processes, such as Çinlar [1975], Cohen [1969], Feller [1968], Karlin [1968], and Khintchine [1969]. (Roughly speaking, a Markov process is a process whose future probabilistic evolution after any time \( t \) depends only on the state of the system at time \( t \), and is independent of the history of the system prior to time \( t \). It should be easy for the reader to satisfy himself that a birth-and-death process is a Markov process.) Hopefully, it is also apparent that the birth-and-death process has sufficient intuitive appeal so that useful insights can be gained from the preceding cursory summary. We shall return to the birth-and-death process in Chapter 3, where we shall develop queueing models by judiciously choosing the birth-and-death coefficients.

2.4. Probability-Generating Functions

Many of the random variables of interest in queueing theory assume only the integral values \( j = 0, 1, 2, \ldots \). Let \( K \) be a nonnegative integer-valued random variable with probability distribution \( \{ p_j \} \), where \( p_j = P \{ K = j \} \) \( (j = 0, 1, 2, \ldots) \). Consider now the power series \( g(z) \),

\[
g(z) = p_0 + p_1 z + p_2 z^2 + \cdots, \tag{4.1}
\]

where the probability \( p_j \) is the coefficient of \( z^j \) in the expansion (4.1). Since \( \{ p_j \} \) is a probability distribution, therefore \( g(1) = 1 \). In fact, \( g(z) \) is convergent for \( |z| < 1 \) [so that the function (4.1) is holomorphic at least on the unit disk—see any standard text in complex-variable or analytic-function theory]. Clearly, the right-hand side of (4.1) characterizes \( K \), since it displays the whole distribution \( \{ p_j \} \). And since the power-series representation of a function is unique, the distribution \( \{ p_j \} \) is completely and uniquely specified by the function \( g(z) \). The function \( g(z) \) is called the probability-generating function for the random variable \( K \). The variable \( z \) has no inherent significance, although, as we shall see, it is sometimes useful to give it a physical interpretation.

The notion of a generating function applies not only to probability distributions \( \{ p_j \} \), but to any sequence of real numbers. However, the generating function is a particularly powerful tool in the analysis of probability problems, and we shall restrict ourselves to probability-generating functions. The generating function transforms a discrete sequence of numbers (the probabilities) into a function of a dummy variable, much the same way the Laplace transform changes a function of a particular variable into another function of a different variable. As with all transform methods, the use of generating functions not only preserves the information while changing its form, but also presents the information in a form that often simplifies manipulations and provides insight. In this section we summarize some important facts about probability-generating functions. For a more complete treatment, see Feller [1968] and Neuts [1973].
As examples of probability-generating functions, we consider those for
the Bernoulli and Poisson distributions. A random variable $X$ has the
Bernoulli distribution if it has two possible realizations, say 0 and 1,
occuring with probabilities $P\{X=0\} = q$ and $P\{X=1\} = p$, where $p+q = 1$.
As discussed in Section 2.1, this scheme can be used to describe a coin
toss, with $X=1$ when a head appears and $X=0$ when a tail appears.
Referring to (4.1), we see that the Bernoulli variable $X$ has probability-
generating function

$$g(z) = q + pz. \quad (4.2)$$

(If this seems somewhat less than profound, be content with the promise
that this simple notion will prove extremely useful.)

In Section 2.2 we considered the random variable $N(t)$, defined as the
number of customers arriving in an interval of length $t$, and we showed
that with appropriate assumptions this random variable is described by the
Poisson distribution

$$P\{N(t)=j\} = \frac{\lambda^j}{j!} e^{-\lambda t} \quad (j = 0, 1, \ldots).$$

If we denote the probability-generating function of $N(t)$ by $g(z)$, then

$$g(z) = \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} e^{-\lambda t} z^j = e^{-\lambda t} \sum_{j=0}^{\infty} \frac{(\lambda tz)^j}{j!},$$

and this reduces to

$$g(z) = e^{-\lambda t(1-z)}. \quad (4.3)$$

Having defined the notion of probability-generating function and given
two important examples, we now discuss the special properties of such
functions that make the concept useful. Since the probability-generating
function $g(z)$ of a random variable $K$ contains the distribution \{\(p_j\)}
implicitly, it therefore contains the information specifying the moments
of the distribution \{\(p_j\)}. Consider the mean $E(K)$,

$$E(K) = \sum_{j=1}^{\infty} j p_j. \quad (4.4)$$

It is easy to see that (4.4) can be obtained formally by evaluating the
derivative \((d/dz)g(z)=\sum_{j=1}^{\infty} j p_j z^{j-1}\) at $z=1$:

$$E(K) = g'(1). \quad (4.5)$$
From (4.2) and (4.5) we see that the Bernoulli random variable \( X \) has mean

\[
E(X) = p, \quad (4.6)
\]

and from (4.3) and (4.5) we see that the Poisson random variable \( N(t) \) has mean

\[
E(N(t)) = \lambda t. \quad (4.7)
\]

