Introduction to Hypothesis Testing

Agenda

- Identify the four steps of hypothesis testing.
- Define null hypothesis, alternative hypothesis, level of significance, test statistic, p value, and statistical significance.
- Define Type I error and Type II error, and identify the type of error that researchers control.
- Calculate the one-independent sample z test and interpret the results.
- Distinguish between a one-tailed and two-tailed test, and explain why a Type III error is possible only with one-tailed tests.
- Explain what effect size measures and compute a Cohen's d for the one-independent sample z test.
- Define power and identify six factors that influence power.
- Summarize the results of a one-independent sample z test in American Psychological Association (APA) format.

Definition

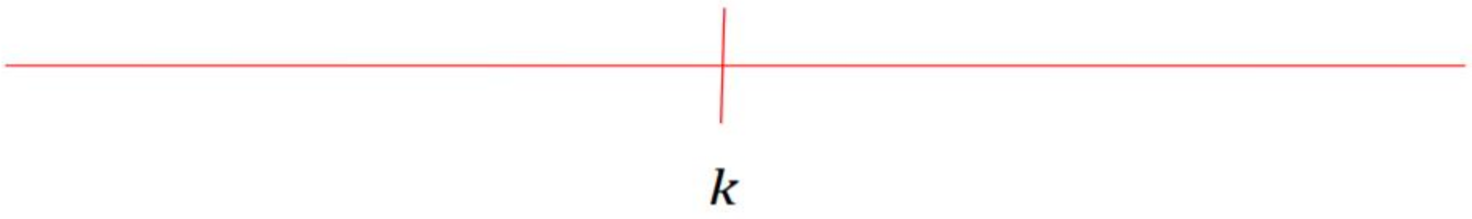
- Hypothesis testing or significance testing is a method for testing a claim or hypothesis about a parameter in a population, using data measured in a sample.
- In this method, we test some hypothesis by determining the likelihood that a sample statistic could have been selected, if the hypothesis regarding the population parameter were true.
- "The main goal in many research studies is to check whether the data collected support certain statements or predictions"
- Hypothesis testing is the method of testing whether claims regarding a population are likely to be true.

Types of hypothesis

- Hypothesis is broadly classified in to Two.
- A) Null Hypothesis: "The null hypothesis ($H0$), stated as the null, is a statement about a population parameter, such as the population mean, that is assumed to be true.
- The null hypothesis is a starting point. We will test whether the value stated in the null hypothesis is likely to be true.
- B) Alternate Hypothesis: "An alternative hypothesis ($H1$) is a statement that directly contradicts a null hypothesis by stating that that the actual value of a population parameter is less than, greater than, or not equal to the value stated in the null hypothesis.
- The alternative hypothesis states what we think is wrong about the null hypothesis."
- Note: H0 will ALWAYS have an equal sign (and possibly a less than or greater than symbol, depending on the alternative hypothesis).

How to differentiate Tests

Two-tailed test	Right-tailed test	Left-tailed test
$H_0: \mu = k$ $H_1: \mu \neq k$	$H_0: \mu = k$ $H_1: \mu > k$	$H_0: \mu = k$ $H_1: \mu < k$



Examples: State the H0 and H1 for each case

- A researcher thinks that if expectant mothers use vitamins, the birth weight of the babies will increase. The average birth weight of the population is 8.6 pounds

  H0: $\mu = 8.6$ H1: $\mu > 8.6$

- An engineer hypothesizes that the mean number of defects can be decreased in a manufacturing process of compact disks by using robots instead of humans for certain tasks. The mean number of defective disks per 1000 is 18.

Important on Null Hypothesis

- When a researcher conducts a study, he or she is generally looking for evidence to support a claim of some type of difference.
- In this case, the claim should be stated as the alternative hypothesis. Because of this, the alternative hypothesis is sometimes called the research hypothesis.

Important Definitions

- Statistical Test – uses the data obtained from a sample to make a decision about whether the null hypothesis should be rejected.
- Test Value (test statistic) – the numerical value obtained from a statistical test.

