# MAT110 REVIEW

Statistics

NUMERICAL STATISTICS

- The **mean** of a set of observations is the average. Add all the numbers together and divide by how many numbers are in the set.

- The **median** is the midpoint of the data set. Put numbers in order from smallest to biggest and find the center. If there are an even number of observations, take the average of the two numbers in the center.

- The **mode** is the number that appears the most in the data set. It is possible to have more than one mode.

- The **lower quartile (Q_1)** is the median of the lower half (numbers less than the median).

- The **upper quartile (Q_3)** is the median of the upper half (numbers greater than the median).

- The **five number summary** consists of the minimum, lower quartile, median, upper quartile, and maximum.
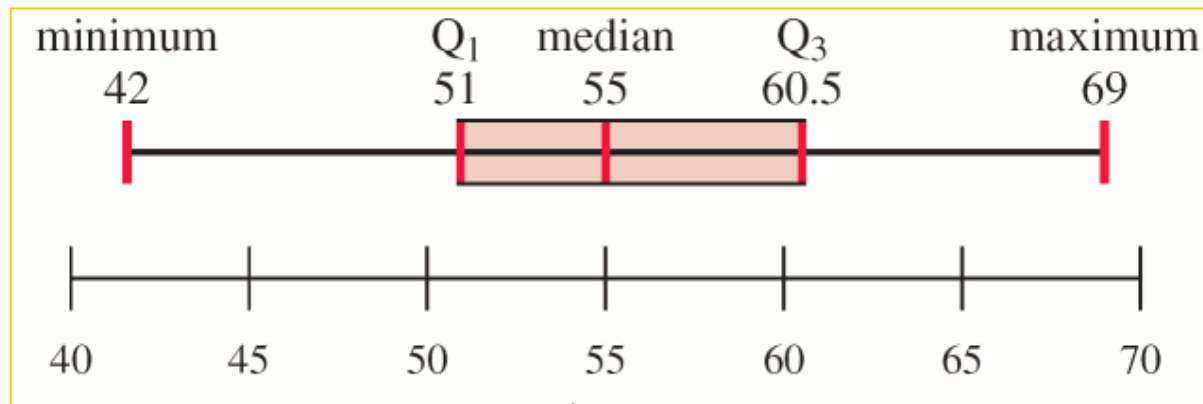
ORGANIZING DATA
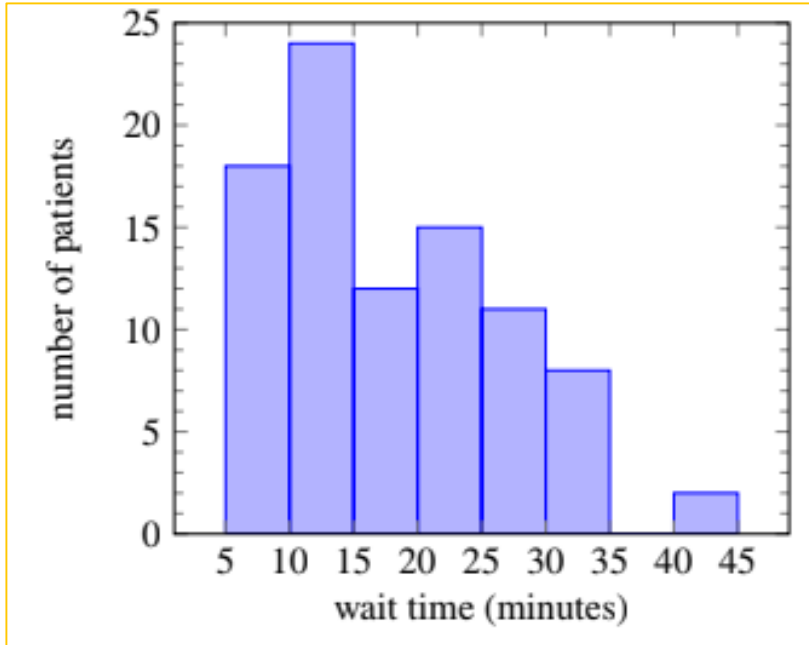
Histogram

Box and Whisker Plot

Stem Plot

Stem	Leaf
0	8 8 9
1	1 2 2 4
2	4 5 8
3	3 9
4	1
5	
6	2 4

EXAMPLE

- Find the mean, median, and mode for the following data:

42, 43, 46, **51, 51, 51,** 52, 54, **55,** 55, 56, 56, 60, 61, 61, 64, 69

- Mean: $\frac{927}{17} = 54.529$

- Median: 55

- Mode: 51

EXAMPLE

- Find the five number summary for the following data:

