Topic 9.4
Normal Distribution

TMA1201 Discrete Structures & Probability
Faculty of Computing & Informatics
Multimedia University

What you will learn in this lecture:

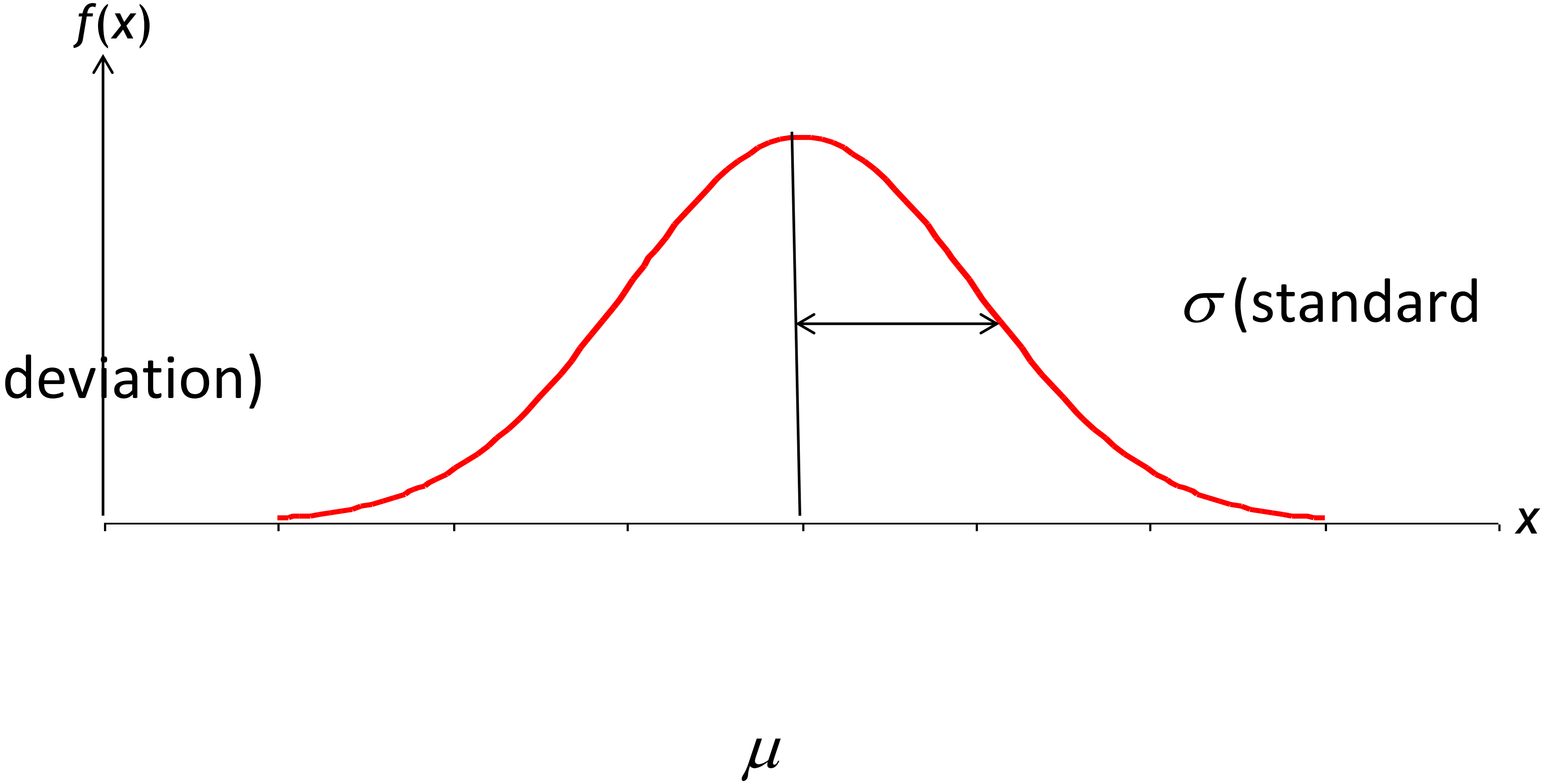
- The Normal distribution
  - Mean and standard deviation
- Standard normal distribution
- Normal approximation to binomial

Normal Distribution: The Bell-Shaped Curve

- The bell-shaped curve, also known as normal curve, is widely used to approximate many phenomena.
- The normal distribution allows us to calculate probabilities associated with observed sample results when we are dealing with CONTINUOUS outcomes.

