Statistical Foundations Unit 2:

## **Central limit theorem**

The [central limit theorem](https://www.scribbr.com/statistics/central-limit-theorem/) is the basis for how normal distributions work in statistics.

In research, to get a good idea of a[population](https://www.scribbr.com/methodology/population-vs-sample/) mean, ideally you’d collect data from multiple [random samples](https://www.scribbr.com/methodology/simple-random-sampling/) within the population. A **sampling distribution of the mean** is the distribution of the means of these different samples.

The central limit theorem shows the following:

* Law of Large Numbers: As you increase sample size (or the number of samples), then the sample mean will approach the population mean.
* With multiple large samples, the sampling distribution of the mean is normally distributed, even if your original variable is not normally distributed.

For example, a **population**follows a [**Poisson distribution**](https://www.scribbr.com/statistics/poisson-distribution/) (left image). If we take 10,000 **samples**from the population, each with a sample size of 50, the sample means follow a normal distribution, as predicted by the **central limit theorem**(right image).

A picture containing diagram, plot, screenshot, line

Description automatically generated

The central limit theorem relies on the concept of a **sampling distribution**, which is the [probability distribution](https://www.scribbr.com/statistics/probability-distributions/) of a **statistic**for a large number of [samples](https://www.scribbr.com/methodology/population-vs-sample/) taken from a population.

Imagining an experiment may help you to understand sampling distributions:

* Suppose that you draw a [random sample](https://www.scribbr.com/methodology/simple-random-sampling/) from a population and calculate a [statistic](https://www.scribbr.com/statistics/parameter-vs-statistic/) for the sample, such as the mean.
* Now you draw another random sample of the same size, and again calculate the [mean](https://www.scribbr.com/statistics/mean/).
* You repeat this process many times, and end up with a large number of means, one for each sample.

The distribution of the sample means is an example of a **sampling distribution.**

The central limit theorem says that the sampling distribution of the mean will always be **normally distributed**, as long as the sample size is large enough. Regardless of whether the population has a normal, Poisson, binomial, or any other distribution, the sampling distribution of the mean will be normal.

A normal distribution is a symmetrical, bell-shaped distribution, with increasingly fewer observations the further from the center of the distribution.

## **Central limit theorem formula**

Fortunately, you don’t need to actually repeatedly sample a population to know the shape of the sampling distribution. The [parameters](https://www.scribbr.com/statistics/parameter-vs-statistic/) of the sampling distribution of the mean are determined by the parameters of the population:

* The [mean](https://www.scribbr.com/statistics/mean/) of the sampling distribution is the mean of the population.

\begin{equation*}\mu_{\bar{x}}=\mu\end{equation*}

* The [standard deviation](https://www.scribbr.com/statistics/standard-deviation/) of the sampling distribution is the standard deviation of the population divided by the square root of the sample size.

\begin{equation*}\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}\end{equation*}

We can describe the sampling distribution of the mean using this notation:

\begin{equation*}\bar{X}\sim N (\mu,\dfrac{\sigma}{\sqrt{n}})\end{equation*}

Where:

* X̄ is the sampling distribution of the sample means
* ~ means “follows the distribution”
* N is the [normal distribution](https://www.scribbr.com/statistics/normal-distribution/)
* µ is the mean of the population
* σ is the standard deviation of the population
* n is the sample size

**Central Limit Theorem: Review**

1. The distribution of sample x¯ 's will, as the sample size increases, approach a normal distribution.
2. The mean of the sample means is the population mean μ.



1. The standard deviation of the distribution of the sample mean is often called the standard error of the mean.



**Important Rules**

1. For samples sizes of *n* larger than 30, the distribution of the sample means can be approximated reasonably well by a normal distribution. The approximation gets better as the sample size *n* becomes larger.
2. If the original population is itself normally distributed, then the sample means will be normally distributed for any sample size n (not just the values of n larger than 30).

