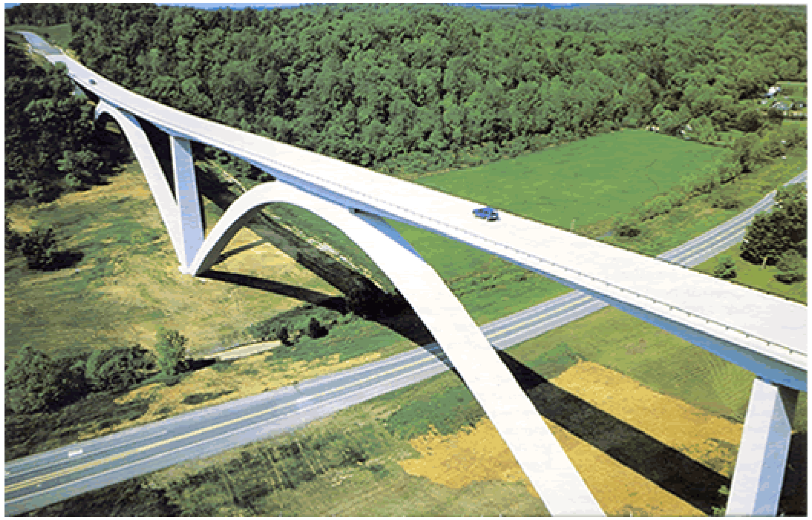
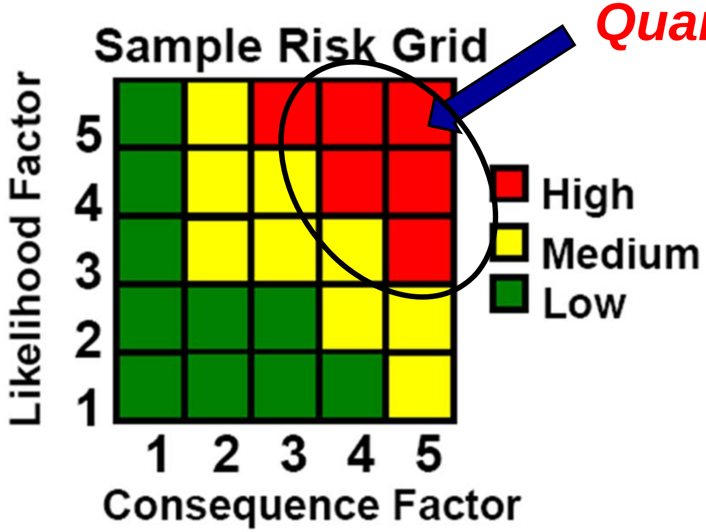
Quantitative Risk Analysis Using Probability Distributions



COORDINATOR:
Dr Lihai Zhang
Department of Infrastructure Engineering
lihzhang@unimelb.edu.au

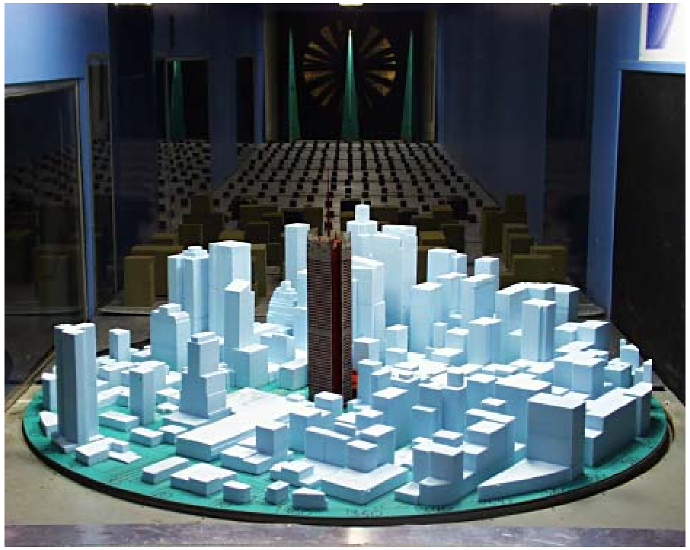
Why Quantitative Risk Analysis ?

Quantitative risk analysis is required !



Qualitative analysis

Hurricane Risks



Wind Tunnel Test

Risks in society

- Selected risks in society (Robert E. Melchers, 2002)

Activity	Approx. death rate (10^{-9} deaths/h exposure)	Typical exposure (h/year)	Typical risk of death (10^{-6}/year)
Construction works	70 ~ 200	2200	150 ~ 440
Coal mining (UK)	210	1500	300
Building fires	1 ~ 3	8000	8 ~ 24
Structural failures	0.02	6000	0.1
Smoking	2500	400	1000
Air travel	1200	20	24
Car travel	700	300	200
Alpine climbing	30,000 ~ 40,000	50	1500 ~ 2000

How an insurance company determines your premiums?

