CHAPTER 2. PROBABILITY SPACES.

Consider the experiment of rolling a single die. We make a mathematical picture of this experiment as follows. First, we decide on what to think of as possible basic outcomes of the experiment. In this case, we take six possible outcomes which we represent by the numbers 1,2,3,4,5,6. We next form the set of all these outcomes. In this case, it will be the set {1,2,3,4,5,6}. We call this set the sample space of the experiment, and the members of the set (which we have taken as the possible basic outcomes of the experiment) are sometimes called the points (or sample points) of the sample space. As we noted in Chapter 1, a set of possible basic outcomes of an experiment is called an event. Thus, in our mathematical picture, the events are represented by the subsets of the sample space, and we shall speak of these subsets themselves as events. An event in a sample space can be described by listing its members. For example, we might speak of the event {1,3,5}. It is often possible to describe an event in other ways also. For example, the phrase "result is odd" describes the same set of outcomes, {1,3,5}. Similarly, "result is less than 3" describes the set {1,2}.

We shall sometimes use capital letters to stand for events in a sample space. Let A be some given event for an experiment and a chosen sample space. If, in a given trial of the experiment, the basic outcome which occurs is in the event A, we say that the event A has occurred on that trial. For example, if we roll a die and get the number 1, then we can say that the event {1,3,5} has occurred. Note that the event {1,2} has also occurred. A single basic outcome may lie in many different events, and whenever that outcome occurs, we can also say that all those events have occurred.

Let us look at a particular basic outcome, say the outcome 1. If we repeat many times the experiment of rolling the die, we expect, by the strong stability of relative frequencies, that values of the relative frequency for this outcome will cluster more and more closely about some fixed value. Similarly for the other five outcomes. If we actually carry out the experiment many times, we find that each of the six basic outcomes has relative frequency close to $\frac{1}{6}$. (Indeed, the physical symmetry of the die, and our own past experience with this kind of physical symmetry, suggest to us ahead of time that the six outcomes will occur about equally often.) For these reasons, we take each point in the sample space and assign to it a numerical value representing the relative frequency that we expect for that point when the experiment is repeated many times. This means, in our present example, that we assign the value $\frac{1}{6}$ to each point in the sample space. We call these values <u>probability values</u>, and the assignment of values to points in the sample space is called a <u>probability function</u>. Once we have a probability function, we can then also give a numerical value to each event in the sample space by simply adding up the probability values of all the outcomes in that event. We call this the <u>probability</u> of the event. If A is an event, P(A) will stand for the probability of the event. Thus

P(A) = <u>the sum of the probability values for outcomes in A.</u>

In our example, P(less than 3) = P({1,2}) = $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$, and P(odd) = P({1,3,5}) = $\frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$. Since the probability values of the basic outcomes represent the expected relative frequencies of those outcomes, it follows that the probability of an event represents the relative frequency that we expect for that event. Thus we expect the event {1,2} to occur about one-third of the time and the event {1,3,5} to occur about half the time. In this way, a probability function on the <u>points</u> of the sample space gives us a probability value for each <u>event</u> in the sample space. (This assignment of probability values to events is sometimes called a <u>probability measure</u> on the sample space.) A sample space, together with a probability function (and hence a probability measure) on that space is called a <u>probability space</u>.

Note that when we choose a probability function, we always choose it so that the sum of the probability values for all points in the sample space is 1. (We do this because we know that on each trial, at least one of the basic outcomes must occur; this means that the event consisting of the entire sample space must occur with relative frequency 1.)

The probability space (that is to say, the sample space with its probability function) forms a mathematical picture or <u>model</u> of the experiment. This chosen model corresponds to the actual experiment in the following way.

Chosen mathematical model		Physical experiment
sample space	⟷	experiment
point in sample space ("outcome")	⟷	basic outcome
subset of sample space ("event")	⟷	set of basic outcomes ("event")
probability value of point	⟷	relative frequency of basic outcome for many repetitions of experiment
probability value of subset (= sum of probability values of points in the subset)	⟷	relative frequency of event for many repetitions of experiment

Note that we use the words "outcome" and "event" in talking about a probability space as well as in talking about the underlying experiment. Note also that, in general, many different models can be chosen for any single given physical experiment. Models for the same experiment can differ because of different choices of probability function. Models for the same experiment can also differ because of different choices of sample space. We discuss this further below.

When we form such a probability model, we cannot always be sure that it fits the experiment exactly, because we cannot be sure that we have chosen the best probability function. Consider, for example, the experiment of tossing a thumbtack described in Chapter 1. For this experiment, the sample space has two points which we call s (for on side) and b (for on back). To get a probability function, we must use our past experience with the experiment. The results described in Chapter 1 for a particular given thumbtack suggest that we take, for that thumbtack, a probability value of about 0.59 for s and a value of about 0.41 for b. We cannot be sure that these values are exactly right. We can only say that they are correct to two decimal places on the basis of our experience.