### 1. Sample size and normality

The larger the sample size, the more closely the sampling distribution will follow a [normal distribution](https://www.scribbr.com/statistics/normal-distribution/).

When the sample size is small, the sampling distribution of the mean is sometimes non-normal. That’s because the central limit theorem only holds true when the sample size is “sufficiently large.”

By convention, we consider a sample size of 30 to be “sufficiently large.”

* **When n < 30**, the central limit theorem doesn’t apply. The sampling distribution will follow a similar distribution to the population. Therefore, the sampling distribution will only be normal if the population is normal.
* **When n ≥ 30**, the central limit theorem applies. The sampling distribution will approximately follow a normal distribution.

### 2. Sample size and standard deviations

The sample size affects the standard deviation of the sampling distribution. Standard deviation is a measure of the [variability](https://www.scribbr.com/statistics/variability/) or spread of the distribution (i.e., how wide or narrow it is).

* **When n is low**, the standard deviation is high. There’s a lot of spread in the samples’ means because they aren’t precise estimates of the population’s mean.
* **When n is high**, the [standard deviation](https://www.scribbr.com/statistics/standard-deviation/) is low. There’s not much spread in the samples’ means because they’re precise estimates of the population’s mean.

**Central Limit Theorem: Recap**

1. As *n* gets larger, no matter what the distribution of the individuals, the sampling distribution of the sample mean becomes more normal.
2. The mean of the sampling distribution of the sample mean is equivalent to the mean of the individuals.



1. The standard deviation of the sampling distribution of the sample mean is often called the standard error of the mean.



A picture containing text, clock, screenshot, design

Description automatically generatedA screenshot of a graph

Description automatically generated with low confidence

## Central limit theorem examples

Applying the central limit theorem to real distributions may help you to better understand how it works.

### Continuous distribution

Suppose that you’re interested in the age that people retire in the United States. The [**population**](https://www.scribbr.com/methodology/population-vs-sample/) is all retired Americans, and the distribution of the population might look something like this:

A picture containing text, screenshot, diagram, plot

Description automatically generated

Age at retirement follows a [left-skewed](https://www.scribbr.com/statistics/skewness/#left-skew) distribution. Most people retire within about five years of the mean retirement age of 65 years. However, there’s a “long tail” of people who retire much younger, such as at 50 or even 40 years old. The population has a standard deviation of 6 years.

Imagine that you take a small **sample**of the population. You randomly select five retirees and ask them what age they retired.

Example: Central limit theorem; sample of n = 5

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 68 | 73 | 70 | 62 | 63 |

The mean of the sample is an [estimate](https://www.scribbr.com/statistics/parameter-vs-statistic/#estimating-parameters-from-statistics) of the population mean. It might not be a very precise estimate, since the sample size is only 5.

Example: Central limit theorem; mean of a small samplemean = (68 + 73 + 70 + 62 + 63) / 5

mean = 67.2 years

Suppose that you repeat this procedure 10 times, taking samples of five retirees, and calculating the mean of each sample. This is a **sampling distribution of the mean**.

Example: Central limit theorem; means of 10 small samples

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 60.8 | 57.8 | 62.2 | 68.6 | 67.4 | 67.8 | 68.3 | 65.6 | 66.5 | 62.1 |

If you repeat the procedure many more times, a histogram of the sample means will look something like this:

A picture containing text, diagram, screenshot, plot

Description automatically generated

Although this sampling distribution is more normally distributed than the population, it still has a bit of a [left skew](https://www.scribbr.com/statistics/normal-distribution/).

Notice also that the spread of the sampling distribution is less than the spread of the population.

The **central limit theorem** says that the sampling distribution of the mean will always follow a normal distribution when the sample size is sufficiently large. This sampling distribution of the mean isn’t normally distributed because its sample size isn’t sufficiently large.

Now, imagine that you take a large sample of the population. You randomly select 50 retirees and ask them what age they retired.

