Continuous Random Variables

Daniel A. Menascé, Ph.D.
Department of Computer Science
George Mason University

1

Relevant Functions

- Probability density function (pdf) of r.v. X: $f_X(x)$

$$P[a \leq X \leq b] = \int_a^b f_X(x)dx$$

- Cumulative distribution function (CDF):

$$F_X(x) = P[X \leq x]$$

- Tail of the distribution (reliability function):

$$R_X(x) = P[X > x] = 1 - F_X(x)$$

2

Moments

- k-th moment: $E[X^k] = \int_{-\infty}^{+\infty} x^k f_X(x)dx$
- Expected value (mean): first moment

$$m = E[X] = \int_{-\infty}^{+\infty} x f_X(x)dx$$

- k-th central moment:

$$E[(X-m)^k] = \int_{-\infty}^{+\infty} (x-m)^k f_X(x)dx$$

- Variance: second central moment

$$s^2 = E[(X-m)^2] = \int_{-\infty}^{+\infty} (x-m)^2 f_X(x)dx$$

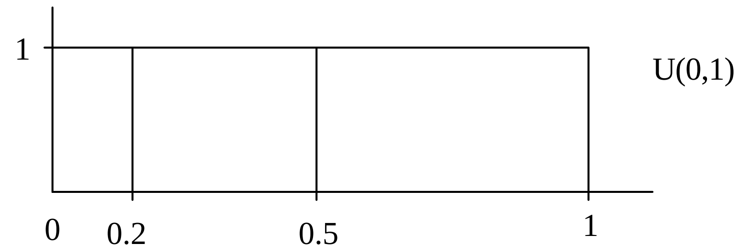
The Uniform Distribution

- pdf: $f_X(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$

- Mean: $m = \dfrac{a+b}{2}$

- Variance: $s^2 = \dfrac{(b-a)^2}{12}$

The Uniform Distribution



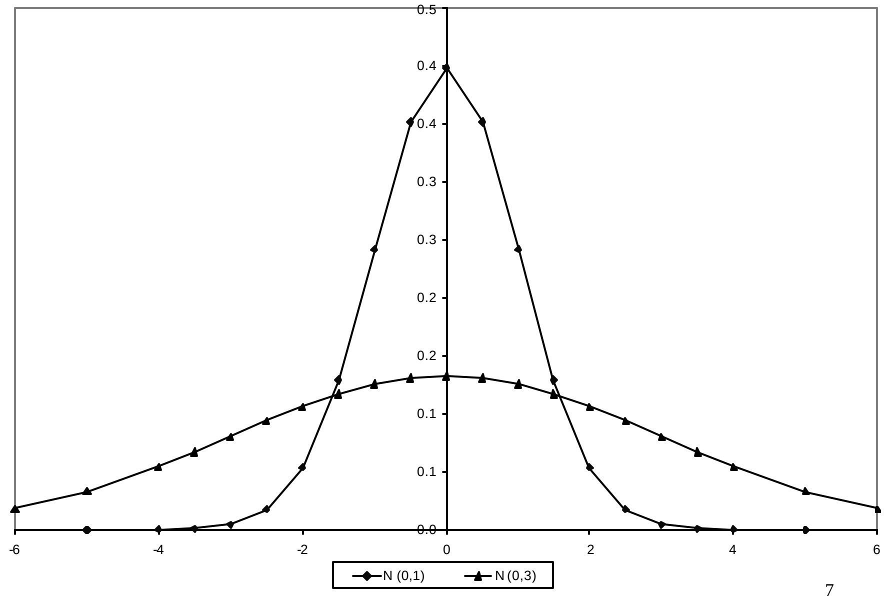
P[0.2<X<0.5]=(0.5-0.2)x1.0=0.3

The Normal Distribution $N(m, s)$

- Many natural phenomena follow a normal distribution.
- The normal distribution can be used to approximate the binomial and the Poisson distributions.
- Two parameters: mean and standard deviation.

