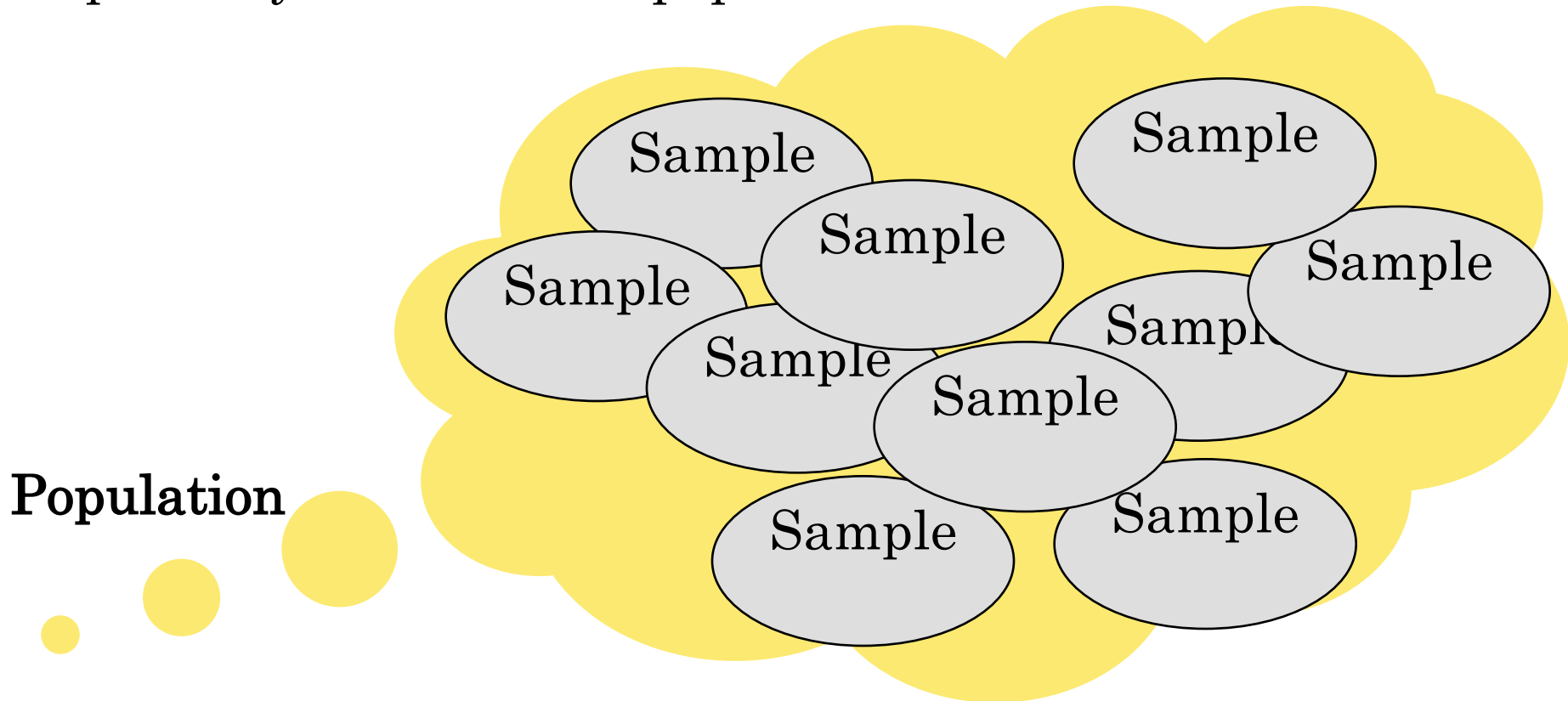# Sampling Distributions and the Central Limit Theorem
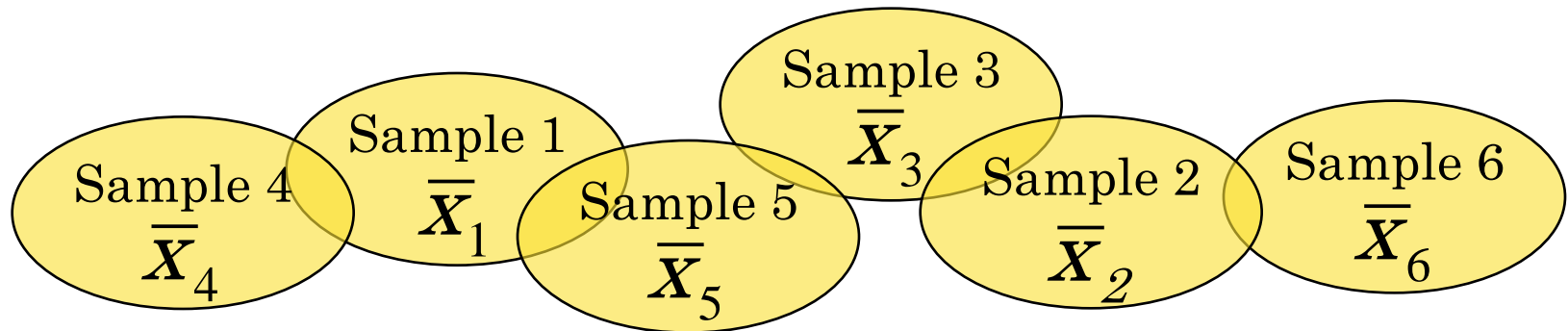
Sampling Distributions