Example: Central limit theorem; sample of n = 50

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 73 | 49 | 62 | 68 | 72 | 71 | 65 | 60 | 69 | 61 |
| 62 | 75 | 66 | 63 | 66 | 68 | 76 | 68 | 54 | 74 |
| 68 | 60 | 72 | 63 | 57 | 64 | 65 | 59 | 72 | 52 |
| 52 | 72 | 69 | 62 | 68 | 64 | 60 | 65 | 53 | 69 |
| 59 | 68 | 67 | 71 | 69 | 70 | 52 | 62 | 64 | 68 |

The mean of the sample is an [estimate](https://www.scribbr.com/statistics/parameter-vs-statistic/#estimating-parameters-from-statistics) of the population mean. It’s a precise estimate, because the sample size is large.

Example: Central limit theorem; mean of a large samplemean = 64.8 years

Again, you can repeat this procedure many more times, taking samples of fifty retirees, and calculating the mean of each sample:

A picture containing text, screenshot, line, diagram

Description automatically generated

In the histogram, you can see that this sampling distribution is normally distributed, as predicted by the central limit theorem.degree

The standard deviation of this sampling distribution is 0.85 years, which is less than the spread of the small sample sampling distribution, and much less than the spread of the population. If you were to increase the sample size further, the spread would decrease even more.

We can use the central limit theorem formula to describe the sampling distribution:

\bar{X} \sim N (\mu,\dfrac{\sigma}{\sqrt{n}})

µ = 65

σ = 6

n = 50

\bar{X} \sim N (65,\dfrac{6}{\sqrt{50}})

**Introducing Confidence Intervals**

* Confidence intervals measure how large or small a parameter value could be.
* Example: Estimate a population mean *μ*from a sample.
  + If we take many multiple samples, we will see slightly different estimates of the mean.
  + The different estimates are a natural result of randomization.
  + Confidence intervals tell us where those answers should be.

**Samples From a Normally Distributed Population**

* Suppose 300 students each draw a random sample of 10 scores from a normally distributed population whose actual mean *μ* = 10 and *σ* = 1.5.
* Normal distribution theory says we should expect:
  + ~68% of the drawn values will be within 1*σ* of *μ*
  + ~95% of the drawn values will be within 2*σ* of *μ*
  + ~99.7% of the drawn values will be within 3*σ* of *μ*
* If we have appropriately obtained sample means, we can say:
  + 95% of the sample mean (X– ) values will be fairly close to the actual population mean (*μ*).
  + 5% of the X– values will be far away from *μ*.

**What the Confidence Interval Tells Us**

* A 95% confidence interval means that, on average, we expect the endpoints of this particular interval to contain the population mean *μ* in 95% of random samples.
  + Remember: The measure is not necessarily confidence in the interval itself but in the procedure that produces the interval.
* In other words, if we draw repeated samples of size *n* from a normally distributed population, 95% of the sample means (X– ) should lie within 1.96 standard errors of the actual *μ*.

**Estimating the Confidence Interval**

* To set up a confidence level estimate, we need to know:
  + A desired *level of confidence* (e.g., 95%, which is the most popular)
  + An estimate of the population's standard deviation
  + The sample mean X–�–
  + The sample size *n*

**Estimating the Confidence Interval**

* To set up a confidence level estimate, we need to know:
  + A desired *level of confidence* (e.g., 95%, which is the most popular)
  + An estimate of the population's standard deviation
  + The sample mean X–
  + The sample size *n*
* <To calculate the confidence, we use this formula:

         CI=X–±tcrit(sn√)

**The Confidence Interval Formula**

* Note that we use *s* instead of *σ*, where *s* is the sample standard deviation.
* *s* is our estimate of *σ*.
* This substitution introduces error.
* So, confidence interval calculated with *t*-distribution instead of Normal distribution.
* *t*-distribution allows for error in estimate of standard deviation.
* *t*-distribution is like Normal distribution but with fatter tails that allow for error.