Risks in society

- Typical collapse failure rates for structures

  (Robert E. Melchers, 2002)

Structural type	Data cover	Average life (years)	Probability of failure
Apartments	Demark	30	0.000003%
Mixed housing	Canada	50	0.1%
Large suspension bridge	World	40	0.3%

OHS Risks

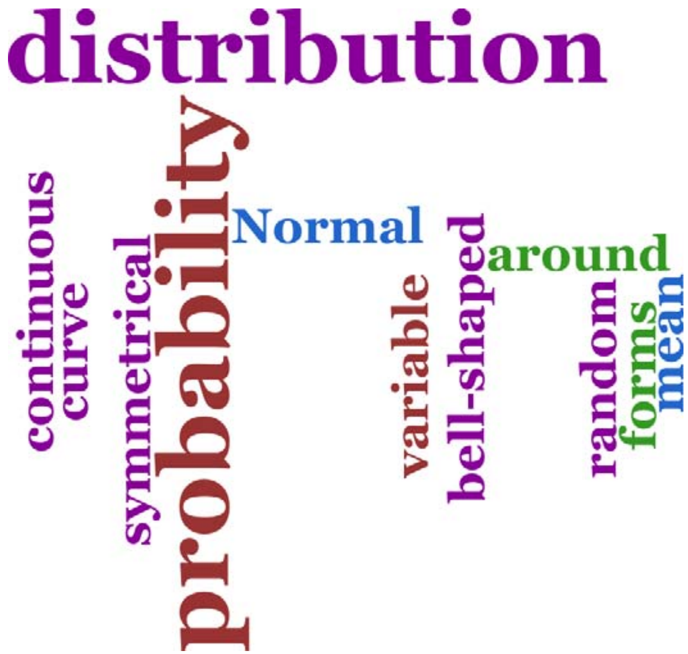
For a large construction project, the contractor estimates that the average rate of on-the-job accidents is three times per year. From past experience, the contractor also estimates that the cost incurred for each accident may be modeled as a lognormal random variable with a median of $6,000 and COV of 20%. The cost of each accident can be assumed to be statistically independent.

(1) What is the probability that there will be no accident in the first month of construction?
(2) What is the probability that an accident will incur a loss exceeding $4,000?

Risk Analysis using Distributions

- **Distributions** where **random variable** takes on a number of specific values with certain probabilities

Discrete **Continuous**



Risk Analysis using Distributions

- **Discrete**
  - **Example: Fair Coin**

    A fair coin is flipped, X to be the random variable, "head" to be 1, and "tail" to be 0. What is the probability that the coin is a head

    P (X=1) =50%, P (X=0) =50%,

- **Continuous**
  - **Example:**

    The number of floods in a given year at a particular location.

    The strength of a concrete cylinder

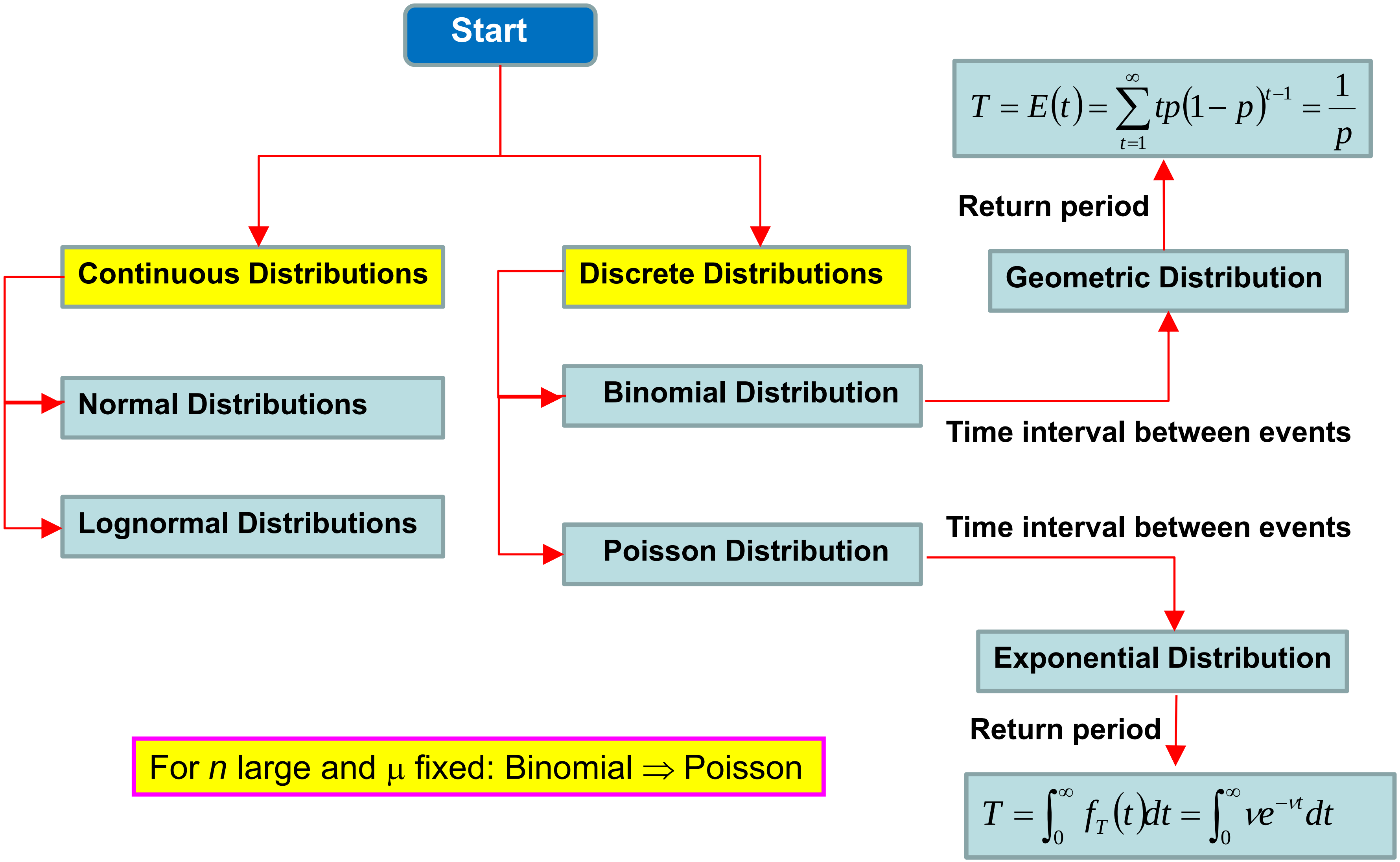
    The distance between cracks in a roadway.

