Chapter 5

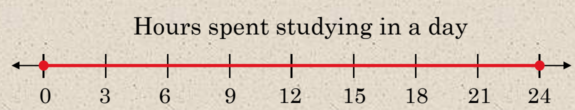
Normal Probability Distributions

§ 5.1

Introduction to Normal Distributions and the Standard Distribution

Properties of Normal Distributions

A **continuous random variable** has an infinite number of possible values that can be represented by an interval on the number line.

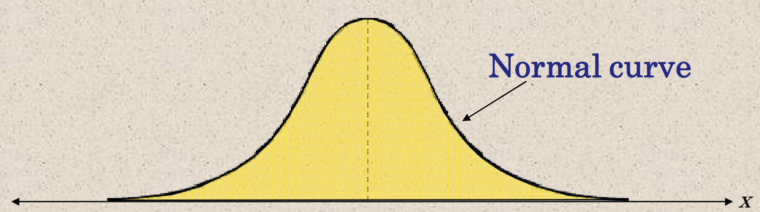


The time spent studying can be any number between 0 and 24.

The probability distribution of a continuous random variable is called a **continuous probability distribution.**

Properties of Normal Distributions

The most important probability distribution in statistics is the **normal distribution**.



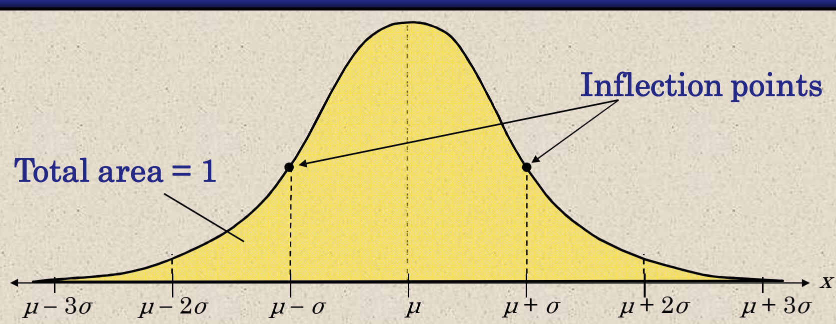
A normal distribution is a continuous probability distribution for a random variable, x. The graph of a normal distribution is called the **normal curve**.

Properties of Normal Distributions

Properties of a Normal Distribution

1. The mean, median, and mode are equal.
2. The normal curve is bell-shaped and symmetric about the mean.
3. The total area under the curve is equal to one.
4. The normal curve approaches, but never touches the *x*-axis as it extends farther and farther away from the mean.
5. Between $\mu - \sigma$ and $\mu + \sigma$ (in the center of the curve), the graph curves downward. The graph curves upward to the left of $\mu - \sigma$ and to the right of $\mu + \sigma$. The points at which the curve changes from curving upward to curving downward are called the *inflection points*.

Properties of Normal Distributions

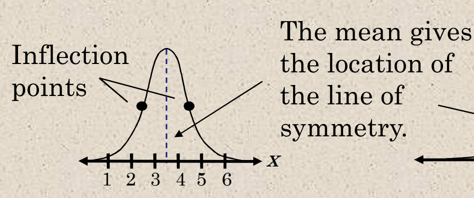


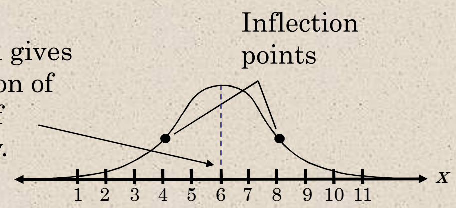
If x is a continuous random variable having a normal distribution with mean μ and standard deviation σ, you can graph a normal curve with the equation

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}. \qquad e = 2.178 \quad \pi = 3.14$$

Means and Standard Deviations

A normal distribution can have any mean and any positive standard deviation.





Mean: $\mu = 3.5$

Standard deviation: $\sigma \approx 1.3$

Mean: $\mu = 6$

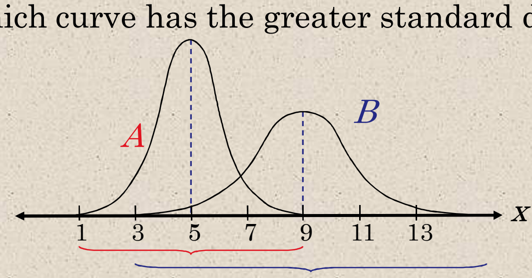
Standard deviation: $\sigma \approx 1.9$

The standard deviation describes the spread of the data.