Type 1 and Type 2 Errors

- When we make a conclusion from a statistical test there are two types of errors that we could make. They are called: Type I and Type II Errors.
- Type I error – reject H0 when H0 is true.
- Type II error – do not reject H0 when H0 is false
- Results of a statistical test:

	H_0 is True	H_0 is False
Reject H_0	**Type I Error**	Correct Decision
Do not Reject H_0	Correct Decision	**Type II Error**

Example

- Example: Decision Errors in a Legal Trial . What are H0 and H1 ?
- H0: Defendant is innocent.
- H1: Defendant is not innocent, i.e., guilty
- If you are the defendant, which is the worse error? Why?
  - The decision of the jury does not prove that the defendant did or did not commit the crime.
  - The decision is based on the evidence presented.
  - If the evidence is strong enough the defendant will be convicted in most cases, if it is weak the defendant will be acquitted.
  - So the decision to reject the null hypothesis does not prove anything

The question is how large of a difference is enough to say we have enough evidence to reject the null hypothesis?

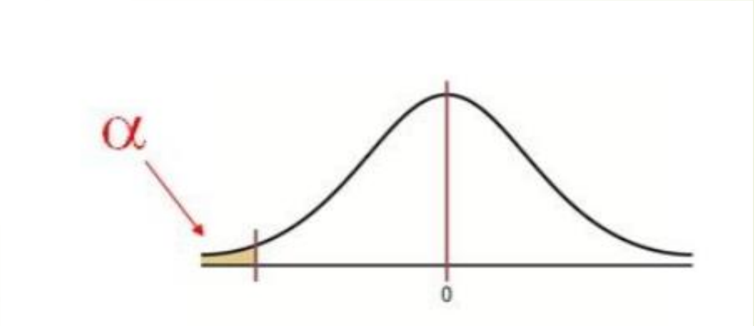
- Significance level - is the maximum probability of committing a Type I error. This probability is symbolized by Alpha   P(Type I error | H0 is true) = Alpha
- Critical or Rejection Region – the range of values for the test value that indicate a significant difference and that the null hypothesis should be rejected.
- Non-critical or Non-rejection Region – the range of values for the test value that indicates that the difference was probably due to chance and that the null hypothesis should not be rejected.

Some Important Definitions

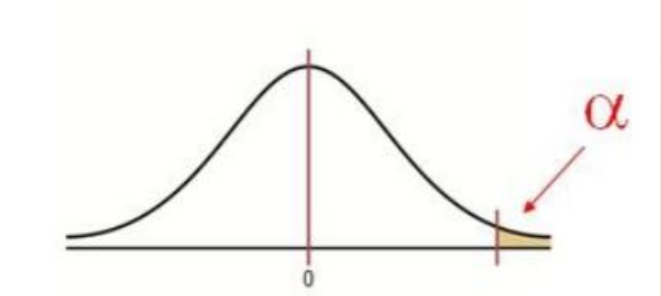
- Critical Value (CV) – separates the critical region from the non-critical region, i.e., when we should reject H0 from when we should not reject H0.
  - The location of the critical value depends on the inequality sign of the alternative hypothesis.
  - Depending on the distribution of the test value, you will use different tables to find the critical value.

Importance on test

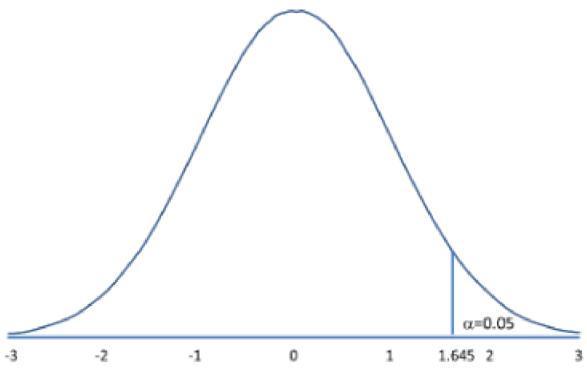
- One-tailed test – indicates that the null hypothesis should be rejected when the test value is in the critical region on one side.
  - Left-tailed test – when the critical region is on the left side of the distribution of the test value.



  - Right-tailed test – when the critical region is on the right side of the distribution of the test value.