    The inches of rainfall during a storm.

Risk Analysis using Distributions

- **Continuous Probability Distributions**
  - **Normal or Gaussian Distribution**
  - **Lognormal Distribution**
- **Discrete Probability Distributions**
  - **Binomial Distribution**
  - **Poisson Distribution**

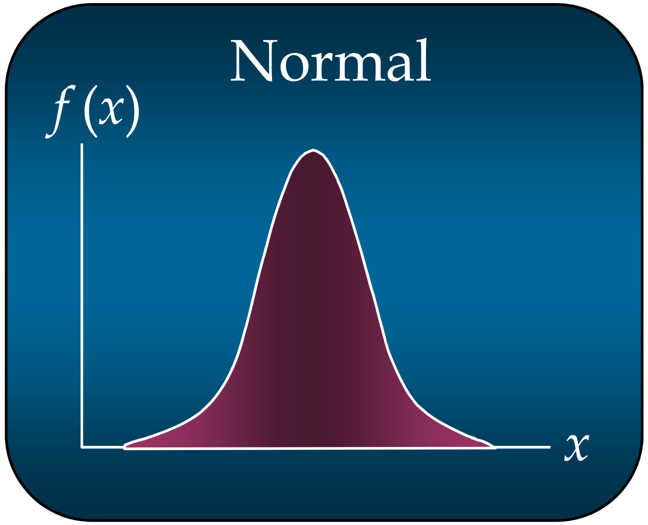
Risk Analysis using Distributions



- **Normal or Gaussian Distribution**

  - The most important distribution for describing a continuous random variable.

  - The normal (or Gaussian) distribution is a continuous probability distribution that has a bell-shaped probability density function.



Quality Risks

Consider the values of the Young's modulus given in below.

Test no.	E (MPa)	Test no.	E (MPa)	Test no.	E (MPa)
1	199,258	12	202,016	23	220.632
2	201,326	13	202,016	24	230,284
3	188.916	14	193,742	25	210,979
4	197,879	15	208,221	26	225,458
5	195,811	16	208,221	27	215,805
6	206,153	17	208,911	28	210,290
7	208,221	18	215,116	29	215,805
8	203,395	19	198,569	30	199,947
9	204,084	20	190,295	31	202,705
10	195,811	21	204,084	32	195,121
11	195,121	22	178,574	33	210,290

Quality Risks

example01.dat

	A VarName1 NUMBER
1	199258
2	201326
3	188916
4	197879
5	195811
6	206153
7	208221
8	203395
9	204084
10	195811
11	195121
12	202016
13	202106
14	193742
15	208221
16	208221
17	208911
18	215116
19	198569
20	[illegible]

EDITOR PUBLISH VIEW

example.m

```matlab
clear all;
close all;
clc;

data = load('example01.dat');

%mean
mu = mean(data);

%standard deviation
stdev=std(data);

%display results
display(mu);
display(stdev);
```

Command Window

```
mu =

   2.0392e+05


stdev =

   1.0390e+04

>>
```

Normal Probability Distributions

- The normal probability density function (PDF) is

$$f(x) = \frac{1}{\boxed{\sigma}\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\boxed{\mu}}{\sigma}\right)^2}$$