Similarly, it is easy (and we leave it as an exercise) to show that the variance \( V(K) \) can be obtained from the probability-generating function as

\[
V(K) = g''(1) + g'(1) - \left[ g'(1) \right]^2. \quad (4.8)
\]

This formula gives for the Bernoulli variable

\[
V(X) = pq \quad (4.9)
\]

and for the Poisson variable

\[
V(N(t)) = \lambda t. \quad (4.10)
\]

Sometimes, as we shall see, it is easier to obtain the probability-generating function than it is to obtain the whole distribution directly. In such cases, the formulas (4.5) and (4.8) often provide the easiest way of obtaining the mean and variance. The higher moments can also be calculated in a similar manner from the probability-generating function, but the complexity of the formulas increases rapidly. If primary interest is in the moments of a distribution rather than the individual probabilities, it is often convenient to work directly with the moment-generating function or the related characteristic function, which generate moments as (4.1) generates probabilities. These topics are covered in most texts on mathematical statistics, such as Fisz [1963].

Another important use of probability-generating functions is in the analysis of problems concerning sums of independent random variables. Suppose \( K = K_1 + K_2 \), where \( K_1 \) and \( K_2 \) are independent, nonnegative, integer-valued random variables. Then \( P \{ K = k \} \ (k = 0, 1, 2, \ldots) \) is given by the convolution

\[
P \{ K = k \} = \sum_{j=0}^{k} P \{ K_1 = j \} P \{ K_2 = k - j \}. \quad (4.11)
\]

Let \( K_1 \) and \( K_2 \) have generating functions \( g_1(z) \) and \( g_2(z) \), respectively:

\[
g_1(z) = \sum_{j=0}^{\infty} P \{ K_1 = j \} z^j \quad \ldots
\]
and
\[ g_2(z) = \sum_{j=0}^{\infty} P\{K_2 = j\} z^j. \]

Then term-by-term multiplication shows that the product \( g_1(z)g_2(z) \) is given by
\[ g_1(z)g_2(z) = \sum_{k=0}^{\infty} \left[ \sum_{j=0}^{k} P\{K_1 = j\} P\{K_2 = k-j\} \right] z^k. \quad (4.12) \]

If \( K \) has generating function \( g(z) = \sum_{k=0}^{\infty} P\{K = k\} z^k \), then (4.11) and (4.12) show that
\[ g(z) = g_1(z)g_2(z). \quad (4.13) \]

Thus we have the important result: The generating function of a sum of mutually independent random variables is equal to the product of their respective generating functions.

For a more concise proof of (4.13), observe that for any random variable \( K \) with probability-generating function \( g(z) \), the random variable \( z^K \) has expected value \( E(z^K) \) given by
\[ E(z^K) = g(z). \quad (4.14) \]

Then if \( K = K_1 + K_2 \), it follows that \( E(z^K) = E(z^{K_1+K_2}) = E(z^{K_1}z^{K_2}) \). If \( K_1 \) and \( K_2 \) are independent, then the random variables \( z^{K_1} \) and \( z^{K_2} \) are also independent, and thus \( E(z^{K_1}z^{K_2}) = E(z^{K_1})E(z^{K_2}) \), from which (4.13) follows.

Consider again the coin-tossing experiment described in Section 2.1. A coin is tossed \( n \) times, with \( X_j = 1 \) if a head appears and \( X_j = 0 \) if a tail appears on the \( j \)th toss. We wish to determine the probability that \( k \) heads appear in \( n \) tosses. Let
\[ S_n = X_1 + \cdots + X_n, \quad (4.15) \]
so that the value of the random variable \( S_n \) is the number of heads appearing in \( n \) tosses. \( S_n \) is the sum of \( n \) mutually independent, identically distributed Bernoulli variables \( \{X_j\} \), each with generating function \( g(z) \) given by (4.2). Therefore, \( S_n \) has generating function \( [g(z)]^n \); hence
\[ (q + pz)^n = \sum_{k=0}^{\infty} P\{S_n = k\} z^k. \quad (4.16) \]
Expanding the left-hand side of (4.16) by the binomial theorem, we obtain

\[(q + pz)^n = \sum_{k=0}^{n} \binom{n}{k} (pz)^k q^{n-k}.\]  

(4.17)

Equating coefficients of \(z^k\) in (4.16) and (4.17) yields

\[P\{S_n = k\} = \begin{cases} \binom{n}{k} p^k q^{n-k} & (k = 0, 1, \ldots, n), \\ 0 & (k > n), \end{cases}\]

(4.18)

which is, of course, the binomial distribution.