Means and Standard Deviations

Example:

1. Which curve has the greater mean?
2. Which curve has the greater standard deviation?

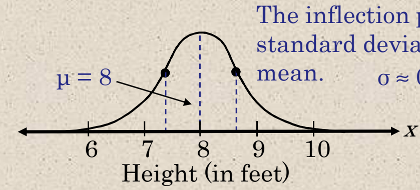


The line of symmetry of curve A occurs at $x = 5$. The line of symmetry of curve B occurs at $x = 9$. Curve B has the greater mean.

Curve B is more spread out than curve A, so curve B has the greater standard deviation.

Interpreting Graphs

Example:
The heights of fully grown magnolia bushes are normally distributed. The curve represents the distribution. What is the mean height of a fully grown magnolia bush? Estimate the standard deviation.



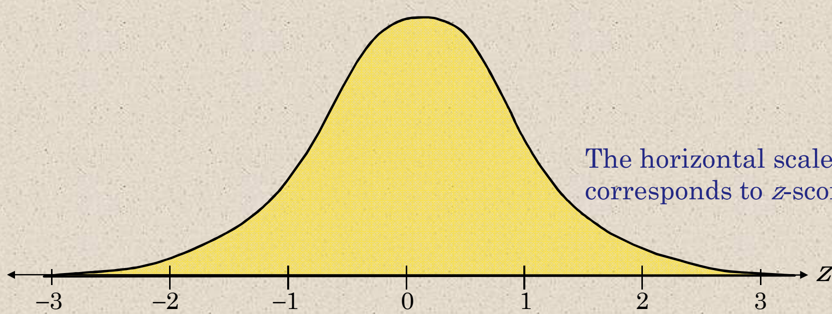
The inflection points are one standard deviation away from the mean.

$\sigma \approx 0.7$

The heights of the magnolia bushes are normally distributed with a mean height of about 8 feet and a standard deviation of about 0.7 feet.

The Standard Normal Distribution

The **standard normal distribution** is a normal distribution with a mean of 0 and a standard deviation of 1.

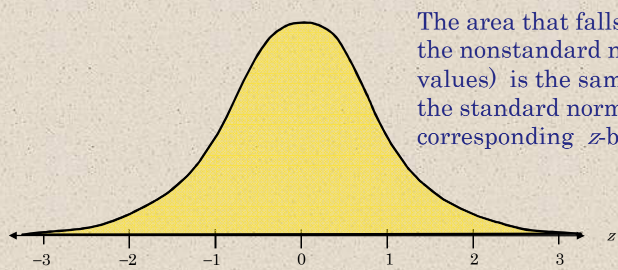


The horizontal scale corresponds to *z*-scores.

Any value can be transformed into a *z*-score by using the formula $z = \dfrac{\text{Value - Mean}}{\text{Standard deviation}} = \dfrac{x - \mu}{\sigma}$.

The Standard Normal Distribution

If each data value of a normally distributed random variable *x* is transformed into a *z*-score, the result will be the standard normal distribution.



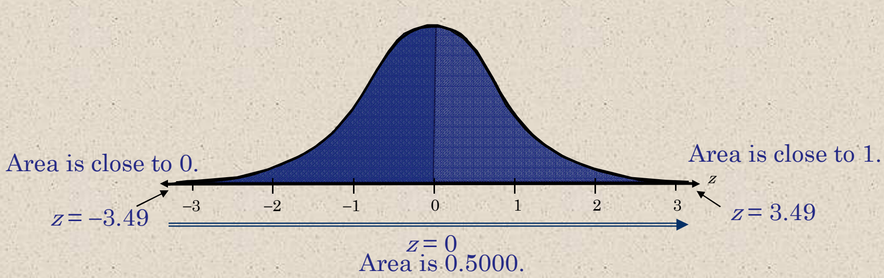
The area that falls in the interval under the nonstandard normal curve (the *x*-values) is the same as the area under the standard normal curve (within the corresponding *z*-boundaries).