Upper tail test Critical values and Z Score

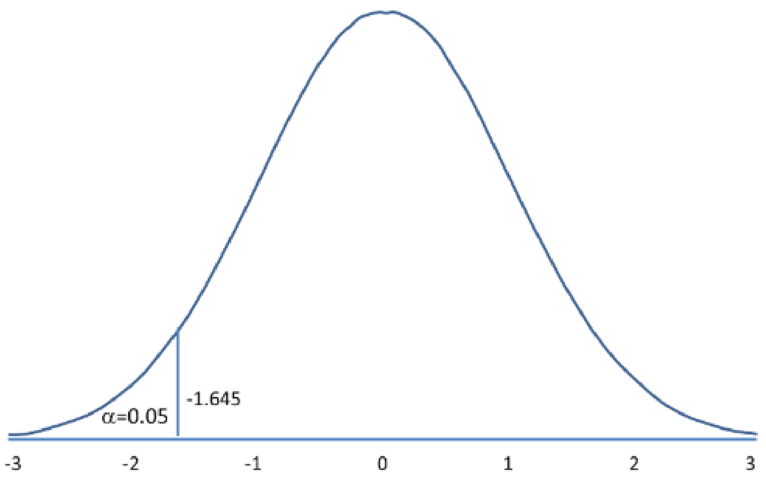


Rejection Region for Upper-Tailed Z Test (H_1: $\mu > \mu_0$) with α=0.05

The decision rule is: Reject H_0 if $Z \geq 1.645$.

Upper-Tailed Test	
α	Z
0.10	1.282
0.05	1.645
0.025	1.960
0.010	2.326
0.005	2.576
0.001	3.090
0.0001	3.719

Lower tail test Critical values and Z Score

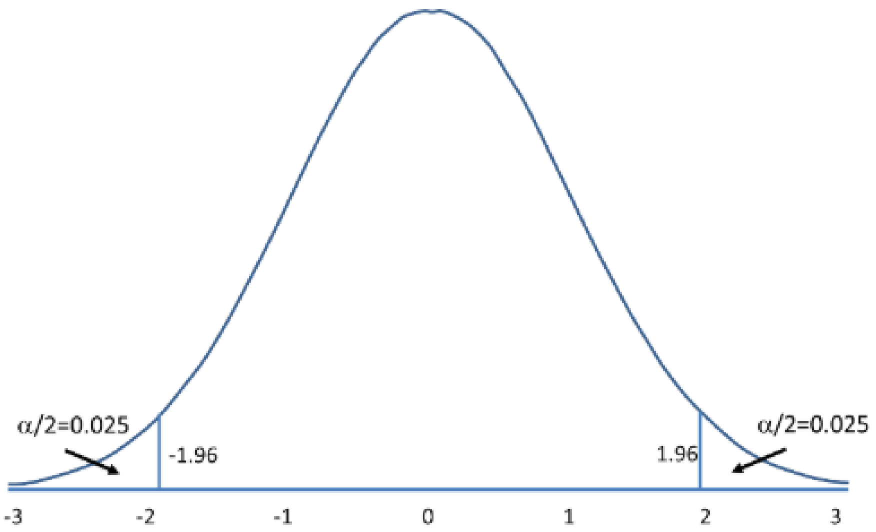


Rejection Region for Lower-Tailed Z Test (H_1: $\mu < \mu_0$) with α =0.05

The decision rule is: Reject H_0 if $Z \leq 1.645$.

Lower-Tailed Test	
a	Z
0.10	-1.282
0.05	-1.645
0.025	-1.960
0.010	-2.326
0.005	-2.576
0.001	-3.090
0.0001	-3.719