The result (4.18) could have been obtained by direct probabilistic reasoning. (Any particular sequence of \(k\) heads and \(n-k\) tails has probability \(p^k q^{n-k}\), and there are \(\binom{n}{k}\) such sequences.) In this example, the direct probabilistic reasoning concerning the possible outcomes of \(n\) tosses is replaced by the simpler probabilistic reasoning concerning the possible outcomes of one toss and the observation (4.15) that leads to the use of probability-generating functions. In a sense, probabilistic or intuitive reasoning has been traded for more formal mathematical manipulation. In the present case both approaches are simple, but this is not always true. Roughly speaking, the difficulties encountered in the generating-function approach, when applicable, are not extremely sensitive to the underlying probabilistic structure. Thus the use of generating functions tends to simplify hard problems and complicate easy ones.

Another example is provided by the Poisson distribution. Let \(N = N_1 + N_2\), where \(N_i\) is a Poisson random variable with mean \(\lambda_i\); that is, \(P\{N_i = j\} = (\lambda_i/j!) e^{-\lambda_i}\). Then \(N\) has probability-generating function

\[g(z) = e^{-\lambda_1(1-z)} e^{-\lambda_2(1-z)} = e^{-(\lambda_1 + \lambda_2)(1-z)}.\]  

(4.19)

Equation (4.19) shows (with no effort) that the sum of two independent Poisson variables with means \(\lambda_1\) and \(\lambda_2\) is itself a Poisson variable with mean \(\lambda = \lambda_1 + \lambda_2\).

Exercise

4. **Compound distributions.** Let \(X_1, X_2, \ldots\) be a sequence of independent, identically distributed, nonnegative, integer-valued random variables with probability-generating function \(f(z)\); and let \(N\) be a nonnegative, integer-valued random variable, independent of \(X_1, X_2, \ldots\), with probability-generating function \(g(z)\). Let \(S_N\) denote the sum of a random number of random variables: \(S_N = X_1 + \cdots + X_N\). (In general, the distribution of the sum of a random number of independent random variables is called a compound distribution.)

**a.** Show that \(S_N\) has probability-generating function \(g(f(z))\).

**b.** Show that \(E(S_N) = E(N) E(X)\) and \(V(S_N) = E(N) V(X) + V(N) E(X)^2\).
Probability-Generating Functions

It was mentioned previously that the variable $z$ in the generating function has no inherent significance, but that it is sometimes useful to give it a probabilistic interpretation. This may allow us to obtain the generating function directly from probabilistic considerations and, in so doing, to replace tedious arguments by elegant ones. This probabilistic interpretation, suggested by D. van Dantzig and called by him the method of collective marks, is: Imagine that the random variable whose probability-generating function we seek has been observed to have realization $K = j$, say. Now suppose we perform $j$ Bernoulli trials, where at each trial we generate a “mark” with probability $1 - z$ (and no mark with probability $z$). Then $z^j$ is the probability that none of the $j$ trials resulted in a mark. Hence, $\sum_{j=0}^{\infty} P(K=j) z^j = g(z)$ is the probability that no marks are generated by the realization of $K$.

For example, consider Exercise 4a, where we wish to determine the probability-generating function of the sum $S_N$ of a random number of random variables. The probability that the realization of $X_1$, say, generates no marks is equal to its probability-generating function $f(z)$. Thus, the probability that the realization of $X_1 + \cdots + X_N$ generates no marks is $[f(z)]^N$. Hence, the probability that the realization of $S_N$ generates no marks is $\sum_{n=0}^{\infty} P(N=n) [f(z)]^n = g(f(z))$.

We shall appeal to the method of collective marks to eliminate messy calculation at various points in the text. For further discussion see Neuts [1973] and Runnenburg [1965] (as well as the accompanying discussion of Runnenburg's paper by van der Vaart).

Exercise

5. Let $N_1$ and $N_2$ be independent, identically distributed, nonnegative integer-valued random variables. We consider two different procedures for distributing $N_1 + N_2$ balls into two cells. In procedure (a), we place $N_1$ balls in the first cell and $N_2$ balls in the second cell. In procedure (b), we consider the $N_1 + N_2 = N$ balls together, and independently place each of the $N$ balls in a cell, with each ball having probability $\frac{1}{2}$ of being placed in the first cell and probability $\frac{1}{2}$ of being placed in the second cell. Define the generating functions

$$g(z) = \sum_{j=0}^{\infty} P(N_r = j) z^j \quad (r = 1, 2),$$

and show that if procedures (a) and (b) are equivalent, then

$$g(x)g(y) = g\left(\frac{x+y}{2}\right).$$