$$f_X(x) = \frac{1}{\sqrt{2p}\, s} e^{-(1/2)[(x-m)/s]^2}$$

The Normal Distribution $N(m, s\ )$



7

The Standard Normal Distribution

- To use tables for computing values related to the normal distribution, we need to standardize a normal r.v. as

standard normal score

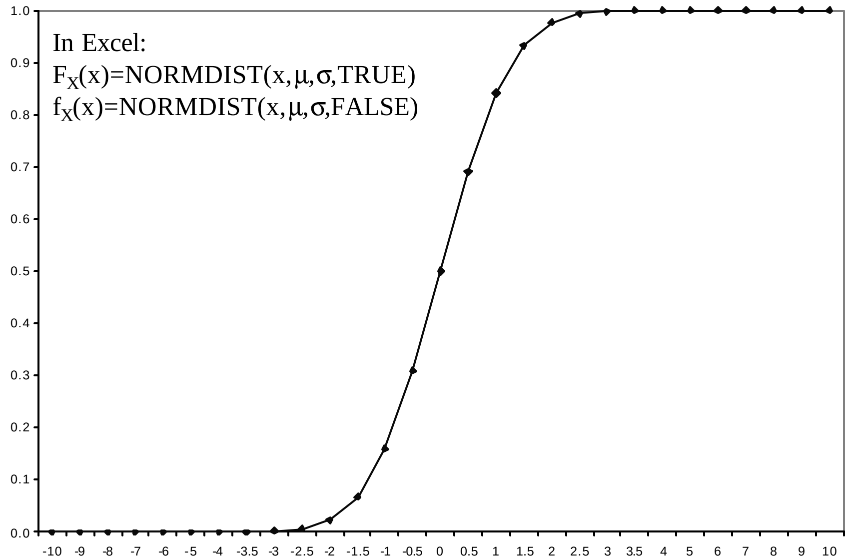
$$Z = \frac{X - m}{s}$$

- Given X, compute a Z value z.
- Find the area value in a Table (Prob [0<Z<z]).

8

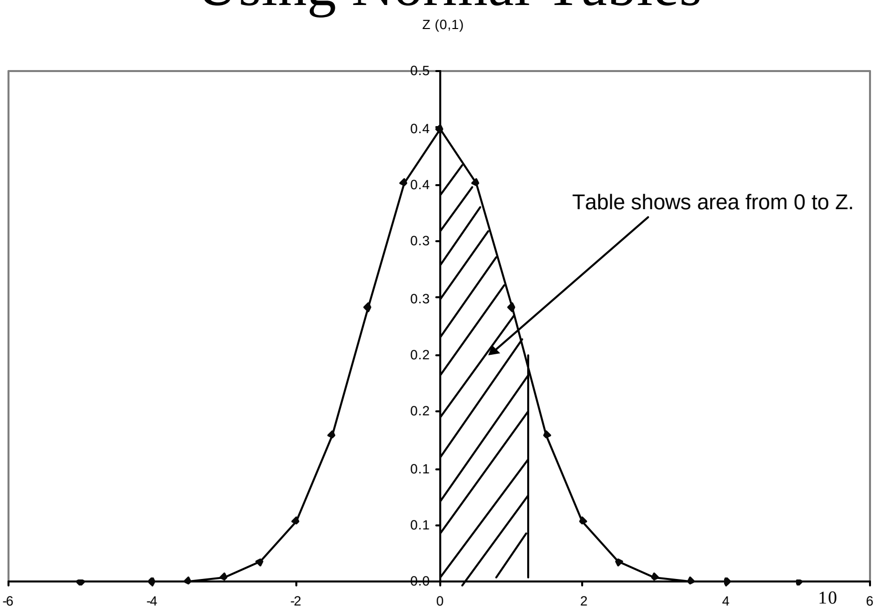
Normal CDF

In Excel:
$F_X(x)$=NORMDIST(x,μ,σ,TRUE)
$f_X(x)$=NORMDIST(x,μ,σ,FALSE)



9

Using Normal Tables



10

The Normal as an Approximation to the Binomial Distribution

- The normal can approximate the binomial if the variance of the binomial

$$np(1-p) \geq 10$$

- Binomial: $m = np$

$$s = \sqrt{np(1-p)}$$

- Transformation: $Z = \dfrac{X - np}{\sqrt{np(1-p)}}$

11

The Normal as an Approximation to the Binomial Distribution

- Consider a binomial r.v. X with average 50 and variance 25. What is $P[50 \leq X \leq 60]$?
- Transformation: $Z = \dfrac{X - 50}{\sqrt{25}} = \dfrac{60 - 50}{5} = 2.0$
- Using the table, the area between 50 and 60 for Z=2.0 is 0.4772. So,

$$P[50 \leq X \leq 60] = 0.4772$$

12

The Normal as an Approximation to the Poisson Distribution

- The normal can approximate the Poisson distribution if λ > 5.