Two tailed test



Rejection Region for Two-Tailed Z Test (H_1: $\mu \neq \mu_0$) with α =0.05

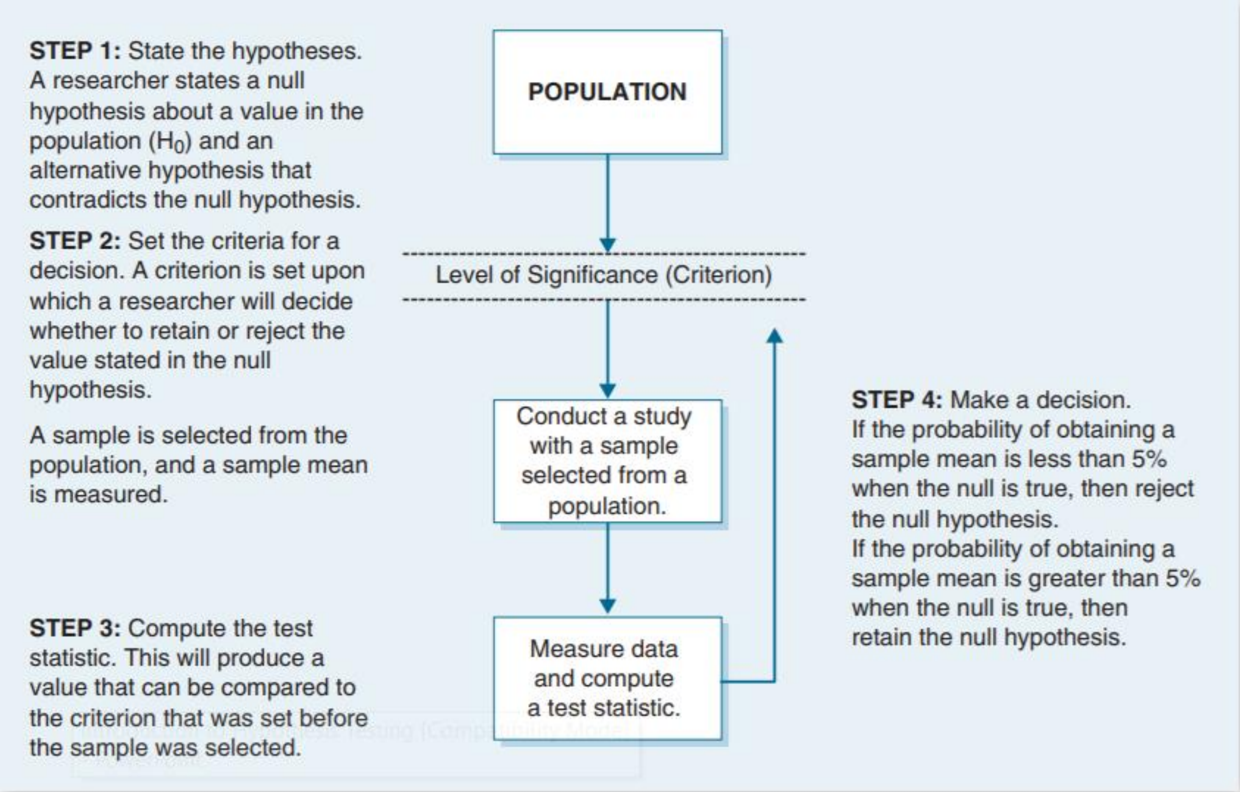
The decision rule is: Reject H_0 if $Z \leq -1.960$ or if $Z \geq 1.960$.

Two-Tailed Test	
α	**Z**
0.20	1.282
0.10	1.645
0.05	1.960
0.010	2.576
0.001	3.291
0.0001	3.819

Hypothesis Test Procedure (Traditional Method)

- Step 1 State the hypotheses and identify the claim.
- Step 2 Find the critical value(s) from the appropriate table.
- Step 3 Compute the test value.
- Step 4 Make the decision to reject or not reject the null hypothesis.
- Step 5 Summarize the results.

Hypothesis Flow



Test 1

- MOTIVATING SCENARIO: It has been reported that the average credit card debt for college seniors is $3262.
- The student senate at a large university feels that their seniors have a debt much less than this, so it conducts a study of 50 randomly selected seniors and finds that the average debt is $2995, and the population standard deviation is $1100.
- Can we support the student senate’s claim using the data collected.