where

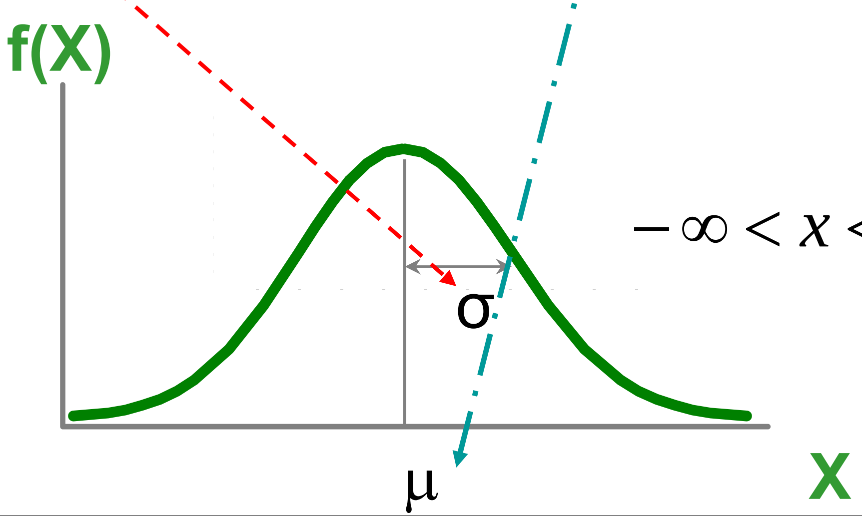
μ : mean

σ : standard deviation

π = 3.14159

e = 2.71828

$-\infty < x < +\infty$

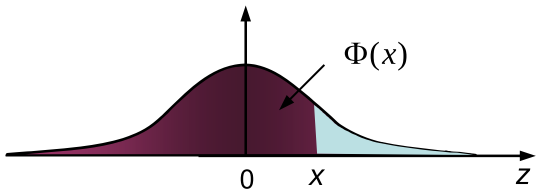


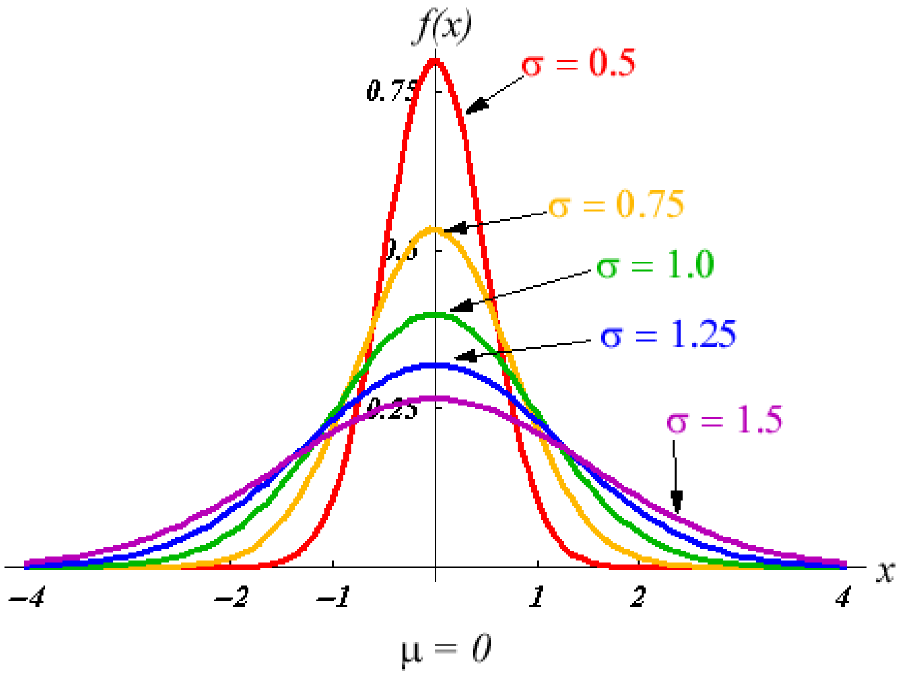
Standard Normal Probability Distribution

- The probability density function (PDF) is

$$f_s(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}z^2}$$

$$Z = \frac{x-\mu}{\sigma}$$



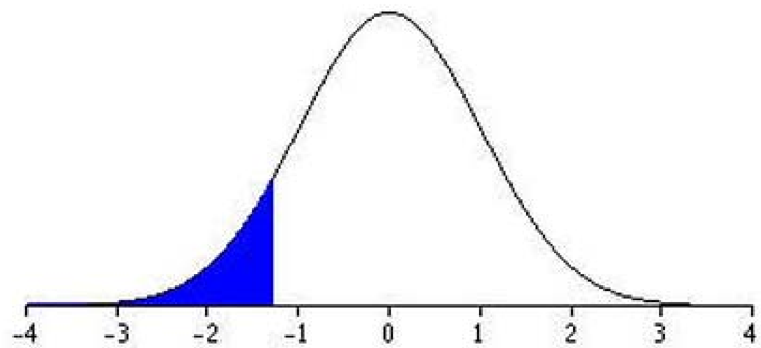


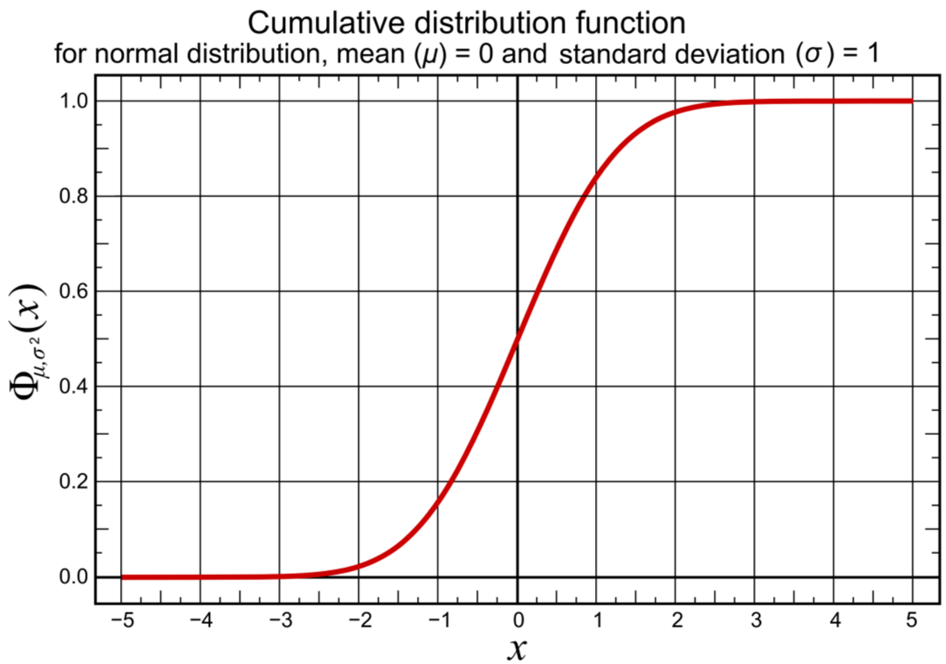
Standard Normal Probability Distribution

