Kishor Trivedi. Preprint used in Fall 1980,

$$P(X<\frac{4}{12} = \frac{1}{3}) = F_X(\frac{1}{3}) = \int_0^{1/3} f_X(x)\ dx$$

$$= \frac{7}{27}$$

or about 26 percent chance.

#

Most random variables we consider will either be discrete (as in Chapter 2) or continuous, but mixed random variables do occur sometimes. For example, there may be a nonzero probability, say, p_0, of initial failure of a component at time 0 due to manufacturing defects. In this case, the time to failure, X, of the component is neither discrete nor a continuous random variable. A possible CDF of X (shown in Figure 3.2) is then,

$$F_X(x) = \begin{cases} 0 & , x<0 \\ p_0 & , x=0 \\ p_0+(1-p_0)(1-e^{-\lambda x}) & , x>0 \end{cases} \qquad (3.2)$$

The CDF of a mixed random variable satisfies properties (F1)-(F3) but it does not satisfy property (F4) of Chapter 2 or the property (F4') above.

III.B. THE EXPONENTIAL DISTRIBUTION

This distribution, sometimes called the negative exponential distribution, occurs in applications such as reliability theory and queuing theory. Reasons for its use include its memoryless (Markov)

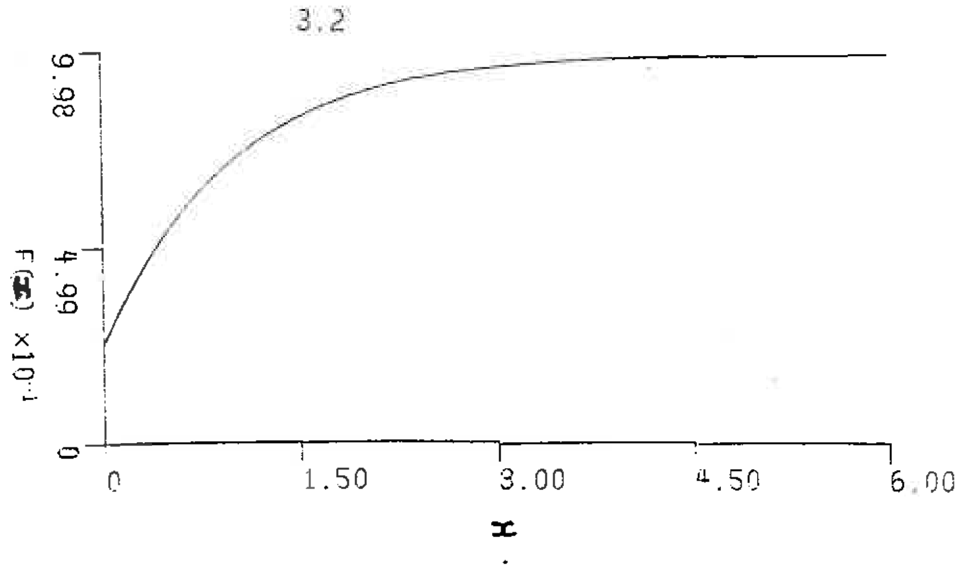

Figure 3.2 : CDF of a Mixed Random Variable.

property (and resulting analytical tractability) and its relation to the (discrete) Poisson distribution. Thus the following random variables will often be modeled to be exponential:

(a) Time between two successive job arrivals to a computing center (often called inter-arrival time).

(b) Service time at a server in a queuing network; the server could be a resource such as the CPU, I/O device, or a communication channel.

(c) Time to failure (lifetime) of a component.

(d) Time required to repair a component that has malfunctioned.

It should be noted that the above distributions are exponential is not a given fact but is an assumption . Experimental verification of this assumption must be sought before relying on the results of the analysis (see Chapter 10 for further elaboration on this topic).

