order to reach four successes, then $F(9;4, 2/3)$ would have been approximately .96. This in turn implies that the probability of ten or more trials required to reach four successes would have been .04, which is less than the criterion agreed upon. In this case, the psychologist would have decided to reject the hypothesis that $p = 2/3$. Either it is true that $p = 2/3$, and a rare event has occurred, or 2/3 is not the correct value of p. The psychologist has already decided to conclude the latter, should such a rare event occur.

Notice what has been assumed here. First of all, we have assumed that $p = 2/3$ for each and every one of the children tested, so that observation of each child represents a trial in a stationary Bernoulli process. Furthermore, we have assumed that the results for each child, or trial, are independent of those for any other child. Finally, we have assumed that the order of selection of the children was completely random. The failure of any of these assumptions to be true could, of course, make a difference in the final conclusions.

5.21 The Poisson distribution

Still another relative of the binomial family of distributions plays a large role in theoretical and applied statistics. This is the Poisson distribution, named after the nineteenth-century French mathematician S. Poisson. A random variable following this rule is referred to as a Poisson variable, and the process generating values of such a random variable is known as a Poisson process. The probability function for a Poisson variable X follows the rule

$$p(x;m) = \begin{cases} \dfrac{e^{-m}m^x}{x!}, & x = 0, 1, 2, 3, \cdots ; m > 0; \\ 0, & \text{otherwise}, \end{cases}$$

where e is the mathematical constant, and m is a constant known as the "intensity" of the Poisson process.

Notice that like a Pascal variable, a Poisson variable X can take on only integral or "whole" values—in this case from zero to an indefinitely large value. This random variable thus can assume any of a countably infinite set of values. It is, however, a discrete variable.

Although Poisson variables and the processes generating them can be given a variety of useful interpretations, perhaps the simplest approach to the study of the Poisson is to regard it as a special case of the binomial. A derivation of the Poisson function from the binomial will now be sketched out. What we are going to show is that if N is allowed to become extremely large while p is made extremely small, and if Np remains constant, the binomial distribution approaches the Poisson distribution. Take note, however, that this is just a sketch of the general outline of a proof. A number of the mathematical qualifications really necessary to such a proof will be omitted here.

Consider a binomial variable X, which takes on values with probabilities depending upon p, the probability of a success, and N, the number of trials. Let us define $m = Np$. Then we can rewrite the binomial probability that $X = x$ as

$$p(x;N, p) = \frac{N!}{x!(N-x)!} p^x(1-p)^{N-x},$$

$$p(x;N, m) = \frac{N!}{x!(N-x)!N^x} (m)^x \left(1 - \frac{m}{N}\right)^{N-x}$$

$$= \frac{(N)}{N} \frac{(N-1)}{N} \frac{(N-2)}{N} \cdots \frac{(N-x+1)}{N} \frac{m^x}{x!} \left(1 - \frac{m}{N}\right)^N \left(1 - \frac{m}{N}\right)^{-x}$$

Now we want to find the limiting value of this probability as N becomes indefinitely large and m remains constant. That is, we want to find

$$\underset{N \to \infty}{\text{limit}}\ p(x;N, m).$$

We can do so by examining what happens to the various terms in the expression above as N approaches the limit. First of all, as N is made indefinitely large, $1 - (1/N)$ approaches 1, $1 - (2/N)$ approaches 1, and so do all of the succeeding terms up through $1 - (x + 1)/N$. The expression $[1 - (m/N)]^{-x}$ must also approach 1. The only other term involving N is $[1 - (m/N)]^N$. In the advanced calculus it is shown that

$$\underset{N \to \infty}{\text{limit}} \left(1 - \frac{m}{N}\right)^N = e^{-m}.$$

Hence, making these substitutions, we have

$$\underset{N \to \infty}{\text{limit}}\ p(x;N, m) = \frac{e^{-m}m^x}{x!}.$$

For fixed $m = Np$, and as N becomes indefinitely large, the limiting value of the binomial probability $p(x;N, m)$ is the Poisson probability $(e^{-m}m^x)/x!$.

Because of this connection between binomial and Poisson probabilities, Poisson sampling can be thought of as the act of making a vast number of trials from a stable and independent Bernoulli process for which the probability of a success is extremely small. The constant m is called the "intensity" of the Poisson process.

