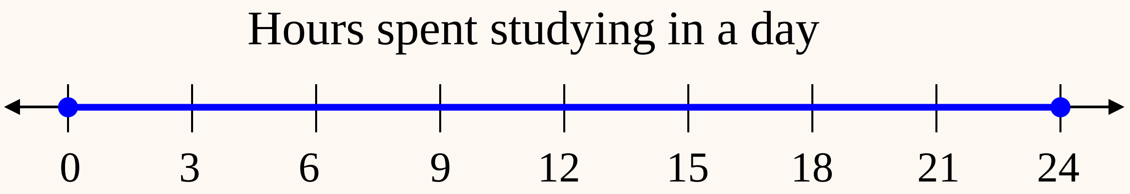
Chapter 6

Introduction to Normal Distributions and the Standard Normal Distributions

Properties of a Normal Distribution

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.

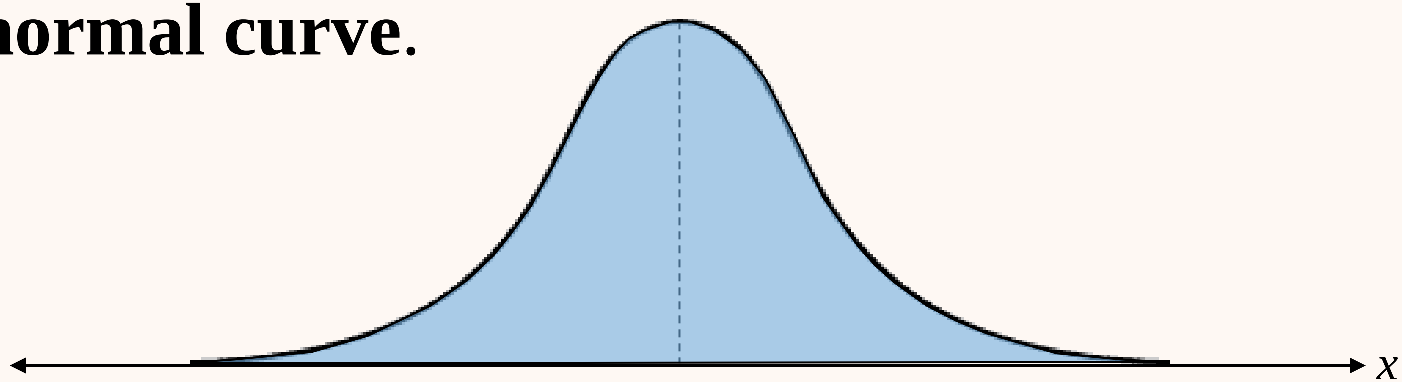
Continuous probability distribution

- The probability distribution of a continuous random variable.

Properties of a Normal Distribution

Normal distribution

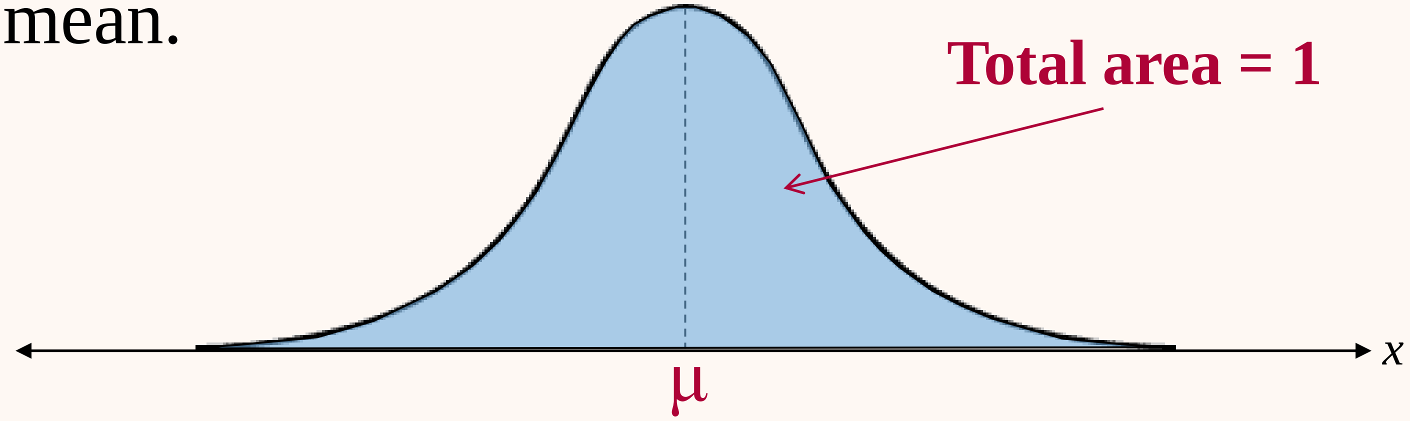
- A continuous probability distribution for a random variable, *x*.
- The most important continuous probability distribution in statistics.
- The graph of a normal distribution is called the **normal curve**.



.

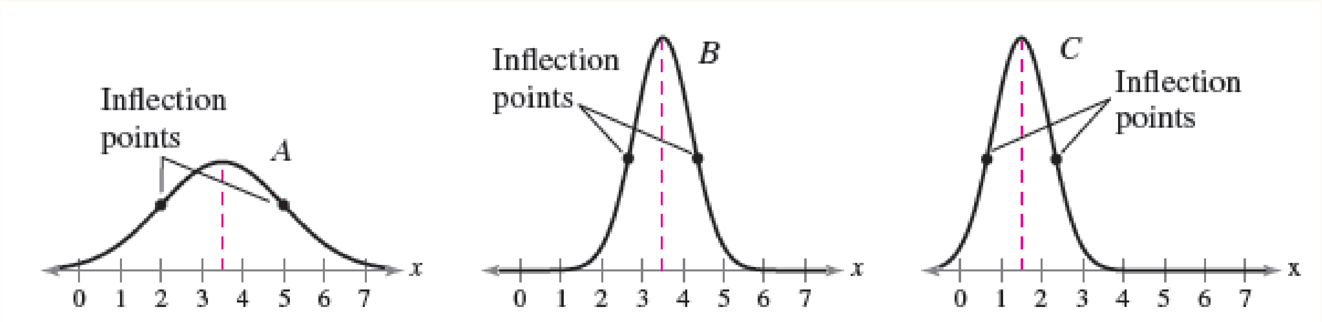
Properties of a Normal Distribution

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.



Means and Standard Deviations

- A normal distribution can have any mean and any positive standard deviation.
- The mean gives the location of the line of symmetry.
- The standard deviation describes the spread of the data.



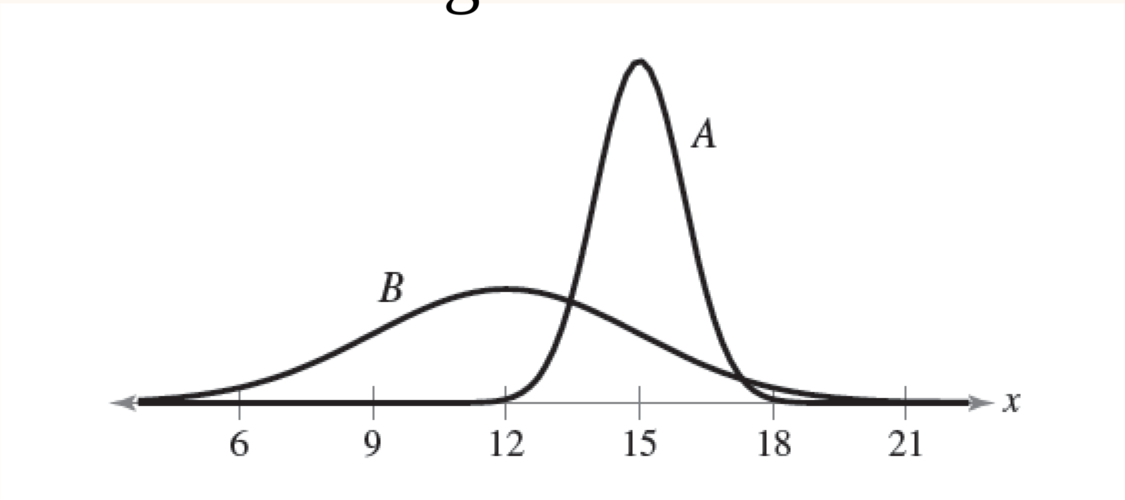
$\mu = 3.5$ $\sigma = 1.5$

$\mu = 3.5$ $\sigma = 0.7$

$\mu = 1.5$ $\sigma = 0.7$

Example: Understanding Mean and Standard Deviation

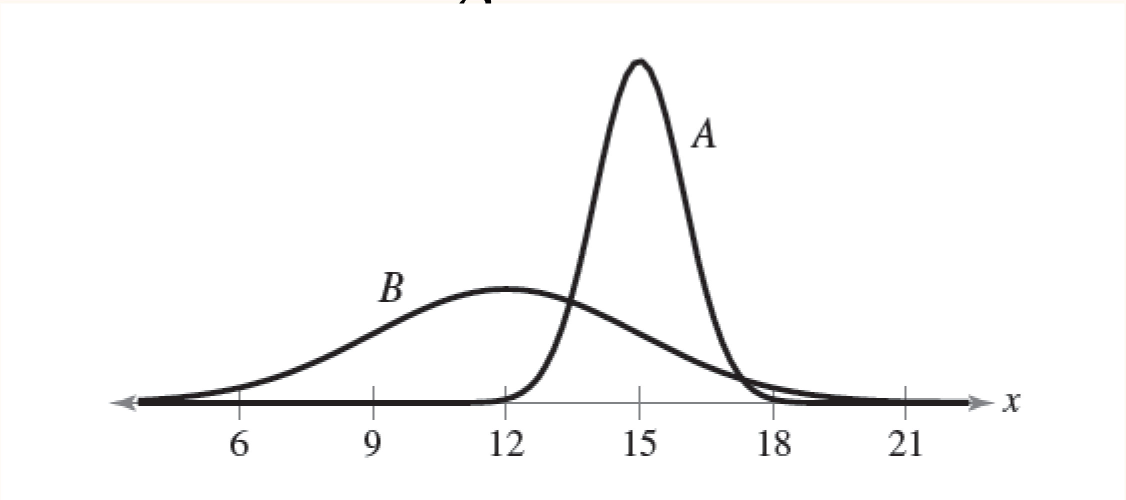
1. Which curve has the greater mean?



.

Example: Understanding Mean and Standard Deviation

2. Which curve has the greater standard deviation?

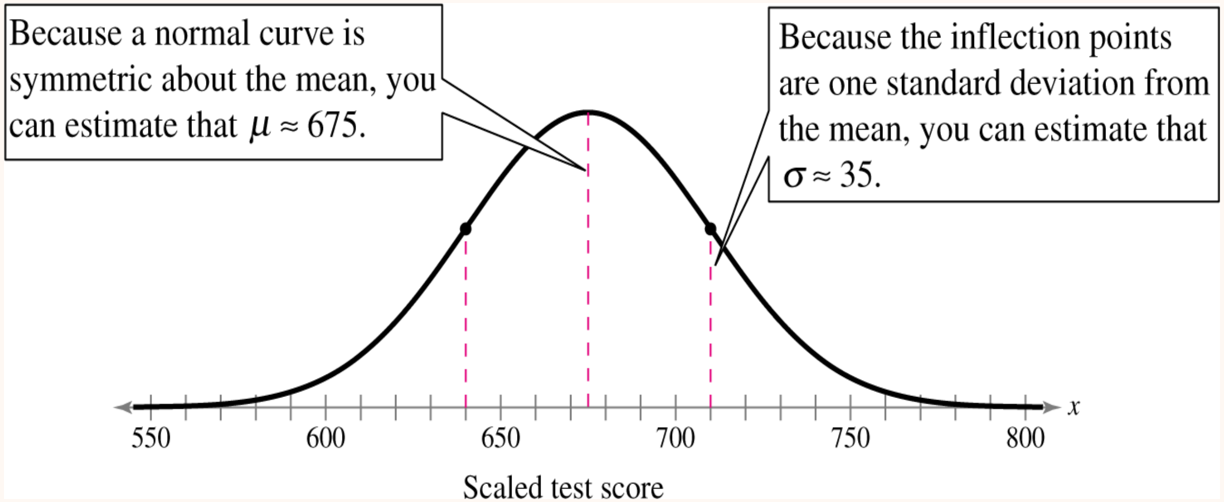


.