The exponential distribution function, shown in Figure 3.3, is given by

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & , \text{ if } 0 \leq x < \infty \\ 0 & , \text{ otherwise} \end{cases} \qquad (3.3)$$

If a random variable X possesses CDF (3.3), we write $X \sim EXP(\lambda)$, for brevity. The pdf of X has the shape shown in Figure 3.4 and is given by,

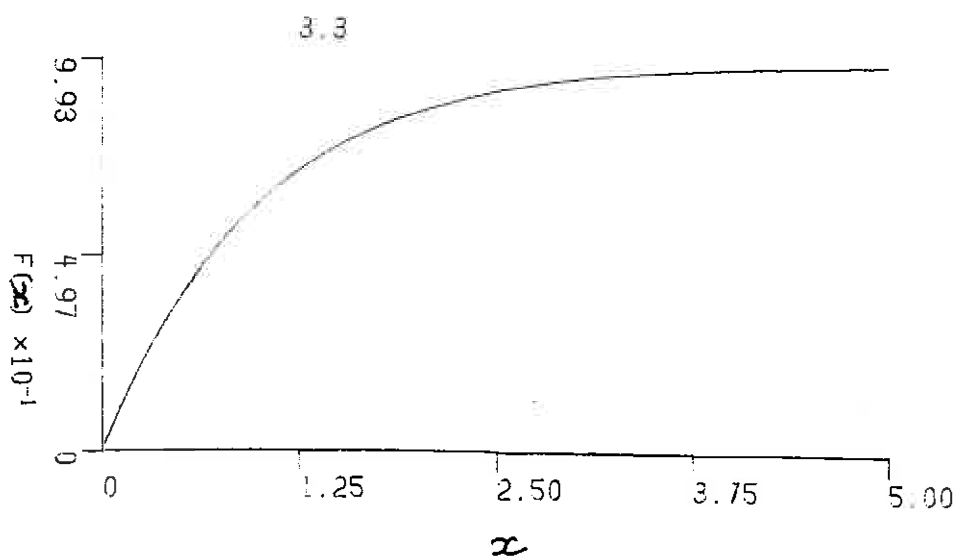

Figure 3.3 : Exponential CDF.

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & , \text{ if } x>0 \\ 0 & , \text{ otherwise} \end{cases} . \qquad (3.4)$$

While specifying a pdf, commonly only the nonzero part is stated, and it is understood that the pdf is zero over any unspecified region. Since $\lim_{x \to \infty} F(x) = 1$, it follows that the total area under the exponential pdf is unity. Also,

$$P(X \geq t) = \int_t^{\infty} f(x)dx$$

$$= e^{-\lambda t} , \qquad (3.5)$$

and

$$P(a \leq X \leq b) = F(b) - F(a)$$

$$= e^{-\lambda a} - e^{-\lambda b} .$$

Now let us investigate the so-called MEMORYLESS or MARKOV PROPERTY of the exponential distribution. Suppose we know that X exceeds some given value t, that is, X>t. For example, if we interpret X as the lifetime of a component, and suppose we have observed that this component has already been operating for t hours. We may then be interested in the distribution of $Y = X - t$, the remaining (residual) lifetime. Let the conditional probability of $Y \leq y$, given that $X > t$, be denoted by $G_t(y)$. Thus for $y >= 0$, we have:

$$G_t(y) = P(Y \leq y | X>t)$$

$$= P(X-t \leq y \mid X>t)$$

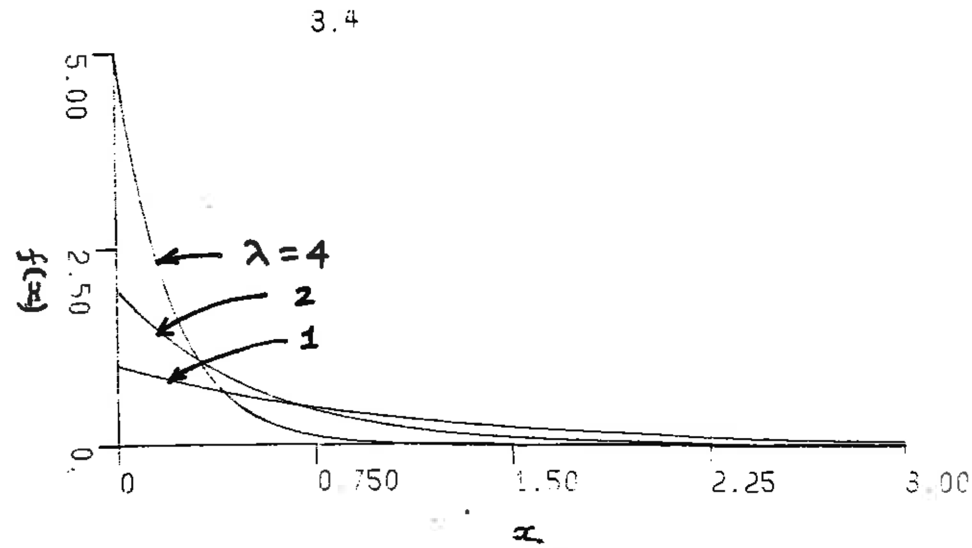

Figure 3.4: Exponential pdf.

