chapter 3

The Arrival Process

Chapter 2 describes the first step in evaluating a queueing system: observation and measurement. Observation is a key part of the formulation stage of systems analysis. It answers the questions: What is the unknown? What are the data? What is the condition? The second stage of systems analysis is modeling. Whereas observation determines what happened, modeling *explains* what happened. It should reveal the underlying process that created the data. Thomas Kuhn, writing on the discovery of oxygen, states:

> Though undoubtedly correct, the sentence, "Oxygen was discovered," misleads by suggesting that discovering something is a single simple act assimilable to our usual (and also questionable) concept of seeing . . . discovering a new sort of phenomenon is necessarily a complex event, one which involves recognizing *that* something is and *what* it is.[1]

Recognizing *what* it is—the act of explanation—is the role of modeling.

Modeling is the subject of this chapter, as well as Chaps. 4 to 6. This chapter concentrates on the customer arrival process, with emphasis on a particular model known as the Poisson process. As early as 1910, with Erlang's paper "The Theory of Probabilities and Telephone Conversations," the Poisson process had been used as a model for customer arrivals. However, it was developed much earlier, in the early nineteenth century, by the French mathematician Denis Poisson. In addition to being a good representation of arrival processes, it turns out that the Poisson process is also a good representation of many physical systems, such as the position of molecules within gases. (See Haight 1967 for more examples.) The Poisson process is a rare example of a model that fulfills the dual objectives of realism and simplicity.

This chapter begins with a description of the conditions that create the Poisson process. Next, it covers the relationship between the Poisson process and the Poisson probability distribution. It then discusses the properties of the Poisson process. The chapter concludes with an examination of some of the ways to determine whether or not an arrival process is Poisson. One should understand random variables, probability distributions, and basic statistics before beginning this chapter.

3.1 CREATION OF THE POISSON PROCESS

The Poisson process is an example of a broader class of stochastic (that is, random) processes known as counting processes. Suppose that events occur at various times, in some random fashion. A ***counting process*** is the function representing the cumulative number of events that have occurred up to any point in time. We have already seen three examples of counting processes—$A(t)$, $D_q(t)$, and $D_s(t)$.

The Poisson process is a type of counting process that applies to customer arrivals, $A(t)$. The definition of the Poisson process will depend on the following:

Definitions 3.1

A counting process has **independent increments** if the numbers of events in any pair of disjoint time intervals are statistically independent.

A counting process has **stationary increments** if the distribution of the number of events in any time interval depends only on the length of the time interval. It does not depend on when the interval occurred.

Formally, the Poisson process is defined as follows:

Definitions 3.2

A counting process $N(t)$ is a *Poisson process* with rate λ if

A. The process has independent increments.

B. The process has stationary increments. And

C.

$$Pr\{[N(t+dt)-N(t)]\begin{Bmatrix} =0 \\ =1 \\ >1 \end{Bmatrix} \begin{matrix} =1-\lambda dt \\ =\lambda dt \\ =0 \end{matrix}$$

where dt is a differential (that is, very small) sized time interval.

The rate λ represents the *expected* number of customers to arrive per unit time. If $\lambda = 10$ customers per hour, then the expected number of customers to arrive in a 60-minute period is 10 and the expected number to arrive in a 30-minute period is 5. The number of customers that actually arrive in any 60-minute period can be very different from 10, and the spacing between customer arrivals does not have to be constant. The significance of condition C is that customers arrive one at a time, and the probability that an arrival occurs in a very short (differential length) time interval equals the arrival rate multiplied by the size of the interval. For example, if $\lambda = 10$ customers per hour, then the probability of one customer arriving within any 1-second interval is approximately $10/3600 = .00278$.

In words, the Poisson process can be summarized as follows:

A. The probability that a customer arrives at any time does not depend on when other customers arrived.

B. The probability that a customer arrives at any time does not depend on the time.

C. Customers arrive one at a time.

The word *description* suggests why the Poisson process is important. Consider a queue for tellers at a bank. Is there any reason to believe that past arrivals influence future arrivals? Does the probability of an arrival vary over time? Do customers arrive in groups? For most customers, the answer to the first and third questions is almost certainly no. Customers do tend to arrive independently of each other, one at a time. There may be exceptions—two people who decide to go to the bank together—but that is what they are, exceptions. The model does not have to be perfect to be useful.

With regard to the second question, the answer is more of a maybe. The arrivals may be fairly constant over short time intervals, but most likely vary over the day and week. For the moment, think of the standard Poisson process as representing what happens over time intervals when the arrival rate is nearly constant. Later on, in Chap. 6, a nonstationary version of the Poisson process is presented that accounts for this variation.

The Poisson process is sometimes viewed in terms of its ***calling population*** the group of *potential* customers. If the calling population is large, customers arrive independently, and the probability that any particular customer arrives during any small time interval is small and constant, then the arrival process will be Poisson. As the size of the calling population declines, the arrival process will look less and less like a Poisson process, in which case an alternative model may be called for (as is discussed in Chap. 5).

3.2 POISSON DISTRIBUTION

The Poisson distribution is a probability distribution that arises from the Poisson process. It is a *discrete distribution*, meaning that its random variable is limited to a set of distinct values. A Poisson random variable is limited to the set of non-negative integers. The Poisson distribution gets its name because it is the probability distribution for the number of arrivals within any time period of a Poisson process.

Derivation. Consider the random variable $N(t)$, representing the number of arrivals over the interval $[0,t]$. Suppose that the interval $[0,t]$ is divided into I segments, where I is a very large number. Consider just one of these segments. From condition C of

Definition 3.2 of the Poisson process, either one customer arrives during the segment or no customers arrive, but never more than one. The probability that exactly one customer arrives is $\lambda\ (t/I)$. The total number of arrivals over the entire interval equals the sum of I Bernoulli (0,1) random variables, each representing whether or not a customer arrived during a segment. Thus $N(t)$ is a binomial random variable with parameters I and $\lambda(t/I)$:

$$P[N(t) = n] = \lim_{I\to\infty} \binom{I}{n} p^n (1 - p)^{I-n} \qquad n = 0, 1, \ldots, \mathrm{I} \tag{3.1}$$

where $P[\ \ ]$ denotes the probability of an event, and $p = \lambda t/I$.

Through a series of calculations, Eq. 3.1 can be reduced to a very simple expression. First, by substituting $\lambda t/I$ for p, and expanding the permutation, we can rewrite Eq. 3.3 as the following:

$$P[N(t) = n] = \lim_{I\to\infty} \left[\frac{I(I - 1)\cdot \ldots \cdot (I - n + 1)}{I^n}\right] \left[\frac{(\lambda t)^n}{n!}\right]\left[\frac{(1 - \lambda t/I)^I}{(1 - \lambda t/I)^n}\right] \tag{3.2}$$

In the limit, as I approaches infinity, the numerator of the first term equals I^n. Thus, the first term approaches one. Also in the limit, the denominator of the third term equals 1^n, or just one. This leaves the following:

$$P[N(t) = n] = \lim_{I\to\infty} \frac{(\lambda t)^n}{n!} (1 - \lambda t/I)^I \tag{3.3}$$

From the definition of the number e, the last term in Eq. 3.3 is $e^{-\lambda t}$. Making this substitution allows Eq. 3.1 to be expressed in final form:

$$P[N(t) = n] = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \qquad n = 0, 1, \ldots \tag{3.4}$$

For a *discrete* random variable, the ***probability function***, $f(x)$, specifies the probability that the random variable *equals* a set value x. Equation 3.4 is the probability function for a Poisson random variable with mean λt. Hence, *the probability distribution for the number of events in any time interval is Poisson.*

Definition 3.3: Poisson Probability Function, Mean λt

$$f(x) = \frac{(\lambda t)^x}{x!} e^{-\lambda t} \qquad x = 0, 1, 2, \ldots \tag{3.5}$$

The ***probability distribution function***, $F(x)$, specifies the probability that a random variable occurs at or below a set value x.

Definition 3.4: Poisson Probability Distribution Function, Mean λt

$$F(x) = \sum_{n=0}^{x} f(n) = \sum_{n=0}^{x} \frac{(\lambda t)^n}{n!} e^{-\lambda t} \qquad x = 0, 1, 2, \ldots \tag{3.6}$$

Example

From past experience, customers are known to arrive at a service station at the rate of $\lambda = 15$ customers/hour. The owner would like to know the probability that more than one customer will arrive during an employee's 5-minute coffee break.

Solution The expected number of customers to arrive during 5 minutes is λt, or 15 customers/hour $\times$ (5/60) hours = 1.25 customers. The probabilities of zero customers arriving and 1 customer arriving are calculated from Eq.(3.4):

$$f(0) = \frac{1.25^0}{0!} e^{-1.25} = e^{-1.25} = .29$$

$$f(1) = \frac{1.25^1}{1!} e^{-1.25} = 1.25\, e^{-1.25} = .36$$

The probability that one or less customers arrive, $F(1)$, equals $f(0) + f(1) = .65$. Therefore, the probability that more than one arrives equals $1 - F(1) = .35$.

Given that the number of arrivals in a time interval has a Poisson distribution, the Poisson process can be defined in the following alternative form:

Definition 3.5

The *Poisson process* is a counting process with the properties:

A. The process has independent increments.

B. Number of events in any time interval of length t has a Poisson distribution with mean λt.

Condition B substitutes for both conditions B and C in Definition 3.2. Therefore, it implies stationarity.

3.3 PROPERTIES OF THE POISSON PROCESS

The Poisson process has a number of important properties. So far, we have seen the following:

Property 1. The distribution for the number of events in any time interval of length t has the Poisson distribution with mean λt.

Here are some of the important characteristics of the Poisson distribution:

Characteristics of Poisson Distribution

1. The Poisson distribution is discrete and defined over the set of non-negative integers.
2. $E[N(t)] = \lambda t$
3. $V[N(t)] = \lambda t$
4. $P[N(t) = 0] = e^{-\lambda t}$

$E[\ ]$ represents the *expectation* of the enclosed random variable, and $V[\ ]$ represents the *variance* of the enclosed random variable.

The Poisson distribution has the distinguishing characteristic that its variance equals its mean (thus, its standard deviation equals the square root of its mean).

Example

Residents of a small city are known to place telephone calls by a Poisson process with rate 1000 per hour. The expected number of phone calls made after a time t is $1000t$, and the standard deviation of the number of phone calls made after a time t is $\sqrt{1000t}$. The *coefficient of variation* (ratio of standard deviation to mean) of the number of calls is $\sqrt{1000t}/1000t = 1/\sqrt{1000t}$, which declines as t increases.

3.3.1 Interarrival Time

The Poisson process is closely related to the exponential probability distribution. The exponential distribution is a ***continuous distribution***, meaning that exponential random variables are not limited to a set of discrete values. Instead of having a probability function, a continuous random variable has a ***probability density function***. The density function, $f(x)$, does not equal the probability that the random variable equals x. Rather, $f(x)$, multiplied by the differential dx, equals the probability that the random variable is contained in an interval of width dx centered around the point x.

In the following characteristics of the exponential distribution, X represents the outcome of a random variable.

Characteristics of the Exponential Distribution

1. The exponential distribution is continuous and defined over the set of non-negative real numbers.
2. The probability density function, $f(x)$, is

$$f(x) = \lambda e^{-\lambda x} \qquad x \geqslant 0$$

3. The probability distribution function, $F(x)$, is

$$F(x) = \int_0^x \lambda e^{-\lambda t}\, dt = 1 - e^{-\lambda x} \qquad x \geqslant 0$$

4. $E(X) = 1/\lambda$
5. $V(X) = 1/\lambda^2 = E^2(X)$

Example

Calculate the probability that an exponential random variable, mean .5, is contained in [.95, 1.05].

Approximate Solution $\lambda = 1/.5 = 2$, and $dx = 1.05 - .95 = .1$. The desired probability is approximately $f(1)dx = 2e^{-2\times 1}(.1) = .027$.

Exact Solution The approximate result is confirmed from the probability distribution function. The probability that $X \leq 1.05$ equals $F(1.05) = .878$ and the probability that $X \leq .95$ equals $F(.95) = .850$. The difference rounds off to .027.

The exponential distribution function is a special case of the gamma distribution, which is discussed in greater detail later. In Fig. 3.1b, the exponential distribution is the gamma distribution with a coefficient of variation of 1. Unlike the Poisson distribution function (Fig. 3.1d), the exponential distribution function is smooth, reflecting the fact that an exponential random variable is not restricted to a discrete set of values. Put another way, time (the exponential variable) is continuous, but counts (the Poisson variable) are discrete.