Example: Interpreting Graphs

The scaled test scores for New York State Grade 8 Mathematics Test are normally distributed. The normal curve shown below represents this distribution. Estimate the standard deviation.

Solution:

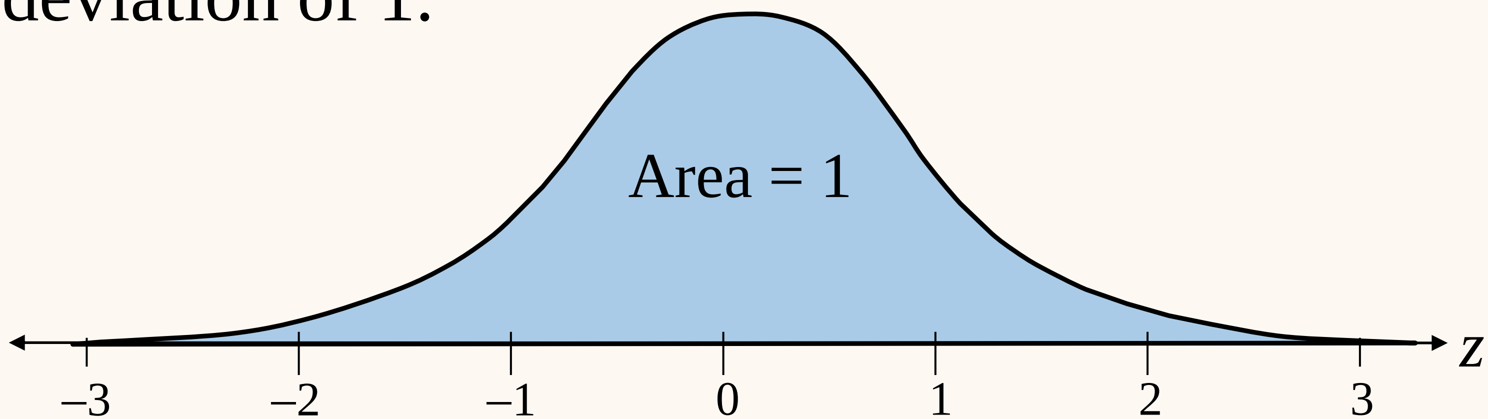


.

The Standard Normal Distribution

Standard normal distribution

- A normal distribution with a mean of 0 and a standard deviation of 1.



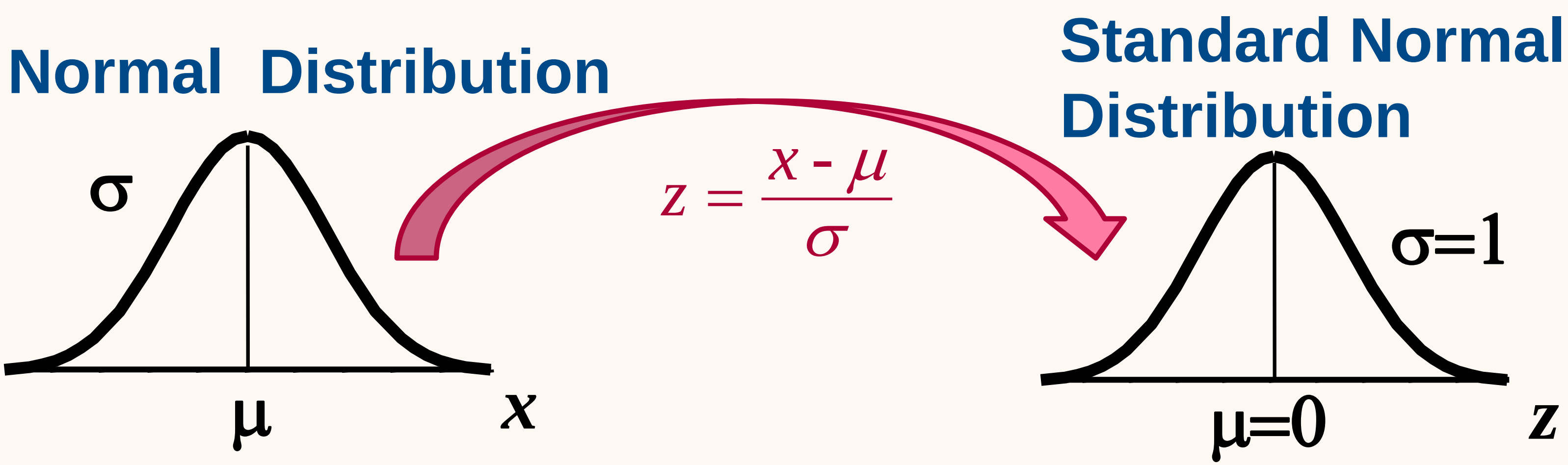
- Any *x*-value can be transformed into a *z*-score by using the formula

$$z = \frac{\text{Value - Mean}}{\text{Standard deviation}} = \frac{x - \mu}{\sigma}$$

.

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.

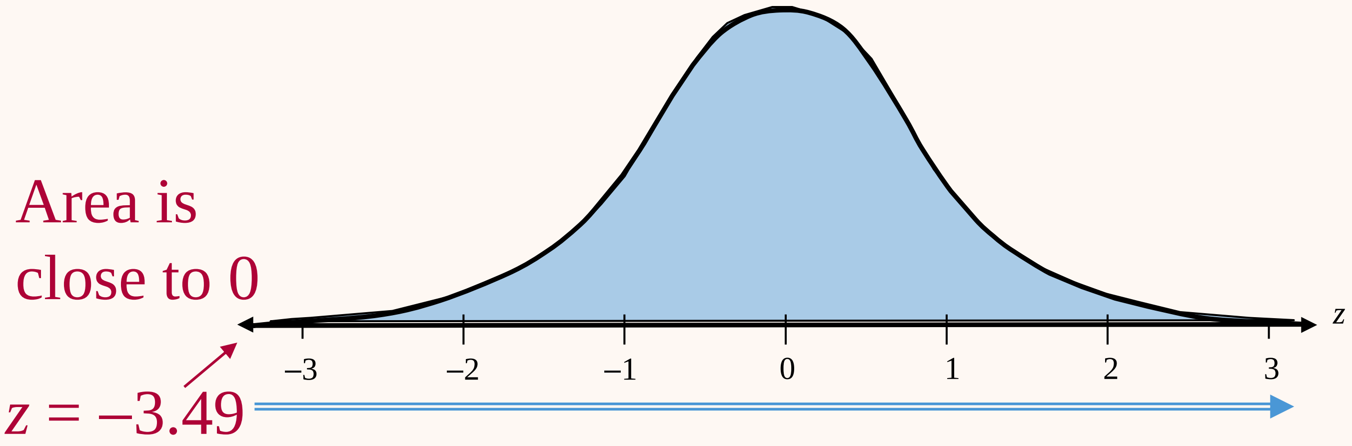


- Use the Standard Normal Table to find the cumulative area under the standard normal curve.

.

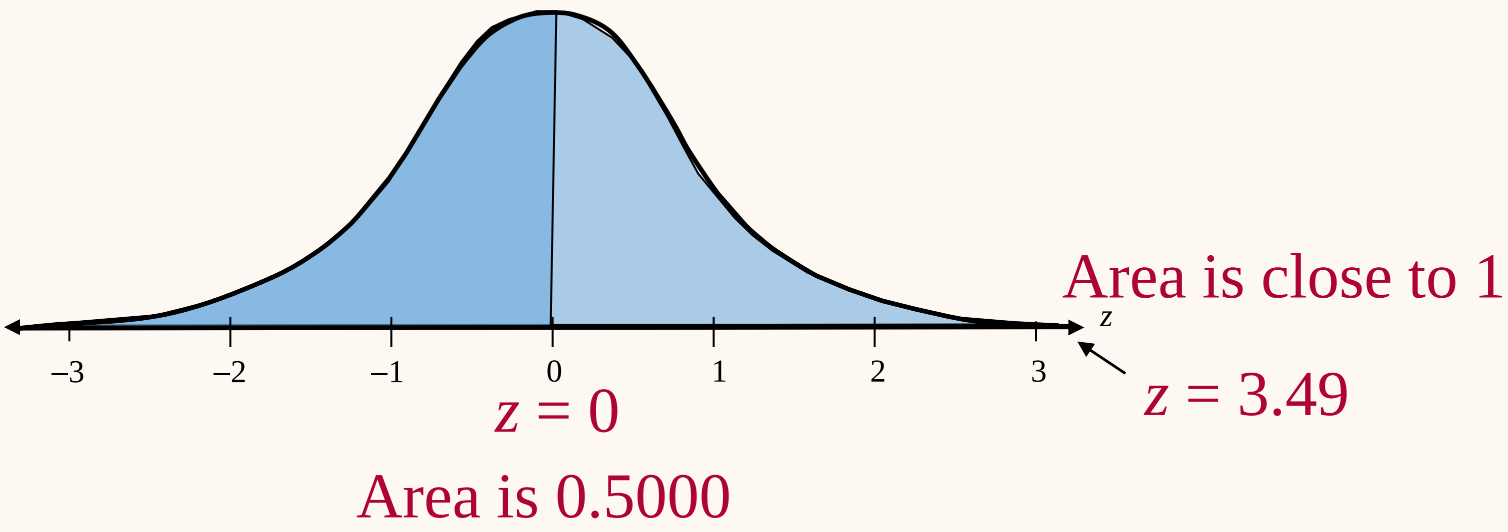
Properties of the Standard Normal Distribution

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.



Properties of the Standard Normal Distribution

3. The cumulative area for $z = 0$ is 0.5000.
4. The cumulative area is close to 1 for z-scores close to $z = 3.49$.



.

Example: Using The Standard Normal Table

Find the cumulative area that corresponds to a *z*-score of 1.15.

z	.00	.01	.02	.03	.04	.05	.06
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279

Solution:

Find 1.1 in the left hand column.

Move across the row to the column under 0.05

The area to the left of *z* = 1.15 is 0.8749.

Example: Using The Standard Normal Table

Find the cumulative area that corresponds to a *z*-score of –0.24.

z	.09	.08	.07	.06	.05	.04	.03
−3.4	.0002	.0003	.0003	.0003	.0003	.0003	.0003
−3.3	.0003	.0004	.0004	.0004	.0004	.0004	.0004
−3.2	.0005	.0005	.0005	.0006	.0006	.0006	.0006
−0.5	.2776	.2810	.2843	.2877	.2912	.2946	.2981
−0.4	.3121	.3156	.3192	.3228	.3264	.3300	.3336
−0.3	.3483	.3520	.3557	.3594	.3632	.3669	.3707
−0.2	.3859	.3897	.3936	.3974	.4013	.4052	.4090
−0.1	.4247	.4286	.4325	.4364	.4404	.4443	.4483
−0.0	.4641	.4681	.4721	.4761	.4801	.4840	.4880

Solution:

Find –0.2 in the left hand column.

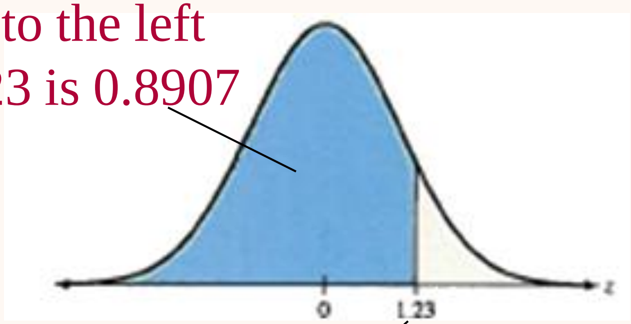
Move across the row to the column under 0.04

The area to the left of *z* = –0.24 is 0.4052.

