15. Poisson Processes

In Lecture 4, we introduced Poisson arrivals as the limiting behavior of Binomial random variables. (Refer to Poisson approximation of Binomial random variables.)
From the discussion there (see (4-6)-(4-8) Lecture 4)

$$P\left\{\begin{array}{c}\text{"}k \text{ arrivals occur in an} \\ \text{interval of duration } \Delta\text{"}\end{array}\right\} = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \cdots \tag{15-1}$$

where

$$\lambda = np = \mu T \cdot \frac{\Delta}{T} = \mu\Delta \tag{15-2}$$

Fig. 15.1

It follows that (refer to Fig. 15.1)

$$P\left\{\begin{array}{l}\text{"}k \text{ arrivals occur in an} \\ \text{ interval of duration } 2\Delta\text{"}\end{array}\right\} = e^{-2\lambda}\frac{(2\lambda)^k}{k!}, \qquad k = 0,\ 1,\ 2,\cdots, \quad (15\text{-}3)$$

since in that case

$$np_1 = \mu T \cdot \frac{2\Delta}{T} = 2\mu\Delta = 2\lambda. \quad (15\text{-}4)$$

From (15-1)-(15-4), Poisson arrivals over an interval form a Poisson random variable whose parameter depends on the duration of that interval. Moreover because of the Bernoulli nature of the underlying basic random arrivals, events over nonoverlapping intervals are independent. We shall use these two key observations to define a Poisson process formally. (Refer to Example 9-5, Text)

Definition: $X(t) = n(0, t)$ represents a Poisson process if

(i) the number of arrivals $n(t_1, t_2)$ in an interval (t_1, t_2) of length $t = t_2 - t_1$ is a Poisson random variable with parameter λt. Thus

$$P\{n(t_1,t_2)=k\}=e^{-\lambda t}\frac{(\lambda t)^k}{k!}, \quad k=0,1,2,\cdots, \quad t=t_2-t_1 \tag{15-5}$$

and

(ii) If the intervals (t_1, t_2) and (t_3, t_4) are nonoverlapping, then the random variables $n(t_1, t_2)$ and $n(t_3, t_4)$ are independent.

Since $n(0, t) \sim P(\lambda t)$, we have

$$E[X(t)]=E[n(0,t)]=\lambda t \tag{15-6}$$

and

$$E[X^2(t)]=E[n^2(0,t)]=\lambda t+\lambda^2 t^2. \tag{15-7}$$

To determine the autocorrelation function $R_{XX}(t_1,t_2)$, let $t_2 > t_1$, then from (ii) above $n(0, t_1)$ and $n(t_1, t_2)$ are independent Poisson random variables with parameters λt_1 and $\lambda(t_2-t_1)$ respectively. Thus

$$E[n(0,t_1)n(t_1,t_2)]=E[n(0,t_1)]E[n(t_1,t_2)]=\lambda^2 t_1(t_2-t_1). \tag{15-8}$$

But

$$n(t_1, t_2) = n(0, t_2) - n(0, t_1) = X(t_2) - X(t_1)$$

and hence the left side if (15-8) can be rewritten as

$$E[X(t_1)\{X(t_2) - X(t_1)\}] = R_{XX}(t_1, t_2) - E[X^2(t_1)]. \tag{15-9}$$

Using (15-7) in (15-9) together with (15-8), we obtain

$$\begin{aligned} R_{XX}(t_1, t_2) &= \lambda^2 t_1 (t_2 - t_1) + E[X^2(t_1)] \\ &= \lambda t_1 + \lambda^2 t_1\, t_2, \quad t_2 \geq t_1\,. \end{aligned} \tag{15-10}$$

Similarly

$$R_{XX}(t_1, t_2) = \lambda t_2 + \lambda^2 t_1\, t_2, \quad t_2 < t_1\,. \tag{15-11}$$

Thus

$$R_{XX}(t_1, t_2) = \lambda^2 t_1\, t_2 + \lambda \min(t_1, t_2). \tag{15-12}$$

From (15-12), notice that the Poisson process $X(t)$ *does not* represent a wide sense stationary process.

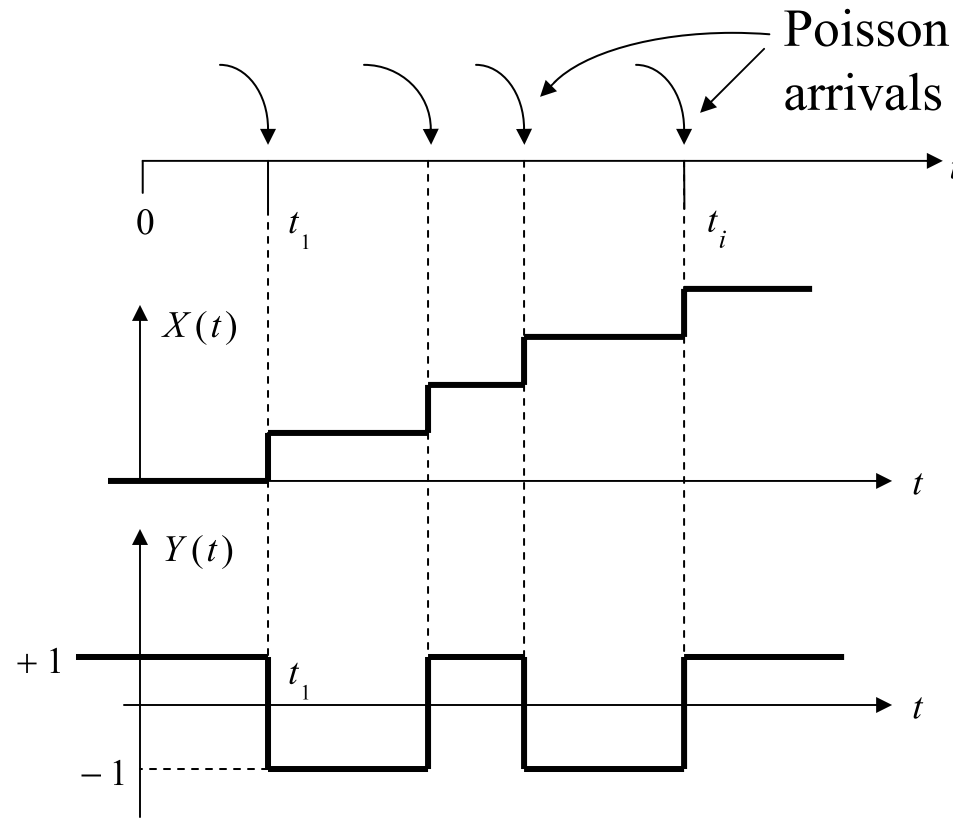


Fig. 15.2

Define a binary level process

$$Y(t) = (-1)^{X(t)} \tag{15-13}$$

that represents a telegraph signal (Fig. 15.2). Notice that the transition instants $\{t_i\}$ are random. (see Example 9-6, Text for the mean and autocorrelation function of a telegraph signal). Although $X(t)$ does not represent a wide sense stationary process, its derivative $X'(t)$ *does* represent a wide sense stationary process.

