

AMA 2634

Introduction to Statistics

2024 Fall

Lecture 6

Normal Distributions



Many continuous variables have distributions that are bell-shaped and are called **approximately normally distributed variables**.

The theoretical curve, called the **bell curve** or the **Gaussian distribution**, can be used to study many variables that are not normally distributed but are approximately normal.

The probability density function of the normal distribution is:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where

$e \approx 2.718281828 \dots$ **Euler's number**

$\pi \approx 3.14159265 \dots$

μ = population mean

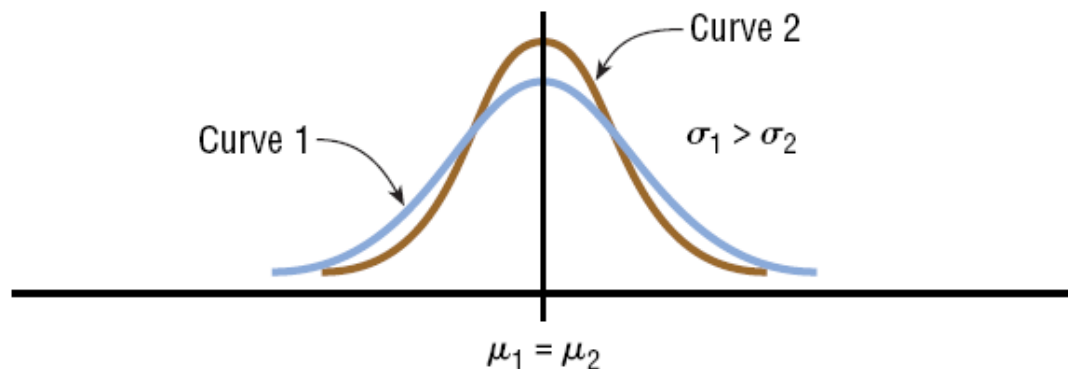
σ = population standard deviation

Normal Distributions



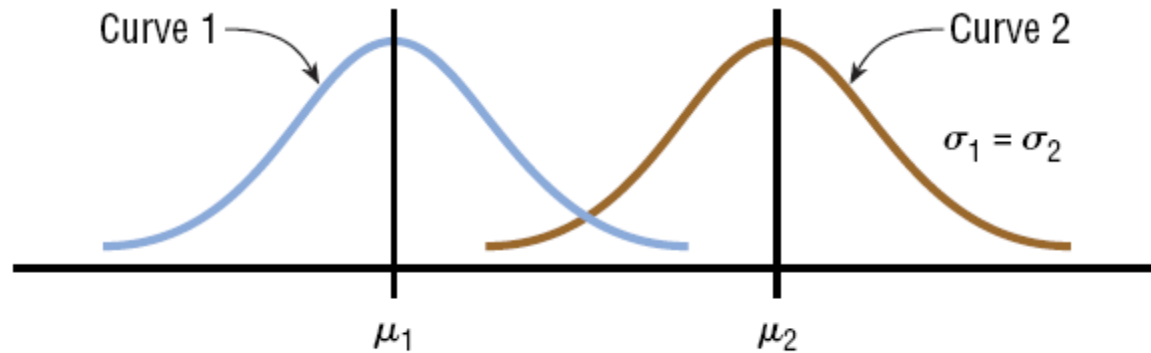
The shape and position of the normal distribution curve depend on two parameters, the **mean** and the **standard deviation**.

Each normally distributed variable has its own normal distribution curve, which depends on the values of the variable's mean and standard deviation.

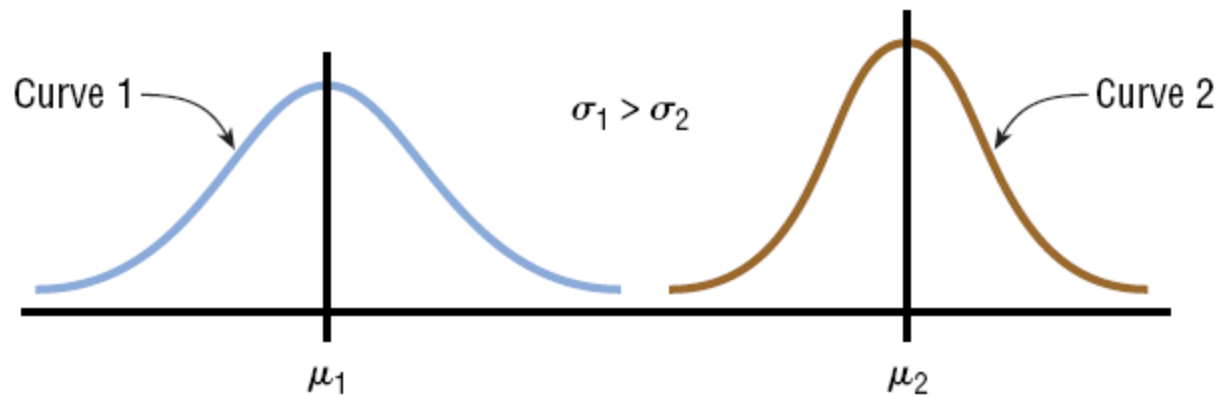


(a) Same means but different standard deviations

Normal Distributions



(b) Different means but same standard deviations



(c) Different means and different standard deviations



The normal distribution curve is bell-shaped.

The mean, median, and mode are equal and located at the center of the distribution.

The normal distribution curve is **unimodal** (i.e., it has only one mode).

The curve is symmetrical about the mean, which is equivalent to saying that its shape is the same on both sides of a vertical line passing through the center.

The density curve is continuous—i.e., there are no gaps or holes. For each value of X , there is a corresponding value of density $f(x)$. For any $x < \infty$,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{(2\sigma^2)}} > 0$$

The curve never touches the x -axis. Theoretically, no matter how far in either direction the curve extends, it never meets the x -axis—but it gets increasingly closer.

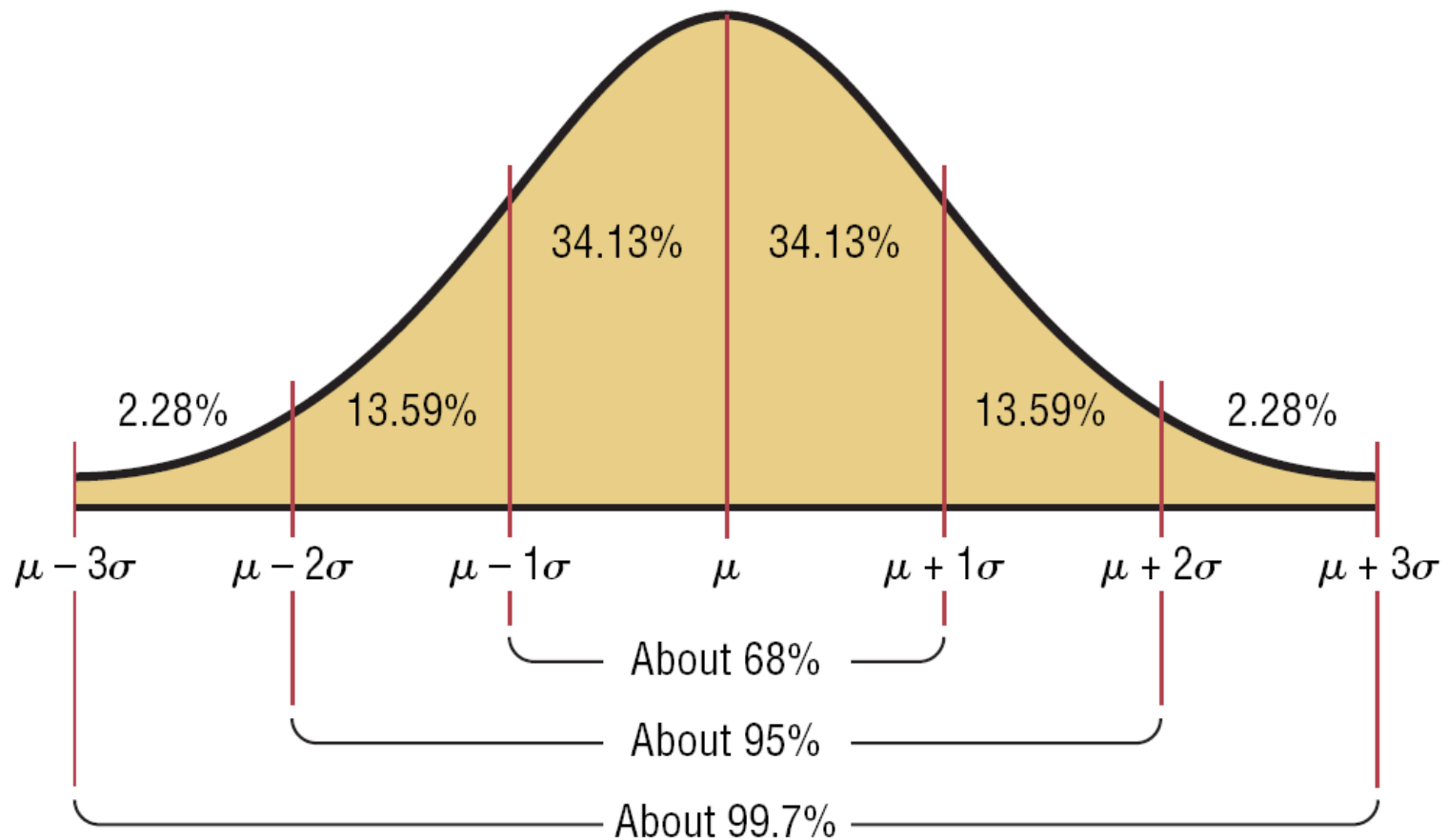
$$\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow \infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{(2\sigma^2)}} = 0$$

The total area under the normal distribution curve is equal to 1.00 or 100%.

The area under the normal curve that lies within

- one standard deviation of the mean is approximately 0.68 (68%).
- two standard deviations of the mean is approximately 0.95 (95%).
- three standard deviations of the mean is approximately 0.997 (99.7%).

Normal Distribution Properties



Why it is “Normal”



Thanks to Mauro Gigli



Imagine that you have to hit a target. You aim at its center but since you aren't a crack shot your throws will departure to some extent from it.

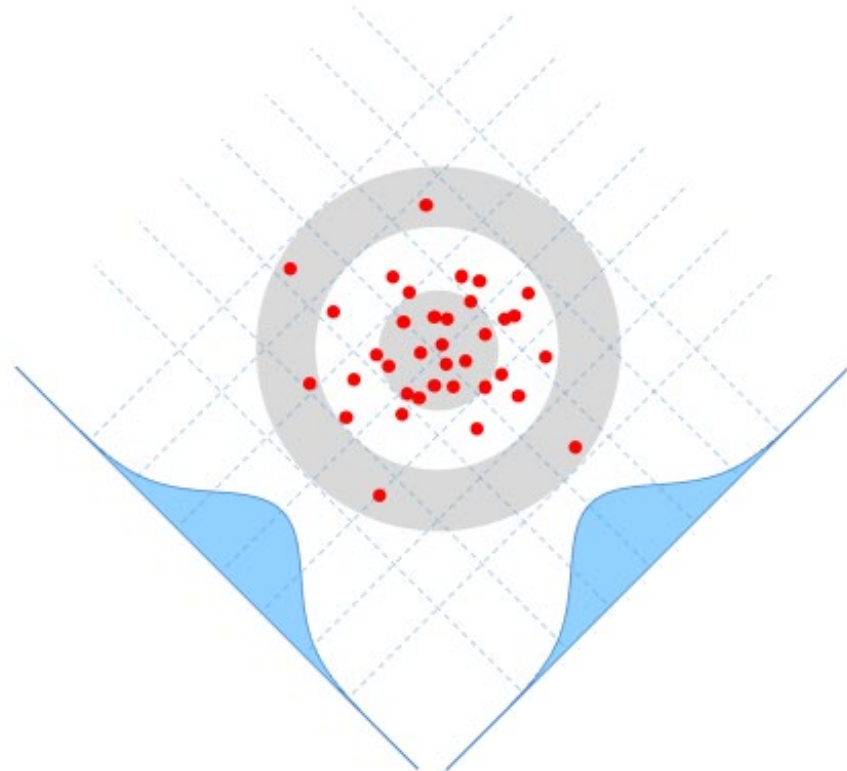
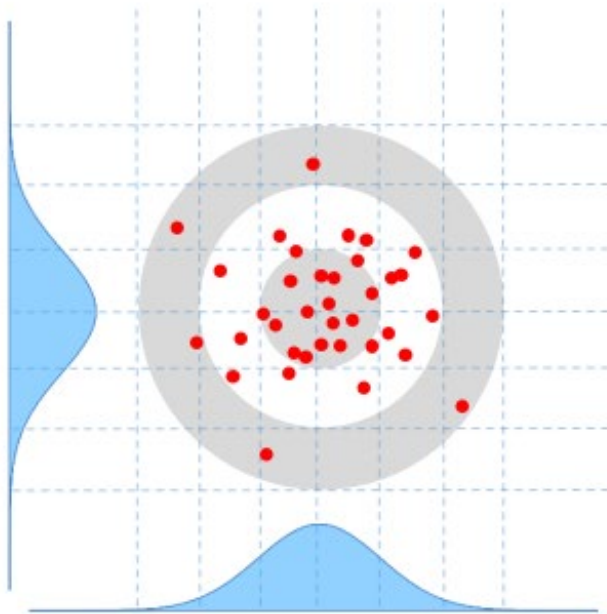
Suppose that:

1. deviates from the center become less likely moving away from it
2. deviates from the center along one direction (for example, horizontally) don't condition ones along the perpendicular direction (vertically)

Example



Then the orthogonal projection of shots on a whatever pair of perpendicular axes (again, for example, left-right and up-down) has **the normal distribution**.