**Example: Emotional Intelligence Scores**

* An estimate of the emotional intelligence (EI) of a random sample of children at a school for the gifted and talented
* The researcher's test:

**Example: Emotional Intelligence Scores**

* An estimate of the emotional intelligence (EI) of a random sample of children at a school for the gifted and talented
* The researcher's test:
  + The sample mean X– = 130
  + The sample standard deviation *s* = 15
  + The sample size *n* = 120
* The confidence interval formula:

         CI=130±tcrit(sn√)

**A Critical Value for the *t*-Distribution**

* The *tcrit* value depends on your chosen confidence interval (e.g., 90%, 95%, 99%, etc.).
* Look up the *tcrit* values in a table if you're calculating by hand.
* *tcrit*is a function of:
  + the degrees of freedom (defined as *n* − 1)
  + the confidence level (which you choose)
* In our example, for a 95% confidence interval and *n*=120:
  + From a table, *tcrit* is 1.979.
  + Note that 1.979 is close to the 1.96 we'd get if we were using a Normal distribution.
  + The values are close because we have a large sample size.

**Calculate the Confidence Interval**

* The confidence interval is:

         CI=130±1.979(15120√)         CI=130±1.979(15120)  
         CI=130±2.71         CI=130±2.71  
         CI=127.29 to 132.71         CI=127.29 to 132.71

* We could say:
  + "95% of the confidence intervals estimated with this procedure for random samples from the same population will contain the population mean *μ*."
  + "The other 5% of confidence intervals will not contain *μ*."
* We should *not*say:
  + "There is a 95% chance that *μ* lies in this interval."
* What we do say:
  + "With 95% confidence, the true mean *μ* is between 127.29 and 132.71, based on a sample *n* = 120 from this population."

If we took many, many samples from the same population and constructed a confidence interval for the mean from each one, then approximately 95 percent of those intervals would contain the true mean. We hope that we have contained the mean in our interval, and thus we are 95 percent confident that our interval contains the mean.

**Caveats**

* The value of the interval—whether it's really a 95% interval, and whether it's really a good estimate of the true population mean *μ*—depends on:
  + How good the random sampling procedure is
  + A screenshot of a business

    Description automatically generatedA screenshot of a math problem

    Description automatically generated with medium confidenceHow representative the students are of the total population

A maths graph on a white background

Description automatically generated with medium confidence

**Hypothesis Test**

* We use a formal hypothesis test to determine whether a sample supports a hypothesized parameter value.
* Recall:
  + We used the central limit theorem to estimate a range of likely values for μ using confidence intervals.
  + We can also use the central limit theorem to estimate a sample distribution.
* Example hypothesis test:
  + Is the normal human body temperature 98.6°F?
  + Can we support this hypothesis by checking a random sample of healthy individuals?

**Example Test: Emotional Intelligence**

* Recall the emotional intelligence (EI) example from earlier.
* What if we knew the μ for regular students (not gifted and talented)?
* How do we determine whether the EI of gifted and talented students is different from the regular students?
* We can do a two-sample test:
  + Sample the EI of gifted and talented students
  + Sample the EI of regular students
  + Is the difference 0, or is it larger than expected by chance?

**Example Test: Sodium Content**

* Quality control at the potato chip factory
* Is your sodium content as advertised?
* To test:
  + Take a random sample from the assembly line.
  + Measure the sodium content.
  + Is the sample mean X– close enough to the advertised target?
  + Is the difference small enough to be chance rather than a problem with the machine?

**Significance Test Conditions**

1. Correct significance level α.
   * α is typically 0.05.
   * α sets probability of **type I error** (when true *H*0 is rejected).
2. Choose **directional** or **nondirectional**test (before looking at data).
   * Directional test: Is the sample value is greater than expected or less than expected?
   * Nondirectional test: Is the sample value different from expected?
   * "One-sided" and "two-sided" tests are alternative names for directional and nondirectional tests.
3. Data must meet required assumptions:
   * Quantitative scores
   * Measured reliably
   * Nearly normally distributed
   * Outliers examined
4. Samples must be drawn randomly and/or representatively.
5. Perform one significance test at a time.