- The cumulative distribution function (CDF) is

$$\Phi(z) = F_z(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}z^2} dz$$

$$P(a < T < b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

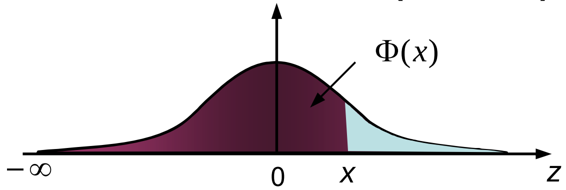




Continuous Distributions

Table of the cumulative distribution function (CFD)

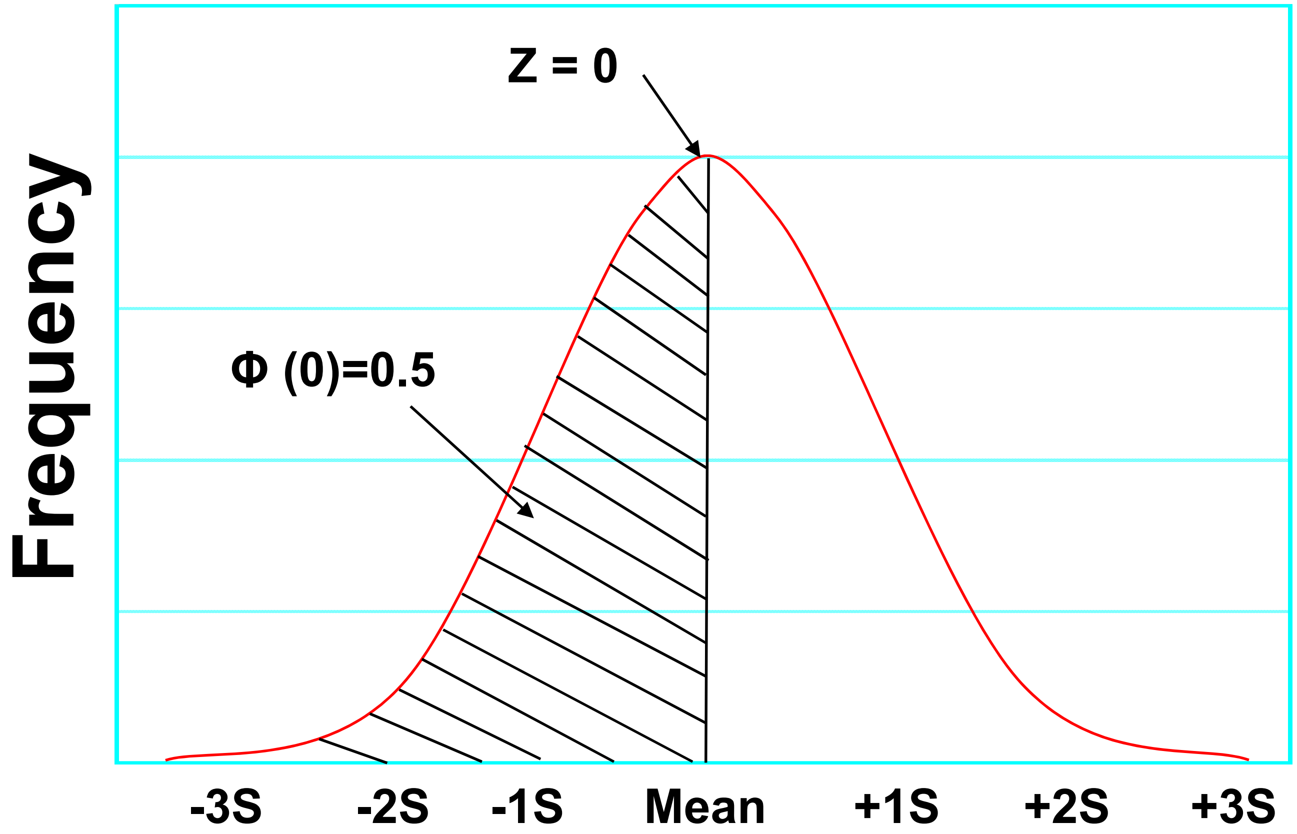
$$\Phi(x) = \frac{1}{\sqrt{2}} \int_{-\infty}^{x} e^{-(z^2/2)} dz$$



x	Φ(x)
0.00	0.50000
0.01	0.50399
0.02	0.50798
0.03	0.51197
0.04	0.51595
0.05	0.51994
0.06	0.52392
0.07	0.52790
0.08	0.53188
0.09	0.53586

x	Φ(x)
1.00	0.84134
1.01	0.84375
1.02	0.84614
1.03	0.84849
1.04	0.85083
1.05	0.85314
1.06	0.85543
1.07	0.85769
1.08	0.85993
1.09	0.86214

Standard Normal Probability Distribution



x	$\Phi(x)$
0.00	0.50000
0.01	0.50399
0.02	0.50798
0.03	0.51197
0.04	0.51595
0.05	0.51994
0.06	0.52392
0.07	0.52790
0.08	0.53188
0.09	0.53586

Example 1

Assume that the randomness in Young's modulus of steel E can be described by normal random variable. Calculate the probability of E having a value between 193,053 MPa and 203,395 MPa.

Quality Risks

Consider the values of the Young's modulus given in below.