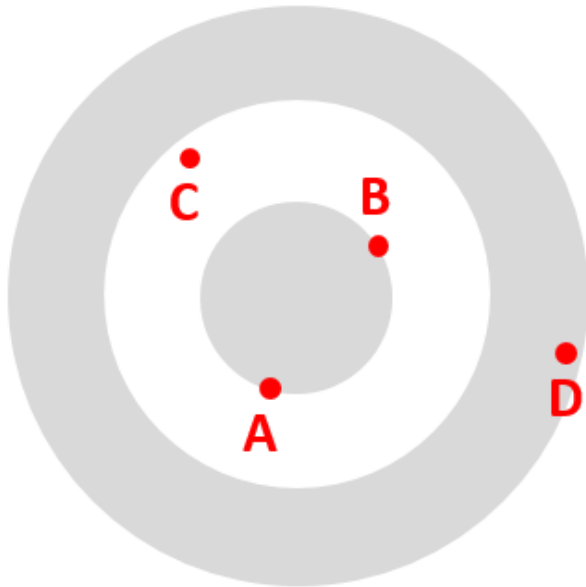
Source: [A playful generator of the normal distribution](#)

Suppose I throw a dart into a dartboard. I aim at the center of the board $(0,0)$, and the dart lands in a random position (X, Y) which has a joint density function $f(x, y)$

Two assumptions about the way I play darts:

1. The density is rotationally invariant so the distribution of where my dart lands only depends on the distance of the dart to the center.
2. The random variables X and Y are independent, how much I miss left and right makes no difference to the distribution of how much I miss up and down.

1. The density is rotationally invariant so the distribution of where my dart lands only depends on the distance of the dart to the center.



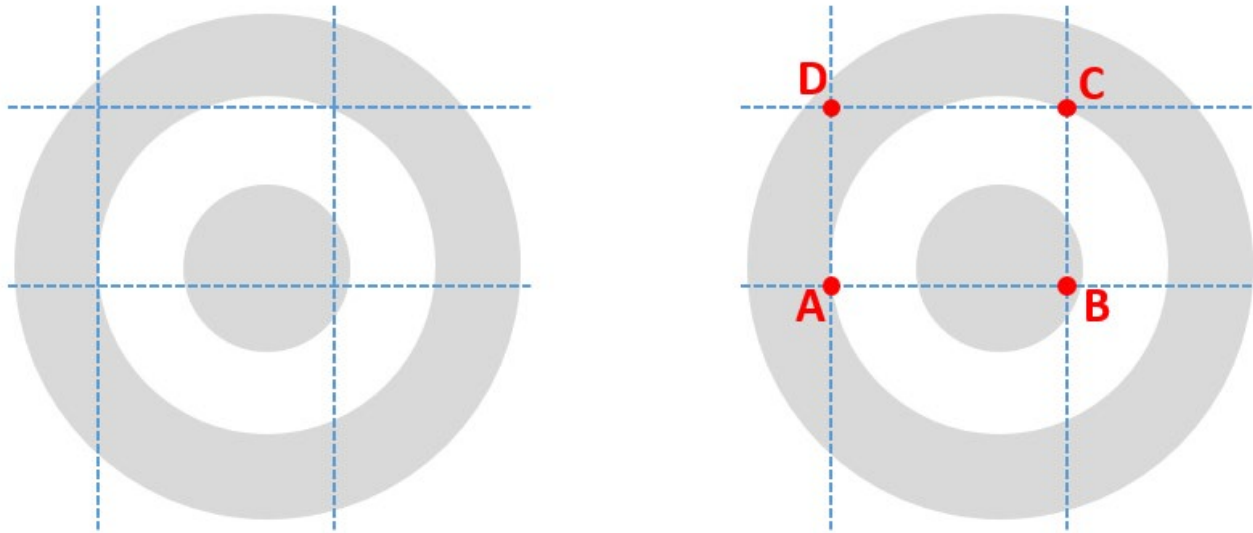
$$f(x, y) = g(x^2 + y^2)$$

Source: [A playful generator of the normal distribution](#)

Normal Density Derivation



2. The random variables X and Y are independent, how much I miss left and right makes no difference to the distribution of how much I miss up and down.



Source: [A playful generator of the normal distribution](#)

X and Y are independent and identically distributed

$$f(x, y) \propto f(x, 0) \cdot f(0, y) \qquad g(x^2 + y^2) \propto g(x^2) \cdot g(y^2)$$

Given $g(x^2 + y^2) \propto g(x^2) \cdot g(y^2)$, we know g must be an exponential function

$$g(t) = A \cdot e^{-Bt}$$

A is some normalizing constant.

B reflects the measuring units. (If I measure the distance in centimeter, B will be 10 times as big as if I measured in micrometer).

B must be positive because the density should be a decreasing function of distance (I'm not that bad at darts.)

Solving $\iint_{\mathbb{R}^2} g(x^2 + y^2) dx dy = 1$ to get A and B .

Standard Normal Distributions



Since each normally distributed variable has its own **mean** and **standard deviation**, the shape and location of these curves will vary. In practical applications, one would like to have a table of areas under the curve for each variable. To simplify this, statisticians use the standard normal distribution.

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

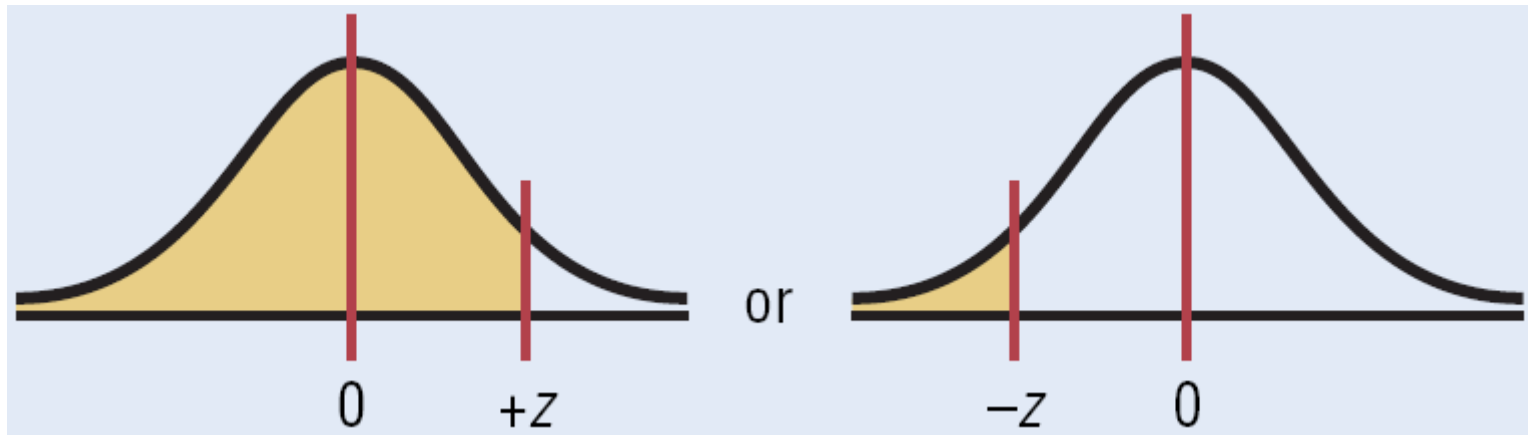
The z value is the number of standard deviations that a particular X value is away from the mean. The formula for finding the z value is:

$$z = \frac{\text{value} - \text{mean}}{\text{standard deviation}}$$

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

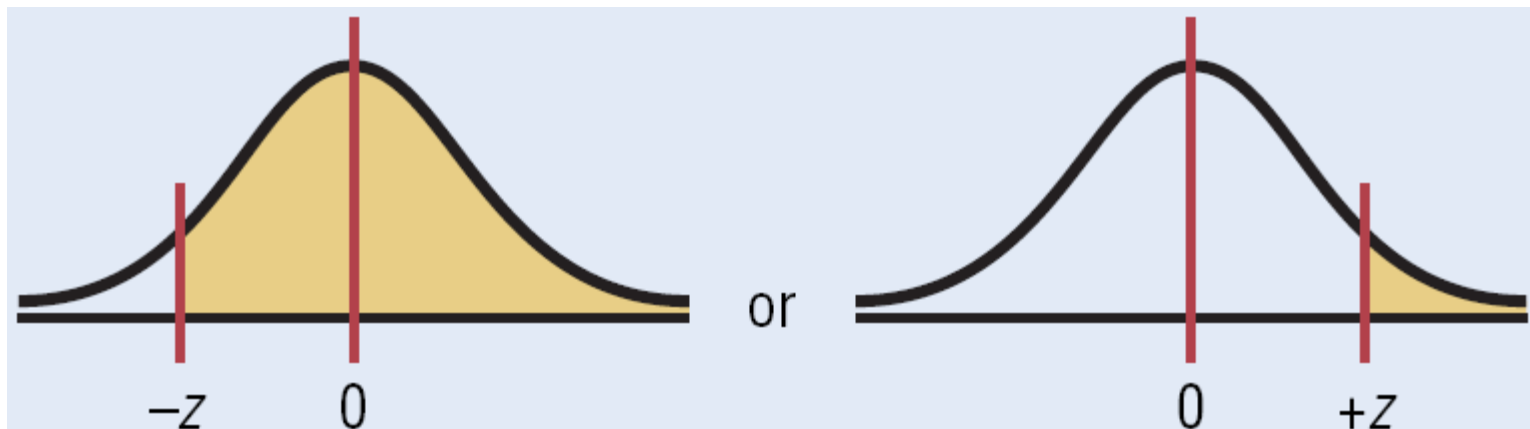
1. To the left of any z value:

Look up the z value in the table and use the area given.



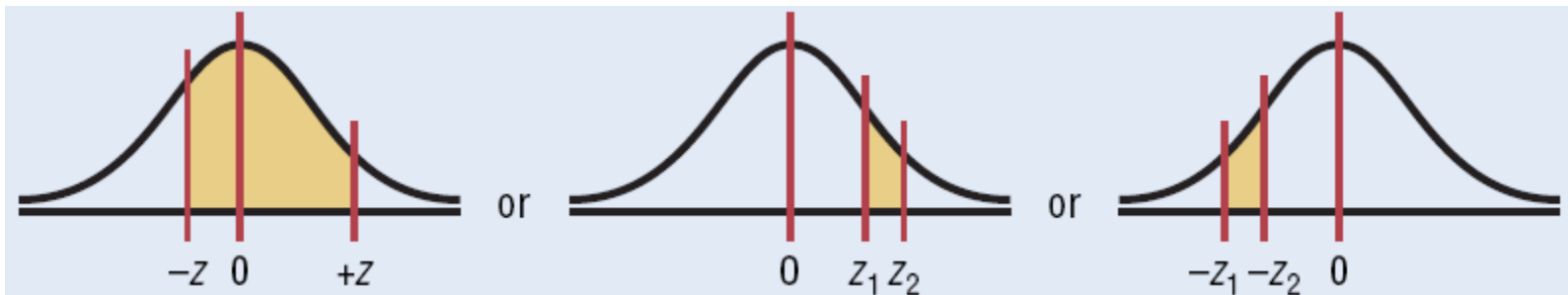
2. To the right of any z value:

Look up the z value and subtract the area from 1.



3. Between two z values:

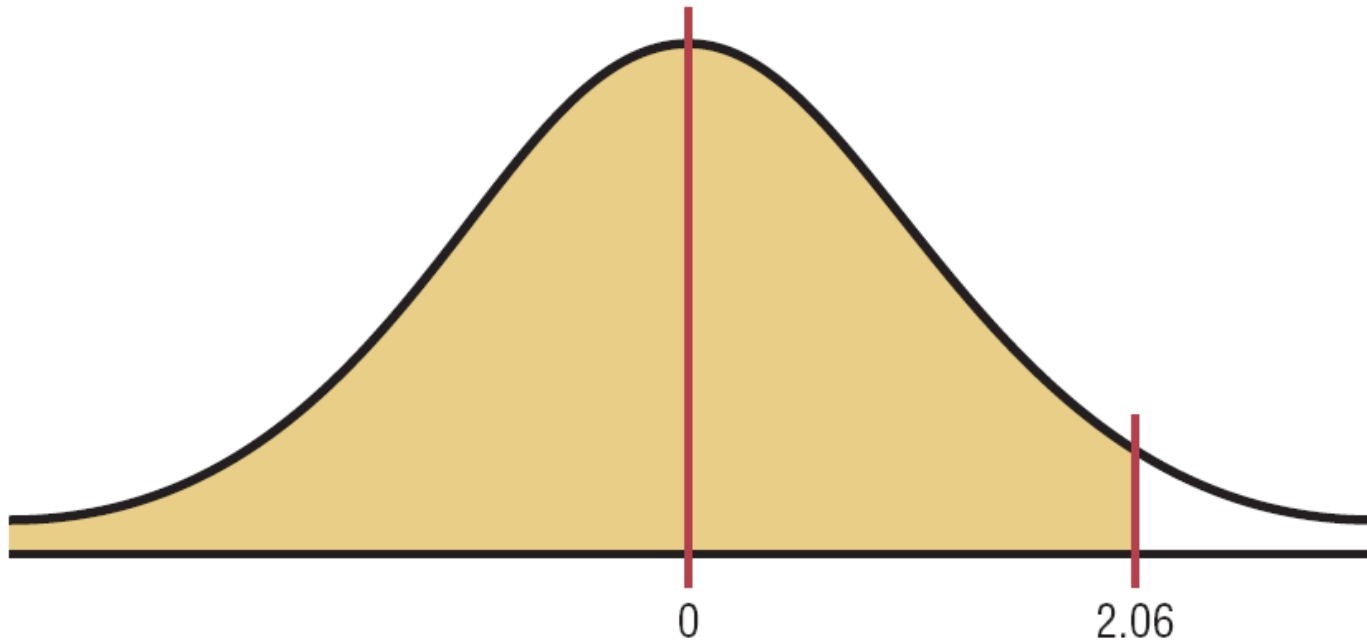
Look up both z values and subtract the corresponding areas.



Example: Area under the Curve



Find the area to the left of $z = 2.06$.