A **sampling distribution** is the probability distribution of a sample statistic that is formed when samples of size n are repeatedly taken from a population.

Population

Sample

Sample

Sample

Sample

Sample

Sample

Sample

Sample

Sample

Sample

Sampling Distributions

If the sample statistic is the sample mean, then the distribution is the **sampling distribution of sample means**.

The sampling distribution consists of the values of the sample means, $\overline{X}_1, \overline{X}_2, \overline{X}_3, \overline{X}_4, \overline{X}_5, \overline{X}_6$.

Properties of Sampling Distributions

Properties of Sampling Distributions of Sample Means

1. The mean of the sample means, $\mu_{\bar{x}}$, is equal to the population mean.

$$\mu_{\bar{x}} = \mu$$

2. The standard deviation of the sample means, $\sigma_{\bar{x}}$, is equal to the population standard deviation, σ, divided by the square root of n.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

The standard deviation of the sampling distribution of the sample means is called the **standard error of the mean**.

Sampling Distribution of Sample Means

Example:
The population values {5, 10, 15, 20} are written on slips of paper and put in a hat. Two slips are randomly selected, with replacement.

    a.   Find the mean, standard deviation, and variance of the population.

<u>Population</u>
    5
    10
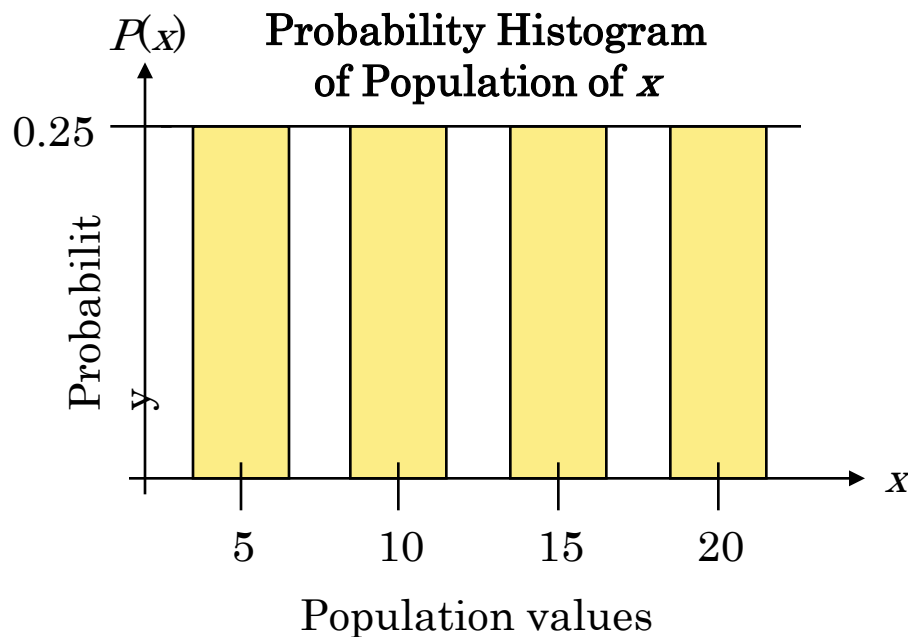   15
    20

$\mu = 12.5$

$\sigma = 5.59$

$\sigma^2 = 31.25$

Sampling Distribution of Sample Means

Example continued:

The population values {5, 10, 15, 20} are written on slips of paper and put in a hat. Two slips are randomly selected, with replacement.

    b.   Graph the probability histogram for the population values.

Probability Histogram of Population of *x*

$P(x)$

Probability

0.25

5   10   15   20

Population values

x

This uniform distribution shows that all values have the same probability of being selected.

Sampling Distribution of Sample Means

Example continued:

The population values {5, 10, 15, 20} are written on slips of paper and put in a hat. Two slips are randomly selected, with replacement.

    c.   List all the possible samples of size $n = 2$ and calculate the mean of each.

