# Module 2-2
# Review of Probability Theory

Chanan Singh
Texas A&M University

# Outline

- Random variable
  - Probability distribution function
  - Survival function
  - Hazard function
  - Exponential distribution function
- Stochastic processes
- Markov process
  - Transition probability
  - Transition rate matrix

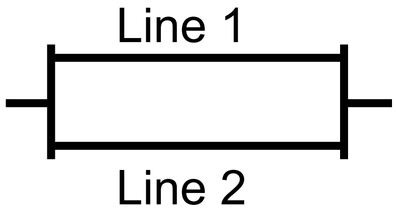
# Random Variable (RV)

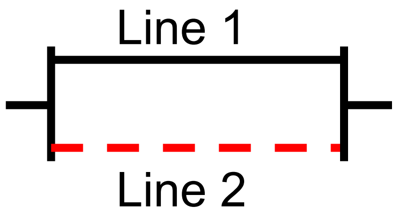
*A function that assigns numerical values to all possible outcomes of a state space*

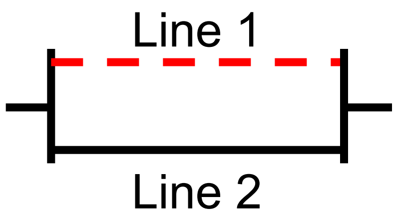
- Discrete RV assigns discrete value.
  - Ex: A number of components down in power system
- Continuous RV assigns continuous value.
  - Ex: Time to failure of a component

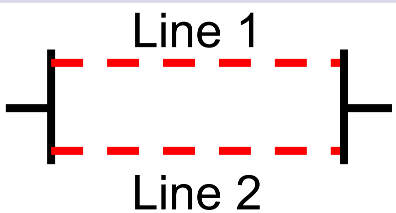
# Discrete RV: Number of T-Lines Down

| Outcomes | Random Variable |
|---|---|
| (1 Up, 2 Up) | 0 |
| (1 Up, 2 Down) | 1 |
| (1 Down, 2 Up) | 1 |
| (1 Down, 2 Down) | 2 |



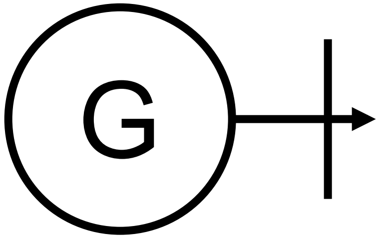



# Continuous RV: Time to Failure

A generator start working at time x = 0




Outcomes: time to failure

It can fail at any time, $x \geq 0$

Random Variable, X

$x \geq 0$

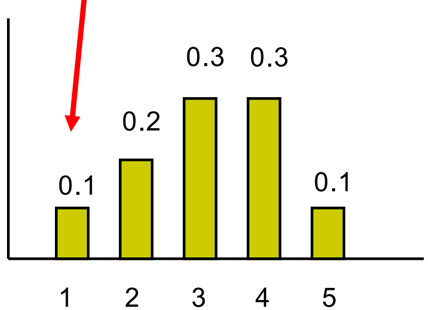
# Probability Distribution Function

*A function that gives probabilities associated with all possible values of a random variable.*

$$f_X(x) = \lim_{\Delta x \to 0} \frac{\Pr(x < X < x + \Delta x)}{\Delta x}$$

- Discrete RV:

$$\Pr(X = 1) = 0.1$$

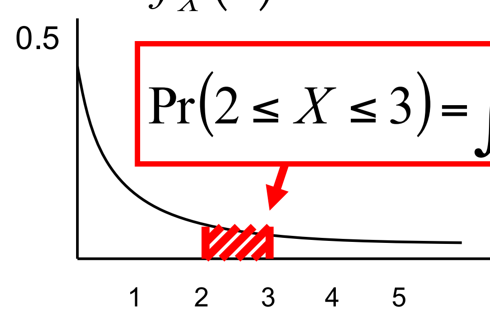



- Continuous RV:

Exponential distribution function

$$f_X(x) = 0.5e^{-0.5x}, x \geq 0$$

$$\Pr(2 \leq X \leq 3) = \int_2^3 0.5e^{-0.5x} dx$$



# Discrete Random Variable

- Probabilty density or mass function

Properties

$p_X(x) = 0$ unless $x$ is one of the values of the discrete rv, $x_1, x_2, \ldots$

$0 \leq p_X(x_i) \leq 1$

$\sum_i p_X(x) = \sum_i P(X = x_i) = 1$

# Continuous Random Variable

$$P[a \le X \le b] = \int_a^b f_X(y)\,dy$$

Properties

i $f_X(x)$ is non-negative

ii $\int_{-\infty}^{+\infty} f(x)\,dx = 1$

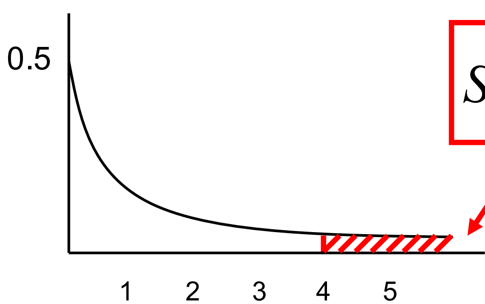
iii The function $f(x)$ is continuous at all but finite number of points, ie, it is piece-wise continuous

# Survival Function

*A function that gives probability of a component survival beyond time x.*

$$S_X(x) = \Pr(X > x)$$

- Time to failure of a component is a random variable.
- This rv is commonly used in reliability theory




$$S_X(4) = \Pr(X > 4) = \int_4^{\infty} 0.5e^{-0.5x}\,dx$$

# Distribution Function

*A function that gives probability of a component failing by time x.*

$$F_X(x) = \Pr(X \leq x)$$

$$F_X(x) = 1 - S_X(x)$$

# Hazard Function

*A function that gives a rate at time x, at which a component fails ( i.e. failure rate), given that it has survived for time x.*

- Denoted by Φ(x),

Probability of a component fails between time x and x+Δx given that it has survived for time x

$$\phi_X(x) = \lim_{\Delta x \to 0} \frac{\Pr\left(x < X < x + \Delta x \mid X > x\right)}{\Delta x}$$

# Hazard Function

- From,

$$\phi_X(x) = \lim_{\Delta x \to 0} \frac{\Pr(x < X < x + \Delta x \mid X > x)}{\Delta x}$$

- We have

$$\phi_X(x) = \lim_{\Delta x \to 0} \frac{\Pr(x < X < x + \Delta x)}{\Delta x} \cdot \frac{1}{\Pr(X > x)}$$