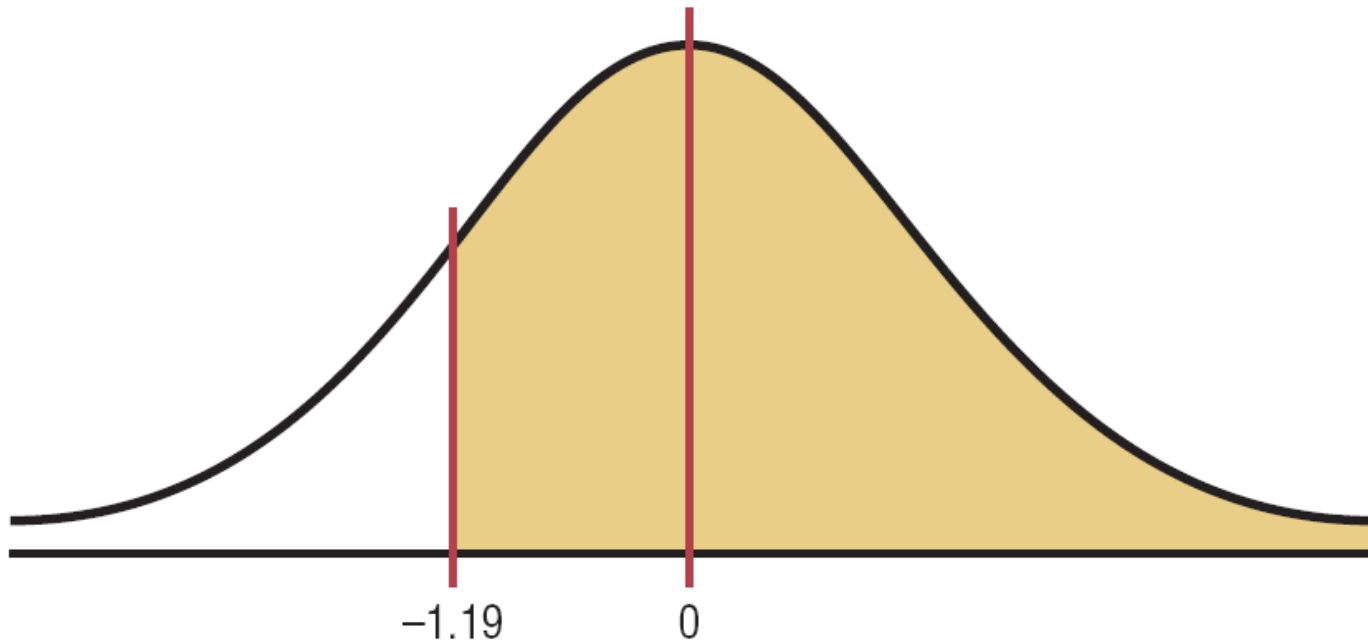


The value in the 2.0 row and the 0.06 column of Table E (of textbook) is 0.9803. The area is 0.9803.

Example: Area under the Curve



Find the area to the right of $z = -1.19$.

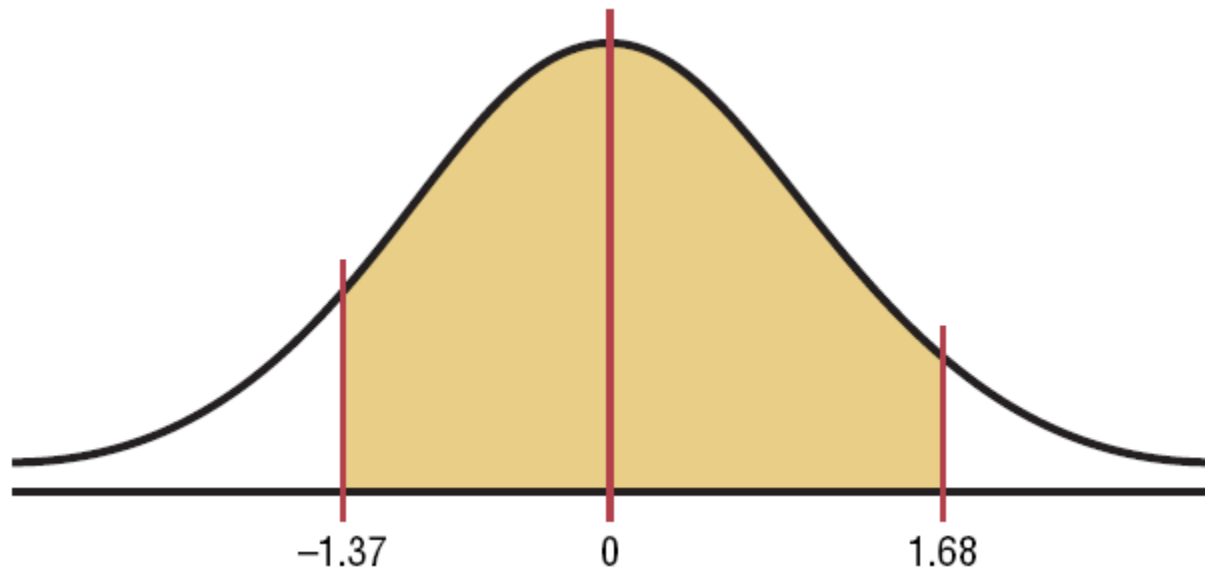


The value in the -1.1 row and the $.09$ column of Table E is 0.1170 . The area is $1 - 0.1170 = 0.8830$.

Example: Area under the Curve

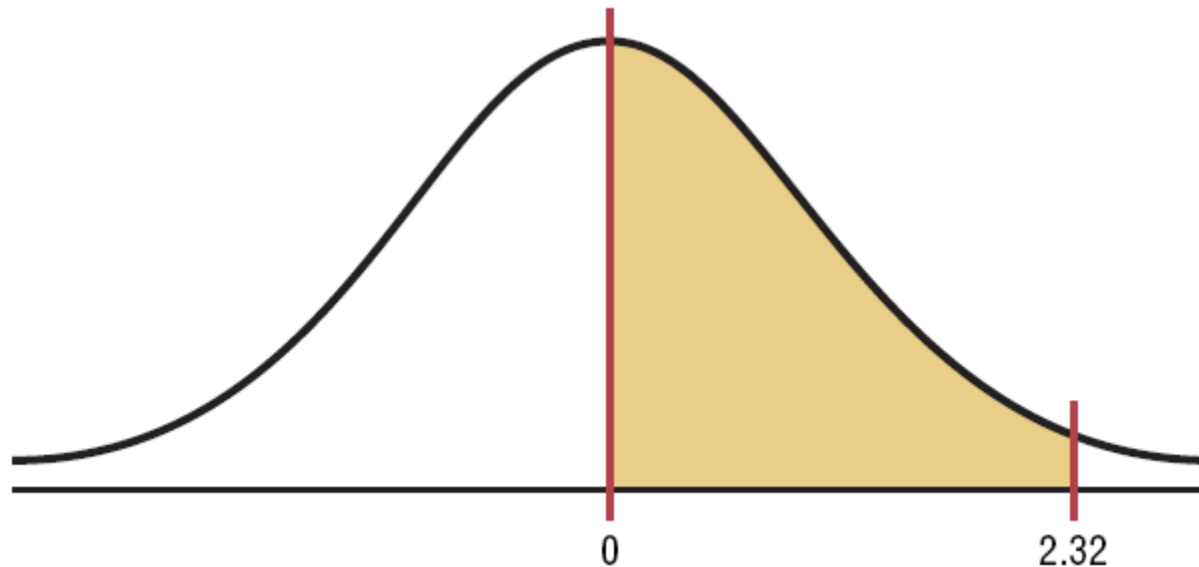


Find the area between $z = +1.68$ and $z = -1.37$.



The values for $z = +1.68$ is 0.9535 and for $z = -1.37$ is 0.0853. The area is $0.9535 - 0.0853 = 0.8682$.

a. Find the probability: $P(0 < z < 2.32)$

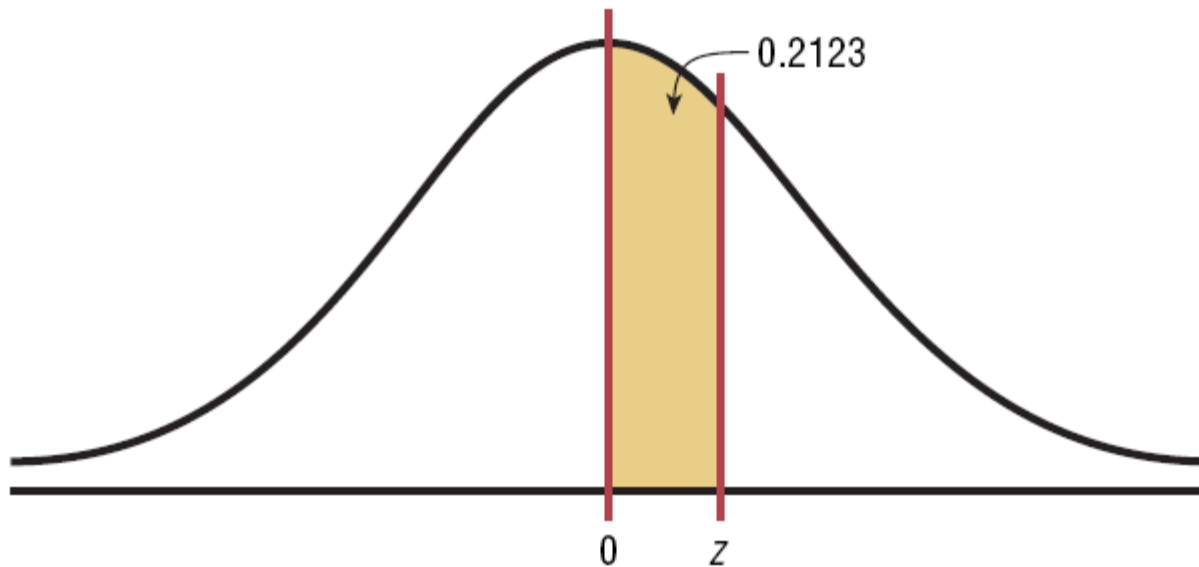


The values for $z = 2.32$ is 0.9898 and for $z = 0$ is 0.5000. The probability is $0.9898 - 0.5000 = 0.4898$.

Example: Probability



Find the z value such that the area under the standard normal distribution curve between 0 and the z value is 0.2123.



Add 0.5000 to 0.2123 to get the cumulative area of 0.7123. Then look for that value inside Table E.

A survey found that women spend on average \$146.21 on beauty products during the summer months.

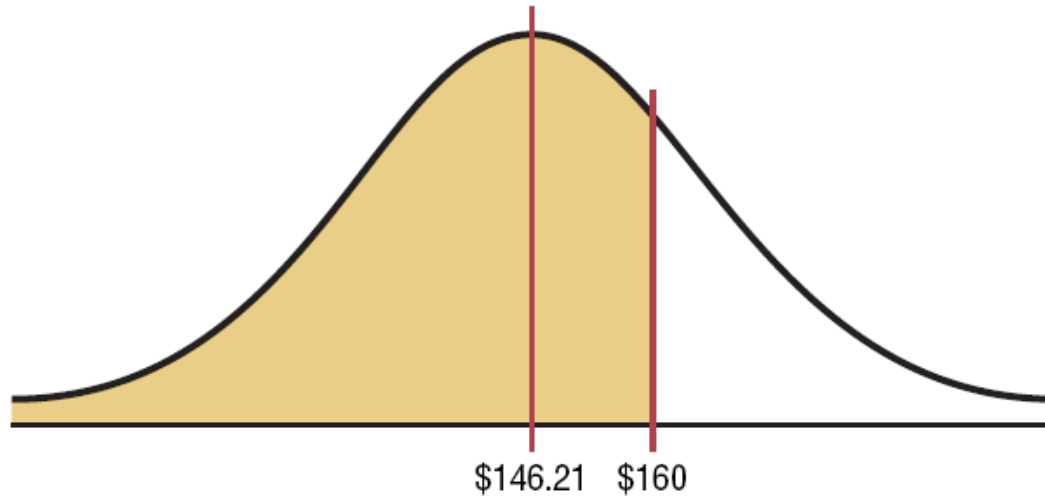
Assume the standard deviation is \$29.44.

Find the percentage of women who spend less than \$160.00. Assume the variable is normally distributed.

Example: Summer Spending



Step 1: Draw the normal distribution curve.



Example: Summer Spending



Step 2: Find the z value corresponding to \$160.00.

$$z = \frac{X - \mu}{\sigma} = \frac{160.00 - 146.21}{29.44} = 0.47$$

Step 3: Find the area to the left of $z = 0.47$.

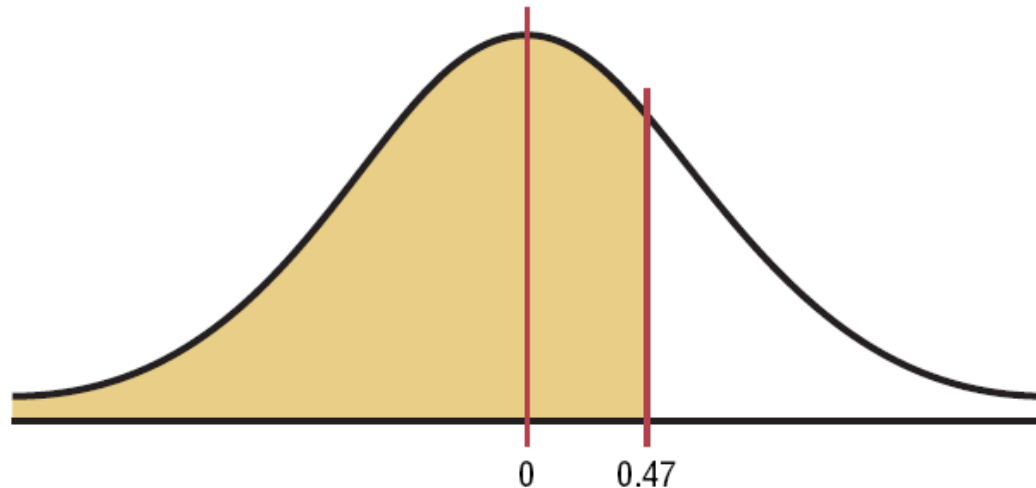


Table E gives us an area of .6808.

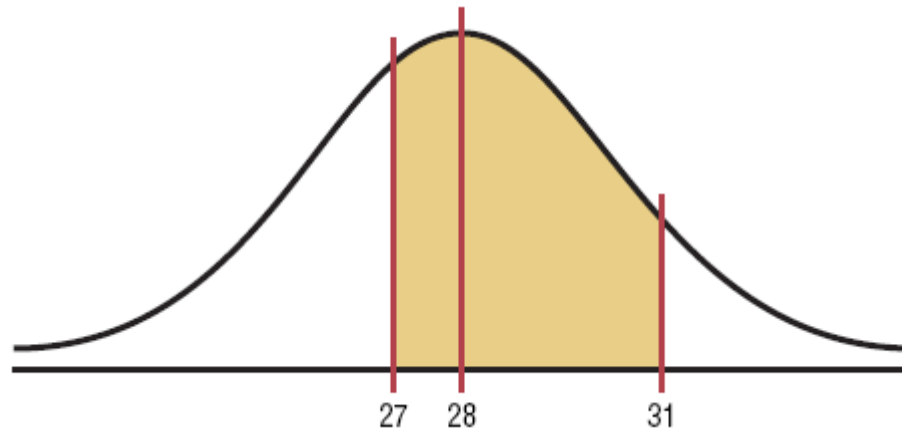
68% of women spend less than \$160.

