Lecture 12

Hypothesis Testing

Today's Updates / Reminders

- PRWA #1 (Graph) – part 2 – closes at the end of the next class day.
  - Recall that you should review each paper. If you just give it a 100 thinking you'll be a "generous grader" and help out your classmates, you will likely lose points since the accuracy of your grading depends on how you compare to your peer graders. Also being too harsh negatively impacts some students.
- Lab 5 using Excel is due on Friday.
- HW 6 is open and will be due end of the next class day.
- Any exam regrade requests must be received by Friday, March 6.
- STAT Hub is open for business! From 5:30-8pm, M-Th in SAS1101. You can park right in front of SAS if driving!

1-sample mean CI by hand and with technology

How accurate are radon detectors of a type sold to homeowners? To answer this question, university researchers placed 12 detectors in a chamber that exposed them to 105 picocuries per liter of radon. The detector readings were as follows:

91.9	97.8	111.4	122.3	105.4	95.0
103.8	99.6	119.3	104.8	101.7	96.6

1-sample mean test and CI by hand and with technology

Find a 90% confidence interval for the population mean.

The mean and standard deviation for this sample of 12 detectors are **104.13** and **9.40**, respectively.

91.9	97.8	111.4	122.3	105.4	95.0
103.8	99.6	119.3	104.8	101.7	96.6

Remember:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}$$

Let's build a 90 % CI

Center: $\bar{x}$ = 104.13

radon.csv

Margin of error: $m = t^* \frac{s}{\sqrt{n}}$

CI at C=90%

s= 9.4

n=12

df = 11

t*=?

90% confidence level; df = 11 → t* = 1.796

CI: $\bar{x} \pm m$

$\bar{x} = 104.13$

$$m = t^* \frac{s}{\sqrt{n}} = 1.796 \frac{9.40}{\sqrt{12}} = 4.87$$

CI: $104.13 \pm 4.87 = (99.26, 109)$

Table entry for p and C is the critical value t^* with probability p lying to its right and probability C lying between $-t^*$ and t^*.

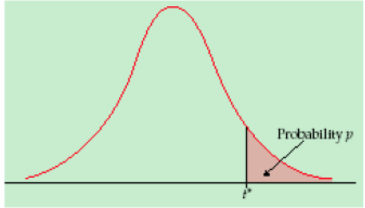
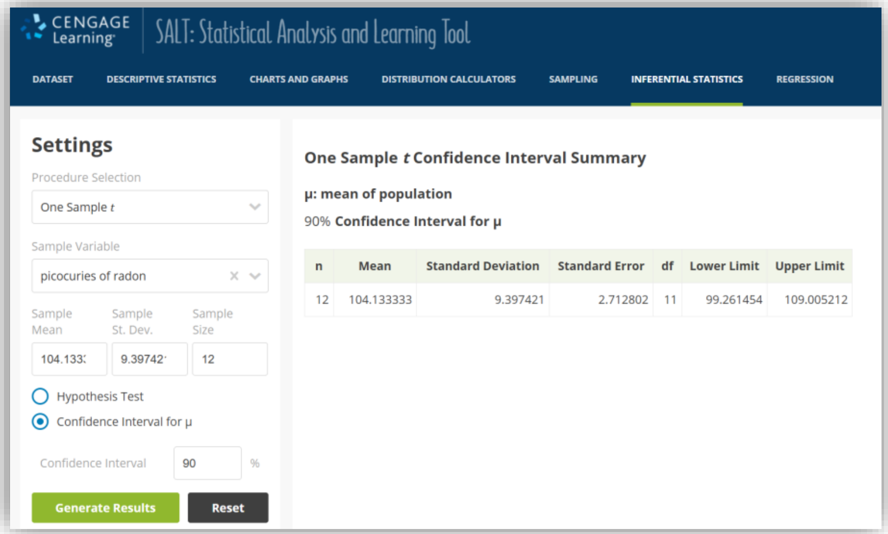


TABLE D t distribution critical values

df	Upper tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	[illegible]	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	[illegible]	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	[illegible]	[illegible]	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
z^*	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level C											

Software Output

Click Dataset. Upload CSV from previous slide. Click Inferential Statistics, One Sample t, Confidence Interval for μ, 90%



Right click and copy link from previous slide. Data, Load, From File, On the Web, Ctrl V, comma delimiter
Confidence Interval for μ, 90%
Stat, T-Stats, One-Sample, With Data
Select "picocuries of radon", Confidence Interval for μ, change "Level" to 0.90