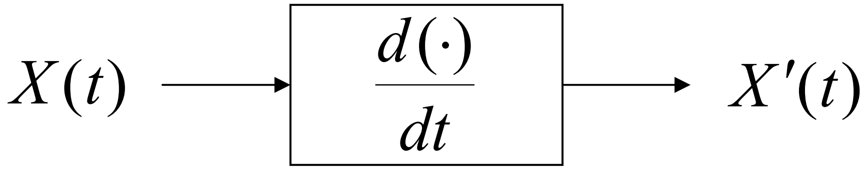


Fig. 15.3 (Derivative as a LTI system)

To see this, we can make use of Fig. 14.7 and (14-34)-(14-37). From there

$$\mu_{X'}(t) = \frac{d\mu_X(t)}{dt} = \frac{d\lambda t}{dt} = \lambda, \quad \textit{a constant} \tag{15-14}$$

and

$$R_{XX'}(t_1, t_2) = \frac{\partial R_{XX}(t_1, t_2)}{\partial t_2} = \begin{cases} \lambda^2 t_1 & t_1 \le t_2 \\ \lambda^2 t_1 + \lambda & t_1 > t_2 \end{cases}$$

$$= \lambda^2 t_1 + \lambda \; U(t_1 - t_2) \tag{15-15}$$

and

$$R_{X'X'}(t_1, t_2) = \frac{\partial R_{XX'}(t_1, t_2)}{\partial t_1} = \lambda^2 + \lambda \, \delta(t_1 - t_2). \tag{15-16}$$

From (15-14) and (15-16) it follows that $X'(t)$ is a wide sense stationary process. Thus nonstationary inputs to linear systems *can* lead to wide sense stationary outputs, an interesting observation.

- **Sum of Poisson Processes:**

  If $X_1(t)$ and $X_2(t)$ represent two independent Poisson processes, then their sum $X_1(t) + X_2(t)$ is also a Poisson process with parameter $(\lambda_1 + \lambda_2)t$. (Follows from (6-86), Text and the definition of the Poisson process in (i) and (ii)).

- **Random selection of Poisson Points:**

  Let $t_1, t_2, \cdots, t_i, \cdots$ represent random arrival points associated with a Poisson process $X(t)$ with parameter λt, and associated with each arrival point, define an independent Bernoulli random variable N_i, where

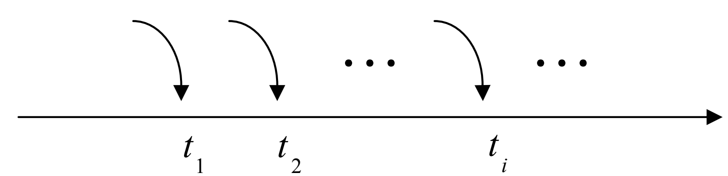


Fig. 15.4

$$P(N_i = 1) = p, \qquad P(N_i = 0) = q = 1 - p. \qquad (15\text{-}17)$$

Define the processes

$$Y(t)=\sum_{i=1}^{X(t)} N_i \quad ; \quad Z(t)=\sum_{i=1}^{X(t)}(1-N_i)=X(t)-Y(t) \qquad (15\text{-}18)$$

we claim that both $Y(t)$ and $Z(t)$ are independent Poisson processes with parameters λpt and λqt respectively.

Proof:

$$Y(t)=\sum_{n=k}^{\infty} P\{Y(t)=k \mid X(t)=n\} P\{X(t)=n)\}. \qquad (15\text{-}19)$$

But given $X(t) = n,$ we have $Y(t)=\sum_{i=1}^{n} N_i \ \sim \ B(n,p)$ so that

$$P\{Y(t)=k \mid X(t)=n\}=\binom{n}{k} p^k q^{n-k}, \qquad 0 \le k \le n, \qquad (15\text{-}20)$$

and

$$P\{X(t)=n\}=e^{-\lambda t}\frac{(\lambda t)^n}{n!}. \qquad (15\text{-}21)$$

Substituting (15-20)-(15-21) into (15-19) we get

$$P\{Y(t)=k\} = e^{-\lambda t} \sum_{n=k}^{\infty} \frac{n!}{(n-k)!k!} p^k q^{n-k} \frac{(\lambda t)^n}{n!} = \frac{p^k e^{-\lambda t}}{k!} (\lambda t)^k \underbrace{\sum_{n=k}^{\infty} \frac{(q\lambda t)^{n-k}}{(n-k)!}}_{e^{q\lambda t}}$$

$$= (\lambda pt)^k \frac{e^{-(1-q)\lambda t}}{k!} = e^{-\lambda pt} \frac{(\lambda pt)^k}{k!}, \qquad k = 0, 1, 2, \cdots$$

$$\sim \quad P(\lambda pt). \tag{15-22}$$

More generally,

$$P\{Y(t)=k, Z(t)=m\} = P\{Y(t)=k, X(t)-Y(t)=m\}$$
$$= P\{Y(t)=k, X(t)=k+m\}$$
$$= P\{Y(t)=k \mid X(t)=k+m\} P\{X(t)=k+m\}$$
$$= \binom{k+m}{k} p^k q^m \cdot e^{-\lambda t} \frac{(\lambda t)^{k+m}}{(k+m)!} = \underbrace{e^{-\lambda pt} \frac{(\lambda pt)^n}{k!}}_{P(Y(t)=k)} \underbrace{e^{-\lambda qt} \frac{(\lambda qt)^n}{m!}}_{P(Z(t)=m)}$$
$$= P\{Y(t)=k\} P\{Z(t)=m\}, \tag{15-23}$$

which completes the proof.

Notice that $Y(t)$ and $Z(t)$ are generated as a result of random Bernoulli selections from the original Poisson process $X(t)$ (Fig. 15.5), where each arrival gets tossed over to either $Y(t)$ with probability p or to $Z(t)$ with probability q. Each such sub-arrival stream is also a Poisson process. Thus random selection of Poisson points preserve the Poisson nature of the resulting processes. However, as we shall see deterministic selection from a Poisson process destroys the Poisson property for the resulting processes.