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.

2. The area to the left of $z = 1.23$ is 0.8907

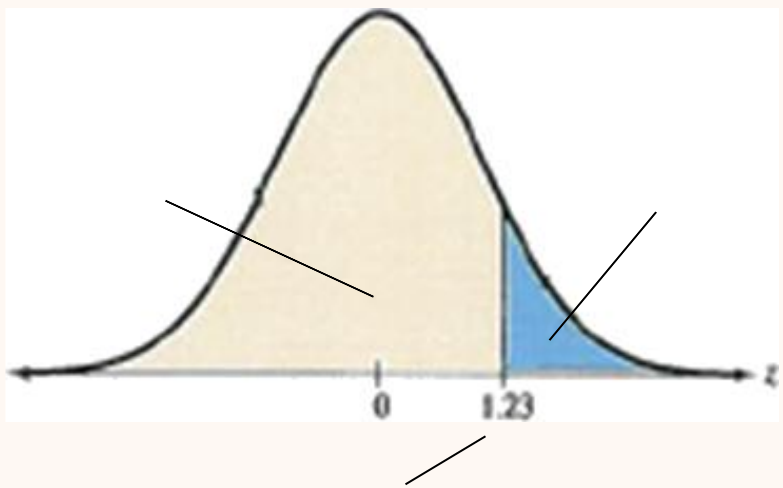


1. Use the table to find the area for the *z*-score

.

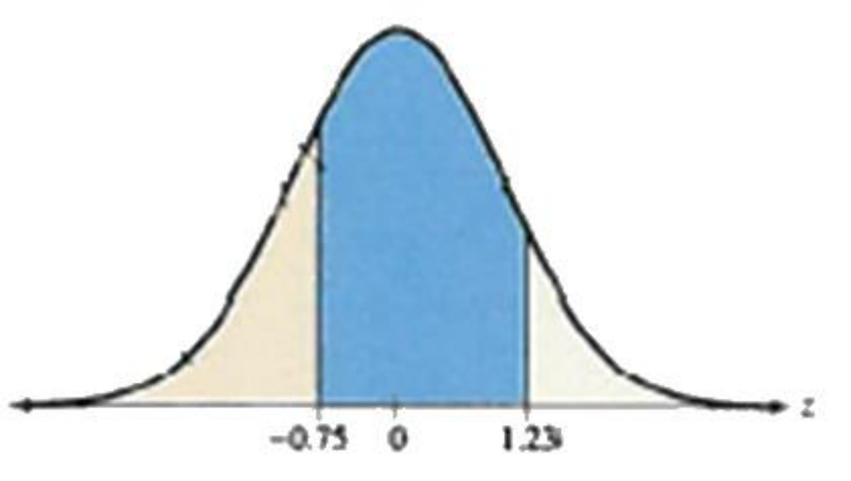
Finding Areas Under the Standard Normal Curve

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.



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.

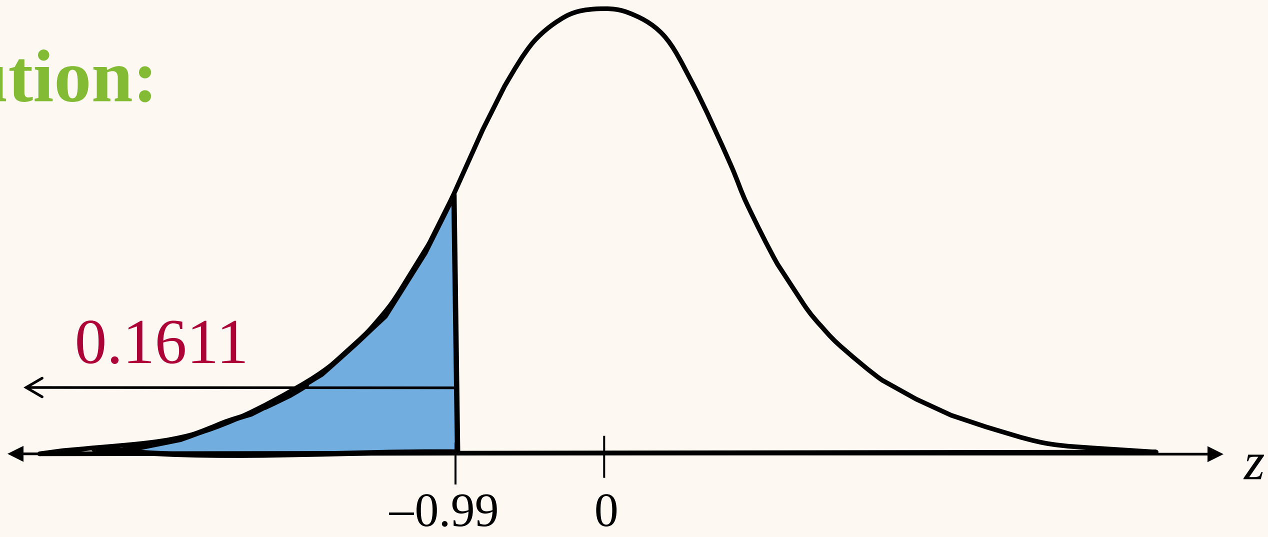


.

Example: Finding Area Under the Standard Normal Curve

Find the area under the standard normal curve to the left of z = -0.99.

Solution:



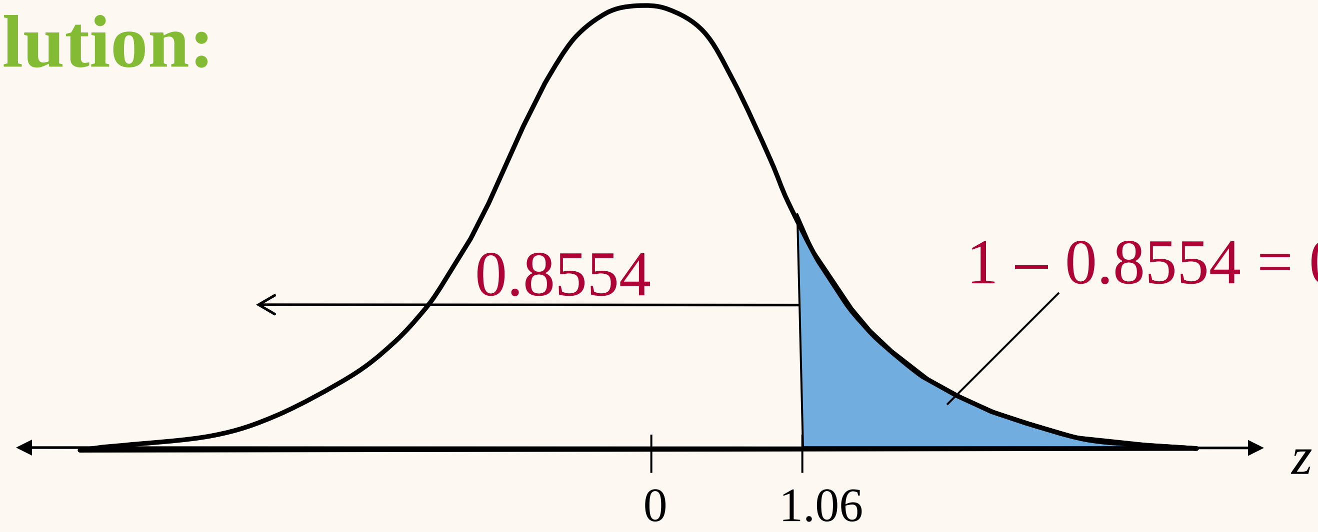
From the Standard Normal Table, the area is equal to 0.1611.

.

Example: Finding Area Under the Standard Normal Curve

Find the area under the standard normal curve to the right of $z = 1.06$.

Solution:



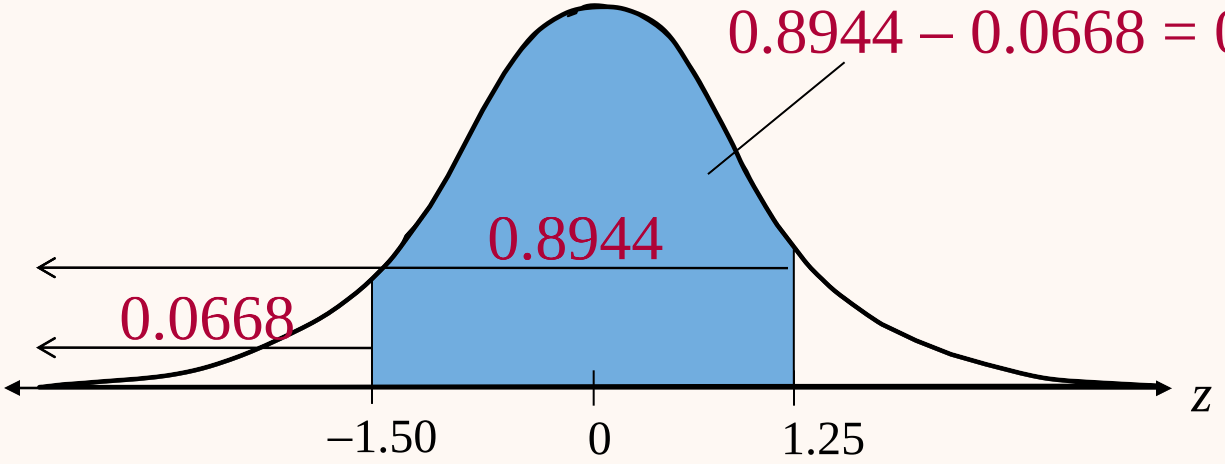
From the Standard Normal Table, the area is equal to 0.1446.

.

Example: Finding Area Under the Standard Normal Curve

Find the area under the standard normal curve between $z = -1.5$ and $z = 1.25$.

Solution:



From the Standard Normal Table, the area is equal to 0.8276.

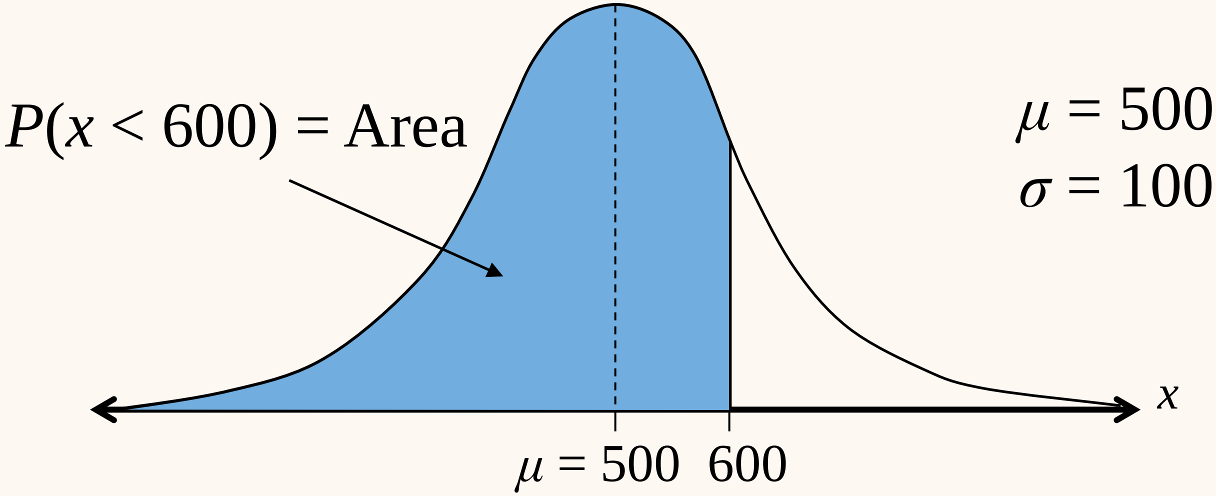
.