The relationship between the Poisson process and the exponential distribution is revealed by the third characteristic of the exponential distribution:

$$P(X \geq t) = 1 - F(t) = e^{-\lambda t} = P[N(t) = 0] \tag{3.7}$$

Thus, the probability that the random variable X is greater than or equal to some value t is identical to the probability that no events occur over an interval of length t. It should be apparent that these are two ways of stating the same thing. Furthermore, the independent increments property guarantees that interarrival times are mutually independent. Hence, we have another property of the Poisson process:

Property 2. The interarrival times for a Poisson process with rate λ are independent exponential random variables with mean $1/\lambda$.

Example

Customers are known to arrive at a medical clinic at the rate of 8 per hour. The receptionist is called away from his desk at 10:00, immediately after a customer arrived. How long can he stay away from his desk if he is willing to take a 50 percent chance of being away when the next customer arrives?

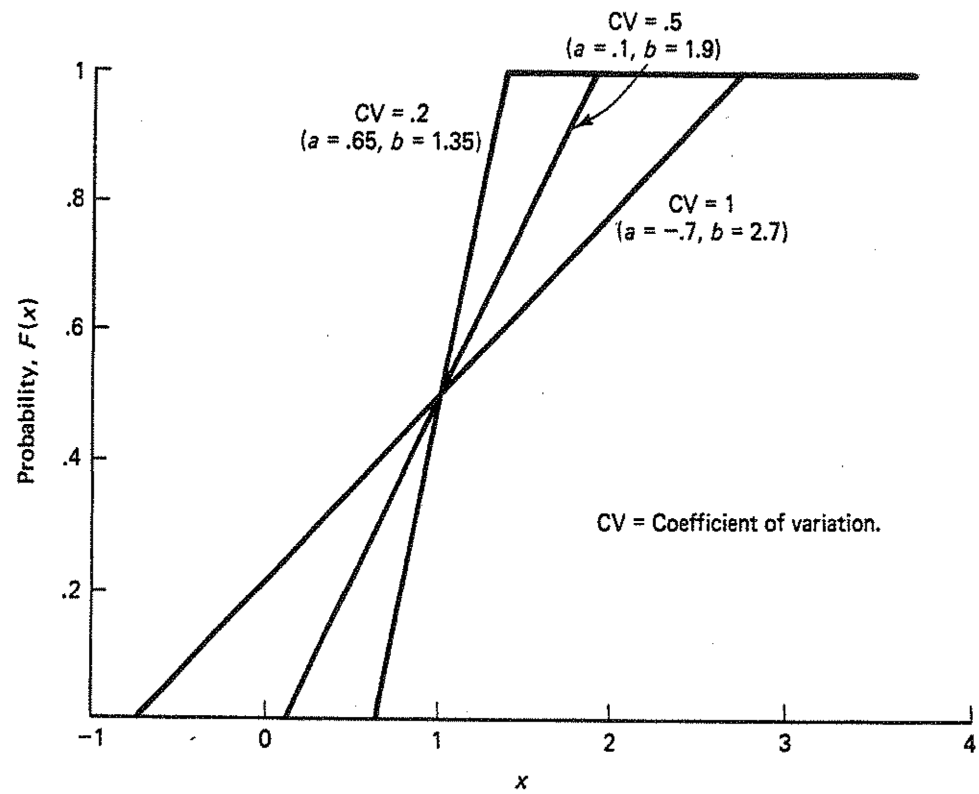
Solution The time until the next arrival has an exponential distribution with mean 1/8 hour. The receptionist would like to determine the value of x for which:

$$F(x) = 1 - e^{-(1/8)x} = .5$$

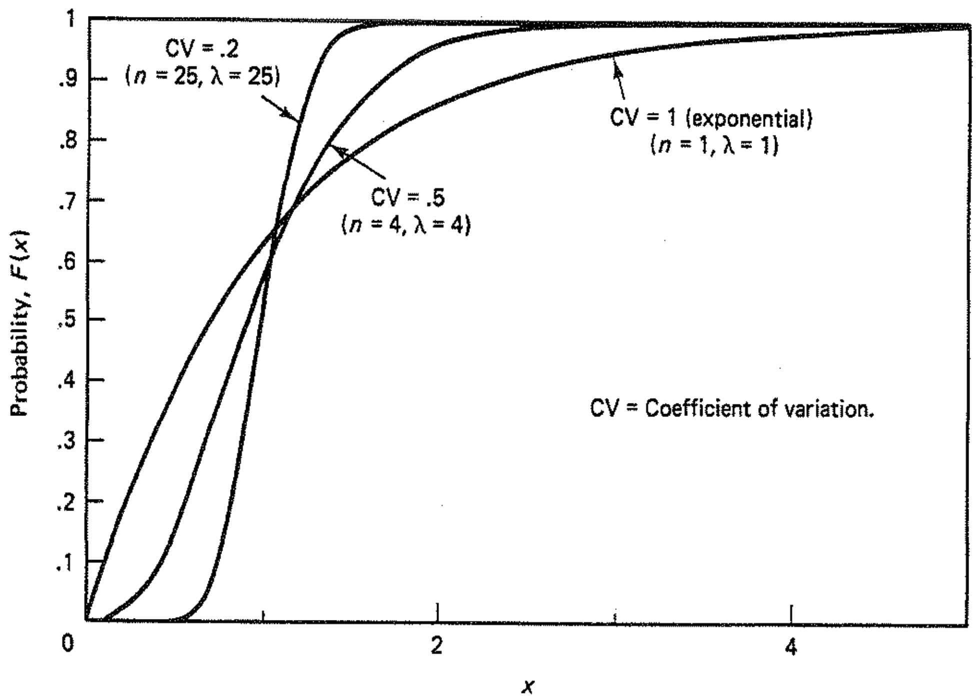
Inverting $F(x)$, this is equivalent to finding the value of x for which

$$x = \frac{-1}{8}\, ln(.5) = .087 \text{ hr} = 5.2 \text{ min}$$

Thus, if the receptionist is gone for 5.2 minutes, there is a 50 percent chance that a customer will arrive before he returns.

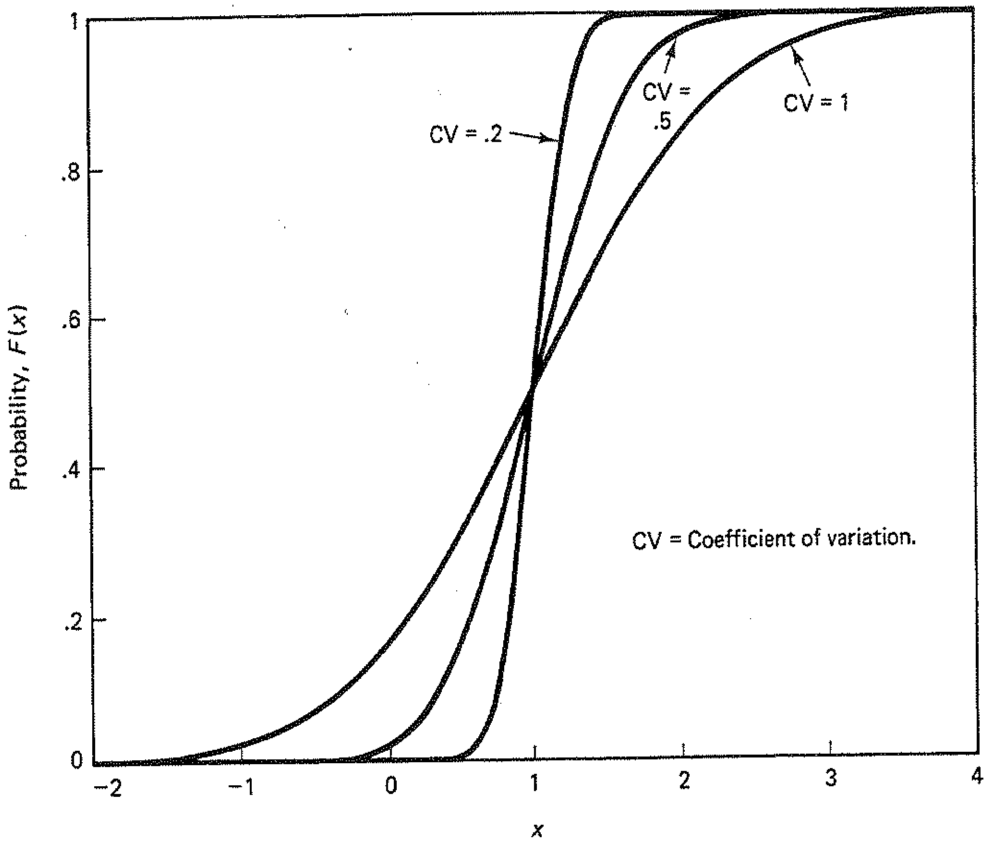


(a) Uniform

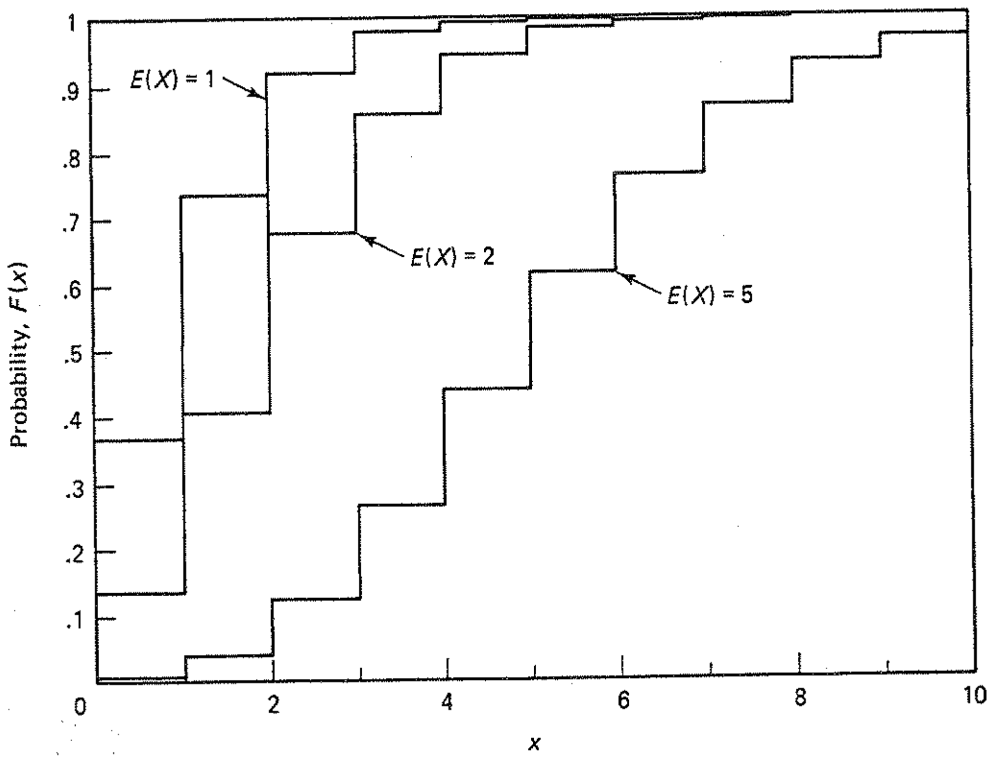


(b) Gamma distribution

Figure 3.1 Comparison of probability distribution functions: (a) uniform, (b) gamma, (c) normal, and (d) Poisson.



(c) Normal



(d) Poisson

Figure 3.1 *(continued)*

3.3.2 Memoryless Property

A random variable is said to be memoryless if the time until the next event does not depend on how much time has already elapsed since the last event:

Definition 3.6

A random variable is ***memoryless*** if:

$$P\{X > s + t \mid X > t\} = P\{X > s\} \qquad \text{for all } s,\ t \geq 0$$

In the context of the Poisson process, the memoryless property applies to the interarrival time. Surely this is a logical consequence of independent increments and stationarity. What consequence should past arrivals have on future arrivals?

The memoryless property is easily proved for the exponential distribution (the distribution governing the time until the next arrival). From the law of conditional probability.

$$P\{X > s + t \mid X > t\} = \frac{P\{X > s + t \text{ and } X > t\}}{P\{X > t\}} = \frac{P\{X > s + t\}}{P\{X > t\}} \tag{3.8}$$

If the random variable X is greater than s plus t, it must also be greater than t. Substituting the exponential distribution function in Eq. (3.8), we obtain the following:

$$P\{X > s + t \mid X > t\} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s} = P(X > s) \tag{3.9}$$

as required. Thus

Property 3. The time until the next event does not depend on the elapsed time since the last event.

Example

Suppose that the receptionist in the previous example left his desk at 10:05, 5 minutes after the last arrival. Then, if he leaves his desk for 5.2 minutes, the probability that no one arrives is still .5, the same as if he had left at 10:00.

3.3.3 Time Until *n*th Event

So far, we know the probability distribution for the number of events within a set time interval, and we know the probability distribution for the length of time until the next event. What about the probability distribution for the time until the second, third, fourth, . . . event?

