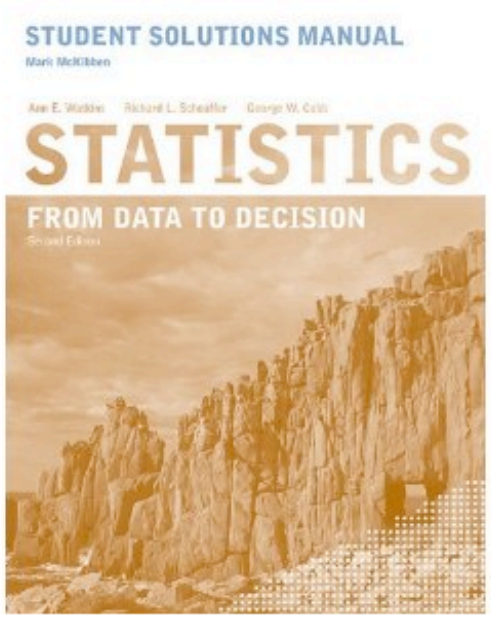
Math 140
Introductory Statistics



Statistics, Student Solutions Manual: From Data to Decision [Paperback]

Ann E. Watkins (Author), Richard L. Scheaffer (Author), George W. Cobb (Author)

No customer reviews yet. Be the first.

Price: $35.00 & this item ships for **FREE with Super Saver Shipping**. Details

In Stock.

Ships from and sold by **Amazon.com**. Gift-wrap available.

Only 1 left in stock--order soon (more on the way).

Want it delivered Thursday, September 16? Order it in the next 23 hours and 30 minutes, and choose **One-Day Shipping** at checkout. Details

6 new from $33.25 **1 used** from $35.95

Join Amazon Student and get FREE Two-Day Shipping for one year with Amazon Prime shipping benefits.

Recentering

Shift everything by a number c

The max will be shifted
The min will be shifted
Q1, Q3 will be shifted
The average and the median will be shifted

Will the range shift?

Recentering

Shift everything by a number c

The max will be shifted
The min will be shifted
Q1, Q3 will be shifted
The average and the median will be shifted

Will the range shift?

No!

Recentering

Standard deviation is

$$\sigma_{n-1} = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$$

If we shift the data x, also $\bar{x}$ will be shifted x+c and $\bar{x}$+c

Does their difference get shifted?

Recentering

Standard deviation is

$$\sigma_{n-1} = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$$

If we shift the data x, also $\bar{x}$ will be shifted x+c and $\bar{x}$+c

Does their difference get shifted?

No!

Recentering

Shift everything by a number c

The max will be shifted
The min will be shifted
Q1, Q3 will be shifted
The average and the median will be shifted

The range WILL NOT shift
The standard deviation WILL NOT shift

Rescaling

Multiply everything by a number d

The max will be rescaled
The min will be rescaled
Q1, Q3 will be rescaled
The average and the median will be rescaled

The range WILL rescale
The standard deviation WILL rescale

Rescaling

All number summaries will be rescaled by d

Recentering

Except for the range and the SD
all number summaries will be shifted by c.

An example

The distribution of the SAT scores for the University of Washington was roughly normal in shape, with mean 1055 and standard deviation 200.

1. What percentage of scores were 920 or below?

2. What SAT score separates the lowest 25% of the SAT scores from the rest?

An example

The distribution of the SAT scores for the University of Washington was roughly normal in shape, with mean 1055 and standard deviation 200.

1. What percentage of scores were 920 or below?

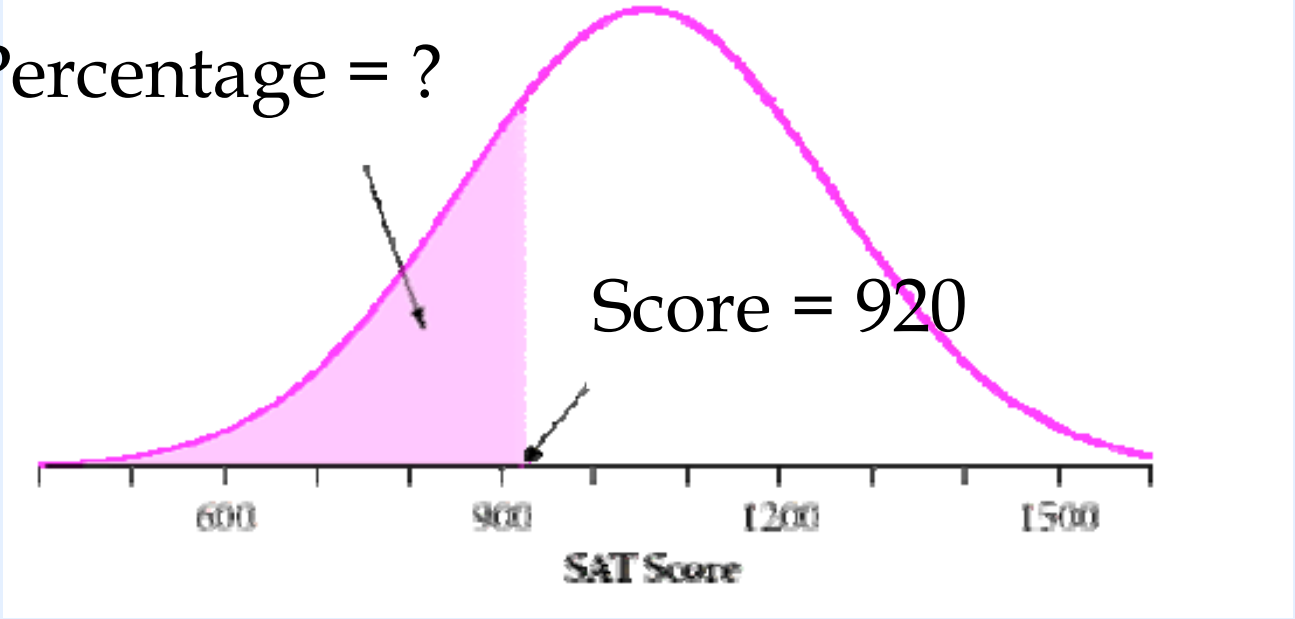
2. What SAT score separates the lowest 25% of the SAT scores from the rest?