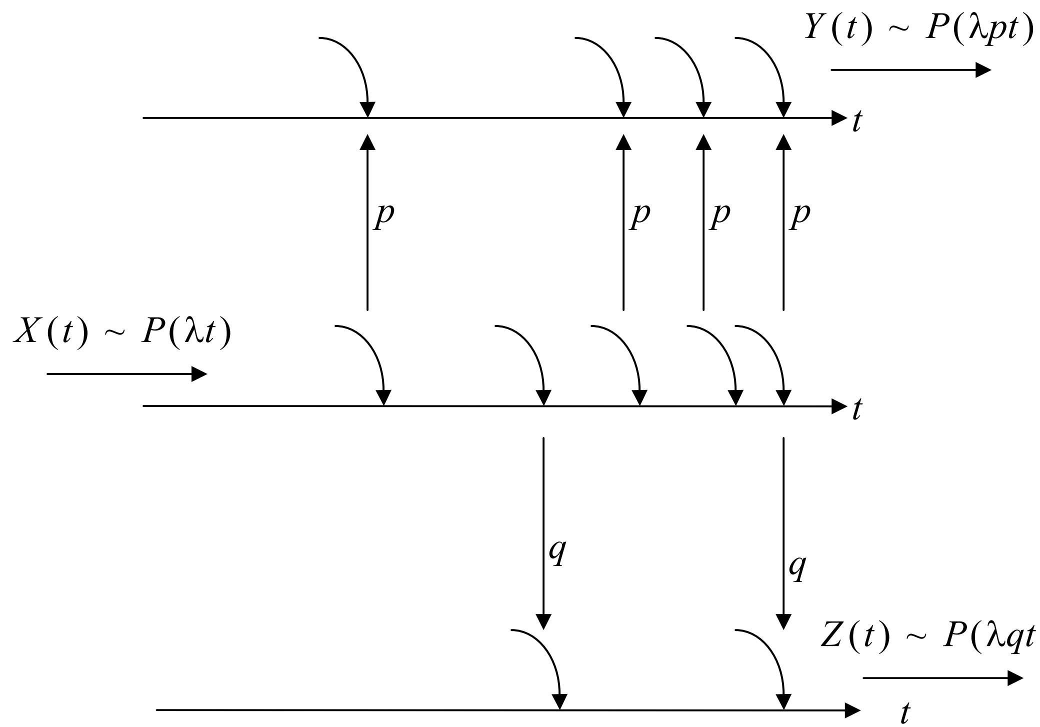


Fig. 15.5

Inter-arrival Distribution for Poisson Processes

Let τ_1 denote the time interval (delay) to the first arrival from *any* fixed point t_0. To determine the probability distribution of the random variable τ_1, we argue as follows: Observe that the event "$\tau_1 > t$" is the same as "$n(t_0, t_0+t) = 0$", or the complement event "$\tau_1 \le t$" is the same as the event "$n(t_0, t_0+t) > 0$".

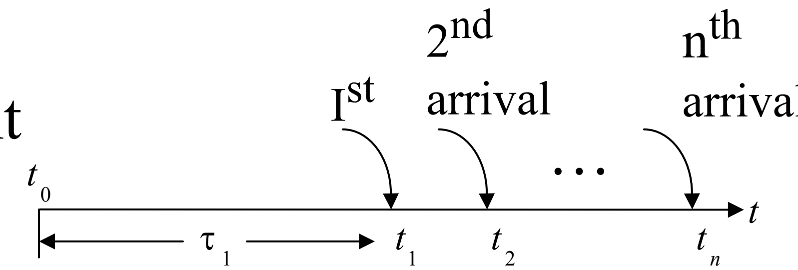


Fig. 15.6

Hence the distribution function of τ_1 is given by

$$F_{\tau_1}(t) \triangleq P\{\tau_1 \le t\} = P\{X(t) > 0\} = P\{n(t_0, t_0 + t) > 0\}$$

$$= 1 - P\{n(t_0, t_0 + t) = 0\} = 1 - e^{-\lambda t} \tag{15-24}$$

(use (15-5)), and hence its derivative gives the probability density function for τ_1 to be

$$f_{\tau_1}(t) = \frac{dF_{\tau_1}(t)}{dt} = \lambda e^{-\lambda t}, \qquad t \ge 0 \tag{15-25}$$

i.e., τ_1 is an exponential random variable with parameter λ so that $E(\tau_1) = 1/\lambda$.

Similarly, let t_n represent the n^{th} random arrival point for a Poisson process. Then

$$F_{t_n}(t) \triangleq P\{t_n \le t\} = P\{X(t) \ge n\}$$

$$= 1 - P\{X(t) < n\} = 1 - \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t} \qquad (15\text{-}26)$$

and hence

$$f_{t_n}(x) = \frac{dF_{t_n}(x)}{dx} = -\sum_{k=1}^{n-1} \frac{\lambda(\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x} + \sum_{k=0}^{n-1} \frac{\lambda(\lambda x)^k}{k!} e^{-\lambda x}$$

$$= \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x}, \qquad x \ge 0 \qquad (15\text{-}27)$$

which represents a gamma density function. i.e., the waiting time to the n^{th} Poisson arrival instant has a gamma distribution. Moreover

$$t_n = \sum_{i=1}^{n} \tau_i$$

where τ_i is the random inter-arrival duration between the $(i-1)^{th}$ and i^{th} events. Notice that τ_i s are independent, identically distributed random variables. Hence using their characteristic functions, it follows that all inter-arrival durations of a Poisson process are independent exponential random variables with common parameter λ. i.e.,

$$f_{\tau_i}(t) = \lambda\, e^{-\lambda t}, \qquad t \geq 0. \tag{15-28}$$

Alternatively, from (15-24)-(15-25), we have τ_1 is an exponential random variable. By repeating that argument after shifting t_0 to the new point t_1 in Fig. 15.6, we conclude that τ_2 is an exponential random variable. Thus the sequence $\tau_1, \tau_2, \cdots, \tau_n, \cdots$ are independent exponential random variables with common p.d.f as in (15-25).

Thus if we systematically tag every m^{th} outcome of a Poisson process $X(t)$ with parameter λt to generate a new process $e(t)$, then the inter-arrival time between any two events of $e(t)$ is a gamma random variable.

Notice that

$$E[e(t)] = m/\lambda, \text{ and if } \lambda = m\mu, \text{ then } E[e(t)] = 1/\mu.$$

The inter-arrival time of $e(t)$ in that case represents an Erlang-m random variable, and $e(t)$ an Erlang-m process (see (10-90), Text). In summary, if Poisson arrivals are randomly redirected to form new queues, then each such queue generates a new Poisson process (Fig. 15.5). However if the arrivals are systematically redirected (Ist arrival to Ist counter, 2nd arrival to 2nd counter, $\cdots$, m^{th} to m^{th} , $(m+1)^{st}$ arrival to Ist counter, $\cdots$), then the new subqueues form Erlang-m processes.

Interestingly, we can also derive the key Poisson properties (15-5) and (15-25) by starting from a simple axiomatic approach as shown below:

Axiomatic Development of Poisson Processes:

The defining properties of a Poisson process are that in any "small" interval Δt, one event can occur with probability that is proportional to Δt. Further, the probability that two or more events occur in that interval is proportional to $o(\Delta t)$, (higher powers of Δt), and events over nonoverlapping intervals are independent of each other. This gives rise to the following axioms.

Axioms:

(i) $P\{n(t, t+\Delta t)=1\} = \lambda \Delta t + o(\Delta t)$

(ii) $P\{n(t, t+\Delta t)=0\} = 1-\lambda \Delta t + o(\Delta t)$

(iii) $P\{n(t, t+\Delta t)\geq 2\} = o(\Delta t)$

and

(iv) $n(t, t+\Delta t)$ is independent of $n(0, t)$

(15-29)

Notice that axiom (iii) specifies that the events occur singly, and axiom (iv) specifies the randomness of the entire series. Axiom(ii) follows from (i) and (iii) together with the axiom of total probability.