Definitions 3.7

X_n = the time between the n − 1st event and the nth event

Y_n = the time of the nth event

$= X_1 + X_2 + \ldots + X_n$

Derivation. The probability that Y_2 occurs before some time t equals the probability that $X_1 + X_2$ is less than t. Suppose that X_1 equals s, where s is some non-negative number less than t. Then Y_2 occurs before t if X_2 is less than $t - s$. Considering all the possible values of s between 0 and t:

$$\begin{aligned} P(Y_2 \leq t) &= \int_0^t P(X_1 + X_2 \leq t \mid X_1 = s) P(X_1 = s) ds \qquad (3.10) \\ &= \int_0^t [1 - e^{-\lambda(t-s)}]\, \lambda e^{-\lambda s}\, ds \\ &= \int_0^t \lambda e^{-\lambda s}\, ds - \int_0^t \lambda e^{-\lambda t}\, ds \\ &= (1 - e^{-\lambda t}) - (\lambda t e^{-\lambda t}) \end{aligned}$$

Equation (3.10) gives the probability distribution function for Y_2. The probability density function, $f_{Y_2}(t)$, is the derivative of Eq. (3.10), $F_{Y_2}(t)$, with respect to t, and equals

$$f_{Y_2}(t) = \frac{dF_{Y_2}(t)}{dt} = \lambda^2 t e^{-\lambda t} \qquad t \geq 0 \qquad (3.11)$$

Equation (3.11) is recognized as the density function of the ***gamma distribution*** with parameters $(2,\lambda)$. In general, the distribution for the sum of n successive interarrival times has the following property:

Property 4. The sum of n independent exponential random variables, each with mean $1/\lambda$, has the gamma distribution with parameters (n,λ).

The probability density function for the gamma distribution is the following:

Gamma Probability Density Function

$$f_{Y_n}(t) = \frac{\lambda^n}{\Gamma(n)} t^{n-1} e^{-\lambda t} \qquad t \geq 0 \qquad (3.12)$$

where $\Gamma(n)$ is the ***gamma function*** (the gamma function is *not* a probability distribution function). When n is an integer (as will be the case in queueing analysis), the gamma function is defined as

$$\Gamma(n) = (n - 1)! \qquad n \geq 1 \text{ and integer} \qquad (3.13)$$

The gamma distribution is continuous and defined over the domain $t \geq 0$. The mean and variance of the gamma distribution are

$$E(T) = \frac{n}{\lambda} \qquad V(T) = \frac{n}{\lambda^2} \tag{3.14}$$

The gamma distribution function is found by integrating Eq. (3.12):

Gamma Probability Distribution Function

$$F_{Y_n}(t) = 1 - \sum_{j=0}^{n-1} \frac{e^{-\lambda t}(\lambda t)^j}{j!} \quad n \geqslant 1 \text{ and integer} \tag{3.15}$$

For noninteger values of n, the integration must be performed numerically. $F_{Y_n}(t)$ is shown in Fig. 3.1b, for $E(T) = 1$, and coefficients of variation of .2, .5, and 1 (the last also being an exponential distribution). Note that the distribution loses symmetry as the coefficient of variation increases.

Although the gamma distribution may appear complicated, it is a natural consequence of the conditions underlying the Poisson process—conditions that are quite plausible.

3.3.4 Event Times Within a Time Interval

The Poisson process has one more property that deserves mention. Suppose that n events are known to have occurred within a time interval. What is the probability distribution for the time of each event?

First, suppose that just one event occurred over the interval $[0,t]$, and let the time of this event be represented by the random variable Y_1. Then the following two statements, representing probabilities of joint events, are equivalent:

$$P\{[Y_1 < s] \text{ and } [N(t) = 1]\} = P\{[1 \text{ event} \in [0,s)] \text{ and } [0 \text{ events} \in [s,t]]\} \tag{3.16}$$

From the law of conditional probability, the first statement can also be expressed as

$$P\{[Y_1 < s] \text{ and } [N(t) = 1]\} = P[Y_1 < s \mid N(t) = 1] \cdot P[N(t) = 1] \tag{3.17}$$

Combining the expressions and substituting the Poisson probability distribution yield the following:

$$P[Y_1 < s \mid N(t) = 1] = \frac{(\lambda s e^{-\lambda s})(e^{-\lambda(t-s)})}{\lambda t e^{-\lambda t}} = \frac{s\, e^{-\lambda t}}{t e^{-\lambda t}} = \frac{s}{t} \tag{3.18}$$

Equation (3.18) is the *uniform* $[0,t]$ probability distribution function, meaning that the event is equally likely to occur at any time during the interval. This should come as no surprise, given the memoryless property of the Poisson process. More generally,

Property 5. If $N(t) = n$, the unordered event times are defined by $N(t)$ independent uniform $[0,t]$ random variables.

The uniform distribution has the following characteristics:

Characteristics of Uniform Distribution Defined over $[a,b]$

1. The uniform distribution is continuous. Depending on a and b, uniform random variables can be either negative or positive.
2. $$f(x) = \begin{cases} 1/(b - a) & x \in [a, b] \\ 0 & \text{elsewhere} \end{cases}$$
3. $$F(x) = \begin{cases} 0, & x < a \\ (x - a)/(b - a) & x \in [a, b] \\ 1, & x > b \end{cases}$$
4. $$E(X) = \frac{b + a}{2}$$
5. $$V(X) = \frac{(b - a)^2}{12}$$

Example distribution functions are shown in Fig. 3.1a. As with the gamma figure, the functions have a mean of 1 and coefficient of variations of .2, .5, and 1. Note that the uniform distribution is always symmetric about its mean and that the slope of the distribution function is constant over $[a,b]$.

Example

Three events are known to have occurred over a 2-hour period of a Poisson process. The probability that any one event occurred in the first half hour is $(1/2)/2 = .25$. The probability that all three events occurred in the first half hour is $.25^3 = 1/64$. The probability that none of the events occurred in the first half hour is $(1 - .25)^3 = 27/64$.

3.3.5 Summary

In this section we have seen that a few plausible assumptions (independence, stationarity, and arriving one at a time) lead to a number of important consequences:

1. The number of events within an interval of length t has a Poisson distribution with mean λt.
2. The time until the next event has the exponential distribution.
3. The time until the next event is independent of the elapsed time since the last event (memoryless property).
4. The time until the nth event has a gamma distribution with parameters (n,λ).

5. If $N(t) = n$, the unordered event times within the interval $[0,t]$ are defined by n independent uniform $[0,t]$ random variables.

These properties are used in the next section to check whether an observed process is indeed Poisson, and in Chap. 4 as the basis for simulating a queue.

3.4 GOODNESS OF FIT

"Goodness of fit" is a term that describes how well a model fits the behavior of a system. Occasionally, a model precisely matches the behavior of a system, as in the binomial distribution matching flipped coins. But, more commonly, there are some differences. Differences are acceptable if the model does not differ appreciably from reality. The amount of difference that is acceptable is a matter of judgment and varies from situation to situation, depending on the needs for accuracy and simplicity.

Above all else, and independent of whatever data are recorded, goodness of fit should be judged according to whether or not the model's assumptions are plausible for the system studied. This standard is best appreciated by considering situations that do *not* conform to the Poisson process.

Example 1

An observer is stationed at the end of the runway at Newark International Airport. Over the 1-hour period from 5:00 to 6:00 P.M., the observer records the time that the front wheels of each airplane touch the runway.

Airplanes certainly land one at a time. And the probability that an airplane arrives at any particular time does not vary appreciably over the hour. But arrivals are not independent of each other. Safety dictates a minimum spacing between planes. If a plane landed at 5:00, the next plane would not land until at least 90 seconds later. The process does not possess independent increments and it is not Poisson.

Example 2

Job candidates are scheduled for interviews with a personnel department. One candidate is scheduled for each hour of the day, beginning at 8:00. An observer records the time that each candidate arrives.

The job candidates most likely arrive independently of each other, one at a time. But the process is not stationary. Arrival times depend on appointment times. The likelihood that a candidate arrives at 7:55 is much greater than the likelihood that a candidate arrives at 8:15. Again, the process is not Poisson. (If the spacing between appointments is small, and arrival times are somewhat random, the process would behave *locally* like a Poisson process. That is, the interarrival times would be approximately exponential, but the number of arrivals over long time intervals would not be Poisson. See Newell 1982.)

Example 3

> An observer records the time that visitors arrive at the San Francisco zoo during the period 10:00 to 11:00 on a weekday morning. Most of the visitors are members of school groups.

If it takes no longer to serve a group of visitors than an individual visitor, then the group would be considered the customer and the arrival process might be Poisson. However, if visitors are served individually, the arrival process would not be Poisson because visitors do not arrive one at a time.

Example 4

> An observer records the arrivals of bank patrons at an automated teller machine over a 1-hour period from 10:00 to 11:00 on a weekday morning.

Bank patrons tend to arrive one at a time, and arrivals tend to be mutually independent. The rate at which they arrive may vary over the day. However, the arrival rate may be fairly constant over a 1-hour period from 10:00 to 11:00. The Poisson process is *plausible*.

The bank is the only one of the four examples for which the Poisson process is plausible. But plausibility does not imply that the process is definitely Poisson. It only means that the process seems to be Poisson. If there is any doubt, final verdict should not be passed until quantitative tests are performed on the data, as discussed in the following section.

If the Poisson process is not plausible, all is not lost. There are many other models for the arrival process, any of which might be appropriate for a given situation. Some of these are described in Chap. 6. The Poisson process is emphasized here because it is a reasonable model for many systems and because its simplicity facilitates evaluation of queueing system performance.

3.5 QUANTITATIVE GOODNESS OF FIT TESTS

Suppose that the Poisson process is a plausible model for customer arrivals, but you are not positive that it is the right model. Your next step is to test the arrival data to see how well they conform to the Poisson process. Quantitative testing seeks to answer two key questions:

1. Whether or not the model is correct, and
2. If the model is not correct, why is it not correct?

The answer to the latter question provides guidance in creating an alternative model that more closely matches the data.

Tests can be performed for any of the properties of the Poisson process. However, because some of the properties are implied by others, there is no need to test them all. For example, if the interarrival times are found to be independent exponential random variables, pairs of interarrival times do not have to be tested to see if they are gamma random variables.

The tests are divided into two categories: graphical tests and statistical tests. Graphical tests are less rigorous than statistical tests. Yet they are very effective at identifying patterns in the data. Statistical tests impose standards for whether or not the data conform to the properties of the Poisson process. In both cases, the tests merely check to see whether the data conform with the properties of the Poisson process. Whether the conditions will recur in the future is a matter of inference.

3.5.1 Graphical Tests

Perhaps the best way to determine whether or not the arrival process is Poisson is to plot and examine the data for patterns. For the graphs described here, absence of pattern generally is an indication that the process is Poisson; presence of pattern generally indicates that the process is not Poisson. If a pattern is found, a cause should be sought. If the process is not stationary, why is it not stationary? If the interarrival times are not exponential, why are they not exponential? If the interarrival times are not independent, why are they not independent? Understanding the process creating the data is by far the most effective means for deciding whether the model is correct or whether to design an alternative model.

Stationarity: Cumulative. Plot the cumulative arrival curve. Draw a straight line connecting the points $A(0)$ and $A(T)$, where T is the end of the time interval observed. There should be no visible pattern to the deviations of A(T) around the straight line, and the difference between $A(T)$ and the straight line should be small for all values of t.

Stationarity/Independence: Interarrival Times. Plot the interarrival times in serial order on an (x,y) graph. The points should be scattered randomly about the line $y = \bar{X}$, where $\bar{X}$ is the average interarrival time. There should be no cyclic patterns.

Independence. Plot the Points $\{(X_n, X_{n-l}), n = 2, 3, \ldots\}$ on graph paper, where X_n represents the nth interarrival time. It should be impossible to approximate the data by a straight line or any other regular curve.

Exponential Interarrival Distribution. Plot the empirical probability distribution for the interarrival times. On the same piece of paper, plot the exponential distribution with $\lambda = A(T)/T$, where T is the length of the time interval observed. There should be no visible pattern of the deviations between the distributions.