In certain cases, because of symmetry, and because of our experience with symmetry in physical experiments, we can assign a probability function without carrying out repetitions of the experiment. This is the case with the die, where we can see ahead of time that each basic outcome should get probability value $\frac{1}{6}$. It is also the case with the coin-tossing experiment. For the coin, there are two points in the sample space, and symmetry shows us that each point should get probability value $\frac{1}{2}$. (The correctness of the value $\frac{1}{2}$ (to several decimal places) is also shown by the data given in Chapter 1 for repeated tosses of a coin.)

The fourth example given in Chapter 1 was the birth-month experiment. In this case, there are twelve points in the sample space. We do not know, ahead of time, what probability function to use. If we try the experiment many times, or if we look up government records of birthdays, we will find that about the same proportion of people is born in each of the twelve months. This suggests that we assign the value $\frac{1}{12}$ to each of the twelve outcomes. (We could seek an even more accurate model by directly using the proportions occurring in government records or by arguing from these records that about the same number of persons are born each day in the year and taking the value $\frac{31}{365.25}$ for January, $\frac{28.25}{365.25}$ for February, and so forth.)

If we let x stand for a point in the sample space S of a model, and if we let p_x stand for the probability value assigned to x in that model, then we can describe the model itself by means of a two-line table where the first line lists the points x and the second line gives the corresponding probability values p_x.

For example, the model described above for the experiment of rolling a single die can be given by the table

x	1	2	3	4	5	6
p_x	1/6	1/6	1/6	1/6	1/6	1/6

In experimenting with a given die, we might find that the observed relative frequencies did not agree with our chosen model. In that case, we would be led to consider other models based on the data actually observed. For example, if we observed that outcomes 2, 3, and 6 each occurred about one-tenth of the time, that 1 and 4 each occurred about two-tenths of the time, and that 5 occurred about three-tenths of the time, we might adopt the model

x	1	2	3	4	5	6
p_x	0.2	0.1	0.1	0.2	0.3	0.1

instead of the model with equal probability values given by the previous table. (In this case, our past experience with physical symmetry would also lead us to conclude that the die was not physically symmetrical in its shape or make-up. Such a die is said to be loaded. If the model with equal probability values is appropriate, we say that the die is fair.)

In talking about events (for some given experiment and sample space) we shall use the standard notations of set theory. If A and B are events, we use $A \cup B$ for the union of A and B, and $A \cap B$ for the intersection of A and B. To say that $A \cup B$ occurs is the same as to say that either A or B or both occur. To say that $A \cap B$ occurs is the same as to say that both A and B occur. We use S for the sample space, $\emptyset$ for the empty set, and $\bar{A}$ for the event which consists of those points in S which are not in A. If every outcome in A is also in B, we say that A is a subset of B, and we write $A \subseteq B$. Thus, in the case of the die, if A is $\{1,3,5\}$ and B is $\{1,2\}$, then

$$A \cup B = \{1,2,3,5\}$$
$$A \cap B = \{1\}$$
$$S = \{1,2,3,4,5,6\}$$
$$\bar{A} = \{2,4,6\}$$
$$\bar{B} = \{3,4,5,6\}$$

We always take $P(\emptyset) = 0$, and, as we have noted before, we always have $P(S) = 1$.

We now state some simple mathematical facts about probability spaces. Although they are simple, these facts are also very useful.

The first fact is that

$$P(A) = 1-P(\overline{A})$$

The second fact is that if $A \cap B = \emptyset$, then

$$P(A \cup B) = P(A) + P(B) .$$

(When $A \cap B = \emptyset$, we say that A and B are <u>disjoint</u>, or <u>mutually exclusive</u>, events.)

The third fact is that for any events A and B,

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

The truth of these facts is immediate and obvious when we recall that the probability of an event is the sum of the probability values of the basic outcomes in that event. For example if we take $P(A) + P(B)$, then each value for an outcome in $A \cap B$ gets counted twice. By subtracting $P(A \cap B)$, we get each value counted once, and we have the third fact above: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. (Note that the second fact is a special case of this third fact.)

We illustrate these facts with some examples from the experiment of rolling a single die. We take the probability space where each probability value is 1/6. Let $A = \{1,3,5\}$ and $B = \{1,2\}$. Then $\overline{A} = \{2,4,6\}$, and "$P(A) = 1 - P(\overline{A})$" asserts that $\frac{1}{2} = 1 - \frac{1}{2}$, which is true; and "$P(A \cup B) = P(A) + P(B) - P(A \cap B)$" asserts that $\frac{2}{3} = \frac{1}{2} + \frac{1}{3} - \frac{1}{6}$, which is also true. The three general facts given above are sometimes called <u>laws of probability</u>.

As we have seen, a probability space serves as a model for an experiment and is made up of a sample space and a probability function. What are some of the uses of such a model? One of the main uses is to calculate the probability of some event we are interested in. Then, provided that the model is a good model, we will know the approximate relative frequency with which we can expect that event to occur. (We sometimes say that we know how likely the event is.)