$$= P(X \leq y+t \mid X>t)$$

$$= \frac{P(X<y+t \text{ and } X>t)}{P(X>t)}$$, by the definition of conditional probability

$$= \frac{P(t<X<y+t)}{P(X>t)} .$$

Thus (see Figure 3.5):

$$G_t(y) = \frac{\int_t^{y+t} f(x)\,dx}{\int_t^{\infty} f(x)\,dx}$$

$$= \frac{\int_t^{y+t} \lambda e^{-\lambda x} dx}{\int_t^{\infty} \lambda e^{-\lambda x} dx}$$

$$= \frac{e^{-\lambda t}(1 - e^{-\lambda y})}{e^{-\lambda t}}$$

$$= 1 - e^{-\lambda y} .$$

Thus $G_t(y)$ is independent of t and is identical to the original exponential distribution of X . The distribution of the remaining life does not depend on how long the component has been operating. The component does not "age" (it is as good as new) or it "forgets" how long it has been operating and its eventual breakdown is the result of some suddenly-appearing failure, not of gradual deterioration.

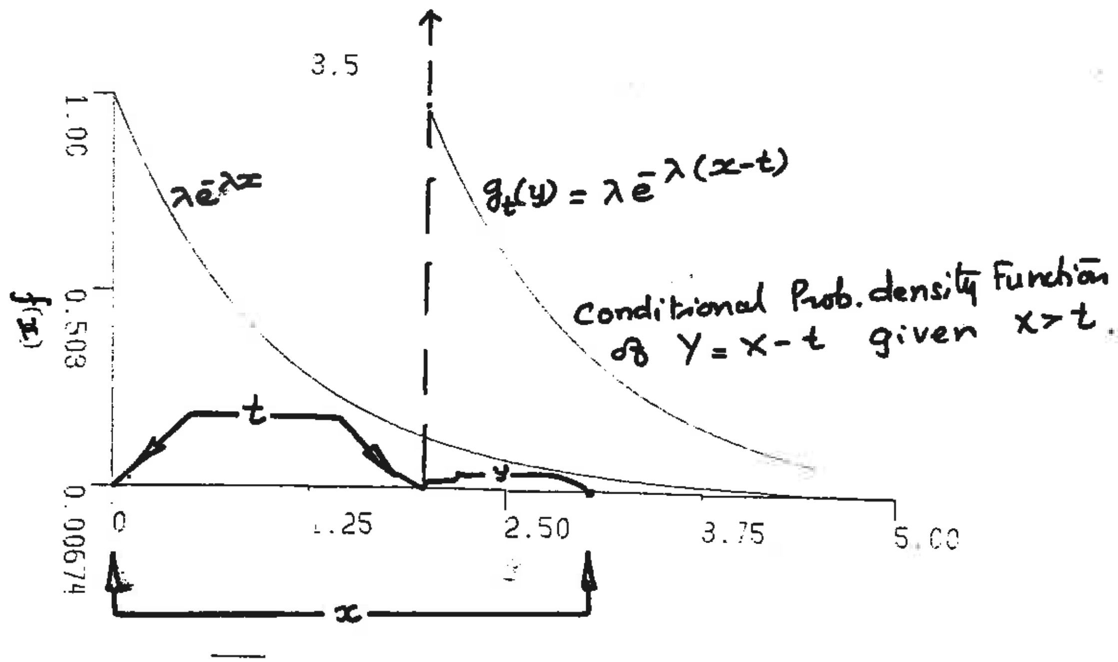

Figure 3.5 : Memoryless Property of the Exponential Distribution.

If the interarrival times are exponentially distributed, then the memoryless property implies that the time we must wait for a new arrival is statistically independent of how long we have already spent waiting for it.

The exponential distribution is the only continuous distribution with the Markov property.