Section 6.2

Normal Distributions: Finding Probabilities

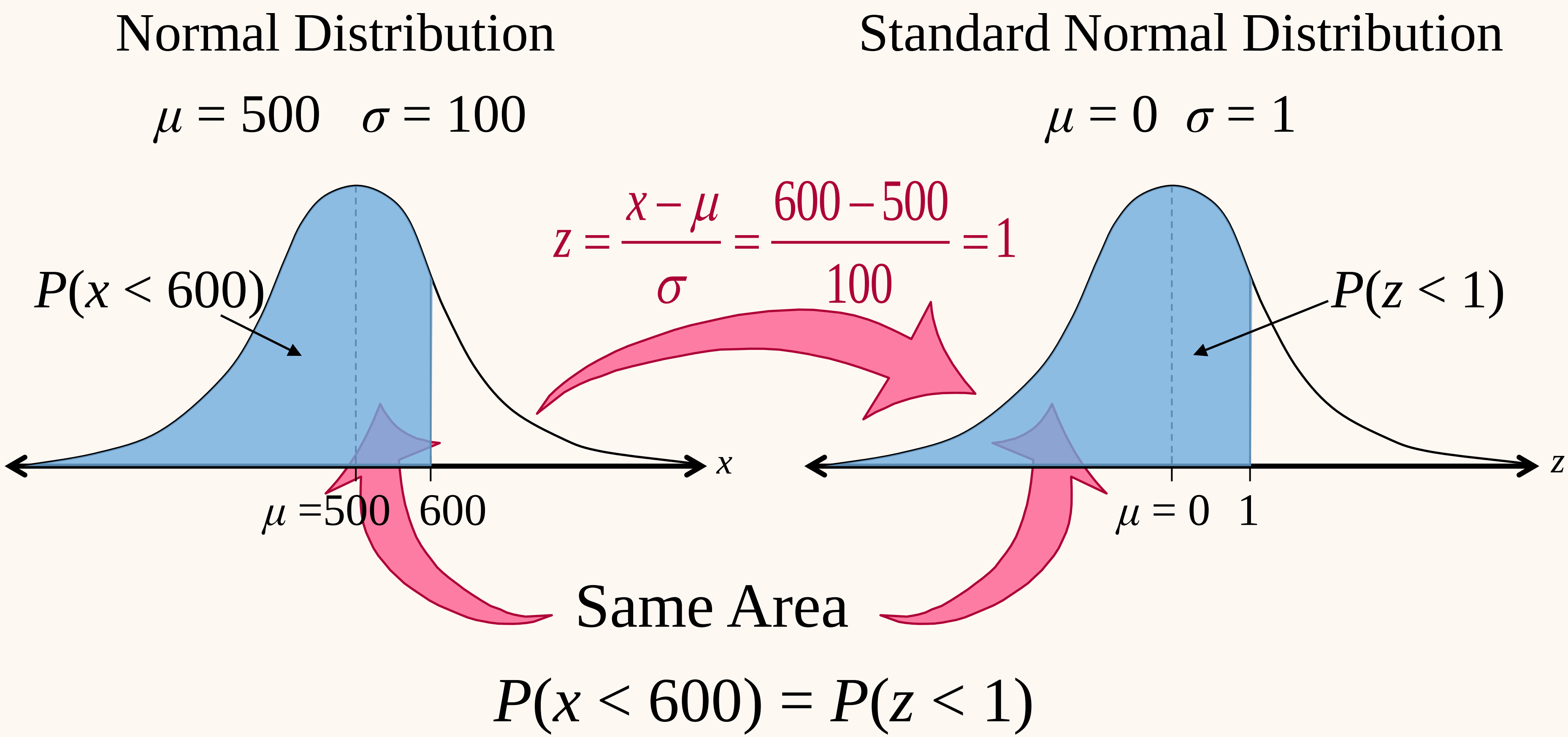
Probability and Normal Distributions

- 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



Example: Finding Probabilities for Normal Distributions

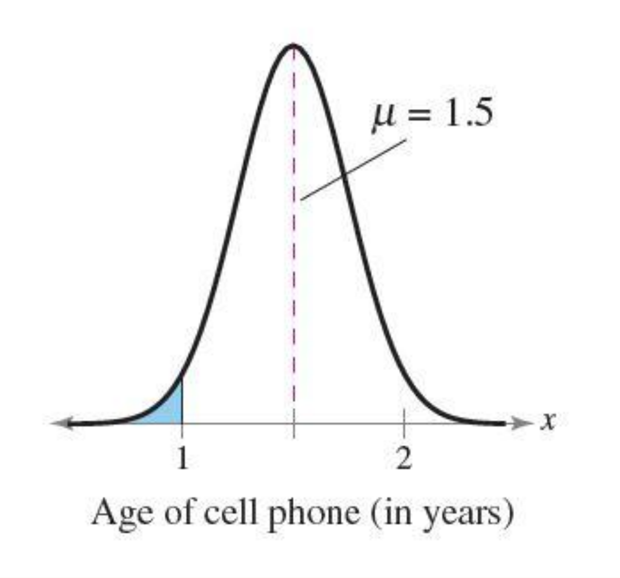
A survey indicates that people use their cellular phones an average of 1.5 years before buying a new one. The standard deviation is 0.25 year. A cellular phone user is selected at random. Find the probability that the user will use their current phone for less than 1 year before buying a new one. Assume that the variable x is normally distributed. *(Source: Fonebak)*

.

Solution: Finding Probabilities for Normal Distributions

Normal Distribution

$\mu = 1.5 \quad \sigma = 0.25$



The z-score for $x = 1$ is

$$z = \frac{x - \mu}{\sigma} = \frac{1 - 1.5}{0.25} = -2$$

The Standard Normal Table shows that $P(z < -2) = 0.0228$. Thus, $P(x < -1) = \mathbf{0.0228}$.

2.28% of cell phone users will keep their phones for less than 1 year before buying a new one.

Example: Finding Probabilities for Normal Distributions

A survey indicates that for each trip to the supermarket, a shopper spends an average of 45 minutes with a standard deviation of 12 minutes in the store. The length of time spent in the store is normally distributed and is represented by the variable x. A shopper enters the store. Find the probability that the shopper will be in the store for between 24 and 54 minutes.

.

Example: Finding Probabilities for Normal Distributions

Find the probability that the shopper will be in the store more than 39 minutes. (Recall μ = 45 minutes and σ = 12 minutes)

.

Solution: Finding Probabilities for Normal Distributions

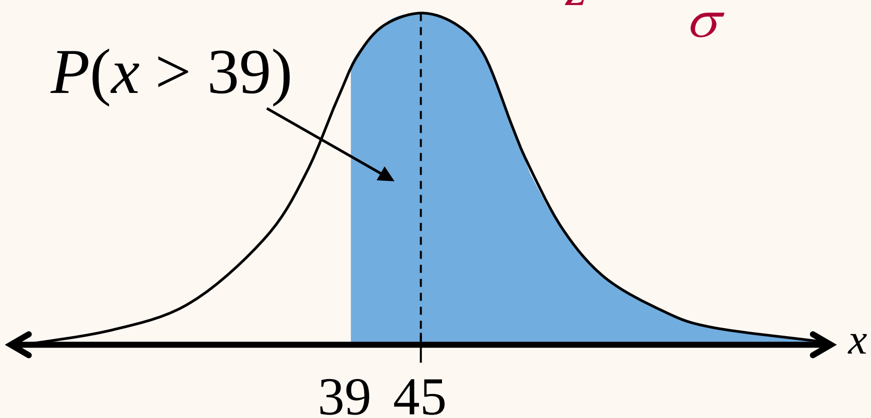
Normal Distribution

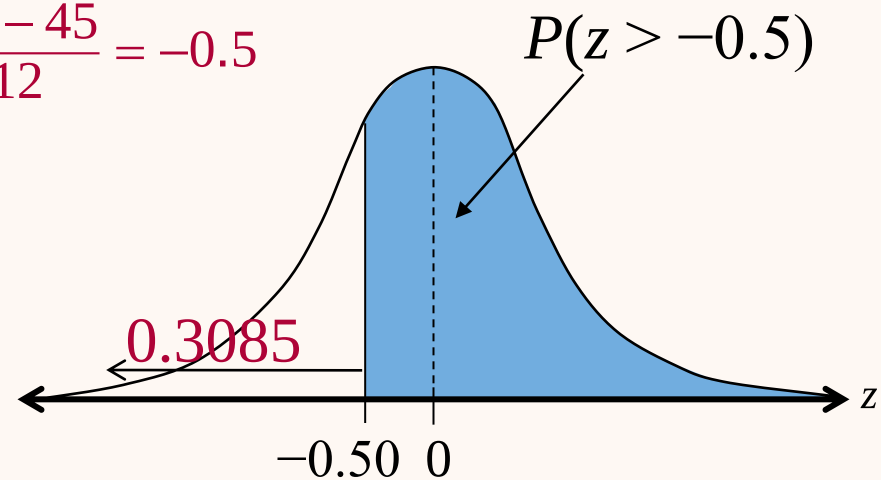
$\mu = 45 \quad \sigma = 12$

Standard Normal Distribution

$\mu = 0 \quad \sigma = 1$

$$z = \frac{x - \mu}{\sigma} = \frac{39 - 45}{12} = -0.5$$





$$P(x > 39) = P(z > -0.5) = 1 - 0.3085 = \mathbf{0.6915}$$

.

Example: Finding Probabilities for Normal Distributions

If 200 shoppers enter the store, how many shoppers would you expect to be in the store more than 39 minutes?

Solution:

Recall $P(x > 39) = 0.6915$.

200(0.6915) =138.3 (or about 138) shoppers

.

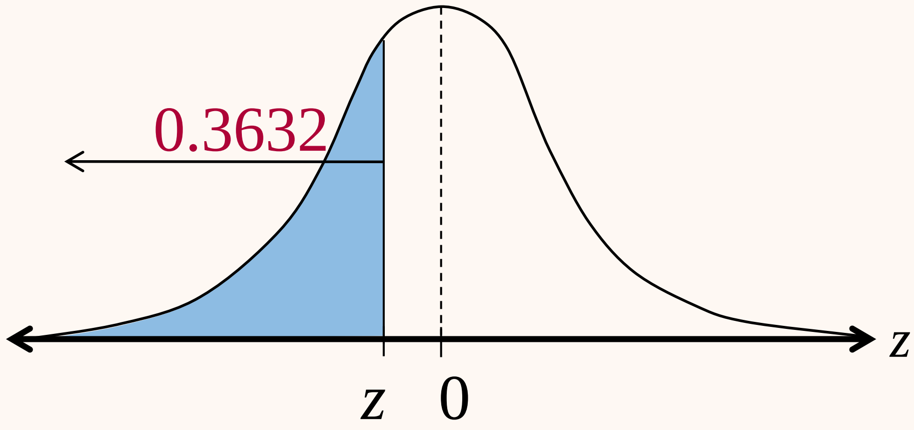
Section 6.3

Normal Distributions: Finding Values

Example: Finding a *z*-Score Given an Area

Find the *z*-score that corresponds to a cumulative area of 0.3632.