Sample	Sample mean, $\bar{x}$
5, 5	5
5, 10	7.5
5, 15	10
5, 20	12.5
10, 5	7.5
10, 10	10
10, 15	12.5
10, 20	15

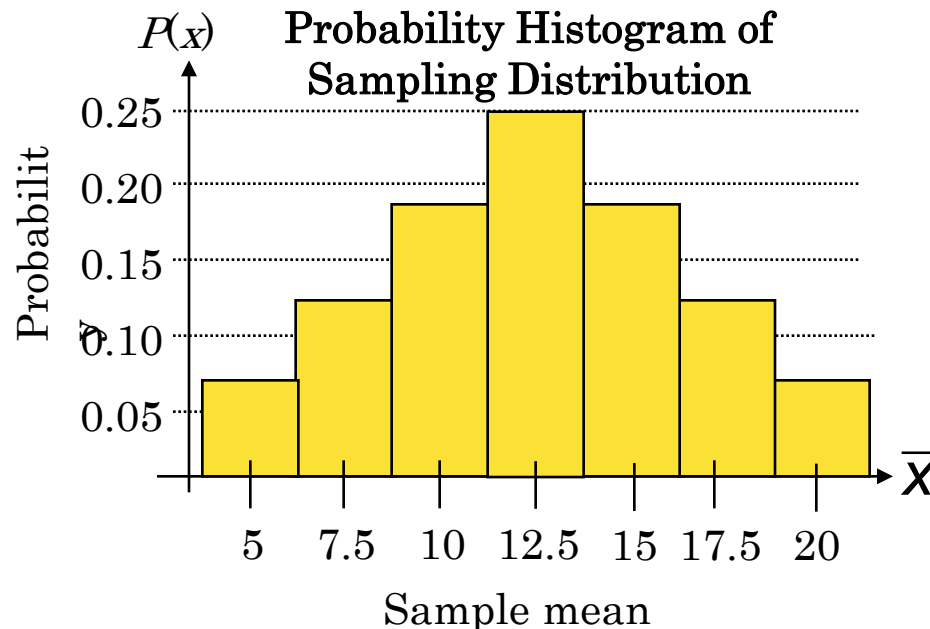| Sample	Sample mean, $\bar{x}$
15, 5	10
15, 10	12.5
15, 15	15
15, 20	17.5
20, 5	12.5
20, 10	15
20, 15	17.5
20, 20	20

These means form the sampling distribution of the sample means.

Continued.

Sampling Distribution of Sample Means

Example continued:
The population values {5, 10, 15, 20} are written on slips of paper and put in a hat. Two slips are randomly selected, with replacement.

    d.   Create the probability distribution of the sample means.

$\bar{x}$	f	Probability
5	1	0.0625
7.5	2	0.1250
10	3	0.1875
12.5	4	0.2500
15	3	0.1875
17.5	2	0.1250
20	1	0.0625

Probability Distribution of Sample Means

Sampling Distribution of Sample Means

Example continued:
The population values {5, 10, 15, 20} are written on slips of paper and put in a hat. Two slips are randomly selected, with replacement.

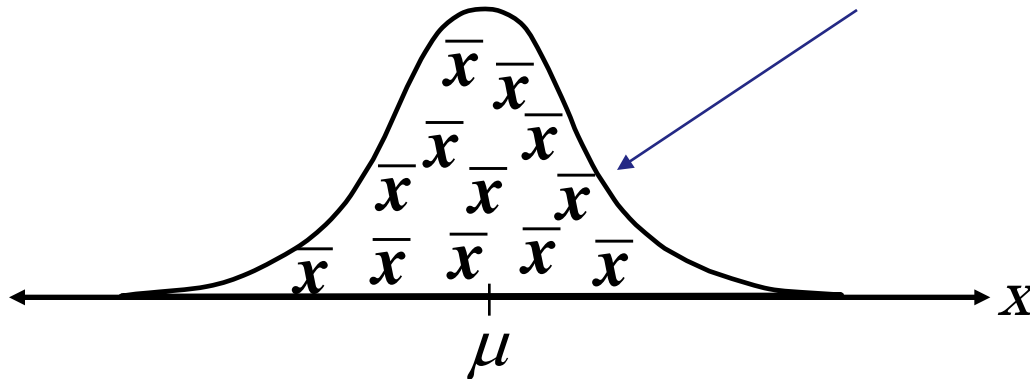   e.   Graph the probability histogram for the sampling distribution.

$P(x)$ **Probability Histogram of Sampling Distribution**

Probability

0.25
0.20
0.15
0.10
0.05

5   7.5   10 12.5 15 17.5 20

Sample mean

$\overline{X}$

The shape of the graph is symmetric and bell shaped. It approximates a normal distribution.

