Model-model Statistika dalam Simulasi

Muhammad Luthfi Shahab

Departemen Matematika

Institut Teknologi Sepuluh Nopember

Review dan Model-Model Statistika Penting

- Buku Banks Bab 5 Halaman 172-183:
  - Variabel acak diskrit
  - Variabel acak kontinu
  - CDF
  - Ekspektasi dan varians
  - Mode
  - Sistem antrian
  - Sistem inventory dan supply-chain
  - Reliability and maintainability

Distribusi Diskrit

- Buku Banks Bab 5 Halaman 183-189:
  - Percobaan Bernoulli dan distribusi Bernoulli
  - Distribusi binomial
  - Distribusi geometric dan negative binomial
  - Distribusi Poisson

Distribusi Poisson

The Poisson probability mass function is given by

$$p(x) = \begin{cases} \dfrac{e^{-\alpha}\alpha^x}{x!}, & x = 0, 1, \ldots \\ 0, & \text{otherwise} \end{cases} \tag{19}$$

where $\alpha > 0$. One of the important properties of the Poisson distribution is that the mean and variance are both equal to α, that is,

$$E(X) = \alpha = V(X)$$

The cumulative distribution function is given by

$$F(x) = \sum_{i=0}^{x} \frac{e^{-\alpha}\alpha^i}{i!} \tag{20}$$

The pmf and cdf for a Poisson distribution with $\alpha = 2$ are shown in Figure 7. A tabulation of the cdf is given in Table A.4 in the Appendix.

