level, and is also rich in interesting cases.
- The practical impact of quantitative analysis on business is emphasized in Refs. [1, 7].
- Last but not least, depending on your personal taste, you might be interested in books covering the application of quantitative analysis to specific application domains:

  Logistics. A wide range of applications related to supply chain management and logistics are covered in Ref. [4], ranging from strategic network design to operational routing of vehicles.

  Operations management. Both descriptive and prescriptive models for production planning and control are covered in Ref. [10]; in Ref. [3] emphasis is given to prescriptive models.

  Pricing and revenue management. Dynamic pricing strategies and revenue management are an essential tool for certain kinds of business application; a relatively informal treatment is given in Ref. [13], whereas Ref. [15] is more challenging and illustrates the application of sophisticated quantitative models.

  Marketing. Interesting applications of statistics to marketing are illustrated in Ref. [12].

  Finance. This is a domain where the role of quantitative analysis has boomed in the last two decades, with a considerable trail of controversy; introductory books that can help you appreciate the role

of quantitative models in financial markets are, e.g., Refs. [11] and [14]; a good reference for corporate finance is Ref. [9]; you may also appreciate the connection between the two fields in Ref. [8], which discusses option valuation for investment analysis.

REFERENCES

1. S. Baker, *The Numerati*, Houghton-Mifflin, Boston, 2008.
2. D. Bertsimas and R.M. Freund, *Data, Models and Decisions: The Fundamentals of Management Science*, Dynamic Ideas, Belmont, MA, 2004.
3. P. Brandimarte and A. Villa, *Advanced Models for Manufacturing Systems Management*, CRC Press, Boca Raton, FL, 1995.
4. P. Brandimarte and G. Zotteri, *Introduction to Distribution Logistics*, Wiley, New York, 2007.
5. L.M.B. Cabral, *Introduction to Industrial Organization*, MIT Press, Cambridge, MA, 2000.
6. J. Curwin and R. Slater, *Quantitative Methods for Business Decisions*, 6th ed., Cengage Learning EMEA, London, 2008.
7. T.H. Davenport, J.G. Harris, and R. Morison, *Analytics at Work: Smarter Decisions, Better Results*, Harvard Business Press, Boston, 2010.
8. M.E. Edleson, *Real Options: Valuing Managerial Flexibility*, Harvard Business Publishing, 2002, (Teaching Note 9294109).
9. D. Hillier, S.A. Ross, R.W. Westerfield, J. Jaffe, and B.D. Jordan, *Corporate Finance*, McGraw-Hill, New York, 2010.
10. W. Hopp and M. Spearman, *Factory Physics*, 3rd ed., McGraw-Hill, New York, 2008.
11. D.G. Luenberger, *Investment Science*, Oxford University Press, New York, 1998.
12. J.H. Myers and G.M. Mullet, *Managerial Applications of Multivariate Analysis in Marketing*, South-Western, Cincinnati, OH, 2003.
13. R.L. Phillips, *Pricing and Revenue Optimization*, Stanford University Press, Stanford, CA, 2005.
14. S.M. Ross, *An Introduction to Mathematical Finance: Options and Other Topics*, Cambridge University Press, Cambridge, UK, 1999.

15. K.T. Talluri and G.J. Van Ryzin, *The Theory and Practice of Revenue Management*, Springer, New York, 2005.