Demonstration. The date in Table 3.1 were collected at the automatic teller machine for a large bank branch in Berkeley, California. The cumulative arrival curve is shown in Fig. 3.2. The arrival curve varies above and below the straight line, in a fairly random fashion. The interarrival times in Fig. 3.3 also show no time varying pattern. Figure 3.4 plots (X_n,X_{n-1}) for 97 data points (number of arrivals minus 1). The downward slope of the *boundary* of the data points makes the data appear to have a negative correlation. However, the data points themselves do not slope downward, and again there is no visible pattern. Finally, the empirical probability distribution for the interarrival time

TABLE 3.1 ARRIVAL, DEPARTURE, AND WAITING TIMES AT AUTOMATIC TELLER MACHINE (MINUTES, TIME 0 = 9:30)

n	$A^{-1}(n)$	$S(n)$	$D_q^{-1}(n)$	$D_s^{-1}(n)$	$W_q(n)$	$W_s(n)$
1	0.61	0.73	0.61	1.34	0.00	0.73
2	0.86	0.85	1.34	2.19	0.48	1.33
3	1.09	1.75	2.19	3.94	1.10	2.85
4	5.61	0.95	5.61	6.56	0.00	0.95
5	6.59	0.77	6.59	7.36	0.00	0.77
6	6.90	0.82	7.36	8.17	0.46	1.27
7	7.71	0.70	8.17	8.87	0.46	1.16
8	7.93	0.72	8.87	9.59	0.94	1.66
9	7.93	0.73	9.59	10.32	1.66	2.39
10	8.39	0.87	10.32	11.19	1.93	2.80
11	9.22	0.73	11.19	11.92	1.97	2.70
12	9.29	0.60	11.92	12.52	2.63	3.23
13	12.92	0.63	12.92	13.55	0.00	0.63
14	14.57	0.53	14.57	15.10	0.00	0.53
15	17.67	0.95	17.67	18.62	0.00	0.95
16	20.15	0.80	20.15	20.95	0.00	0.80
17	20.39	0.87	20.95	21.82	0.56	1.43
18	20.83	0.75	21.82	22.57	0.99	1.74
19	20.99	1.12	22.57	23.68	1.58	2.69
20	21.31	1.88	23.68	25.57	2.37	4.26
21	21.39	0.88	25.57	26.45	4.18	5.06
22	22.96	0.78	26.45	27.23	3.49	4.27
23	24.03	0.68	27.23	27.92	3.20	3.89
24	24.15	0.48	27.92	28.40	3.77	4.25
25	24.32	0.82	28.40	29.22	4.08	4.90
26	25.40	0.90	29.22	30.12	3.82	4.72
27	25.77	0.72	30.12	30.83	4.35	5.06
28	30.39	0.52	30.83	31.35	0.44	0.96
29	30.91	1.30	31.35	32.65	0.44	1.74
30	34.43	0.73	34.43	35.16	0.00	0.73
31	34.54	1.15	35.16	36.31	0.62	1.77
32	35.27	0.62	36.31	36.93	1.04	1.66
33	35.39	0.47	36.93	37.40	1.54	2.01
34	36.38	0.77	37.40	38.16	1.02	1.78
35	40.39	0.73	40.39	41.12	0.00	0.73
36	41.61	0.83	41.61	42.44	0.00	0.83
37	44.77	0.53	44.77	45.30	0.00	0.53
38	46.09	0.62	46.09	46.71	0.00	0.62
39	48.85	0.67	48.85	49.52	0.00	0.67
40	51.40	1.45	51.40	52.85	0.00	1.45
41	52.87	0.63	52.87	53.50	0.00	0.63
42	53.84	0.77	53.84	54.61	0.00	0.77
43	55.09	1.08	55.09	56.17	0.00	1.08
44	55.64	0.92	56.17	57.09	0.53	1.45
45	55.83	0.57	57.09	57.66	1.26	1.83
46	57.56	0.85	57.66	58.51	0.10	0.95
47	58.65	0.63	58.65	59.28	0.00	0.63
48	59.83	0.72	59.83	60.55	0.00	0.72
49	61.91	0.57	61.91	62.48	0.00	0.57

TABLE 3.1 *(continued)*

n	$A^{-1}(n)$	$S(n)$	$D_q^{-1}(n)$	$D_s^{-1}(n)$	$W_q(n)$	$W_s(n)$
50	64.71	0.57	64.71	65.28	0.00	0.57
51	64.95	0.95	65.28	66.23	0.33	1.28
52	70.20	1.37	70.20	71.57	0.00	1.37
53	71.18	1.48	71.57	73.05	0.39	1.87
54	72.15	0.68	73.05	73.73	0.90	1.58
55	72.72	0.72	73.73	74.45	1.01	1.73
56	72.94	0.65	74.45	75.10	1.51	2.16
57	74.74	0.63	75.10	75.73	0.36	0.99
58	75.12	0.58	75.73	76.32	0.61	1.20
59	75.64	0.65	76.32	76.97	0.68	1.33
60	75.77	0.68	76.97	77.65	1.20	1.88
61	76.95	0.72	77.65	78.37	0.70	1.42
62	78.36	0.85	78.37	79.22	0.01	0.86
63	79.98	0.80	79.98	80.78	0.00	0.80
64	80.21	0.87	80.78	81.65	0.57	1.44
65	81.93	0.67	81.93	82.60	0.00	0.67
66	82.75	0.73	82.75	83.48	0.00	0.73
67	82.95	0.60	83.48	84.08	0.53	1.13
68	83.35	0.65	84.08	84.73	0.73	1.38
69	84.55	0.68	84.73	85.42	0.18	0.87
70	84.72	0.70	85.42	86.12	0.70	1.40
71	85.28	1.00	86.12	87.12	0.84	1.84
72	86.17	0.60	87.12	87.72	0.95	1.55
73	87.58	0.87	87.72	88.58	0.14	1.00
74	87.85	1.25	88.58	89.83	0.73	1.98
75	88.33	0.45	89.83	90.28	1.50	1.95
76	88.42	1.00	90.28	91.28	1.86	2.86
77	92.17	0.67	92.17	92.84	0.00	0.67
78	92.97	0.85	92.97	93.82	0.00	0.85
79	93.63	0.83	93.82	94.65	0.19	1.02
80	96.34	0.58	96.34	96.92	0.00	0.58
81	96.53	1.85	96.92	98.77	0.39	2.24
82	97.39	1.00	98.77	99.77	1.38	2.38
83	97.88	0.77	99.77	100.54	1.89	2.66
84	99.64	0.90	100.54	101.44	0.90	1.80
85	103.82	0.77	103.82	104.59	0.00	0.77
86	104.88	0.65	104.88	105.53	0.00	0.65
87	105.12	1.35	105.53	106.88	0.41	1.76
88	105.66	0.67	106.88	107.55	1.22	1.89
89	106.92	0.93	107.55	108.48	0.63	1.56
90	107.55	1.00	108.48	109.48	0.93	1.93
91	110.02	0.90	110.02	110.92	0.00	0.90
92	113.14	2.35	113.14	115.49	0.00	2.35
93	114.12	0.73	115.49	116.22	1.37	2.10
94	115.01	0.48	116.22	116.71	1.21	1.70
95	116.32	0.58	116.71	117.29	0.39	0.97
96	117.50	0.55	117.50	118.05	0.00	0.55
97	119.18	0.55	119.18	119.73	0.00	0.55
98	119.71	0.85	119.73	120.58	0.02	0.87

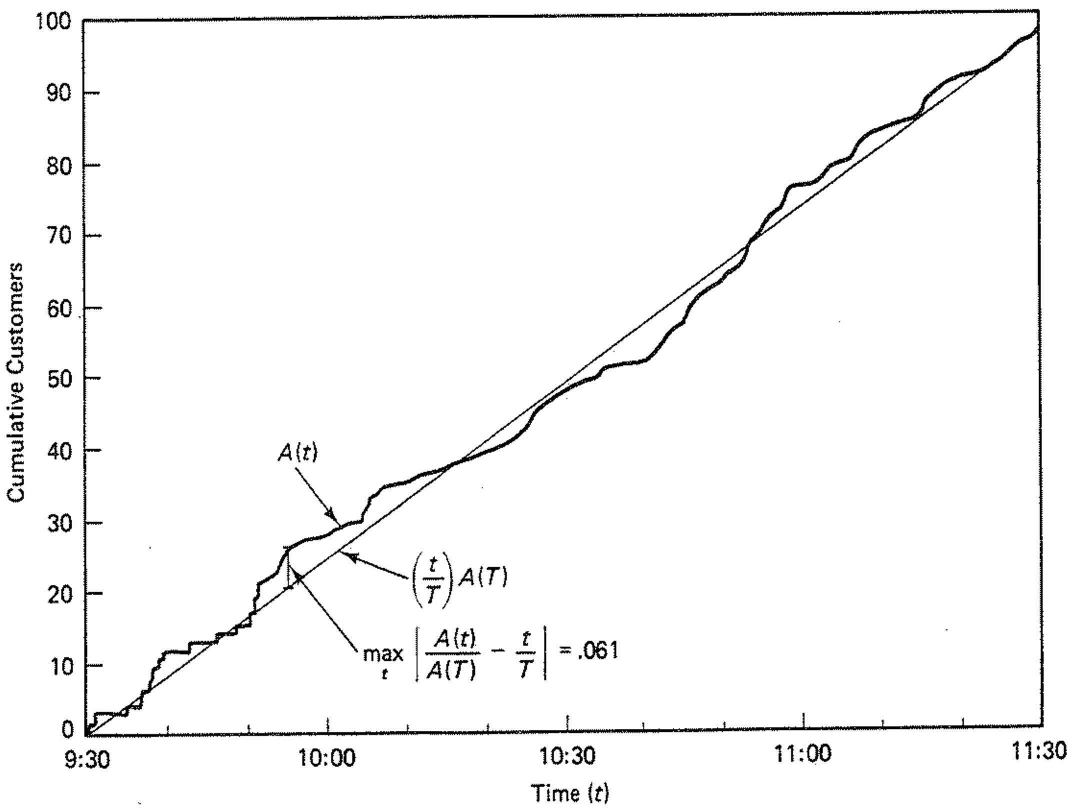


Figure 3.2 Cumulative arrivals recorded at automated teller machine. If arrivals are stationary, the deviation from the diagonal line should be small (see Kolmogorov-Smirnov test), as shown.

is shown in Fig. 3.5. The deviations are small and again show no pattern. Based on these four graphs, there is no reason to believe that the process is not Poisson.

Consider a second example, which is not Poisson. Table 3.2 provides data on arrival times of elevators at the lobby of a high-rise building on the University of California campus. The plot of the interarrival time distribution in Fig. 3.6 appears to be exponential. Though not shown, the cumulative arrival plot also appears to be stationary. However, the plot of paired interarrival times (Fig. 3.7) reveals a subtle dependency. It appears that short interarrival times tend to be followed by long interarrival times, and vice versa (note that fewer points are concentrated near the origin and more points are further out near the axes). There is a negative correlation between (X_n, X_{n-1}).

Careful observation of the arrival process should reveal the cause of the pattern. There is a common phenomenon in many modes of transportation (buses and elevators, for example) known as "bunching." The number of people who board a vehicle depends on how much time has elapsed since the previous vehicle arrived (the longer the time, the more people). If this separation happens to decline, fewer people will board the second vehicle, and the vehicle will begin to travel faster, further reducing the separation. Eventually, the second vehicle will catch up with the first vehicle, and from then on the two will travel in a pair, perhaps "leap frogging" from stop to stop. Bunching is common in elevators and explains why the interarrival times are not independent. Hence, the arrival process is not Poisson, and a different model is called for.