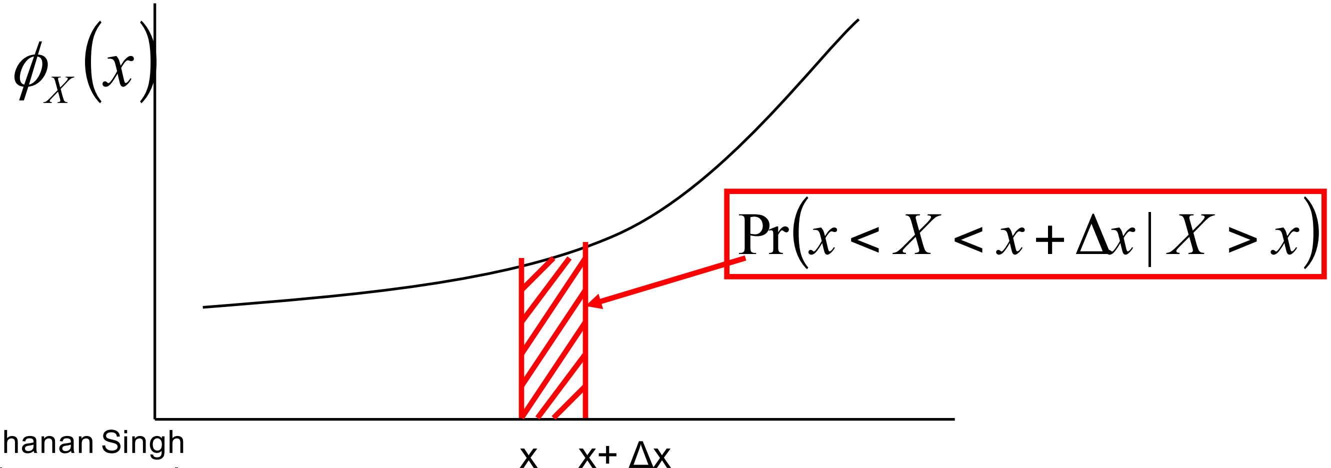
$$\phi_X(x) = \frac{f_X(x)}{S_X(x)}$$

# An Important Relationship

- From hazard function as $\Delta x \rightarrow 0$,

$$\phi_X(x)\Delta x = \Pr(x < X < x + \Delta x \mid X > x)$$

- This gives probability of failure of a component in interval (x, x+Δx).



# Important Note

- Although , for simplicity, survival function and hazard rate function have been described with respect to a component failure, they apply to any random variable.
- For example if the random variable is time to repair, then $\Phi(x)$ represents the repair rate

# Expected Value

For discrete random variable

$$E(X) = \sum_i x_i \, P(X = x_i)$$

$$= \sum_i x_i \, p(x_i)$$

For continuous random variable

$$E(X) = \int_{-\infty}^{\infty} x \, f(x) dx$$

## Mean and Average

Consider a random sample of n observations of X. Then the sample mean

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

According to the law of large numbers, for any constant c>0

$$\lim_{n \to \infty} P[|\bar{X} - m| > c] = 0$$

Where m is the mean of x

Thus as sample size increases, the sample mean approaches the mean value of rv.

# Sum of Random Variables

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$

This result holds even when the random variables are not independent.

# Variance

$$V(X) = E\left[\left(X - E(X)\right)^2\right]$$

$$= \sum_i (x_i - m)^2 \, p_i(x) \quad \text{for discrete rv}$$

$$= \int_{-\infty}^{\infty} (x - m)^2 f(x)\,dx \quad \text{for continuous rv}$$

Variance provides a measure of spread of the distribution

# Variance

$$
\begin{aligned}
V(X) &= E\left[\left(X - E(X)\right)^2\right] \\
&= E\left[X^2 + \left(E(X)\right)^2 - 2XE(X)\right] \\
&= E(X^2) - (E(X))^2
\end{aligned}
$$

This relationship is easier to implement

# Covariance

A measure of the tendency of two random variables to vary together.

$$Cov(X,Y) = E\left[\left(X - E(X)\right)\left(Y - E(Y)\right)\right]$$
$$= E(XY) - E(X)E(Y)$$

Correlation Coefficient

$$C_{XY} = \frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}}$$

Range (-1, 1)

# Covariance

If X, Y are independent

$$C_{XY} = 0$$

But reverse is not always true

# Expectation of a Real Valued Function g(X)

$$E[g(X)] = \sum_i g(x_i)\, p(x_i) \qquad \text{for drv}$$

$$= \int_{-\infty}^{\infty} g(x) f(x) dx \qquad \text{for crv}$$

# Moments

Let $g(X) = X^k$

Then the kth initial moment of X

$$m_k(X) = m_k$$

$$= \sum_i x_i^k \, p(x_i) \qquad \text{for drv}$$

$$= \int_{-\infty}^{\infty} x^k f(x) dx \qquad \text{for crv}$$

# Moments around Mean or Central Moments

$$M_k(X) = M_k$$

$$= \sum_i (x_i - m)^k \, p(x_i) \qquad \text{for drv}$$

$$= \int_{-\infty}^{\infty} (x - m)^k \, f(x) dx \qquad \text{for crv}$$

# Transform Methods
# Characteristic Function of X

$$\phi_X(\theta) = E\left[e^{i\theta X}\right]$$
$$= \int_{-\infty}^{\infty} e^{i\theta x} f(x)\,dx$$

where $i = \sqrt{-1}$

f(x) = pdf of x

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi(\theta) e^{-i\theta x}\,d\theta$$

One to one correspondence between pdf & CF

# CF and Moments
## Differentiate CF k times

$$\phi^{(k)}(\theta) = \frac{d^k\phi(\theta)}{d\theta^k} = i^k \int_{-\infty}^{\infty} x^k e^{i\theta x} f(x)dx$$