  42, 43, 46, 51, 51, 51, 52, 54, 55, 55, 56, 56, 60, 61, 61, 64, 69

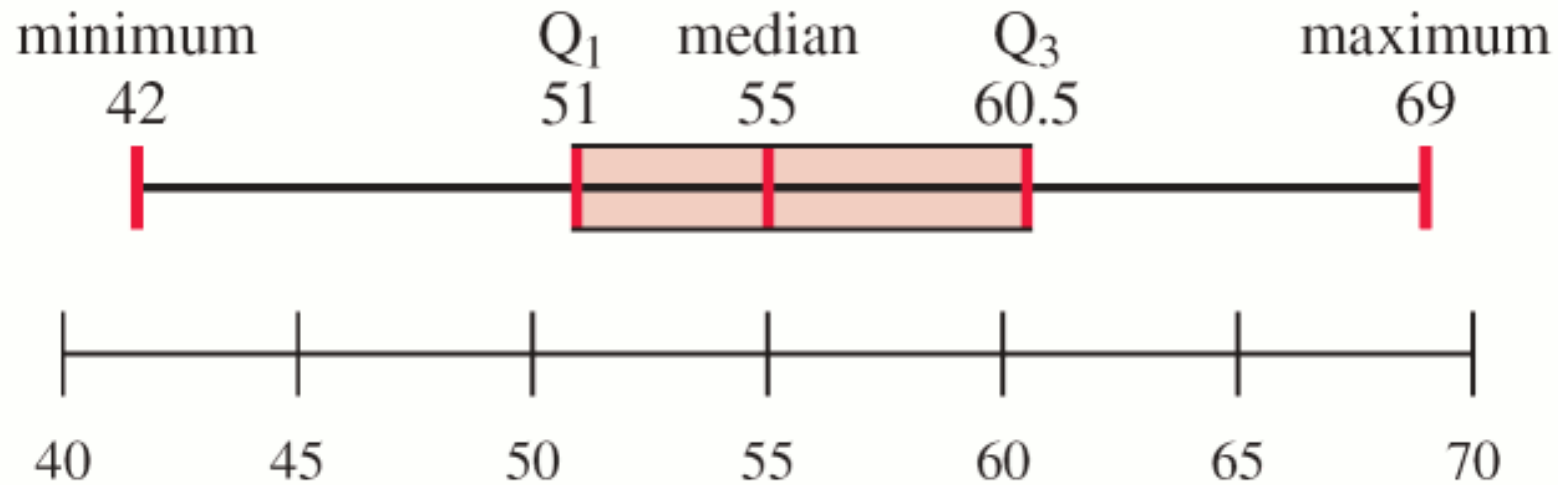- Median: 55

- Lower Quartile: $\frac{51+51}{2} = 51$

- Upper Quartile: $\frac{60+61}{2} = 60.5$

- Minimum: 42

- Maximum: 69

- Five Number Summary: **42, 51, 55, 60.5, 69**

EXAMPLE

▪ Sketch a box and whisker plot for the following data:
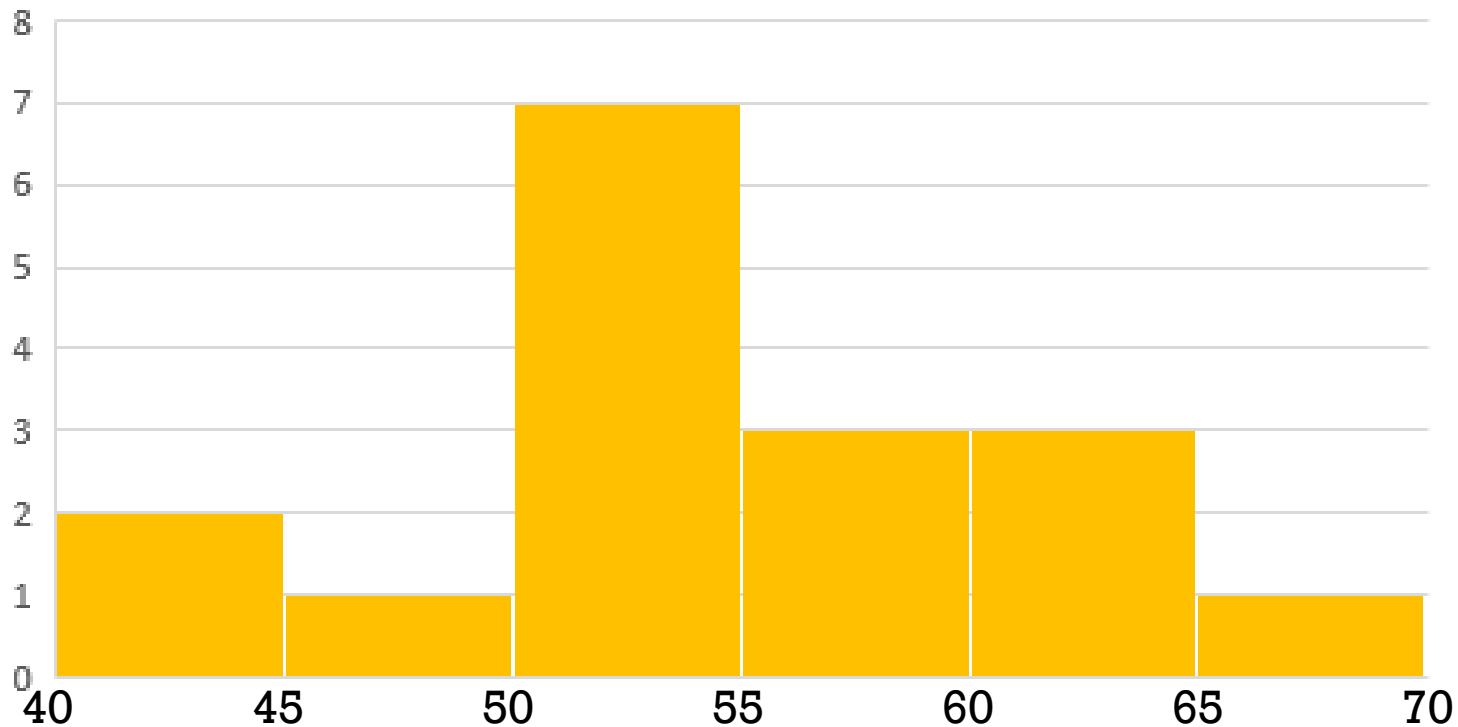
42, 43, 46, 51, 51, 51, 52, 54, 55, 55, 56, 56, 60, 61, 61, 64, 69

EXAMPLE

- Sketch a histogram with bins of size 5 for the following data:

42, 43, 46, 51, 51, 51, 52, 54, 55, 55, 56, 56, 60, 61, 61, 64, 69

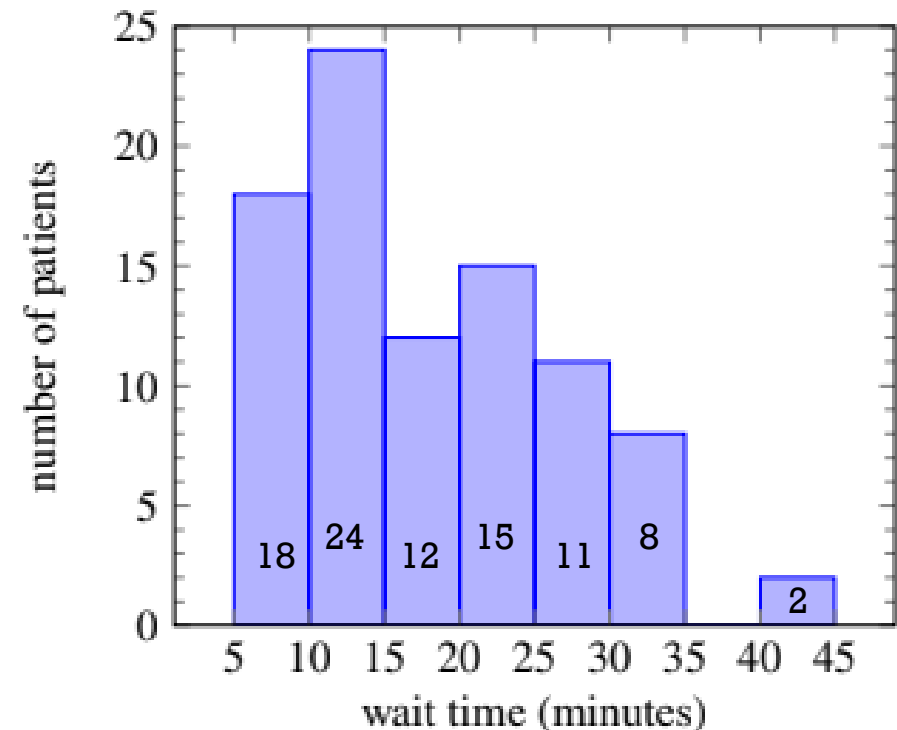| Bins	Frequency
41-45	2
46-50	1
51-55	7
56-60	3
61-65	3
66-70	1

EXAMPLE

- A researcher surveyed 90 patients at a certain hospital. The histogram below gives the length of time patients waited to see a doctor at the hospital. The bins in minutes are 5-9.9, 10-14.9, 15-19.9, …, etc and the vertical axis represents the number of patients.

  - What percent of patients waited more than 20 minutes to see a doctor?

$$\frac{15+11+8+2}{90} = \frac{36}{90}$$

$$= 0.4$$

40% of patients waited more than 20 minutes.

EXAMPLE

- A researcher surveyed 90 patients at a certain hospital. The histogram below gives the length of time patients waited to see a doctor at the hospital. The bins in minutes are 5-9.9, 10-14.9, 15-19.9, …, etc and the vertical axis represents the number of patients.

  - Determine the interval that contains the median wait time.