Here is an example of such a probability calculation. The experiment consists of stopping three people on the street in turn and asking each one the month of his or her birth. What is the probability that at least two of the three people stopped have the same birth-month? The sample space must have $12 \times 12 \times 12 = 1728$ basic outcomes. (Try listing all the possibilities to see why this is true.) The probability function is chosen so that each point gets the same probability value: $\frac{1}{1728}$. (Reasons for this choice will be given in Chapter 4.) Let A be the event: at least two of the three people have the same birth-month. To get our answer, $P(A)$, we must find out how many points of the sample space are in the event A. At first, this appears to be a long and complex problem, but the laws of probability come to our rescue as follows. $\overline{A}$ is the event: all three people have different birth-months. It is quite easy to get $P(\overline{A})$, because there are easily seen to be $(12)(11)(10) = 1320$ basic outcomes in $\overline{A}$. Thus $P(\overline{A})$ must be $\frac{1320}{1728} = \frac{55}{72}$. Using the law $P(A) = 1 - P(\overline{A})$, we get $P(A) = 1 - \frac{55}{72} = \frac{17}{72} = 0.24$ (to two decimal places), and we have our answer. We can expect the event A to occur about one-fourth of the time.

The same method can be used if we change the experiment and stop five people (instead of three) to ask for their birth-months. What is the probability, in this new experiment, of the event B that at least two of the five people have the same birth-month? We have $P(B) = 1 - P(\bar{B}) = 1 - \frac{(12)(11)(10)(9)(8)}{12^5} = 1 - \frac{55}{144} = \frac{89}{144} = 0.62.$ Thus we find, to our surprise, that we can expect the event B to occur over 60% of the time.

In most of the above examples, we have used the probability function which assigns the same probability value to each basic outcome. This assignment is called the equiprobable probability function. When we use this function (for reasons of physical symmetry, or for other reasons) the probability of an event A must be the same as the ratio of the number of points in A to the number of points in S. Calculating this ratio is not always easy, especially when the sample space has a large number of points. The probability laws are often helpful in such calculations.

There are also situations, as we shall see in Chapter 4, where we have only limited information about an experiment and do not know a probability function to begin with. Instead, we only know probabilities for some of the events. It is often possible in such cases to go on and calculate probabilities of other events by using laws of probability.

Remark. Perhaps the most important thing for a beginner in probability theory to remember, when he or she works on a problem, is that one must have in mind a single, definite experiment, a single, definite sample space, and a single, definite probability function (even though one may not be sure, ahead of time, of all the probability values.) Many of the difficulties that trouble a beginner in probability arise because one has confused two different experiments, two different sample spaces, or two different probability functions in the same problem. Before one starts on a problem, one should always ask oneself: what is the experiment, what is the sample space, and what is the probability function?

Perhaps the second most important thing for a beginner in probability theory to do is to keep separate in his or her mind the experiment (the physical procedure) from the model (the sample space and probability function) chosen to give a mathematical picture of the experiment. As we get more data about an experiment, we may decide to change our model. We illustrated this above in the case of a loaded die. Indeed, there may be no single obvious model to begin with when we study an experiment. (This is the case, for example, with tossing a thumb tack.) We shall later see that much of the subject of mathematical statistics has to do with the problem of choosing a best model, or else a set of good models, for a given experimental situation.

Usually, when we calculate a probability $P(A)$ for an event A in a given experiment, we do so from a particular model (probability space) that we have chosen for that experiment. If we used a different model, we might have to give a different value for that probability. Let μ and μ' be two different models for an experiment, and let A be an event for that experiment. We shall occasionally use notations such as "$P_\mu(A)$" or

"$P_{\mu'}(A)$" in which we indicate the model that is being used to give the probability of A.

In setting up a model for a given experiment, several different choices of sample space may be possible. Which we choose will depend upon mathematical convenience and upon the aspects of the experiment that are of interest to us. For example, if we roll a die on a table, we are ordinarily interested only in the number appearing, and in this case we take the sample space {1,2,3,4,5,6}. We might, however, be interested as well in the position upon the table at which the die comes to rest. If so, we would construct a more complex sample space (where each point corresponds to a particular position on the table together with a particular value on the die.) The sample space we choose is part of the _model_ we decide to use. When we use probability theory to study an experiment, we must choose an appropriate sample space (for our purposes) and then stay with that choice throughout the study.

Odds. Probability statements are sometimes given in the form of _true odds against an event_. If we say that the true odds against the event A are 3 to 2, and if we are talking about an experiment that can be repeated under the same general conditions, then we are stating that we expect the event, on the average, to occur two times for every three times that it fails to occur. In other words, we expect a relative frequency of 2/5, and we have in mind a model where P(A) = 2/5. True odds are usually given in terms of whole numbers. If the true odds against an event are m to n, we sometimes say that the true odds <u>in favor of</u> the event are n to m. In general, to say that the true odds against A are m to n is to say that $P(A) = \frac{n}{m+n}$. Thus, if $P(A) = p$ is given, the true odds against A can be stated as m to n, where m and n are whole numbers such that the value $\frac{n}{m+n}$ is equal, or approximately equal, to p.