To show this assume that X is a non-negative continuous random variable with the above property. Then

$$\frac{P(t<X<y+t)}{P(X>t)} = P(X\leq y) = P(0<X\leq y)$$

or

$$F_X(y+t) - F_X(t) = (1-F_X(t))(F_X(y)-F_X(0))$$

Since $F_X(0) = 0$, we rearrange the above equation to get

$$\frac{F_X(y+t) - F_X(y)}{t} = \frac{F_X(t)(1-F_X(y))}{t}$$

Taking the limit as t approaches zero, we get

$$F_X'(y) = F_X'(0) \; [1-F_X(y)]$$

where F_X' denotes the derivative of F_X. Let $R_X(y) = 1 - F_X(y)$, then the above equation reduces to

$$R_X'(y) = R_X'(0) \; R_X(y)$$

The solution to this differential equation is given by

$$R_X(y) = K \; e^{R_X'(0)y}$$

where K is a constant of integration and $-R_X'(0) = F_X'(0) = f_X(0)$, the pdf evaluated at 0. Noting that the reliability $R_X(0) = 1$, and denoting the constant $\lambda = f_X(0)$, we get

$$R_X(y) = e^{-\lambda y}$$

and hence

$$F_X(y) = 1 - e^{-\lambda y} \quad , \; y > 0$$

Therefore X must have the exponential distribution.

The exponential distribution can be obtained from the Poisson distribution by considering the interarrival times rather than the number of arrivals.

Example 3.2 - Let the discrete random variable Y_t denote the number of jobs arriving to a computer system, in the interval $(0,t)$. Let X be the time of the first arrival. Further assume Y_t is Poisson distributed with parameter λt, so that λ is the arrival rate. Then:

$$P(X>t) = P(Y_t=0)$$

$$= \frac{e^{-\lambda t}(\lambda t)^0}{0!}$$

$$= e^{-\lambda t}$$

and

$F_X(t) = 1 - e^{-\lambda t}$. Therefore, interarrival times are exponentially distributed.

#

Example 3.3 - Consider a university computer center with an average rate of job submission $\lambda = 0.1$ jobs per second. Assuming that the number of arrivals per unit time is Poisson distributed, the interarrival time, X, is exponentially distributed with parameter λ. The probability that an interval of 10 seconds elapses without job submission is then given by

$$P(X \geq 10) = \int_{10}^{\infty} 0.1e^{-0.1t}dt = \lim_{t \to \infty} [-e^{-0.1t} - e^{-1}] = e^{-1} = 0.368.$$

#

III.C. THE RELIABILITY, FAILURE DENSITY, AND HAZARD FUNCTION

Let the random variable X be the lifetime or the time to failure of a unit. The probability that the unit survives until some time t is called the <u>reliability</u> R(t) of the unit. Thus, $R(t) = P(X>t) = 1 - F(t)$, where F is the distribution function of the component lifetime, X. The unit (component) is assumed to be working properly at time $t = 0$ (i.e., $R(0) = 1$) and no component can work forever without failure (i.e., $\lim_{t \to +\infty} R(t) = 0$). Also, R(t) is a monotone non-increasing function of t. For t less than zero, reliability has no meaning, but we let $R(t) = 1$ for $t < 0$. F(t) will often be called the unreliability.

Consider a fixed number of components, N_0, under test. After time t, $N_f(t)$ have failed and $N_s(t)$ have survived with $N_f(t) + N_s(t) = N_0$. The estimated probability of survival may be written as (using the frequency interpretation of probability):

$$\hat{P}(\text{survival}) = \frac{N_s(t)}{N_0} \ .$$

In the limit as $N_0 \to \infty$, we expect $\hat{P}$(survival) to approach R(t). As the test progresses, $N_s(t)$ gets smaller and R(t) decreases.

$$R(t) \cong \frac{N_s(t)}{N_0}$$

$$= \frac{N_0 - N_f(t)}{N_0}$$

$$= 1 - \frac{N_f(t)}{N_0}$$

The total number of components N_0 is constant while the number of failed components N_f increases with time. Taking derivatives on both sides of the above equation, we get:

$$R'(t) \cong -\frac{1}{N_0} N_f'(t) \qquad (3.6)$$

In equation (3.6), $N_f'(t)$ is the rate at which components fail. Therefore, as $N_0 \to \infty$, the right hand side may be interpreted as the negative of the failure density function, $f_X(t)$:

$$R'(t) = -f_X(t) \quad . \qquad (3.7)$$

→ obvious, because $f_X(t) = F'(t)$ (derivative of unreliability) and $F(t) = 1 - R(t)$