- Poisson: $$m = l$$ $$s = \sqrt{l}$$

- Transformation: $$Z = \frac{X - l}{\sqrt{l}}$$

13

The Lognormal Distribution

- It is a random variable such that its natural logarithm has a normal distribution.

$$f_X(x) = \frac{1}{x\sqrt{2p}\,s_{\ln X}} e^{-(1/2)[(\ln x - m_{\ln X})/s_{\ln X}]^2} \qquad x > 0$$

Y = ln X (X and Y are r.v.'s) and Y = N(μ,σ)

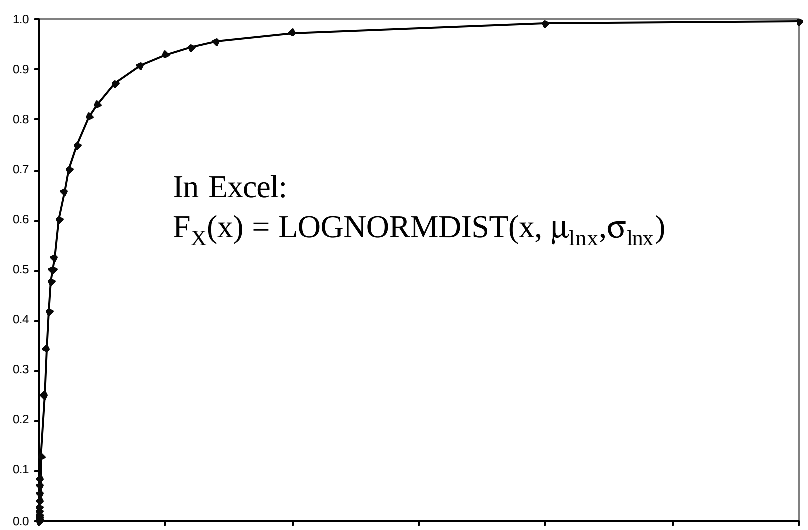
14

The Lognormal distribution

- Mean: $E[X] = e^{m_{\ln X} + s^2_{\ln X}/2}$
- Standard Deviation: $s = \sqrt{e^{2m_{\ln X} + s^2_{\ln X}}.(e^{s^2_{\ln x}} - 1)}$

15

Lognormal CDF



In Excel:
$F_X(x)$ = LOGNORMDIST(x, μ_{lnx},σ_{lnx})

16

The Exponential Distribution

- Widely used in queuing systems to model the inter-arrival time between requests to a system.
- If the inter-arrival times are exponentially distributed then the number of arrivals in an interval *t* has a Poisson distribution and vice-versa.

$$f_X(x) = l e^{-l.x} \qquad F_X(x) = 1 - e^{-l.x} \qquad x \geq 0$$

17

The Exponential Distribution

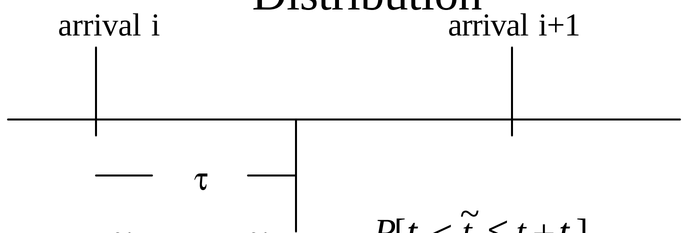
- Mean and Standard Deviation:

$$m = s = 1/\,l$$

- The COV is 1. The exponential is the only continuous r.v. with COV=1.
- The exponential distribution is "memoryless." The distribution of the residual time until the next arrival is also exponential with the same mean as the original distribution.

18

Memoryless Property of the Exponential Distribution



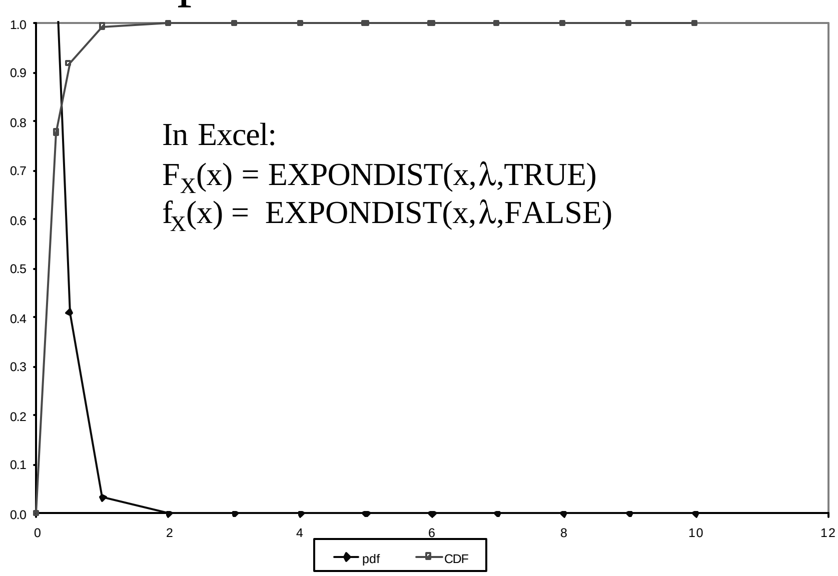
$$P[\tilde{t} \leq t + t \mid \tilde{t} > t] = \frac{P[t < \tilde{t} \leq t + t]}{P[\tilde{t} > t]}$$