Normal Distribution: The Bell-Shaped Curve

- The random variable with normal distribution is characterized by its mean μ and variance σ^2. $X \sim N(\mu, \sigma^2)$



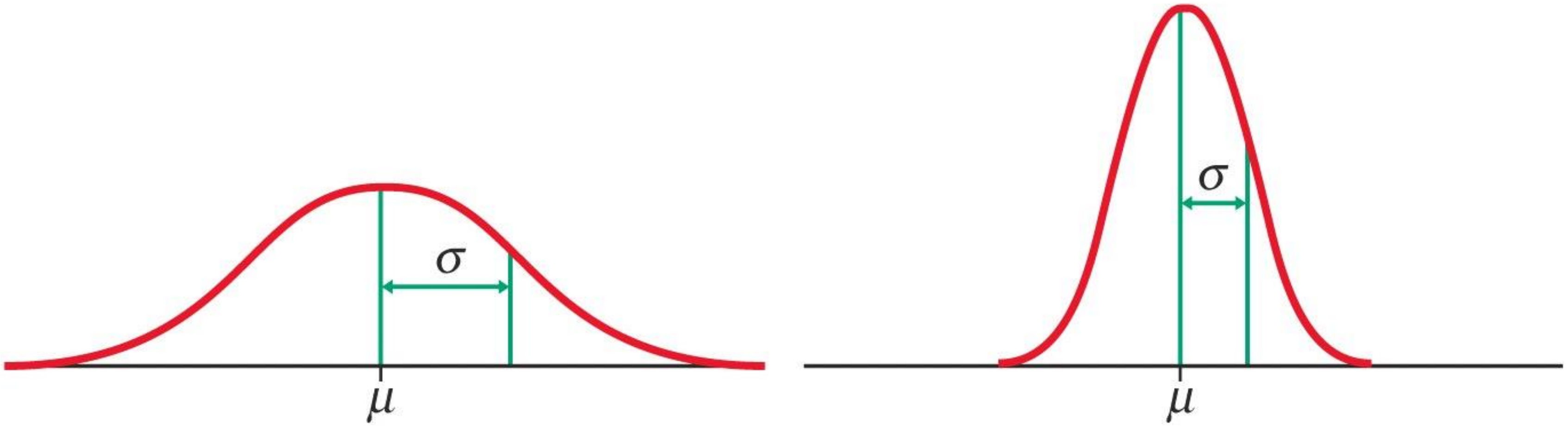
Normal Distribution

- The *pdf* for a Normal random variable *X* is given by

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

Normal Distribution

- Knowing the mean (μ) and standard deviation (σ) allows us to make various conclusions about Normal distributions.



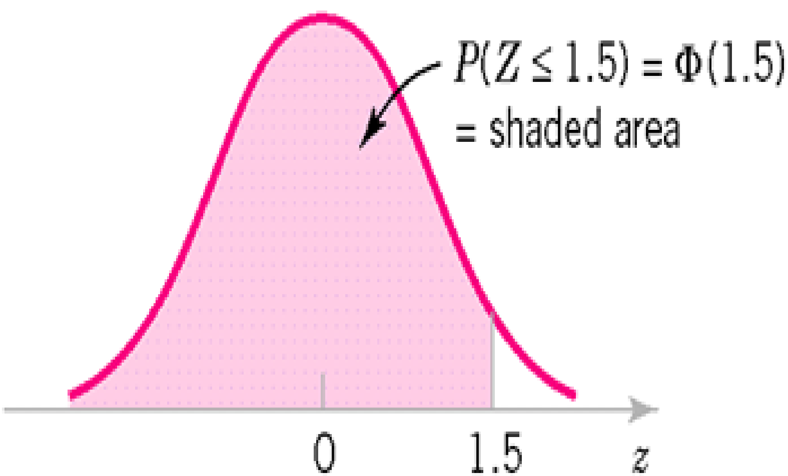
Standard Normal Distribution ($Z \sim N(0, 1)$)