Test no.	E (MPa)	Test no.	E (MPa)	Test no.	E (MPa)	Test no.	E (MPa)
1	199,258	12	202,016	23	220.632	34	214,426
2	201,326	13	202,016	24	230,284	35	202,016
3	188.916	14	193,742	25	210,979	36	188,916
4	197,879	15	208,221	26	225,458	37	202,016
5	195,811	16	208,221	27	215,805	38	202,016
6	206,153	17	208,911	28	210,290	39	215,805
7	208,221	18	215,116	29	215,805	40	189,605
8	203,395	19	198,569	30	199,947	41	202,705
9	204,084	20	190,295	31	202,705		
10	195,811	21	204,084	32	195,121		
11	195,121	22	178,574	33	210,290		

Solution

Example (continued)

Assume the design value of Young's modulus for steel is 199,947 MPa, calculate

(1) The probability of the Young's modulus being **less than** the design value.

(2) The probability that Young's modulus will be **at least** the design value.

Example 2

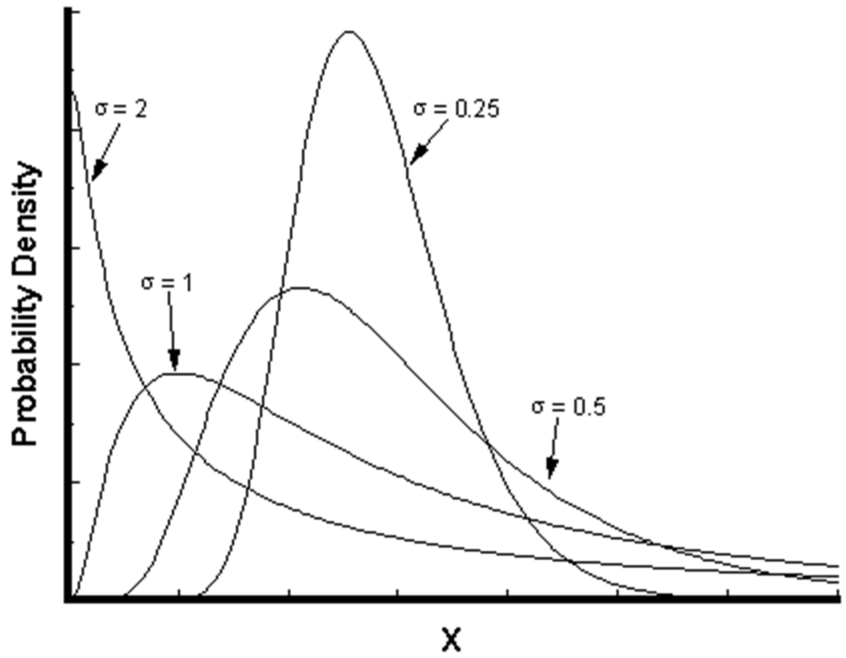
Suppose a steel cable has to carry a weight of **44.5 kN**. Information on the strength of similar cables indicates that the strength of the cable, *R*, can be modeled by a normal random variable with a **mean of 111.2kN** and a **standard deviation of 22.2 kN**. Calculate the probability that the cable will break.

Solution

• Lognormal Distribution

- A log-normal distribution is a probability distribution of a random variable whose logarithm is normally distributed

- In many engineering problems, a random variable cannot have negative value due to the physical aspects of the problem.



Lognormal Probability Distributions

- The Lognormal probability density function (PDF) is

$$f(x) = \begin{cases} \dfrac{1}{\sigma x\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\frac{\ln x - \mu}{\sigma})^2} & if \quad x > 0 \\ 0 & if \quad x \leq 0 \end{cases}$$

where

μ : mean

σ : standard deviation

π = 3.14159

e = 2.71828

Lognormal Probability Distribution

- The cumulative distribution function (CDF) is

$$Z = \frac{\ln x - \lambda_X}{\xi_X}$$

$$P(a < T < b) = F(b) - F(a)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{\ln a - \lambda_X}{\xi_X}}^{\frac{\ln b - \lambda_X}{\xi_X}} e^{-\frac{1}{2}z^2} dz = \Phi\left(\frac{\ln b - \lambda_X}{\xi_X}\right) - \Phi\left(\frac{\ln a - \lambda_X}{\xi_X}\right)$$

$$\lambda_X = \ln \mu_X - \frac{1}{2}\xi_X^2 \qquad \xi_X^2 = \ln\left[1 + \left(\frac{\sigma_X}{\mu_X}\right)^2\right] = \ln\left(1 + COV^2\right)$$

If $COV = \sigma / \mu \le 0.3$ $\qquad \xi_X = COV$

Example 3

The Young's modulus example with a mean of 29,576 Pa and a standard deviation of 1507 Pa can be considered.

It is assumed that the Young's modulus is log-normally distributed. Calculate

(1) The probability of *E* being less than the design value of 29,500 Pa.

(2) The probability of *E* having a value between 28,000 Pa and 29,500 Pa.