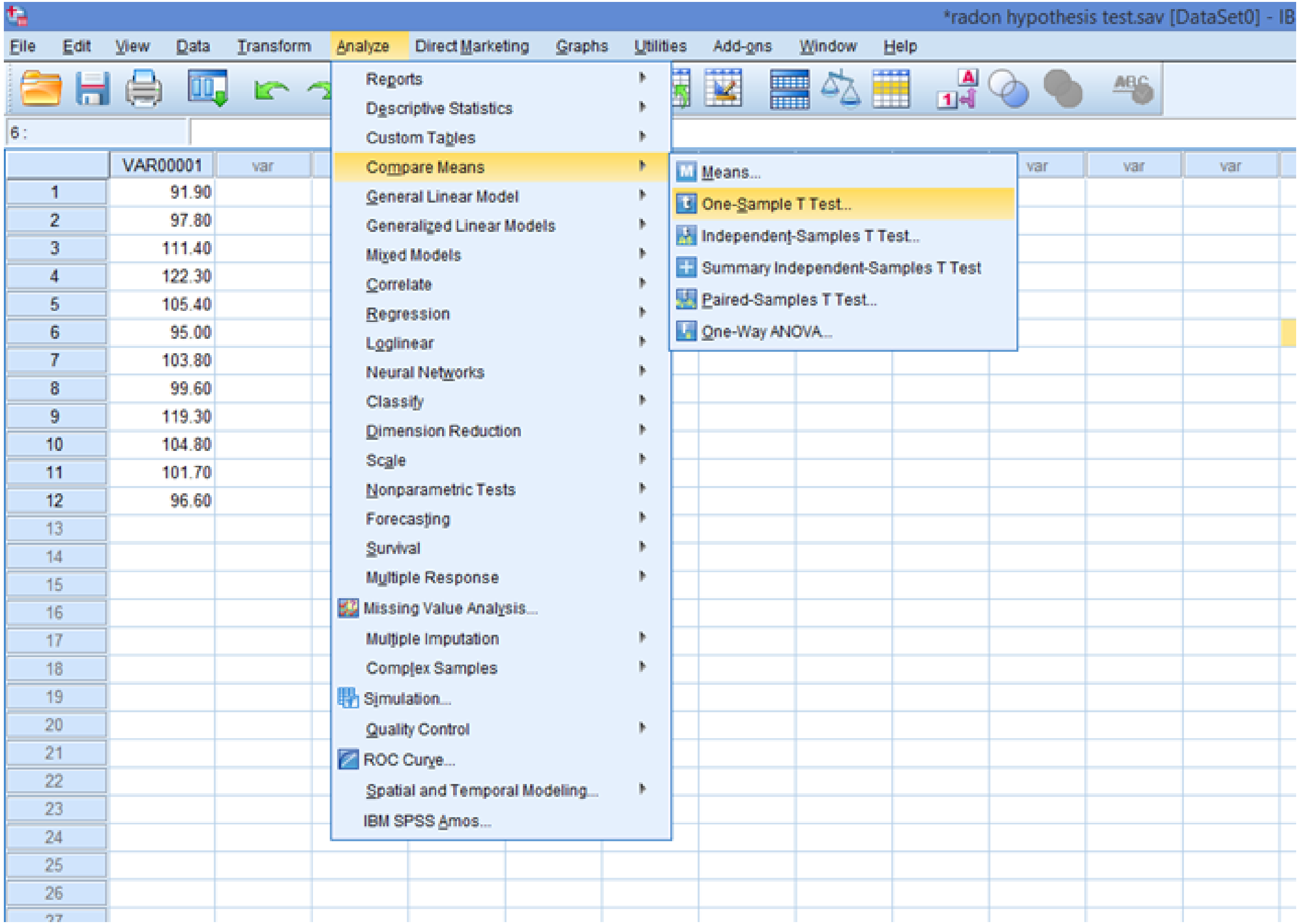
One sample T confidence interval:
μ : Mean of variable

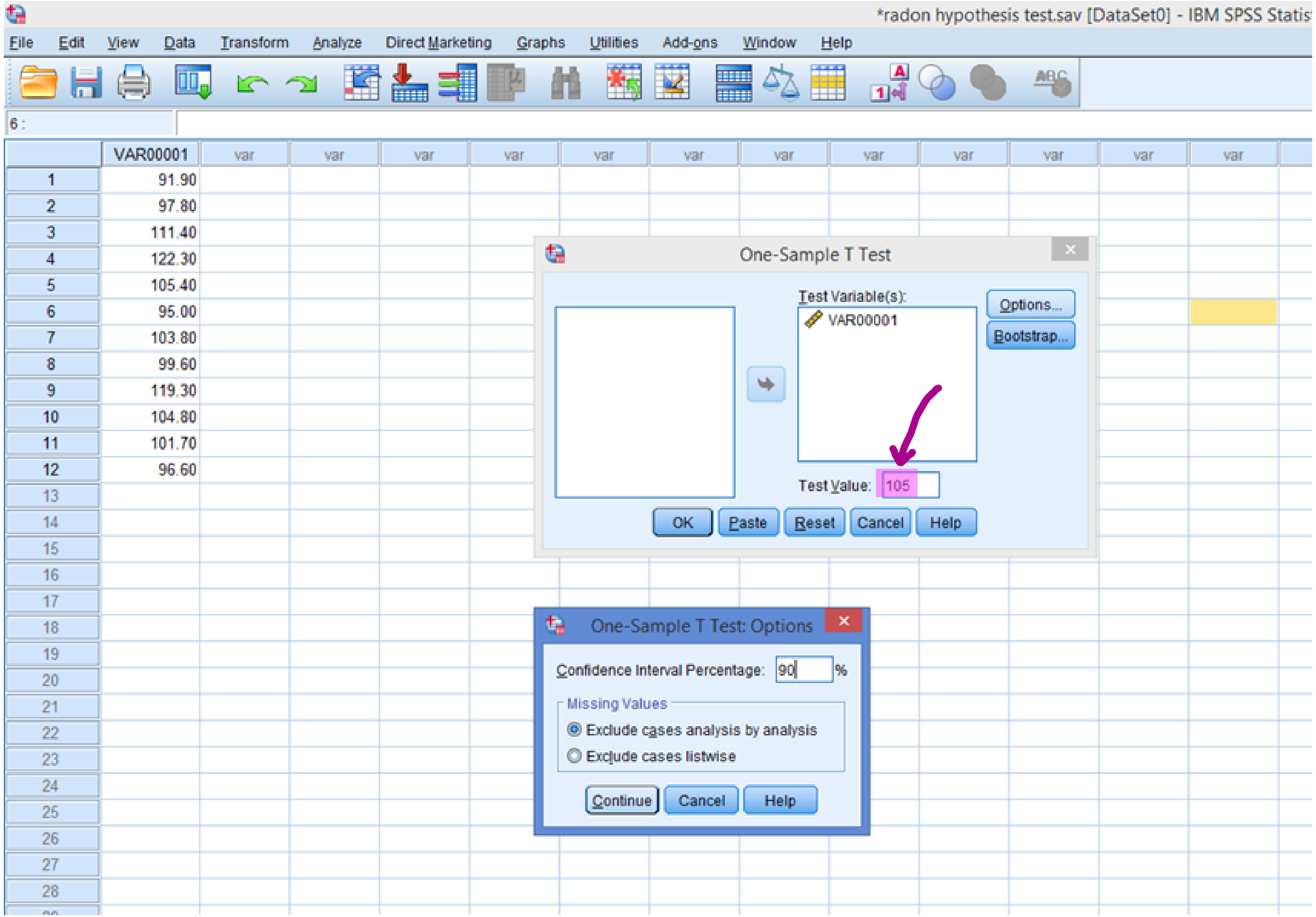
90% confidence interval results:

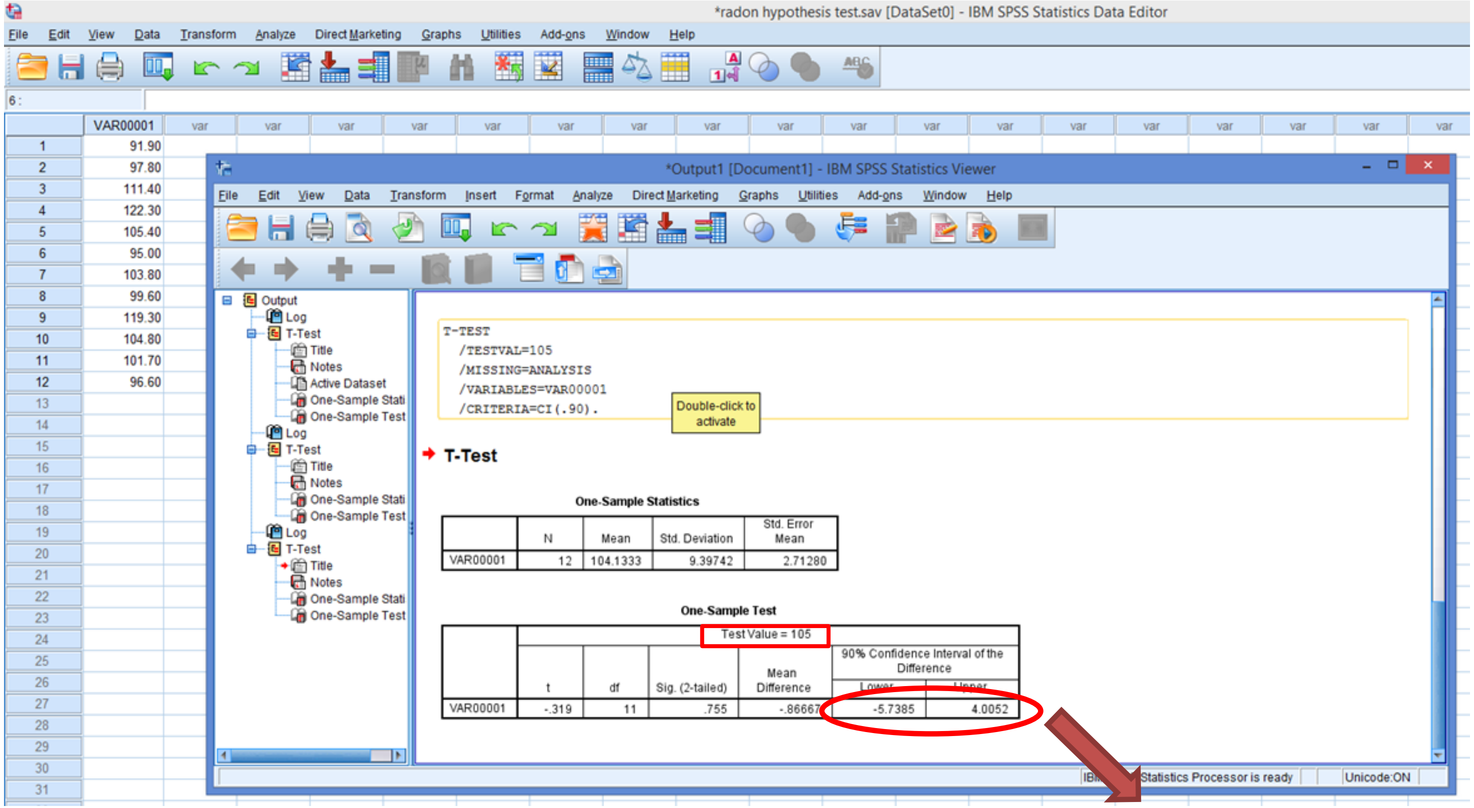
Variable	Sample Mean	Std. Err.	DF	L. Limit	U. Limit
picocuries of radon	104.13333	2.7128017	11	99.261454	109.00521

SPSS

- SPSS is an IBM product, but it is not free for students. I will show here just to give some exposure to different software and how others approach doing statistics.


*radon hypothesis test.sav [DataSet0] - IB
File
Edit
View
Data
Transform
Analyze
Direct Marketing
Graphs
Utilities
Add-ons
Window
Help
6 :
VAR00001
var
1 91.90
2 97.80
3 111.40
4 122.30
5 105.40
6 95.00
7 103.80
8 99.60
9 119.30
10 104.80
11 101.70
12 96.60
Reports
Descriptive Statistics
Custom Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
ROC Curve...
Spatial and Temporal Modeling...
IBM SPSS Amos...
Means...
One-Sample T Test...
Independent-Samples T Test...
Summary Independent-Samples T Test
Paired-Samples T Test...
One-Way ANOVA...


*radon hypothesis test.sav [DataSet0] - IBM SPSS Statis
File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help
VAR00001
91.90
97.80
111.40
122.30
105.40
95.00
103.80
99.60
119.30
104.80
101.70
96.60
One-Sample T Test
Test Variable(s):
VAR00001
Options...
Bootstrap...
Test Value: 105
OK Paste Reset Cancel Help
One-Sample T Test: Options
Confidence Interval Percentage: 90 %
Missing Values
Exclude cases analysis by analysis
Exclude cases listwise
Continue Cancel Help



Need to add test value to both:
(-5.7385 + 105 , 4.0052 + 105)

Guidelines for One-Sample t-Test

- When is it OK to use the t procedures?