There are 90 patients total, so the median is between Patient 45 and 46 when they are lined up smallest to biggest.
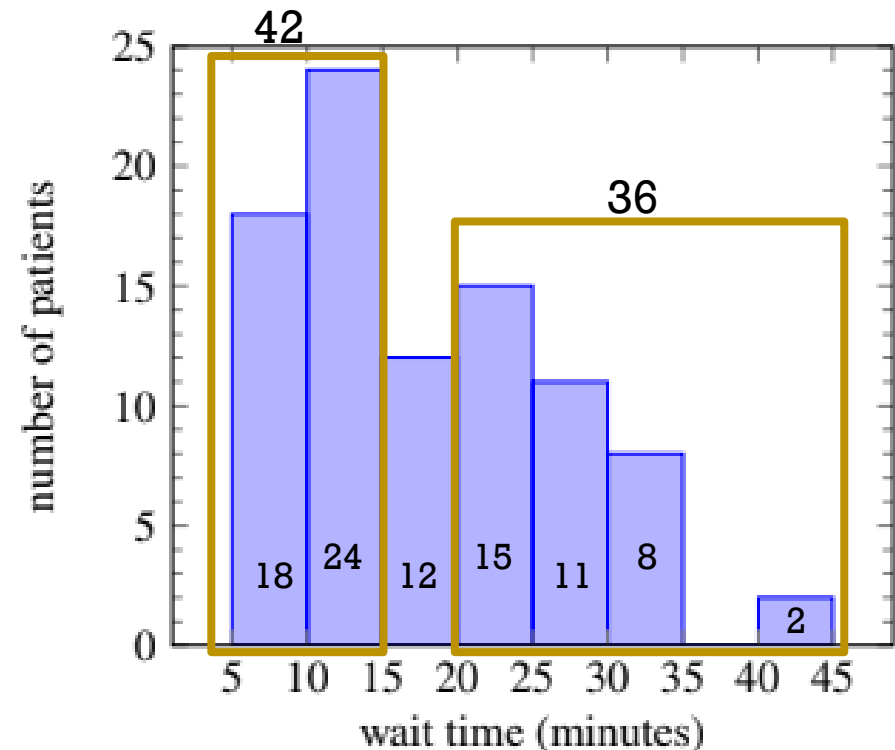
The median is in bin 15-20

EXAMPLE

▪ A researcher surveyed 90 patients at a certain hospital. The histogram below gives the length of time patients waited to see a doctor at the hospital. The bins in minutes are 5-9.9, 10-14.9, 15-19.9, …, etc and the vertical axis represents the number of patients.

▪ Sketch a box and whisker plot to represent the data in the histogram.

▪ Median: 15-20

▪ Minimum: 5-10

▪ Maximum: 40-45

▪ Lower Quartile: 10-15

▪ Upper Quartile: 20-25
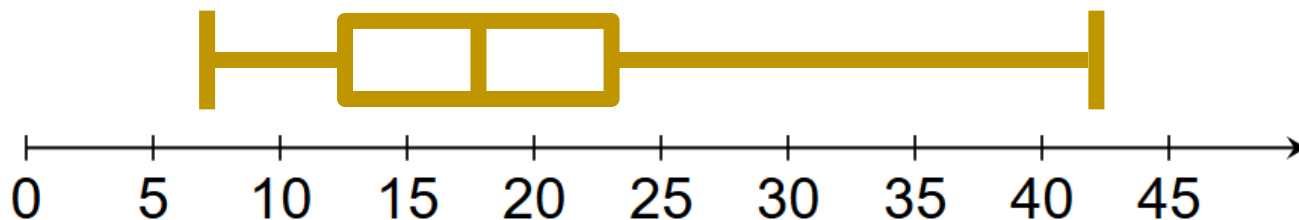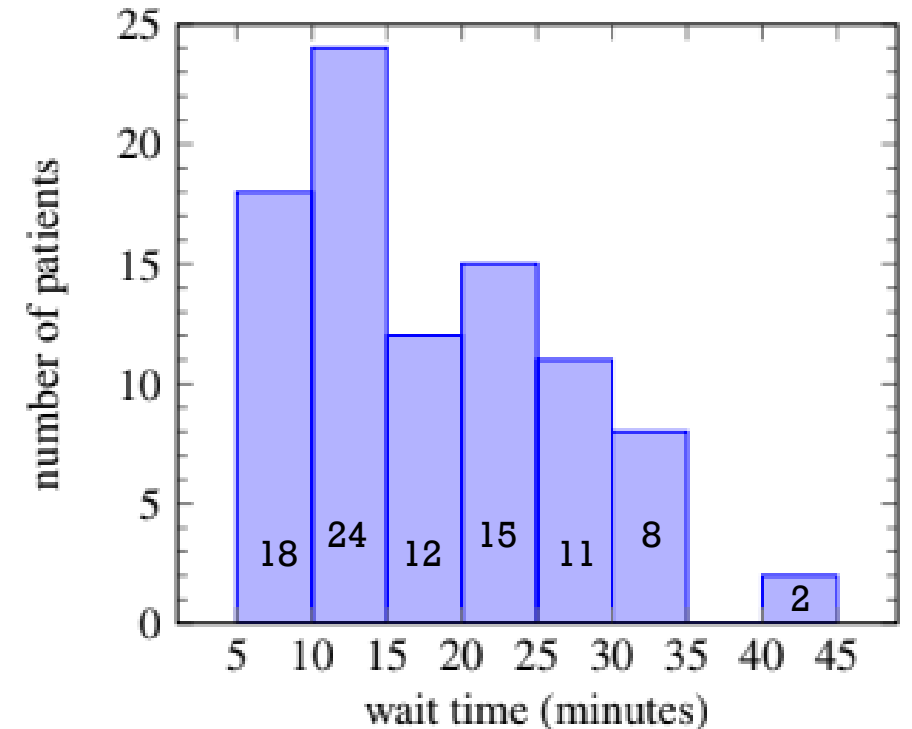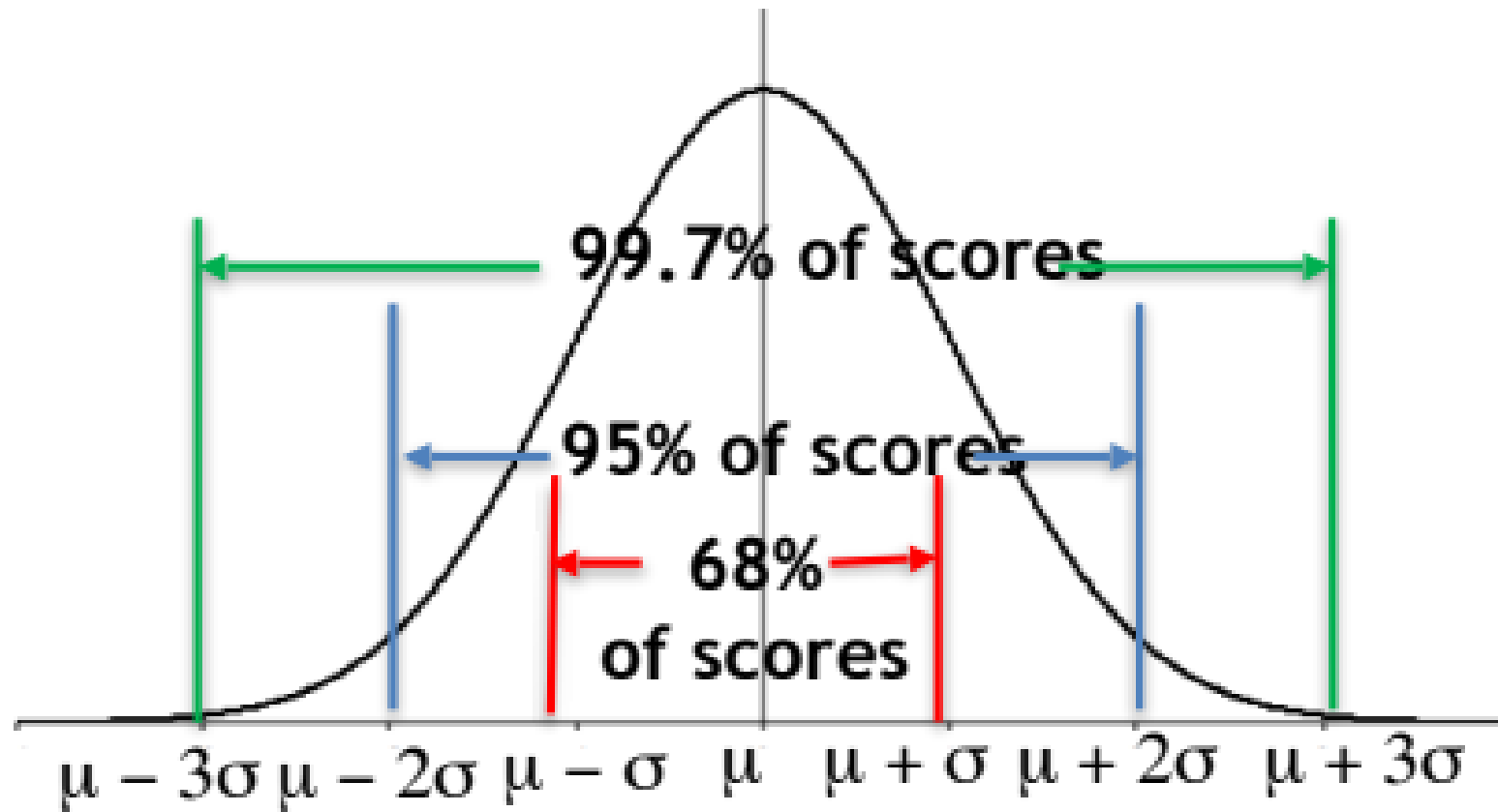
THE NORMAL DISTRIBUTION