How….the z Test for a Mean

- **A statistical test** uses the data obtained from a sample to make a decision about whether the null hypothesis should be rejected.
- The numerical value obtained from a statistical test is called the **test value**.
- You will notice that our statistical tests will resemble the general formula for a z-score:
- Test Value = observed value – expected value /standard error

The **z test** for Means

- The z test is a statistical test for the mean of a population.
- It can be used when $n \geq 30$, or when the population is normally distributed and σ is known.
- The formula for the z-test is:

Z Test Statistic

for Test of Significance for Mean

$$Z_0 = \left| \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \right|$$

$Z_0 \longrightarrow$ Z statistic

$\bar{x} \longrightarrow$ sample mean

$\mu \longrightarrow$ mean of expectation

$\sigma \longrightarrow$ standard deviation of population

$n \longrightarrow$ number of elements in large sample

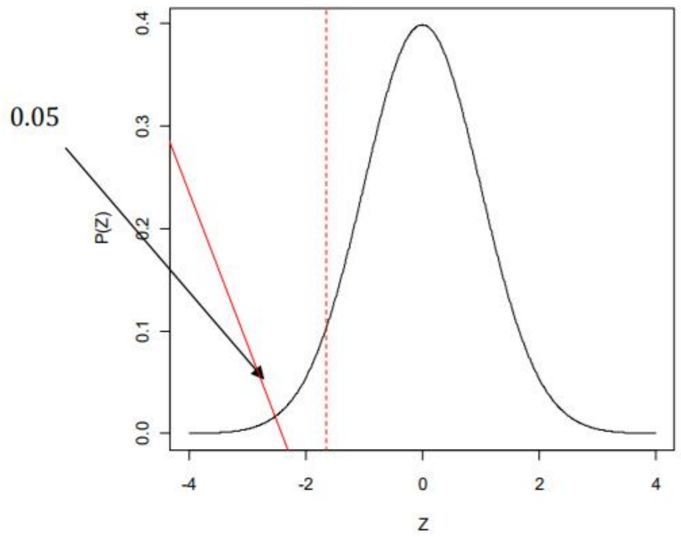
Example :

- Example: It has been reported that the average credit card debt for college seniors is \$3262. The student senate at a large university feels that their seniors have a debt much less than this, so it conducts a study of 50 randomly selected seniors and finds that the average debt is \$2995, and the population standard deviation is \$1100.
- Let's conduct the test based on a Type I error of alpha=0.05
- Step 1: State the hypotheses and identify the claim.

$H_0: \mu=\$3262$ $\quad\quad$ $H_1: \mu<\$3262$

CLAIM

- Step 2 : Find the critical value(s) from the appropriate table.
- Left-tailed test, alpha=0.05 , Z will be negative and have probability 0.05 underneath it



Z = -1.645 or
Z = -1.65

Table E The Standard Normal Distribution

Cumulative Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05
−3.4	.0003	.0003	.0003	.0003	.0003	.0003
−3.3	.0005	.0005	.0005	.0004	.0004	.0004
−3.2	.0007	.0007	.0006	.0006	.0006	.0006
−3.1	.0010	.0009	.0009	.0009	.0008	.0008
−3.0	.0013	.0013	.0013	.0012	.0012	.0011
−2.9	.0019	.0018	.0018	.0017	.0016	.0016
−2.8	.0026	.0025	.0024	.0023	.0023	.0022
−2.7	.0035	.0034	.0033	.0032	.0031	.0030
−2.6	.0047	.0045	.0044	.0043	.0041	.0040
−2.5	.0062	.0060	.0059	.0057	.0055	.0054
−2.4	.0082	.0080	.0078	.0075	.0073	.0071
−2.3	.0107	.0104	.0102	.0099	.0096	.0094
−2.2	.0139	.0136	.0132	.0129	.0125	.0122
−2.1	.0179	.0174	.0170	.0166	.0162	.0158
−2.0	.0228	.0222	.0217	.0212	.0207	.0202
−1.9	.0287	.0281	.0274	.0268	.0262	.0256
−1.8	.0359	.0351	.0344	.0336	.0329	.0322
−1.7	.0446	.0436	.0427	.0418	.0409	.0401
−1.6	.0548	.0537	.0526	.0516	.0505	.0495