As $\theta \to 0$

$$\phi^{(k)}(0) = i^k \int_{-\infty}^{\infty} x^k f(x)dx$$

$$= i^k m_k$$

or

$$m_k = i^{-k}\phi^{(k)}(0)$$

# CF of Sum of Random Variables

$$Y = X_1 + X_2 + \cdots + X_n$$

Where $X_i$ are independent

$$\begin{aligned}
\phi_Y(\theta) &= E\left(e^{i\theta Y}\right) \\
&= E\left(e^{i\theta(X_1+X_2+\cdots+X_n)}\right) \\
&= E\left(e^{i\theta X_1}\right)E\left(e^{i\theta X_2}\right)\ldots E\left(e^{i\theta X_n}\right) \\
&= \phi_{X_1}(\theta)\phi_{X_2}(\theta)\ldots\phi_{X_n}(\theta)
\end{aligned}$$

# Moment Generating Function

For real numbers

$$\Psi(\theta) = E\left(e^{\theta X}\right)$$

$$= \int_{-\infty}^{\infty} e^{\theta x} f(x)\,dx \qquad \text{for crv}$$

$$= \sum_{i} e^{\theta x_i} p(x_i) \qquad \text{for drv}$$

Here

$$m_k = \Psi^{(k)}(0)$$

# Moment Generating Function
## About mean m (or any point)

$$\Psi_m(\theta) = E\left(e^{(X-m)\theta}\right)$$

$$= e^{-m\theta}\Psi(\theta)$$

The kth moment about mean

$$M_k = \Psi_m^{(k)}(0)$$

# Laplace transform
## (If X is non-negative)

$$L\big(g(t)\big) = \bar{g}(s) = \int_0^{\infty} g(t)e^{-st}dt$$

Compare with moment generating function

$$\theta = -s$$

$$\bar{f}(s) = E(e^{-sX})$$

$$= E[\sum_{k=0}^{\infty}(-1)^k \frac{s^k X^k}{k!}]$$

$$= \sum_{k=0}^{\infty}(-1)^k \frac{s^k}{k!} m_k$$

# Some Special Distributions
## Exponential Distribution

A non-negative rv is said to have negative exponential distribution if probability density function is

$$f(x) = \rho e^{-\rho x}$$

ρ is a positive constant

$$F(x) = \int_0^x \rho e^{-\rho u} du = 1 - e^{-\rho x}$$

$$S(x) = 1 - F(x) = e^{-\rho x}$$

# Some Special Distributions

## Exponential Distribution

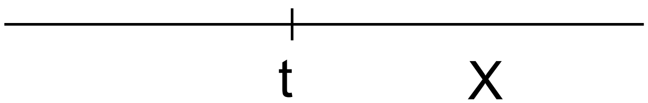
$$\phi(x) = \frac{f(x)}{S(x)} = \rho$$

$$Mean = \frac{1}{\rho}$$

$$Variance = \frac{1}{\rho^2}$$

$$Standard\ dev = \sqrt{Variance} = \frac{1}{\rho}$$

# Distribution of Residual Life Time




$$Y = X - t$$

$$F_Y(x) = P[(X - t) \le x \mid X > t]$$

$$= P[X \le x + t \mid X > t]$$

$$= P[t < X \le x + t] / P[\,X > t]$$

$$= \frac{\int_t^{x+t} \rho e^{-\rho u} du}{e^{-\rho t}}$$

$$= 1 - e^{-\rho x}$$

$$= F_X(x)$$

# Poisson Distribution

If the number of events is given by Poisson distribution, in time t

$$p_k(t) = P(No\ of\ events = k)$$

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

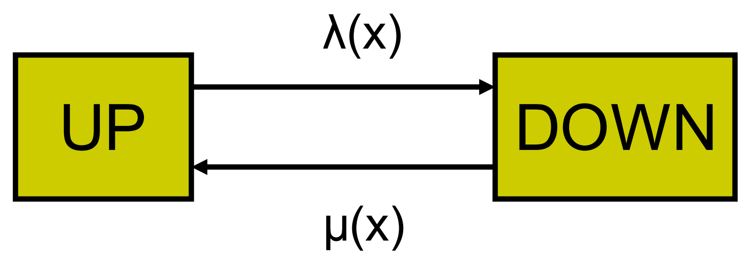
So probability of no failures during (0,t)