After the formula is used to transform an *x*-value into a *z*-score, the Standard Normal Table in Appendix B is used to find the cumulative area under the curve.

The Standard Normal Table

Properties of the Standard Normal Distribution

1. The cumulative area is close to 0 for *z*-scores close to $z = -3.49$.
2. The cumulative area increases as the *z*-scores increase.
3. The cumulative area for $z = 0$ is 0.5000.
4. The cumulative area is close to 1 for *z*-scores close to $z = 3.49$



The Standard Normal Table

Example:
Find the cumulative area that corresponds to a *z*-score of 2.71.

Appendix B: Standard Normal Table

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981

Find the area by finding 2.7 in the left hand column, and then moving across the row to the column under 0.01.

The area to the left of $z = 2.71$ is 0.9966.

The Standard Normal Table

Example:
Find the cumulative area that corresponds to a *z*-score of –0.25.

Appendix B: Standard Normal Table

z	.09	.08	.07	.06	.05	.04	.03	.02	.01	.00
–3.4	.0002	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003
–3.3	.0003	.0004	.0004	.0004	.0004	.0004	.0004	.0005	.0005	.0005
–0.3	.3483	.3520	.3557	.3594	.3632	.3669	.3707	.3745	.3783	.3821
–0.2	.3859	.3897	.3936	.3974	.4013	.4052	.4090	.4129	.4168	.4207
–0.1	.4247	.4286	.4325	.4364	.4404	.4443	.4483	.4522	.4562	.4602
–0.0	.4641	.4681	.4724	.4761	.4801	.4840	.4880	.4920	.4960	.5000

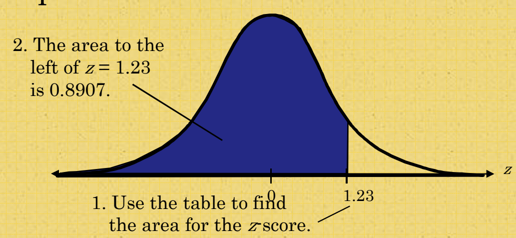
Find the area by finding –0.2 in the left hand column, and then moving across the row to the column under 0.05.

The area to the left of $z = -0.25$ is 0.4013

Guidelines for Finding Areas

Finding Areas Under the Standard Normal Curve

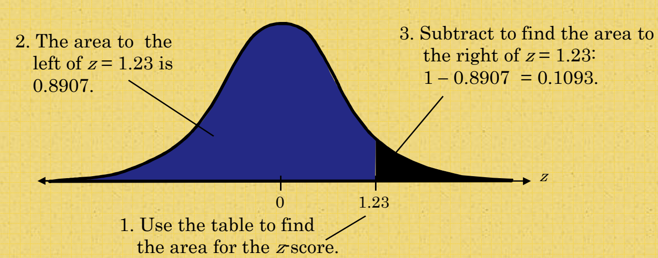
1. Sketch the standard normal curve and shade the appropriate area under the curve.
2. Find the area by following the directions for each case shown.
   a. To find the area to the *left* of z, find the area that corresponds to z in the Standard Normal Table.



Guidelines for Finding Areas

Finding Areas Under the Standard Normal Curve

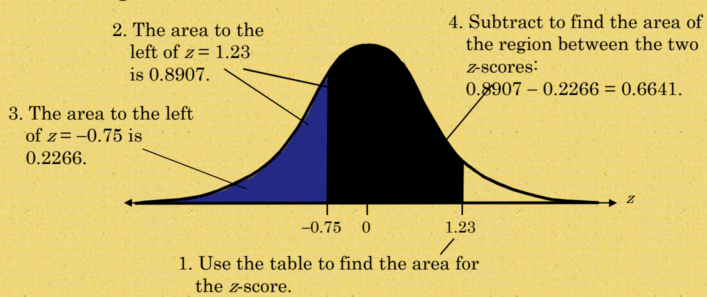
b. To find the area to the *right* of z, use the Standard Normal Table to find the area that corresponds to z. Then subtract the area from 1.