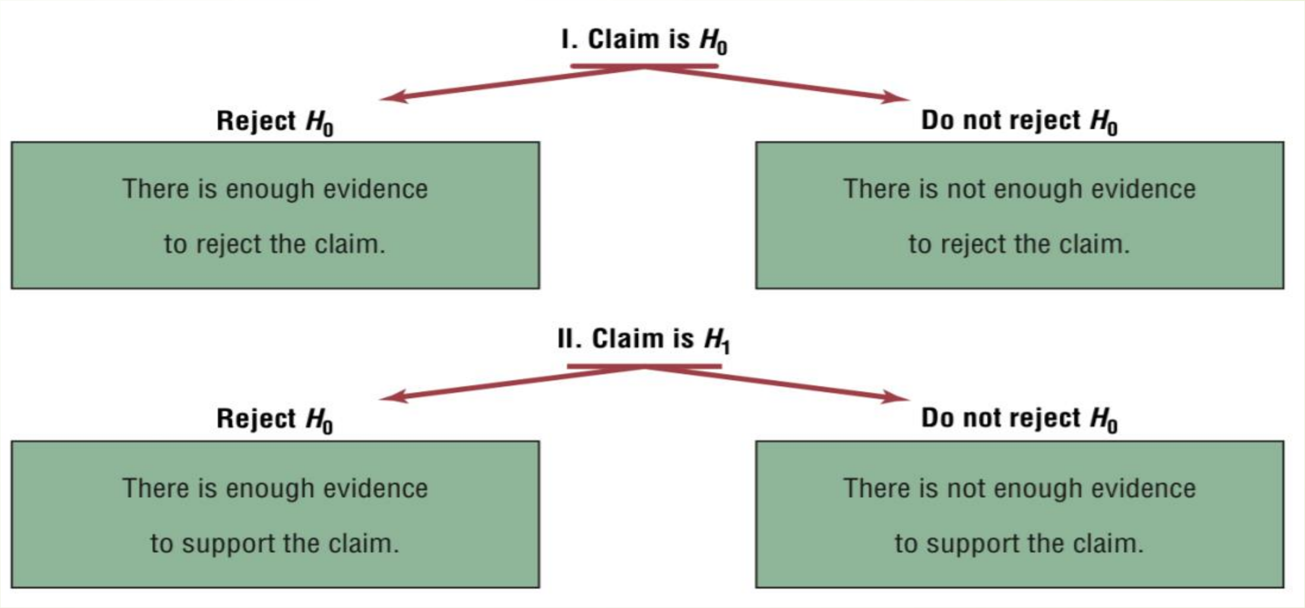
- Step 3 : Compute the test value

$$z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{2995 - 3262}{\frac{1100}{\sqrt{50}}} = \text{-}1.716341$$

- Step 4: Make the decision to reject or not reject the null hypothesis..
- Since this is a left-tailed test, our rejection region consists of values of Z that are smaller than our critical value of Z = - 1.645. Since our test value (-1.716341) is less than our critical value (-1.645), we reject the null hypothesis.
- Step 5: Summarize the results. We have evidence to support the student senate claim that the university's seniors have credit card debt that is less than the reported average debt. This is based on a Type I error rate of 0.05. This means we falsely make the claim above 5% of the time.

IMPORTANT NOTE:

- **When the null hypothesis is not rejected, we do not accept it as true. There is merely not enough evidence to say that it is false.**
- We conclude the alternative hypothesis (when we reject the null) because the data clearly support that conclusion.

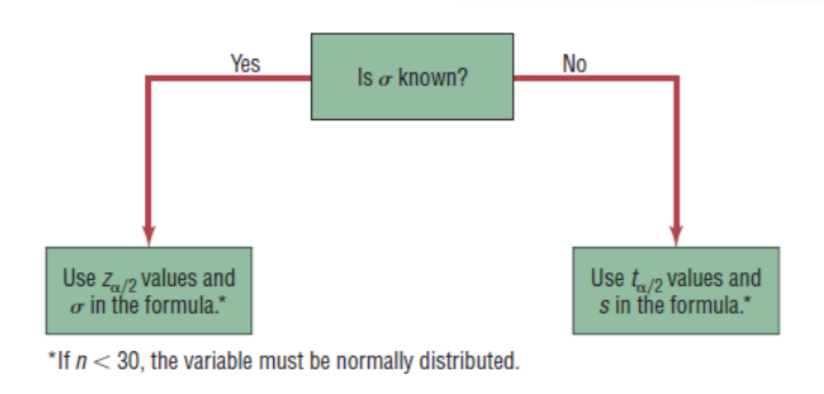


P-Value Method for Hypothesis Testing

- We often test hypotheses at common levels of significance (a = 0.05, or 0.01). Recall that the choice of alpha depends on the seriousness of the Type I error. There is another approach that utilizes a P-value.
- The P-Value (or probability value) is the probability of getting a sample statistic (such as the mean) or a more extreme sample statistic in the direction of the alternative hypothesis when the null hypothesis is true.
- The P-value is the actual area under the standard normal distribution curve of the test value or a more extreme value (further in the tail).