Sample size	Use t procedures
n < 15	if your data looks like a Normal distribution (you can check this with a Normal Quantile Plot)
15 ≤ n < 40	except in the presence of outliers or strong skewness. Some skewness is ok!
n ≥ 40	except in the presence of outliers, ok even if data are very skewed.

Use different procedure

Investigate the cause: data wrongly recorded, equipment malfunction, response bias

Z vs. T

Summary Confidence Intervals and Hypothesis Tests

* If df = n-1 is not on table → round down df

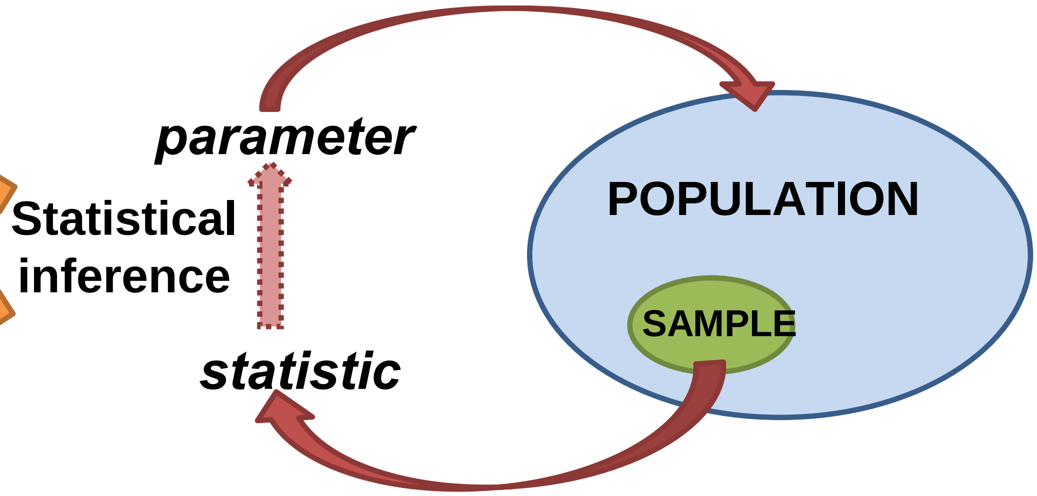
σ is known: Z-test	σ is not known: T-test
From sample (size n) we find $\bar{x}$	From sample (size n) we find $\bar{x}$ and s
$z = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}} \rightarrow N(0,1)$ (standard normal distr.)	$t = \frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n}}} \rightarrow t(n-1)$ (t-Distribution, df = n-1)
CI with C% confidence level: $\left(\bar{x} - z^* \frac{\sigma}{\sqrt{n}}, \bar{x} + z^* \frac{\sigma}{\sqrt{n}}\right)$ z^* : from bottom t-table, column C%	CI with C% confidence level: $\left(\bar{x} - t^* \frac{s}{\sqrt{n}}, \bar{x} + t^* \frac{s}{\sqrt{n}}\right)$ t^* : from t-table, row df = n-1, column C% if df = n-1 is not in t-table → round down df
Remember the 4 steps for HT, use z as test-statistic and the standard normal table (table A) to calculate the P-value if P-value < α → reject H_0; if P-value > α → don't reject H_0	Remember the 4 steps for HT, use t as test-statistic and the t-table (row df = n-1) to calculate the P-value if df = $n-1$ is not on t-table → round down df if P-value < α → reject H_0; if P-value > α → don't reject H_0
$H_a: \mu < \mu_0$ → P-value = P(Z < z) → z-table	$H_a: \mu < \mu_0$ → P-value = P(T < t) if t < 0 P-value = P(T < t) = P(T > | t |) → t-Table, row df = n-1 if t > 0 P-value = P(T < t) = 1- P(T > t) → t-Table, row df = n-1 *
$H_a: \mu > \mu_0$ → P-value = P(Z > z) = 1 - P(Z < z) → z-table	$H_a: \mu > \mu_0$ → P-value = P(T > t) if t < 0 P-value = P(T > t) = 1 – P(T > | t |) → t-table, row df = n-1 if t > 0 P-value = P(T > t) → t-table, row df = n-1 *
$H_a: \mu \neq \mu_0$ → if z < 0: P-value = 2 P(Z < z) → z-table if z > 0: P-value = 2 P(Z < -z) → z-table	$H_a: \mu \neq \mu_0$ → if t < 0 P-value = 2 P(T > | t |) → t-table, row df = n-1 if t > 0 P-value = 2 P(T > t) → t-table, row df = n-1 *

Objectives

- Learn how to perform a test of significance (or hypothesis test)
- Describe the four steps to follow in any significance test
- Define what is the null hypothesis and the alternative hypothesis
- Describe a common form for the test statistic
- Define what is a P-value and a significance level
- Learn how to draw a conclusion from a significance test based on the P-value and the significance level
- Describe the relationship between a two-sided significance test and a confidence interval

Confidence Interval
to estimate or infer the value of the parameter
(You have no clue of the value of the parameter, so you build an interval to estimate it)

Significance Test to infer the evidence of a claim regarding the parameter
(There is a certain value of the parameter that you want to test)



Tests of Significance

- **When will a significance test be useful?**
  - You have been told that the average grade in a certain course (population mean) is 60/100. You take an SRS of students taking that course and collect the grades of all of them. You calculate the statistic: sample mean and obtain 90/100. This looks like a pretty high grade!
  - Assuming that $\mu = 60$, is this $\bar{x} = 90$ just a rare case?
  - How rare is it? Is there any evidence that maybe the average grade of the course is larger than 60/100?