We already know that 68% of data is between

855 and 1255

Unknown percentage problem

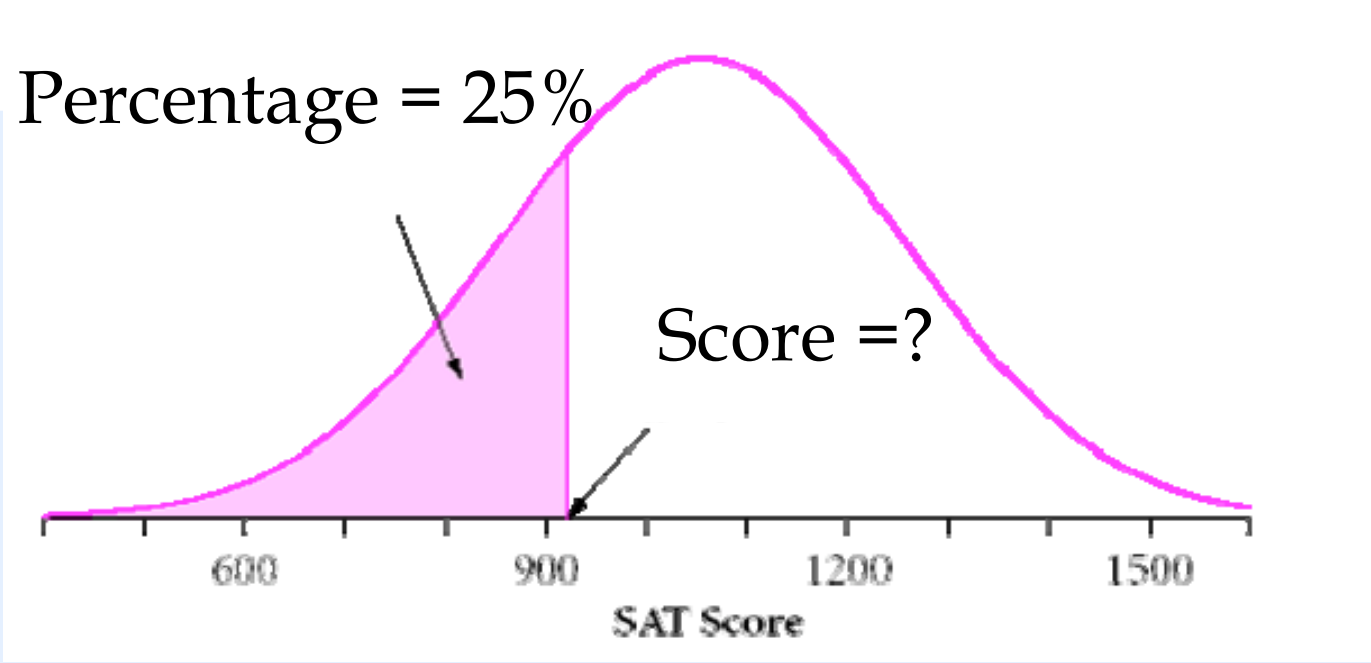
1. What percentage of scores were 920 or below?



Given x (a score), find the percentage

Unknown value problem

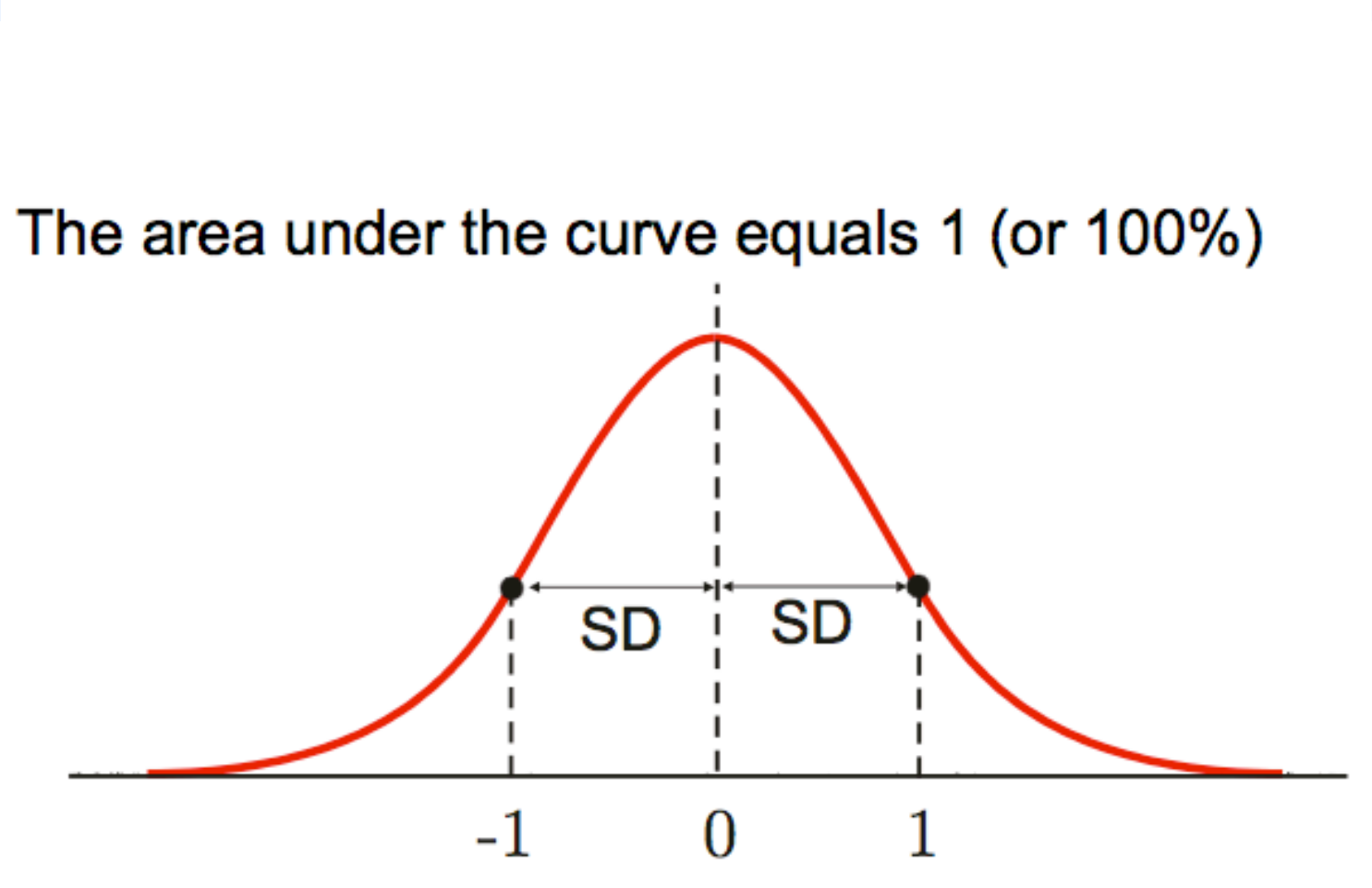
2. What SAT score separates the lowest 25% of the scores from the rest?