$$= \frac{P[\tilde{t} \leq t + t] - P[\tilde{t} \leq t]}{P[\tilde{t} > t]}$$

$$= \frac{1 - e^{-l.(t+t)} - (1 - e^{-l.t})}{1 - (1 - e^{-l.t})}$$

$$= 1 - e^{-l.t}$$

Exponential Distribution



In Excel:
$F_X(x)$ = EXPONDIST(x,λ,TRUE)
$f_X(x)$ = EXPONDIST(x,λ,FALSE)

Pareto Distribution

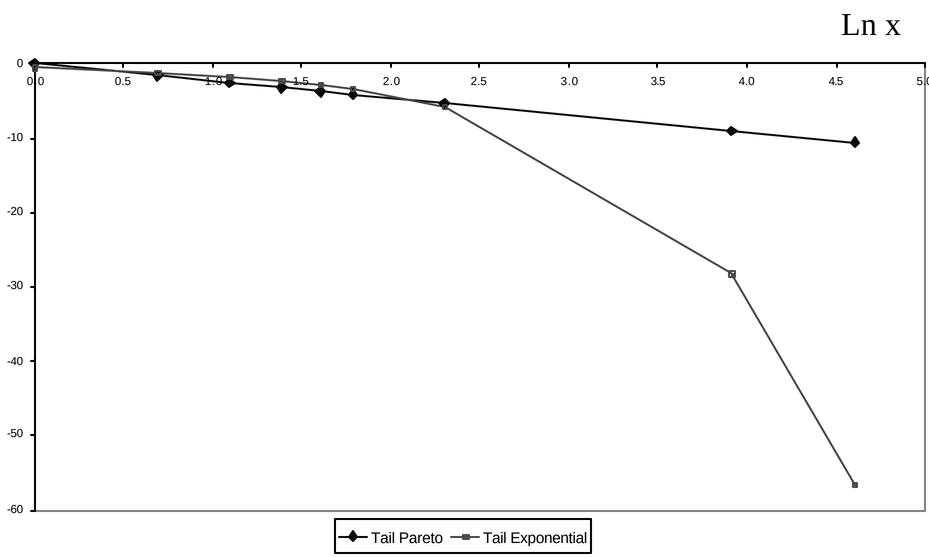
- A case of a heavy-tailed distribution.
- The probability of large values is not negligible.

$$f_X(x) = \frac{a}{x^{1+a}} \qquad a > 0 \quad , \quad x \geq 1$$

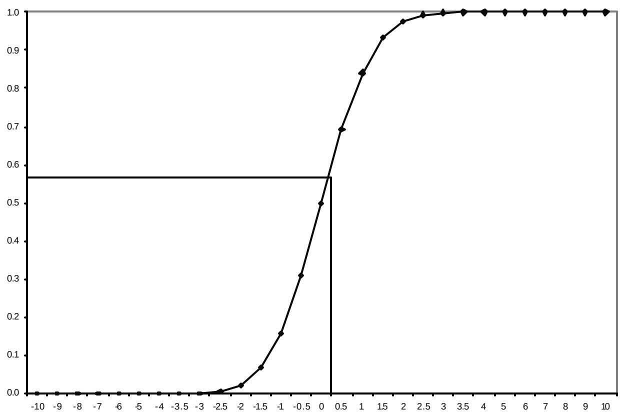
$$F_X(x) = 1 - \frac{1}{x^a} \qquad a > 0 \quad , \quad x \geq 1$$

- Mean: $\frac{a}{a-1} \qquad a > 1$
- Variance: $\frac{a}{(a-1)^2(a-2)} \qquad a > 2$

Tail of the Pareto and Exponential Distributions



Generation of Random Variables



- randomly generate a number u = U(01,)
- $x = F^{-1}(u)$ where F is the CDF

23