Example: Newspaper Recycling



Each month, an American household generates an average of 28 pounds of newspaper for garbage or recycling. Assume the standard deviation is 2 pounds. If a household is selected at random, find the probability of its generating between 27 and 31 pounds per month. Assume the variable is approximately normally distributed.

Step 1: Draw the normal distribution curve.



Example: Newspaper Recycling



Step 2: Find z values corresponding to 27 and 31.

$$z = \frac{27 - 28}{2} = -0.5 \qquad z = \frac{31 - 28}{2} = 1.5$$

Step 3: Find the area between $z = -0.5$ and $z = 1.5$.

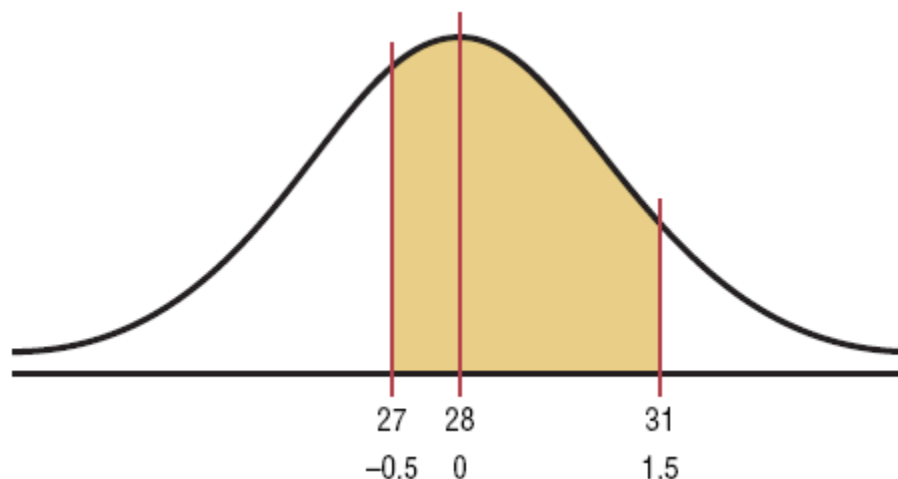


Table E gives us an area of $0.9332 - 0.3085 = 0.6247$.
The probability is 62%.

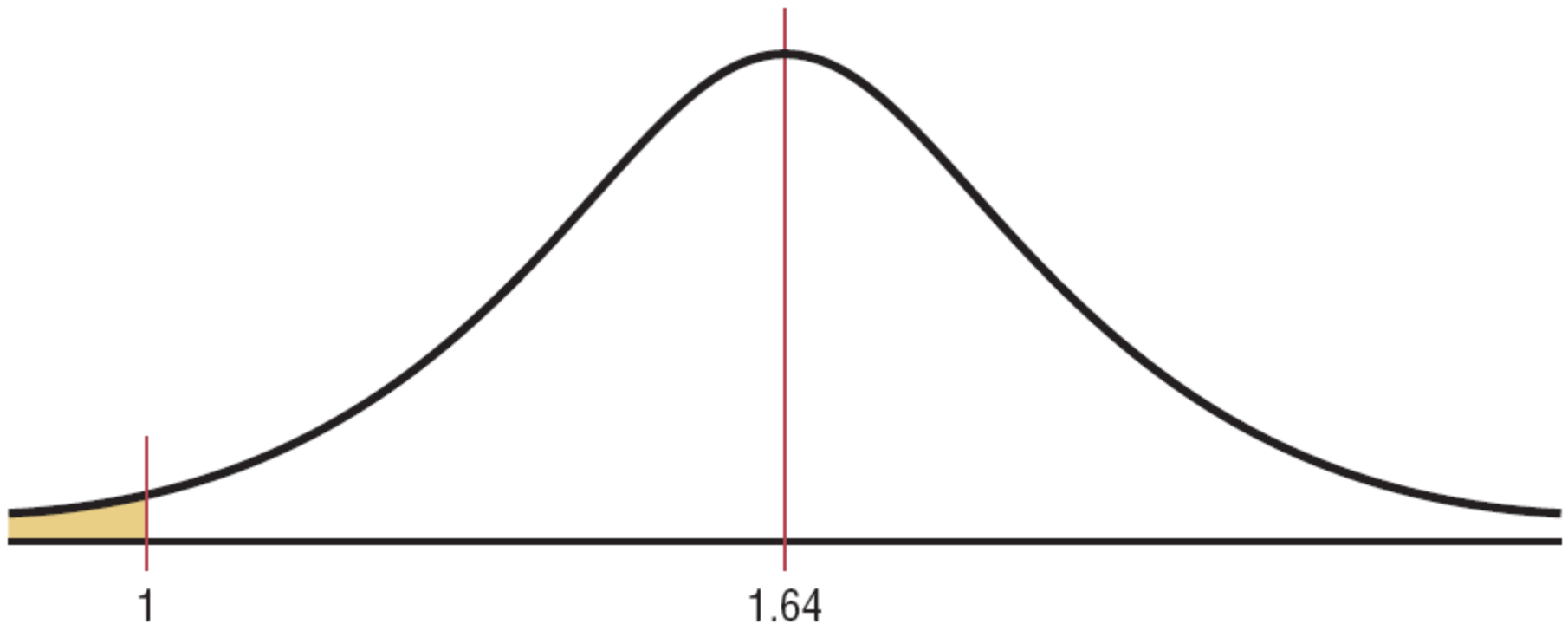
Americans consume an average of 1.64 cups of coffee per day. Assume the variable is approximately normally distributed with a standard deviation of 0.24 cup.

If 500 individuals are selected, approximately how many will drink less than 1 cup of coffee per day?

Example: Coffee Consumption



Step 1: Draw the normal distribution curve.



Step 2: Find the z value for 1.

$$z = \frac{1 - 1.64}{0.24} = -2.67$$

Step 3: Find the area to the left of $z = -2.67$. It is 0.0038.

Step 4: To find how many people drank less than 1 cup of coffee, multiply the sample size 500 by 0.0038 to get 1.9.

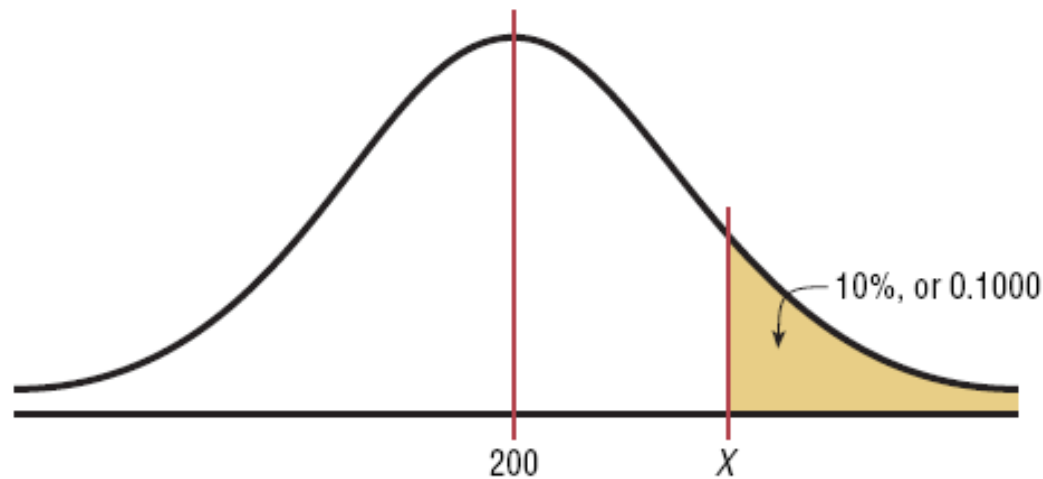
Since we are asking about people, round the answer to 2 people. Hence, approximately 2 people will drink less than 1 cup of coffee a day.

Example: Police Academy



To qualify for a police academy, candidates must score in the top 10% on a general abilities test. The test has a mean of 200 and a standard deviation of 20. Find the lowest possible score to qualify. Assume the test scores are normally distributed.

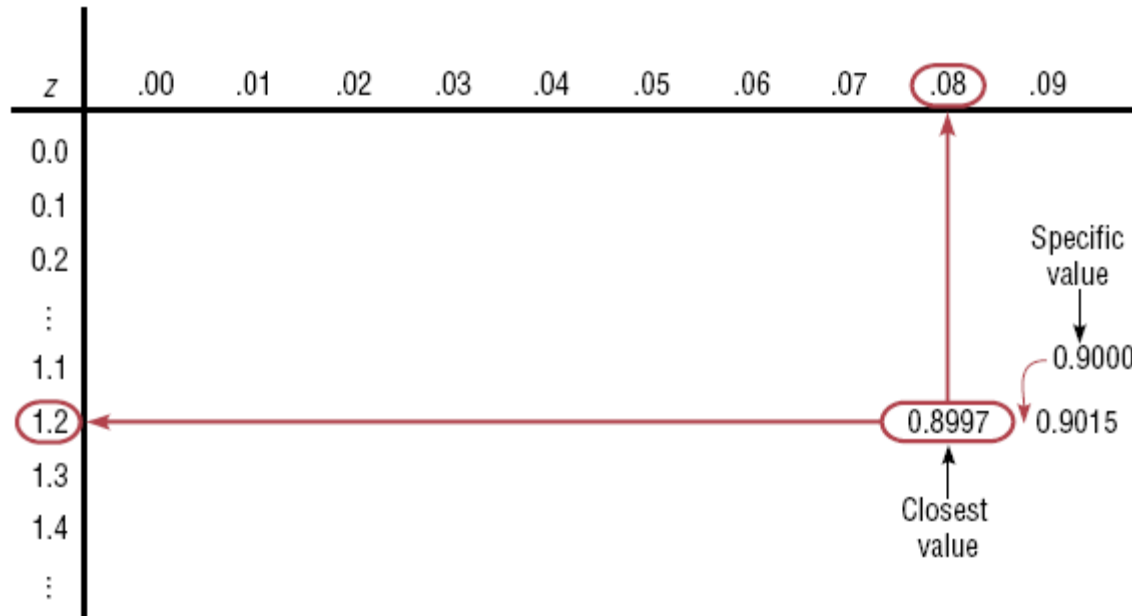
Step 1: Draw the normal distribution curve.



Example: Police Academy



Step 2: Subtract $1 - 0.1000$ to find area to the left, 0.9000. Look for the closest value to that in Table E.



Step 3: Find X .

$$X = \mu + z\sigma = 200 + 1.28(20) = 225.60$$

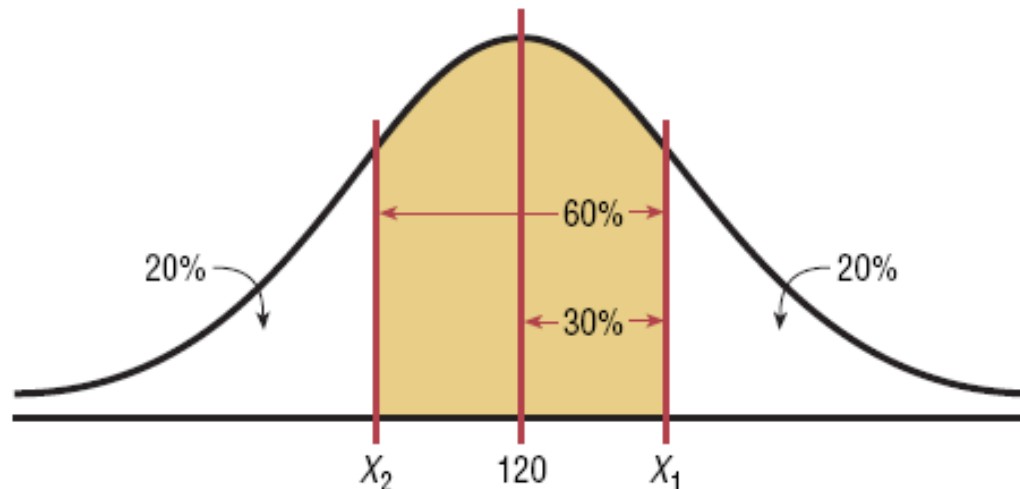
The cutoff, the lowest possible score to qualify, is 226.

Example: Systolic Blood Pressure

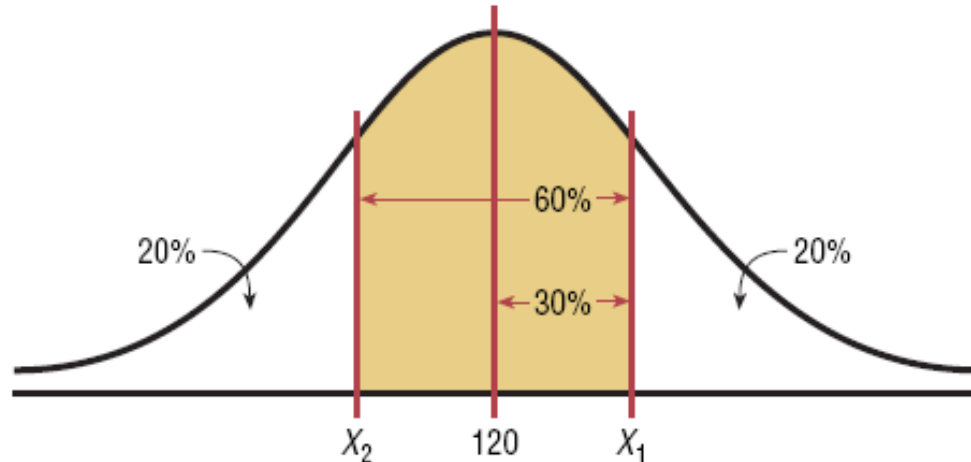


For a medical study, a researcher wishes to select people in the middle 60% of the population based on blood pressure. If the mean systolic blood pressure is 120 and the standard deviation is 8, find the upper and lower readings that would qualify people to participate in the study.