Given the percentage P, find the score z

Standard normal distribution

The normal distribution with mean =0 and SD = 1



Standard normal distribution

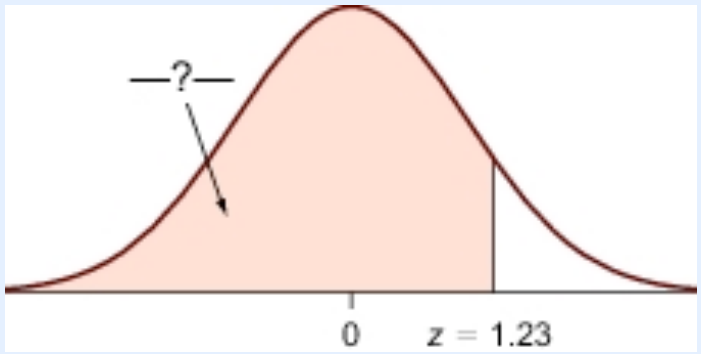
Any normal distribution can be rescaled
or recentered to give you the normal distribution

STANDARDIZING or
CONVERTING TO STANDARD UNITS

Given the score z find P
Unknown percentage

Table A. Page 759
Use the units and the first decimal to locate the row
and the closest hundredths digits to
locate the column.

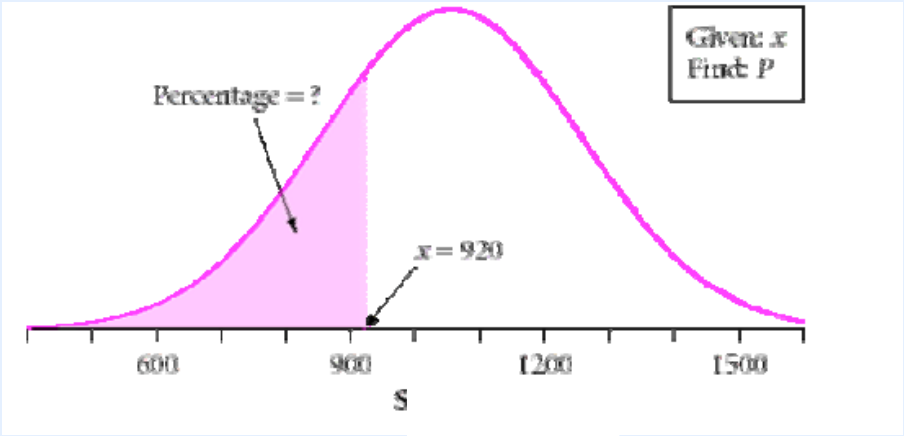
The number found is the percentage
to the left of z.



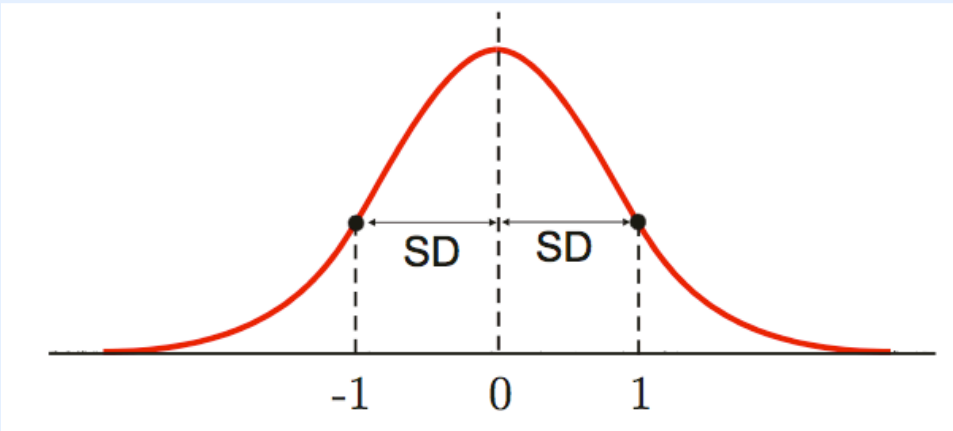
Symbols

Data prior to standardization is indicated with x

Data after standardization is indicated with z



SAT scores x



Standardized scores z

How do we "standardize?"