**State the Hypothesis**

* Two types: the null hypothesis (*H*0) and the alternative hypothesis (*H*a)
* A **hypothesis** is an assumption about the characteristics of one or more variables in one or more populations
* **Null hypothesis:**the statement being tested
  + "The mean human body temperature is 98.6°F."
* **Alternative hypotheses:**the statement we suspect is true instead of the null hypothesis
  + The true value is >98.6.
  + The true value is <98.6.
  + The true value is ≠ 98.6 (in either direction).
  + Note: The first two alternative hypotheses are one sided; the third is two sided.

**The *p*-Value**

* ***p*-value:** the chance of obtaining a particular sample result if the null hypothesis is true
* We want a *p*-value that tells us—based on the condition of the test—whether to reject the *H*0.
* We always express our result in terms of whether or not we reject the *H*0.
* The *p*-value tells us whether random variation alone can account for the difference between between the null hypothesis and observations in the random sample.
* Given the sample size and the standard deviation, is the *p*-value close enough to the expected value to support the null hypothesis?

**Interpreting the *p*-Value**

* We calculate the *p*-value with the assumption that the *H*0 is true.
* The *p*-value is a probability between 0 and 1.
* A small *p*-value implies random variation is *not*likely to account for the differences we see:
  + Evidence *against* the *H*0
* A large *p*-value suggests that the *H*0 is valid.

**How Small Is Small?**

* "Small" depends on the significance level you choose.
* A typical significance level is 0.05.
  + If the *p*-value is much smaller than 0.05, you reject the *H*0.
  + If the *p*-value is 0.05 or larger, you do not reject the *H*0.
* What if the *p*-value is very close to 0.05?
  + If the *p*-value is 0.049, do you reject the *H*0?
  + What if the *p*-value is 0.051, which is not much larger than 0.05?
* The *p*-value is a probability, not a hard cutoff.
* State the *p*-value and let the reader decide whether your evidence is close enough to reject the *H*0.

***p*-Value Example: Salt Content**

* We want the sodium content to be as advertised, therefore:
  + We want to support the *H*0 (i.e., we want to see a large *p*-value).
  + *n* = 40
  + *H*0: μ = 75
  + *H*a: μ ≠ 75 **OR** μ < 75 **OR** μ > 75
* Ha: μ ≠ 75 (nondirectional test)
  + X– = 76.1 and *s* = 2.2878
* Is this consistent with our advertised claim (i.e., consistent with *H*0)?
  + *t* = 2.95, which yields *p*-value = 0.0052.
  + 0.5% chance (52 out of 10,000) we would see a value as large as 76.1 out of a sample of 40 if the *H*0 were true.
* We reject the null hypothesis (i.e., the *p*-value is too small to support *H*0).
* We need to fix the manufacturing process.

**Reminder About Conditions**

* Conditions must be met.
  + The data must be a random sample.
  + The data must be close to normally distributed without too many outliers.

**One-Sided and Two-Sided Tests**

* A **one-tail** or **one-sided**test of a population mean has these null and alternative hypotheses:
  + H0:μ=[a specific number]     Ha:μ<[a specific number]     OR�0:μ=[a specific number]     �a:μ<[a specific number]     OR
  + H0:μ=[a specific number]     Ha:μ>[a specific number]�0:μ=[a specific number]     �a:μ>[a specific number]
* A **two-tail** or **two-sided** test of a population mean has these null and alternative hypotheses:
  + H0:μ=[a specific number]     Ha:μ≠[a specific number]