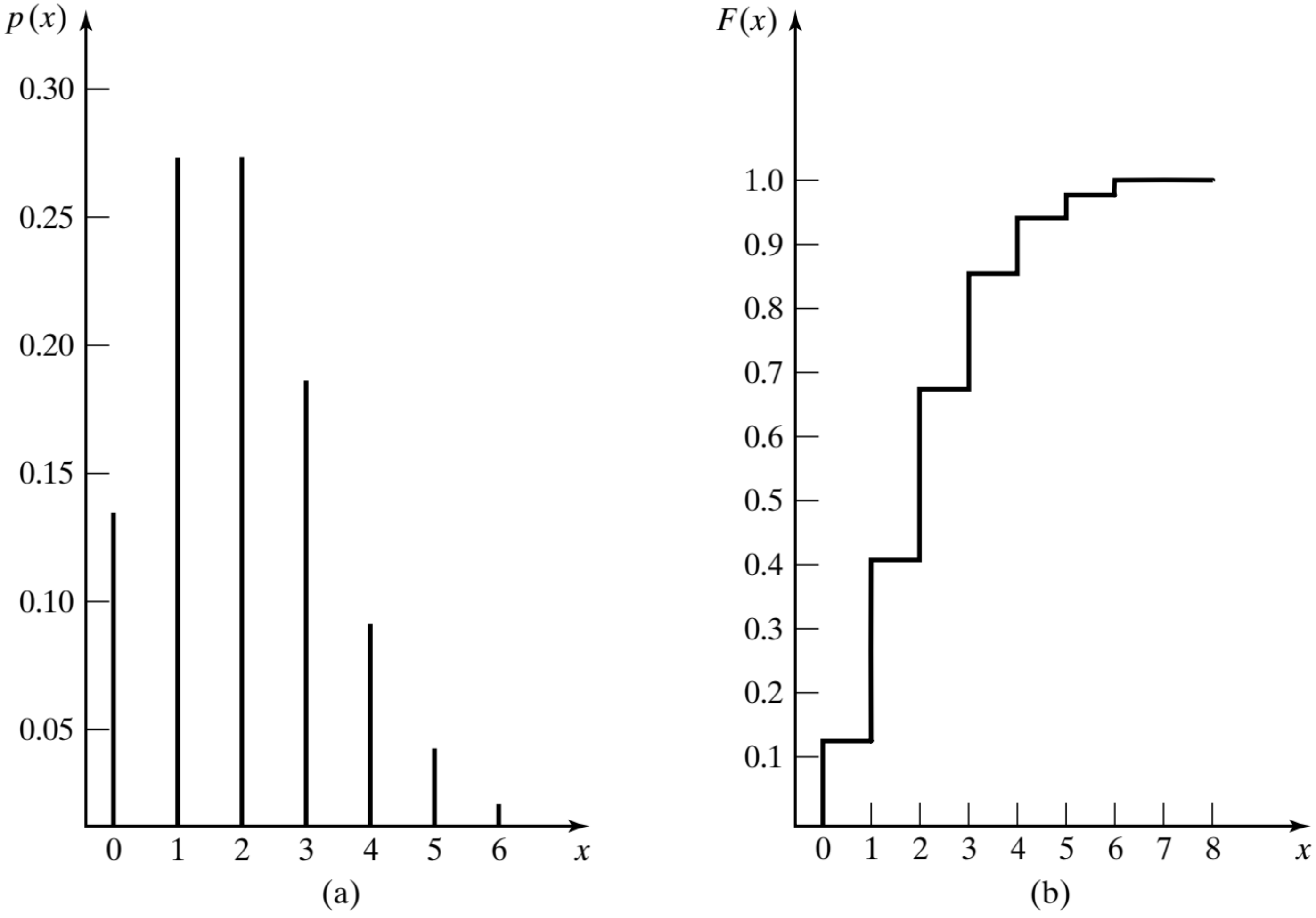

Figure 7 Poisson pmf and cdf.

Distribusi Kontinu

- Buku Banks Bab 5 Halaman 189-211:
  - Distribusi uniform
  - Distribusi eksponensial
  - Distribusi gamma
  - Distribusi Erlang
  - Distribusi normal
  - Distribusi Weibull
  - Distribusi segitiga
  - Distribusi lognormal
  - Distribusi beta

Distribusi Eksponensial

A random variable X is said to be exponentially distributed with parameter $\lambda > 0$ if its pdf is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{elsewhere} \end{cases} \tag{26}$$

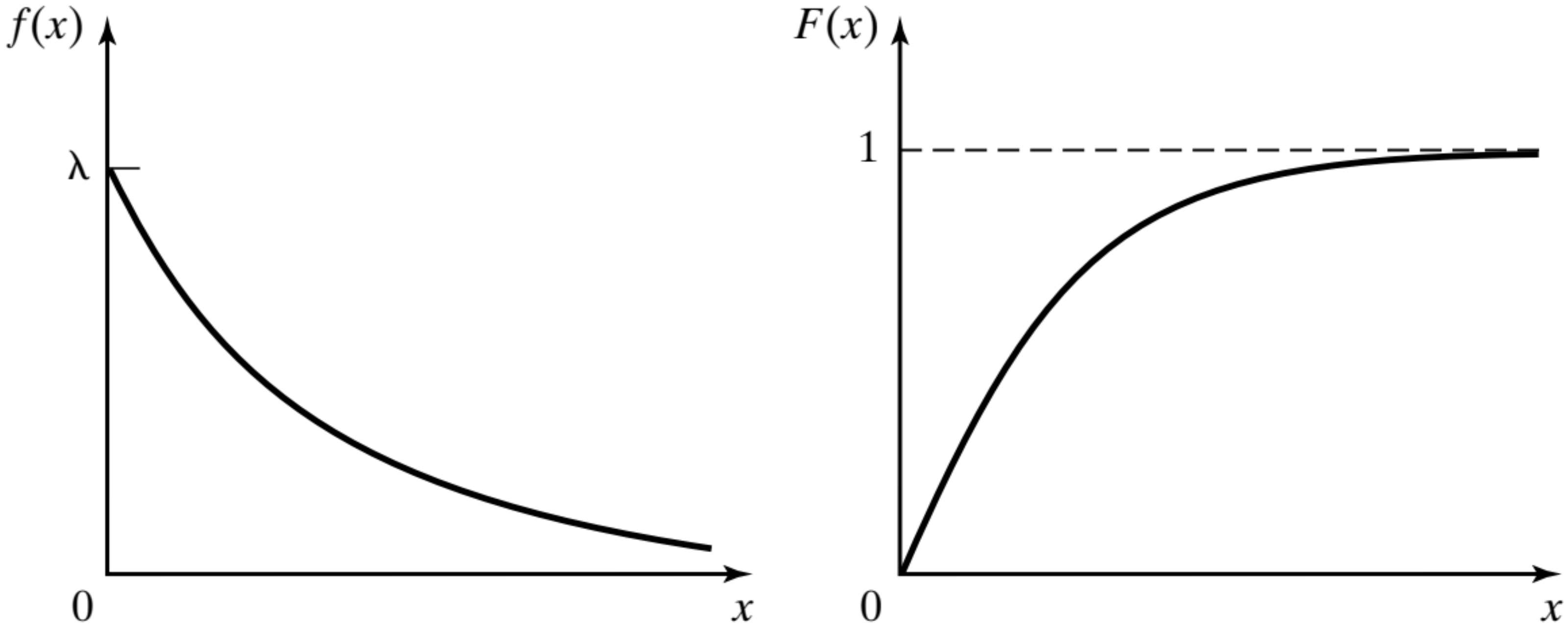

Figure 9 Exponential density function and cumulative distribution function.