Step 1: Draw the normal distribution curve.



Example: Systolic Blood Pressure



Area to the left of the positive z : $0.5000 + 0.3000 = 0.8000$.

Using Table E, $z \approx 0.84$.

$$X = 120 + 0.84(8) = 126.72$$

Area to the left of the negative z : $0.5000 - 0.3000 = 0.2000$.

Using Table E, $z \approx -0.84$.

$$X = 120 - 0.84(8) = 113.28$$

The middle 60% of readings are between 113 and 127.

A normally shaped or bell-shaped distribution is only one of many shapes that a distribution can assume; however, it is very important since many statistical methods require that the distribution of values (shown in subsequent chapters) be normally or approximately normally shaped.

There are a number of ways statisticians check for normality. We will focus on three of them.

Histogram

Pearson's Index PI of Skewness

Outliers

Normal Quantile Plot

Kolmogorov-Smirnov Test

Other Tests

- Chi-Square Goodness-of-Fit Test
- Lilliefors Test

Example: Technology Inventories

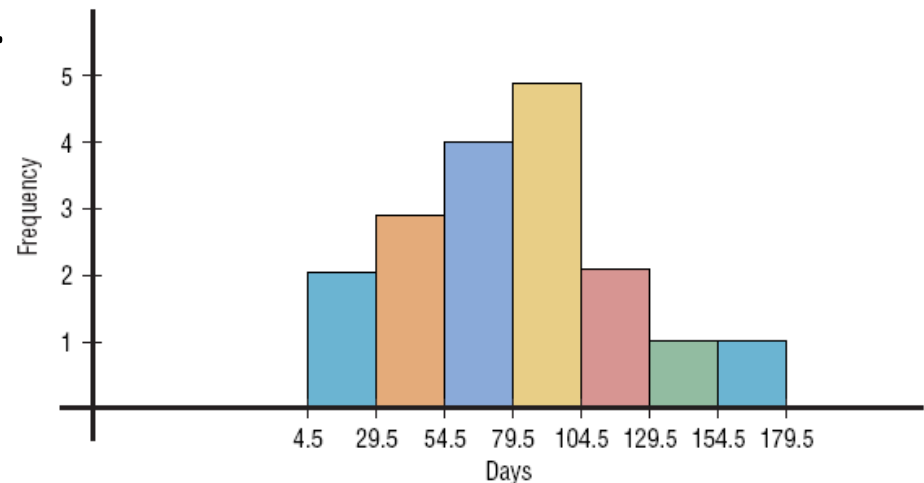


A survey of 18 high-technology firms showed the number of days' inventory they had on hand. Determine if the data are approximately normally distributed.

5 29 34 44 45 63 68 74 74
81 88 91 97 98 113 118 151 158

Method 1: Construct a Histogram.

Class	Frequency
5–29	2
30–54	3
55–79	4
80–104	5
105–129	2
130–154	1
155–179	1



The histogram is approximately bell-shaped.

Method 2: Check for Skewness.

$$\bar{X} = 79.5, MD = 77.5, s = 40.5$$

$$PI = \frac{3(\bar{X} - MD)}{s} = \frac{3(79.5 - 77.5)}{40.5} = 0.148$$

The PI is not greater than 1 or less than -1 , so it can be concluded that the distribution is not significantly skewed.

Method 3: Check for Outliers.

Five-Number Summary: 5 - 45 - 77.5 - 98 - 158

$$Q1 - 1.5(IQR) = 45 - 1.5(53) = -34.5$$

$$Q3 + 1.5(IQR) = 98 + 1.5(53) = 177.5$$

No data below -34.5 or above 177.5 , so no outliers.

A survey of 18 high-technology firms showed the number of days' inventory they had on hand. Determine if the data are approximately normally distributed.

5 29 34 44 45 63 68 74 74
81 88 91 97 98 113 118 151 158

Conclusion:

The histogram is approximately bell-shaped.

The data are not significantly skewed.

There are no outliers.

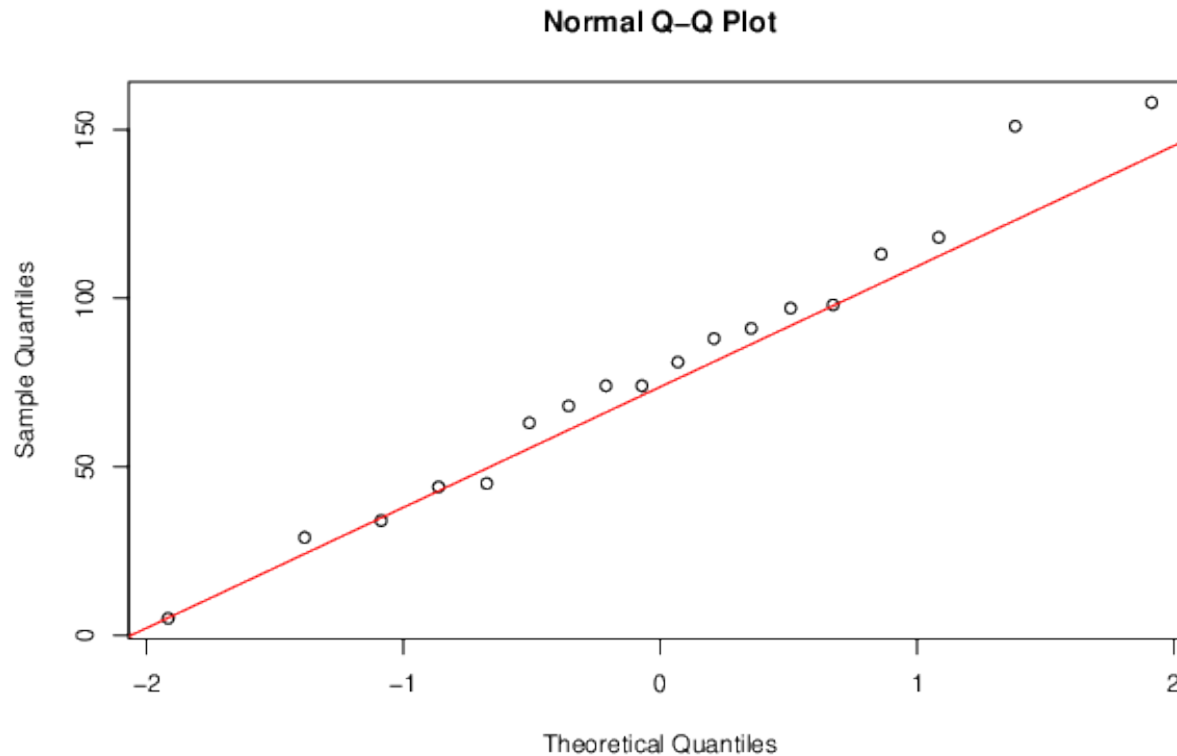
Thus, it can be concluded that the distribution is approximately normally distributed.

Normal Quantile Plot



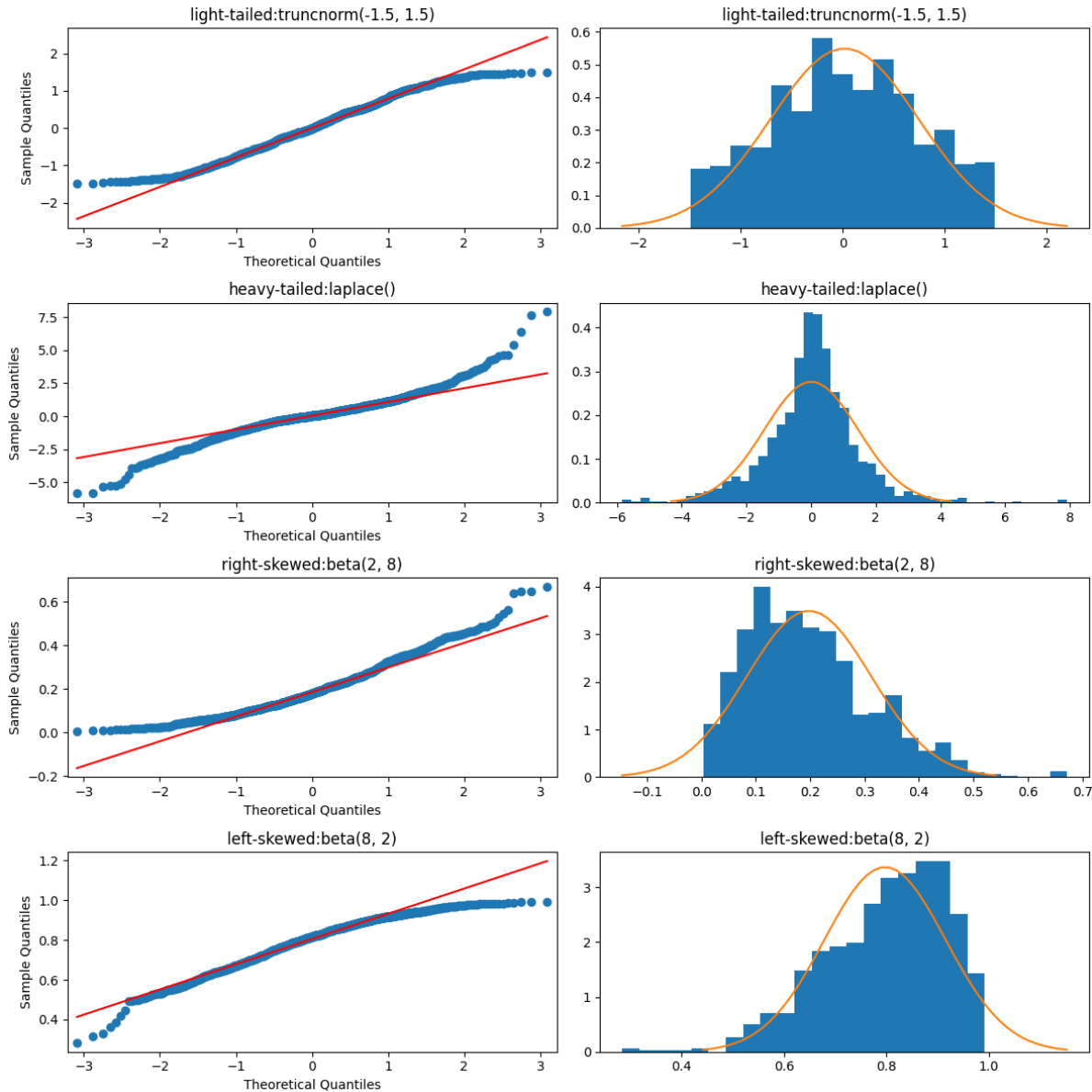
A survey of 18 high-technology firms showed the number of days' inventory they had on hand. Determine if the data are approximately normally distributed.

5 29 34 44 45 63 68 74 74 81 88 91 97 98 113 118 151 158



Free Statistics Calculator: https://www.wessa.net/rwasp_varia1.wasp

Normal Quantile Plot

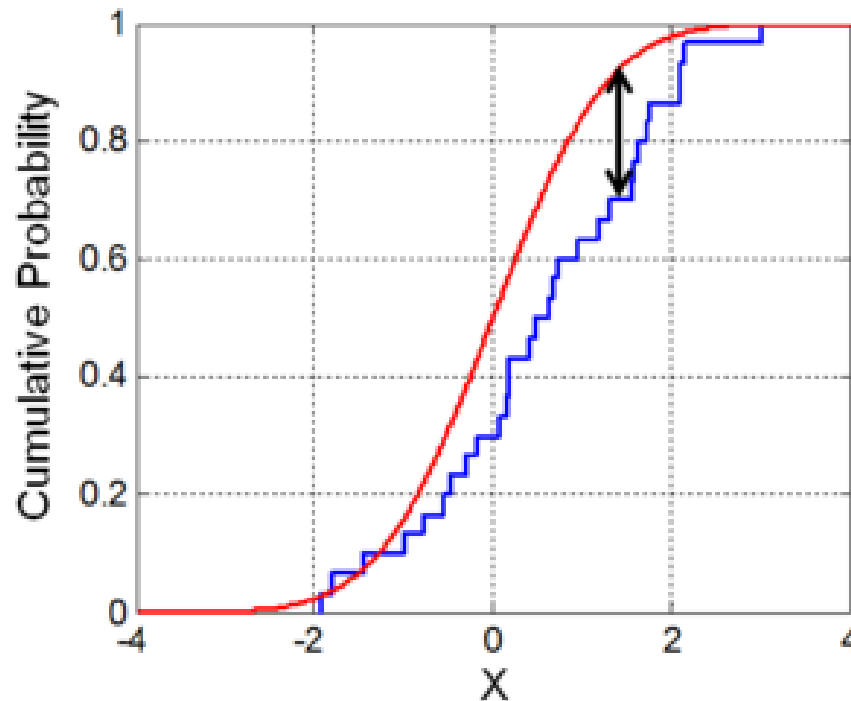


Source: [How to interpret a QQ plot?](#)

Kolmogorov–Smirnov test



Kolmogorov–Smirnov test (K–S test or KS test) is a nonparametric test of the equality of continuous (or discontinuous), one-dimensional probability distributions that can be used to test whether a sample came from a given reference probability distribution.