Recenter and Rescale

Recentering = Make the mean zero

How do we "standardize?"

Recenter and Rescale

Recentering = Make the mean zero

1. Subtract the mean from all values

$x-\bar{x}$

first step towards new values z

How do we "standardize?"

Recenter and Rescale

Rescale = Make the SD one

2. Divide all previous values by the SD

$$z = (x-\bar{x})/SD$$

Final step towards new values z

We say that z is the z-score is the number
of standard deviations above or below the mean

How do we "standardize?"

$$z = (x - \bar{x})/SD$$

How do we "un standardize?"

Do the reverse operation

$$x = \bar{x} + z * SD$$

An example

The distribution of the SAT scores for the University of Washington was roughly normal in shape, with mean 1055 and standard deviation 200.

1. What percentage of scores were 920 or below?

2. What SAT score separates the lowest 25% of the SAT scores from the rest?

YOU TRY

An example

Mean = $\bar{x}$= 1055 SD = 200.

1. What percentage of scores were 920 or below?

$z = (x - 1055)/200$

The x we care about is 920

$z = (920 - 1055)/200 = -0.675$

Check on table ~ about 25%

An example

Does this also answer question 2?

2. What SAT score separates the lowest 25% of the SAT scores from the rest?

z- scores

moving from x to z

Example: Mean is 610 and the SD is 69

What is the z score for a value of 560?

z- scores

Is the z we get from our initial x
(moving from x to z)

Example: Mean is 610 and the SD is 69

What is the z score for a value of 560?

$z = (x-610)/69$

The x we care about is 560

$z = (560-610)/69 = -0.725$

x- scores

Example: Mean is 610 and the SD is 69

What is the x score if the z-score is 1.6?

$x = 610 + z*69$

The z we care about is 1.6

$x = 610 + 1.6*69 = 720$

z-scores are useful for comparisons

	Mean	SD
Heart disease	219	46
Cancer	194	30

Death rates per 100,000 people in the entire US

Alaska: 94 deaths per 100,000 people - heart disease
110 deaths per 100,000 people - cancer

Which is more extreme with respect to the nation?

You calculate the z-scores

	mean	SD
Heart disease	219	46
Cancer	194	30

Death rates per 100,000 people in the entire US

Alaska: 94 deaths per 100,000 people - heart disease
110 deaths per 100,000 people - cancer

Which is more extreme with respect to the nation?

Let's calculate the z-scores

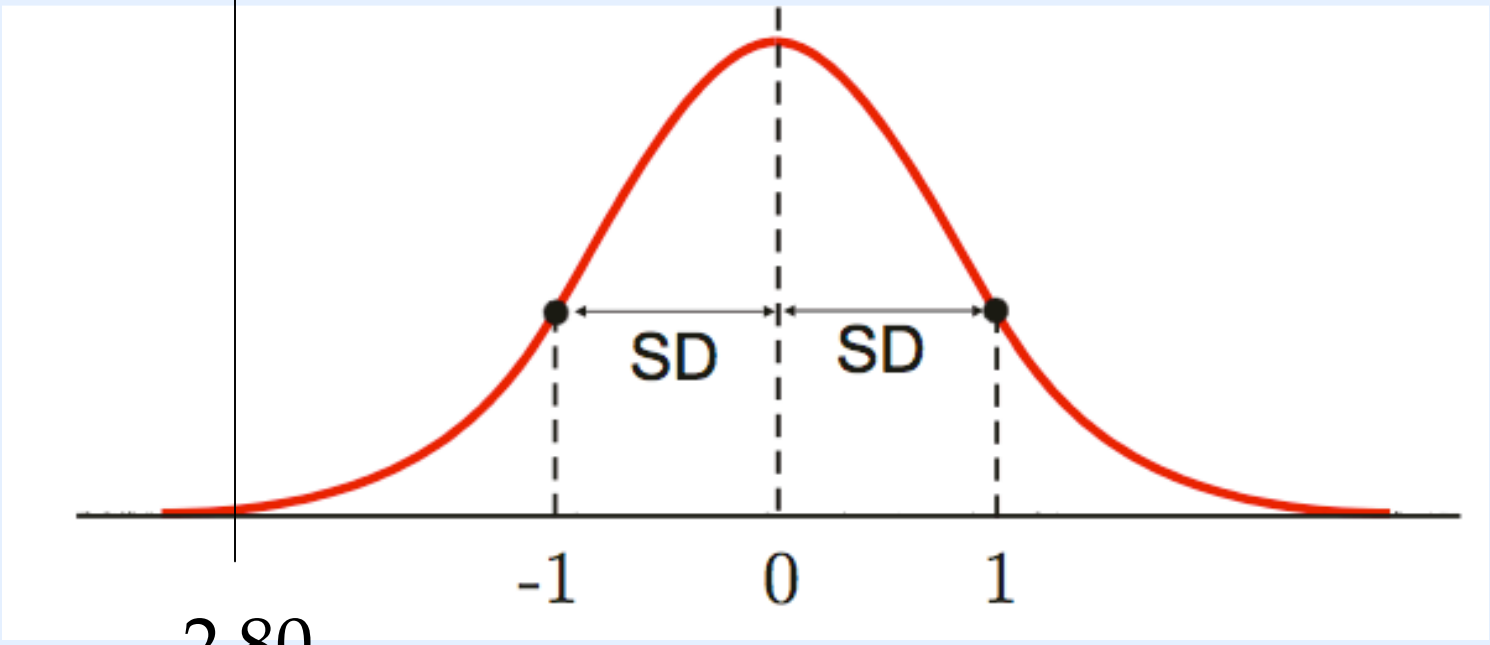
	mean	SD
Heart disease	219	46
Cancer	194	30