The exponential distribution has been used to model interarrival times when arrivals are completely random and to model service times that are highly variable. In these instances, λ is a rate: arrivals per hour or services per minute. The exponential distribution has also been used to model the lifetime of a component that fails catastrophically (instantaneously), such as a light bulb; then λ is the failure rate.

The exponential distribution has mean and variance given by

$$E(X) = \frac{1}{\lambda} \quad \text{and} \quad V(X) = \frac{1}{\lambda^2} \tag{27}$$

Thus, the mean and standard deviation are equal. The cdf can be exhibited by integrating Equation (26) to obtain

$$F(x) = \begin{cases} 0, & x < 0 \\ \int_0^x \lambda e^{-\lambda t}\, dt = 1 - e^{-\lambda x}, & x \geq 0 \end{cases} \tag{28}$$

Poisson Process

Consider random events such as the arrival of jobs at a job shop, the arrival of e-mail to a mail server, the arrival of boats to a dock, the arrival of calls to a call center, the breakdown of machines in a large factory, and so on. These events may be described by a counting function $N(t)$ defined for all $t \geq 0$. This counting function will represent the number of events that occurred in $[0, t]$. Time zero is the point at which the observation began, regardless of whether an arrival occurred at that instant. For each interval $[0, t]$, the value $N(t)$ is an observation of a random variable where the only possible values that can be assumed by $N(t)$ are the integers $0, 1, 2, \ldots$.

The counting process, $\{N(t), t \geq 0\}$, is said to be a Poisson process with mean rate λ if the following assumptions are fulfilled:

a. Arrivals occur one at a time.
b. $\{N(t), t \geq 0\}$ has stationary increments: The distribution of the number of arrivals between t and $t+s$ depends only on the length of the interval s, not on the starting point t. Thus, arrivals are completely at random without rush or slack periods.
c. $\{N(t), t \geq 0\}$ has independent increments: The number of arrivals during nonoverlapping time intervals are independent random variables. Thus, a large or small number of arrivals in one time interval has no effect on the number of arrivals in subsequent time intervals. Future arrivals occur completely at random, independent of the number of arrivals in past time intervals.

If arrivals occur according to a Poisson process, meeting the three preceding assumptions, it can be shown that the probability that $N(t)$ is equal to n is given by

$$P[N(t) = n] = \frac{e^{-\lambda t}(\lambda t)^n}{n!} \quad \text{for } t \geq 0 \text{ and } n = 0, 1, 2, \ldots \tag{63}$$

Comparing Equation (63) to Equation (19), it can be seen that $N(t)$ has the Poisson distribution with parameter $\alpha = \lambda t$. Thus, its mean and variance are given by

$$E[N(t)] = \alpha = \lambda t = V[N(t)]$$

For any times s and t such that $s < t$, the assumption of stationary increments implies that the random variable $N(t) - N(s)$, representing the number of arrivals in the interval from s to t, is also Poisson-distributed with mean $\lambda(t - s)$. Thus,

$$P[N(t) - N(s) = n] = \frac{e^{-\lambda(t-s)}[\lambda(t - s)]^n}{n!} \quad \text{for } n = 0, 1, 2, \ldots$$

and

$$E[N(t) - N(s)] = \lambda(t - s) = V[N(t) - N(s)]$$

Now, consider the time at which arrivals occur in a Poisson process. Let the first arrival occur at time A_1, the second occur at time $A_1 + A_2$, and so on, as shown in Figure 25.