Guidelines for Finding Areas

Finding Areas Under the Standard Normal Curve

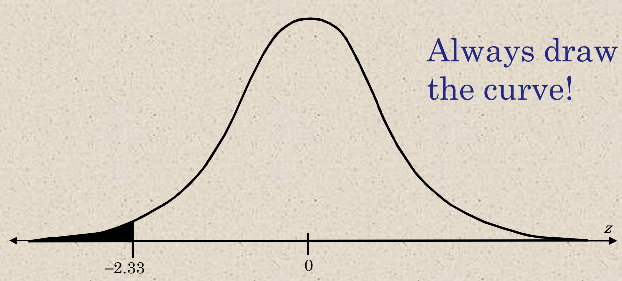
c. To find the area *between* two *z*-scores, find the area corresponding to each *z*-score in the Standard Normal Table. Then subtract the smaller area from the larger area.



Guidelines for Finding Areas

Example:

Find the area under the standard normal curve to the left of $z = -2.33$.



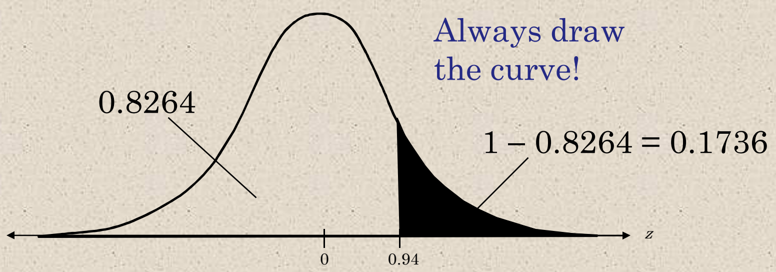
Always draw the curve!

From the Standard Normal Table, the area is equal to 0.0099.

Guidelines for Finding Areas

Example:
Find the area under the standard normal curve to the right of $z = 0.94$.

Always draw the curve!

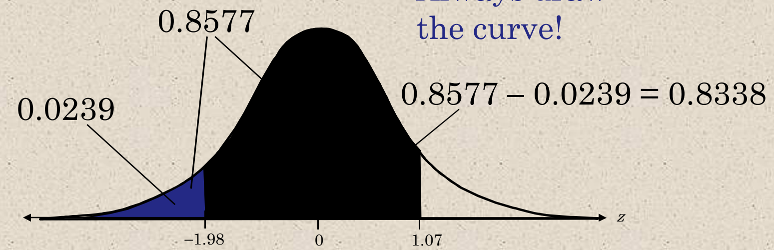


From the Standard Normal Table, the area is equal to 0.1736.

Guidelines for Finding Areas

Example:
Find the area under the standard normal curve between $z = -1.98$ and $z = 1.07$.

Always draw the curve!



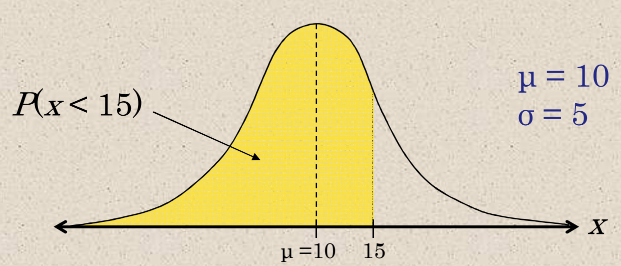
From the Standard Normal Table, the area is equal to 0.8338.

§ 5.2

Normal Distributions: Finding Probabilities

Probability and Normal Distributions

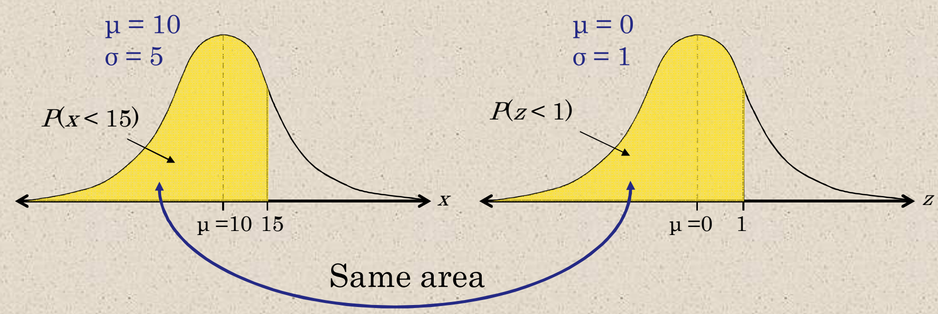
If a random variable, *x*, is normally distributed, you can find the probability that *x* will fall in a given interval by calculating the area under the normal curve for that interval.