**Calculating *p*-Value**

* We compare the *p*-value with a level that we regard as decisive called the **significance level**, α. Usually, α = 0.05.
* Everything is calculated on the basis of the *t*-distribution.
* The ***p*-value** is the probability, if *H*0 is true, of randomly drawing a sample like the one obtained or more extreme, in the direction of *H*α.
* The ***p*-value** is calculated as the corresponding area under the curve, one-tailed or two-tailed depending on *H*α.
  + I.e., the areas to the right and the left of the *t*-statistic
* Recall the *t*-statistic formula:  
         t=(x–−μ0)(s/n√)

   t=(x–−μ0)(s/n√)=(76.1−75)(2.287/40√)=3.04

### Two-Sided p-Value

A diagram of a normal distribution

Description automatically generated

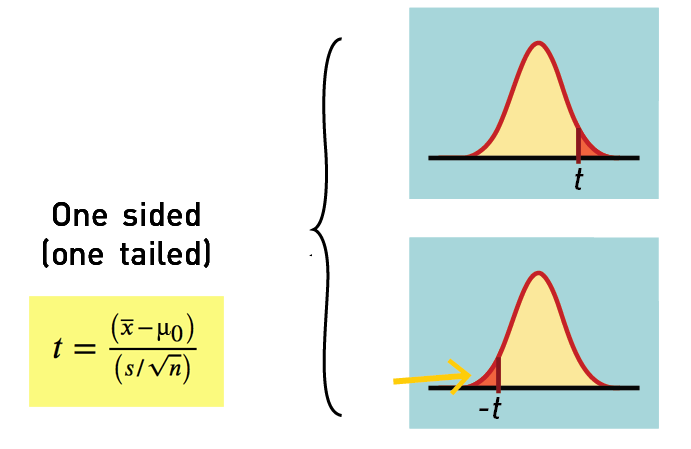
**Two-Sided *p*-Value**

A black and yellow logo

Description automatically generated with medium confidence

* The *t*-values are marked on the *x*-axis (in our case, 2.95 on either side of the center).
* The *H*a probability corresponds to the area under both tails.
* The software tells us the total area under both right and left tails = 0.0052.
  + The area under either single tail is 0.0052/2 because the function is symmetric.
* Rephrased, we want the areas to the left of −2.95 and to the right of +2.95.
* These two areas total 0.0052, which is our *p*-value.

**One-Sided *p*-Value**



* If your *H*a is that the true mean is >75 mg, the *p*-value is represented by the shaded area right of *t*in the upper graph.
  + This *p*-value is 0.0026 (the two-sided *p*-value, 0.0052, divided by 2).
* If your *H*a is that the true mean is < 75 mg, the *p*-value is represented by the shaded area left of *t*in the bottom graph.
  + *p* = 1 − 0.0026 = 0.9974
  + We would *not* reject the *H*0 in that case.
  + But here the area left of the *t* is actually very small.

**Keep Track of Signs**

* For a one-sided hypothesis, pay attention to the sign of the numerator.
  + A negative value indicates that X– is smaller than μ0.
  + If *H*a: μ < μ0, you want to see a *negative* value.
  + If *H*a: μ > μ0, you want to see a *positive* value.
* It's always the sample mean minus the population mean in the numerator.

**Summary**

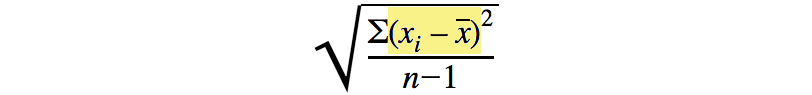
* Remember: Software gives the two-sided *p*-value.
  + For the one-sided *p*-value, divide by two.
* Remember: The difference between the sample mean and hypothesized mean keeps with the direction of the alternative hypothesis.

**Standard Deviation**

The **standard deviation** is the average amount of [variability](https://www.scribbr.com/statistics/variability/) in your dataset. It tells you, on average, how far each value lies from the mean.