In Chapter 1, we considered bets in which a bettor stands to lose the amount ℓ if a certain event does not occur and to win the amount w if the event does occur. Whole numbers m and n such that $\frac{m}{n} = \frac{w}{\ell}$ are called <u>betting odds</u> for this bet. In Chapter 1, we said that a bet is <u>fair</u> if $\frac{\ell}{w+\ell}$ = the limiting relative frequency for the event. Hence, for a probability model which accurately reflects observed relative frequencies, true odds against an event are just betting odds that give a fair bet on the event. Betting odds are sometimes used in connection with experiments that cannot be repeated ("the odds are 3 to 1 against there being life on Mars") in order to describe how strongly a person holds a certain belief. We later return (in Chapter 20) to the topic of non-repeatable experiments and degrees of belief.

<u>More on notation</u>. As is common in mathematics, we can use the symbol "$\in$" to mean "is a member of". Thus $x \in A$ means that x is a member of the set A. We can also use the symbol Σ to indicate a sum in the following way: given a model, $\sum_{x \in A} p_x$

will stand for the sum of the probability values for all outcomes in the event A. Thus the definition of P(A) can be given as

$$P(A) = \sum_{x \in A} p_x ,$$

and the requirement that probability values total to 1 can be given as

$$\sum_{x \in S} p_x = 1 .$$

Infinite sample spaces. In the examples above, all sample spaces have a finite number of points. Sample spaces with infinitely many points are also important and useful. For example, consider the experiment of tossing a coin until a head appears, and let the outcome be the number of tosses required. This number can be any positive integer. Hence, it is natural to take, as model, a probability space with infinitely many points. If the coin is well-balanced, and if the results of successive single tosses are "independent" (in a sense to be defined precisely in Chapter 4), then it can be shown from rules to be given in Chapter 4 that the probability of outcome x is $(1/2)^x$. Hence we have, as model, the probability space given in the following table:

x	1	2	3	4	5	6	...	n	...
p_x	1/2	1/4	1/8	1/16	1/32	1/64	...	$1/2^n$	...

Some events now have an infinite number of points, and calculation of their probabilities requires that we form corresponding infinite sums. For example,

$$P(x \text{ even}) = 1/4 + 1/16 + 1/64 + \ldots = 1/3$$

and $$P(x \text{ odd}) = 1/2 + 1/8 + 1/32 + \ldots = 2/3.$$

Such infinite sums are usually called infinite series, and ways for calculating such sums are given in the theory of infinite series. [Footnote: we do not give this theory here, but only make the technical remark, for those acquainted with the theory, that in an infinite probability space like the one above, the series for any event will be absolutely convergent and hence give a sum which does not depend upon the order of the terms in the series.] A probability space is discrete if it has a sample space in which every point, taken as an event by itself, has a positive probability. Finite probability spaces are discrete and so is the infinite probability space given in the example above.

A different kind of infinite probability space arises in experiments where, for purposes of a model, we take possible outcomes to be all points in some interval of real numbers (or in some continuous region of a higher dimensional space such as the plane). Such probability spaces are said to be continuous. (Older texts refer to such spaces as "geometrical".) For example, consider the experiment of monitoring telephone calls through a telephone exchange at a given time of day and measuring the length of time between the beginning of one call and the beginning of the next. It is natural to take, as a sample space, the infinite interval of all non-negative real numbers. Unfortunately, it is no longer mathematically possible to give a positive probability value to each point. (Hence the space is not discrete). Instead, we assign probability values to certain <u>events</u> by introducing a non-negative function $f(x)$, called the <u>probability density function</u>. We then calculate the probability of any event of the form $a \leqslant x \leqslant b$, for given a and b in this space, by taking

$$P(a \leqslant x \leqslant b) = \int_a^b f(x)\, dx.$$

Thus $P(a \leqslant x \leqslant b)$ is given by the area between $x = a$ and $x = b$ under the curve for $f(x)$.

In the above example of the telephone exchange, a good model (for most exchanges) is got by taking $f(x) = me^{-mx}$ where m is the average number of calls coming in per unit time. (Here, and throughout the book, $e = 2.718\ldots$, the base of the natural logarithms.) Thus for an exchange averaging two calls every minute, the probability that the time between two successive calls will be 5 minutes or less is given by

$$P(0 \leqslant x \leqslant 5) = \int_0^5 2e^{-2x}\, dx = 1 - e^{-10} = 0.99995.$$

See Figure 2.1.