Heart disease = (94-219)/46 = -2.72

Cancer = (110-194)/30 = -2.80

They are both very distant from the mean but Death from cancer is slightly more extreme

Cancer z-score in Alaska

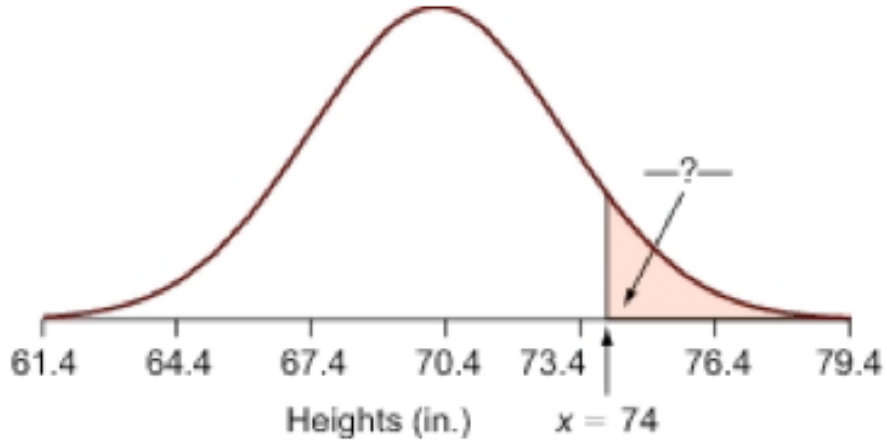


-2.80

How tall are US men?

Heights are normally distributed
Mean = 70.4 inches SD = 3.0 inches

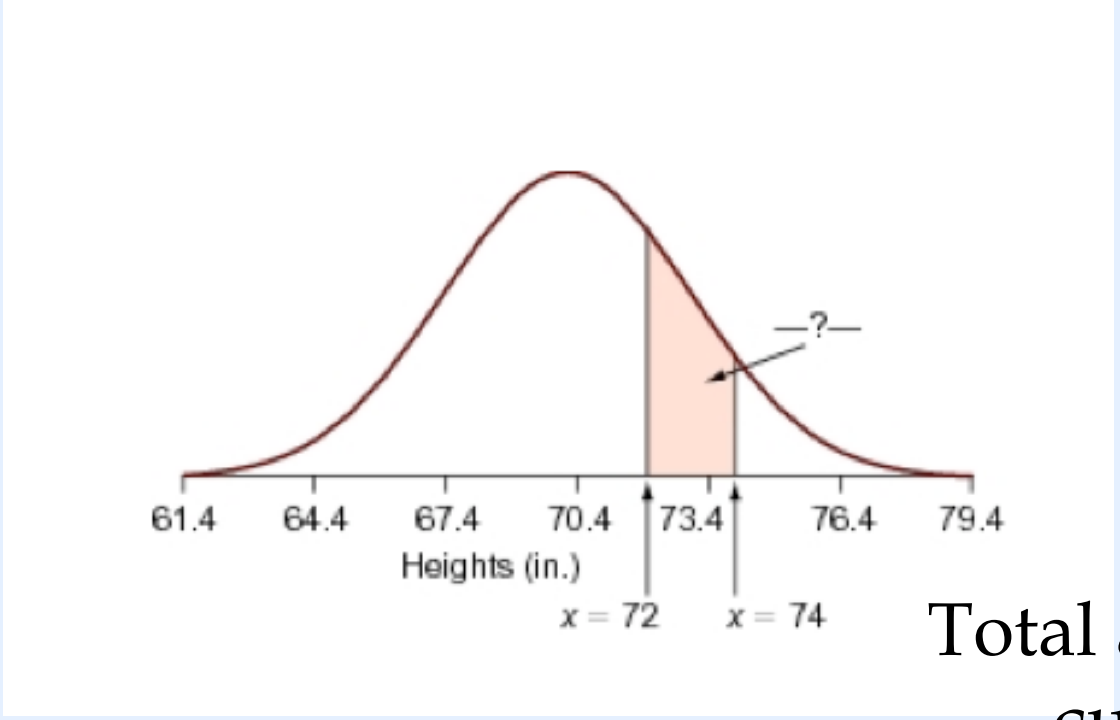
1. What percentage are more than 74 inches tall?



How tall are US men?

Heights are normally distributed
Mean = 70.4 inches SD = 3.0 inches

2. What percentage is between 72 and 74 inches tall?



Total area under curve is 1

How tall are US men?

Heights are normally distributed
Mean = 70.4 inches SD = 3.0 inches

Above 74:

$z = (74-70.4)/3 = 1.20$

Check table

Area to the left is about 0.885.
Area to the right is 1 - 0.885 = 0.115

The percentage is 11.5%

How tall are US men?