We shall use these axiom to rederive (15-25) first:
Let t_0 be *any* fixed point (see Fig. 15.6) and let $t_0+\tau_1$ represent the time of the first arrival after t_0 . Notice that the random variable τ_1 is independent of the occurrences prior to the instant t_0 (Axiom (iv)). With $F_{\tau_1}(t)=P\{\tau_1\le t\}$ representing the distribution function of τ_1, as in (15-24) define $Q(t)\triangleq 1-F_{\tau_1}(t)=P\{\tau_1>t\}$. Then for $\Delta t>0$

$$\begin{aligned}Q(t+\Delta t)&=P\{\tau_1>t+\Delta t\}\\&=P\{\tau_1>t,\ \text{ and no event occurs in } (t_0+t,\ t_0+t+\Delta t)\}\\&=P\{\tau_1>t,\ n(t_0+t,\ t_0+t+\Delta t)=0\}\\&=P\{n(t_0+t,\ t_0+t+\Delta t)=0\,|\,\tau_1>t\}P\{\tau_1>t\}.\end{aligned}$$

From axiom (iv), the conditional probability in the above expression is not affected by the event $\{\tau_1>t\}$ which refers to $\{n(t_0,\, t_0+t)=0\}$, i.e., to events before t_0+t, and hence the unconditional probability in axiom (ii) can be used there. Thus

$$Q(t+\Delta t)=[1-\lambda\Delta t+o(\Delta t)]Q(t)$$

or

$$\lim_{\Delta t\to 0}\frac{Q(t+\Delta t)-Q(t)}{\Delta t}=Q'(t)=-\lambda Q(t)\quad\Rightarrow\quad Q(t)=ce^{-\lambda t}.$$

But $c = Q(0) = P\{\tau_1 > 0\} = 1$ so that

$$Q(t) = 1 - F_{\tau_1}(t) = e^{-\lambda t}$$

or

$$F_{\tau_1}(t) = 1 - e^{-\lambda t}, \qquad t \geq 0$$

which gives

$$f_{\tau_1}(t) = \frac{dF_{\tau_1}(t)}{dt} = \lambda e^{-\lambda t}, \qquad t \geq 0 \qquad (15\text{-}30)$$

to be the p.d.f of τ_1 as in (15-25).

Similarly (15-5) can be derived from axioms (i)-(iv) in (15-29) as well. To see this, let

$$p_k(t) \triangleq P\{n(0,t) = k)\}, \quad k = 0,\ 1,\ 2, \cdots$$

represent the probability that the total number of arrivals in the interval (0, t) equals k. Then

$$p_k(t + \Delta t) = P\{n(0, t + \Delta t) = k)\} = P\{X_1 \cup X_2 \cup X_3\}$$

where the events

$$X_1 \triangleq \text{"}n(0,t)=k, \text{ and } n(t,\ t+\Delta t)=0\text{"}$$
$$X_2 \triangleq \text{"}n(0,t)=k-1, \text{ and } n(t,\ t+\Delta t)=1\text{"}$$
$$X_3 \triangleq \text{"}n(0,t)=k-i, \text{ and } n(t,\ t+\Delta t)=i\geq 2\text{"}$$

are mutually exclusive. Thus

$$p_k(t+\Delta t)=P(X_1)+P(X_2)+P(X_3).$$

But as before

$$P(X_1)=P\{n(t,t+\Delta t)=0\,|\,n(0,t)=k\}P\{n(0,t)=k\}$$
$$=P\{n(t,t+\Delta t)=0\}P\{n(0,t)=k\}$$
$$=(1-\lambda\Delta t)\,p_k(t)$$
$$P(X_2)=P\{n(t,t+\Delta t)=1\,|\,n(0,t)=k-1\}P\{n(0,t)=k-1\}$$
$$=\lambda\Delta t\; p_{k-1}\Delta t$$

and

$$P(X_3)=0$$

where once again we have made use of axioms (i)-(iv) in (15-29). This gives

$$p_k(t+\Delta t) = (1-\lambda\Delta t)\, p_k(t) + \lambda\Delta t p_{k-1}(t)$$

or with

$$\lim_{\Delta t \to 0} \frac{p_k(t+\Delta t) - p_k(t)}{\Delta t} = p_k'(t)$$

we get the differential equation

$$p_k'(t) = -\lambda p_k(t) - \lambda p_{k-1}(t), \qquad k = 0, 1, 2, \cdots$$

whose solution gives (15-5). Here $p_{-1}(t) \equiv 0.$ [Solution to the above differential equation is worked out in (16-36)-(16-41), Text]. This is completes the axiomatic development for Poisson processes.

Poisson Departures between Exponential Inter-arrivals

Let $X(t) \sim P(\lambda t)$ and $Y(t) \sim P(\mu t)$ represent two independent Poisson processes called *arrival* and *departure* processes.

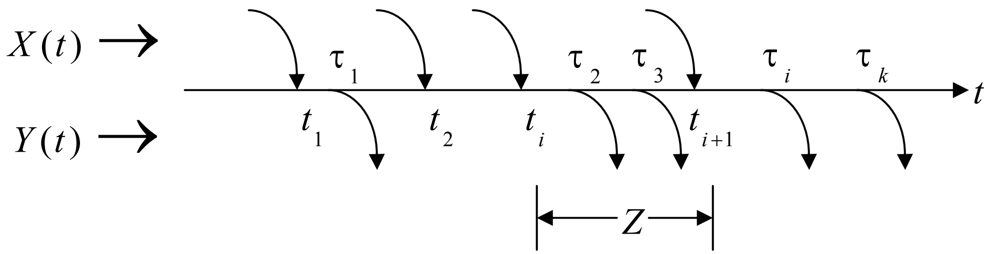


Fig. 15.7

Let Z represent the random interval between *any* two successive arrivals of $X(t)$. From (15-28), Z has an exponential distribution with parameter λ. Let N represent the number of "departures" of $Y(t)$ between *any* two successive arrivals of $X(t)$. Then from the Poisson nature of the departures we have

$$P\{N = k \mid Z = t\} = e^{-\mu t}\,\frac{(\mu t)^k}{k!}.$$

Thus

$$P\{N=k\}=\int_0^\infty P\{N=k\mid Z=t\}f_Z(t)dt$$
$$=\int_0^\infty e^{-\mu t}\,\frac{(\mu t)^k}{k!}\lambda e^{-\lambda t}dt$$
$$=\frac{\lambda}{k!}\int_0^\infty (\mu t)^k e^{-(\lambda+\mu)t}dt$$
$$=\frac{\lambda}{\lambda+\mu}\left(\frac{\mu}{\lambda+\mu}\right)^k \frac{1}{k!}\underbrace{\int_0^\infty x^k e^{-x}dx}_{k!}$$
$$=\left(\frac{\lambda}{\lambda+\mu}\right)\left(\frac{\mu}{\lambda+\mu}\right)^k, \quad k=0,\,1,\,2,\,\cdots \tag{15-31}$$

i.e., the random variable N has a geometric distribution. Thus if customers come in and get out according to two independent Poisson processes at a counter, then the number of arrivals between any two departures has a geometric distribution. Similarly the number of departures between *any* two arrivals also represents another geometric distribution.

Stopping Times, Coupon Collecting, and Birthday Problems

Suppose a cereal manufacturer inserts a sample of one type of coupon randomly into each cereal box. Suppose there are n such distinct types of coupons. One interesting question is that how many boxes of cereal should one buy on the average in order to collect at least one coupon of each kind?

We shall reformulate the above problem in terms of Poisson processes. Let $X_1(t), X_2(t), \cdots, X_n(t)$ represent n *independent* identically distributed Poisson processes with common parameter $\lambda t.$ Let $t_{i1}, t_{i2}, \cdots$ represent the first, second, $\cdots$ random arrival instants of the process $X_i(t), \;\; i = 1, 2, \cdots, n.$ They will correspond to the first, second, $\cdots$ appearance of the i^{th} type coupon in the above problem. Let

$$X(t) \triangleq \sum_{i=1}^{n} X_i(t), \qquad (15\text{-}32)$$

so that the sum $X(t)$ is also a Poisson process with parameter μt, where

$$\mu = n\lambda. \qquad (15\text{-}33)$$

From Fig. 15.8, $1/\lambda$ represents The average inter-arrival duration between any two arrivals of $X_i(t), i=1,2,\cdots,n,$ whereas $1/\mu$ represents the average inter-arrival time for the combined sum process $X(t)$ in (15-32).