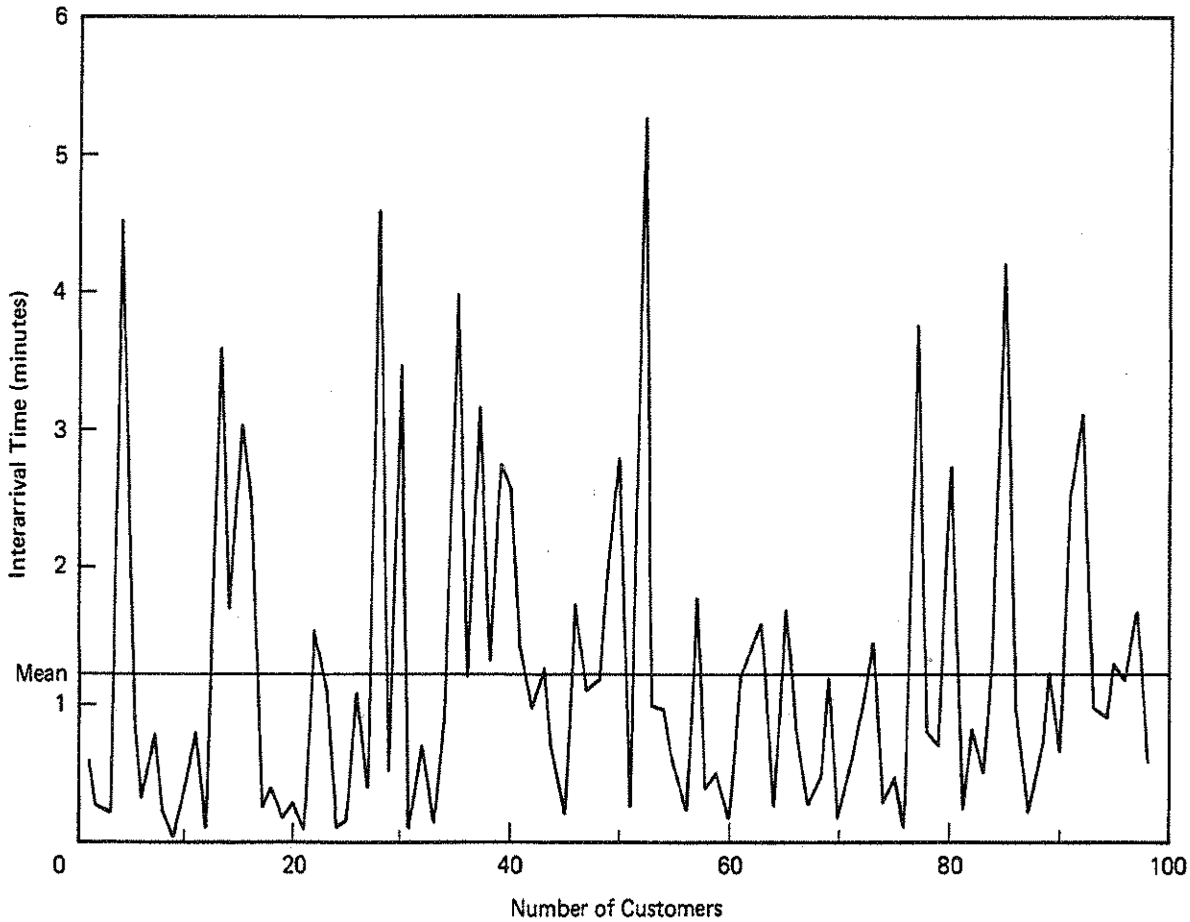


Figure 3.3 Interarrival times at automated teller machine. If arrival process possesses independent increments, there should be no cyclic pattern to the data, as shown.

The point to recognize is that an unexpected pattern is reason for further investigation. The system should be observed to identify what factors are causing the pattern. If the factors are significant, they should be incorporated into the model.

3.5.2 Statistical Tests

Graphs alone may leave doubt as to whether the process is Poisson. Perhaps a small deviation was found, but you have no idea what caused it. You may wonder how much deviation is acceptable. Where is the line drawn between what is a Poisson process and what is not? Statistical tests help in these situations.

Definition 3.8

A ***statistic*** is a function of the data. Like the data, a statistic is itself a random variable. The sample average, sample standard deviation, and maximum and minimum are all examples of statistics.

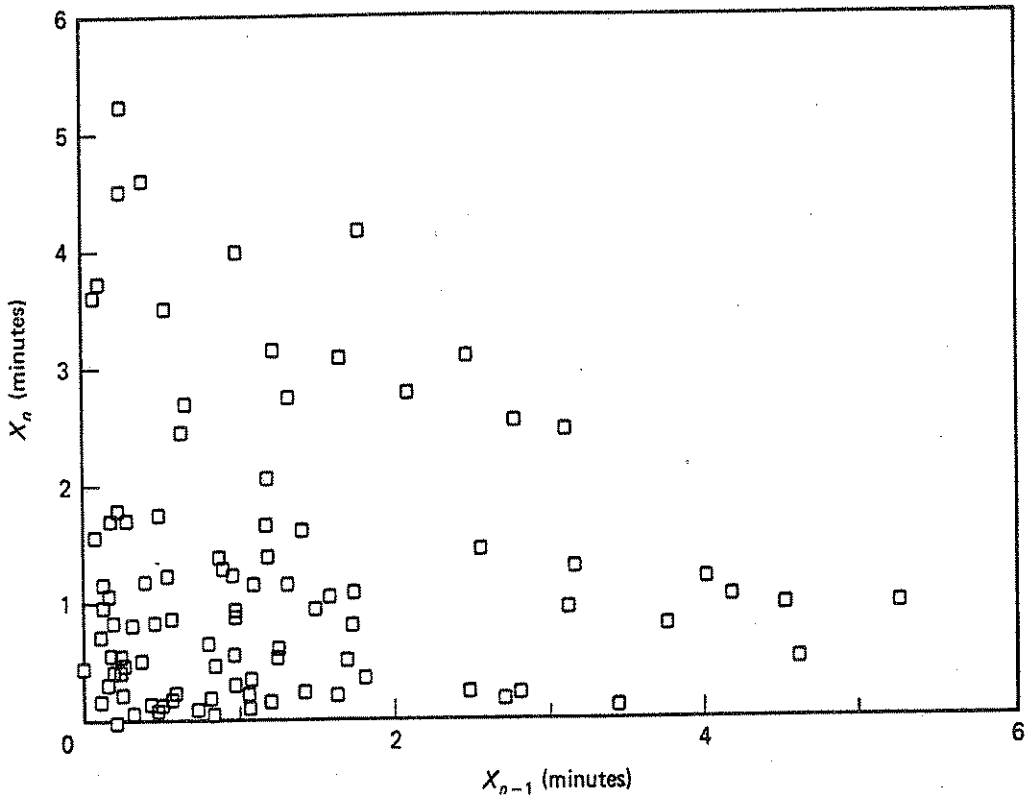


Figure 3.4 Paired successive interarrival times at automated teller machine. If arrival process possesses independent increments, the data cannot be approximated by a straight line, as shown.

Statistical tests are phrased in the form of a ***hypothesis*** (denoted by the letter H). For example, suppose that you have made a bet with a friend. Each time she flips a coin and it turns up heads, she wins a dollar. Each time it turns up tails, you win a dollar. Now suppose that after ten flips the coin has turned up heads nine times. You may then wonder whether the coin she flipped is fair. That is, you may wonder whether the following hypothesis, H, is true:

H: Probability of heads = .5

or whether antihypothesis, A, is true instead;

A: Probability of heads > .5

The statistical test only suggests whether or not the hypothesis is true; it does not provide conclusive evidence. It does this by *calculating the probability of obtaining the observed data, given that the hypothesis is true*. In the case of the coin flip, you would be interested in the probability of obtaining one or fewer heads in ten trials, given the probability of .5. From the binomial distribution, this probability can be expressed as follows:

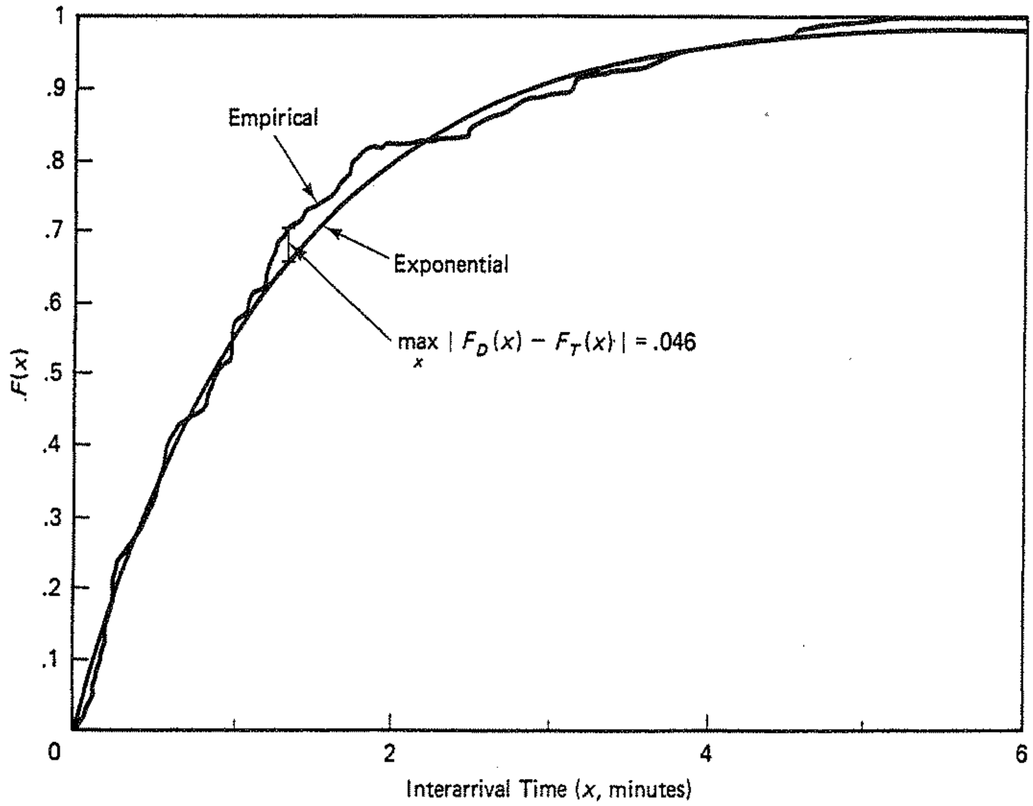


Figure 3.5 Interarrival time distribution at automated teller machine. For Poisson process, deviation from the exponential distribution should be small, as shown.

$$\begin{aligned} P(\text{data} \mid H \text{ is true}) &= P(1 \text{ or fewer heads in 10 trials} \mid p = .5) \\ &= \binom{10}{0} .5^0 (1 - .5)^{10} + \binom{10}{1} .5^1 (1 - .5)^9 \\ &= .011 \end{aligned}$$

The probability of seeing one or fewer heads is only .011, so there is reason to be suspicious that the coin (and perhaps your friend) is unfair. Nevertheless, because the probability is greater than zero, one cannot say unequivocally that the coin is unfair.

In statistics, the $P(\text{data} \mid H \text{ is true})$ is expressed as a ***significance level***, which is merely a way of rounding off the probability. In the example, the probability would be rounded up to .02, and one would say that "the hypothesis is rejected at the 2% significance level." This statement does not imply that the hypothesis is absolutely

TABLE 3.2 ARRIVAL TIMES OF ELEVATORS (MINUTES)

.8	1.2	2.8	3.2	4.1	7.5	8.7	9.7	10.2	11.1	13.5
15.0	16.2	16.2	19.1	21.9	22.1	23.5	24.1	26.0	27.0	27.1
27.9	20.7	31.0	31.4	33.3	33.8	36.2	36.3	40.0	40.2	41.1
42.6	44.8	45.1	45.6	45.8	49.0	51.2	51.8	53.5	53.8	54.5
55.0	56.5	56.8								

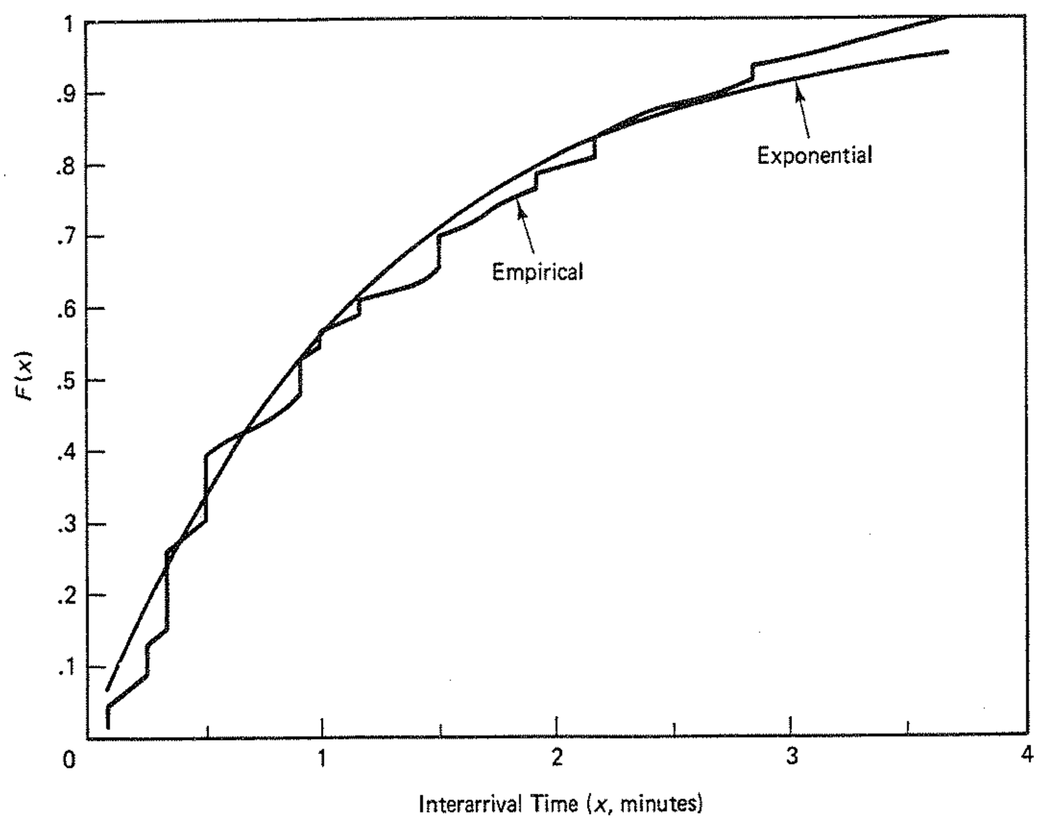


Figure 3.6 Interarrival time distribution for elevators appears to be exponential.

rejected. It only means that the data have a likelihood of less than 2%. As a matter of custom, $P(\text{data} \mid H \text{ is true})$ is compared to a significance level of 1%, 2%, or 5%. If the probability falls below .05, then it is rounded up to the next highest significance level and the hypothesis is rejected at that level. If the probability falls above .05, then we would say that "the hypothesis is not rejected at the 5% significance level." Ordinarily, one never states that a hypothesis is accepted, because one can never be completely sure.