- The statement being tested is that the mean of the population μ (parameter) is 60 → **Null Hypothesis**

- The **significance test** is designed to assess the strength of evidence against the null hypothesis.

> **Rare Event Rule for Inferential Statistics**
>
> If, under a given assumption, the probability of a particular observed event is extremely small, we conclude that the assumption is probably not correct.

First step in a test of significance:

State a claim that we'll try to find evidence against → **Null Hypothesis**
State an alternative statement we suspect is true → **Alternative Hypothesis**

H_0 (Null Hypothesis): **Statement being tested. It is about the population parameter. ($\mu = \mu_0$, μ_0 is the test value)**

H_a (Alternative Hypothesis): **Statement we suspect is true instead of H_0. It is also about the population parameter.**

H_a

- **One-sided:** Parameter differs from its null hypothesis value in a specific direction ($\mu < \mu_0$, or $\mu > \mu_0$)
- **Two-sided:** Parameter differs from its null hypothesis value in either direction ($\mu \neq \mu_0$)

<u>One-sided hypothesis tests</u> would be:

H_0: $\mu = 60$ H_0: $\mu = 60$

H_a: $\mu > 60$ H_a: $\mu < 60$

A <u>two-sided hypothesis test</u> uses a ≠ sign:

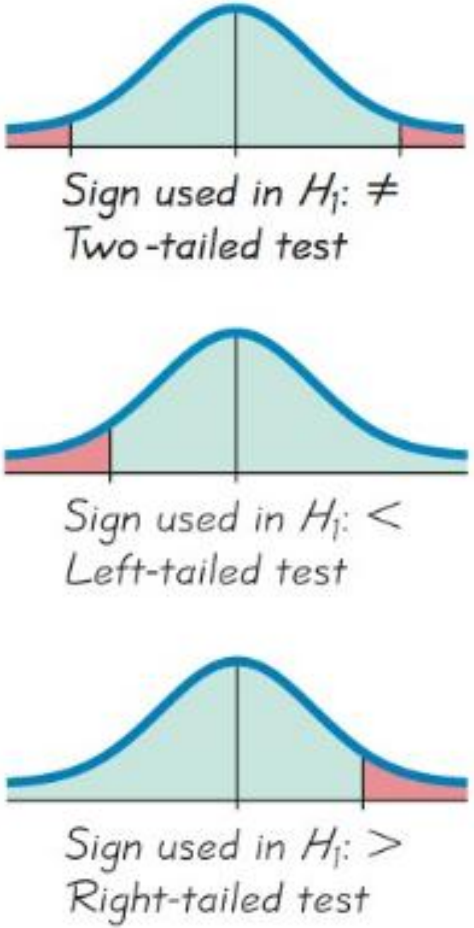
H_0: $\mu = 60$

H_a: $\mu \neq 60$

- **Two-tailed test:** The critical region is in the two extreme regions (tails) under the curve (as in the top graph in Figure 8-3).
- **Left-tailed test:** The critical region is in the extreme left region (tail) under the curve (as in the middle graph in Figure 8-3).
- **Right-tailed test:** The critical region is in the extreme right region (tail) under the curve (as in the bottom graph in Figure 8-3).

HINT To determine whether a test is two-tailed, left-tailed, or right-tailed, look at the alternative hypothesis and identify the region that supports that alternative hypothesis and conflicts with the null hypothesis. A useful check is summarized in Figure 8-3. *See that the inequality sign in H_1 points in the direction of the critical region*. The symbol ≠ is sometimes expressed in programming languages as <>, and this reminds us that an alternative hypothesis such as $p \neq 0.5$ corresponds to a two-tailed test.

Example: With H_0: $p = 0.5$ and H_1: $p > 0.5$, we reject the null hypothesis and support the alternative hypothesis only if the sample proportion is greater than 0.5 by a significant amount, so the hypothesis test in this case is *right-tailed*.



Steps in Significance Test

1. State the null and alternative hypothesis

2. Calculate a test statistic to measure the compatibility between the null hypothesis and the data.

3. Calculate the probability of the estimate (the statistic you measured from the sample) under the null hypothesis (**P-value**).

4. State a conclusion regarding evidence against the null hypothesis. Determine the required significance level, α. Reject the null hypothesis, H_0, if the p-value is LESS than the critical value. FAIL TO REJECT the null hypothesis, H_0, if the p-value is greater than α.

→ ***Let's go over 1 with examples and then show to calculate 2, 3 and 4***

Hypothesis Test Examples

- Null hypothesis, **H_0** – this is status quo. Before somebody proved that smoking causes cancer, the null hypothesis was what everybody assumed to be true up to that point, “Smoking has no effect on cancer.”
  - In statistical terms, we would state that the proportion of girl births equals 50%.
  - We would state that the mean fuel economy of some car is equal to what is stated by the car manufacturer.
  - We would state that your favorite bar stores their beer at the ideal temperature of 41°F. Symbolically:
- H_0: $p = 0.5$ (proportion of girl births is 50%)
- H_0: $\mu = 35$ miles per gallon (fuel mileage as stated by the car manufacturer)
- H_0: $\mu = 41°$ F (ideal beer temperature)

Hypothesis Test Examples

- Alternative hypothesis, $\mathbf{H_a}$ (sometimes H_1) – this is what we are trying to prove that is different from common knowledge. We wish to prove that smoking DOES cause cancer. In mathematical terms, we use >, <, or ≠. We may wish to prove that the proportion of girl births is greater than 50%. If we are trying to disprove the car company, we would say the mean fuel economy is less than what it stated by the manufacturer. We would state that your favorite beer is storing the beer either too cold or too hot. Symbolically:
- H_a: $p > 0.5$ (proportion of girl births is **greater than** 50%)
- H_a: $\mu < 35$ miles per gallon (fuel mileage **is less than what was** stated by the car manufacturer)
- H_a: $\mu \neq 41°$ F (the beer is either too cold or too hot)

Step 2: Test Statistics

Table 8-2

Parameter	Sampling Distribution	Requirements	Test Statistic
Proportion p	Normal (z)	$np \geq 5$ and $nq \geq 5$	$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$
Mean μ	t	σ not known and normally distributed population *or* σ not known and $n > 30$	$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$
Mean μ	Normal (z)	σ known and normally distributed population *or* σ known and $n > 30$	$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

<u>One Sample Z Test for a Population Mean</u>: we want to test whether we have evidence against the mean of the population having a certain value, assuming that **we know the population's standard deviation σ**.