Probability and Normal Distributions

Normal Distribution

Standard Normal Distribution

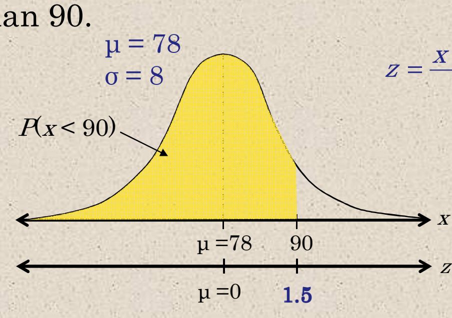


$P(x < 15) = P(z < 1) =$ Shaded area under the curve

$= 0.8413$

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 a student receives a test score less than 90.



$$z = \frac{x - \mu}{\sigma} = \frac{90 - 78}{8}$$
$$= 1.5$$

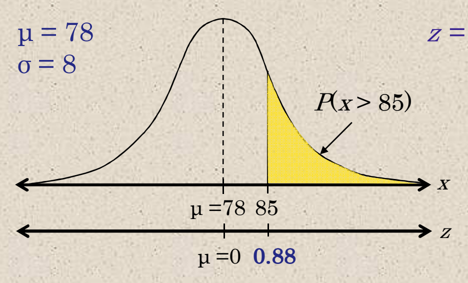
The probability that a student receives a test score less than 90 is 0.9332.

$P(x < 90) = P(z < 1.5) = 0.9332$

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 a student receives a test score greater than than 85.

$\mu = 78$
$\sigma = 8$



$$z = \frac{x - \mu}{\sigma} = \frac{85 - 78}{8}$$
$$= 0.875 \approx 0.88$$

The probability that a student receives a test score greater than 85 is 0.1894.

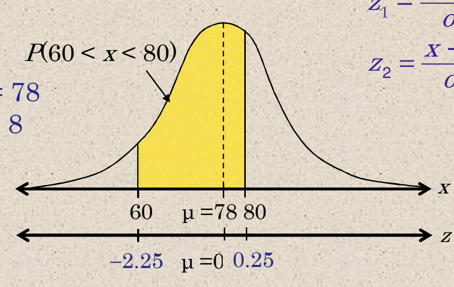
$$P(x > 85) = P(z > 0.88) = 1 - P(z < 0.88) = 1 - 0.8106 = 0.1894$$

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 a student receives a test score between 60 and 80.

$$z_1 = \frac{x - \mu}{\sigma} = \frac{60 - 78}{8} = -2.25$$
$$z_2 = \frac{x - \mu}{\sigma} = \frac{80 - 78}{8} = 0.25$$

$\mu = 78$
$\sigma = 8$



The probability that a student receives a test score between 60 and 80 is 0.5865.

$$P(60 < x < 80) = P(-2.25 < z < 0.25) = P(z < 0.25) - P(z < -2.25)$$
$$= 0.5987 - 0.0122 = 0.5865$$

§ 5.3

Normal Distributions: Finding Values

Finding z-Scores

Example:
Find the z-score that corresponds to a cumulative area of 0.9973.

Appendix B: Standard Normal Table

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981

Find the z-score by locating 0.9973 in the body of the Standard Normal Table. The values at the beginning of the corresponding row and at the top of the column give the z-score.

The z-score is 2.78.

Finding z-Scores

Example:
Find the z-score that corresponds to a cumulative area of 0.4170.

Appendix B: Standard Normal Table

z	.09	.08	.07	.06	.05	.04	.03	.02	.01	.00
–3.4	.0002	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003
–0.2	.0003	.0004	.0004	.0004	.0004	.0004	.0004	.0005	.0005	.0005

z	.09	.08	.07	.06	.05	.04	.03	.02	.01	.00
–0.3	.3483	.3520	.3557	.3594	.3632	.3669	.3707	.3745	.3783	.3821
–0.2	.3859	.3897	.3936	.3974	.4013	.4052	.4090	.4129	.4168	.4207
–0.1	.4247	.4286	.4325	.4364	.4404	.4443	.4483	.4522	.4562	.4602
–0.0	.4641	.4681	.4724	.4761	.4801	.4840	.4880	.4920	.4960	.5000

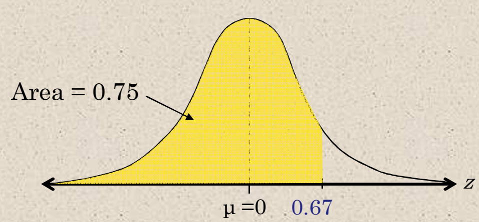
Use the closest area.