A high standard deviation means that values are generally far from the [mean](https://www.scribbr.com/statistics/mean/), while a low standard deviation indicates that values are clustered close to the mean.

* How far each value is from the center (i.e., the mean)



### Sample Standard Deviation

* The sample standard deviation notation is s.
* Standard deviation of a set of sample values is a measure of variation of values about the mean.

The **sample standard deviation** formula looks like this:

| **Formula** | **Explanation** |
| --- | --- |
| A picture containing screenshot, black  Description automatically generated | * s = sample standard deviation * \sum = sum of… * X = each value * \bar{x} = sample mean * n = number of values in the sample |

**Standard Deviation: Key Points**

* Standard deviation is a measure of variation of all values from the mean.
* It is sensitive to extreme values.
* The value of the standard deviation *s* is positive.
* The value of s can increase dramatically with the inclusion of one or more outliers (data values far away from all others).
* The units of s are the same as the units of the original data values.

**The Population Standard Deviation**

* The population standard deviation is σ (lower case sigma).
* Sample standard deviation (*s*) estimates population standard deviation (σ).

The**population standard deviation** formula looks like this:

| **Formula** | **Explanation** |
| --- | --- |
|  | * \sigma = population standard deviation * \sum = sum of… * X = each value * \mu = population mean * N = number of values in the population |

**Variance**

The **variance** is a measure of [variability](https://www.scribbr.com/statistics/variability/). It is calculated by taking the average of squared deviations from the mean.

Variance tells you the degree of spread in your data set. The more spread the data, the larger the variance is in relation to the [mean](https://www.scribbr.com/statistics/mean/).

* A measure of variation equal to the square of the standard deviation
* Sample variance: *s2* (square of the sample standard deviation)
* Population variance: σ*2* (square of the population standard deviation)

Since the units of variance are much larger than those of a typical value of a data set, it’s harder to interpret the variance number intuitively. That’s why standard deviation is often preferred as a main measure of variability.

However, the variance is more informative about variability than the standard deviation, and it’s used in making [statistical inferences](https://www.scribbr.com/statistics/inferential-statistics/).

## Steps for calculating the standard deviation by hand

The standard deviation is usually calculated automatically by whichever software you use for your statistical analysis. But you can also calculate it by hand to better understand how the formula works.

There are six main steps for finding the standard deviation by hand. We’ll use a small data set of 6 scores to walk through the steps.

| **Data set** | | | | | |
| --- | --- | --- | --- | --- | --- |
| 46 | 69 | 32 | 60 | 52 | 41 |

### **Step 1**: Find the mean

To [find the mean](https://www.scribbr.com/statistics/mean/), add up all the scores, then divide them by the number of scores.

| **Mean (x̅)** |
| --- |
| \bar{x} = \dfrac{(46 + 69 + 32 + 60 + 52 + 41)}{6} = 50 |

### **Step 2**: Find each score’s deviation from the mean

Subtract the mean from each score to get the deviations from the mean.

Since x̅ = 50, here we take away 50 from each score.

| **Score** | **Deviation from the mean (x1 -****x̅)** |
| --- | --- |
| 46 | 46 – 50 = **-4** |
| 69 | 69 – 50 = **19** |
| 32 | 32 – 50 = **-18** |
| 60 | 60 – 50 = **10** |
| 52 | 52 – 50 = **2** |
| 41 | 41 – 50 = **-9** |

### **Step 3**: Square each deviation from the mean

Multiply each deviation from the mean by itself. This will result in positive numbers.

| **Squared deviations from the mean (x1 -****x̅) 2** |
| --- |
| (-4)2 = 4 × 4 = **16** |
| 192 = 19 × 19 = **361** |
| (-18)2 = -18 × -18 = **324** |
| 102 = 10 × 10 = **100** |
| 22 = 2 × 2 = **4** |
| (-9)2 = -9 × -9 = **81** |

### **Step 4**: Find the sum of squares

Add up all of the squared deviations. This is called the sum of squares.

| **Sum of squares** |
| --- |
| 16 + 361 + 324 + 100 + 4 + 81 = **886** |

### **Step 5:** Find the variance

Divide the sum of the squares by n – 1 (for a[sample](https://www.scribbr.com/methodology/population-vs-sample/)) or N (for a population) – this is the [variance](https://www.scribbr.com/statistics/variance/).