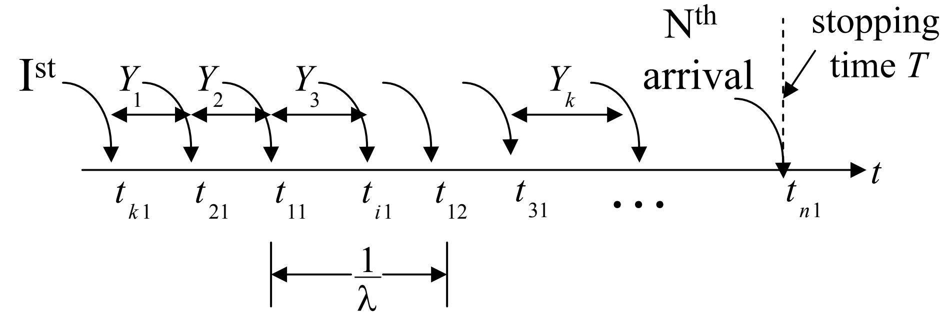


Fig. 15.8

Define the **stopping time** T to be that random time instant by which at least one arrival of $X_1(t), X_2(t), \cdots, X_n(t)$ has occurred. Clearly, we have

$$T = \max\,(t_{11}, t_{21}, \cdots, t_{i1}, \cdots, t_{n1}). \tag{15-34}$$

But from (15-25), $t_{i1}, \; i=1,2,\cdots,n$ are independent exponential random variables with common parameter λ. This gives

$$\begin{aligned} F_T(t) &= P\{T \le t\} = P\{\max\,(t_{11}, t_{21}, \cdots, t_{n1}) \le t\} \\ &= P\{t_{11} \le t,\; t_{21} \le t,\; \cdots,\; t_{n1} \le t\} \\ &= P\{t_{11} \le t\}P\{t_{21} \le t\}\cdots P\{t_{n1} \le t\} = [F_{ti}(t)]^n. \end{aligned}$$

Thus

$$F_T(t) = (1 - e^{-\lambda t})^n, \qquad t \geq 0 \tag{15-35}$$

represents the probability distribution function of the stopping time random variable T in (15-34). To compute its mean, we can make use of Eqs. (5-52)-(5-53) Text, that is valid for nonnegative random variables. From (15-35) we get

$$P(T > t) = 1 - F_T(t) = 1 - (1 - e^{-\lambda t})^n, \qquad t \geq 0$$

so that

$$E\{T\} = \int_0^\infty P(T > t)dt = \int_0^\infty \{1 - (1 - e^{-\lambda t})^n\}dt. \tag{15-36}$$

Let $1 - e^{-\lambda t} = x,$ so that $\lambda\, e^{-\lambda t} dt = dx,$ or $dt = \frac{dx}{\lambda(1-x)},$ and

$$E\{T\} = \frac{1}{\lambda}\int_0^1 (1 - x^n)\frac{dx}{1-x} = \frac{1}{\lambda}\int_0^1 \frac{1-x^n}{1-x}\,dx$$

$$= \frac{1}{\lambda}\int_0^1 (1 + x + x^2 + \cdots + x^{n-1})dx = \frac{1}{\lambda}\sum_{k=1}^{n} \frac{x^k}{k}\Bigg|_0^1$$

$$E\{T\} = \frac{1}{\lambda}\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right) = \frac{n}{\mu}\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right)$$

$$\simeq \frac{1}{\mu} n(\ln n + \gamma) \qquad (15\text{-}37)$$

where $\gamma \simeq 0.5772157\cdots$ is the Euler's constant[1].

Let the random variable N denote the total number of all arrivals up to the stopping time T, and $Y_k,\ k = 1, 2, \cdots, N$ the inter-arrival random variables associated with the sum process $X(t)$ (see Fig 15.8).

---

[1] **Euler's constant:** The series $\{1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \ln n\}$ converges, since

$$u_n \triangleq \int_0^1 \frac{x}{n(n+x)} dx = \int_0^1 \left(\frac{1}{n} - \frac{1}{n+x}\right) dx = \frac{1}{n} - \ln \frac{n+1}{n} > 0 \qquad (1)$$

and $u_n \le \int_0^1 \frac{x}{n^2} dx \le \int_0^1 \frac{1}{n^2} dx = \frac{1}{n^2}$ so that $\sum_{n=1}^{\infty} u_n < \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} < \infty.$ Thus the series $\{u_n\}$ converges to some number $\gamma > 0$. From (1) we obtain also $\sum_{k=1}^{n} u_k = \sum_{k=1}^{n} \frac{1}{k} - \ln(n+1)$ so that

$$\lim_{n\to\infty} \{1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \ln n\} = \lim_{n\to\infty} \left\{\sum_{k=1}^{n} u_k + \ln \frac{n+1}{n}\right\} = \sum_{k=1}^{\infty} u_k = \gamma = 0.5772157\cdots.$$

Then we obtain the key relation

$$T = \sum_{i=1}^{N} Y_i \qquad (15\text{-}38)$$

so that

$$E\{T \mid N = n\} = E\{\sum_{i=1}^{n} Y_i \mid N = n\} = E\{\sum_{i=1}^{n} Y_i\} = nE\{Y_i\} \qquad (15\text{-}39)$$

since $\{Y_i\}$ and N are independent random variables, and hence

$$E\{T\} = E[E\{T \mid N = n\}] = E\{N\}E\{Y_i\} \qquad (15\text{-}40)$$

But $Y_i \sim \text{Exponential}(\mu)$, so that $E\{Y_i\} = 1/\mu$ and substituting this into (15-37), we obtain

$$E\{N\} \simeq n(\ln n + \gamma). \qquad (15\text{-}41)$$

Thus on the average a customer should buy about $n \ln n$, or slightly more, boxes to guarantee that at least one coupon of each type has been collected.

Next, consider a slight generalization to the above problem: What if two kinds of coupons (each of n type) are mixed up, and the objective is to collect one complete set of coupons of either kind?

Let $X_i(t)$ and $Y_i(t)$, $i = 1, 2, \cdots, n$ represent the two kinds of coupons (independent Poisson processes) that have been mixed up to form a single Poisson process $Z(t)$ with normalized parameter unity. i.e.,

$$Z(t) = \sum_{i=1}^{n} [X_i(t) + Y_i(t)] \sim P(t). \tag{15-42}$$

As before let $t_{i1}, t_{i2}, \cdots$ represent the first, second, $\cdots$ arrivals of the process $X_i(t)$, and $\tau_{i1}, \tau_{i2}, \cdots$ represent the first, second, arrivals of the process $Y_i(t)$, $i = 1, 2, \cdots, n$.