The Central Limit Theorem

If a sample of size $n \geq 30$ is taken from a population with *any type of distribution* that has a mean = μ and standard deviation = σ,
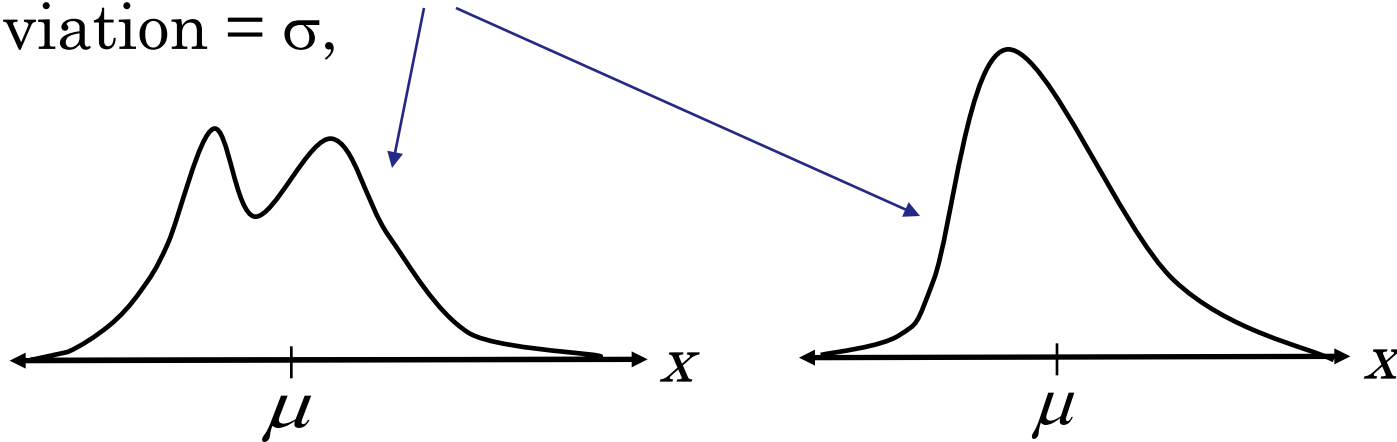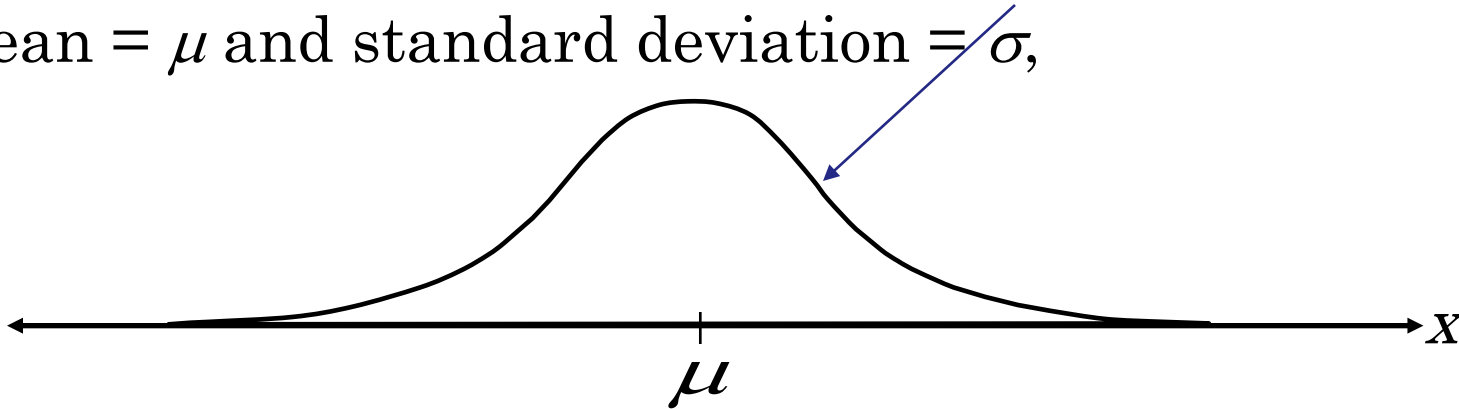
the *sample means* will have a **normal distribution**.