Heights are normally distributed
Mean = 70.4 inches SD = 3.0 inches

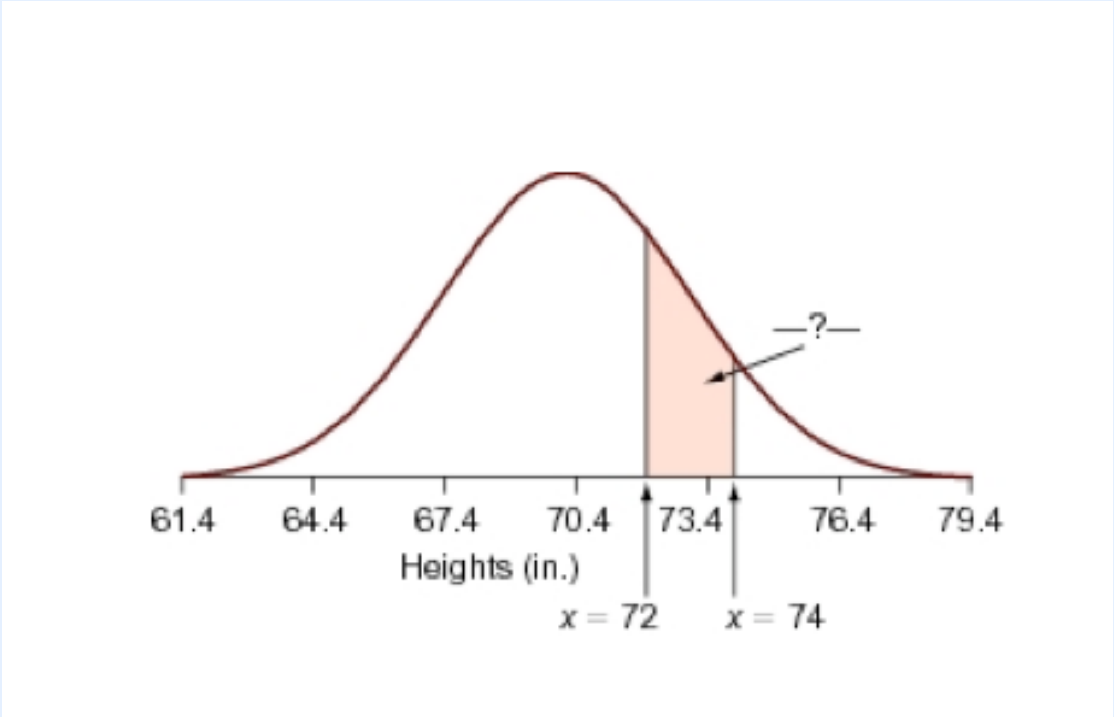
Above 72:

$z = (72-70.4)/3 = 0.53$

Check table

Area to the left is about 0.702.
Area to the right is 1 - 0.702= 0.298

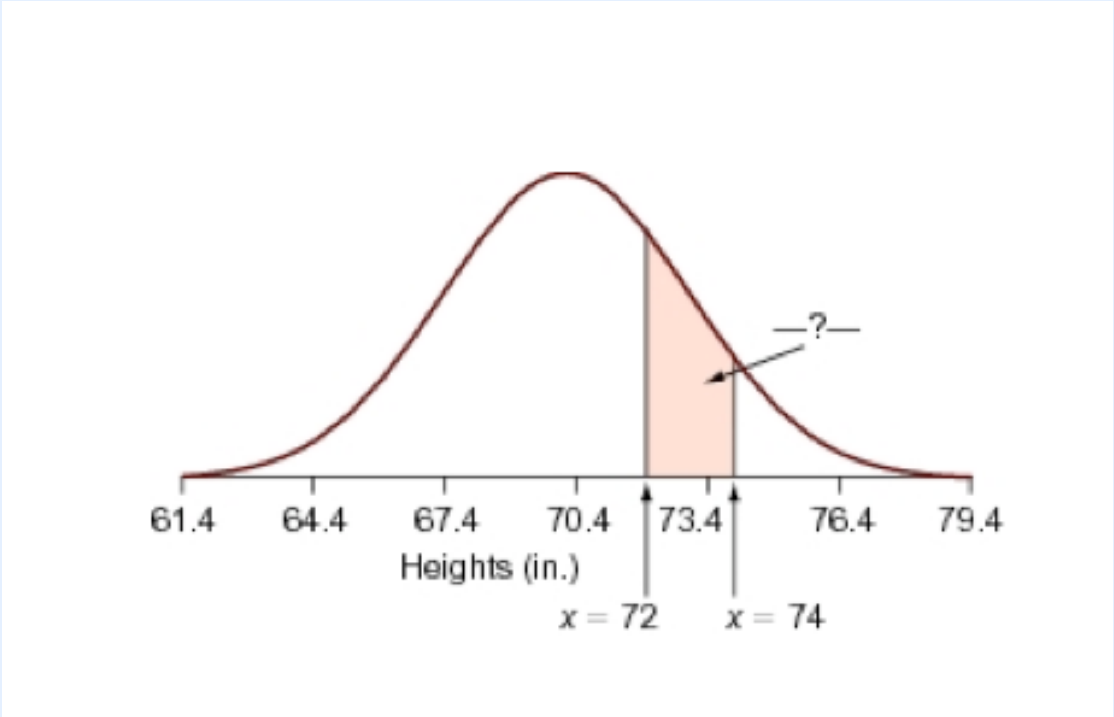
How tall are US men?