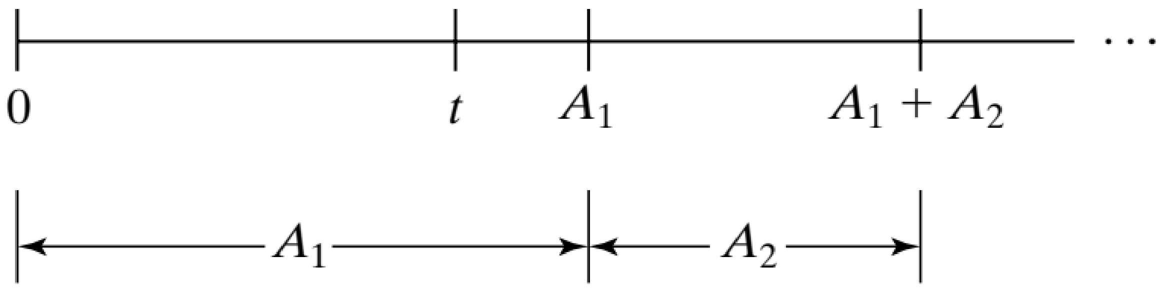

Figure 25 Arrival process.

Thus, $A_1, A_2, \ldots$ are successive interarrival times. The first arrival occurs after time t if and only if there are no arrivals in the interval $[0, t]$, so it is seen that

$$\{A_1 > t\} = \{N(t) = 0\}$$

and, therefore,

$$P(A_1 > t) = P[N(t) = 0] = e^{-\lambda t}$$

the last equality following from Equation (63). Thus, the probability that the first arrival will occur in $[0, t]$ is given by

$$P(A_1 \leq t) = 1 - e^{-\lambda t}$$

which is the cdf for an exponential distribution with parameter λ. Hence, A_1 is distributed exponentially with mean $E(A_1) = 1/\lambda$.

It can also be shown that all interarrival times, $A_1, A_2, \ldots,$ are exponentially distributed and independent with mean $1/\lambda$. As an alternative definition of a Poisson process, it can be shown that, if interarrival times are distributed exponentially and independently, then the number of arrivals by time t, say $N(t)$, meets the three previously mentioned assumptions and, therefore, is a Poisson process.

Recall that the exponential distribution is memoryless—that is, the probability of a future arrival in a time interval of length s is independent of the time of the last arrival. The probability of the arrival depends only on the length of the time interval s. Thus, the memoryless property is related to the properties of independent and stationary increments of the Poisson process.

Example 30

The jobs at a machine shop arrive according to a Poisson process with a mean of $\lambda = 2$ jobs per hour. Therefore, the interarrival times are distributed exponentially, with the expected time between arrivals being $E(A) = 1/\lambda = \frac{1}{2}$ hour.

Properties of Poisson Process

Several properties of the Poisson process, discussed by Ross [2002] and others, are useful in discrete-system simulation. The first of these properties concerns random splitting. Consider a Poisson process $\{N(t), t \geq 0\}$ having rate λ, as represented by the left side of Figure 26.

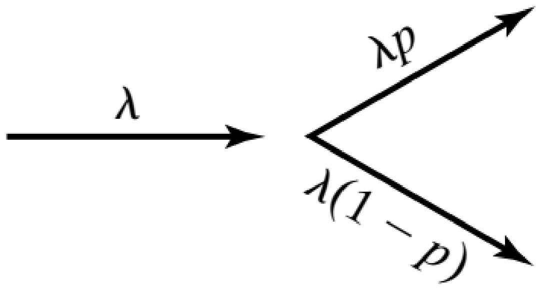

Figure 26 Random splitting.

Suppose that, each time an event occurs, it is classified as either a type I or a type II event. Suppose further that each event is classified as a type I event with probability p and type II event with probability $1 - p$, independently of all other events.

Let $N_1(t)$ and $N_2(t)$ be random variables that denote, respectively, the number of type I and type II events occurring in $[0, t]$. Note that $N(t) = N_1(t) + N_2(t)$. It can be shown that $N_1(t)$ and $N_2(t)$ are both Poisson processes having rates λp and $\lambda(1 - p)$, as shown in Figure 26. Furthermore, it can be shown that the two processes are independent.

Example 31: Random Splitting

Suppose that jobs arrive at a shop in accordance with a Poisson process having rate λ. Suppose further that each arrival is marked "high priority" with probability 1/3 and "low priority" with probability 2/3. Then a type I event would correspond to a high-priority arrival and a type II event would correspond to a low-priority arrival. If $N_1(t)$ and $N_2(t)$ are as just defined, both variables follow the Poisson process, with rates $\lambda/3$ and $2\lambda/3$, respectively.