Not only does this connection with the binomial distribution permit one interpretation of the Poisson distribution; there are practical consequences as well. In cases in which N is large and p is relatively small, the binomial probabilities may be very laborious to calculate. In this instance, it is a much simpler matter to approximate the exact binomial probabilities through use of Poisson probabilities for the various values of x.

When a Poisson probability for some specific x value is desired, one simply calculates $m = Np$, and then applies the Poisson rule

$$p(x;m) = \frac{e^{-m}m^x}{x!}$$

for the value $X = x$ of interest. Table X in Appendix C gives selected values of e^{-m}, and Table VIII gives various values of factorials. For example, suppose that for $N = 3000$ and $p = .001$, we wished to find the probability for $X = 5$. Then, $m = (3000)(.001) = 3$, and we calculate the probability by taking

$$p(10;3) = \frac{e^{-3}3^5}{5!}$$
$$= \frac{(.0498)(243)}{120}$$
$$= .1008.$$

This is an exact Poisson probability, $p(10;3)$. It is also an approximation of the binomial probability $p(10;\ 3000,\ .001)$. The Poisson probability will be approximately equal to the actual binomial probability only for very large N and very small p, of course. Nevertheless, the approximation is good enough to be useful even when N is only moderately large and p only relatively small. Table 5.21.1 contrasts binomial and Poisson probabilities for $N = 20$ and $p = .10$, and for $N = 30$, $p = .01$. In the first instance the Poisson probabilities approach but certainly do not equal the binomial values. In the second instance, the fit is somewhat better. The approximation would grow steadily better if we increased N and reduced p. Later we will have more examples of the use of the Poisson distribution to approximate the binomial distribution. However, first it will be of value to examine some other interpretations of a Poisson process.

A great many illustrations of Poisson processes occur in the physical and the biological sciences, as well as in everyday life. For example, the degeneration of a radioactive substance is regarded as a Poisson process. At any given instant the probability is very small that an alpha-particle will be emitted, while there are vast numbers of opportunities for such an event to occur. The distribution of bacteria on a Petri plate can be viewed as a Poisson process. Each tiny area on the plate can be viewed as a trial, and a bacterium may or may not occur on such an area. The probability of such an occurrence on any given area is very

Table 5.21.1

A COMPARISON OF BINOMIAL PROBABILITIES WITH POISSON PROBABILITIES WHERE $m = Np$

	$p = .10, N = 20$		$p = .01, N = 20$	
	Binomial	*Poisson*	*Binomial*	*Poisson*
r				
0	.1216	.1353	.7397	.7408
1	.2702	.2707	.2242	.2223
2	.2852	.2707	.0328	.0333
3	.1901	.1804	.0031	.0033
4	.0898	.0902	.0002	.0003
5	.0319	.0361	.0000	.0000
6	.0089	.0120		
7	.0020	.0034		
8	.0004	.0009		
9	.0001	.0002		
10 or more	.0000	.0000		

small indeed, but there are very many areas on such a plate. The distribution of misprints in a book can be studied as a Poisson process, as can the occurrence of accidents of a certain kind in a manufacturing plant.

The Poisson distribution is thus important in its own right, quite apart from its connection with the binomial distribution. A typical situation that is regarded as a Poisson process involves a continuum of time, which can be broken down into arbitrary small segments, or "instants." At any given instant an event or success may occur. The occurrence of a success at any instant is quite independent of the occurrence or nonoccurrence of a success on any prior or following instant. However, the successes occur at a given and constant rate, which is the intensity of the process. This rate is stated in terms of a time interval longer than an instant, although the interval may be of any size. Thus, the intensity may be stated in terms of the expected number of successes per minute, or the expected number of successes per hour, and so on. The random variable is then the actual number of successes in a minute, or the actual number of successes in an hour, or any other fixed time span of interest. This random variable, number of successes in a time interval, may take on any whole value from zero to an indefinitely large value.

As a simple example, imagine a checkout counter at a grocery store. The counter is open all the time, and a customer may or may not arrive at a given instant. Experience has shown that the customers tend to arrive at a rate of, say, 2 per minute. On the other hand, during any given minute, no customers may come, one may come, 500 may come, and so forth. Each possible number of customers that might arrive in a given minute has a probability, and this is given by the Poisson rule with $m = 2$, the intensity of the process.