- A normal curve is bell shaped.

- The highest point on the curve is at the mean of the distribution.

- The mean and median of the distribution are equal.

- The curve is symmetric with respect to the mean.

- The total area under the curve is 1.

- We represent the population mean by μ and standard deviation by σ
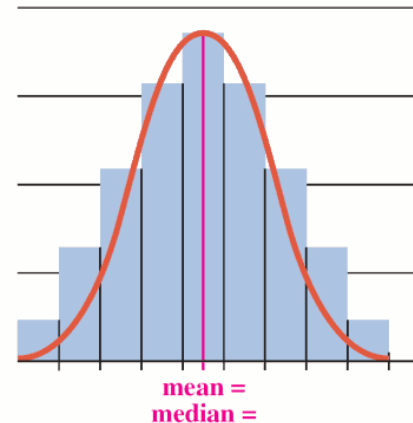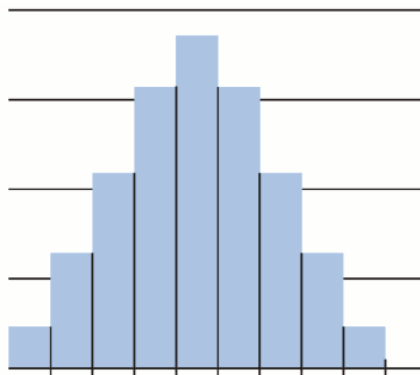
mean =
median =

THE 68-95-99.7 RULE

EXAMPLE

- Suppose the distribution of scores of 1,000 students who take a standardized intelligence test is a approximately normal. If the distribution's mean is $\mu = 450$ and its standard deviation is $\sigma = 25$,
  - Approximately how many scores do we expect to fall between 425 and 475?

  - Note: 425 and 475 are each 1 standard deviation from the mean. Approximately 68% of the scores lie within 1 standard deviation of the mean.

  - We expect about $0.68 \times 1,000 = $ **680** scores to fall in the range 425 to 475.

68%
of scores

mean
450

425 = one
standard deviation
below mean

475 = one
standard deviation
above mean
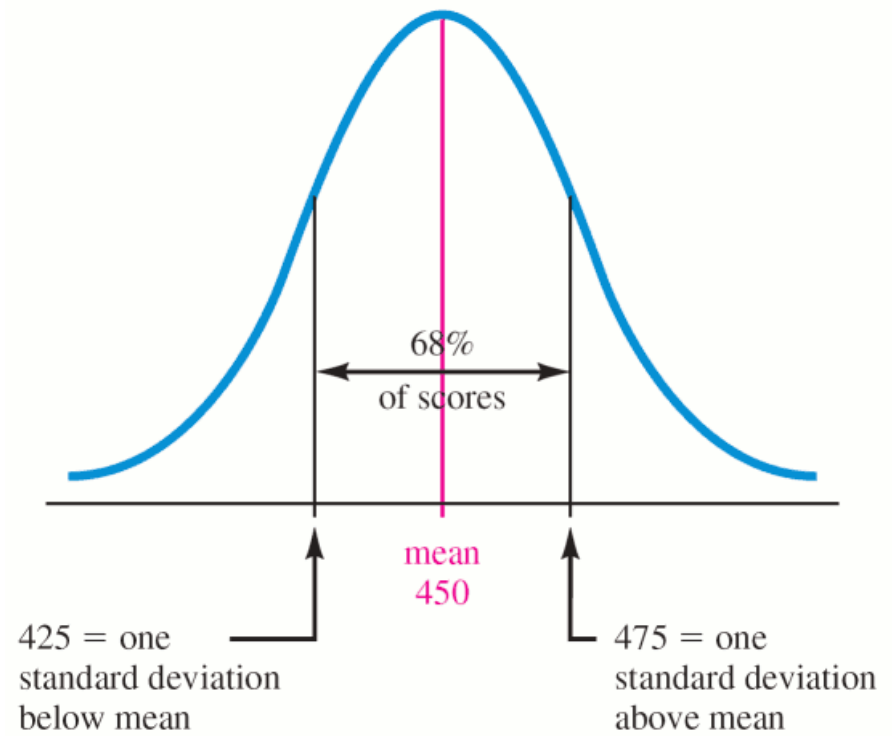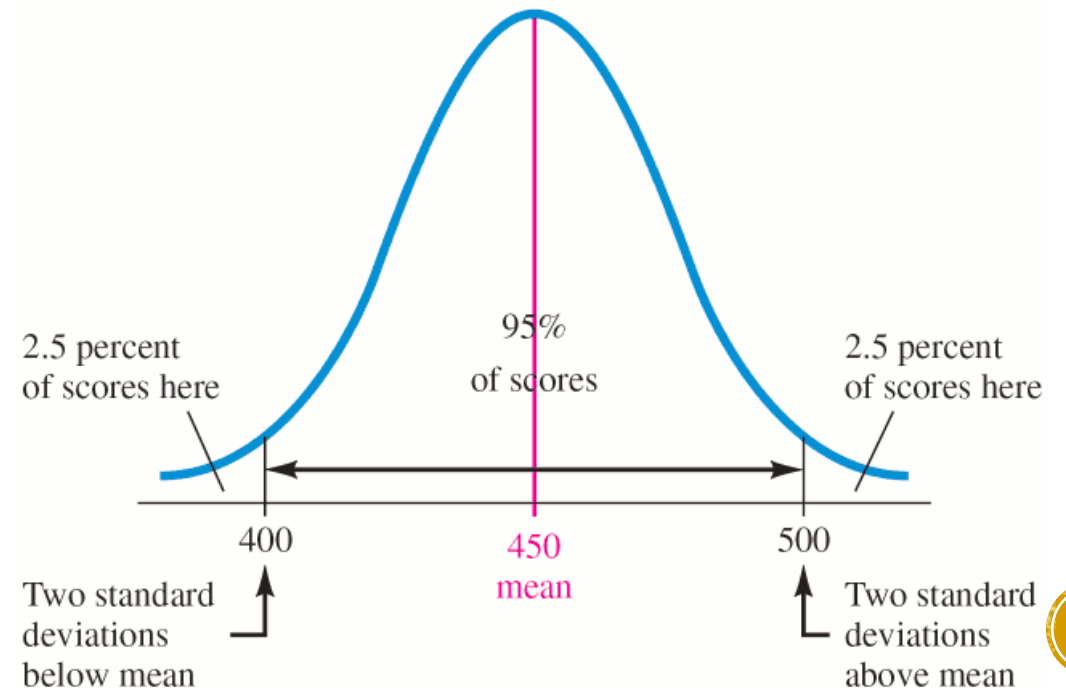
EXAMPLE