t Test for a Mean

- When a population is normally or approximately normally distributed, but the population standard deviation is unknown, the z test is inappropriate for testing hypotheses involving means.
- Instead we will use the t test when sigma is unknown and the distribution of the variable is approximately normal.



- The one-sample t test is a statistical test for the mean of a population and is used when the population is normally or approximately normally distributed and σ is unknown.
- The formula for the test value of the one-sample t test is:

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \qquad d.f. = n-1$$

T-Test Example:

- Example: Find the critical t value for alpha= 0.01 with sample size of 13 for a left-tailed test.
- Left tailed means the critical t value will be negative
- n=13 means the degrees of freedom are n-1 = 12
- The critical value is -2.681

Table F The *t* Distribution

	Confidence intervals	80%	90%	95%	98%	99%
	One tail, α	0.10	0.05	0.025	0.01	0.005
d.f.	Two tails, α	0.20	0.10	0.05	0.02	0.01
1		3.078	6.314	12.706	31.821	63.657
2		1.886	2.920	4.303	6.965	9.925
3		1.638	2.353	3.182	4.541	5.841
4		1.533	2.132	2.776	3.747	4.604
5		1.476	2.015	2.571	3.365	4.032
6		1.440	1.943	2.447	3.143	3.707
7		1.415	1.895	2.365	2.998	3.499
8		1.397	1.860	2.306	2.896	3.355
9		1.383	1.833	2.262	2.821	3.250
10		1.372	1.812	2.228	2.764	3.169
11		1.363	1.796	2.201	2.718	3.106
12		1.356	1.782	2.179	2.681	3.055
13		1.350	1.771	2.160	2.650	3.012
14		1.345	1.761	2.145	2.624	2.977

Example:

- We wish to check that normal body temperature may be less than 98.6 degrees. In a random sample of n =18 individuals, the sample mean was found to be 98.217 and the standard deviation was .684. Assume the population is normally distributed. Use alpha=0.05.
- Step 1: State the hypotheses and identify the claim.

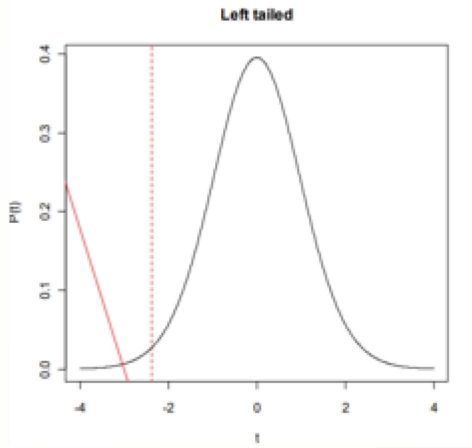
$$H_0: \mu = 98.6$$
$$H_1: \mu < 98.6 \quad \Leftarrow \quad \text{CLAIM}$$

- Step 2 : Find the critical value(s) from the appropriate table.

Left tailed, $\alpha = 0.05$, df=18-1=17 $\Rightarrow$ t critical value = 1.740

- Step 3 : Compute the test value and determine the P-value

$$t = \frac{98.217 - 98.6}{\frac{0.684}{\sqrt{18}}} = -2.375631 \approx -2.38$$



- p-value is between 0.01 and 0.025 p-value = 0.0146

- Step 4 : Make the decision to reject or not reject the null hypothesis
- Since our p-value is less than our alpha = 0.05, we reject the null hypothesis. The same conclusion is reached by looking at the critical value.
- Our test value is smaller than the critical value of -1.74.
- You only need to do it one way. The decision will always match
- Step 5 : Summarize the results. We have enough evidence to support the claim that average body temperature is less than 98.6 degrees.

- Thank You