Statistical tests are effective at identifying when something is amiss but not at finding solutions. They offer little guidance as to what is right. Therefore, they are no substitute for plotting the data and checking to see how the data actually behave.

Statistical tests are provided for the following hypotheses, all of which must be true if the arrival process is Poisson:

H_1: The interarrival times have an exponential probability distribution.

H_2: The interarrival times are independent.

H_3: The unordered arrival times have a uniform distribution over the period of observation.

If none of the three hypothesis is rejected *and* the graphical tests show no abnormal pattern *and* the assumptions of the Poisson process are plausible, then it is safe to assume that the observed arrival process was indeed Poisson. However, just because a data set does not

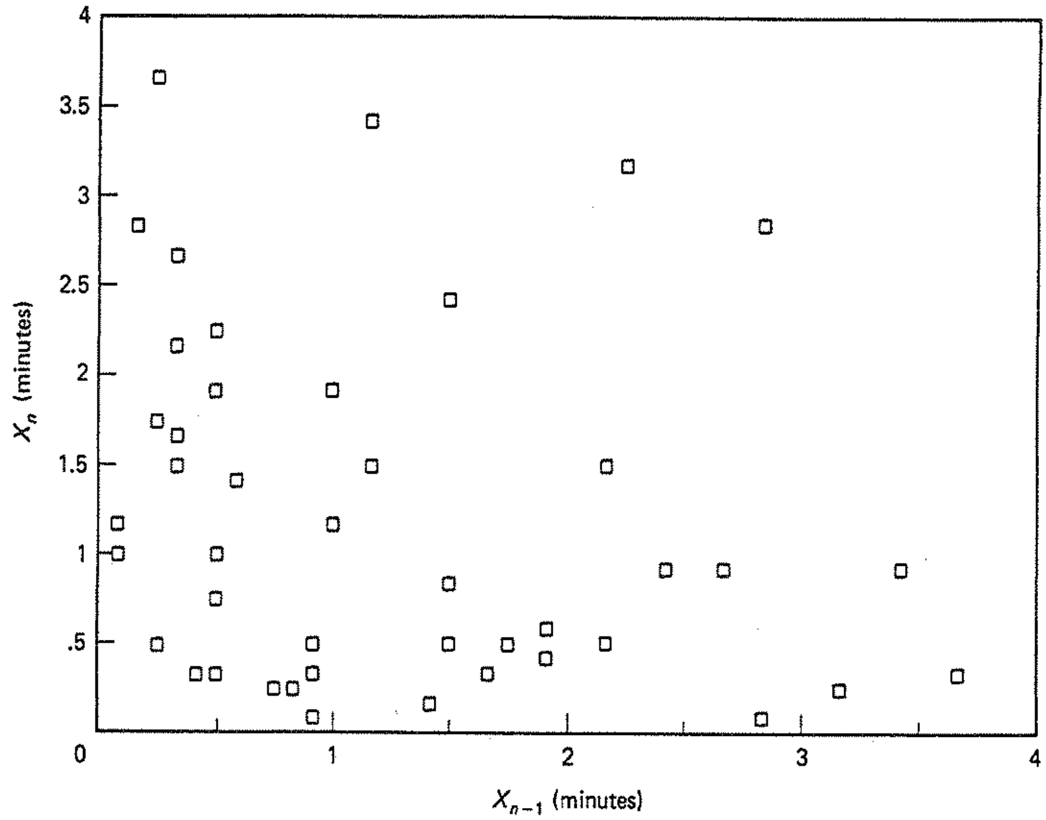


Figure 3.7 Paired successive interarrival times of elevators. Negative correlation revealed by points along axes.

conform exactly to the assumptions of the Poisson process does not imply that the Poisson process is not a reasonable model. The simplicity of the Poisson model may outweigh any gain in accuracy from using a more complicated model.

3.5.2.1 H_1: Interarrival Time Distribution. The two most common tests for whether a data set conforms to a theoretical probability distribution are the Kolmogorov-Smirnov (K-S) test and the chi-squared test. The K-S test is based on deviations between the empirical probability distribution function and the theoretical distribution function. The chi-squared test is based on deviations between a data histogram and the theoretical density function. Because the K-S test is the more powerful of the two (and also the simpler), it will be the only test presented here. Books listed at the end of this chapter provide further information on the chi-squared and other distribution tests.

The Kolmogorov-Smirnov test can be applied to any set of continuous, independent, random variables. It is based on the maximum deviation between the empirical distribution function and a theoretical distribution function. If the theoretical distribution is correct, this deviation should be small. The K-S test is an example of a *one-tail test* because a large deviation suggests that the model is incorrect and a small deviation suggests that the model is correct—in contrast to a *two-tail test* in which either a small or large value would cause concern.

Let

$F_D(x)$ be the empirical probability distribution function.
$F_T(x)$ be the theoretical probability distribution function.

The specific shape of the theoretical distribution function is defined by parameters, which may be derived from the data. For the exponential distribution function, the parameter λ can set equal to $A(T)/T$, which is to say, the observed arrival rate (this is discussed in greater detail in a section 3.6). The *K-S statistic*, D, is the maximum deviation between the two distribution functions:

$$D = \max_x |F_D(x) - F_T(x)| \tag{3.19}$$

The maximum must be computed over the entire range of values for x, including $x = 0$. Also, because $F_D(x)$ is a step curve, $F_D(x) - F_T(x)$ must be computed at both the top and the bottom of each step. The K-S statistic has the distribution provided in Table A.4 in the appendix to this book. Reading from the table, if there are 20 points in a data set, and the hypothesis is true, there is a .01 probability that D is greater than .356. This is to say, if D is greater than .356 the hypothesis would be rejected at the 1% significance level. Also from Table A.4, if D is less than .294 for a sample of 20 points, the hypothesis would not be rejected at the 5% significance level.

Example

The maximum deviation between the empirical distribution function and the theoretical distribution function in Fig. 3.5 is .046 (when $x = 1.3$). There are 97 interarrival times in the data set. From Table A.4 in the appendix, .046 is below the critical value .138 ($1.36/\sqrt{97}$) and the hypothesis is not rejected at the 5% significance level. The data are not unusual for a Poisson process.

3.5.2.2 H_2: Interarrival Time Independence.

Independence is one of the most difficult characteristics to test because it has such broad implications. It is possible for an interarrival time, X_n, to be independent of X_{n-1} but not independent of X_{n-2}, X_{n-3}, ... Thus, each statement X_n is independent of X_{n-1} ($n = 2, 3, \ldots$), X_n is independent of X_{n-2} ($n = 3, 4, \ldots$), and so on, might require a separate test. There is really no limit to the types of interdependencies that might be checked. However, unless there is some reason to believe that a complicated form of dependency exists, the test for independence is usually limited to checking whether X_n is independent of X_{n-1}.

Just as there is a chi-squared test for distributions, there is a chi-squared test for independence. The test is based on categorizing the data into a contingency table and comparing the number of observations in each category to the expected number of observations. The numbers should be similar if the hypothesis is correct. The unfortunate part of this approach is that diffusing the data into categories necessitates that a large data set be collected for the test to be powerful.

An alternative to the chi-squared test for independence is the *t-test*. This checks the following weaker hypothesis: *the correlation coefficient between X_n and X_{n-1} equals* *zero*. Independent random variables must have a correlation coefficient of zero, but the reverse is not necessarily true. Figure 3.8 plots data points that are not independent, yet still have zero correlation. However, such a pattern would be most unusual in observing queues, and even if it did materialize, it would be easily identified by plotting the data. Hence, if it can be shown that the correlation coefficient is zero between pairs of successive interarrival times, then the data are likely independent.

The correlation coefficient between two data sets, $\{X_1, X_2, \ldots, X_N\}$ and $\{Y_1, Y_2, \ldots, Y_N\}$, is the ratio of the covariance to the product of the standard deviations.

Definition 3.9

r_{xy} is the *sample correlation coefficient*

$$= \frac{\sum_{n=1}^{N} (X_n - \bar{X})(Y_n - \bar{Y})}{(N-1)s_x s_y} \tag{3.20}$$

where $\bar{X}$ and $\bar{Y}$ are sample averages, and s_x and s_y are *sample standard deviations:*

$$s_x = \sqrt{\frac{\sum_{n=1}^{N} (X_n - \bar{X})^2}{N-1}} \tag{3.21a}$$

$$s_y = \sqrt{\frac{\sum_{n=1}^{N} (Y_n - \bar{Y})^2}{N-1}} \tag{3.21b}$$

$N - 1$ is used instead of N in the denominator to obtain an unbiased estimate of the standard deviation. That is, the expectation of s_x and s_y equals σ_x and σ_y. The correlation coeffficent can be any value between -1 and $+1$.

The statistical test for correlation is not performed directly on r_{xy}. Rather, it is performed on the transformation of r_{xy} into the t statistic, as shown below (Note: this value of t has nothing to do with the value of t used in Sec. 3.3):

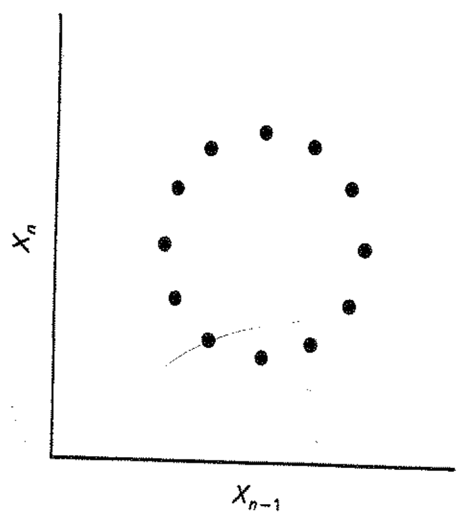


Figure 3.8 Dependent random variables might not be correlated, as shown here.

$$t = \frac{r_{xy}\sqrt{N-2}}{\sqrt{1-r_{xy}^2}} \tag{3.22}$$

If the random variables X and Y are independent and have normal distributions, then the statistic t should have the ***t probability distribution*** with $N - 2$ degrees of freedom (see Table A.3 in the appendix to this book). In the case of the Poisson process, X and Y have the exponential distribution, so the t-test is only approximate.

The t-test is a two-tail test because either positive correlation (r_{xy} close to 1) or negative correlation (r_{xy} close to -1) is cause for concern. The significance level is defined by the probability that $|t|$ is greater than or equal to the observed value of t, given that the hypothesis is true.

Example

The correlation coefficient (X_n, X_{n-1}) for the teller data set was calculated as $-.032$. There are 96 *pairs* of interarrival times in the data set (one less than the number of interarrival times), so $N = 96$. Substitution of $-.032$ for r_{xy} and 96 for N in Eq. (3.22) gives $t = -.31$.

Rounding N to 120 in Table A.3, $P(t < 1.98)$ equals .975. Because the t distribution is symmetric, the $P(|t| > 1.98)$ equals .05. Because $|-.31|$ is well below 1.98, the hypothesis that the correlation coefficient equals zero *cannot be rejected* at the 5% significance level.

When the data set has 50 or more data points, the normal distribution (Table A.2) can be used in place of the t distribution. The normal distribution is identical to the last line of Table A.3 for the t distribution ($N = \infty$).

3.5.2.3 Arrival Time Uniformity. If the arrival process is stationary, the unordered arrival times should conform to the uniform distribution. That is, the empirical probability distribution should be approximately a straight line passing through the points (0,0) and (T,1), where T is the length of the time period observed. Though it was not introduced as such, the empirical probability distribution for arrival times is simply the cumulative arrival diagram (Fig. 3.2), with the vertical axis rescaled from 0 to 1.

The Kolmogorov-Smirnov test can be used again. The K-S statistic, D, is calculated as follows:

$$D = \max_t \left| \frac{A(t)}{A(T)} - t/T \right| \tag{3.23}$$

A large value of D would suggest that the arrival process is not stationary.