Discrete Probability Distributions

- Binomial distribution
- Poisson distribution

Discrete Distribution – Binomial

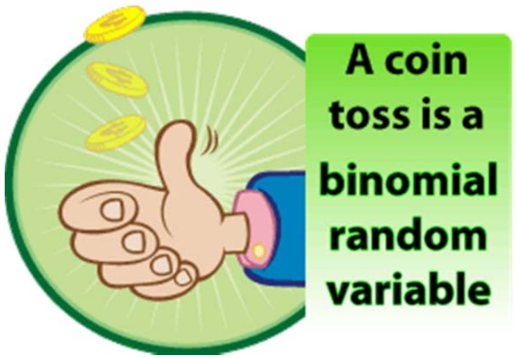
A random trail

Two possible outcomes: **success** | **failure**

probability: **p** | **1- p**

X – total number of successes, $X \geq 0$

S – sample space, $S = \{ 0, 1, 2, . . . , n\}$



The probability distribution of random variable **X** is given by the

binomial distribution.

Discrete Distribution – Binomial

- Binomial Mean $= n \times p$
- Binomial Standard Deviation $= \sqrt{n \times p \times (1 - p)}$
- Binomial probability formula

$$P(X = x, n | p) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad x = 0, 1, 2, \ldots, n$$

By definition $$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

where $n! = n(n-1)(n-2), \ldots 1$, and $0! = 1$

n: Bernoulli trials p: probability

x: exactly x successes out of n Bernoulli trials

Example 4

A random sample of 15 valves is observed. From past experience, it is known that the probability of a given failure within 500 hours following maintenance is 0.18.

Calculate the probability that these valves will experience 0, 1, and 2 independent failures within 500 hours following their maintenance.

Discrete Distribution – Binomial

Solution:

Risk of Building Collapsing

Example 5

Suppose the probability of failure of a structure due to earthquakes is estimated as 10^{-5} per year. Assuming that the design life of the structure is 50 years and the probability of failure in each year remains constant and independent during its lifetime.

Estimate the probability of no failure using the binomial distribution.

Risk of Building Collapsing

Solution:

Risk of Flooding

Example 6

The drainage system of a city has been designed for a rainfall intensity that will be exceeded on an average once in 50 years.

What is the probability that the city will be flooded in 2 out of 10 years?

Risk of Flooding

Solution:

Geometric Distribution

- The first occurrence time of an event is of great interest in Engineering.
  - The first time the design wind speed will be exceeded in an area
  - The first time a structure will be damaged by earthquakes

Geometric Distribution

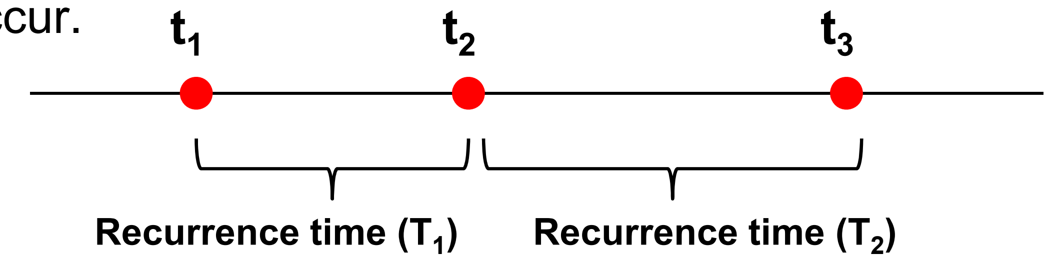
$$P(T = t) = p(1 - p)^{t-1}, t = 1,2,.....$$

- The events occur in a Bernoulli sequence
- *p* is the probability of occurrence in the each trial, then the probability that the event will occur for the first time at the *t*-th trial.
- This is no occurrence in the previous (*t* -1) trials.

THE UNIVERSITY OF MELBOURNE

Return Period - Binomial

- Return Period (also known as a recurrence interval)
  - An estimate of the likelihood of an event, such as an earthquake, flood or a river discharge flow to occur.



  - Assuming a Bernoulli sequence, recurrence time must follow the probabilistic characteristics of the first occurrence (i.e. the geometric distribution)
  - Mean recurrence time (**Return period**)

$$T = E(t) = \sum_{t=1}^{\infty} t p_T(t) = \sum_{t=1}^{\infty} t p (1-p)^{t-1} = p \underbrace{\left[1 + 2(1-p) + 3(1-p)^2 + 4(1-p)^3 + \ldots..\right]}_{1/p^2} = p \times \frac{1}{p^2} = \frac{1}{p}$$

Example: A return period of 50 years for the design flood level indicates that on average there will be a flood once every 50 years. However, there is a probability that no flood will occur in the next 50 years.

Example 7

Suppose the probability of failure of a structure due to earthquakes is estimated as 10^{-5} per year. Assuming that the probability of failure in each year remains constant and independent during its lifetime.

Estimate the probability of the failure of the structure due to earthquakes for the first time in the 10th year using the geometric distribution.

Risk of Structure Collapsing

Solution:

Discrete Distribution – Poisson

- Poisson distribution fits cases of **rare events** that occur in a fixed amount of time **OR** in a specified region
- Poisson random variable X is successes in a time interval.
- The Probability Mass Function (PMF) is

$$P(x\ occurences\ in\ time\ t) = \frac{(vt)^x}{x!} e^{-vt} \qquad x = 0,1,2...$$