Solution:



Solution: Finding a *z*-Score Given an Area

- Locate 0.3632 in the body of the Standard Normal Table.

z	.09	.08	.07	.06	.05	.04	.03
−3.4	.0002	.0003	.0003	.0003	.0003	.0003	.0003
−3.3	.0003	.0004	.0004	.0004	.0004	.0004	.0004
−3.2	.0005	.0005	.0005	.0006	.0006	.0006	.0006
−0.5	.2776	.2810	.2843	.2877	.2912	.2946	.2981
−0.4	.3121	.3156	.3192	.3228	.3264	.3300	.3336
−0.3	.3483	.3520	.3557	.3594	.3632	.3669	.3707
−0.2	.3859	.3897	.3936	.3974	.4013	.4052	.4090
−0.1	.4247	.4286	.4325	.4364	.4404	.4443	.4483
−0.0	.4641	.4681	.4721	.4761	.4801	.4840	.4880

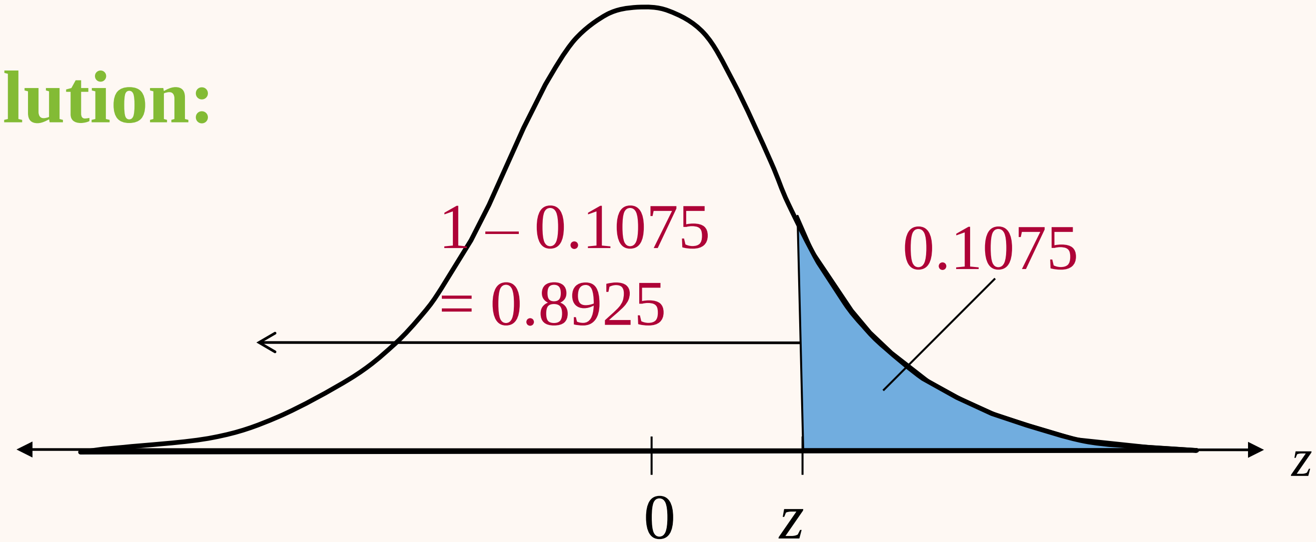
The *z*-score is –0.35.

- The values at the beginning of the corresponding row and at the top of the column give the *z*-score.

Example: Finding a *z*-Score Given an Area

Find the *z*-score that has 10.75% of the distribution's area to its right.

Solution:



Because the area to the right is 0.1075, the cumulative area is 0.8925.

Solution: Finding a *z*-Score Given an Area

- Locate 0.8925 in the body of the Standard Normal Table.

z	.00	.01	.02	.03	.04	.05	.06
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279

The *z*-score is 1.24.

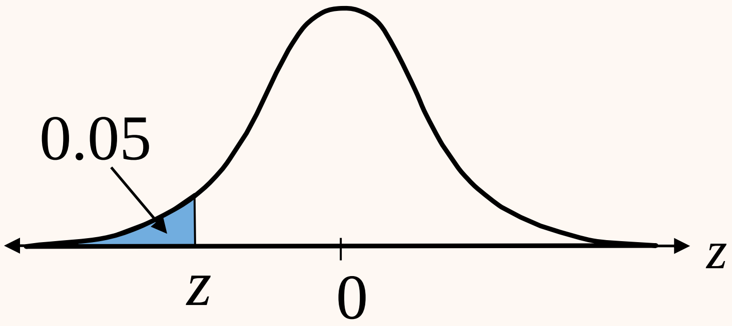
- The values at the beginning of the corresponding row and at the top of the column give the *z*-score.

Example: Finding a z-Score Given a Percentile

Find the z-score that corresponds to P_5.

Solution:

The z-score that corresponds to P_5 is the same z-score that corresponds to an area of 0.05.



The areas closest to 0.05 in the table are 0.0495 (z = –1.65) and 0.0505 (z = – 1.64). Because 0.05 is halfway between the two areas in the table, use the z-score that is halfway between –1.64 and –1.65. **The *z*-score is -1.645**.

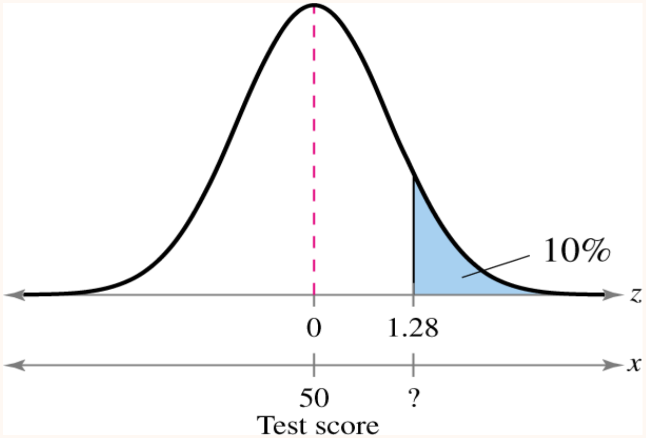
Example: Finding an x-Value

A veterinarian records the weights of cats treated at a clinic. The weights are normally distributed, with a mean of 9 pounds and a standard deviation of 2 pounds. Find the weights x corresponding to z-scores of 1.96, -0.44, and 0.

Example: Finding a Specific Data Value

Scores for the California Peace Officer Standards and Training test are normally distributed, with a mean of 50 and a standard deviation of 10. An agency will only hire applicants with scores in the top 10%. What is the lowest score you can earn and still be eligible to be hired by the agency?

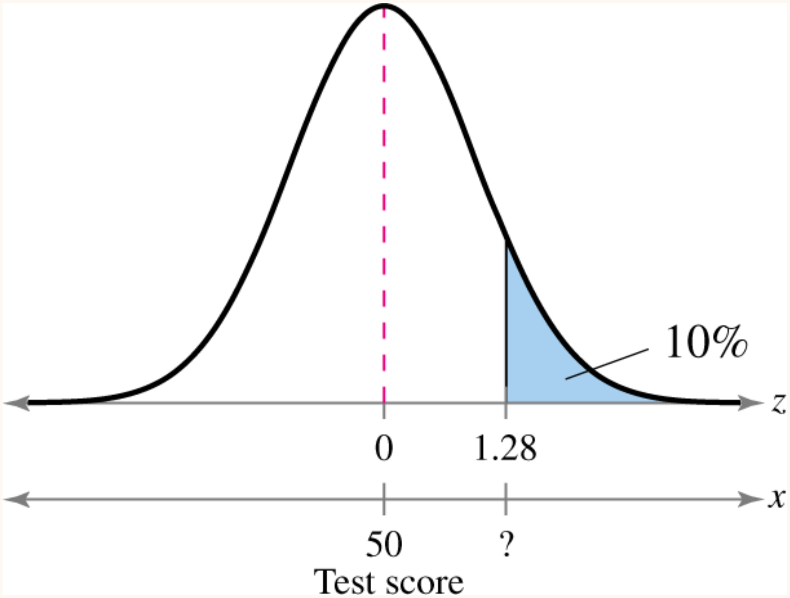
Solution:



An exam score in the top 10% is any score above the 90th percentile. Find the *z*-score that corresponds to a cumulative area of 0.9.

Solution: Finding a Specific Data Value

From the Standard Normal Table, the area closest to 0.9 is 0.8997. So the z-score that corresponds to an area of 0.9 is $\mathbf{z = 1.28}$.



Chapter 7

Sampling Distributions and the Central Limit Theorem

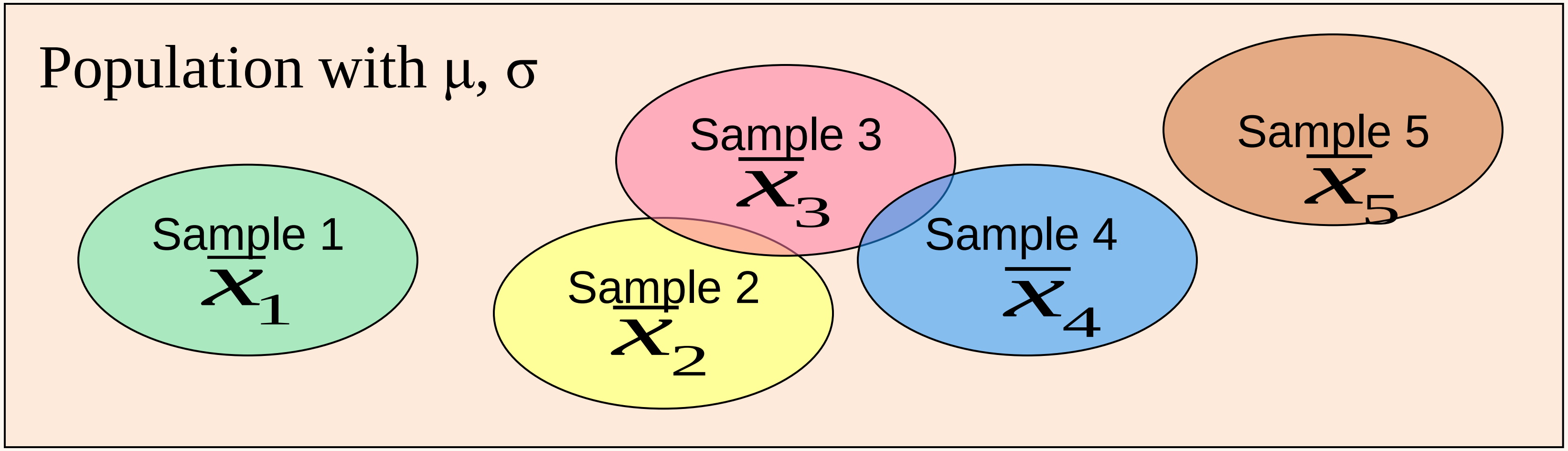
The Central Limit Theorem is one of the most important and useful theorems in statistics. This theorem forms the foundation for the inferential branch of statistics. The Central Limit Theorem describes the relationship between the sampling distribution of sample means and the population that the samples are taken from.

Sampling Distributions

Sampling distribution

- The probability distribution of a sample statistic.
- Formed when samples of size n are repeatedly taken from a population.
- e.g. Sampling distribution of sample means

Sampling Distribution of Sample Means



The sampling distribution consists of the values of the sample means, $\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4, \bar{x}_5, \ldots$

Properties of Sampling Distributions of Sample Means

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

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

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 the sample size, *n*.

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

- Called the **standard error of the mean**.

Example: Sampling Distribution of Sample Means

The population values {1, 3, 5, 7} are written on slips of paper and put in a box. Two slips of paper are randomly selected, with replacement.

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

Solution: Mean: $\mu = \frac{\Sigma x}{N} = 4$

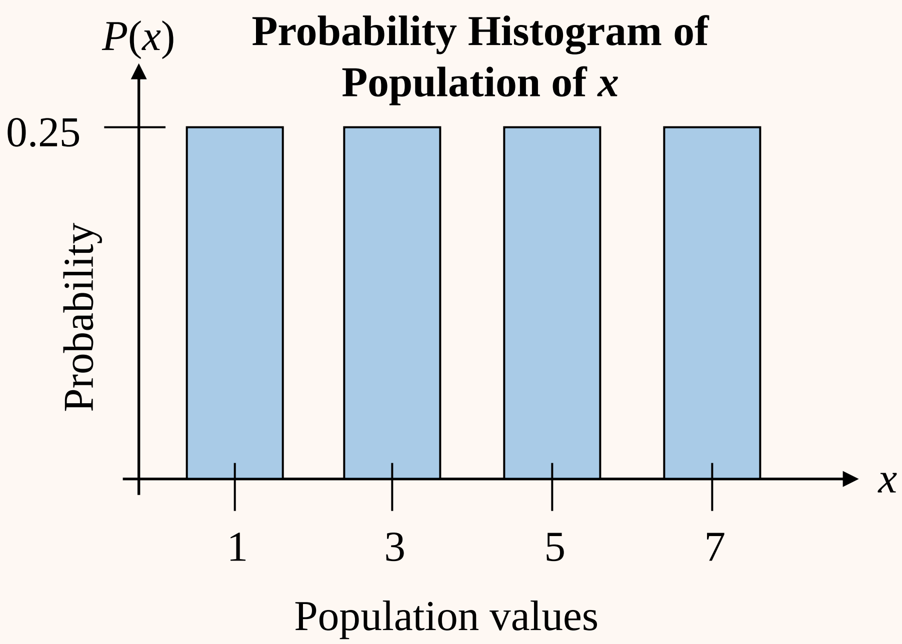
Variance: $\sigma^2 = \frac{\Sigma(x-\mu)^2}{N} = 5$

Standard Deviation: $\sigma = \sqrt{5} \approx 2.236$

Example: Sampling Distribution of Sample Means

b. Graph the probability histogram for the population values.