H_0: $\mu=\mu_0$ (Null Hypothesis)

From a sample of size n we measure the sample mean $\bar{x}$ (estimate).

To assess how far is the estimate from the hypothesized value,

we standardize. In general, a test statistic will be:

$$\text{Test-statistic} = \frac{\text{estimate - hypothesized value}}{\text{standard deviation of the estimate}}$$

Based on the Central Limit Theorem:

Hypothesized value of the population mean

$$\bar{X} \text{ approx } N(\mu_0, \frac{\sigma}{\sqrt{n}})$$

Then, the **Z test-statistic** is:

$$z = \frac{\text{estimate - hypothesized value}}{\text{standard deviation of the estimate}} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

We use Z test-statistic when we know σ, the population std dev.

Step 3: P-value

P-value: **It is the probability (assuming H_0 true) that the test statistic would take a value as extreme or more extreme than the actually observed (remember $0 \leq P\text{–value} \leq 1$)**

If a P-value is small (i.e. close to 0), it means that the probability of observing that difference between the estimate μ_0 and the hypothesized value $\overline{X}$ is small; i.e. it is very rare to measure an estimate that is that different from the hypothesized value→ **the data is providing strong evidence against H_0**

If a P-value is large (i.e. close to 1), then it is not rare to measure an estimate that is that different from the hypothesized value → **there isn't enough evidence against H_0**

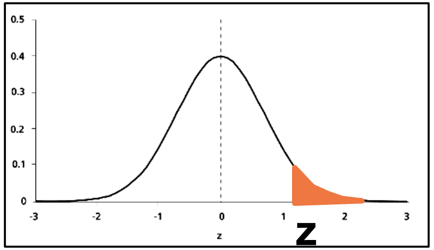
The P-value we calculate depends on the alternative hypothesis!!

(We'll use the Standard Normal Table to calculate P-values for z-Tests; σ is known)

Remember z is the test statistic calculated in step 2!

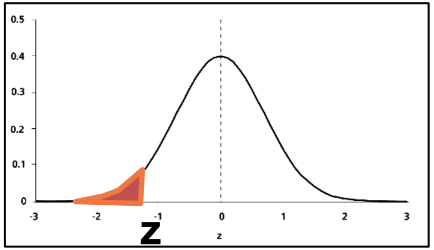
$$H_a: \mu > \mu_0 \rightarrow P(Z > z)$$

(One-sided Ha) $$= 1 - P(Z < z)$$



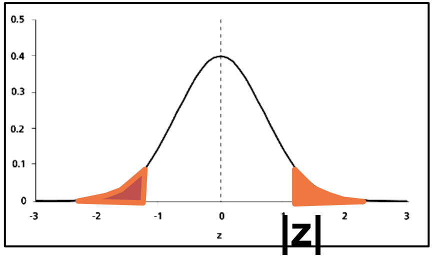
$$H_a: \mu < \mu_0 \rightarrow P(Z < z)$$

(One-sided Ha)



$$H_a: \mu \neq \mu_0 \rightarrow 2P(Z > |z|)$$

(Two-sided Ha) $$2P(Z < -|z|)$$



Step 4: Conclusions

4^{th} Step: State a conclusion regarding evidence against the null hypothesis

Compare the P-value to a fixed value that we regard as decisive → **Significance Level**

Significance Level α : Decisive value that announces in advance how much evidence against H_0 we will require to reject it. Typical values for α are 0.05 (it's the default), 0.01, 0.1, etc.

If the <u>P-value is as small or smaller than α</u>, we say that the data are <u>statistical significant at level α</u> → <u>we have enough evidence to reject</u> H_0

P-value

> α → **Do not reject H_0**

We do not have enough evidence against H_0. We do not have enough evidence in favor of H_a

≤ α → **Reject H_0**

We have enough evidence against H_0.
We have enough evidence in favor of H_a

Important: If the problem doesn't state the value of α, we consider α = 0.05!!

The smaller the α the stronger the evidence to reject H_0: If α = 0.01, you'll need P-val ≤ 0.01 to reject the null; if α=0.05, you'll only need P-val ≤ 0.05 to reject the null.

WARNING!

P−value > α does not mean we accept H_0

We cannot prove or accept H_0 with this kind of test

We can only reject H_0 in favor of H_a

Hypothesis Test #1

Universetoday.com M101

A theory predicts that the average size of spiral galaxies (like the Milky Way) is 50,000 light years (diameter).