The red line is a model CDF

The blue line is an empirical CDF

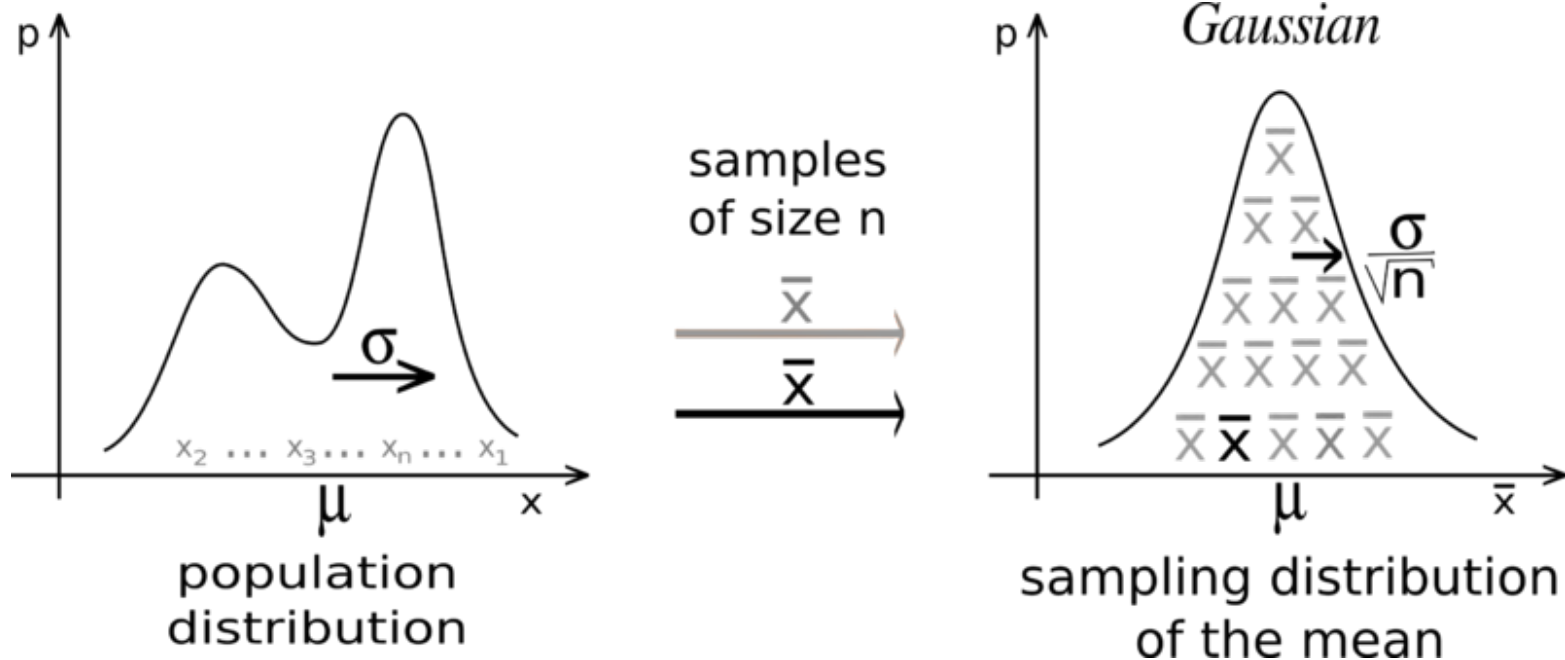
The black arrow is the KS statistic.

The Central Limit Theorem

Sample Mean



In addition to knowing how individual data values vary about the mean for a population, statisticians are interested in knowing how the **means of samples** of the same size taken from the same population vary about the population mean.



Let $X_1, \dots, X_n$ be an independent and identically distributed (i.i.d.) sample of size n drawn from a population with mean μ and standard deviation σ

A **sampling distribution of sample means** is a distribution obtained by using the means computed from random samples of a specific size taken from a population. The distribution of $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$.

Sampling error is the difference between the sample measure and the corresponding population measure due to the fact that the sample is not a perfect representation of the population. $\bar{X} - \mu$.



The mean of the sample means will be the same as the population mean.

$$\mathbb{E}[\bar{X}] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n X_i\right] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i] = \mu.$$

The standard deviation of the sample means will be equal to the population standard deviation divided by the square root of the sample size.

$$\text{Var}[\bar{X}] = \text{Var}\left[\frac{1}{n}\sum_{i=1}^n X_i\right] = \frac{1}{n^2}\text{Var}\left[\sum_{i=1}^n X_i\right] = \frac{1}{n}\sigma^2.$$



As the sample size n increases, the shape of the distribution of the sample means taken **with replacement** from a population with mean μ and standard deviation σ will approach a normal distribution.

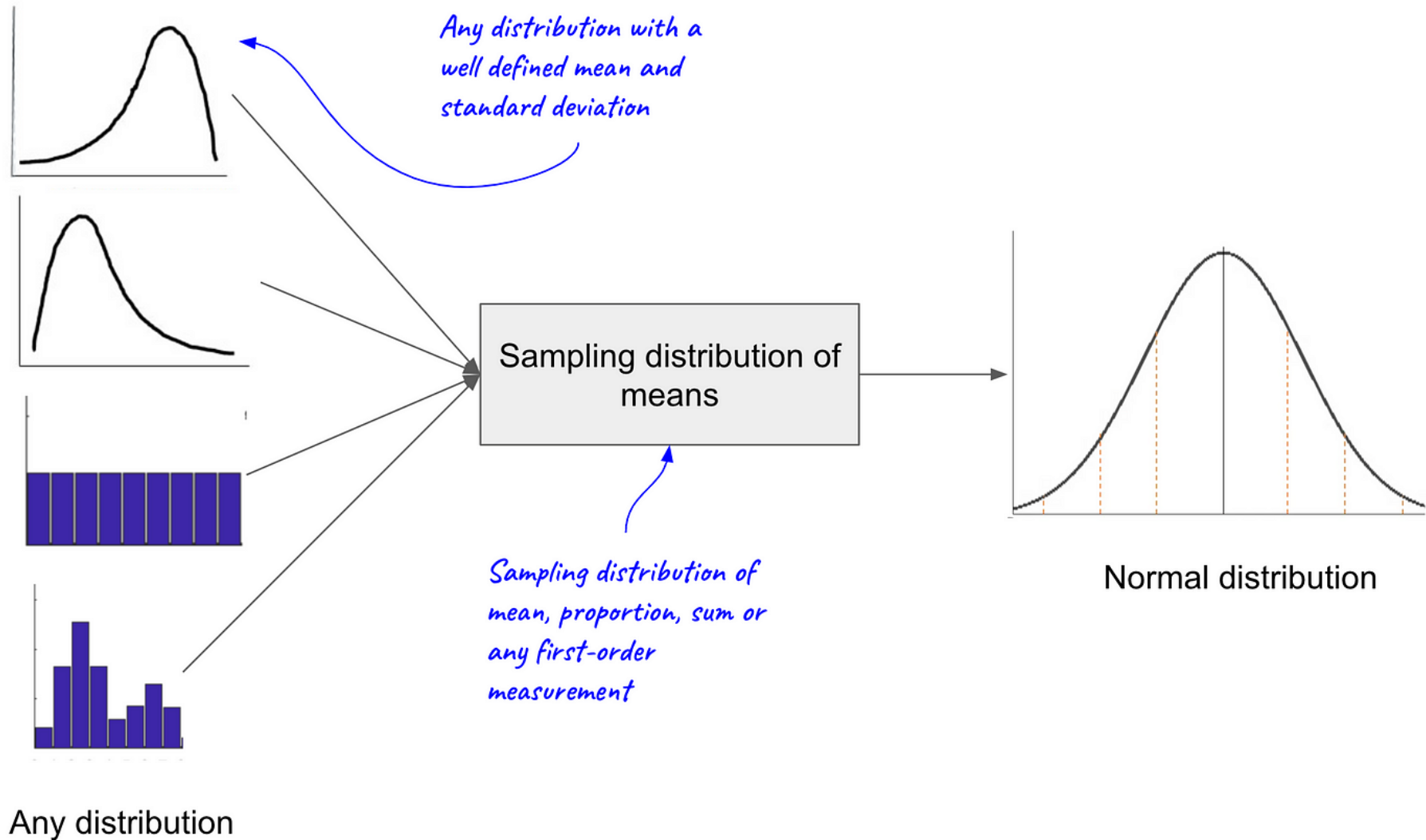
The mean of the sample means equals the population mean.

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

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

$$\sigma_{\bar{X}} = \sigma / \sqrt{n}.$$

The Central Limit Theorem



Source: [Stat Digest: What is the intuition behind Central Limit Theorem?](#)

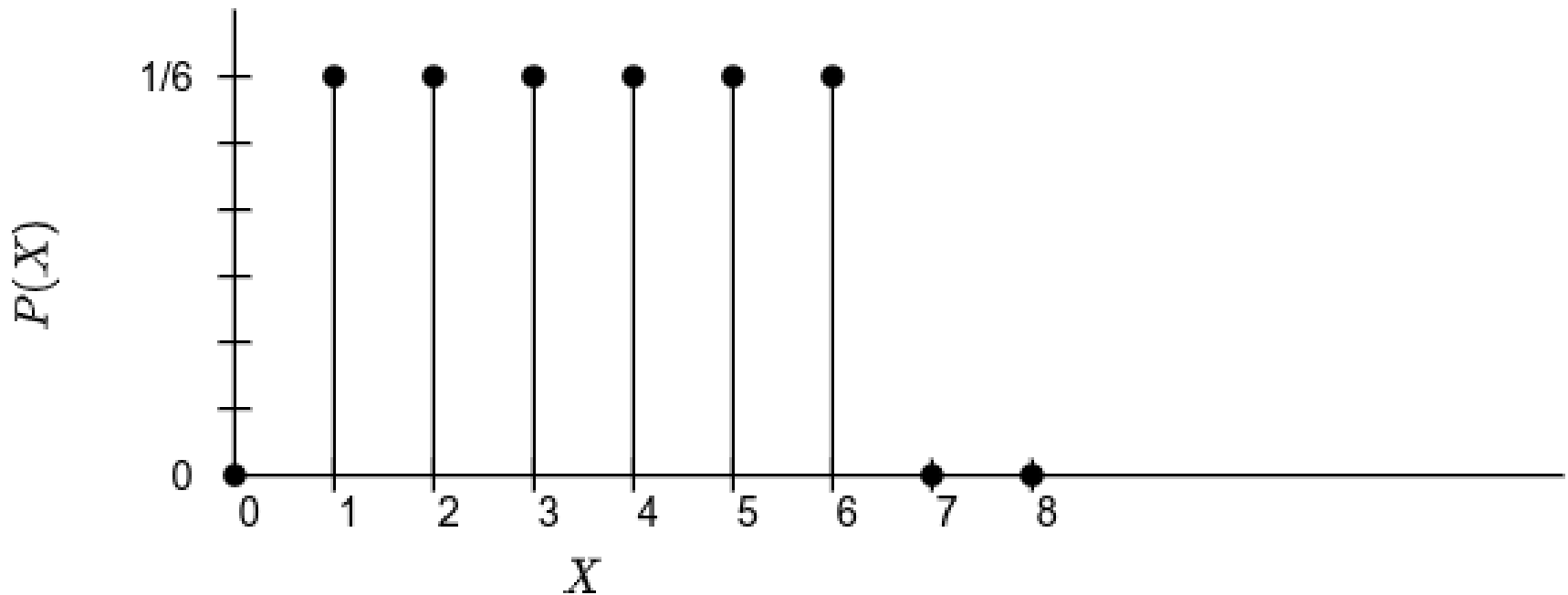
Convolution of Discrete Random Variables: If X and Y are independent random variables,

$$P(X + Y = n) = \sum_k P(X = k) \cdot P(Y = n - k)$$

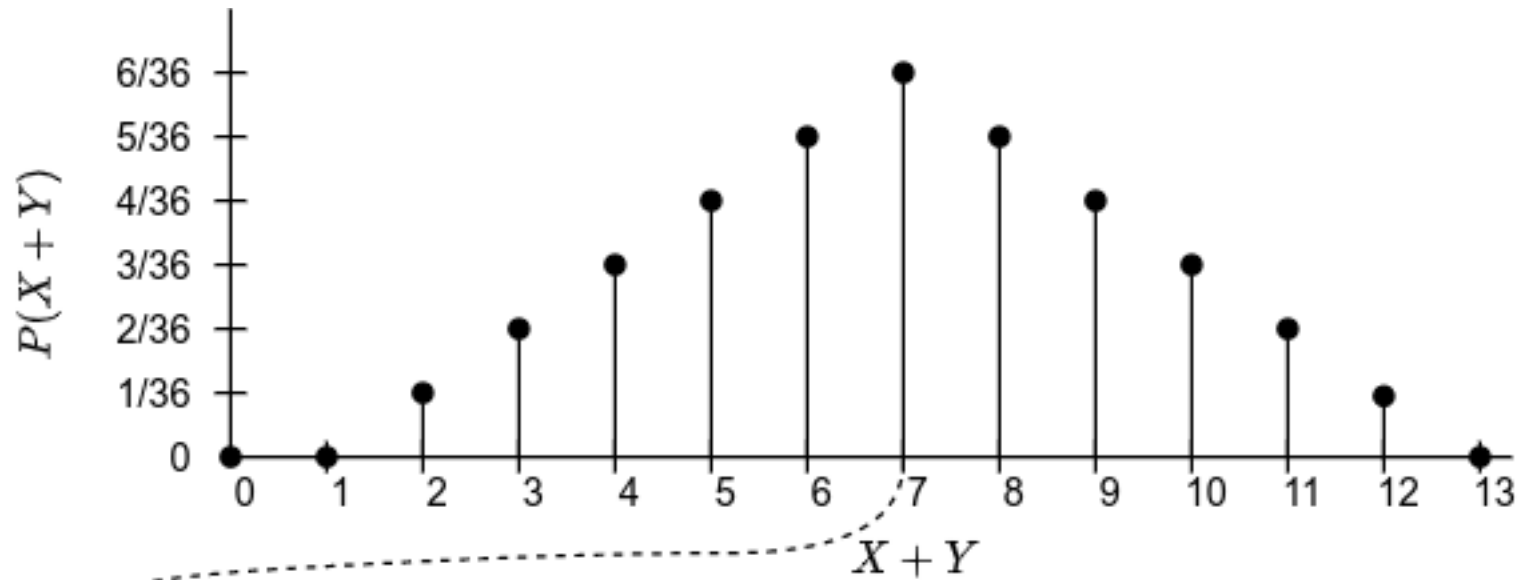
Convolution of Continuous Random Variables: If X and Y are independent random variables,

$$f_{X+Y}(t) = \int_{-\infty}^{+\infty} f_X(s) \cdot f_Y(t - s) ds$$

Example: Rolling Dice



Example: Rolling Dice



$$\begin{aligned}
 & \xrightarrow{\text{dashed arrow}} P(7) \\
 &= P(\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}) \\
 &= P((1, 6)) + P((2, 5)) + P((3, 4)) + P((4, 3)) + P((5, 2)) + P((6, 1)) \\
 &= P(1)P(6) + P(2)P(5) + P(3)P(4) + P(4)P(3) + P(5)P(2) + P(6)P(1) \\
 &= \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} \\
 &= \frac{6}{36}
 \end{aligned}$$