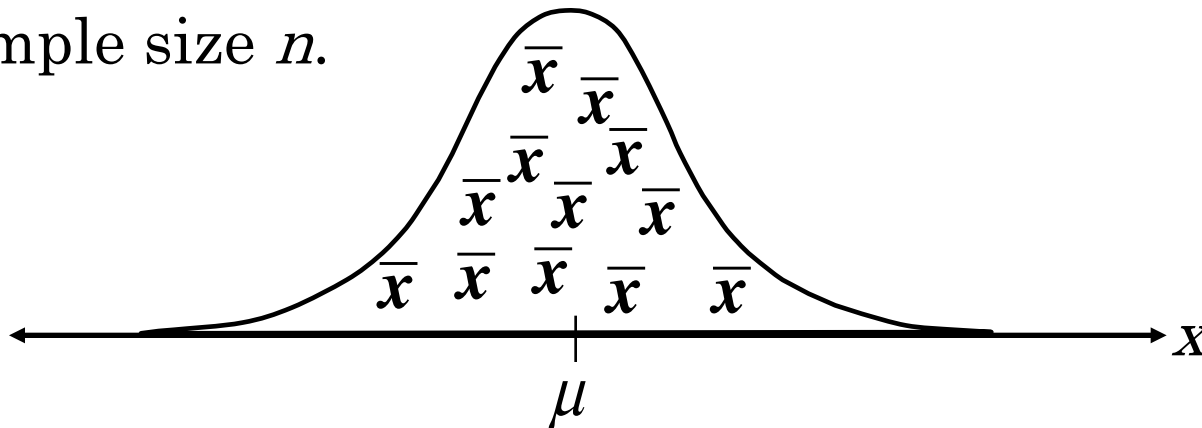# The Central Limit Theorem

If the population itself is *normally distributed*, with mean $= \mu$ and standard deviation $= \sigma$,

the *sample means* will have a **normal distribution** for *any* sample size n.

The Central Limit Theorem

In either case, the sampling distribution of sample means has a mean equal to the population mean.

$$\mu_{\bar{x}} = \mu$$

Mean of the sample means

The sampling distribution of sample means has a standard deviation equal to the population standard deviation divided by the square root of *n.*

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Standard deviation of the sample means

This is also called the **standard error of the mean**.

The Mean and Standard Error

Example:
The heights of fully grown magnolia bushes have a mean height of 8 feet and a standard deviation of 0.7 feet. 38 bushes are randomly selected from the population, and the mean of each sample is determined. Find the mean and standard error of the mean of the sampling distribution.

Mean

$$\mu_{\bar{x}} = \mu$$

$$= 8$$

Standard deviation (standard error)

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$
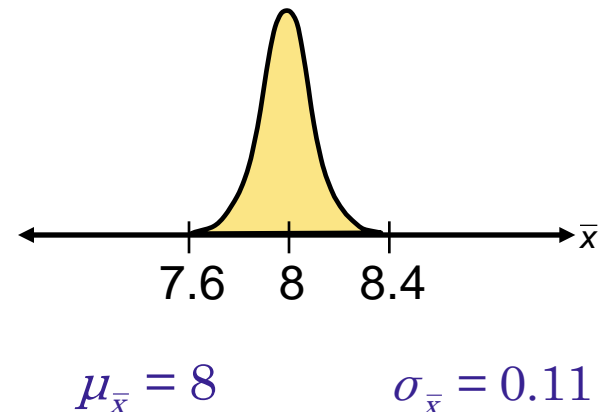
$$= \frac{0.7}{\sqrt{38}} = 0.11$$

Interpreting the Central Limit Theorem

Example continued:
The heights of fully grown magnolia bushes have a mean height of 8 feet and a standard deviation of 0.7 feet. 38 bushes are randomly selected from the population, and the mean of each sample is determined.

The mean of the sampling distribution is 8 feet ,and the standard error of the sampling distribution is 0.11 feet.

From the Central Limit Theorem, because the sample size is greater than 30, the sampling distribution can be approximated by the normal distribution.

7.6   8   8.4

$\mu_{\bar{x}} = 8$ $\sigma_{\bar{x}} = 0.11$

Finding Probabilities

Example:
The heights of fully grown magnolia bushes have a mean height of 8 feet and a standard deviation of 0.7 feet. 38 bushes are randomly selected from the population, and the mean of each sample is determined.

The mean of the sampling distribution is 8 feet, and the standard error of the sampling distribution is 0.11 feet.