Find the z-score by locating 0.4170 in the body of the Standard Normal Table. Use the value closest to 0.4170.

The z-score is –0.21.

Finding a z-Score Given a Percentile

Example:
Find the z-score that corresponds to P_{75}.



The z-score that corresponds to P_{75} is the same z-score that corresponds to an area of 0.75.

The z-score is 0.67.

Transforming a z-Score to an x-Score

To transform a standard z-score to a data value, x, in a given population, use the formula

$$x = \mu + z\sigma.$$

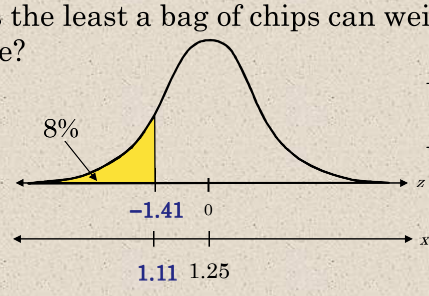
Example:
The monthly electric bills in a city are normally distributed with a mean of $120 and a standard deviation of $16. Find the *x*-value corresponding to a *z*-score of 1.60.

$$\begin{aligned} x &= \mu + z\sigma \\ &= 120 + 1.60(16) \\ &= 145.6 \end{aligned}$$

We can conclude that an electric bill of $145.60 is 1.6 standard deviations above the mean.

Finding a Specific Data Value

Example:
The weights of bags of chips for a vending machine are normally distributed with a mean of 1.25 ounces and a standard deviation of 0.1 ounce. Bags that have weights in the lower 8% are too light and will not work in the machine. What is the least a bag of chips can weigh and still work in the machine?



$P(z < ?) = 0.08$

$P(z < -1.41) = 0.08$

$$\begin{aligned} x &= \mu + z\sigma \\ &= 1.25 + (-1.41)0.1 \\ &= 1.11 \end{aligned}$$

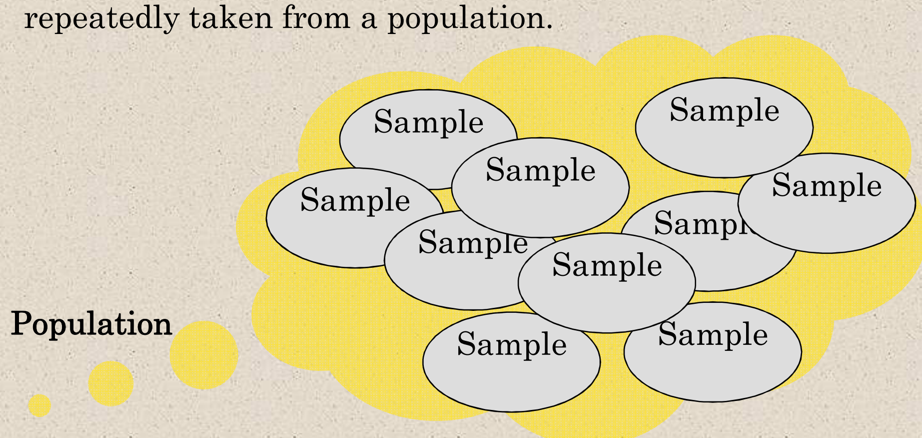
The least a bag can weigh and still work in the machine is 1.11 ounces.

§ 5.4

Sampling Distributions and the Central Limit Theorem

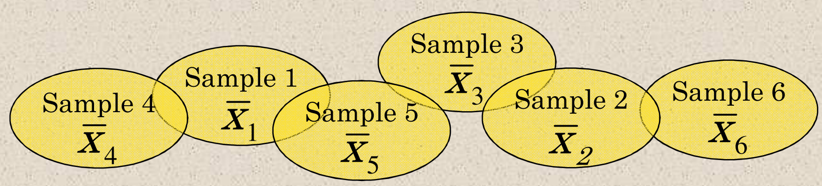
Sampling Distributions

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.