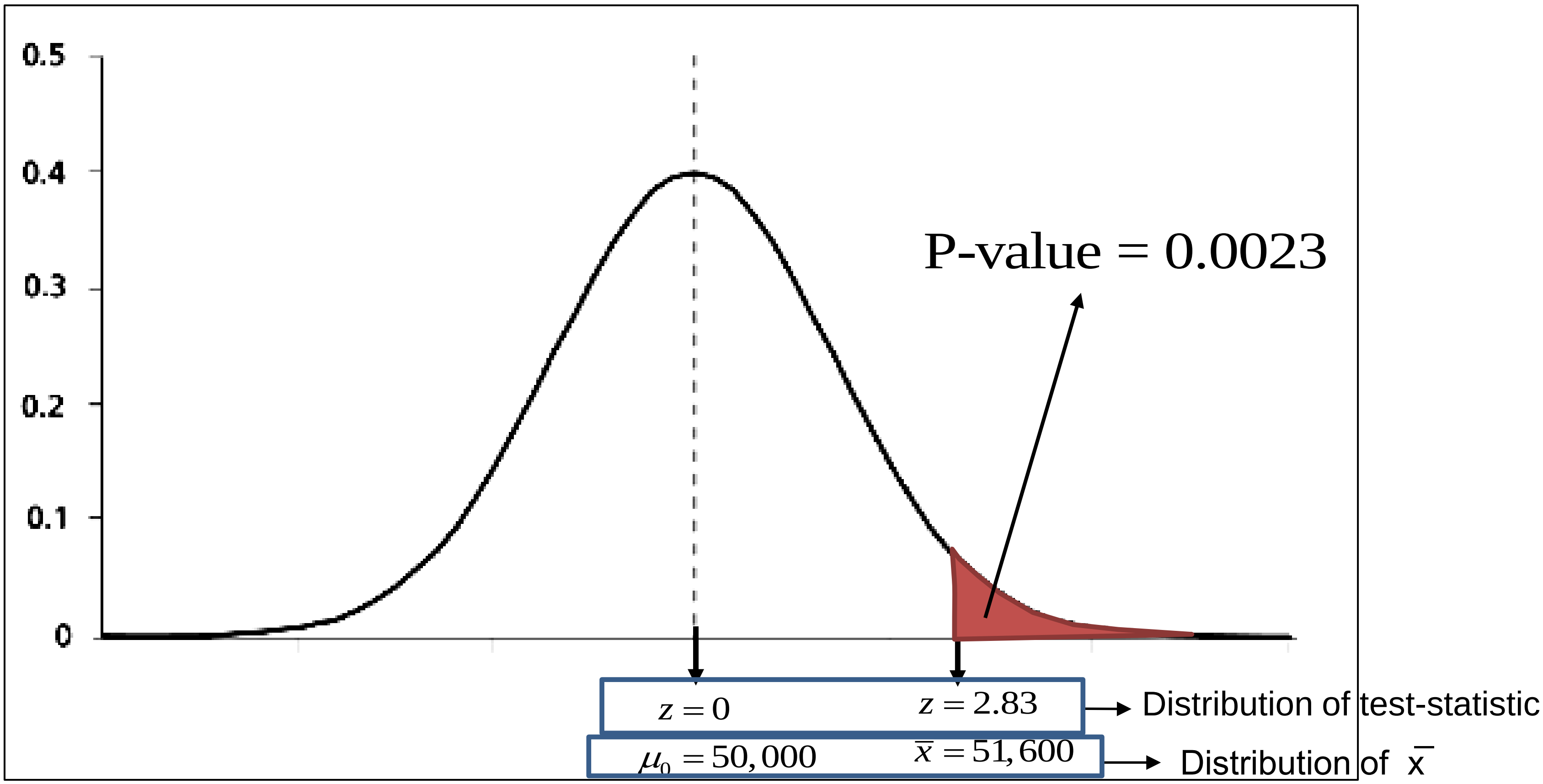
A research group selected a SRS of 50 spiral galaxies from a catalog containing data of spiral galaxies in the neighborhood of the Milky Way. They concluded in their study that the sample average was 51,600 light years. Assuming that the std of the population is σ=4,000 light years, conduct a significance test to see if there is evidence that the true population mean is more than the value predicted by the theory (the claim). (Use α=0.01)

1

2

3

4


0.5
0.4
0.3
0.2
0.1
0
P-value = 0.0023
$z = 0$
$z = 2.83$
Distribution of test-statistic
$\mu_0 = 50,000$
$\bar{x} = 51,600$
Distribution of $\bar{x}$

Indianapublicmedia.org

Hypothesis test #2

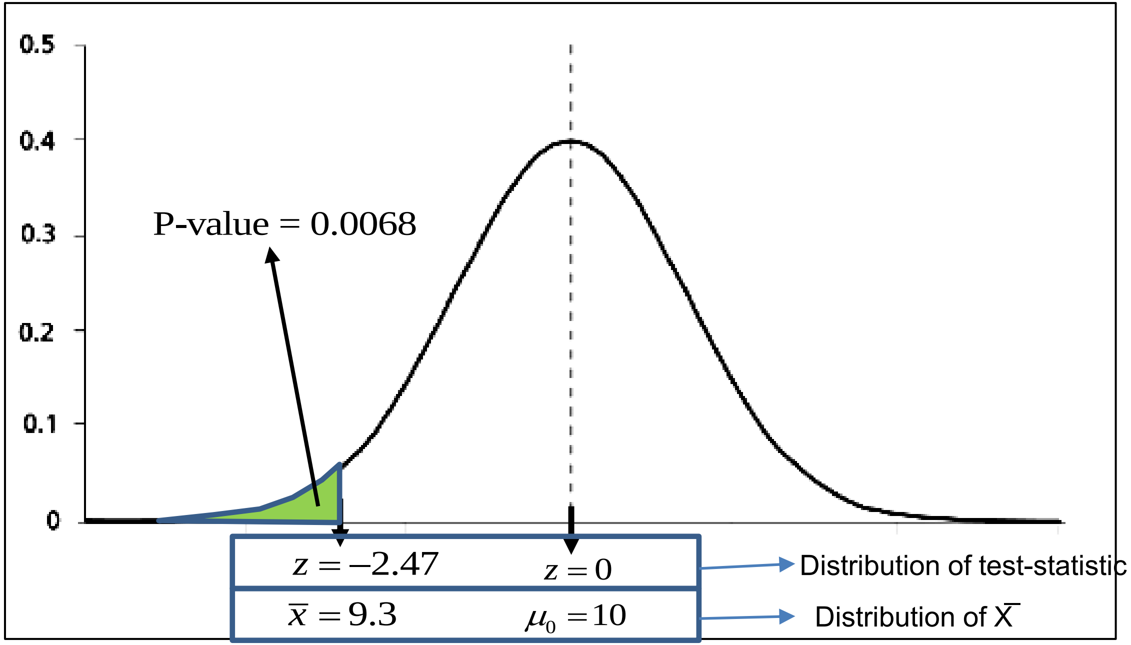
In a recent publication a claim was made that the average amount of candy eaten per year by children between five and eight years old in the country is 10 lbs. A stratified random sample of 50 children within that age group was gathered in order to check the evidence for that claim. The study showed the average amount of candy eaten per year was 9.3lbs. Given a σ=2 lbs, perform a significance test to determine whether there is any evidence that the average amount of candy eaten per year by children 5-8 years old is smaller than 10 lbs. Assume α=0.03.