Find the probability that the mean height of the 38 bushes is less than 7.8 feet.
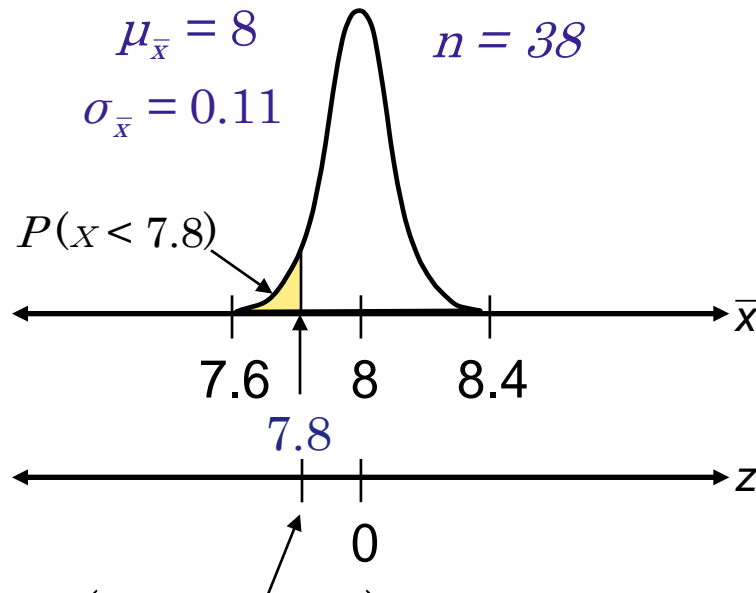
$\mu_{\bar{x}} = 8$ $n = 38$

$\sigma_{\bar{x}} = 0.11$

7.6 8 8.4

7.8

$\bar{x}$
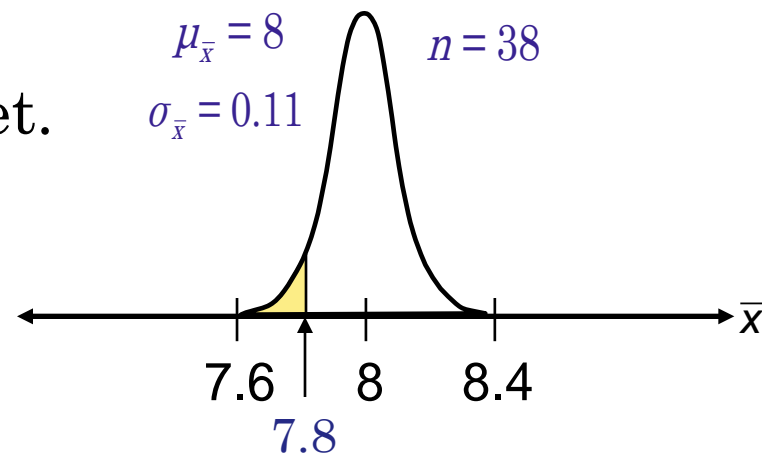
Continued.

Finding Probabilities

Example continued:
Find the probability that the mean height of the 38 bushes is less than 7.8 feet.

$$\mu_{\bar{x}} = 8$$
$$\sigma_{\bar{x}} = 0.11$$

$n = 38$

$P(\bar{x} < 7.8)$

7.6   8   8.4
7.8

0

$$Z = \frac{\bar{X} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

$$= \frac{7.8 - 8}{0.11}$$

$$= -1.82$$

$$P(\bar{x} < 7.8) = P(z < \underline{\quad ? \quad}) = 0.0344$$

The probability that the mean height of the 38 bushes is less than 7.8 feet is 0.0344.

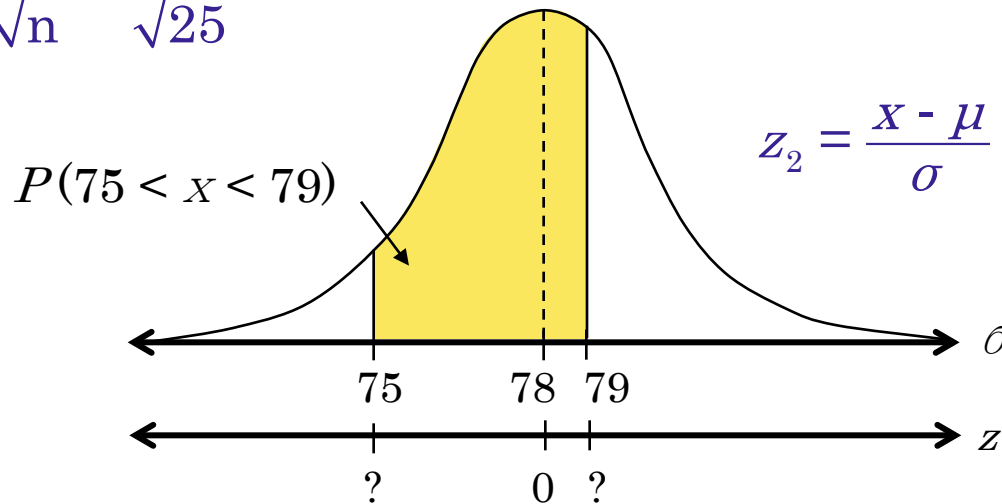# Probability and Normal Distributions

Example:

The average on a statistics test was 78 with a standard deviation of 8. If the test scores are normally distributed, find the probability that the mean score of 25 randomly selected students is between 75 and 79.

$$\mu_{\bar{x}} = 78$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{8}{\sqrt{25}} = 1.6$$

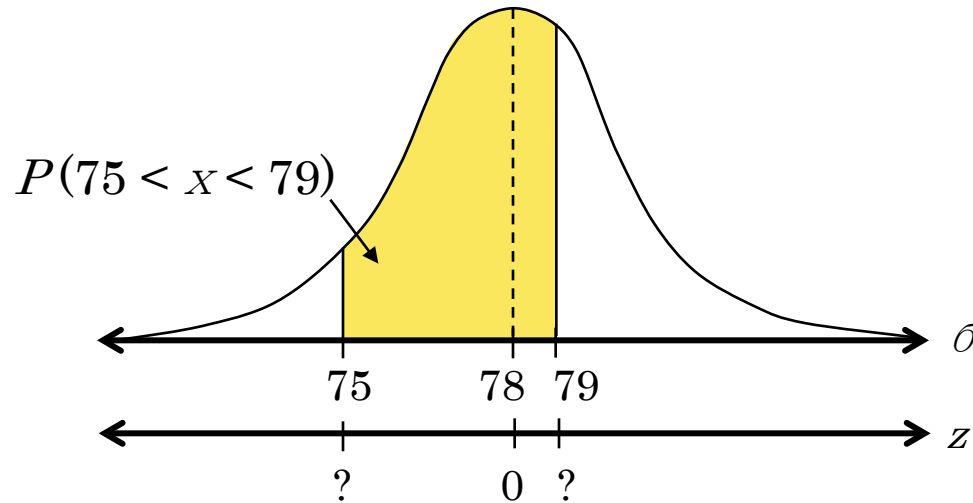$$z_1 = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{75 - 78}{1.6} = -1.88$$

$$z_2 = \frac{x - \mu}{\sigma} = \frac{79 - 78}{1.6} = 0.63$$

$P(75 < x < 79)$

75   78 79 σ

?   0   ? z

Continued.

Probability and Normal Distributions

Example continued:

$$P(75 < x < 79) = P(-1.88 < z < 0.63) = P(z < 0.63) - P(z < -1.88)$$

$$= 0.7357 - 0.0301 = 0.7056$$

Approximately 70.56% of the 25 students will have a mean score between 75 and 79.
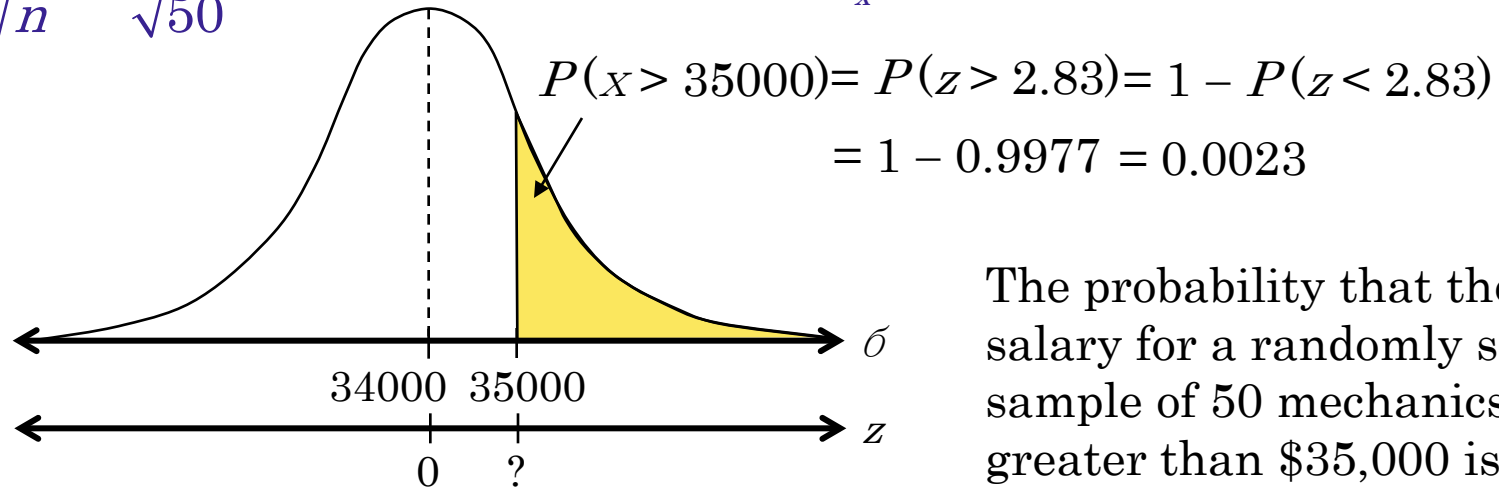
Probabilities of x and $\bar{x}$