Area to the right of 74 is about 0.115
Area to the right is 72 is about 0.298

The difference is 0.298-0.115 = 0.183

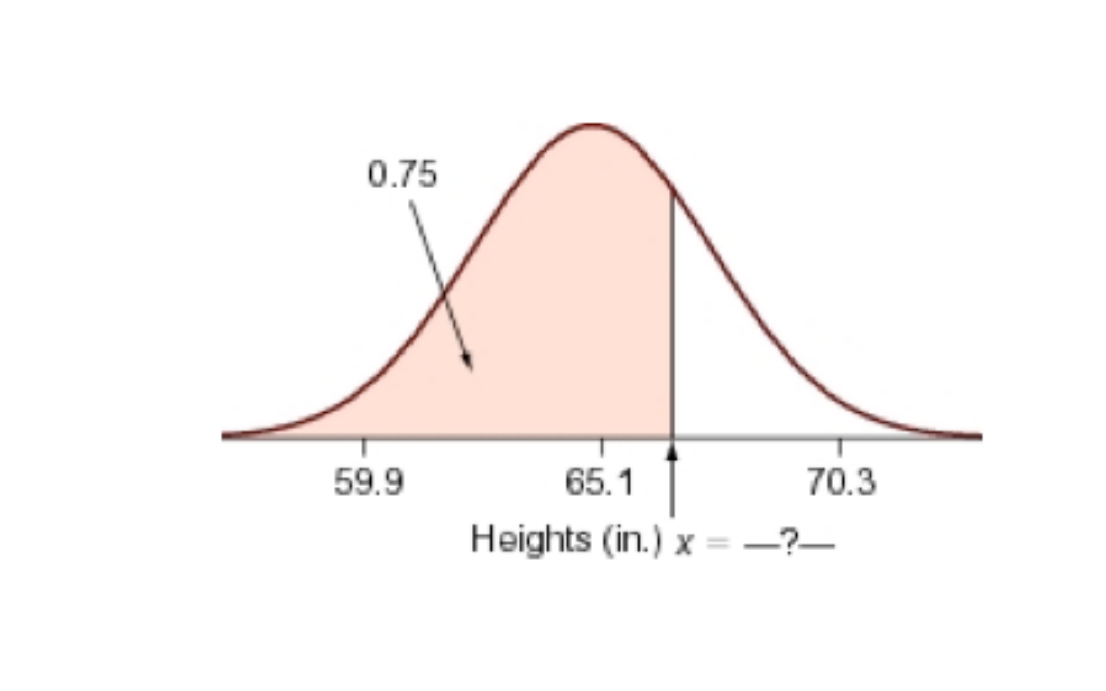
How tall are US men?



The percentage of US men whose height is between 72 and 74 inches is

18.3%

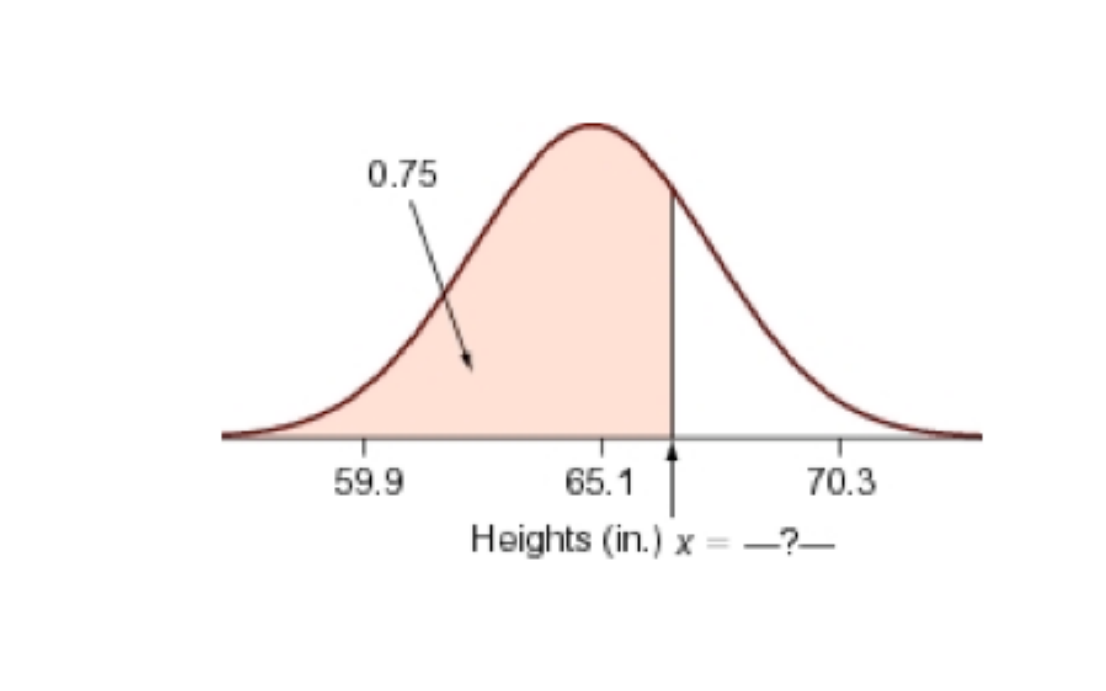
How tall are US women?



Heights are normally distributed
Mean = 65.1 inches SD = 2.6 inches

What value separates the lowest 75%
from the highest 25%?

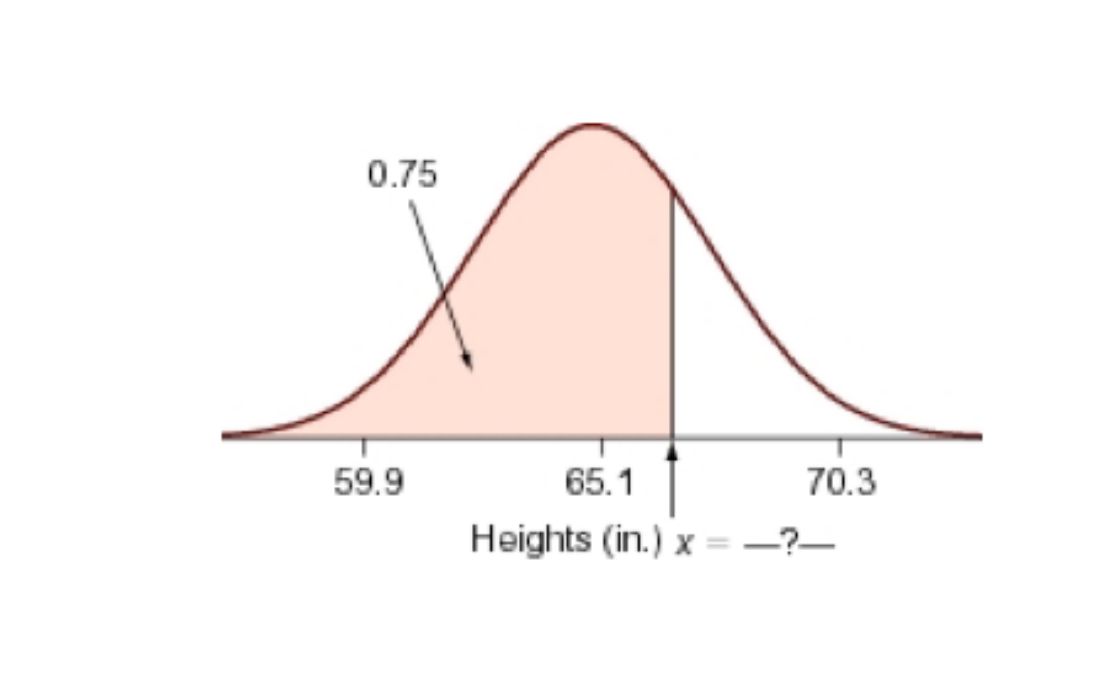
How tall are US women?