- Suppose the distribution of scores of 1,000 students who take a standardized intelligence test is a approximately normal. If the distribution's mean is μ = 450 and its standard deviation is σ = 25,
  - How many scores do we expect to fall above 500?

  - Note: 95% of scores fall between 400 and 500, which leaves 5% of the scores in the tails. Because the normal curve is symmetric, there must be 5% ÷ 2 = 2.5% in either tail.

  - We expect $0.025 \cdot 1000 = $ **25** scores to be above 500.

2.5 percent of scores here

95% of scores

2.5 percent of scores here

400

450 mean

500

Two standard deviations below mean

Two standard deviations above mean

Z-SCORES

- The **standard normal distribution** has a mean of 0 and a standard deviation of 1.

- A **z-score** represents the number of standard deviations a data value is from the mean.
  - For example, for a normal distribution with mean 450 and standard deviation 25, the value 500 is 2 standard deviations above the mean; that is, the value 500 corresponds to a *z*-score of 2.

$$z = \frac{x - \mu}{\sigma}$$

Z-SCORES

- Below is a portion of a table that gives the area under the standard normal curve between the mean and a z-score.
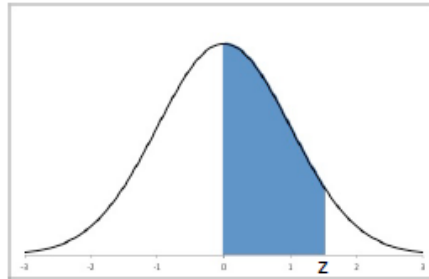
Confidence level C	90%	95%	99%
Critical value z*	1.645	1.960	2.576

$$z = \frac{x - \mu}{\sigma}$$

$$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$$

| Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389

EXAMPLE

- Suppose you take a standardized test. Assume that the distribution of scores is normal and you received a score of 72 on the test, which had a mean of $\mu = 65$ and a standard deviation of $\sigma = 4$. What percentage of those who took this test had a score below yours?

- First, find the z-score for 72.

$$z = \frac{72 - 65}{4} = 1.75$$

- Then use the z-score table to find the area.

- $0.5 + 0.4599 = 0.96$

96% of scores are below yours.
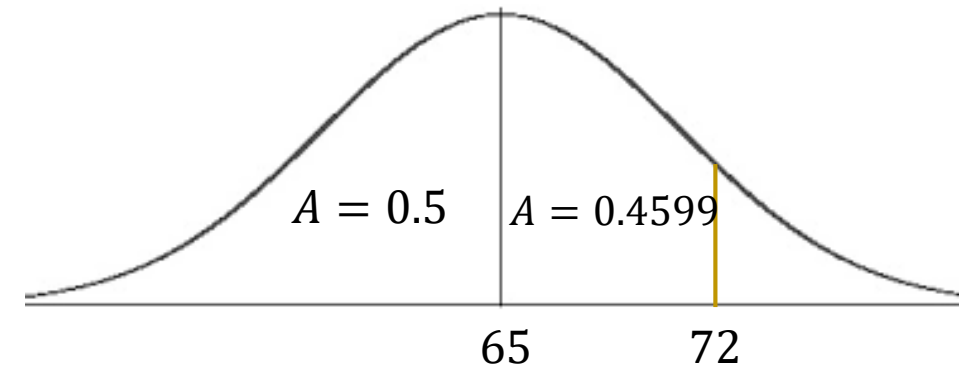
$A = 0.5$ $A = 0.4599$

65   72

EXAMPLE

- The scores of students on an exam are normally distributed with a mean of 516 and a standard deviation of 36.
  - Find the upper and lower quartiles for this exam.
  - Note: There's 25% of the data between the mean and each quartile.

  - First, find the closest area to 0.25 to determine the z-score.

  - The closest area to 0.25 is 0.2486, which has a z-score of 0.67.

  - *The lower quartile will have a negative z-score, and the upper quartile will have a positive z-score.*
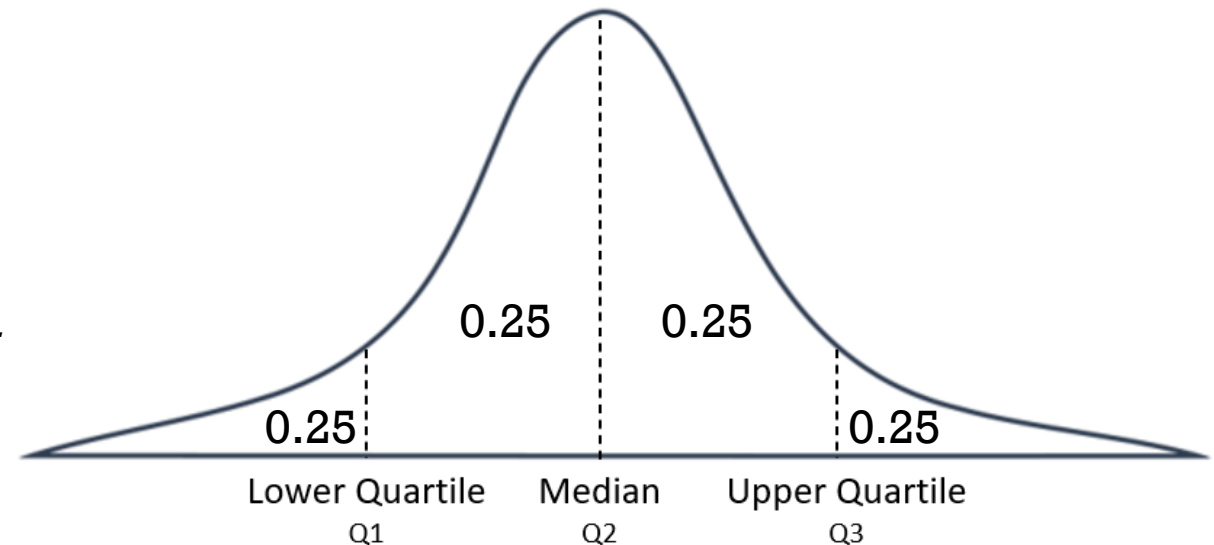
0.25   0.25