Since we’re working with a sample size of 6, we will use  n – 1, where n = 6. You only subtract 1 in the sample. With the population, use the whole of n.

**Variance**

### **Step 6**: Find the square root of the variance

To find the standard deviation, we take the square root of the variance.

\dfrac{886}{(6 - 1)} = \dfrac{886}{5} = 177.2A mathematical equation with black text

Description automatically generated

A picture containing screenshot, black

Description automatically generated**Standard Deviation**

\sqrt{177.2} = 13.31**A picture containing line, plot, diagram, slope

Description automatically generated**

From learning that SD = 13.31, we can say that each score deviates from the mean by 13.31 points on average.

Range = 37 R = H – L (37 = 69 -32)

S = 13.31

Sigma σ = √886/6 = √147.6 = 12.15

S2 = 177.2

σ*2 = 147.6*

**Introduction**

* **Statistical significance** means a *p*-value less than some chosen α (e.g., 0.05).
* We tend to think statistical significance implies there is something more than random sampling error.
* Small *p*-values suggest results have meaning within the study context.
  + This is usually true but sometimes false.
* Statistical significance is not the same as **practical significance**.

**Sample Size Influences *p*-Values**

* One challenge: the *p*-value is related to sample size.
* A small difference between the hypothesized mean (μ0) and sample mean ( X–�– ) can yield a small *p*-value, even if the difference is *practically* insignificant.
  + A large sample size (*n*) can cause this erroneous conclusion.
  + Conversely, a small *n* can cause the opposite, erroneous conclusion.
* A large difference between μ0 and X–�– can still yield a large *p*-value if *n*is small.
  + If *n* is too small, you might not have collected enough data to show the difference between the sample statistic and the hypothesized population parameter.

**Practical Significance**

* One aspect of practical significance is **effect size**.
* Effect size relates to the practical value of the effect.
* Example: emotional intelligence study
  + μ0 = 120
  + *n* = 130
  + X– = 122
  + The *p*-value is very close to zero.
  + Difference between X– (122) and μ0 (120) is just 2 points.
  + Is this practically significant? What does this mean in terms of EI?
  + If not practically significant, statistical significance doesn't matter.

**A Practical Significance Benchmark**

* What makes a value practically significant?
* For emotional intelligence scores, we might review the literature:
  + Imagine that research says a 5-point EI difference can affect a person's ability to function emotionally.
  + This provides a benchmark for practical significance.
  + The benchmark tells us that our 2-point difference isn't meaningful, regardless of the small *p*-value.

**Effect Size and Cohen's *d***

* Effect size is a way to measure practical significance.
* **Cohen's *d*** is a statistical method that allows us to calculate an effect size.

         d=∣∣(x–−μ0)s∣∣

* For the EI example:

         d=∣∣(122−120)15∣∣=0.133

* Cohen provided a framework to categorize effects as small, medium, or large.
* *d*≤ 0.2 signifies a "small" effect.
  + Our *d* = 0.133 is below that threshold, so it is a "small" effect.
  + Any statistical significance might be the result of a large *n*.
* *d*> 0.8 signifies a "large" effect.
  + Even without statistical significance, *d* ≥ 0.8 suggests a large effect.
* 0.2 < *d*< 0.8 signifies a "medium" effect.
  + This doesn't clearly suggest whether an effect is practically significant.

**Summary**

* A statistically significant *p*-value does not imply useful results.
* We want useful results to help us make decisions about practical issues.