Look up the chart for the z-score corresponding to 75%

z=0.67 approximately

How tall are US women?

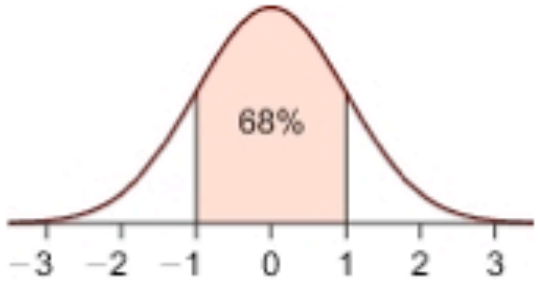


calculate x

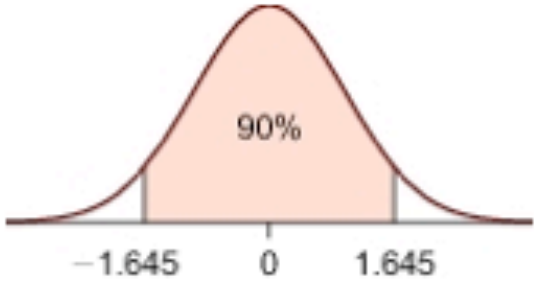
x=65.1 + 0.67 * 2.6 = 66.8 inches

Central Intervals for Normal

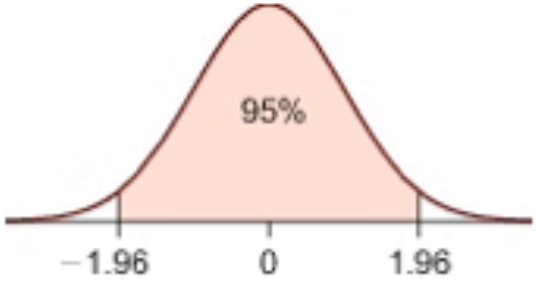
Distributions 68% of the values lie within 1 standard deviation of the mean.



90% of the values lie within 1.645 standard deviations of the mean.



95% of the values lie within 1.96 (or about 2) standard deviations of the mean.

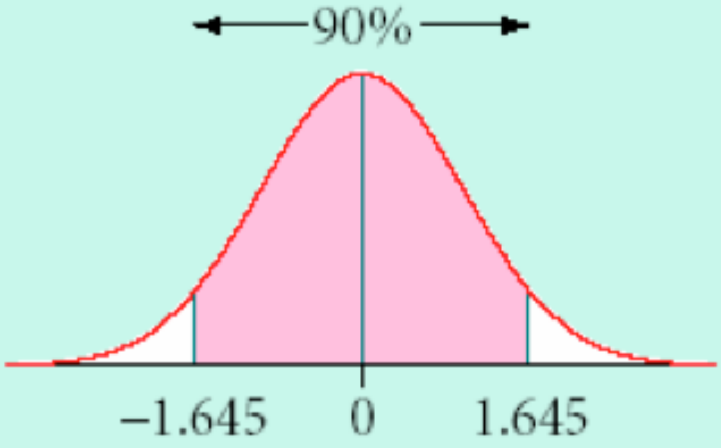


99.7% (or almost all) of the values lie within 3 standard deviations of the mean.

Where are the middle 90% of data?

Cancer: mean =194, SD = 30

90% of the values lie within 1.645 standard deviations of the mean.



90% of data for the standardized distribution are between - 1.645 and 1.645

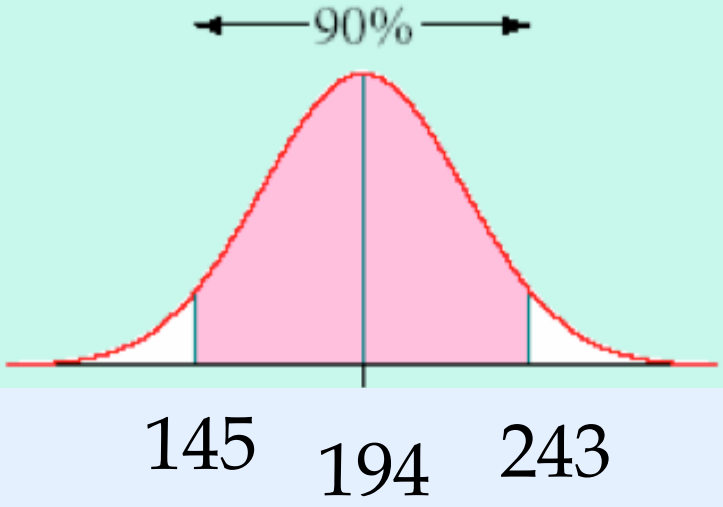
Where are the middle 90% of data?

From z to x

$x = 194 + 1.645*30 = 243$

$x = 194 - 1.645*30 = 145$

Cancer data



Hk

Page 84 E59, E61, E62, E63, E64, E71, E69, E66a, E65