Example

For the sample teller data, $D = .061$, which occurs at time 9:55. There are 98 arrival times in the data set. Because .061 is less than the critical value, .137 ($1.36/\sqrt{98}$), the hypothesis cannot be rejected at the 5% significance level.

3.5.2.4 A Quick Test. As mentioned earlier, the coefficient of variation (ratio of standard deviation to mean) for the exponential distribution is 1. One of the easiest checks for whether an arrival process is Poisson is to calculate this value and compare it to 1.

Example

The standard deviation of the interarrival times is 1.18 and the average is 1.22. Therefore, the coefficient of variation is .97, which is nearly 1.

3.5.2.5 Interpretation of Statistical Tests. Models are rarely perfect representations of reality. Occasional disturbances might disrupt the usual arrival pattern—some customers might arrive in pairs or the arrival rate might fluctuate slightly. Small imperfections such as these are difficult to detect in small data samples, yet will certainly arise if the sample is large enough. This means that a slightly imperfect model will invariably be statistically rejected if the data sample is sufficiently large.

For practical purposes, one should not discard slightly imperfect models, only grossly imperfect models. Just because a model is statistically rejected does not mean it should not be used, particularly if the data sample is large. It does mean though, that the cause of the imperfection should be identified and considered for inclusion in the model. If the gain in accuracy does not justify the added complexity, then the cause should not be included. Like many aspects of modeling, this is a matter of judgment.

3.6 PARAMETER ESTIMATION

Section 3.5 discusses how to determine whether the Poisson process is a good model for an observed arrival process. Suppose that the arrival process passes the test of plausibility, passes the graphical tests, and passes the statistical tests, as do the teller data. There is still one more task to undertake before the model is complete: ***parameter estimation***. The Poisson process is not a single model but actually a family of models, defined by different values of the parameter λ. The model is complete when the value of λ is estimated.

Parameter estimation is somewhat different from assessing goodness of fit because it is not a matter of answering a simple yes or no question. λ can be any real positive number, and there usually is no a priori reason to believe that λ should be any particular value. So parameter estimation normally does not appeal to one's knowledge of the underlying arrival process. Rather, it depends almost exclusively on analysis of the data. Along these lines, the accuracy of a parameter estimate is a matter of degree rather than simply true or false. If the parameter estimate is 10, the true value (the value of λ that created the data) might be 10.1, 9.5, or even 13.6. This is because the parameter estimate, or the ***estimator***, is itself a random variable.

Parameter estimation is the process of determining the best estimate of a distribution's parameter or parameters, given the observed data. Usually, this means that the parameter estimate should be precise (that is, the standard deviation of the estimator is small) and the parameter estimate is ***unbiased*** (that is, the expected value of the estimator equals the true vale of the parameter).

The simplest estimation technique is called the ***method of moments (MOM)***, which amounts to equating sample moments to population moments, such as $E(X)$ or $E(X^2)$. The method of moments was used in the previous section to estimate λ, the arrival rate. An estimator is denoted by the caret symbol. Thus, $\hat{\lambda}$ is the estimate for the true value of λ. For the method of moments

$$\hat{\lambda} = \frac{A(T)}{T} \tag{3.24}$$

For the teller data, $\hat{\lambda}$ equals $98/120 = .817$. This estimator is the guess for the true value of λ, the value of λ that actually created the data.

A more rigorous technique for estimating parameters is the ***maximum likelihood method.*** This approach determines the parameter values that are most likely to generate the observed data sample. The maximum likelihood estimator (MLE) tends to be more precise than the method of moments estimator. However, determining the MLE may be more difficult, partly because it depends on the probability distribution for the random variable.

There are two ways to determine the MLE for λ. The first is based on the Poisson distribution for the number of events, and the second is based on the exponential distribution for the interarrival times.

Definition 3.10

The ***likelihood function,*** $L(\textbf{data} \mid \lambda)$, specifies the probability density function for the observed data, given that the parameter equals λ.

Suppose that $N = A(T)$ arrivals are observed over a time interval of length T. Then the likelihood of observing these data is defined by the Poisson probability function as follows:

$$L(N \text{ arrivals} \mid \lambda) = \frac{(\lambda T)^N}{N!} e^{-\lambda T} \tag{3.25}$$

The MLE is the value of λ that maximizes the likelihood function in Eq. (3.25). For the particular expression above, the MLE is found by determining the point where the derivative *with respect to the parameter* λ equals 0:

$$\frac{\partial L}{\partial \lambda} = N \frac{\lambda^{N-1} T^N}{N!} e^{-\lambda T} - T \frac{(\lambda T)^N}{N!} e^{-\lambda T} = 0$$

$$[N/\lambda - T] \frac{(\lambda T)^N}{N!} e^{-\lambda T} = 0$$

$$\hat{\lambda} = \frac{N}{T} = \frac{A(T)}{T} \tag{3.26}$$

For the Poisson process, the method of moments estimator and the Poisson distribution MLE estimator happen to be the same. This will not be the case for all estimators, such as the estimator for the standard deviation of a normal distribution.

The second likelihood function, based on the exponential distribution, requires more data than the first, because it depends on the interarrival times, not just the number of events. It also differs because X_n is a continuous, rather than a discrete, random variable. Because the likelihood function is a probability density function, it is not a true probability for a continuous random variable. As before, let X_n represent the nth interarrival time. Then

$$L(X_1, \ldots, X_n \mid \lambda) = \left[\prod_{n=1}^{N} \lambda e^{-\lambda X_n} \right] e^{-\lambda(T - \Sigma X_n)} \tag{3.27}$$

The first term is the likelihood associated with the first N interarrival times. The second term is the likelihood associated with the last interval, length $T - \Sigma X_n$, during which no arrival occurred. Equation (3.27) can be simplified to

$$L(X_1, \ldots, X_N \mid \lambda) = \lambda^N e^{-\lambda T} \tag{3.28}$$

Surprisingly, the likelihood reduces to a function of just two observations, N and T, not the entire set of interarrival times. As before, the estimator is found by taking the derivative *with respect to the parameter* λ. It should come as no surprise that the MLE estimator is the same as the Poisson distribution MLE: N/T.

Both the method of moments and the maximum likelihood method are general techniques for estimating distribution parameters from a data set and can be used for virtually any probability distribution (not just Poisson or exponential). For the Poisson process, the two techniques happen to yield the same result. In general, the MLE is guaranteed to be the more precise estimator for large data sets. However, in choosing an estimator, the added precision (which tends to be small) must be weighed against the effort needed to determine the estimator.

3.6.1 Confidence Intervals

The accuracy of a parameter estimate can be measured by way of a confidence interval. While the parameter estimate is the best guess for λ, the ***confidence interval*** specifies a range of plausible values for λ. If this range is large, then one has little confidence that the estimate is correct. This situation can only be corrected by collecting more data and/or changing the method for collecting the data.

The confidence interval is based on the ***standard error*** of the parameter being estimated, which is just another name for the standard deviation of the estimator. The standard error for the mean of a sample of N *independent* random variables ($\bar{X}$) equals

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}} \tag{3.29}$$

where σ is the standard deviation of the random variable.

The standard error for the statistic $\bar{X}$ is smaller than the standard deviation for the random variable, X, because averaging reduces fluctuations. As N becomes large, fluctuations in the average become smaller and $\sigma_{\bar{X}}$ approaches zero.

Just as the sample mean has an estimator, so does the standard deviation. Usually this is defined by the following:

$$\hat{\sigma} = \sqrt{\frac{\sum_{n=1}^{N} (X_n - \bar{X})^2}{N - 1}} \tag{3.30}$$

To provide an unbiased estimate of σ, $N - 1$, not N, appears in the denominator (note that this in not a MOM estimator).

Recall that a property of the exponential distribution is that the mean equals the standard deviation. For the exponential distribution, it turns out that the best estimator (that is, the MLE) for σ is not the sample standard deviation; the best estimator for σ is the sample mean (T/N):

Exponential Random Variable

$$\hat{\sigma}_{\bar{X}} = \frac{\hat{\sigma}}{\sqrt{N}} = \frac{\mathrm{T/N}}{\sqrt{N}} = T \cdot N^{-1.5} \tag{3.31}$$

The confidence interval is always specified for a given level of confidence. For example, a 95% confidence interval means that there is a 95% probability that the true value of the parameter falls inside the confidence interval. To determine the confidence interval, one usually refers to the normal probability distribution. From the ***central limit theorem***, we know that probability distribution for a sum of n independent, identically distributed random variables will approach the normal distribution as n becomes large. And because the limiting distribution for the sum is normal, the limiting distribution for the ***sample mean*** (that is, average) must also be normal.

In terms of the arrival process, the confidence interval for $1/\lambda$ (mean of exponential distribution) can be derived from tables of the normal distribution (see Table A.2 in the appendix to this book) if the sample is large (greater than 50 data points). The normal distribution does not apply to λ directly because it is not a sample mean.

$$P(\widehat{1/\lambda} - 1.96\hat{\sigma}_{\bar{X}} \leq 1/\lambda \leq \widehat{1/\lambda} + 1.96\hat{\sigma}_{\bar{X}}) = .95 \tag{3.32}$$

$$P(\widehat{1/\lambda} - 2.58\hat{\sigma}_{\bar{X}} \leq 1/\lambda \leq \widehat{1/\lambda} + 2.58\hat{\sigma}_{\bar{X}}) = .99 \tag{3.33}$$

The symbol $1/\lambda$ represents the true mean of the random variable; $\widehat{1/\lambda}$ and $\hat{\sigma}_{\bar{X}}$ represent the estimators for the mean and mean standard error, based on the observed data.

The first line gives the 95% confidence region and the second gives the 99% confidence region. Both regions are symmetric around $\widehat{1/\lambda}$, indicating that $\widehat{1/\lambda}$ could be either be larger or smaller than the true value. More specifically, the confidence region for $1/\lambda$ in a Poisson process can be derived from the exponential distribution:

$$P[T/N - 1.96(T \cdot N^{-1.5}) \leq 1/\lambda \leq T/N + 1.96(T \cdot N^{-1.5})] = .95 \quad (3.34)$$

$$P[T/N - 2.58(T \cdot N^{-1.5}) \leq 1/\lambda \leq T/N + 2.58(T \cdot N^{-1.5})] = .99 \quad (3.35)$$

Example

For the automatic teller machine data in Table 3.1, $T/N = 120/98 = 1.22$ minutes and $N = 98$. Thus, the confidence intervals are:

95% confidence $1.22 - 1.96\,(1.22/\sqrt{98}) \leq 1/\lambda \leq 1.22 + 1.96\,(1.22/\sqrt{98})$

$.98 \leq 1/\lambda \leq 1.46$

99% confidence $1.22 - 2.58\,(1.22/\sqrt{98}) \leq 1/\lambda \leq 1.22 + 2.58\,(1.22/\sqrt{98})$

$.90 \leq 1/\lambda \leq 1.54$

Based on the data sample, there is a 95% probability that the true value of $1/\lambda$ (the value that created the data) is between .98 and 1.46 and a 99% probability that the true value of $1/\lambda$ is between .90 and 1.54. The 95% and 99% confidence intervals for λ are found by inverting these bounds, (1.02, .68) and (1.11, .65), respectively. These confidence intervals are not particularly small, which suggests that more data need to be collected to obtain a more precise estimate.

3.6.2 Sample Size

One final issue needs to be addressed: sample size. The sample should be sufficiently large to provide a precise estimate of the model parameters. But data collection is expensive, and no more data should be collected than necessary.

A reasonable approach for selecting the sample size is to begin by specifying a desired level of accuracy and then calculate how long the process must be observed in order to obtain the desired level of accuracy. For the teller example, one might specify that the 95% confidence interval should have a width of no more than .2 minute. This means that

$$P(|\widehat{1/\lambda} - 1/\lambda| \leq .1) \geq .95 \quad (3.36)$$

The width of the 95% confidence interval is 2×1.96 multiplied by the standard error for $\widehat{1/\lambda}$ (the number 2 accounts for the two tails of the distribution). This quantity must be less than or equal to the desired width, .2 minute.

$$2\,[1.96\; 1/(\lambda\sqrt{N})] \leq .2 \quad (3.37)$$

Simplifying the equation leads to

$$N \geq \left[\frac{2 \cdot 1.96}{.2\lambda}\right]^2 = \frac{384}{\lambda^2} \quad (3.38)$$

From the teller data in Table 3.1, $1/\lambda$ is estimated to be 1.22 minutes. Substituting this value in Eq. (3.38) yields 572 arrivals (11.6 hours). This is approximately how many observations would be needed to obtain a 95% confidence interval of width .2. Fortunately, the arrival rate is very easy to measure, and only involves counting arrivals. So observing the process for 11.6 hours may not be difficult. In general, if

$$w = \text{desired width of confidence interval (in time units)}$$

the desired accuracy requires a sample size of

95% Confidence	**99% Confidence**	
$N \geq \left[\dfrac{3.92}{w\lambda}\right]^2$	$N \geq \left[\dfrac{5.16}{w\lambda}\right]^2$	(3.39)

Of course, λ is not known until the sample is taken. But one usually can roughly guess the arrival rate beforehand. This prior guess can be the basis for the sample size calculation. The calculation might also be based on a "presample"—a sample over a short time interval used to obtain a preliminary guess for the parameter. In either case, determining the sample size does not require absolute precision.