0.25   0.25

Lower Quartile   Median   Upper Quartile

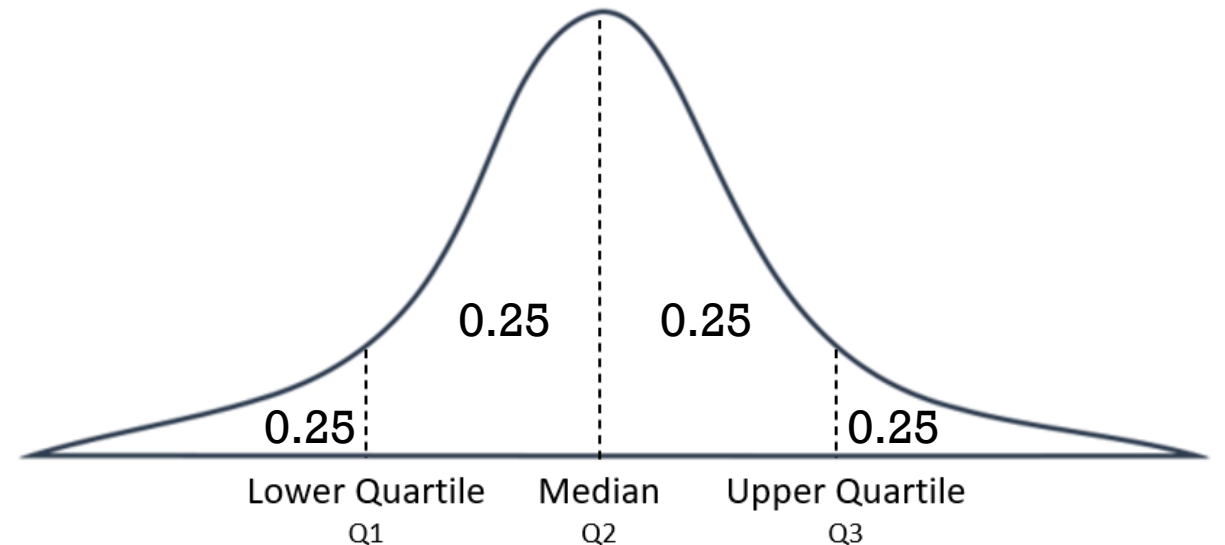Q1   Q2   Q3

EXAMPLE

- The scores of students on an exam are normally distributed with a mean of 516 and a standard deviation of 36.
  - Find the upper and lower quartiles for this exam.
  - Note: There's 25% of the data between the mean and each quartile.

Lower Quartile: $-0.67 = \frac{x-516}{36}$
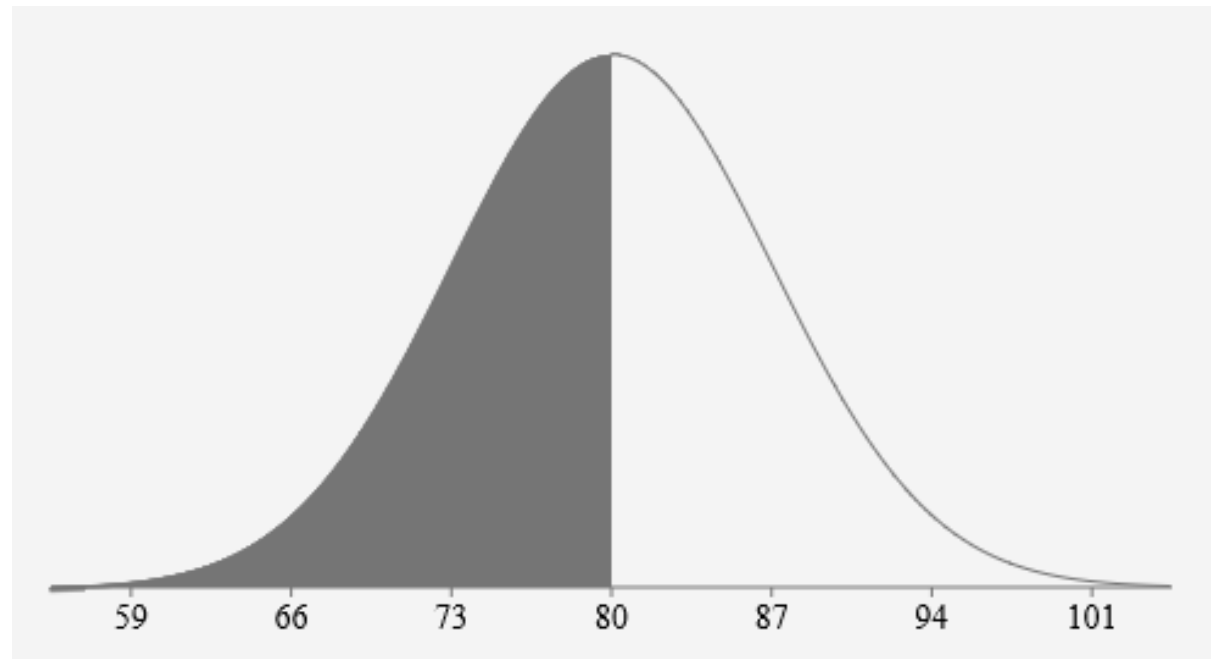
So $Q_1 = 491.88$

Upper Quartile: $0.67 = \frac{x-516}{36}$

So $Q_3 = 540.12$

EXAMPLE

- A normal distribution has mean $\mu = 80$ and standard deviation $\sigma = 7$.
  - What percent of observations are below 80?

50%

EXAMPLE

- A normal distribution has mean $\mu = 80$ and standard deviation $\sigma = 7$.
  - What percent of observations are between 72 and 91?

  - For 72:

    $z = \frac{72-80}{7} = -1.14$

    Area: 0.3729

  - For 91:

    $z = \frac{91-80}{7} = 1.57$

    Area: 0.4418

  Total area: $0.3729 + 0.4418 = 0.8147$

  81.47%

EXAMPLE

- A normal distribution has mean $\mu = 80$ and standard deviation $\sigma = 7$.
  - What percent of observations are between 63 and 72?

- For 63:

$z = \dfrac{63-80}{7} = -2.43$

Area: 0.4925

- For 72:

$z = \dfrac{72-80}{7} = -1.14$

Area: 0.3729

Total area: $0.4925 - 0.3729 = 0.1196$

11.96%

EXAMPLE

- A normal distribution has mean $\mu = 80$ and standard deviation $\sigma = 7$.
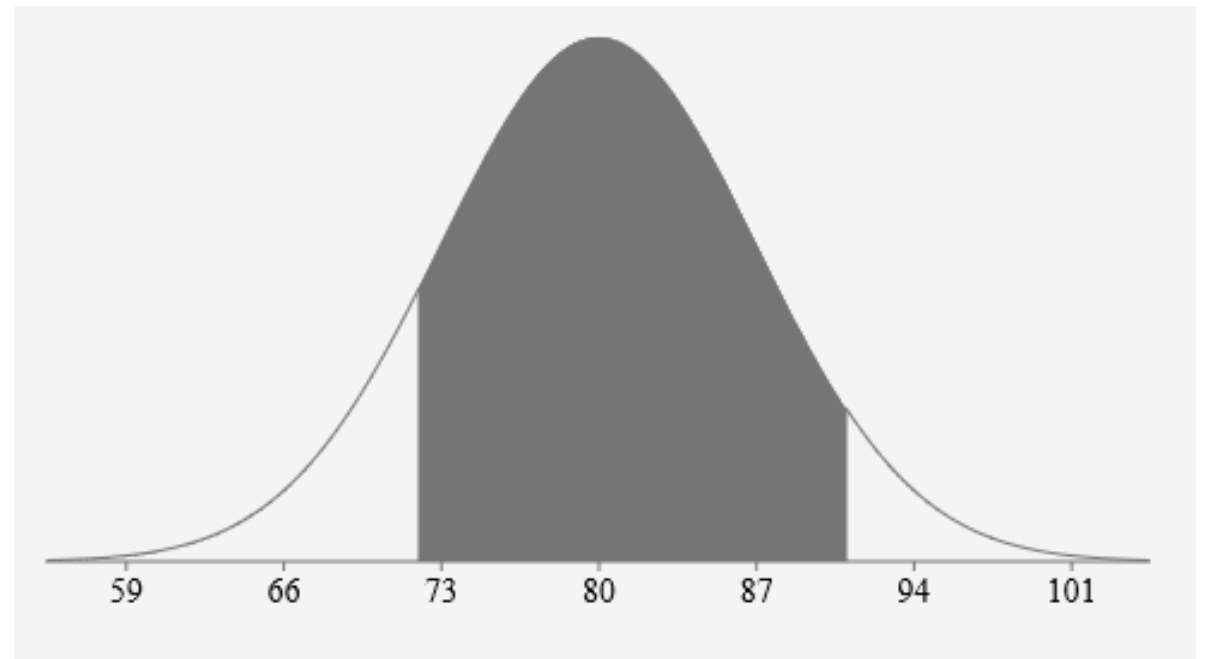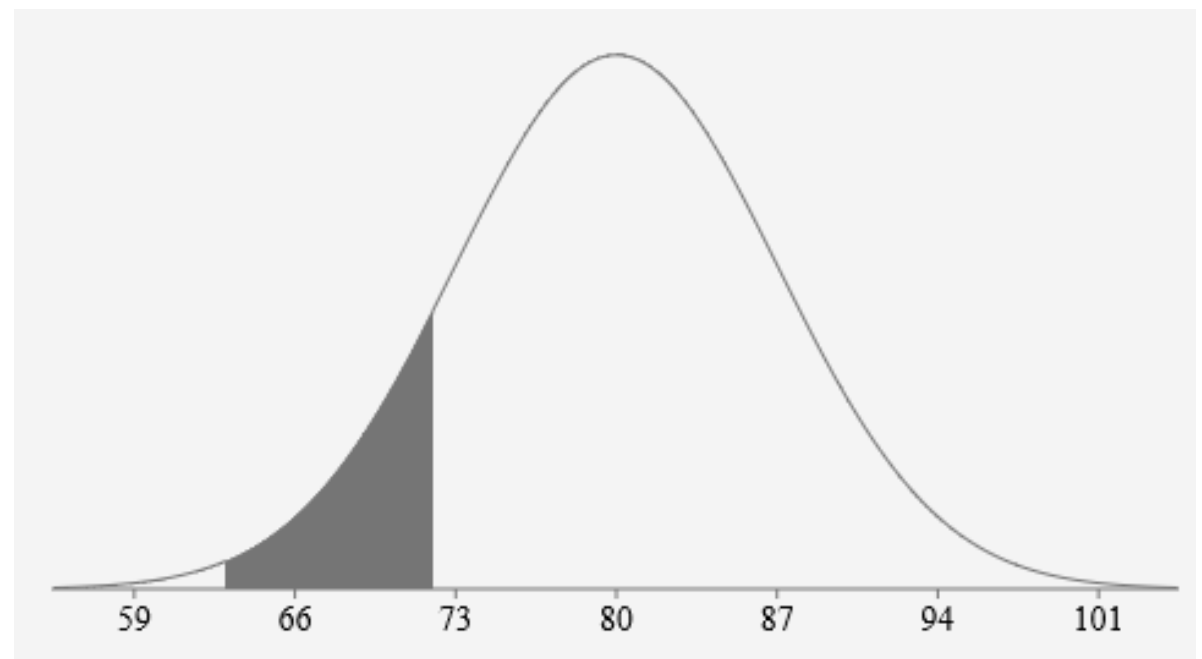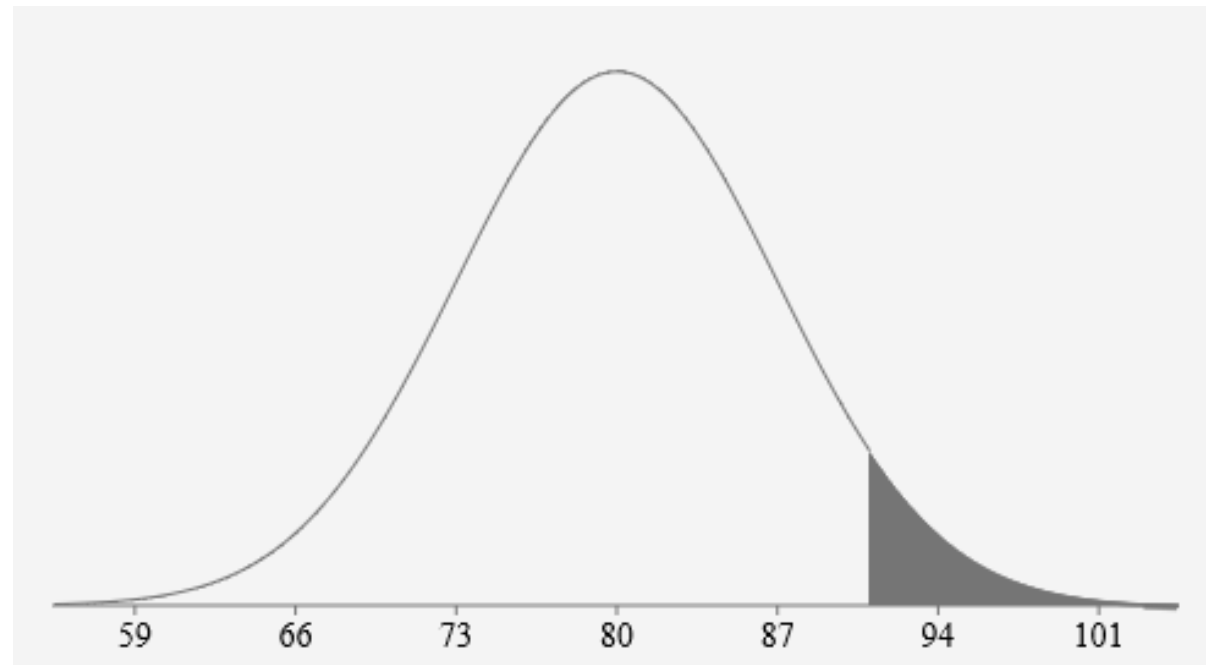 - What percent of observations are above 91?

  - For 91:

  $z = \frac{91-80}{7} = 1.57$

  Area: 0.4418

  Total area: $0.5 - 0.4418 = 0.0582$

  5.82%

EXAMPLE

▪ A normal distribution has mean $\mu = 80$ and standard deviation $\sigma = 7$.

  ▪ What percent of observations are below 89?