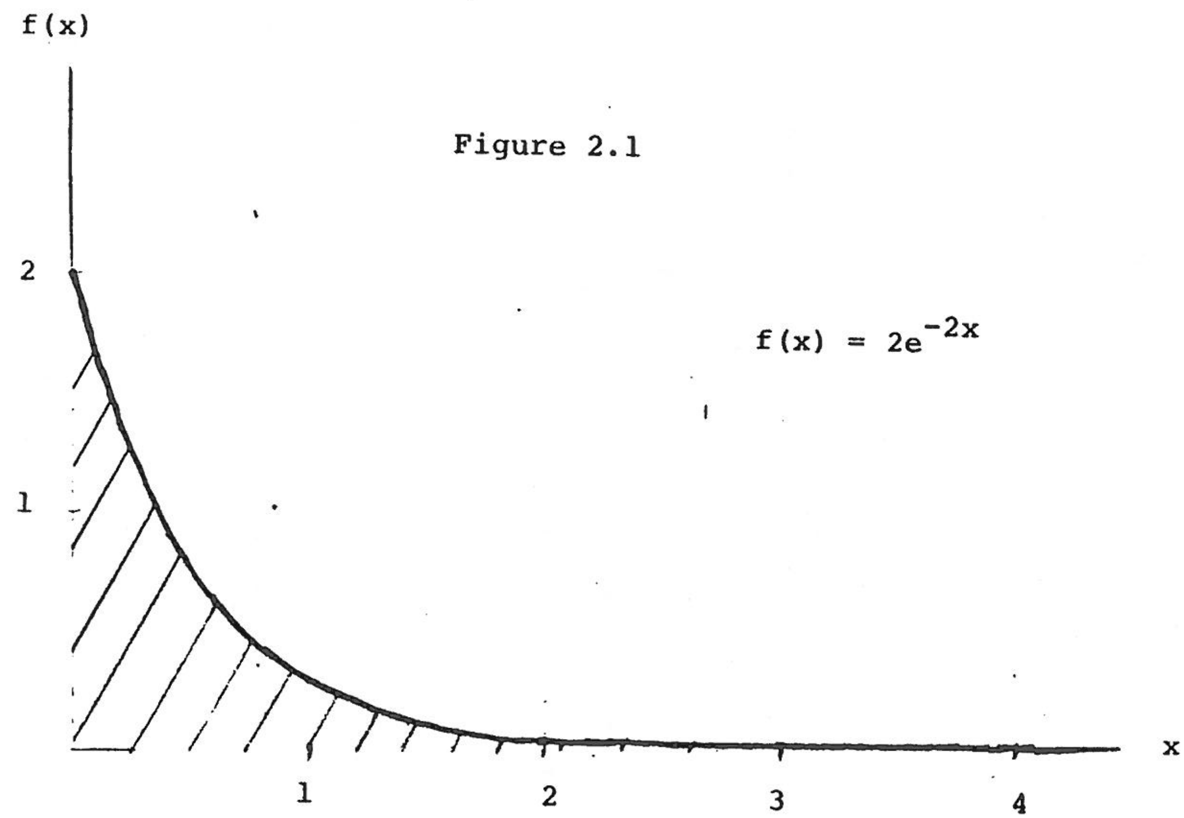


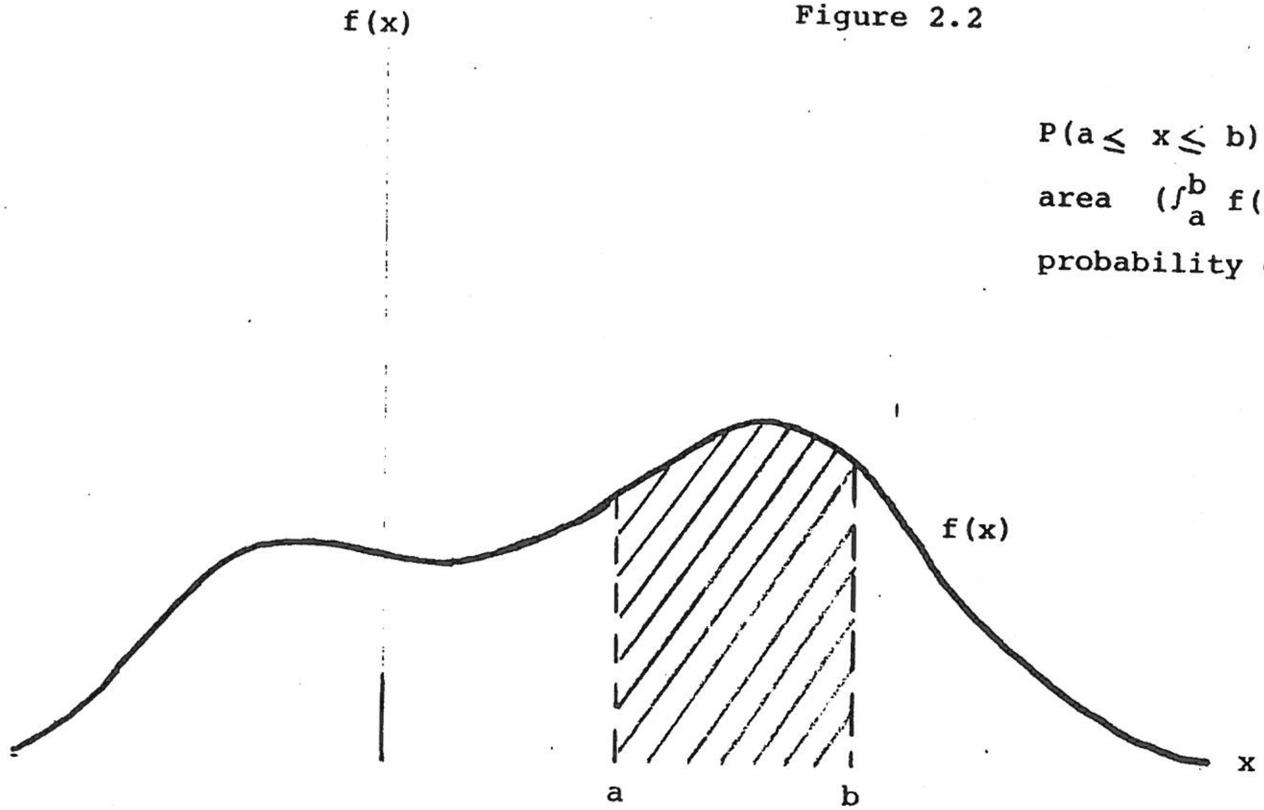
Figure 2.1

When a continuous probability space is given as an interval (possibly infinite) of the real line, we obtain probabilities by using a probability density function as in the above example. See Figure 2.2. The choice of this function (like the choice of the values p_x in a discrete space) will depend upon physical circumstances and past experience with the experiment. Clearly, we must have $f(x) \geq 0$ for all $x \in S$, and we must have $\int_S f(x)\, dx = 1$ (just as, in a discrete space, we must have $p_x \geq 0$ and $\sum_{x \in S} p_x = 1$.) (The reader can check, in the telephone example, that $\int_0^\infty me^{-mx}\, dx = 1$.)

In the present text, we shall use infinite discrete spaces from time to time. We shall use continuous spaces less often, but certain special examples of continuous spaces will be important in