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.

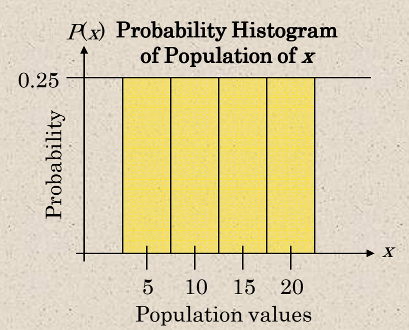
Population	
5	$\mu = 12.5$
10	$\sigma = 5.59$
15	$\sigma^2 = 31.25$
20	

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.

b. Graph the probability histogram for the population values.



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

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.

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.

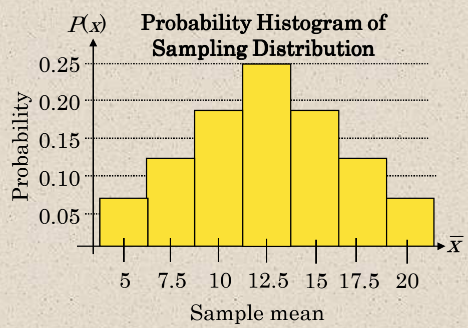
$\bar{x}$	f	Probability
5	1	0.0625
7.5	2	0.1250
10	3	0.1875
12.5	4	0.2500
15	3	0.1875
17.5	2	0.1250
20	1	0.0625

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.

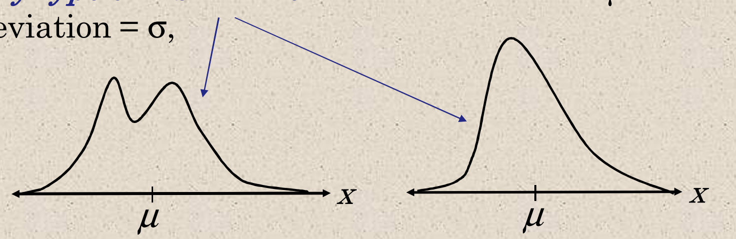
e. Graph the probability histogram for the sampling distribution.



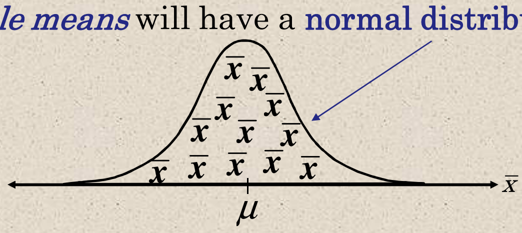
The shape of the graph is symmetric and bell shaped. It approximates a normal distribution.

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,

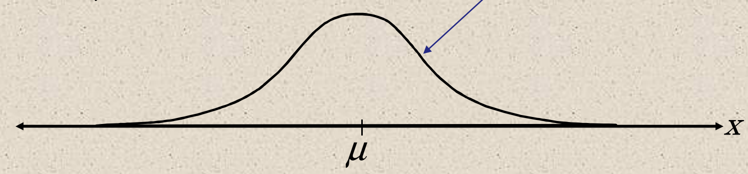


the ***sample means*** will have a **normal distribution**.

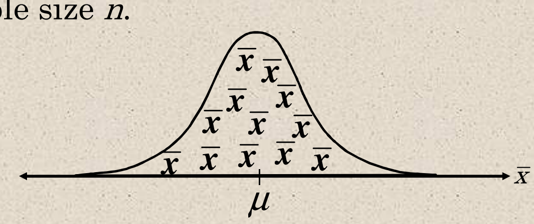


The Central Limit Theorem

If the population itself is ***normally distributed***, with mean = μ and standard deviation = σ,



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)

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

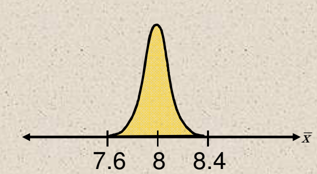
Continued.

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.



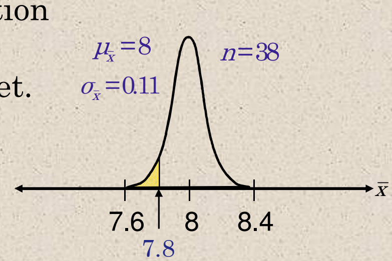
$\mu_{\bar{x}} = 8$ $\sigma_{\bar{x}} = 0.11$

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.