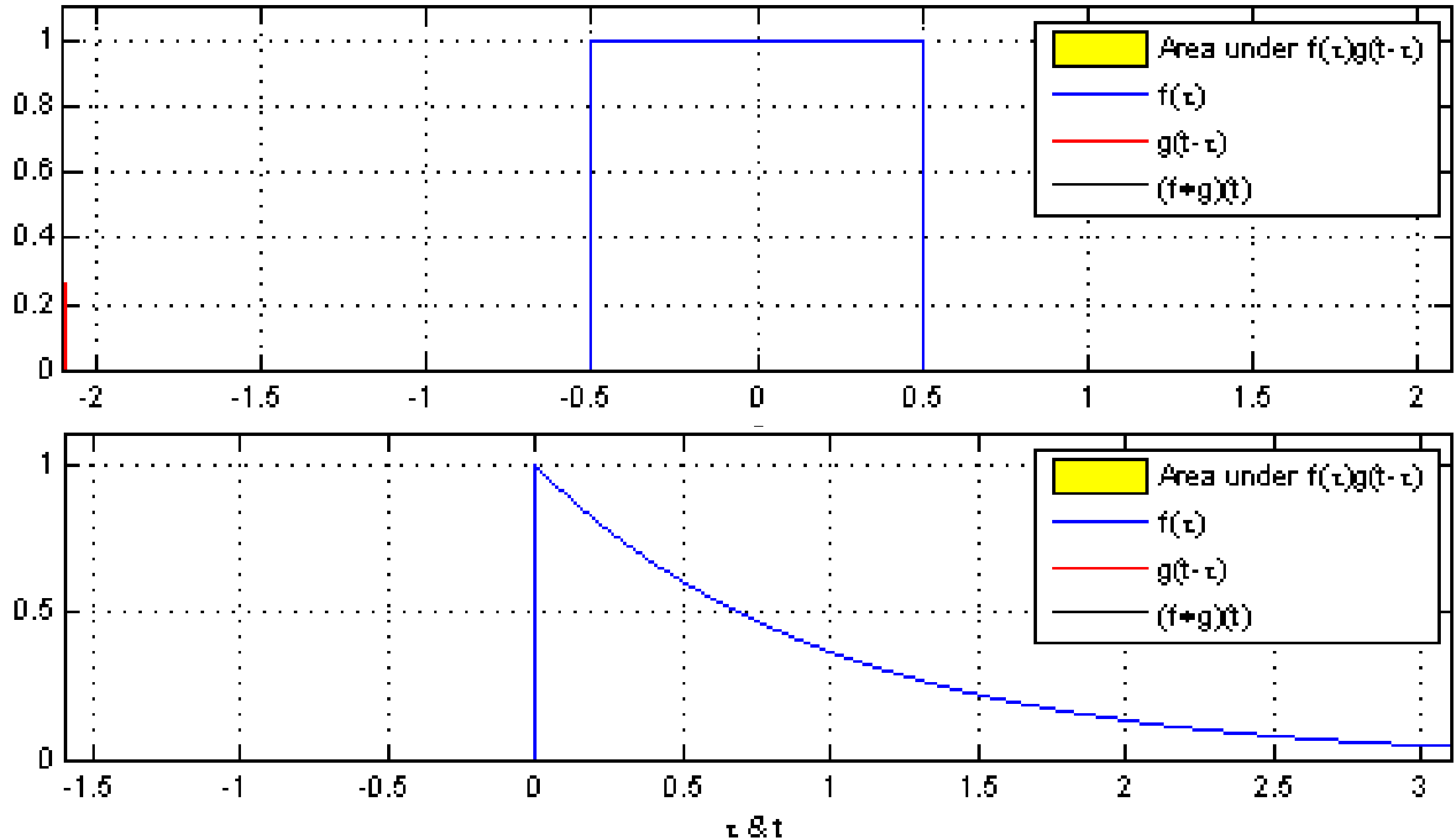
Example: Rolling Dice

Die 1	Die 2					
	1	2	3	4	5	6
1	(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
2	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
3	(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
4	(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
5	(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
6	(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

Convolution

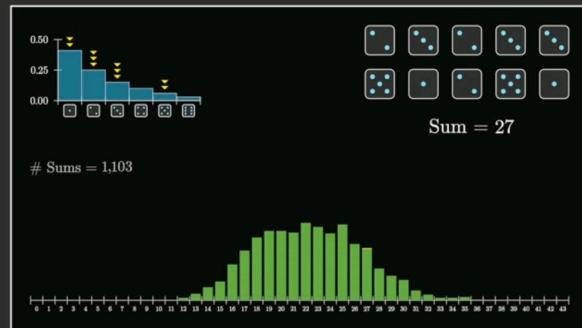


Example: Convolution for Continuous Distribution

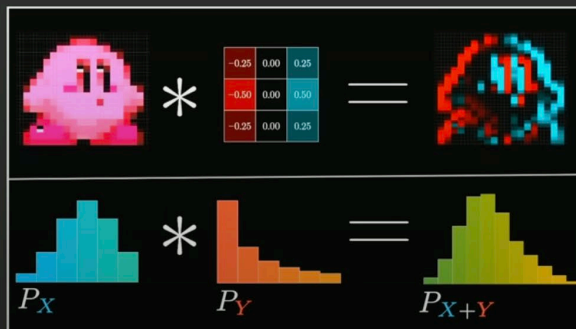


Source: [Convolution and probability theory](#)

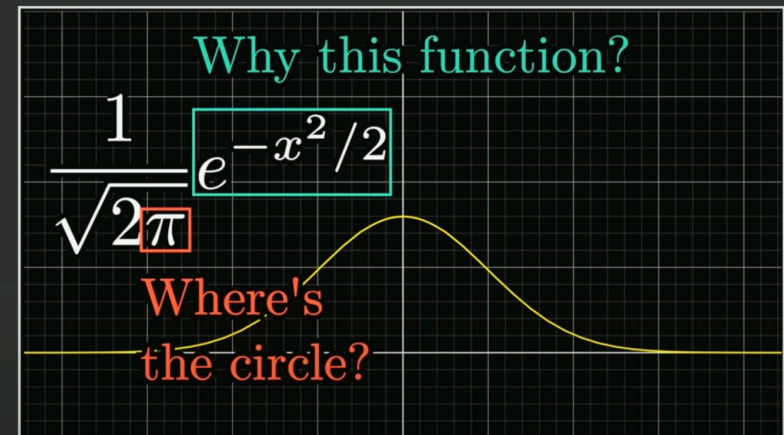
The Central Limit Theorem



Convolutions



Diving deeper



Thanks to [3Blue1Brown](#)

Convolution: <https://youtu.be/KuXjwB4LzSA?si=2fnZnmWKTW9loFkS>

Central Limit Theorem: <https://www.youtube.com/watch?v=zeJD6dqJ5lo&t=649s>

The central limit theorem can be used to answer questions about sample means in the same manner that the normal distribution can be used to answer questions about individual values.

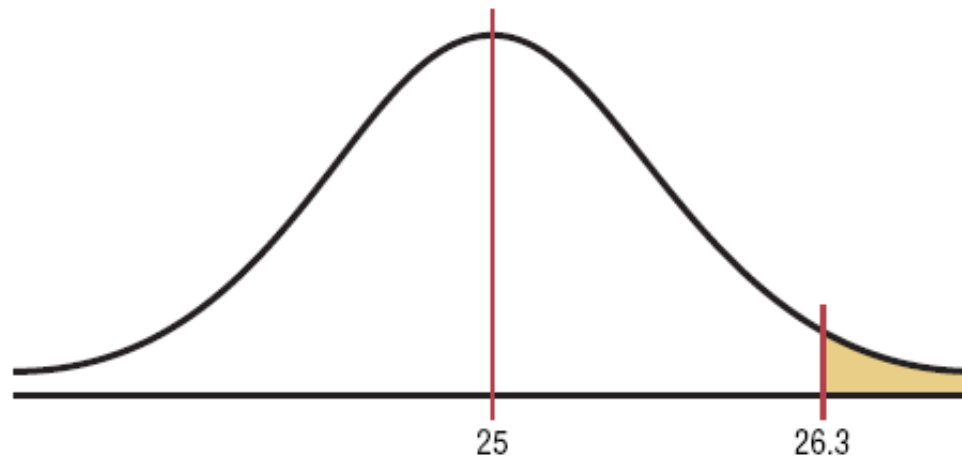
A new formula must be used for the z values:

$$z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

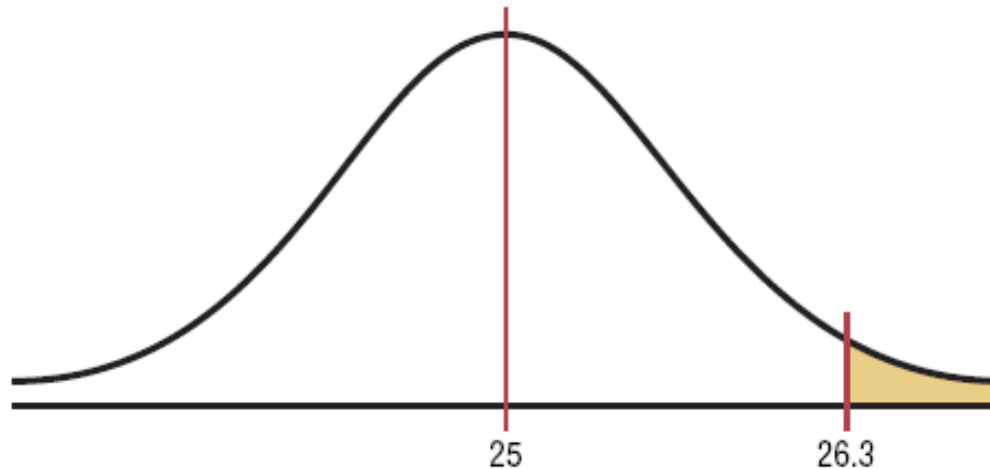
Example: Hours of Television



A. C. Nielsen reported that children between the ages of 2 and 5 watch an average of 25 hours of television per week. Assume the variable is normally distributed and the standard deviation is 3 hours. If 20 children between the ages of 2 and 5 are randomly selected, find the probability that the mean of the number of hours they watch television will be greater than 26.3 hours.



Example 6-13: Hours of Television



Since we are calculating probability for a sample mean, we need the Central Limit Theorem formula

$$z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{26.3 - 25}{3 / \sqrt{20}} = 1.94$$

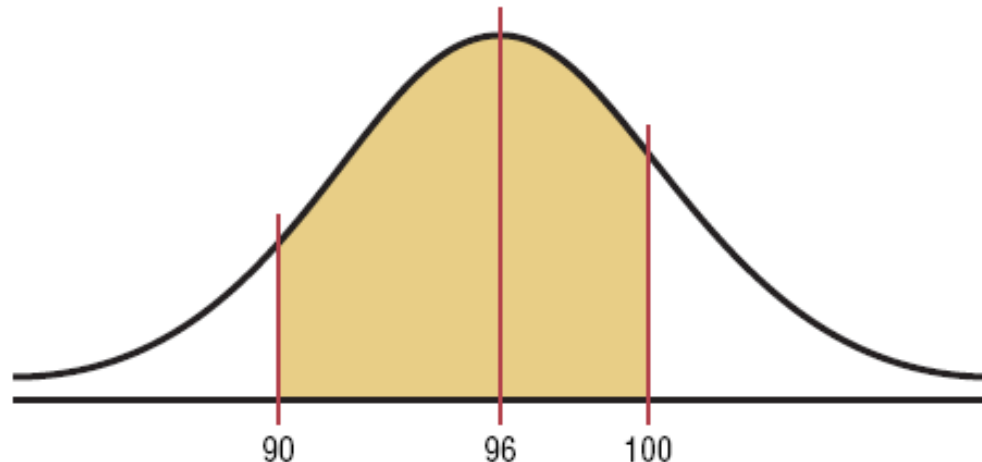
The area is $1.0000 - 0.9738 = 0.0262$. The probability of obtaining a sample mean larger than 26.3 hours is 2.62%.

Example: Vehicle Age

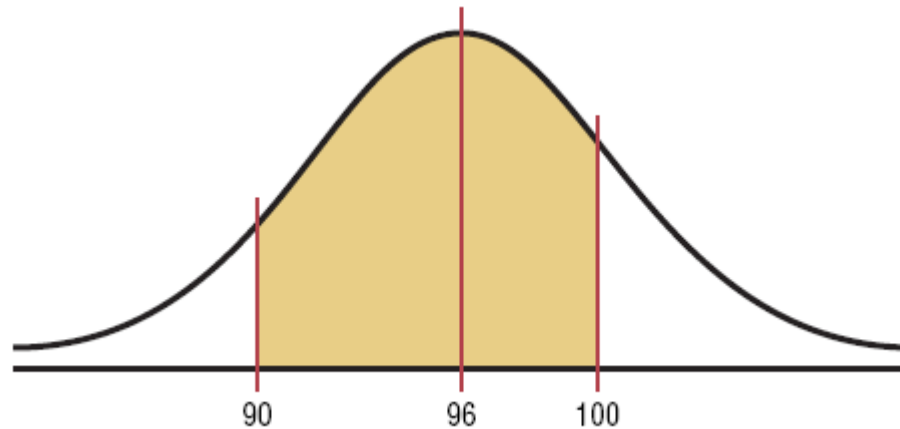


The average age of a vehicle registered in the United States is 8 years, or 96 months. Assume the standard deviation is 16 months. If a random sample of 36 vehicles is selected, find the probability that the mean of their age is between 90 and 100 months.

Since the sample is 30 or larger, the normality assumption is not necessary.



Example: Vehicle Age



$$z = \frac{90 - 96}{16/\sqrt{36}} = -2.25$$

$$z = \frac{100 - 96}{16/\sqrt{36}} = 1.50$$

Table E gives us areas 0.9332 and 0.0122, respectively.

The desired area is $0.9332 - 0.0122 = 0.9210$.