Example 32

The rate in Example 31 is $\lambda = 3$ per hour. The probability that no high-priority jobs will arrive in a 2-hour period is given by the Poisson distribution with parameter $\alpha = \lambda pt = 2$. Thus,

$$P(0) = \frac{e^{-2}2^0}{0!} = 0.135$$

Now, consider the opposite situation from random splitting, namely the pooling of two arrival streams. The process of interest is illustrated in Figure 27. It can be shown that, if $N_i(t)$ are random variables representing independent Poisson processes with rates λ_i, for $i = 1$ and 2, then $N(t) = N_1(t) + N_2(t)$ is a Poisson process with rate $\lambda_1 + \lambda_2$.

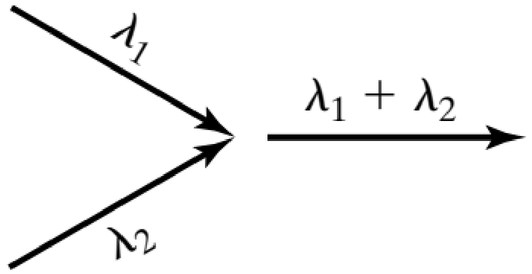

Figure 27 Pooled process.

Example 33: Pooled Process

A Poisson arrival stream with $\lambda_1 = 10$ arrivals per hour is combined (or pooled) with a Poisson arrival stream with $\lambda_2 = 17$ arrivals per hour. The combined process is a Poisson process with $\lambda = 27$ arrivals per hour.

Nonstationary Poisson Process

If we keep the Poisson Assumptions 1 and 3 (Section 5), but drop Assumption 2 (stationary increments) then we have a *nonstationary Poisson process* (NSPP), which is characterized by $\lambda(t)$, the arrival rate at time t. The NSPP is useful for situations in which the arrival rate varies during the period of interest, including meal times for restaurants, phone calls during business hours, and orders for pizza delivery around 6 P.M.

The key to working with an NSPP is the expected number of arrivals by time t, denoted by

$$\Lambda(t) = \int_0^t \lambda(s)\, ds$$

To be useful as an arrival-rate function, $\lambda(t)$ must be nonnegative and integrable. For a stationary Poisson process with rate λ we have $\Lambda(t) = \lambda t$, as expected.

Let $T_1, T_2, \ldots$ be the arrival times of stationary Poisson process $N(t)$ with $\lambda = 1$, and let $\mathcal{T}_1, \mathcal{T}_2, \ldots$ be the arrival times for an NSPP $\mathcal{N}(t)$ with arrival rate $\lambda(t)$. The fundamental relationship for working with NSPPs is the following:

$$T_i = \Lambda(\mathcal{T}_i)$$

$$\mathcal{T}_i = \Lambda^{-1}(T_i)$$

In other words, an NSPP can be transformed into a stationary Poisson process with arrival rate 1, and a stationary Poisson process with arrival rate 1 can be transformed into an NSPP with rate $\lambda(t)$, and the transformation in both cases is related to $\Lambda(t)$.

Example 34

Suppose that arrivals to a post office occur at a rate of 2 per minute from 8 A.M. until 12 P.M., then drop to 1 every 2 minutes until the day ends at 4 P.M. What is the probability distribution of the number of arrivals between 11 A.M. and 2 P.M?

Let time $t = 0$ correspond to 8 A.M. Then this situation could be modeled as an NSPP $\mathcal{N}(t)$ with rate function

$$\lambda(t) = \begin{cases} 2, & 0 \leq t < 4 \\ \dfrac{1}{2}, & 4 \leq t \leq 8 \end{cases}$$

The expected number of arrivals by time t is therefore

$$\Lambda(t) = \begin{cases} 2t, & 0 \leq t < 4 \\ \frac{t}{2} + 6, & 4 \leq t \leq 8 \end{cases}$$

Notice that computing the expected number of arrivals for $4 \leq t \leq 8$ requires that the integration be done in two parts:

$$\Lambda(t) = \int_0^t \lambda(s)\, ds = \int_0^4 2\, ds + \int_4^t \frac{1}{2}\, ds = \frac{t}{2} + 6$$

Since 2 P.M. and 11 A.M. correspond to times 6 and 3, respectively, we have

$$
\begin{aligned}
P[\mathcal{N}(6) - \mathcal{N}(3) = k] &= P[N(\Lambda(6)) - N(\Lambda(3)) = k] \\
&= P[N(9) - N(6) = k] \\
&= \frac{e^{-(9-6)}(9-6)^k}{k!} \\
&= \frac{e^{-3}(3)^k}{k!}
\end{aligned}
$$

where $N(t)$ is a stationary Poisson process with arrival rate 1.