The stopping time T_1 in this case represents that random instant at which *either* all X – type *or* all Y – type have occurred at least once. Thus

$$T_1 = \min\{X, Y\} \tag{15-43}$$

where

$$X \triangleq \max\ (t_{11}, t_{21}, \cdots, t_{n1}) \tag{15-44}$$

and

$$Y \triangleq \max\ (\tau_{11}, \tau_{21}, \cdots, \tau_{n1}). \tag{15-45}$$

Notice that the random variables X and Y have the same distribution as in (15-35) with λ replaced by $1/2n$ (since $\mu = 1$ and there are $2n$ independent and identical processes in (15-42)), and hence

$$F_X(t) = F_Y(t) = (1 - e^{-t/2n})^n, \qquad t \geq 0. \tag{15-46}$$

Using (15-43), we get

$$
\begin{aligned}
F_{T_1}(t) = P(T_1 \le t) &= P(\min\{X,Y\} \le t) \\
&= 1 - P(\min\{X,Y\} > t) = 1 - P(X > t, Y > t) \\
&= 1 - P(X > t)P(Y > t) \\
&= 1 - (1 - F_X(t))(1 - F_Y(t)) \\
&= 1 - \{1 - (1 - e^{-t/2n})^n\}^2, \qquad t \ge 0 \qquad (15\text{-}47)
\end{aligned}
$$

to be the probability distribution function of the new stopping time T_1. Also as in (15-36)

$$
\begin{aligned}
E\{T_1\} &= \int_0^\infty P(T_1 > t)dt = \int_0^\infty \{1 - F_{T_1}(t)\}dt \\
&= \int_0^\infty \{1 - (1 - e^{-t/2n})^n\}^2 dt.
\end{aligned}
$$

Let $1 - e^{-t/2n} = x$, or $\frac{1}{2n} e^{-t/2n} dt = dx$, $dt = \frac{2n dx}{1-x}$.

$$
\begin{aligned}
E\{T_1\} &= 2n \int_0^1 (1 - x^n)^2 \frac{dx}{1-x} = 2n \int_0^1 \left(\frac{1 - x^n}{1 - x} \right) (1 - x^n)\, dx \\
&= 2n \int_0^1 (1 + x + x^2 + \cdots + x^{n-1})(1 - x^n) dx
\end{aligned}
$$

$$
\begin{aligned}
E\{T_1\} &= 2n\int_0^1 \Big(\sum_{k=0}^{n-1} x^k - \sum_{k=0}^{n-1} x^{n+k}\Big)\, dx \\
&= 2n\left\{ \left(1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}\right) - \left(\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{2n}\right) \right\} \\
&\simeq 2n(\ln(n/2)+\gamma). \qquad (15\text{-}48)
\end{aligned}
$$

Once again the total number of random arrivals N up to T_1 is related as in (15-38) where $Y_i \sim$ Exponential (1), and hence using (15-40) we get the average number of total arrivals up to the stopping time to be

$$
E\{N\} = E\{T_1\} \simeq 2n(\ln(n/2)+\gamma). \qquad (15\text{-}49)
$$

We can generalize the stopping times in yet another way:

Poisson Quotas

Let

$$
X(t) = \sum_{i=1}^{n} X_i(t) \sim P(\mu t) \qquad (15\text{-}50)
$$

where $X_i(t)$ are independent, identically distributed Poisson processes with common parameter $\lambda_i t$ so that $\mu = \lambda_1 + \lambda_2 + \cdots + \lambda_n$. Suppose integers $m_1, m_2, \cdots, m_n$ represent the preassigned number of arrivals (quotas) required for processes $X_1(t), X_2(t), \cdots, X_n(t)$ in the sense that when m_i arrivals of the process $X_i(t)$ have occurred, the process $X_i(t)$ satisfies its "quota" requirement.

The **stopping time** T in this case is that random time instant at which *any* r processes have met their quota requirement where $r \le n$ is given. The problem is to determine the probability density function of the stopping time random variable T, and determine the mean and variance of the total number of random arrivals N up to the stopping time T.

Solution: As before let $t_{i1}, t_{i2}, \cdots$ represent the first, second, $\cdots$ arrivals of the i^{th} process $X_i(t)$, and define

$$Y_{ij} = t_{ij} - t_{i,j-1} \tag{15-51}$$

Notice that the inter-arrival times Y_{ij} and independent, exponential random variables with parameter λ_i, and hence

$$t_{i,m_i} = \sum_{j=1}^{m_i} Y_{ij} \sim \text{Gamma}(m_i, \lambda_i)$$

Define T_i to the **stopping time** for the i^{th} process; i.e., the occurrence of the m_i^{th} arrival equals T_i . Thus

$$T_i = t_{i,m_i} \sim \text{Gamma}(m_i, \lambda_i), \quad i = 1, 2, \cdots, n \tag{15-52}$$

or

$$f_{T_i}(t) = \frac{t^{m_i - 1}}{(m_i - 1)!} \lambda_i^{m_i} e^{-\lambda_i t}, \qquad t \geq 0. \tag{15-53}$$

Since the n processes in (15-49) are independent, the associated stopping times $T_i, \quad i = 1, 2, \cdots, n$ defined in (15-53) are also *independent* random variables, a key observation.

Given *independent* gamma random variables $T_1, T_2, \cdots, T_n,$ in (15-52)-(15-53) we form their order statistics: This gives

$$T_{(1)} < T_{(2)} < \cdots < T_{(r)} < \cdots < T_{(n)}. \tag{15-54}$$

Note that the two extremes in (15-54) represent

$$T_{(1)} = \min\,(T_1, T_2, \cdots, T_n) \tag{15-55}$$

and

$$T_{(n)} = \max\,(T_1, T_2, \cdots, T_n). \tag{15-56}$$

The desired stopping time T when r processes have satisfied their quota requirement is given by the r^{th} order statistics $T_{(r)}$. Thus

$$T = T_{(r)} \tag{15-57}$$

where $T_{(r)}$ is as in (15-54). We can use (7-14), Text to compute the probability density function of T. From there, the probability density function of the stopping time random variable T in (15-57) is given by

$$f_T(t) = \frac{n!}{(r-1)!(n-r)!} F_{T_i}^{k-1}(t)[1-F_{T_i}(t)]^{n-k} f_{T_i}(t) \tag{15-58}$$

where $F_{T_i}(t)$ is the distribution of the i.i.d random variables T_i and $f_{T_i}(t)$ their density function given in (15-53). Integrating (15-53) by parts as in (4-37)-(4-38), Text, we obtain

$$F_{T_i}(t) = 1 - \sum_{k=0}^{m_i-1} \frac{(\lambda_i t)^k}{k!} e^{-\lambda_i t}, \qquad t \geq 0. \tag{15-59}$$

Together with (15-53) and (15-59), Eq. (15-58) completely specifies the density function of the stopping time random variable T, where r types of arrival quota requirements have been satisfied.