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

This is exponential distribution

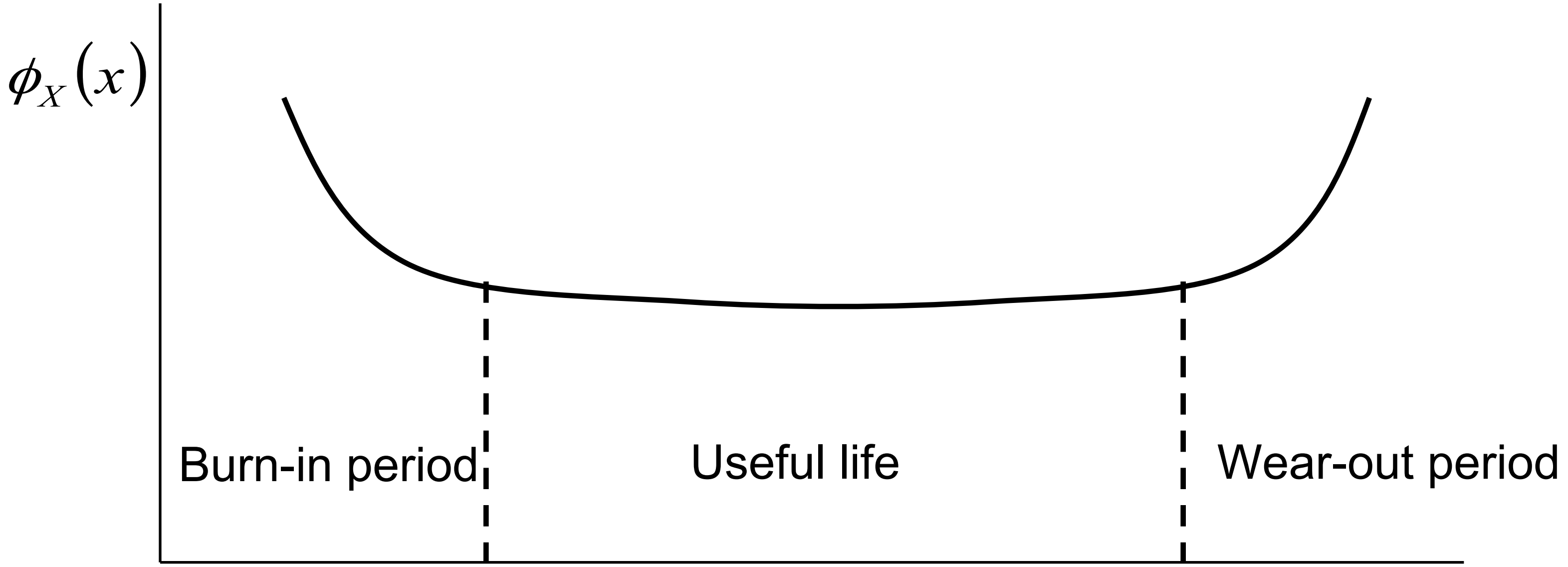
# A Two-State Component

- Consider a two-state component
- $\lambda(x)$ is a hazard function of the up time, called failure rate.
- $\mu(x)$ is a hazard function of the down time, called repair rate.
- Generally, hazard function is called '*transition rate*' in reliability work.



# A Bath Tub Curve

- Typical hazard function of a component.




- It is fairly common to assume constant transition rates in reliability modeling.

# Exponential Distribution Function

- Non-negative continuous random variable
- Commonly used to represent up time of a component

$$f_X(x) = \lambda e^{-\lambda x}, x \geq 0$$

$$S_X(x) = e^{-\lambda x}, x \geq 0$$

$$\boxed{\phi_X(x) = \lambda, x \geq 0}$$

- If we assume up time and repair time of a component are exponentially distributed, the failure rate and repair rate are *constant*.