later sections on mathematical statistics. Note that in a continuous space, we cannot calculate the probability for every event, but only for events having the form of an interval (or for events which are unions of intervals). Note also the special and curious feature of a continuous space that an event which consists of a single point gets probability zero. (Thus every individual outcome has, as an event, probability zero.) This is related to the fact that actually performing such an experiment will involve a measurement of some kind, that this measurement can never be fully precise, and hence that the measurement can only determine a small interval rather than a single point.

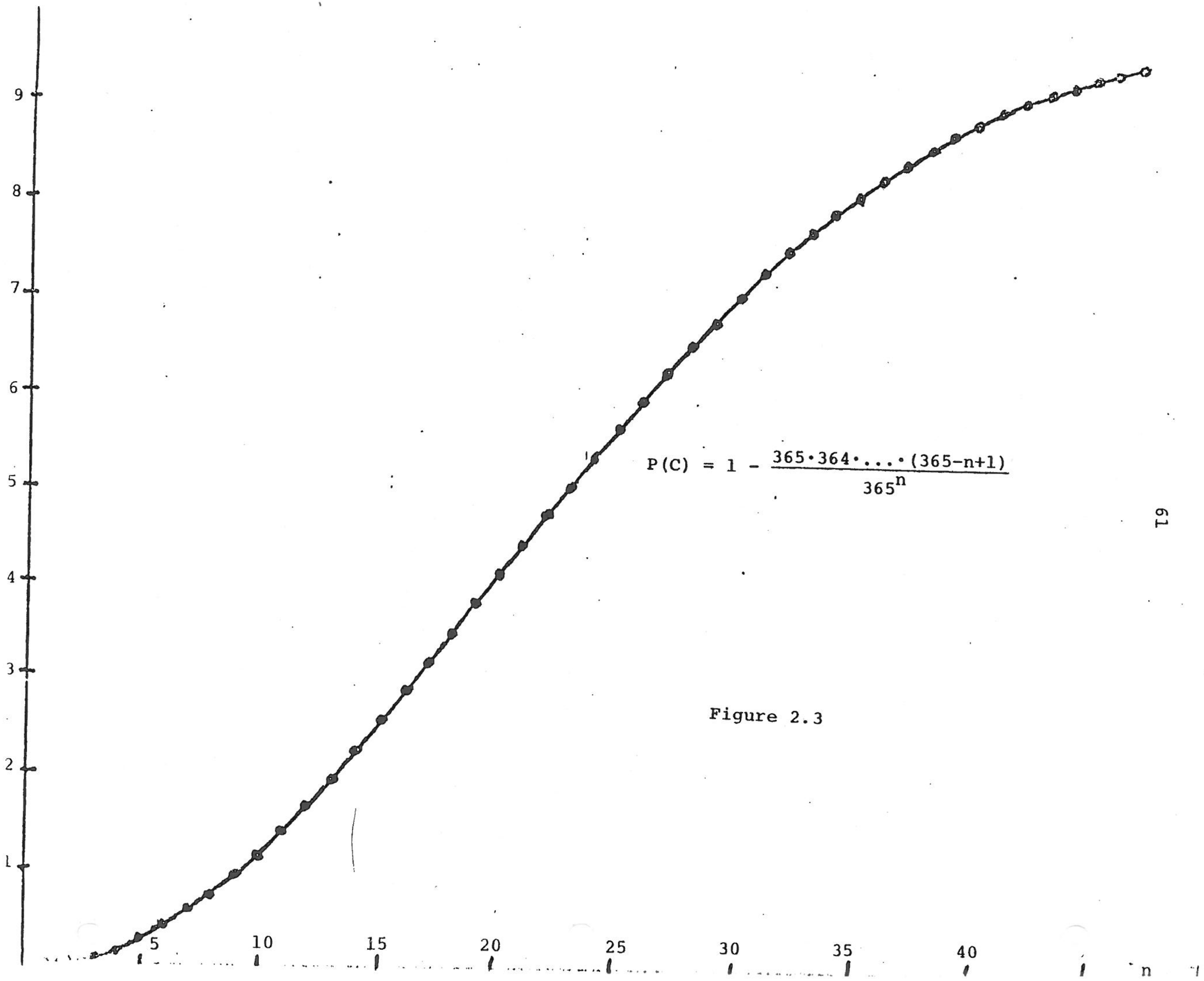
In more advanced work in probability, still other, more complex, forms of probability space can arise. For example, in studies of the path of a microscopically visible particle under random molecular