Example:

The population mean salary for auto mechanics is $\mu = \$34,000$ with a standard deviation of $\sigma = \$2,500$. Find the probability that the <u>mean</u> salary for a randomly selected <u>sample of 50</u> mechanics is greater than $35,000.

$$\mu_{\bar{x}} = 34000$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{2500}{\sqrt{50}} = 353.55$$

$$Z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{35000 - 34000}{353.55} = 2.83$$

$$P(x > 35000) = P(z > 2.83) = 1 - P(z < 2.83)$$

$$= 1 - 0.9977 = 0.0023$$

The probability that the mean salary for a randomly selected sample of 50 mechanics is greater than $35,000 is 0.0023.
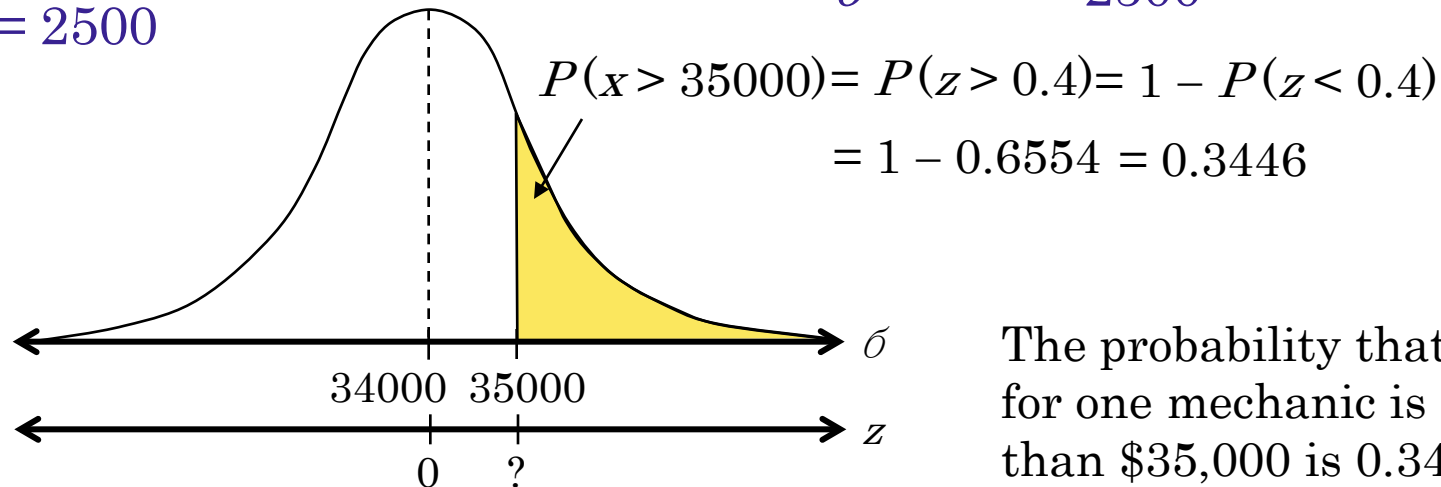
Probabilities of x and x̄

Example:
The population mean salary for auto mechanics is $\mu = \$34,000$ with a standard deviation of $\sigma = \$2,500$. Find the probability that the salary for <u>one</u> randomly selected mechanic is greater than $35,000.

(Notice that the Central Limit Theorem does not apply.)

$\mu = 34000$

$\sigma = 2500$

$z = \dfrac{x - \mu}{\sigma} = \dfrac{35000 - 34000}{2500} = 0.4$

$P(x > 35000) = P(z > 0.4) = 1 - P(z < 0.4)$

$= 1 - 0.6554 = 0.3446$

34000 35000

0   ?

The probability that the salary for one mechanic is greater than $35,000 is 0.3446.

Probabilities of x and $\bar{x}$

Example:
The probability that the salary for <u>one</u> randomly selected mechanic is greater than $35,000 is 0.3446. In a group of 50 mechanics, approximately how many would have a salary greater than $35,000?

$$P(x > 35000) = 0.3446$$

This also means that 34.46% of mechanics have a salary greater than $35,000.

$$34.46\% \text{ of } 50 = 0.3446 \times 50 = 17.23$$

You would expect about 17 mechanics out of the group of 50 to have a salary greater than $35,000.