1

2

3

4


0.5
0.4
0.3
0.2
0.1
0
P-value = 0.0068
$z = -2.47$
$z = 0$
$\bar{x} = 9.3$
$\mu_0 = 10$
Distribution of test-statistic
Distribution of $\bar{X}$

Hypothesis test #3

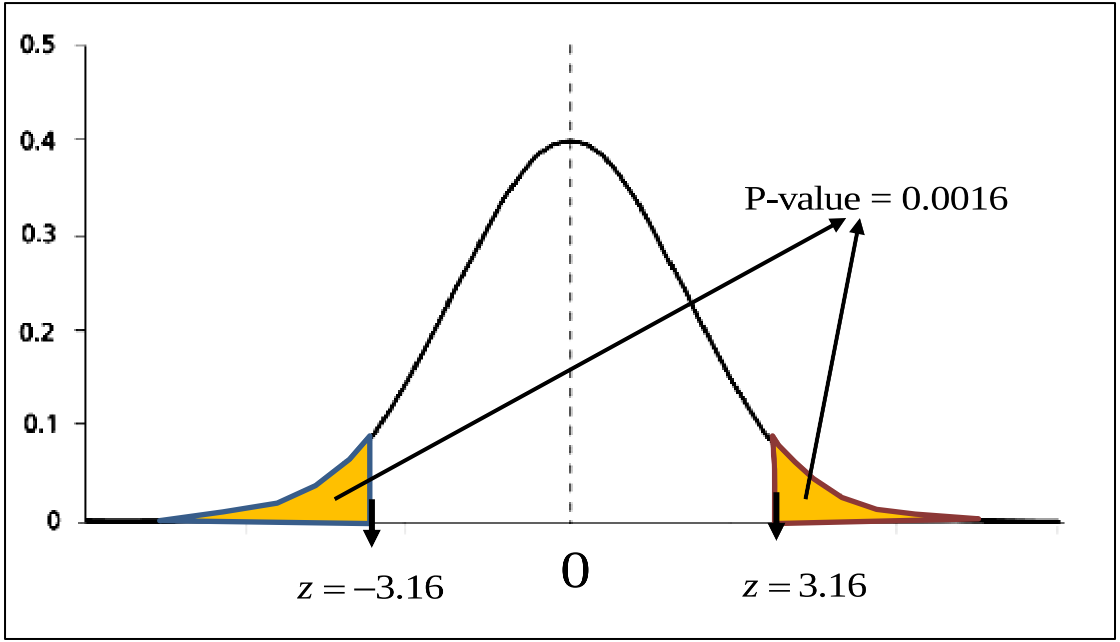
Newspapers reported that within the county school system, the average time spent per week by first graders doing HW is 50 minutes. Such a claim started several discussions and before taking any further steps the school system decided to initiate its own study. A SRS of 40 first graders registered in that school system were selected, and their parents/guardians were asked to report the number of minutes per week that the student spent doing HW. The average time/week out of that sample was 55 minutes/week. If $\sigma=10$ minutes/week, perform a significance test, at the 10% significance level, to establish whether there is any evidence that the mean of the population is different from 50 minutes/week.

1

2

3

4


0.5
0.4
0.3
0.2
0.1
0
P-value = 0.0016
z = −3.16
0
z = 3.16

Another way of performing a significance test: Using a confidence interval to draw a conclusion about a hypothesis test.

To do that you need two things:

1. Your alternative hypothesis from the test must be two-sided (has a ≠)

2. Your confidence level and your significance level add to 100% (Example: $\alpha = 0.1\,(10\%)\,;\text{C}=90\%$)

How do we use a confidence interval (CI) to draw a conclusion about a Hypothesis Test (HT)?

You need to ask yourself the following question:

Is the test value from null hypothesis (μ_0) contained in the CI?

Yes → Do not reject H_0

No → Reject H_0

Hypothesis test #3 – HW example

$H_0 : \mu = 50$
$H_a : \mu \neq 50$
at $\alpha = 0.1$, we rejected the null.

Find the 90% C.I. for the mean time/week spent by first graders in the county school system doing HW.

$\text{CI} \rightarrow \bar{x} \pm m; \quad \bar{x}=55; \; m = z^* \frac{\sigma}{\sqrt{n}};$

Use bottom part of t-Table (Table D):

$z^* = 1.645$ (for C=90%); σ=10; n=40

$$\text{margin of error} \rightarrow m = z^* \frac{\sigma}{\sqrt{n}} = 1.645 \frac{10}{\sqrt{40}} = 2.6$$

$$\text{confidence interval} \rightarrow (\bar{x} - m, \bar{x} + m) = (55 - 2.6, 55 + 2.6) = (52.4, 57.6)$$

Another way of thinking about the relationship between CI and HT

Suppose you are told:

H_0: μ = 37

You gather a SRS and build a 95% CI.

CI: [34.61, 39.12]

What would be the conclusion of a HT based on this CI? What is the significance level that you should consider?