- The standard normal distribution is a normal probability distribution that has a mean of 0 and a standard deviation of 1.
- All the observations of any normal variable *X* can be transformed to a new set of observations of standard normal variable *Z* with mean 0 and variance 1.

$$\boxed{z = \frac{\text{Observed value } - \text{Mean}}{\text{Standard deviation}} = \frac{x - \mu}{\sigma}}$$

Standard Normal Distribution

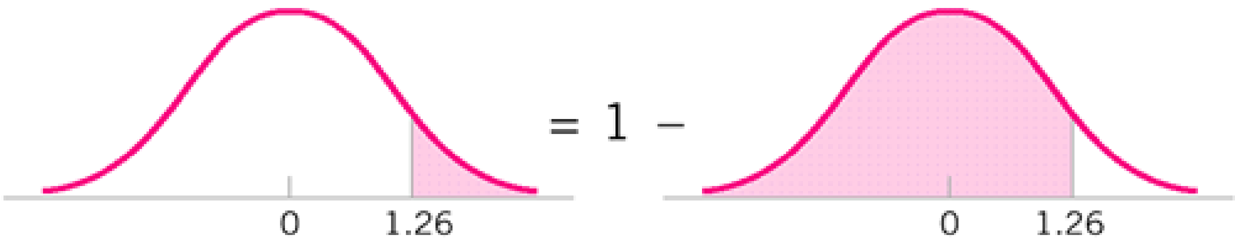
- We can find the areas under the standard normal curve by referring to Standard Normal Table which gives cumulative probabilities $\Phi(z)$.
- Standard Normal curve areas, $\Phi(z) = P(Z \leq z)$



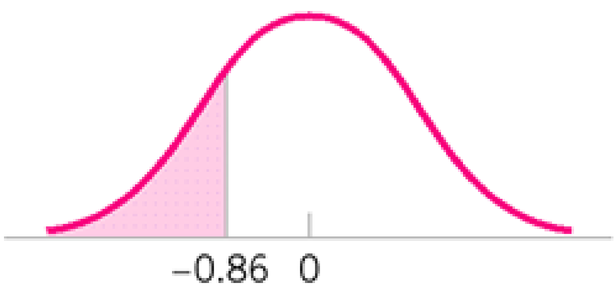
z	0.00	0.01	0.02	0.03
0	0.50000	0.50399	0.50398	0.51197
⋮		⋮		
1.5	0.93319	0.93448	0.93574	0.93699

Standard Normal Distribution

(1) $P(Z > 1.26) = 1 - P(Z \leq 1.26) = 1 - 0.89616 = 0.10384$

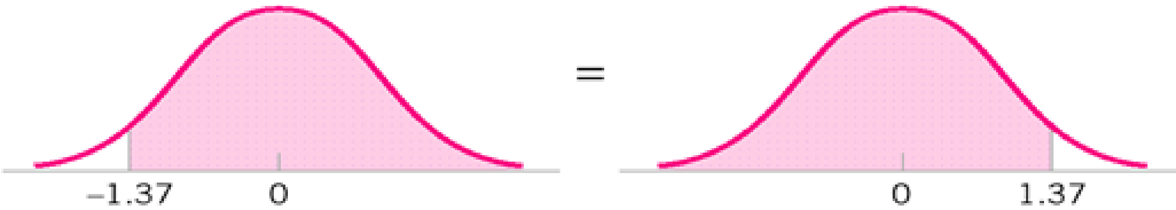


(2) $P(Z < -0.86) = 0.19490.$

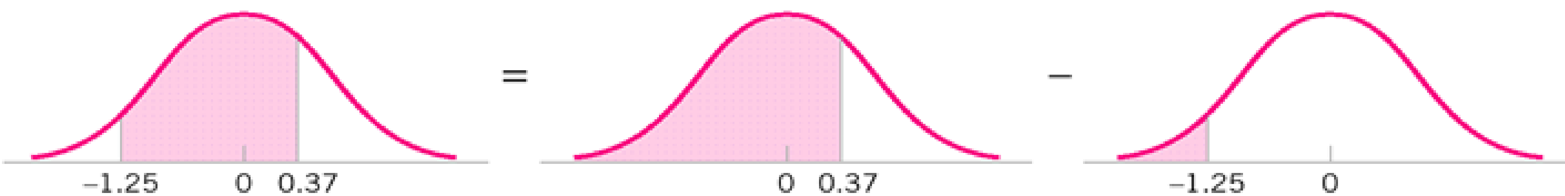


Standard Normal Distribution

(3) $P(Z > -1.37) = P(Z < 1.37) = 0.91465$



(4) $P(-1.25 \le Z \le 0.37)$



Converting Non-Standard Normal Distribution

- If X has a normal distribution with mean μ and standard deviation σ, then $z = \dfrac{x-\mu}{\sigma}$ has a standard normal distribution.