Solution:



All values have the same probability of being selected (uniform distribution)

Example: Sampling Distribution of Sample Means

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

Solution:

Sample	$\bar{x}$
1, 1	1
1, 3	2
1, 5	3
1, 7	4
3, 1	2
3, 3	3
3, 5	4
3, 7	5

Sample	$\bar{x}$
5, 1	3
5, 3	4
5, 5	5
5, 7	6
7, 1	4
7, 3	5
7, 5	6
7, 7	7

These means form the sampling distribution of sample means.

Example: Sampling Distribution of Sample Means

d. Construct the probability distribution of the sample means.

Solution:

$\bar{x}$	f	Probability
1	1	0.0625
2	2	0.1250
3	3	0.1875
4	4	0.2500
5	3	0.1875
6	2	0.1250
7	1	0.0625

Example: Sampling Distribution of Sample Means

e. Find the mean, variance, and standard deviation of the sampling distribution of the sample means.

Solution:

The mean, variance, and standard deviation of the 16 sample means are:

$$\mu_{\bar{x}} = 4 \qquad \sigma_{\bar{x}}^2 = \frac{5}{2} = 2.5 \qquad \sigma_{\bar{x}} = \sqrt{2.5} \approx 1.581$$

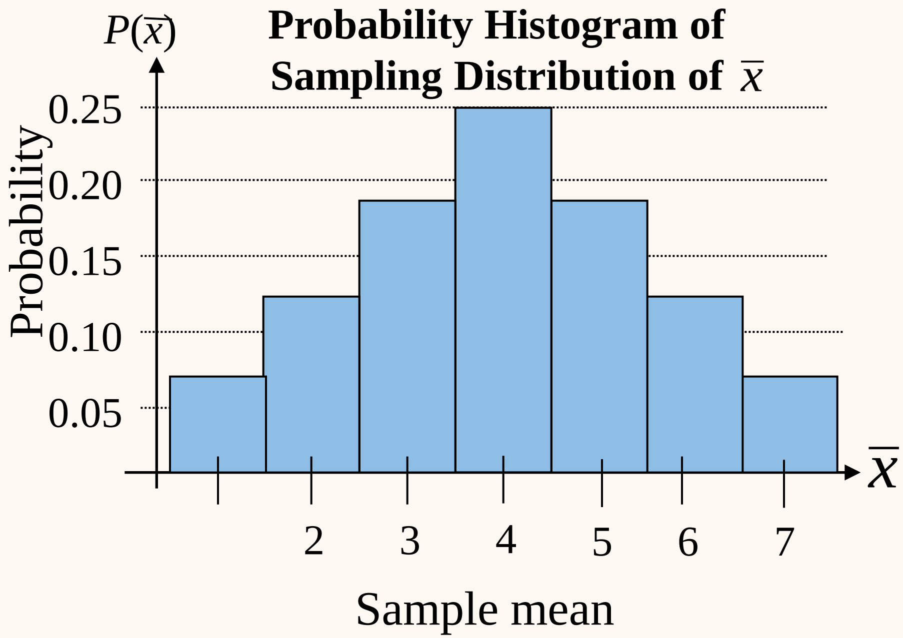
These results satisfy the properties of sampling distributions of sample means.

$$\mu_{\bar{x}} = \mu = 4 \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{5}}{\sqrt{2}} \approx \frac{2.236}{\sqrt{2}} \approx 1.581$$

Example: Sampling Distribution of Sample Means

f. Graph the probability histogram for the sampling distribution of the sample means.

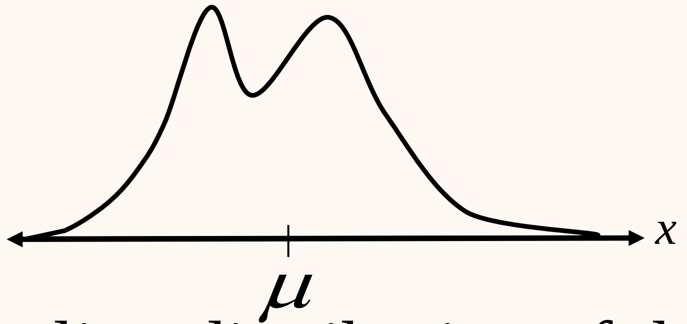
Solution:



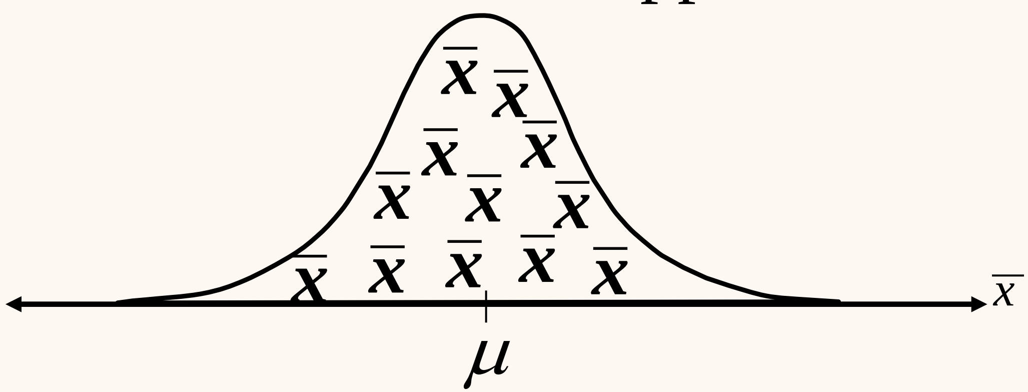
The shape of the graph is symmetric and bell shaped. It approximates a normal distribution.

The Central Limit Theorem

1. If samples of size $n \geq 30$, are drawn from any population with mean = μ and standard deviation = σ,

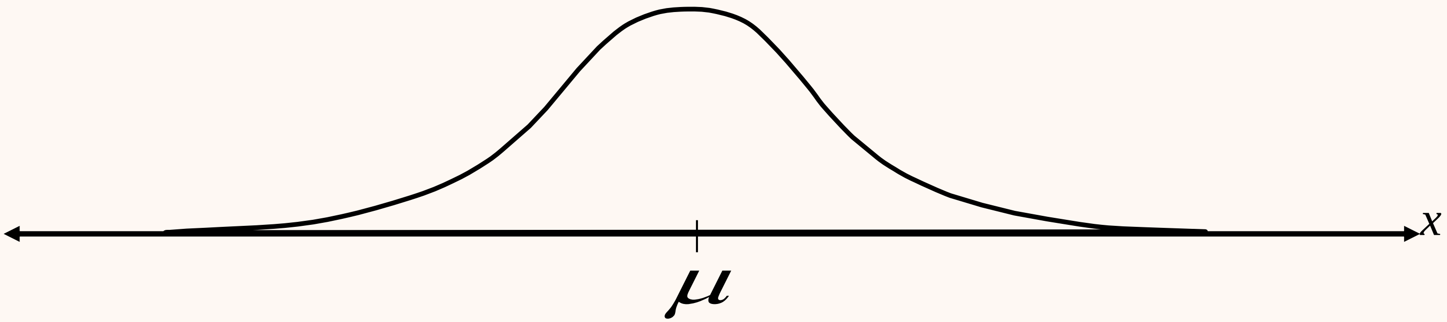


then the sampling distribution of the sample means approximates a normal distribution. The greater the sample size, the better the approximation.

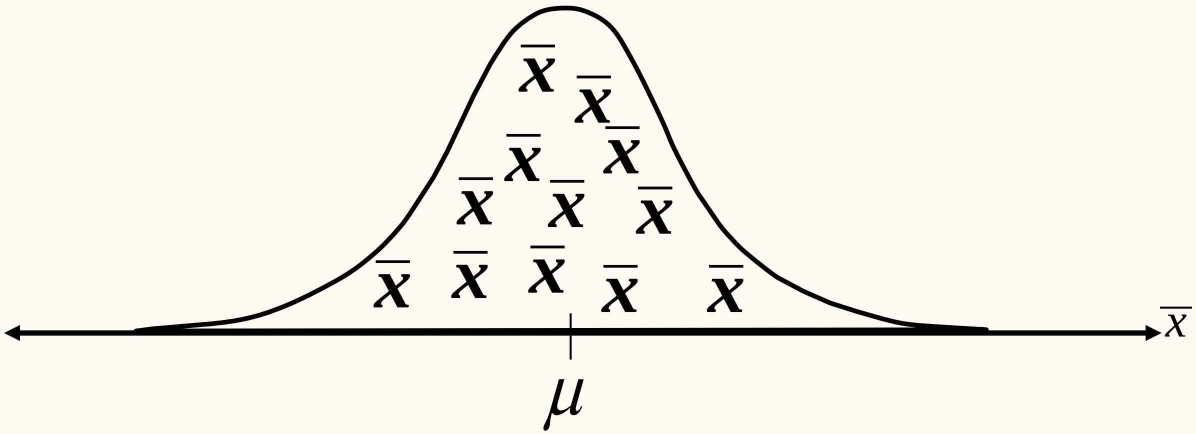


The Central Limit Theorem

2. If the population itself is normally distributed,



the sampling distribution of the sample means is normally 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.

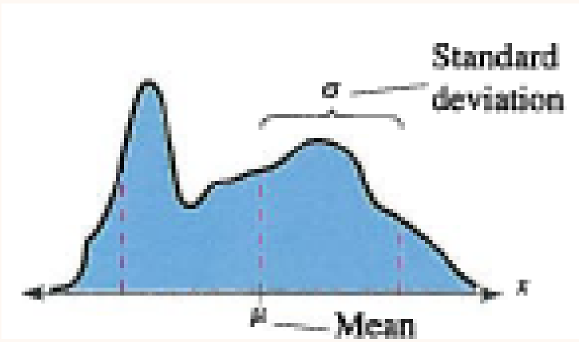
$$\mu_{\bar{x}} = \mu$$

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

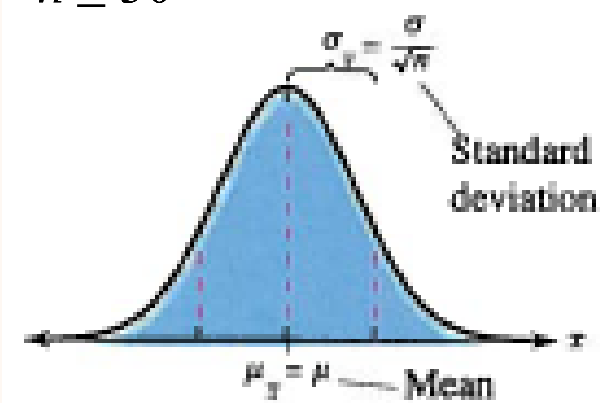
Standard deviation (**standard error of the mean)**