Figure 2.2



$P(a \leq x \leq b)$ given by shaded area $(\int_a^b f(x)\,dx)$, for probability density function $f(x)$.

(Note that total area under $f(x)$ must $= 1$.)



$$P(C) = 1 - \frac{365\cdot 364\cdot \ldots \cdot (365-n+1)}{365^{n}}$$

Figure 2.3

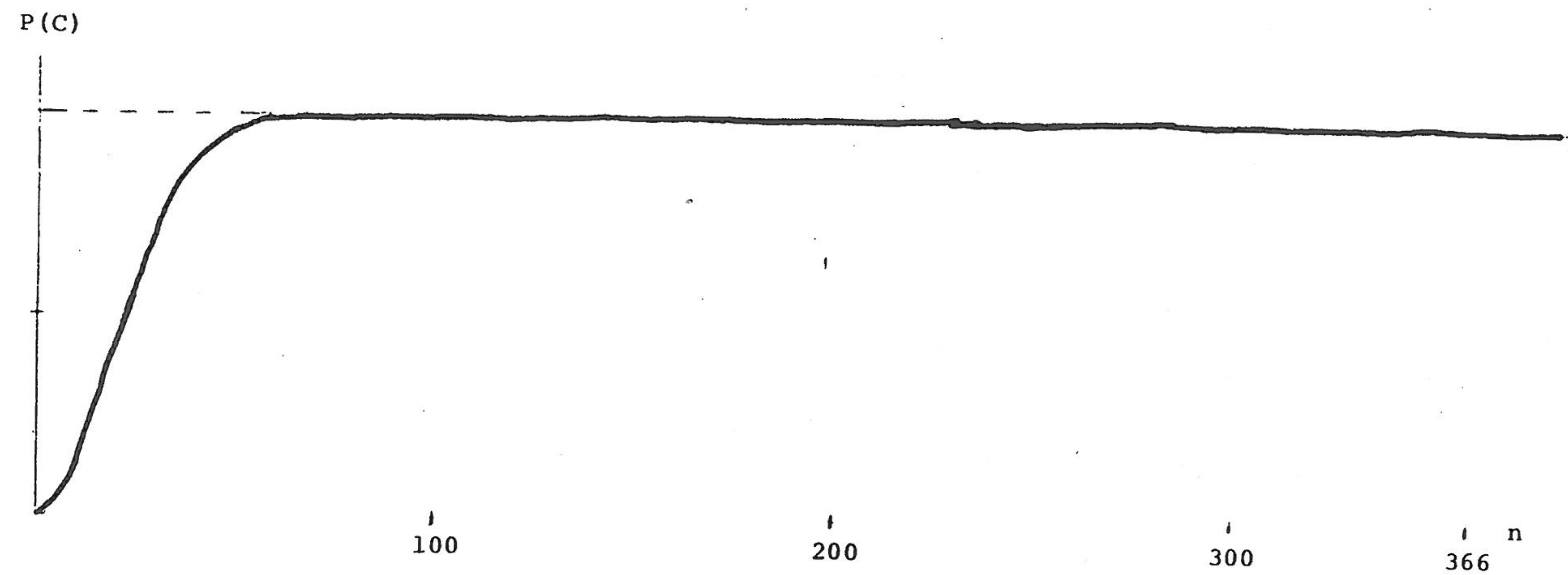


Figure 2.4

$$P(C) = 1 - \frac{365 \cdot 364 \quad \cdots \quad (365 - n+1)}{365^n}$$

EXERCISES FOR CHAPTER 2.