If N represents the total number of *all* random arrivals up to T, then arguing as in (15-38)-(15-40) we get

$$T = \sum_{i=1}^{N} Y_i \tag{15-60}$$

where Y_i are the inter-arrival intervals for the process $X(t)$, and hence

$$E\{T\} = E\{N\}E\{Y_i\} = \frac{1}{\mu} E\{N\} \tag{15-61}$$

with normalized mean value $\mu (=1)$ for the sum process $X(t)$, we get

$$E\{N\} = E\{T\}. \tag{15-62}$$

To relate the higher order moments of N and T we can use their characteristic functions. From (15-60)

$$E\{e^{j\omega T}\} = E\{e^{j\omega \sum_{i=1}^{N} Y_i}\} = E\{\underbrace{E[e^{j\omega \sum_{i=1}^{n} Y_i} \mid N = n]}_{[E\{e^{j\omega Y_i}\}]^n}\}$$

$$= E[\{E\{e^{j\omega Y_i}\} \mid N = n\}^n]. \tag{15-63}$$

But $Y_i \sim$ Exponential (1) and independent of N so that

$$E\{e^{j\omega Y_i} \mid N=n\} = E\{e^{j\omega Y_i}\} = \frac{1}{1-j\omega}$$

and hence from (15-63)

$$E\{e^{j\omega T}\} = \sum_{n=0}^{\infty} [E\{e^{j\omega Y_i}\}]^n P(N=n) = E\left\{ \left(\tfrac{1}{1-j\omega}\right)^N \right\} = E\{(1-j\omega)^{-N}\}$$

which gives (expanding both sides)

$$\sum_{k=0}^{\infty} \frac{(j\omega)^k}{k!} E\{T^k\} = \sum_{k=0}^{\infty} \frac{(j\omega)^k}{k!} E\{N(N+1)\cdots(N+k-1)\}$$

or

$$E\{T^k\} = E\{N(N+1)\cdots(N+k-1)\}, \qquad (15\text{-}64)$$

a key identity. From (15-62) and (15-64), we get

$$\mathrm{var}\{N\} = \mathrm{var}\{T\} - E\{T\}. \qquad (15\text{-}65)$$

As an application of the Poisson quota problem, we can reexamine *the birthday pairing* problem discussed in Example 2-20, Text.

"Birthday Pairing" as Poisson Processes:

In the birthday pairing case (refer to Example 2-20, Text), we may assume that $n = 365$ possible birthdays in an year correspond to n independent identically distributed Poisson processes each with parameter $1/n$, and in that context each individual is an "arrival" corresponding to his/her particular "birth-day process". It follows that the birthday pairing problem (i.e., two people have the same birth date) corresponds to the first occurrence of the 2nd return for *any* one of the 365 processes. Hence

$$m_1 = m_2 = \cdots m_n = 2, \qquad r = 1 \tag{15-66}$$

so that from (15-52), for each process

$$T_i \sim \text{Gamma}\,(2, 1/n). \tag{15-67}$$

Since $\lambda_i = \mu / n = 1/n,$ and from (15-57) and (15-66) the stopping time in this case satisfies

$$T = \min\,(T_1, T_2, \cdots, T_n). \tag{15-68}$$

Thus the distribution function for the "birthday pairing" stopping time turns out to be

$$\begin{aligned} F_T(t) &= P\{T \le t\} = 1 - P\{T > t\} \\ &= 1 - P\{\min\ (T_1, T_2, \cdots, T_n) > t\} \\ &= 1 - [P\{T_i > t\}]^n = 1 - [1 - F_{T_i}(t)]^n \\ &= 1 - (1 + \tfrac{t}{n})^n e^{-t} \qquad (15\text{-}69) \end{aligned}$$

where we have made use of (15-59) with $m_i = 2$ and $\lambda_i = 1/n.$

As before let N represent the number of random arrivals up to the stopping time T. Notice that in this case N represents the number of people required in a crowd for at least two people to have the same birth date. Since T and N are related as in (15-60), using (15-62) we get

$$\begin{aligned} E\{N\} &= E\{T\} = \int_0^\infty P(T > t)dt \\ &= \int_0^\infty \{1 - F_T(t)\}dt = \int_0^\infty (1 + \tfrac{t}{n})^n e^{-t} dt. \qquad (15\text{-}70) \end{aligned}$$

To obtain an approximate value for the above integral, we can expand $\ln(1 + \frac{t}{n})^n = n \ln(1 + \frac{t}{n})$ in Taylor series. This gives

$$\ln\left(1+\frac{t}{n}\right)=\frac{t}{n}-\frac{t^2}{2n^2}+\frac{t^3}{3n^3}$$

and hence

$$\left(1+\frac{t}{n}\right)^n=e^{(t-\frac{t^2}{2n}+\frac{t^3}{3n^2})}$$

so that

$$\left(1+\frac{t}{n}\right)^n e^{-t}=e^{-\frac{t^2}{2n}}\, e^{\frac{t^3}{3n^2}}\approx e^{-\frac{t^2}{2n}}\left(1+\frac{t^3}{3n^2}\right) \qquad (15\text{-}71)$$

and substituting (15-71) into (15-70) we get the mean number of people in a crowd for a two-person birthday coincidence to be

$$\begin{aligned} E\{N\} &\approx \int_0^\infty e^{-t^2/2n}\,dt+\frac{1}{3n^2}\int_0^\infty t^3 e^{-t^2/2n}\,dt \\ &= \frac{1}{2}\sqrt{2\pi\, n}+\frac{2n^2}{3n^2}\int_0^\infty x e^{-x}\,dx=\sqrt{\frac{\pi\, n}{2}}+\frac{2}{3} \\ &= 24.612. \end{aligned} \qquad (15\text{-}72)$$

On comparing (15-72) with the mean value obtained in Lecture 6 (Eq. (6-60)) using entirely different arguments $(E\{X\}\approx 24.44)$, we observe that the two results are essentially equal. Notice that the probability that there will be a coincidence among 24-25 people is about 0.55. To compute the variance of T and N we can make use of the expression (see (15-64))

$$E\{N(N+1)\} = E\{T^2\} = \int_0^\infty 2tP(T>t)dt$$

$$= \int_0^\infty 2t\left(1+\frac{t}{n}\right)^n e^{-t}dt \approx 2\int_0^\infty t\ e^{-t^2/2n}dt + \frac{2}{3n^2}\int_0^\infty t^4\ e^{-t^2/2n}dt$$

$$= 2n\int_0^\infty e^{-x}dx + \frac{2}{3n^2}\sqrt{\pi}\ \ \frac{3}{8}(2n)^2\sqrt{2n} = 2n + \sqrt{2\pi\ n} \qquad \text{(15-73)}$$

which gives (use (15-65))

$$\sigma_T \approx 13.12, \qquad \sigma_N \approx 12.146. \qquad \text{(15-74)}$$

The high value for the standard deviations indicate that in reality the crowd size could vary considerably around the mean value.

Unlike Example 2-20 in Text, the method developed here can be used to derive the distribution and average value for a variety of "birthday coincidence" problems.

Three person birthday-coincidence:

For example, if we are interested in the average crowd size where three people have the same birthday, then arguing as above, we obtain

$$m_1 = m_2 = \cdots = m_n = 3, \qquad r = 1, \tag{15-75}$$

so that

$$T_i \sim \text{Gamma}\,(3, 1/n) \tag{15-76}$$

and T is as in (15-68), which gives

$$F_T(t) = 1 - [1 - F_{T_i}(t)]^n = 1 - \left(1 + \frac{t}{n} + \frac{t^2}{2n^2}\right)^n e^{-t}, \qquad t \geq 0 \tag{15-77}$$

(use (15 - 59) with $m_i = 3,\ \lambda_i = 1/n$) to be the distribution of the stopping time in this case. As before, the average crowd size for three-person birthday coincidence equals

$$E\{N\} = E\{T\} = \int_0^\infty P(T > t)dt = \int_0^\infty \left(1 + \frac{t}{n} + \frac{t^2}{2n^2}\right)^n e^{-t} dt.$$

By Taylor series expansion

$$\ln\left(1+\frac{t}{n}+\frac{t^2}{2n^2}\right)=\left(\frac{t}{n}+\frac{t^2}{2n^2}\right)-\frac{1}{2}\left(\frac{t}{n}+\frac{t^2}{2n^2}\right)^2+\frac{1}{3}\left(\frac{t}{n}+\frac{t^2}{2n^2}\right)^3$$

$$\approx \frac{t}{n}-\frac{t^3}{6n^3}$$

so that

$$E\{N\}=\int_0^\infty e^{-t^3/6n^2}\,dt=\frac{6^{1/3}n^{2/3}}{3}\int_0^\infty x^{\left(\frac{1}{3}-1\right)}e^{-x}\,dx$$

$$=6^{1/3}\,\Gamma(4/3)\,n^{2/3}\approx 82.85. \qquad (15\text{-}78)$$

Thus for a three people birthday coincidence the average crowd size should be around 82 (which corresponds to 0.44 probability).