  Thus,

$$\boxed{P(a \le X \le b) = P\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right) = P\left(Z \le \frac{b-\mu}{\sigma}\right) - P\left(Z < \frac{a-\mu}{\sigma}\right)}$$

$$\boxed{P(X \le a) = P\left(Z \le \frac{a-\mu}{\sigma}\right)}$$

$$\boxed{P(X \ge b) = P\left(Z \ge \frac{b-\mu}{\sigma}\right) = 1 - P\left(Z < \frac{b-\mu}{\sigma}\right)}$$

Example 1

- The lengths of certain items follow a normal distribution with mean μ cm and standard deviation 6 cm. It is known that 4.78% of the items have a length greater than 82 cm. Find the value of the mean μ and $P(45 < X < 62)$.

Solutions:

- $P(X > 82) = 0.0478$,

$$P\left(Z > \frac{82-\mu}{6}\right) = 0.0478$$

$$P\left(Z < \frac{82-\mu}{6}\right) = 0.9522$$

$$\frac{82-\mu}{6} = 1.67;\ \mu = 71.98.$$

- $P(45 < X < 62)$
  $= P(-4.497 < Z < -1.663)$
  $= 0.0485$

Example 2

- The masses of articles produced in a particular workshop are normally distributed with mean μ and standard deviation σ.

  5% of the articles have a mass greater than 85g and 10% have a mass less than 25g. Find

  1) μ and σ.

  2) $P(X > 60)$

  3) a, given that $P(a < X < 74.65) = 0.75$

Example 2 (Cont.)

Solutions:

1) Let X denotes the mass of article produced by a particle workshop.

$P(X > 85) = 0.05$ and $P(X < 25) = 0.1$

or $P\left(Z \le \frac{85-\mu}{\sigma}\right) = 0.95$ and $P\left(Z < \frac{25-\mu}{\sigma}\right) = 0.1$

From table we get

$$\frac{85-\mu}{\sigma} = 1.645 \text{ and } \frac{25-\mu}{\sigma} = -1.28.$$

Solve for μ and σ we get μ = 51.26 and σ = 20.51.

2) P(X> 60) =

=1- P(Z $\le$ 0.4261) = 1 – 0.6664 = 0.3336

Example 2 (Cont.)

Solutions:

3) $P(a < X < 74.65) = 0.75$,

$$P(X < 74.65) - P(X \le a) = 0.75$$

$$P\left(Z \le \frac{a - 51.26}{20.51}\right) = P\left(Z < \frac{74.65 - 51.26}{20.51}\right) - 0.75$$

$$P\left(Z \le \frac{a - 51.26}{20.51}\right) = P(Z < 1.1404) - 0.75$$

$$P\left(Z < \frac{a - 51.26}{20.51}\right) = 0.8729 - 0.75 = 0.1229$$

$$\frac{a - 51.26}{20.51} = -1.16;$$

$$\therefore a = 27.47$$

Normal Approximation to the Binomial

- The binomial distribution shape is close to symmetrical for large n and p close to 0.5.

- If a large sample is selected from a population of binary values, the probabilities of the observed outcomes can be approximated using the normal distribution $N(\mu, \sigma^2)$ where $\mu = np$, and $\sigma^2 = np(1 - p)$.

Approximating the Binomial Distribution

Let X be a binomial rv based on n trials with success probability p. Then if the binomial probability histogram is not too skewed, X has approximately a normal distribution with $\mu = np$ and $\sigma = \sqrt{npq}$. In particular, for $x =$ a possible value of X,

$$P(X \le x) = B(x, n, p) \approx \begin{pmatrix} \text{area under the normal curve} \\ \text{to the left of } x + .5 \end{pmatrix}$$
$$= \Phi\left(\frac{x + .5 - np}{\sqrt{npq}}\right)$$

In practice, the approximation is adequate provided that both $np \ge 10$ and $nq \ge 10$, since there is then enough symmetry in the underlying binomial distribution.