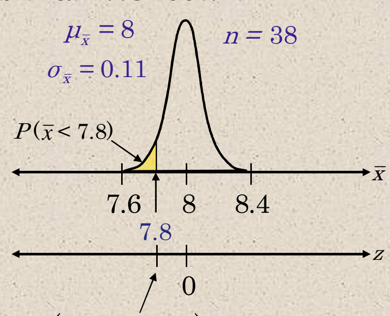
Find the probability that the mean height of the 38 bushes is less than 7.8 feet.



Continued.

Finding Probabilities

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



$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

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

$$= -1.82$$

$$P(\bar{x} < 7.8) = P(z < -1.82) = 0.0344$$

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

Probability and Normal Distributions

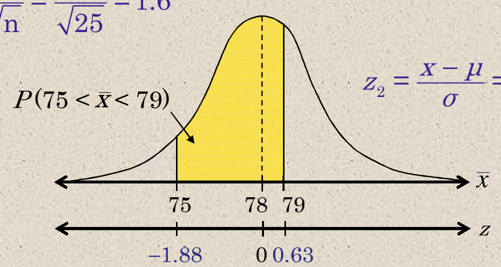
Example:
The average on a statistics test was 78 with a standard deviation of 8. If the test scores are normally distributed, find the probability that the mean score of 25 randomly selected students is between 75 and 79.

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

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

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

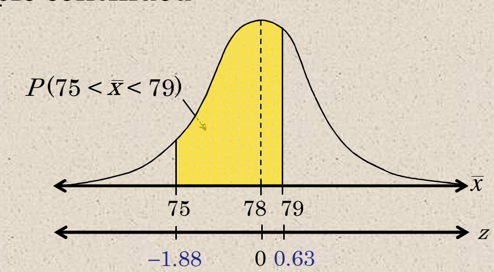
$$z_2 = \frac{x - \mu}{\sigma} = \frac{79 - 78}{1.6} = 0.63$$



Continued.

Probability and Normal Distributions

Example continued:



$$P(75 < \bar{x} < 79) = P(-1.88 < z < 0.63) = P(z < 0.63) - P(z < -1.88)$$

$$= 0.7357 - 0.0301 = 0.7056$$

Approximately 70.56% of the 25 students will have a mean score between 75 and 79.

Probabilities of x and $\bar{x}$

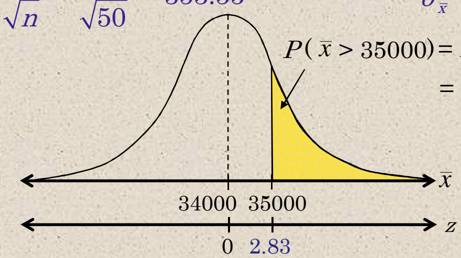
Example:
The population mean salary for auto mechanics is μ = \$34,000 with a standard deviation of σ = \$2,500. Find the probability that the <u>mean</u> salary for a randomly selected <u>sample of 50</u> mechanics is greater than \$35,000.

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

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

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

$P(\bar{x} > 35000) = P(z > 2.83) = 1 - P(z < 2.83)$
$= 1 - 0.9977 = 0.0023$



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 $\bar{x}$

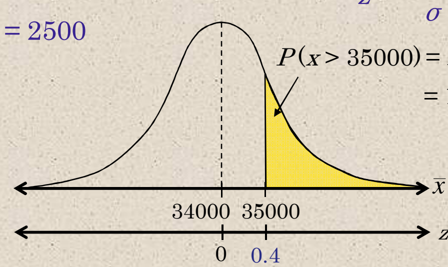
Example:
The population mean salary for auto mechanics is μ = \$34,000 with a standard deviation of σ = \$2,500. Find the probability that the salary for <u>one</u> randomly selected mechanic is greater than \$35,000.
(Notice that the Central Limit Theorem does not apply.)

$\mu = 34000$

$\sigma = 2500$

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

$P(x > 35000) = P(z > 0.4) = 1 - P(z < 0.4)$
$= 1 - 0.6554 = 0.3446$



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 one 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.