Suppose that we wish to know the probability that 5 customers arrive at the counter in one minute. Then we must find

$$p(5;2) = \frac{e^{-2}2^5}{2!},$$

since $m = 2$, $X = 5$. with the aid of Table X of Appendix C we know that e^{-2} is about .135. The probability is then given by

$$\begin{aligned} p(5;2) &= \frac{(.135)(32)}{5!} \\ &= 4.32/120 \\ &= .036. \end{aligned}$$

Only in about 36 out of a thousand minutes should we expect exactly 5 customers to arrive at the checkout counter.

If we wished to know the probability of *five or fewer* customers, we would find

$$\begin{aligned} F(5;2) &= \frac{e^{-2}2^0}{0!} + \frac{e^{-2}2}{1!} + \frac{e^{-2}2^2}{2!} + \frac{e^{-2}2^3}{3!} + \frac{e^{-2}2^4}{4!} + \frac{e^{-2}2^5}{5!} \\ &= e^{-2}\left(1 + 2 + \frac{4}{2} + \frac{8}{6} + \frac{16}{24} + \frac{32}{120}\right) \end{aligned}$$

$$= (.135)(7.266)$$
$$= .983, \quad \text{approximately.}$$

It also follows that the probability of *six or more* customers is given by

$$1 - F(5;2) = 1 - .983 = .017, \quad \text{approximately.}$$

Poisson probabilities and cumulative probabilities can be determined relatively easily in the way just shown, given Table VIII for the values of factorials and Table X for values of e^{-m}. Many books on advanced statistics give extensive tables of Poisson probabilities, particularly when they are designed to be used in fields such as the physical sciences or industry. However, since our use of the Poisson will not be extensive, space will not be given to such tables here. Rather a method will be given for using another table to find cumulative probabilities that will be useful in solving problems involving Poisson variables.

Suppose that we wish to find the cumulative probability that a Poisson variable X takes on a value less than or equal to some specific value a. That is, we wish to find $F(a;m)$ for a Poisson variable, where the process has intensity m. We can proceed as follows: In Table IV, Appendix C, look in the column at the extreme left under the symbol "ν." Find a value as close as possible to the value of $2(a + 1)$. Next, in the corresponding row, find in the body of the table the two values that fall to either side of $2m$. Then the cumulative Poisson probability lies between the two Q values given at the top of the table, corresponding to the two values to either side of $2m$. Thus, for example, suppose that we wished to know the value of $F(6;3)$. Here $a = 6$ and $m = 3$. In Table IV we look along the left-hand column until we locate $\nu = 2(6 + 1)$ or 14. Then in the row corresponding to 14 we look in the body of the table and find that the two values to either side of $2m = 6$ are 5.62872 and 6.57063. The first lies in a column headed $Q = .975$ and the second in the column headed $Q = .950$. Then we know that

$$.950 \leq F(6;3) \leq .975.$$

(As it happens, the actual value here is .97.) If a somewhat more accurate approximation is desired, the method of linear interpolation may be used. We shall employ this approximate method for finding cumulative Poisson probabilities in the next section.

5.22 Solving problems through use of the Poisson distribution

First we will consider an example in which the Poisson approximation to the binomial distribution is employed. Then, we will consider another example in which the random variable itself is presumed to follow a Poisson rule.

Suppose that in a study of the effect of training upon the ability of students to solve a difficult abstract problem, a psychologist used a problem that ordinarily can be solved only by about one in fifty students of the given age group.

He chose a random sample of 300 students and gave the prior training to each student. After the training, each student attempted to solve the problem. Twelve students solved it correctly. The psychologist wishes to know if results this far from what one ordinarily expects without training (about 6 students correct) cast any doubt on the hypothesis that the training has no effect. In particular, he decides to adopt the following rule for deciding: If the probability is .05 or less of getting 12 or more successes, given that p is actually .02, he will reject the hypothesis that $p = .02$; otherwise, he will not reject the hypothesis. In other words, if 12 or more successes given $p = .02$ is a rather rare event, then the psychologist will feel that sufficient doubts are cast by his results on the hypothesis that $p = .02$ to lead him to reject that hypothesis for his group of trained students.