Notice that other generalizations such as "two distinct birthdays to have a pair of coincidence in each case" (m_i =2, r = 2) can be easily worked in the same manner.

We conclude this discussion with the other extreme case, where the crowd size needs to be determined so that "all days in the year are birthdays" among the persons in a crowd.

All days are birthdays:

Once again from the above analysis in this case we have

$$m_1 = m_2 = \cdots = m_n = 1, \qquad r = n = 365 \tag{15-79}$$

so that the stopping time statistics T satisfies

$$T = \max\,(T_1, T_2, \cdots, T_n), \tag{15-80}$$

where T_i are independent exponential random variables with common parameter $\lambda = \frac{1}{n}$. This situation is similar to the *coupon collecting problem* discussed in (15-32)-(15-34) and from (15-35), the distribution function of T in (15-80) is given by

$$F_T(t) = (1 - e^{-t/n})^n, \qquad t \geq 0 \tag{15-81}$$

and the mean value for T and N are given by (see (15-37)-(15-41))

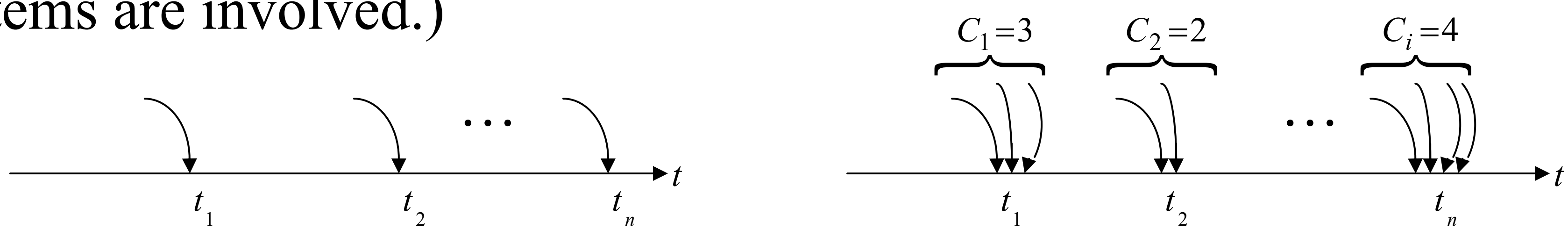
$$E\{N\} = E\{T\} \approx n(\ln n + \gamma) = 2{,}364.14. \tag{15-82}$$

Thus for "everyday to be a birthday" for someone in a crowd, the average crowd size should be 2,364, in which case there is 0.57 probability that the event actually happens.

For a more detailed analysis of this problem using Markov chains, refer to Examples 15-12 and 15-18, in chapter 15, Text. From there (see Eq. (15-80), Text) to be quite certain (with 0.98 probability) that all 365 days are birthdays, the crowd size should be around 3,500.

Bulk Arrivals and Compound Poisson Processes

In an ordinary Poisson process $X(t)$, only one event occurs at any arrival instant (Fig 15.9a). Instead suppose a random number of events C_i occur simultaneously as a cluster at every arrival instant of a Poisson process (Fig 15.9b). If $X(t)$ represents the total number of all occurrences in the interval $(0, t)$, then $X(t)$ represents a **compound Poisson process**, or a **bulk arrival process**. Inventory orders, arrivals at an airport queue, tickets purchased for a show, etc. follow this process (when things happen, they happen in a bulk, or a bunch of items are involved.)



(a) Poisson Process (b) Compound Poisson Process

Fig. 15.9

Let

$$p_k = P\{C_i = k\}, \quad k = 0, 1, 2, \cdots \tag{15-83}$$

represent the common probability mass function for the occurrence in any cluster C_i. Then the compound process $X(t)$ satisfies

$$X(t) = \sum_{i=1}^{N(t)} C_i, \qquad (15\text{-}84)$$

where $N(t)$ represents an ordinary Poisson process with parameter λ. Let

$$P(z) = E\{z^{C_i}\} = \sum_{k=0}^{\infty} p_k z^k \qquad (15\text{-}85)$$

represent the moment generating function associated with the cluster statistics in (15-83). Then the moment generating function of the compound Poisson process $X(t)$ in (15-84) is given by

$$\begin{aligned}
\phi_X(z) &= \sum_{n=0}^{\infty} z^n P\{X(t)=n\} = E\{z^{X(t)}\} \\
&= E\{E[z^{X(t)} \mid N(t)=k]\} = E[E\{z^{\sum_{i=1}^{k} C_i} \mid N(t)=k\}] \\
&= \sum_{k=0}^{\infty} (E\{z^{C_i}\})^k P\{N(t)=k\} \\
&= \sum_{k=0}^{\infty} P^k(z) e^{-\lambda t} \frac{(\lambda t)^k}{k!} = e^{-\lambda t(1-P(z))} \qquad (15\text{-}86)
\end{aligned}$$

If we let

$$P^k(z) \triangleq \left(\sum_{n=0}^{\infty} p_n z^k \right)^k = \sum_{n=0}^{\infty} p_n^{(k)} z^n \qquad (15\text{-}87)$$

where $\{p_n^{(k)}\}$ represents the k fold convolution of the sequence $\{p_n\}$ with itself, we obtain

$$P\{X(t)=n\} = \sum_{k=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} p_n^{(k)} \qquad (15\text{-}88)$$

that follows by substituting (15-87) into (15-86). Eq. (15-88) represents the probability that there are n arrivals in the interval $(0, t)$ for a compound Poisson process $X(t)$.

Substituting (15-85) into (15-86) we can rewrite $\phi_X(z)$ also as

$$\phi_X(z) = e^{-\lambda_1 t(1-z)} e^{-\lambda_2 t(1-z^2)} \cdots e^{-\lambda_k t(1-z^k)} \cdots \qquad (15\text{-}89)$$

where $\lambda_k = p_k \lambda,$ which shows that the compound Poisson process can be expressed as the sum of integer-secaled independent Poisson processes $m_1(t), m_2(t), \cdots.$ Thus

$$X(t) = \sum_{k=1}^{\infty} k\, m_k(t). \qquad (15\text{-}90)$$

More generally, every linear combination of independent Poisson processes represents a compound Poisson process. (see Eqs. (10-120)-(10-124), Text).

Here is an interesting problem involving compound Poisson processes and coupon collecting: Suppose a cereal manufacturer inserts *either* one *or* two coupons randomly – from a set consisting of n types of coupons – into every cereal box. How many boxes should one buy on the average to collect at least one coupon of each type? We leave it to the reader to work out the details.