  ▪ For 89:

  $z = \frac{89-80}{7} = 1.29$

  Area: 0.4015

  Total area: $0.5 + 0.4015 = 0.9015$

  90.15%
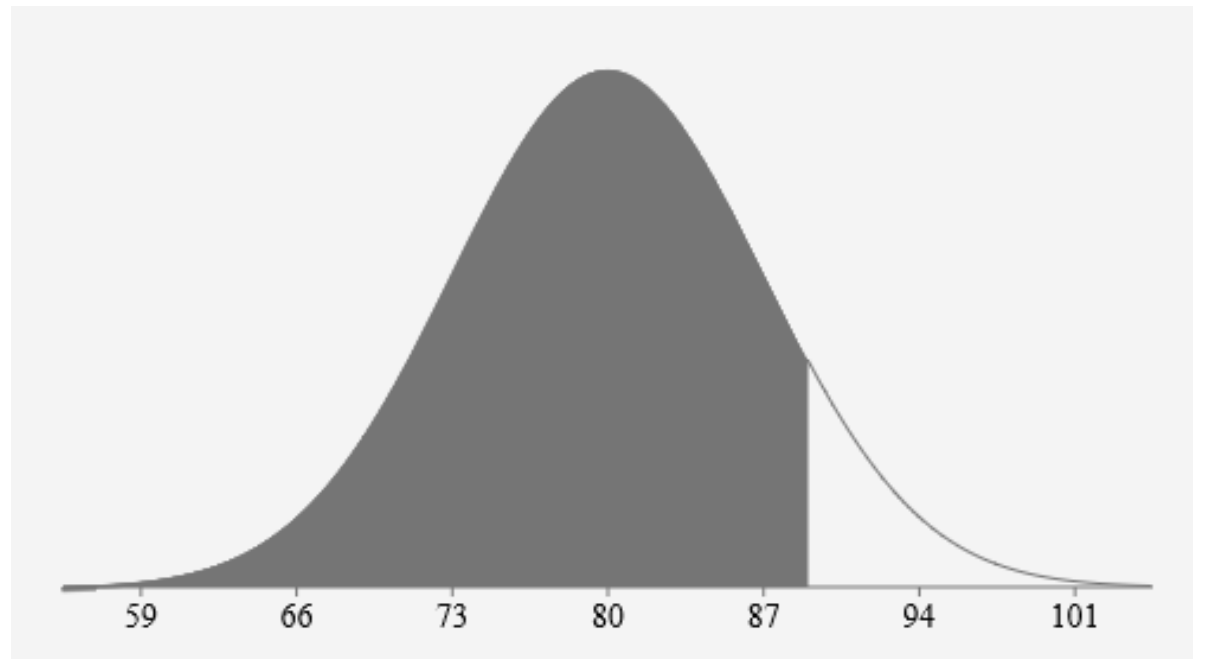
CONFIDENCE INTERVALS

- A **Confidence Interval** gives a range of plausible values for the population mean.

- The formula for a confidence interval is: $\bar{x} \pm z^* \dfrac{\sigma}{\sqrt{n}}$

  - Where $\bar{x}$ is the sample mean, and $z^* \dfrac{\sigma}{\sqrt{n}}$ is the **margin of error.**

- *To find the critical value, z^* for a level C confidence interal:*

  - *Write C as a decimal and find $\dfrac{C}{2}$ This is the area between $z = 0$ and z^* on the standard normal curve.*

  - *Find the z-score in your table with the area closest to $\dfrac{C}{2}$. This is your z^*.*

EXAMPLE

- Scores on a standardized intelligence test are known to follow a normal distribution with a standard deviation $\sigma = 10$. A researcher finds a random sample of 100 students test scores has a sample mean of $\bar{x} = 90$.
  - Use this to find a 95% confidence interval for the mean of all students taking the test.

  - The z^* for a 95% confidence level is given: $z^* = 1.96$.
  - The left end of the interval is: $90 - 1.96 \cdot \frac{10}{\sqrt{100}} = 88.04$
  - The right end of the interval is: $90 + 1.96 \cdot \frac{10}{\sqrt{100}} = 91.96$

- So the 95% confidence interval is $(\mathbf{88.04}, \mathbf{91.96})$, OR we are 95% confident the *population* mean is between 88.04 and 91.96

EXAMPLE

- Scores on a standardized exam are known to follow a normal distribution with a standard deviation of $\sigma = 6.$ A researcher finds a random sample of 26 exam scores has a mean $\bar{x} = 81.$

  - Use this to find a 92% confidence interval for the mean exam score of all students.

  - First find z^*.

  $\frac{0.92}{2} = 0.46$

  The closest area to 0.46 is 0.4599, so $z^* = 1.75$.

  - Then use the confidence interval formula: $81 \pm 1.75 \frac{6}{\sqrt{26}}$

- The 92% confidence interval is $(\mathbf{78.941, 83.059}).$

EXAMPLE

- Scores on a standardized exam are known to follow a normal distribution. A researcher estimates the mean score on a standardized exam to be between 83 and 92 with a 95% confidence interval. What is the margin of error?

  - *Remember the confidence interval is found by adding or subtracting the margin of error to/from the mean.*

  So $83 + m = \bar{x}$ and $92 - m = \bar{x}$

  You can also say $\frac{92-83}{2} = m$

So the Margin of Error is **4.5**.

EXAMPLE

- The heights of Christmas trees are known to follow a normal distribution with a standard deviation of 8 inches. A researcher wants to estimate the mean height of all Christmas trees. If she wants to estimate this with an 81% confidence interval, what size sample should she use to estimate the mean tree height to within $\pm$ 5 inches.

  - If we want the mean tree height to be within 5 inches, then the margin of error is 5. We will use $z^* \dfrac{\sigma}{\sqrt{n}} = m$

  - First, find z^*

  $\dfrac{.81}{2} = .405$

  The closest area to 0.405 is 0.4049, so $z^* = 1.31$

EXAMPLE

- The heights of Christmas trees are known to follow a normal distribution with a standard deviation of 8 inches. A researcher wants to estimate the mean height of all Christmas trees. If she wants to estimate this with an 81% confidence interval, what size sample should she use to estimate the mean tree height to within ± 5 inches.

    - Using the margin of error formula, $m = z^* \cdot \dfrac{\sigma}{\sqrt{n}}$ and our critical value $z^* = 1.31$

$$5 = 1.31 \cdot \frac{8}{\sqrt{n}}$$

$$5\sqrt{n} = 1.31 \cdot 8$$

$$\sqrt{n} = \frac{1.31 \cdot 8}{5}$$

$$n = \left(\frac{1.31 \cdot 8}{5}\right)^2$$

$$n = 4.39$$

We cannot have 0.39 of a tree, so we will round our answer to 5.

Note: Rounding down will cause our margin of error to be greater than 5.

We need a sample of **5** trees.

EXAMPLE

- The absorption time of a new drug is known to follow a normal distribution with a standard deviation of $\sigma = 17$ minutes. A researcher wants to estimate the mean absorption time with a 91% confidence interval. How large a sample does she need to estimate the mean absorption time to within 8 minutes?

  - If we want the mean absorption time to be within 8 minutes, then the margin of error is 8. We will use $z^* \dfrac{\sigma}{\sqrt{n}} = m$

  - First, find z^*

  $\dfrac{.91}{2} = .455$

  The closest area to 0.455 is 0.4554, so $z^* = 1.70$

EXAMPLE

- The absorption time of a new drug is known to follow a normal distribution with a standard deviation of $\sigma = 17$ minutes. A researcher wants to estimate the mean absorption time with a 91% confidence interval. How large a sample does she need to estimate the mean absorption time to within 8 minutes?

  - Using the margin of error formula, $m = z^* \cdot \dfrac{\sigma}{\sqrt{n}}$ and our critical value $z^* = 1.70$

$$8 = 1.70 \cdot \frac{17}{\sqrt{n}}$$

$$8\sqrt{n} = 1.70 \cdot 17$$

$$\sqrt{n} = \frac{1.70 \cdot 17}{8}$$

$$n = \left(\frac{1.70 \cdot 17}{8}\right)^2$$

$$n = 13.050$$

We need a sample of **14** people.

CENTER FOR ACADEMIC SUPPORT

Hearnes 213

816-271-4524