Example

Twelve arrivals are observed over a 1-hour period. The sample size that would produce a 99% confidence interval of width .05 hour is desired.

Solution Substituting the value .05 hour for w and 12/hour for λ, we get a sample size of 74 arrivals. This translates into 6.2 hours of observation.

There is no right way to set the width of the confidence interval or the confidence level. Both the cost of collecting the data and the need for precision affect the answer.

One final comment: The sample size calculations presume that it is possible to observe the arrival process for the required length of time. But what if the process is short-lived, lasting only one or two hours? It would then be impossible to observe the process long enough to obtain the desired precision. But if the process is observed over its entire lifetime, then there is no need to collect more data. The data already collected should accurately depict what happened, and since the process has ended, there is no need to infer what will happen in the future.

3.7 CHAPTER SUMMARY

A model is a way to explain the underlying process creating the data. Sometimes the model is a perfect representation of the underlying process, as in the binomial distribution representing flipped coins. More often, the model is a plausible representation that closely matches observations.

The Poisson process represents the customer arrival process of systems for which

1. The probability of a customer arriving at any time does not depend on when other customers arrived.

2. The probability that a customer arrives at any time does not depend on the time.
3. Customers arrive one at a time.

The Poisson process is a particularly important model because it accurately represents many real arrival processes and because it is fairly simple to analyze. The Poisson process possesses a number of unique properties summarized at the end of Sec. 3.3.

Determining whether or not an arrival process is Poisson always begins with an assessment of plausibility. The most important question is whether the conditions underlying the Poisson process accurately represent the situation. If they do, then further quantitative tests—graphical and statistical—can be performed to see whether the data displays the properties of the model. In the case of the Poisson process, those properties included stationarity, exponential interarrival times, and independent interarrival times. However, only a few of the many possible statistical tests were provided. For a more complete survey, consult Bhat (1978), Green and Kolesar (1989), or one of the statistics texts cited at the end of this chapter.

If the model passes the plausibility test and quantitative tests, then the next step is to determine the exact shape of the model through parameter estimation. Once this is done, a confidence interval can be calculated to estimate the precision of the parameter estimator. If the precision is not sufficiently accurate, further data should be collected. The exact amount of data to collect can be calculated with the method provided for determining sample size in Sec. 3.6.2.

Keep in mind that just because a process is Poisson with rate λ today does not mean that it will be Poisson with rate λ tomorrow, or at any other time. Predicting the future must always rely on inference—inference as to whether the conditions that created the historical data will recur in the future. The subject of prediction is addressed in the following chapter.

Although this chapter focuses on the Poisson process, the general steps of

1. Assessing the conditions creating the data
2. Gauging model plausibility
3. Testing data for goodness of fit
4. Estimating parameters and confidence intervals

apply to most any situation. The model does not have to be Poisson, but it does have to be a reasonable representation of reality. Just what this representation should be varies from situation to situation.

FURTHER READING

Allen, A. O. 1979. *Probability, Statistics and Queueing Theory*, New York: Academic Press.

Bhat, U. N. 1978. "Theory of Queues." in *Handbook of Operations Research, Foundations and Fundamentals*, ed. J. J. Moder and S. E. Elmaghraby. New York: Van Nostrand Reinhold.

GREEN, L., and P. KOLESAR. 1989. "Testing the Validity of a Queueing Model of Police Patrol," *Management Science*, 35, 127–148.

GREENSHIELDS, B. D., and F. M. WEIDA. 1978. *Statistics with Applications to Highway Traffic Analysis*. Westport, Connecticut: Eno Foundation for Transportation.

HAIGHT, F. A. 1967. *Handbook of the Poisson Distribution*, New York: John Wiley.

HAYS, W. L., and R. L. WINKLER. 1971. *Statistics, Probability, Inference and Decision*, New York: Holt, Rinehart, Winston.

MOOD, A. M., F. A. GRAYBILL, and D. C. BOES. 1974. *Introduction to the Theory of Statistics*, New York: McGraw-Hill.

NEWELL, G. F. 1982. *Applied Queueing Theory*, London: Chapman and Hall.

ROSS, S. M. 1972. *Introduction to Probability Models*, New York: Academic Press.

TUFTE, E. R. 1988. *The Visual Display of Quantitative Information*, Cheshire, Conn.: Graphics Press.

WONNACOTT, T. H., and R. J. WONNACOTT. 1969. *Introductory Statistics*, New York: John Wiley.

NOTE

1. KUHN, T. THOMAS, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1970), p. 55.

PROBLEMS

Probability Distributions

1. In each of 10 minutes, the probability of exactly one arrival equals .1 and the probability of no arrivals equals .9.

- **(a)** Using the binomial distribution, calculate the probability of 0 arrivals over 10 minutes. Repeat for 1, 2, 3, and 4 arrivals.
- **(b)** Now assume that arrivals occur by a Poisson process at the rate of .1 per minute. Repeat your calculations for part a, based on the Poisson distribution. Why are your results similar, or different?

***2.** Suppose that an arrival comprises either a pair of customers or a single customer. The probability that any arrival has 2 customers equals .5 and the probability that any arrival has 1 customer equals .5. In each of 10 minutes, the probability of exactly one arrival equals .0667 and the probability of no arrivals equals .9333.

- **(a)** Calculate the probability that 0 customers arrive over 10 minutes. Repeat for 1, 2, and 3 customers. (Hint: For any number of customers n, sum the probabilities of n customers given 1 arrival, n customers given 2 arrivals, ...)
- **(b)** Why are your results similar, or different, from those in part b of Prob. 1?

3. Telephone calls are known to arrive by a Poisson process with rate 20 per hour between 1:00 and 3:00 P.M. Determine the following:

- **(a)** The probability function for the number of arrivals in a 5-minute interval (for up to 5 arrivals).

*Difficult problem

(b) Probability that no customer arrives over a 10-minute interval.
(c) Probability that the second arrival after 1:00 occurs before 1:10.
(d) Given that three arrivals occurred between 1:00 and 1:30, the probability that the second arrival occurred after 1:15.

4. The chancellor of a university has determined that student complaints arrive by a Poisson process, with a rate of 25 per year. Determine the following:
(a) Probability that one month passes with no more than one complaint received (assume that a month is 1/12 of a year).
(b) The probability that the first complaint of the year occurs during the second month.
(c) The probability that the third complaint of the year occurs during March (use the gamma distribution).

5. Based on Definition 3.2 for the Poisson process, prove that the variance for the Poisson distribution equals λt. (Hint: Derive the result from the variance of a binomial random variable.)

*6. Individual customers arrive by a Poisson process with rate of .0333 per minute, and pairs of customers also arrive by a Poisson process, with the same rate. Determine the probability function for the number of customers arriving during a 10-minute period, for 0 to 3 customers. Compare your result to the calculations in Prob. 2. Explain why your answer is the same, or different.

7. The Poisson distribution is discrete, but the gamma and exponential distributions are continuous. Is this inconsistent, given that all three define properties of the Poisson process? Explain.

8. A desperate gambler has flown to Lake Tahoe to win his fortune. He has chosen a $1 slot machine, which will pay a prize of $1 million, with a probability of 1/2,000,000 (success is independent among tries).
(a) Assuming that there is only one possible prize, what is the probability of winning before 500,000 tries? (Approximate from Poisson process.)
(b) Explain why your answer to part a is *not* 500,000/2,000,000.
(c) Suppose that the gambler knows that the machine has not paid out in the last 2 million tries. Will this information affect your answer to part a? Explain.

9. From any familiar application, give examples of three continuous random variables and three discrete random variables. Based on your intuition alone, plot the probability distribution function for each example. Which, if any, of your examples conforms to a familiar theoretical distribution?

10. Collect ten observations of any continuous random, then ten observations of any discrete random variable. Plot the empirical distribution function for each. By examining these distribution functions, can you tell which random variable was discrete and which was continuous? (If possible, show these functions to a classmate, and see whether he or she can tell which is discrete and which is continuous.)

Goodness of Fit

*11. The data below represent the times that people stopped to deliver letters at a mailbox over a 255-minute period.

3.1	48.6	117.0	177.5
6.2	50.8	123.3	183.0
6.5	72.2	128.8	195.0

*Difficult problem

10.1	76.0	131.1	200.0
13.9	78.2	145.7	204.2
29.3	90.3	147.7	207.6
34.4	91.5	150.7	207.9
35.2	104.2	156.2	239.7
39.2	113.1	162.3	240.2
45.0	114.4	169.1	251.6

(a) Is it plausible that the arrival process of people is Poisson?
(b) Is it plausible that the arrival process of letters is Poisson?
(c) Derive an MOM estimate for the arrival rate of people. Calculate a 95% confidence interval for your estimate. Is your estimate also a maximum likelihood estimator? Briefly discuss.
(d) Plot the empirical probability distribution function for the interarrival times. On the same graph, plot the distribution function for an exponential distribution with mean defined by part c. Perform the K-S goodness of fit test. Do you believe the interarrival times are exponential random variables?
(e) Plot successive interarrival times, as in Fig. 3.4. Then perform the t-test for correlation. Do you believe that the interarrival times are independent?
(f) Plot the cumulative arrival curve. On the same graph, plot a straight line connecting $A(T)$ and $A(0)$. Next, perform the K-S test for goodness of fit. Do you believe that the arrival process is stationary?
(g) Based on all the evidence collected, do you believe the arrival process is Poisson?

***12.** The data below are the arrival times of westbound BART trains at the Montgomery Street station in San Francisco, between time 6.85 A.M. and time 8.85 A.M. (in hours).

6.861	6.926	7.063	7.159	7.208	7.271	7.356	7.415	7.484	7.532	7.613
7.665	7.739	7.793	7.860	7.905	7.965	8.060	8.110	8.155	8.237	8.296
8.359	8.403	8.489	8.569	8.640	8.732	8.818				

(a) Is it plausible that the arrival of trains is a Poisson process?
(b) Is it plausible that the arrival of people is a Poisson process?
(c–g) Repeat parts c–g from Prob. 11 for the new data set.

13. BART trains are scheduled to arrive every 3.75 minutes at the Montgomery Street Station, beginning at time 7.35 hours and ending at time 8.48 hours. Based on this information and the data given in Prob. 12, develop a new model for the arrival process (other than the Poisson process). That is, define a probability distribution function for the arrival time of the nth train.

14. For each of the three examples in Sec. 3.4 that do not conform to the Poisson process, describe a mathematical model that represents the arrival process.

15. Using the data from Prob. 11, estimate the minimum time that the system would have to be observed to obtain a 95% confidence interval for $1/\lambda$ of width no more than 30 seconds.

*Difficult problem

EXERCISE: GOODNESS OF FIT

The purpose of this exercise is to test the data recorded in Chap. 2 to see whether a Poisson process was observed.

1. Discuss whether the arrival process matches the conditions underlying the Poisson process, referring to the three properties:
   (a) Stationarity
   (b) Independent increments
   (c) Customers arrive one at a time.

   If the process does not precisely match the Poisson process, discuss whether the differences are appreciable.
2. As a quick check, calculate the coefficient of variation for the interarrival times. Does the arrival process seem to be Poisson?
3. On a graph of cumulative arrivals, draw a straight line connecting $A(T)$ to $A(0)$. Is there any pattern in the difference between the two curves? Discuss whether the process appears to be stationary.
4. On a graph for the empirical probability distribution for interarrival times, drawn an exponential distribution function with identical mean. Is there any pattern in the difference between the two curves? Discuss whether the interarrival times appear to be exponential.
5. Plot the paired intervals (X_n, X_{n-1}) on graph paper. Is there any pattern to the data? Discuss whether the data appear to be independent.
6. Perform the following statistical tests at the 5% significance level:
   (a) Kolmogorov-Smirnov test for stationarity of arrivals
   (b) Kolmogorov-Smirnov test for an exponential interarrival distribution
   (c) Correlation t-test between X_n and X_{n-1}

   Can you conclude that the arrival process was Poisson?
7. Calculate 95% and 99% confidence intervals for $1/\lambda$. Convert these intervals into 95% and 99% confidence intervals for λ.
8. Determine how long (in minutes) the process would have to be observed to obtain a 95% confidence interval for $1/\lambda$ of width $(1/\lambda)/20$. (Remember that the confidence interval extends on both sides of $1/\lambda$).