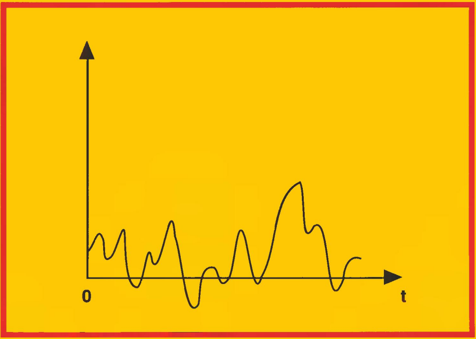
Undergraduate Texts in Mathematics

Readings in Mathematics

Richard Isaac

The Pleasures of Probability



Springer

Contents

1

Cars, Goats, and Sample Spaces

Behold, there stand the caskets noble prince:
If you choose that wherein I am contain'd,
Straight shall our nuptial rites be solemnized:
But if you fail, without more speech, my lord,
You must be gone from hence immediately.
William Shakespeare, *Portia* in *The Merchant of Venice*

1.1 Getting your goat

It's a critical moment for you. The master of ceremonies confronts you with three closed doors, one of which hides the car of your dreams, new and shiny and desirable. Behind each of the other two doors, however, is standing a pleasant but not so shiny and somewhat smelly goat. You will choose a door and win whatever is behind it. You decide on a door and announce your choice, whereupon the host opens one of the other two doors and reveals a goat. He then asks you if you would like to switch your choice to the unopened door that you did not at first choose. Is it to your advantage to switch (assuming, of course, that you are after the car, not the goat)?

This popular puzzler created a stir in 1991 when it appeared in the newspaper and (see [32] [1]) received a lot of wrong answers from readers, even from some who were mathematicians. How do we think about a problem

[1] Numbers in square brackets refer to the references at the end of the book.

like this, and why is it so tricky? (The most common wrong answer was that switching is irrelevant because each of the two unopened doors would hide the car with equal probability.) What I'd like to do is use this problem to introduce you to the branch of mathematics called probability. When you finish this chapter you should be able to think about the car-goat problem and many other probability problems in a reasonable way. Let's begin with a capsule description of probability theory, its significance and history, in a few action-packed paragraphs.

1.2 Nutshell history and philosophy lesson

Probability can be described as the mathematical theory of uncertainty. Its origins are undoubtedly ancient, since an early cave dweller looking at the sky for some clue about the weather was using primitive notions of probability. In fact, one could argue that all of us use probability daily to assess risks; the probabilities used are rough estimates based on previous experience. (Dark clouds today mean rain is likely since it has rained in the past when the clouds had that look. Better carry an umbrella.) Primitive or instinctive probability, however, is very different from a developed mathematical discipline. Officially, probability as a formal theory is sometimes said to have begun in the seventeenth century with a famous correspondence between the two French mathematicians Blaise Pascal and Pierre Fermat. The gambling halls of Paris were giving life to the new science. In a sense, a casino is almost a perfect laboratory of probability in action; a serious gambler has to have a pretty good idea of the risks in order to bet rationally. After a while the gambler either has to become a mathematician or consult one.

From these somewhat frivolous beginnings, the theory developed to its present status, with applications to all branches of science, technology, and even to that citadel of uncertainty, the stock market. Moreover, the twentieth century provided a new and rather startling star role for probability ideas within the framework of modern physics. In the physics of the eighteenth century, Newton's era, it was supposed that if you only had all the data you could use the equations of physics to predict the position and velocity of a particle exactly. Physicists viewed probability as a useful tool, mainly because it was often too hard to get all the data input for a problem. So probability was tolerated, in a sense, as a lesser discipline, because if our ignorance were only eliminated we wouldn't need probability, there wouldn't be any uncertainty, it was argued. For example, if we knew all about how a coin was tossed, the accelerations, angles, and forces involved, we could in principle predict whether a head or tail would come up. That was fine, until the new physics came along and Werner Heisenberg postulated the "uncertainty principle," that for very small particles it was impossible to know both the position and velocity exactly; the better you know the position, the fuzzier your idea of the velocity becomes, and vice versa, and there isn't anything you can do about it. This idea revolutionized the foundations of physics. Here was Heisenberg now saying that in principle you could not make exact predictions; the best you could do would be to make probability statements no matter how much data you collected. It was all very distressing to Einstein, who rejected Heisenberg's theory with his famous statement "God does not play dice." However, modern physicists now believe that Heisenberg was correct.

1.3 Let those dice roll. Sample spaces

Let's begin rolling dice, tossing coins, and other such things, because this is where the heart of probability lies. Since probability measures uncertainty, we have to measure something, and these objects probabilists like to call *events* since this is a reasonable name to give to the something that happens. Now suppose we are interested in what happens when we roll a pair of dice once. Assume one die is red and the other is green. When the red die falls it can come up in six ways, and the same holds for the green die. Each possible result can be represented by an ordered pair (a, b), where a is one of the numbers from 1 to 6 and represents what the red die's number is, and b is also one of the numbers 1 to 6 representing the green die's number. So what actually happens when you roll the pair of dice once? Well, there are 36 ordered pairs (a, b) where a and b vary between 1 and 6 (just write these all out to see that for $a = 1$ there are six possibilities for b, for $a = 2$, another six possibilities for b, etc.). What happens when you roll the dice can be conveniently described by exactly one of the 36 ordered pairs possible. Each of those 36 possible ordered pairs we call an *outcome*. Outcomes are the simplest kind of event. More complicated events contain a number of outcomes. For example, the event defined by the phrase "rolling a seven" contains six outcomes; it can be described as the event