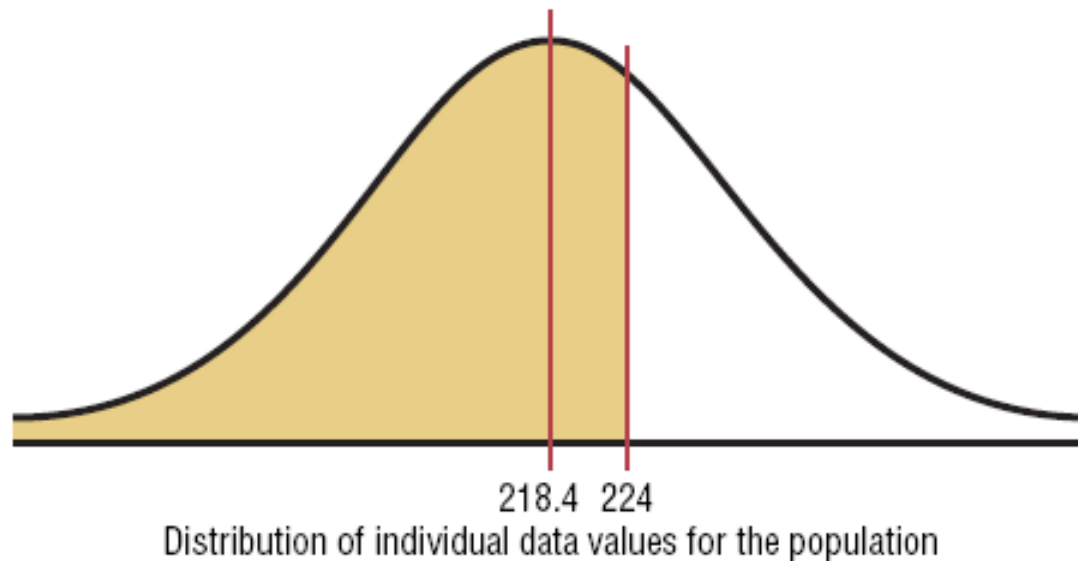
The probability of obtaining a sample mean between 90 and 100 months is 92.1%.

Example: Meat Consumption

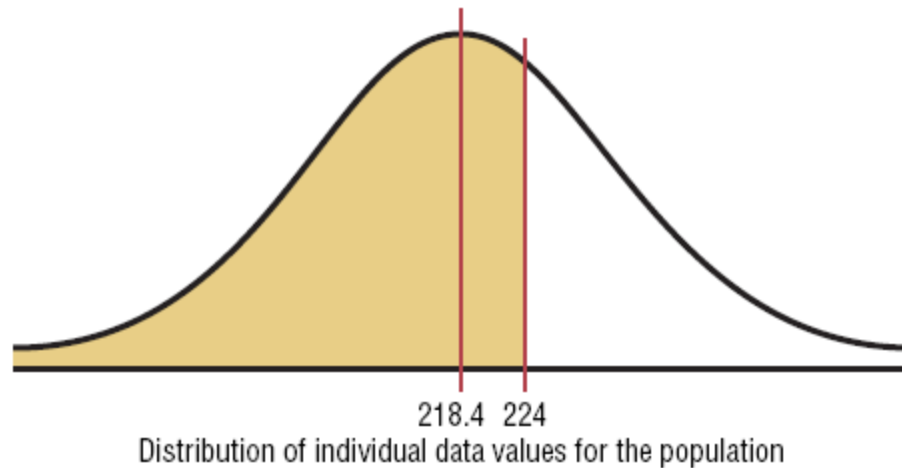


The average number of pounds of meat that a person consumes per year is 218.4 pounds. Assume that the standard deviation is 25 pounds and the distribution is approximately normal.

- a. Find the probability that a person selected at random consumes less than 224 pounds per year.



Example: Meat Consumption



$$z = \frac{X - \mu}{\sigma} = \frac{224 - 218.4}{25} = 0.22$$

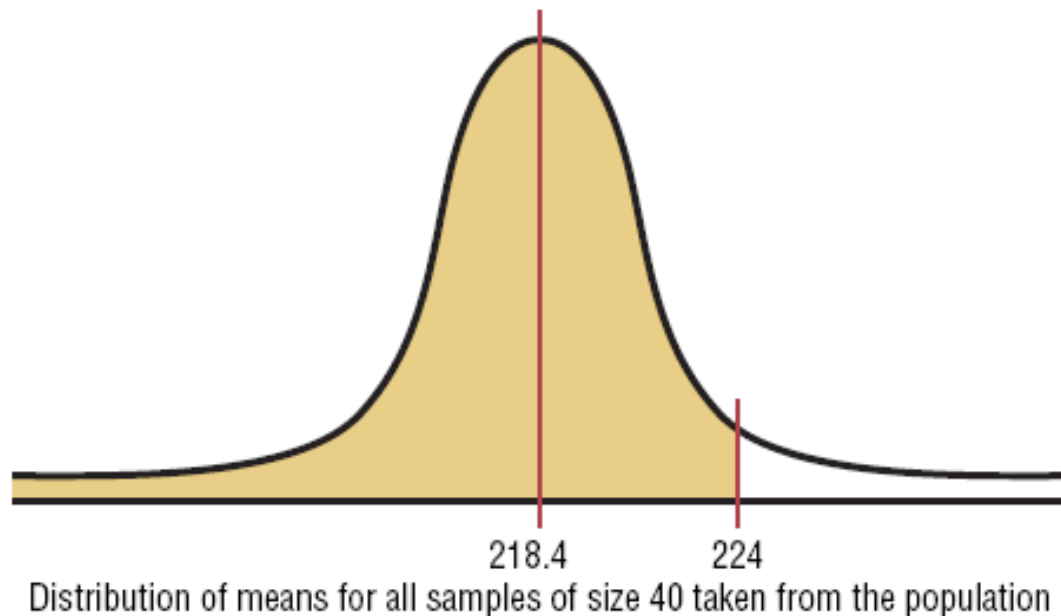
The area to the left of $z = 0.22$ is 0.5871. Hence, the probability of selecting an individual who consumes less than 224 pounds of meat per year is 0.5871, or 58.71%.

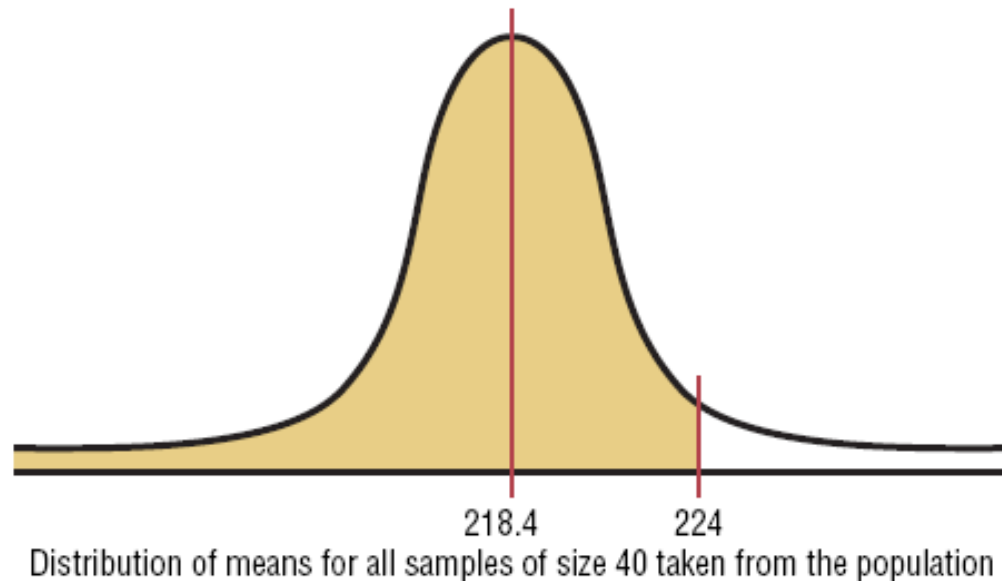
Example: Meat Consumption



The average number of pounds of meat that a person consumes per year is 218.4 pounds. Assume that the standard deviation is 25 pounds and the distribution is approximately normal.

- b. If a sample of 40 individuals is selected, find the probability the sample mean will be less than 224 pounds per year.





$$z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{224 - 218.4}{25 / \sqrt{40}} = 1.42$$

The area to the left of $z = 1.42$ is 0.9222. Hence, the probability that the mean of a sample of 40 individuals is less than 224 pounds per year is 0.9222, or 92.22%.



The formula for standard error of the mean is accurate when the samples are drawn with replacement or are drawn without replacement from a very large or infinite population.

A **correction factor** is necessary for computing the standard error of the mean for samples drawn without replacement from a finite population.



The correction factor is computed using the following formula:

$$\sqrt{\frac{N - n}{N - 1}}$$

where N is the population size and n is the sample size.

The standard error of the mean must be multiplied by the correction factor to adjust it for large samples taken from a small population.



The standard error for the mean must be adjusted when it is included in the formula for calculating the z values.

$$\frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}}$$



A normal distribution is often used to solve problems that involve the binomial distribution since when n is large (say, 100), the calculations are too difficult to do by hand using the binomial distribution.





The normal approximation to the binomial is appropriate when $np \geq 5$ and $nq \geq 5$.

In addition, a **correction for continuity** may be used in the normal approximation to the binomial.

The continuity correction means that for any specific value of X , say 8, the boundaries of X in the binomial distribution (in this case, 7.5 to 8.5) must be used.



Binomial	Normal
When finding:	Use:
$P(X = a)$	$P(a - 0.5 < X < a + 0.5)$
$P(X \geq a)$	$P(X > a - 0.5)$
$P(X > a)$	$P(X > a + 0.5)$
$P(X \leq a)$	$P(X < a + 0.5)$
$P(X < a)$	$P(X < a - 0.5)$

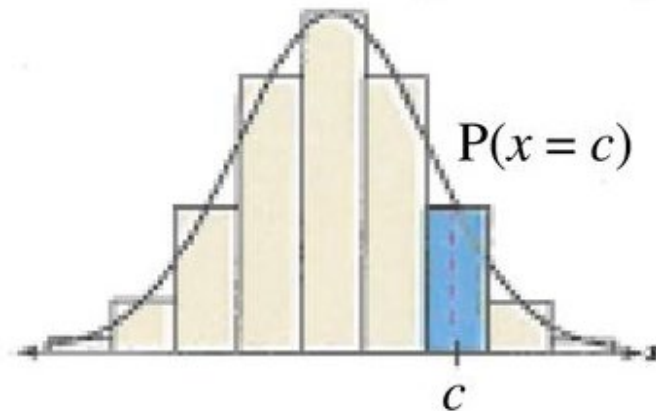
For all cases, $\mu = np$, $\sigma = \sqrt{npq}$, $np \geq 5$, $nq \geq 5$

Correction for continuity

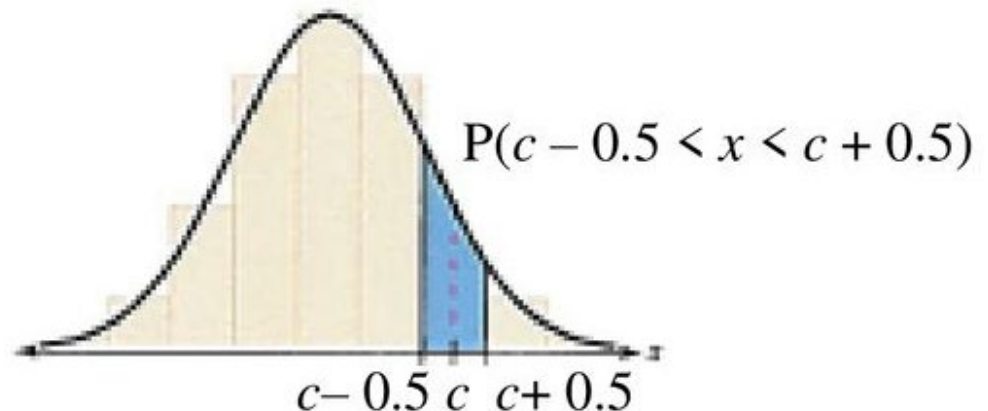


When using a continuous normal distribution to approximate a binomial probability, you need to move 0.5 unit to the left and right of the midpoint to include all possible x -values in the interval.

Exact binomial probability



Normal approximation



A magazine reported that 6% of American drivers read the newspaper while driving. If 300 drivers are selected at random, find the probability that exactly 25 say they read the newspaper while driving.

Here, $p = 0.06$, $q = 0.94$, and $n = 300$.

Step 1: Check to see whether a normal approximation can be used.

$$np = (300)(0.06) = 18 \text{ and } nq = (300)(0.94) = 282$$

Since $np \geq 5$ and $nq \geq 5$, we can use the normal distribution.

Step 2: Find the mean and standard deviation.

$$\mu = np = (300)(0.06) = 18$$

$$\sigma = \sqrt{npq} = \sqrt{300(0.06)(0.94)} = 4.11$$

Example: Reading While Driving



Step 3: Write in probability notation. $P(X = 25)$

Step 4: Rewrite using the continuity correction factor.

$$P(24.5 < X < 25.5)$$

Step 5: Find the corresponding z values.

$$z = \frac{24.5 - 18}{4.11} = 1.58, \quad z = \frac{25.5 - 18}{4.11} = 1.82$$

Step 6: Find the solution

The area between the two z values is $0.9656 - 0.9429 = 0.0227$, or 2.27%.

Hence, the probability that exactly 25 people read the newspaper while driving is 2.27%.

Example: Widowed Bowlers



Of the members of a bowling league, 10% are widowed. If 200 bowling league members are selected at random, find the probability that 10 or more will be widowed.

Here, $p = 0.10$, $q = 0.90$, and $n = 200$.

Step 1: Check to see whether a normal approximation can be used.

$$np = (200)(0.10) = 20 \text{ and } nq = (200)(0.90) = 180$$

Since $np \geq 5$ and $nq \geq 5$, we can use the normal distribution.

Step 2: Find the mean and standard deviation.

$$\mu = np = (200)(0.10) = 20$$

$$\sigma = \sqrt{npq} = \sqrt{200(0.10)(0.90)} = 4.24$$

Example: Widowed Bowlers



Step 3: Write in probability notation. $P(X \geq 10)$

Step 4: Rewrite using the continuity correction factor.

$$P(X > 9.5)$$

Step 5: Find the corresponding z values.

$$z = \frac{9.5 - 20}{4.24} = -2.48$$

Step 6: Find the solution

The area to the right of the z value is $1.0000 - 0.0066 = 0.9934$, or 99.34%.

The probability of 10 or more widowed people in a random sample of 200 bowling league members is 99.34%.