The Central Limit Theorem

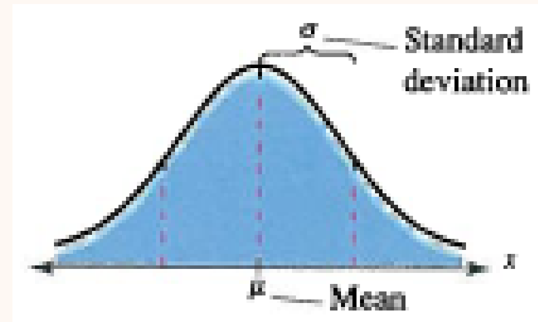
1. Any Population Distribution



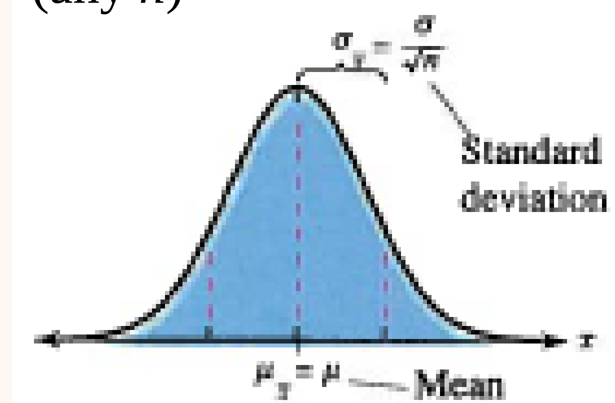
Distribution of Sample Means,
$n \geq 30$



2. Normal Population Distribution

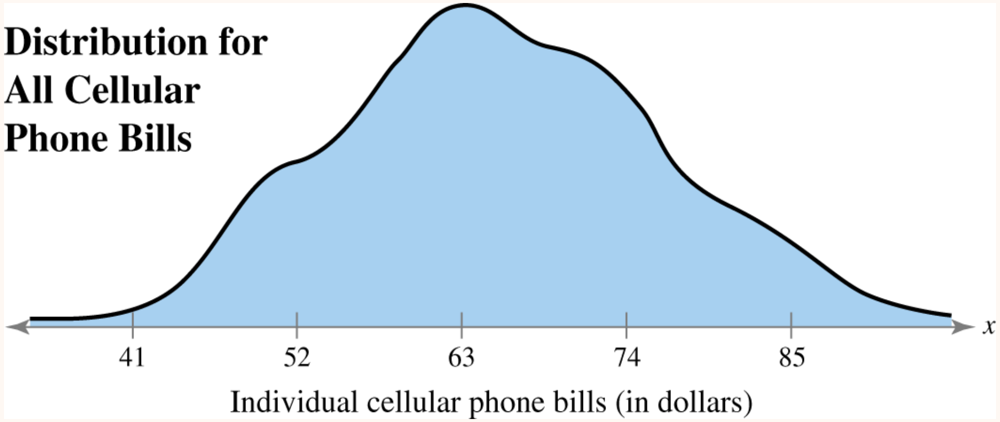


Distribution of Sample Means,
(any n)



Example: Interpreting the Central Limit Theorem

Cellular phone bills for residents of a city have a mean of $63 and a standard deviation of $11. Random samples of 100 cellular phone bills are drawn from this population and the mean of each sample is determined. Find the mean and standard error of the mean of the sampling distribution. Then sketch a graph of the sampling distribution of sample means.



Solution: Interpreting the Central Limit Theorem

- The mean of the sampling distribution is equal to the population mean

$$\mu_{\bar{x}} = \mu = 63$$

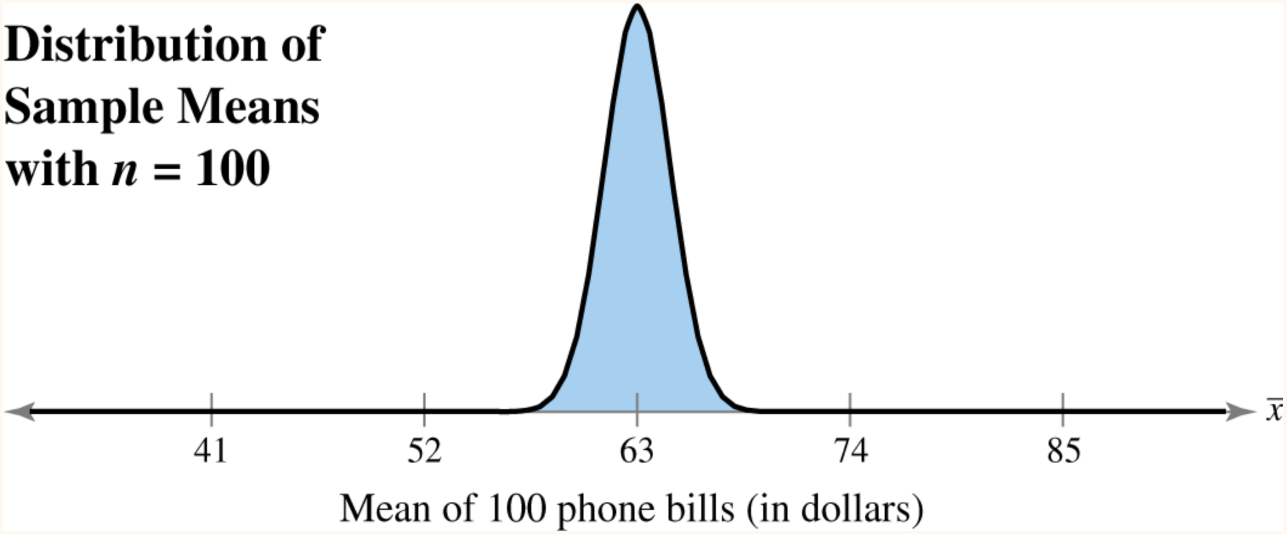
- The standard error of the mean is equal to the population standard deviation divided by the square root of *n*.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{11}{\sqrt{100}} = 1.1$$

Solution: Interpreting the Central Limit Theorem

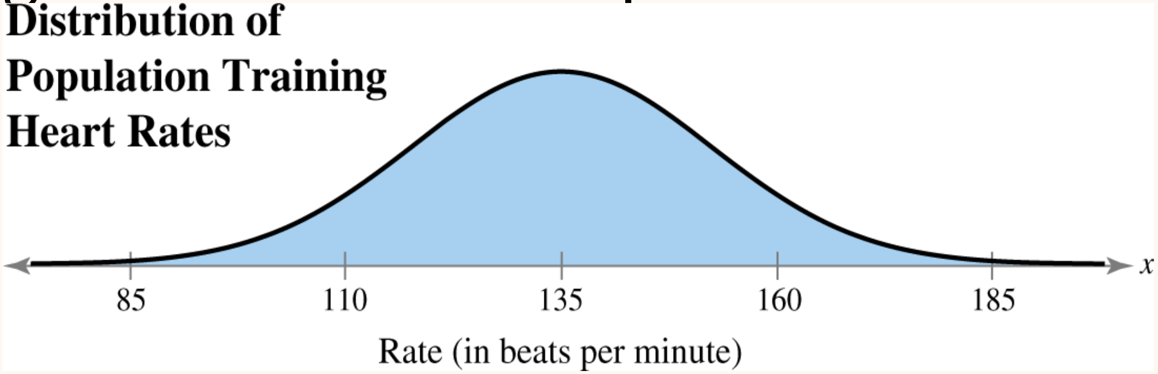
- Since the sample size is greater than 30, the sampling distribution can be approximated by a normal distribution with

$$\mu_{\bar{x}} = \$63 \qquad \sigma_{\bar{x}} = \$1.10$$



Example: Interpreting the Central Limit Theorem

The training heart rates of all 20-years old athletes are normally distributed, with a mean of 135 beats per minute and standard deviation of 18 beats per minute. Random samples of size 4 are drawn from this population, and the mean of each sample is determined. Find the mean and standard error of the mean of the sampling distribution. Then sketch a graph of the sampling distribution of sample means.



Solution: Interpreting the Central Limit Theorem

- The mean of the sampling distribution is equal to the population mean

$$\mu_{\overline{x}} = \mu = 135$$

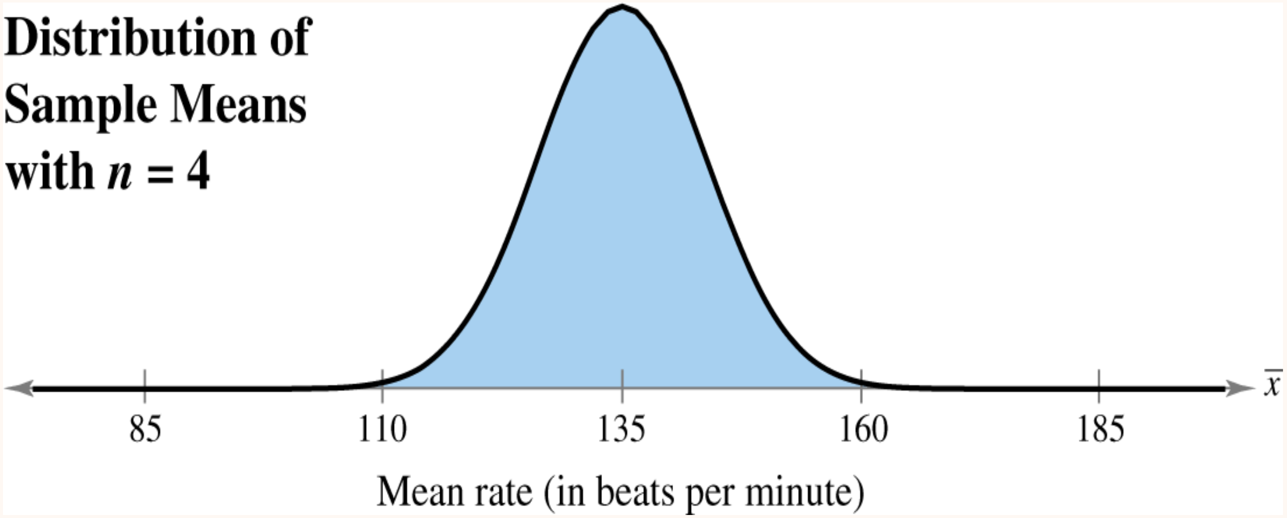
- The standard error of the mean is equal to the population standard deviation divided by the square root of *n*.

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{18}{\sqrt{4}} = 9$$

Solution: Interpreting the Central Limit Theorem

- Since the population is normally distributed, the sampling distribution of the sample means is also normally distributed.

$$\mu_{\bar{x}} = 135 \qquad \sigma_{\bar{x}} = 9$$



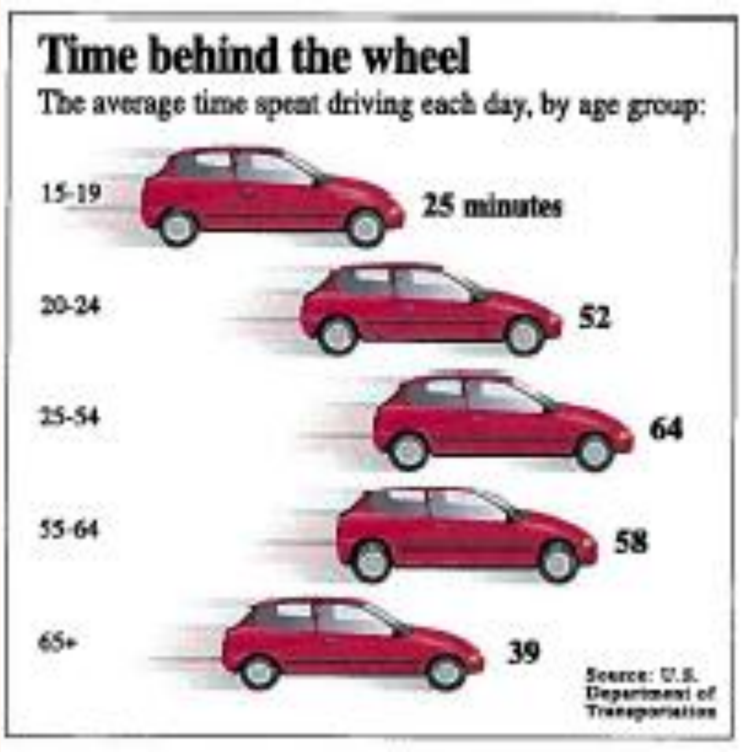
Probability and the Central Limit Theorem

- To transform $\bar{x}$ to a *z*-score

$$z = \frac{\text{Value-Mean}}{\text{Standard Error}} = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

Example: Probabilities for Sampling Distributions

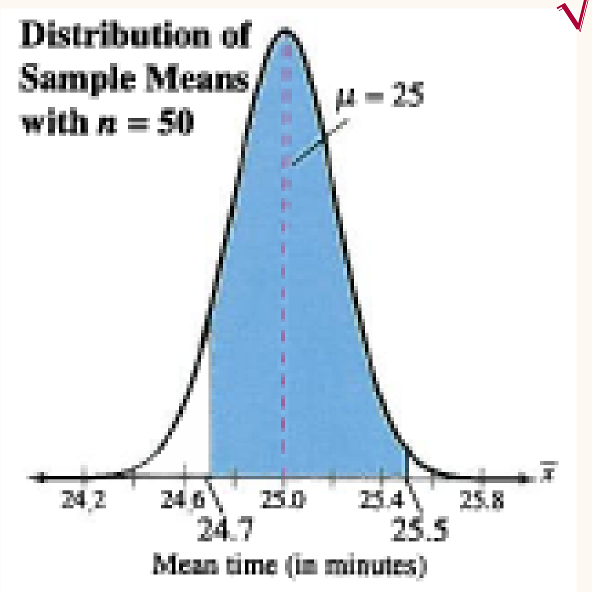
The graph shows the length of time people spend driving each day. You randomly select 50 drivers age 15 to 19. What is the probability that the mean time they spend driving each day is between 24.7 and 25.5 minutes? Assume that $\sigma = 1.5$ minutes.