$$A = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}.$$

Here the curly braces tell us to regard the enclosed items as being lumped together to form the event called A. We can say that the "experiment" of rolling a pair of dice once gives rise to a *sample space* S, which is just the set of all 36 of the possible outcomes, and that any event is simply a set of some collection of these 36 elementary outcomes or building blocks for events. For instance, the event

$$B = \{(1,1), (1,2), (1,3)\}$$

could be described in words as "rolling 1 with the red die and rolling between 1 and 3 inclusive with the green die." Figure 1.1 shows how the sample space S and the event B can be represented in a sketch, with the

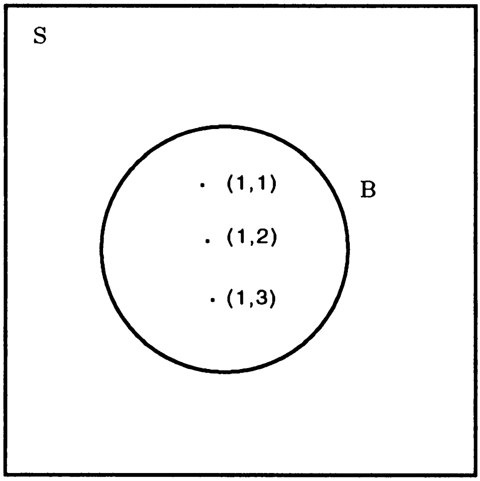


FIGURE 1.1. Venn diagram of a set B in sample space S

outcomes designated as points in the picture. Such a pictorial representation of sets is called a *Venn diagram.*

As we have seen above, an event is just a suggestive word probabilists use to talk about a set, namely, a collection of objects which, in the probability situation, is a collection of outcomes from a random experiment (the word *random* here means you can't predict the outcome in advance). As another example, the experiment of tossing a coin twice gives us a sample space S with four outcomes where, if we use H and T for head and tail, respectively, we can write:

$$S = \{(H,H), (H,T), (T,H), (T,T)\}$$

where the first entry in the ordered pair represents what happens on the first toss and the second entry what happens on the second toss. The event C, "at least one head occurs," can be written, for example, as

$$C = \{(H,H), (H,T), (T,H)\}.$$

There are a number of important points we should keep in mind about sample spaces. First, we used the word "experiment" to describe the happening that gives rise to the sample space. Experiments are usually things that can be repeated, and this is appropriate. That is because we will be considering probabilities for the most part for the kind of events that can arise only from some repeatable circumstance such as rolling dice or tossing coins. Suppose we are at a trial by jury; we will not consider an event like "the defendant is guilty" to be the kind of event to which we are going to attach a probability (at least for the moment) because it is not the kind of event arising from a repeatable experiment like rolling dice or tossing coins.

Another point is that a sample space provides what is called a *mathematical model* of the real-life situation for which it is supposed to be an abstraction. The reason for constructing this abstraction is that mathematical analysis can only be performed on the ideal structure of the sample space, not on the real-life situation itself. Once you have this model you may derive some nice mathematical relationships about the ideal structure, the abstraction. Since the abstraction resembles the real world, you might think that the mathematical relationships you found say something about the real world. You can now perform scientific experiments to check out the real-world situation. If you were clever and lucky, the mathematical model helped you decipher the real world; you know this because the results of your experiments are consistent with the mathematical relationships you obtained from the model. It could also happen that your model was too simple or otherwise in error and did not give a true picture of the real-world situation. In this case, the mathematical relationships, while true for the model, cannot be verified by laboratory experiment. Then it's back to the drawing board to look for a more accurate model.

It follows that since a sample space is constructed to model a real-life situation and is therefore only a construct, a figment of the imagination of the observer of that situation, it depends on what that observer thinks is important. For example, let us say that every now and then when you roll the dice, your dog jumps up on the table and grabs the red die in his jaws and runs under the couch with it. If you wanted, you could consider the sample space including with the 36 outcomes in S another six outcomes which could be represented as $(D,1),(D,2),(D,3),(D,4),(D,5),(D,6)$. Here $(D,5)$, for example, means that the dog has run off with the red die so no number has turned up on it but the green die came up with 5. Similarly, if the dog occasionally runs off with the green die or with both dice and we want to include sample points for these occurrences, we could add points to denote this (the geometric word "point" is a convenient and suggestive word for an outcome; it derives from the practice of drawing a picture of a sample space as in Fig 1.1, with the list of all possible outcomes as a scattering of dots or points inside it). The sample space representing what happens when a pair of dice is rolled is therefore not unique; it may be considerably more complicated than the one originally given by S. It all depends on what the problem is and what you judge to be the relevant information.

1.4 Discrete sample spaces. Probability distributions and spaces

So far, as you have noticed, we don't have the idea of probability at all in our mathematical structure, the sample space. All we have is the list of all possible outcomes that can be generated by the performance of some