Example 3

A variety of studies suggest that 11% of the world population is left-handed. Consider the situation where the sample size is two.

- $n = 2$, $p = 0.11$

  where $np = 2 \cdot (0.11) = 0.22 \leq 10$

- As such it is inappropriate to use the normal approximation to the probability of the binomial random variable.

Example 4

A variety of studies suggest that 11% of the world population is left-handed. Consider the situation where the sample size is increased to 100 instead of 2.

- *Given n* = 100, *p* = 0.11

  We have $np = 100 \cdot (0.11) = 11 \geq 10$ and
  $nq = 100 \cdot (0.89) = 89 \geq 10$

- We can thus use normal approximation to the binomial distribution.

Example 5

Assuming that the proportion of left-handed people in a population is 11%, find the probability that 70 or more from a sample of 500 people selected from this population are left-handed.

Solution:

- Let *X* be no of left-handed among the 500 randomly selected people. We thus have $X \sim Bin(n = 500, p = 0.11)$
- As *np* = 55 >10 and n(1-p) = 445 > 10, we can safely approximate the probability of *X* by normal distribution:
- $P(X \leq x) \approx \Phi\left(\frac{x+0.5-np}{\sqrt{np(1-p)}}\right) = \Phi\left(\frac{x+0.5-55}{\sqrt{48.95}}\right)$
- $P(X \geq 70) = 1 - P(X < 70) = 1 - P(X \leq 69) \approx 1 - \Phi\left(\frac{69+0.5-55}{6.996}\right)$
  $= 1 - \Phi(2.07) = 1 - 0.9808 = 0.0192$

Example 6

Suppose that 25% of all students at a large public university receive financial aid. Let X be the number of students in a random sample of size 50 who receive financial aid, so that $p = .25$. Then $\mu = 12.5$ and $\sigma = 3.06$. Since $np = 50(.25) = 12.5 \geq 10$ and $nq = 37.5 \geq 10$, the approximation can safely be applied. The probability that at most 10 students receive aid is

$$\begin{aligned} P(X \leq 10) = B(10; 50, .25) &\approx \Phi\left(\frac{10 + .5 - 12.5}{3.06}\right) \\ &= \Phi(-.65) = .2578 \end{aligned}$$

Similarly, the probability that between 5 and 15 (inclusive) of the selected students receive aid is

$$\begin{aligned} P(5 \leq X \leq 15) &= B(15; 50, .25) - B(4; 50, .25) \\ &\approx \Phi\left(\frac{15.5 - 12.5}{3.06}\right) - \Phi\left(\frac{4.5 - 12.5}{3.06}\right) = .8320 \end{aligned}$$

Summary

Materials discussed in this lecture

- The Normal distribution
- The Standard Normal distribution
- Normal approximation to Binomial

Exercise 1

X is a normally distributed variable with $\mu = 10$ and $\sigma = 2$.

(a) Find

(i) $P(X \leq 20)$

(ii) $P(X > 12)$

(iii) $P(X \leq 7.6)$

(iv) $P(7 \leq X \leq 13)$

(b) Find x when

(i) $z = 1.4$

(ii) $z = -0.5$

Exercise 2

For a certain type of computers, the length of time between charges of the battery is normally distributed with a mean of 50 hours and a standard deviation of 15 hours. James owns one of these computers and wants to know the probability that the length of time will be between 50 and 70 hours.

Exercise 3

The length of life of an instrument produced by a machine has a normal distribution with a mean of 12 months and standard deviation of 2 months. Find the probability that an instrument produced by this machine will last

(a) less than 7 months

(b) between 7 and 12 months

Exercise 4

50 percent of 12th graders attend school in a particular urban school district. A sample of 100 12th grade children are selected.

(a) Find p, q and n.

(b) Determine if you can approximate normal distribution to the binomial distribution.

(c) Find the mean, μ and standard deviation, σ.

(d) Find the probability that at least 80 are actually enrolled in school.