where

t : time period

v : mean occurrence rate of events at a location

x: occurrences

e = 2.71828

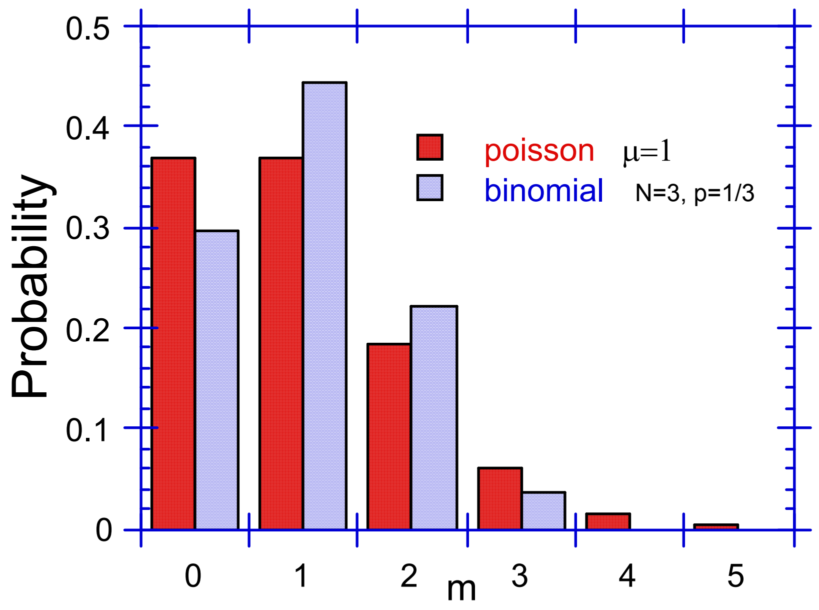
Siméon Denis Poisson
(1781-1840)

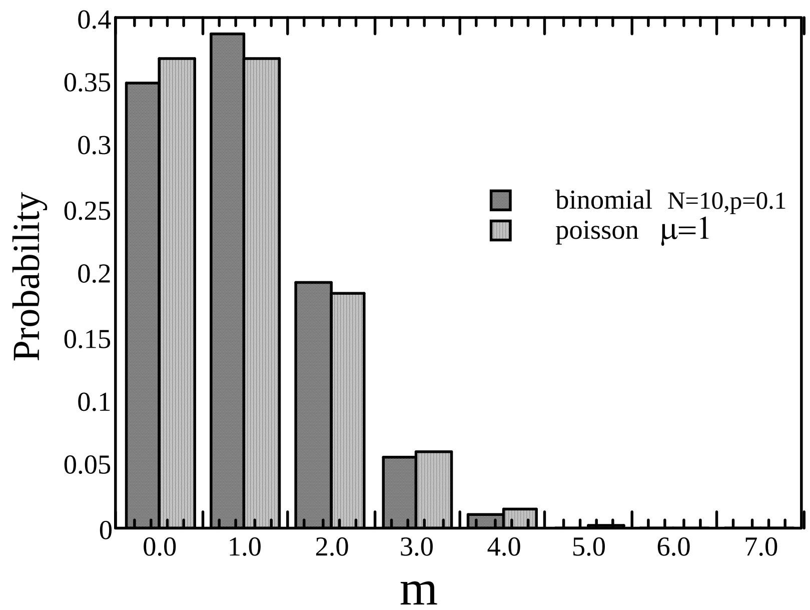
Difference between Binomial and Poisson distributions

- Poisson distribution can be derived by taking the appropriate limits of the Binomial distribution

For large *n* and fixed μ: Binomial $\Rightarrow$ Poisson

$$\text{Binomial Mean} = n \times p \qquad \text{Poisson Mean} = v$$





***m* special events (success)**

Example 8

From records of past 50 years, it is observed that tornadoes occur in a particular area an average of two times a year.

Calculate the probability of no tornadoes in the next year.

Solution

Risk of Earthquakes

Example 9

The safety of a building in an earthquake-prone area is under consideration.

The past 100 years of data indicate that there were four strong earthquakes in the area. Assume that damage event for different earthquakes are statistically independent.

(a) What is the probability that there will be no strong earthquakes in the area in 50 years, during the service life of the building?

(b) What is the probability that there will only two strong earthquakes in 50 years?

Solution

Risk of Power Supply Failure

Example 10

A nuclear plant receives its electric power from a utility grid outside of the plant. From past experience, it is know that loss of grid power occurs at a rate of once a year.

What is the probability that over a period of 3 years no power outage will occur?


Solution

Example 11

For a large construction project, the contractor estimates that the average rate of on-the-job accidents is three times per year. From past experience, the contractor also estimates that the cost incurred for each accident may be modeled as a lognormal random variable with a median of \$6,000 and COV of 20%. The cost of each accident can be assumed to be statistically independent.

OHS Risks

(a) What is the probability that there will be no accident in the first month of construction?

(b) What is the probability that an accident will incur a loss exceeding $4,000?

Solution

Exponential Distribution

- The probability distribution that describes the time between events in a Poisson process.
- It is the continuous analogue of the Geometric Distribution.
- Relationship between Exponential and Poisson distribution
  - Poisson is a discrete random variable that measures the number of occurrence of some given event over a specific interval (time, distance)
  - Exponential describes the length of the interval between occurrence.

Exponential Distribution

- PDF of the exponential distribution is

$$f_T(t) = \nu e^{-\nu t}$$

where

t : time period

v : the average rate of occurrences

e = 2.71828

Return Period (Poisson Distribution)

$$T = \int_0^\infty t f_T(t) dt = \int_0^\infty t \nu e^{-\nu t} dt$$

Exponential Distribution

Example 12

Strong earthquakes in an area are assumed to occur according to the Poisson distribution with the average rate of occurrences of 0.04 per year. Assuming the time between two consecutive occurrences of strong earthquakes can be modeled by an Exponential distribution, determine the return period of strong earthquakes.

Solution