Solution: Probabilities for Sampling Distributions

From the Central Limit Theorem (sample size is greater than 30), the sampling distribution of sample means is approximately normal with

$$\mu_{\bar{x}} = \mu = 25 \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.5}{\sqrt{50}} \approx 0.21213$$



Solution: Probabilities for Sampling Distributions

Normal Distribution

$\mu = 25 \quad \sigma = 0.21213$

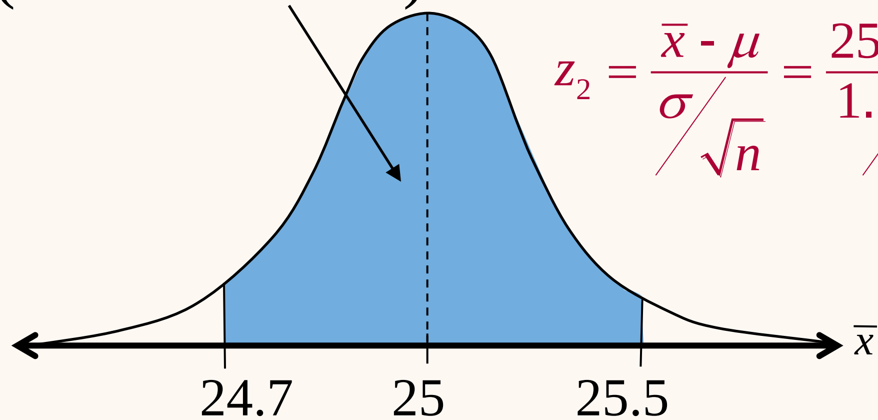
Standard Normal Distribution

$\mu = 0 \quad \sigma = 1$

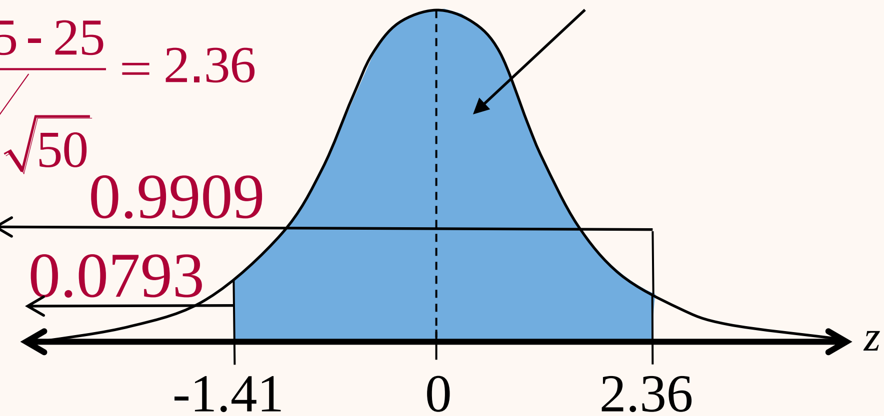
$$z_1 = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{24.7 - 25}{1.5/\sqrt{50}} = -1.41$$

$$z_2 = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{25.5 - 25}{1.5/\sqrt{50}} = 2.36$$

$P(24.7 < \bar{x} < 25.5)$



$P(-1.41 < z < 2.36)$



$$P(24 < \bar{x} < 54) = P(-1.41 < z < 2.36)$$
$$= 0.9909 - 0.0793 = \mathbf{0.9116}$$

Example: Probabilities for x and $\bar{x}$

An education finance corporation claims that the average credit card debts carried by undergraduates are normally distributed, with a mean of \$3173 and a standard deviation of \$1120. *(Adapted from Sallie Mae)*

1. What is the probability that a randomly selected undergraduate, who is a credit card holder, has a credit card balance less than \$2700?

Solution:

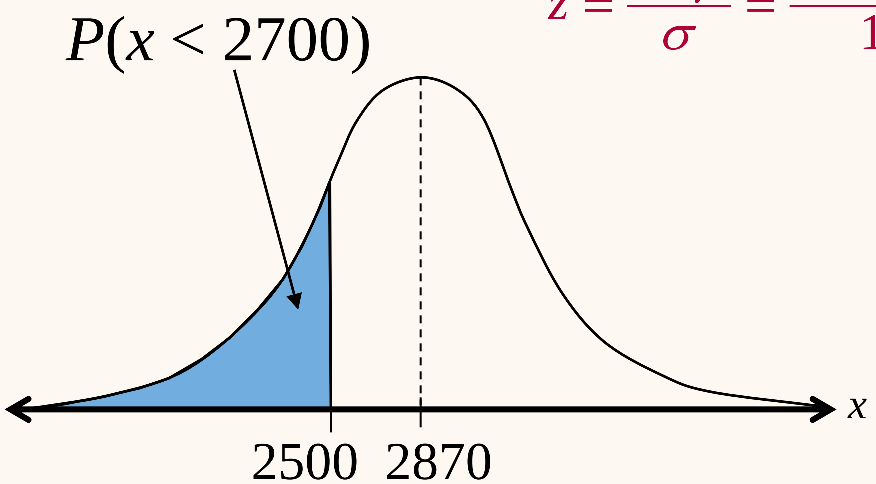
You are asked to find the probability associated with a certain value of the random variable *x*.

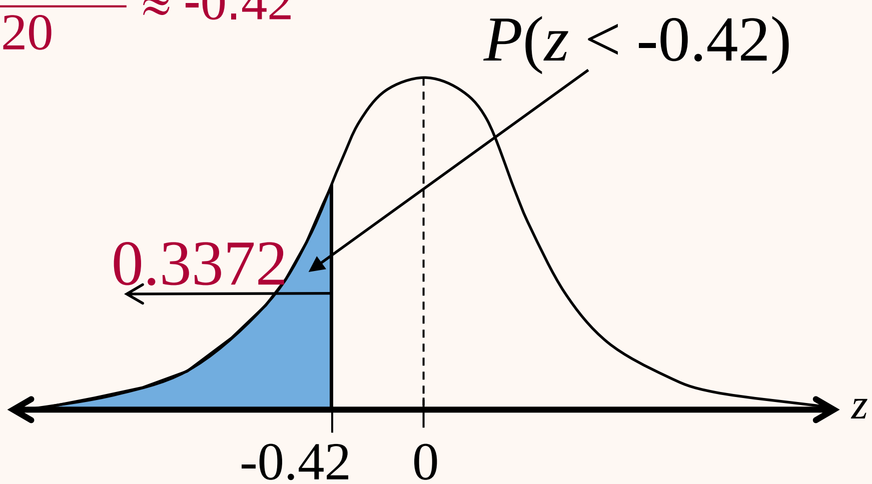
Solution: Probabilities for *x* and $\bar{x}$

Normal Distribution
$\mu = 3173 \quad \sigma = 1120$

Standard Normal Distribution
$\mu = 0 \quad \sigma = 1$

$$z = \frac{x - \mu}{\sigma} = \frac{2700 - 3173}{1120} \approx -0.42$$





$P(x < 2700) = P(z < 0.42) =$ **0.3372**

Example: Probabilities for *x* and $\bar{x}$

2. You randomly select 25 undergraduates who are credit card holders. What is the probability that their mean credit card balance is less than $2700?

Solution:

You are asked to find the probability associated with a sample mean $\bar{x}$.

$$\mu_{\bar{x}} = \mu = 3173 \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1120}{\sqrt{25}} = 224$$

Solution: Probabilities for *x* and $\overline{x}$

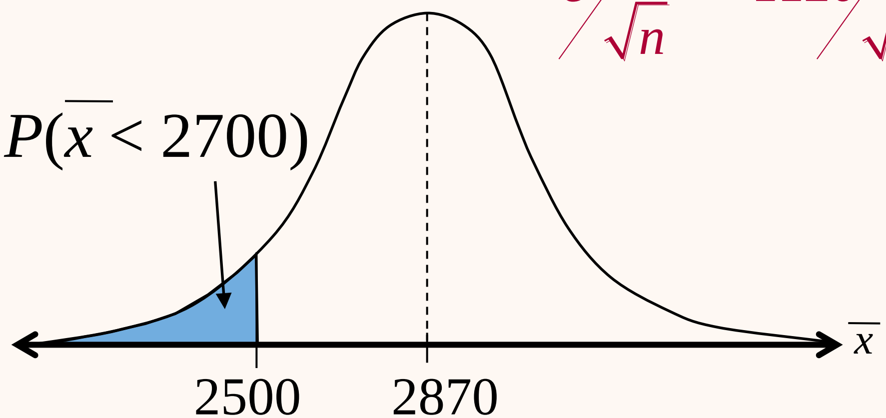
Normal Distribution

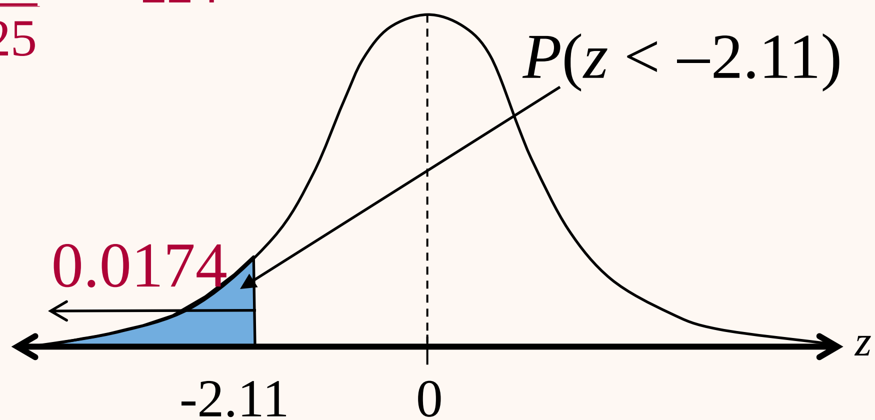
$\mu = 3173 \quad \sigma = 1120$

Standard Normal Distribution

$\mu = 0 \quad \sigma = 1$

$$z = \frac{\overline{x} - \mu}{\sigma / \sqrt{n}} = \frac{2700 - 3173}{1120 / \sqrt{25}} = \frac{-473}{224} \approx -2.11$$





$P(\overline{x} < 2700) = P(z < -2.11) =$ **0.0174**

Solution: Probabilities for x and $\bar{x}$

- There is a 34% chance that an undergraduate will have a balance less than $2700.
- There is only a 2% chance that the mean of a sample of 25 will have a balance less than $2700 (unusual event).
- It is possible that the sample is unusual or it is possible that the corporation's claim that the mean is $3173 is incorrect.