As you can readily see, this has all of the elements of a binomial problem, since there is a fixed number of trials, $N = 300$, a fixed probability of a success, $p = .02$ under the hypothesis, and an obtained number of actual successes, $X = 12$. However, since N is relatively large, and p is relatively small, we can apply the Poisson approximation to the binomial. What we want to find is the probability of 12 or more successes. This is the same as finding 1.00 minus the probability of 11 or fewer successes, or

$$p(X \geq 12;300, .02) = 1 - F(11;300, .02).$$

In Poisson terms, letting $m = Np$, or $m = 300(.02) = 6$, we need to find

$$p(X \geq 12;m = 6) = 1 - F(11;6).$$

Using Table IV of Appendix C as described above, we locate the entry $2(x + 1)$, or 24 in the column labeled ν, and then in that row locate the two values closest to $2m$, or 12. These are 10.8564 and 12.4011. The Q value for the first is .990 and that for the second is .975. Thus, we can say that $F(11;6)$ lies between .99 and .975. This means that $p(X \geq 12;6)$ lies between .01 and .025. Since this probability is less than .05, our decision rule says that we reject the hypothesis that the probability of a success $p = .02$. The rule for deciding *how* to decide indicates that a result as rare as that actually obtained, given $p = .02$, casts sufficient doubt on the hypothesis to lead to its rejection. The psychologist would assert he had obtained significant evidence that the training improved performance on the problem.

Consider now this example of a rather different type. In an industrial plant, long experience has shown that the rate of accidents per standard working month is 8. However, it was decided that a new safety program should be instituted, with each worker receiving intensive safety training; the program would be evaluated in the month following its completion. The question was whether or not the number of accidents in that month was sufficiently low to permit the conclusion that the safety program was having an effect.

During the month after the program was completed, the number of accidents turned out to be 4. Thus the question was: "Is the occurrence of 4 or fewer accidents sufficient evidence to permit the conclusion that the rate is no longer 8 per month?" The decision was made to reject the hypothesis that $m = 8$ if the cumulative probability $F(4;8)$ was .05 or less.

Again the probability will be evaluated by use of Table IV. The left-hand column gives us the entry $2(4 + 1) = 10$, and then in that row we find the two values closest to $2m$ or 16. These values are 15.9871 and 18.3070, corresponding to Q values of .10 and .05. Thus we can see that the probability of 4 or fewer successes (i.e., accidents), given that $m = 8$, is between .05 and .10. In fact, the value under $Q = .10$ is almost exactly equal to 16, so that in this instance we are safe in saying that $F(4;8) = .10$. This is not a small enough probability to permit us to reject the hypothesis that $m = 8$, and so we would conclude that the training program has not been effective enough to permit our saying that the accident rate has changed.

5.23 The multinomial distribution

The basic rationale underlying the binomial distribution can be generalized to situations with more than two event classes. This generalization is known as the "multinomial distribution," having the following rule:

Consider K classes, mutually exclusive and exhaustive, and with probabilities $p_1, p_2, \cdots, p_K$. If N observations are made independently and at random, then the probability that exactly n_1 will be of kind 1, n_2 of kind 2, $\cdots$, and n_K of kind K, where $n_1 + n_2 + \cdots + n_K = N$, is given by

$$\frac{N!}{n_1!\, n_2! \cdots n_K!} (p_1)^{n_1}(p_2)^{n_2} \cdots (p_K)^{n_K}.$$

Think once again of colored marbles mixed together in a box, where the following probability distribution holds:

Color	p
Black	.40
Red	.30
White	.20
Blue	.10
	1.00

Now suppose that 10 balls were drawn at random and with replacement. The sample shows 2 black, 3 red, 5 white, and 0 blue. What is the probability of a sample distribution such as this? On substituting into the multinomial rule, we have

$$\frac{10!}{(2!)(3!)(5!)(0!)} (.4)^2(.3)^3(.2)^5(.10)^0$$

$$= \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10}{(1 \cdot 2)(1 \cdot 2 \cdot 3)(1 \cdot 2 \cdot 3 \cdot 4 \cdot 5)(1)} (.4)^2(.3)^3(.2)^5$$

since $0!$ and $(.1)^0$ are both equal to 1. Working out this number, we find that .087