2-1. The game of roulette is described in Exercise 1-4. On a roulette wheel, the red pockets are numbered 1, 3, 5, 7, 9, 12, 14, 16, 18, 19, 21, 23, 25, 27, 30, 32, 34, 36. The event even is said to occur if one of the even numbers from 1 to 36 occurs. Similarly for odd. Use the equiprobable measure to find the following probabilities: $P(\{9\})$, $P(\text{red})$, $P(\text{black or 9})$, $P(\text{even})$, $P(\text{red and even})$, $P(\text{red or even})$, $P(\text{black})$, $P(\text{black and odd})$.

2-2. A sample space has four points: a, b, c, d. Let μ_1 be the probability space obtained by using the equiprobable measure, and let μ_2 be the probability space obtained by using a measure such that $P(\{a\}) = P(\{b,c,d\})$, $P(\{b\}) = P(\{c,d\})$, and $P(\{c\}) = P(\{d\})$. Find $P_{\mu_1}(\{a,b,c\})$ and $P_{\mu_2}(\{a,b,c\})$.

2-3. For certain events A and B in a certain probability space, you are informed that $P(A) = 1/2$ and $P(A\cap B) = 1/3$. What can you conclude about $P(A\cap\overline{B})$, about $P(\overline{A}\cap B)$, and about $P(B)$?

2-4. Find an expression for $P(A\cup B\cup C)$ in terms of $P(A)$, $P(B)$, $P(C)$, $P(A\cap B)$, $P(A\cap C)$, $P(B\cap C)$, and $P(A\cap B\cap C)$.

2-5. Four people are chosen and the day of the week on which each was born is determined.

(a) Using the equiprobable measure as in the examples in the text, find the probability that at least two of the four were born on the same day of the week.

(b) Do the same for chosen groups of 2, 3, 5, 6, 7, and 8 people, and make a figure similar to Figure 2.3.

(c) Find the probability that at least two of the four people have the same day but no three have the same day.

(d) Find the probability that at least three have the same day.

2-6. A hustler offers a bettor even money that in a group of 35 strangers at least two were born on the same day of the year. (The bettor loses if this event occurs.) Use Figure 2.3 to estimate true odds against the bettor winning. If $\ell = 1$, what should w be for the bet to be fair? (Give an approximate value.)

2-7. (a) A hustler offers a better even money that the bettor cannot roll at least one <u>six</u> in three rolls of a die. Use an appropriate probability space to verify the value of λ given on page 28. What are true odds against rolling at least one <u>six</u>?

(b) Verify the value of λ and find true odds for the sucker bet described in Exercise 1-6.

2-8. Use the probability space given in the text for the experiment of tossing a coin until a head appears.

(a) Verify the probability given in the text that the number x of tosses is even.

(b) Find the probability that x is divisible by 3.

(c) Find the probability that either x is even or divisible by 3 (or both).

(Hint. Use the fact that the sum of the series $a + ar + ar^2 + ar^3 + \ldots$ is $\frac{a}{1-r}$. Such a series is called a geometric series.)

2-9. A pair of dice is rolled until either a six appears or a seven appears. Use a probability space with the sample space $S = A \cup B$, where A consists of all outcomes of the form: six appears before seven and for the first time on roll x $(x = 1,2,3,\ldots)$; and B consists of all outcomes of the form: seven appears before six and for the first time on roll y $(y = 1,2,3,\ldots)$. In Chapter 4, we shall see that an appropriate probability value for each outcome in A is $\left(\frac{25}{36}\right)^{x-1} \frac{5}{36}$, and that an appropriate value for each outcome in B is $\left(\frac{25}{36}\right)^{y-1} \frac{6}{36}$. Use the hint in Exercise 2-8 to find $P(A)$ and $P(B)$ in this probability space, and to verify that $P(S) = 1$.

2-10. Let the sample space S consist of all real numbers in the interval $2 \leq x \leq 5$. We form a continuous probability space μ_1 by using the probability density function $f_1(x) = 1/3$. We then form a different probability space μ_2 by using the probability density function $f_2(x) = \frac{2x}{21}$.

(a) Verify that $P_{\mu_1}(S) = 1$ and $P_{\mu_2}(S) = 1$.

(b) Find $P_{\mu_1}(2 \leq x \leq 4)$ and $P_{\mu_2}(2 \leq x \leq 4)$.

2-11. At the busiest time of day, the telephone exchange in a certain small town handles an average of 3 new calls per minute.

(a) Use a continuous probability space of the form given in the text to find the probability that the time between the beginnings of two successive calls is less than one minute,

(b) What is the probability that the time is more than two minutes?