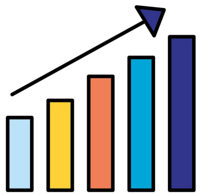
Intro to Applied Statistics and Hypothesis Testing

CSE 407 - Week 8

LEC RAIYAN RAHMAN
Dept of CSE, MIST
raiyan@cse.mist.ac.bd

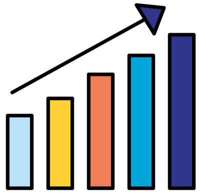
What’s Statistics?



- The branch of mathematics that deals with the collection, organization, analysis, interpretation and presentation of data.

- “Statistics” as defined by the American Statistical Association (ASA) “is the science of learning from data, and of measuring, controlling and communicating uncertainty.”

What’s Statistics?



- A collection of well-documented data
- Methods to properly collect and organize data
- Methods to analyze gathered data
- Methods to interpret that analysis
  - or Study of the techniques and methods used to gather insights from that data.
- Then the presentation of the statistical data and results.
- It’s where we study statistical inference, statistical modelling, probability, applied statistics, queuing theory etc.

What’s Applied Statistics?

- Use of statistical tools, methods and techniques to analyze and get insights into a particular dataset.

- The root of modern data science and data analysis.

- Computers and Programming are widely used to analyze huge masses of data.

Examples

Data: Say we have the final marks of the 84 students of CSE 407. The average mark is 68.5 for CSE 17.

Question: How likely is it that the average mark of CSE 407 for CSE 18 will be within 65-70?

Data: You're working as maintenance engineer at Samsung. The new SSDs that they are making have an advertised read speed of 550mb/s. But after purchasing some users are reporting much lower read speeds at an avg of 535mb/s.

Question: Is the advertised speed by Samsung wrong? Do you need to revise it?

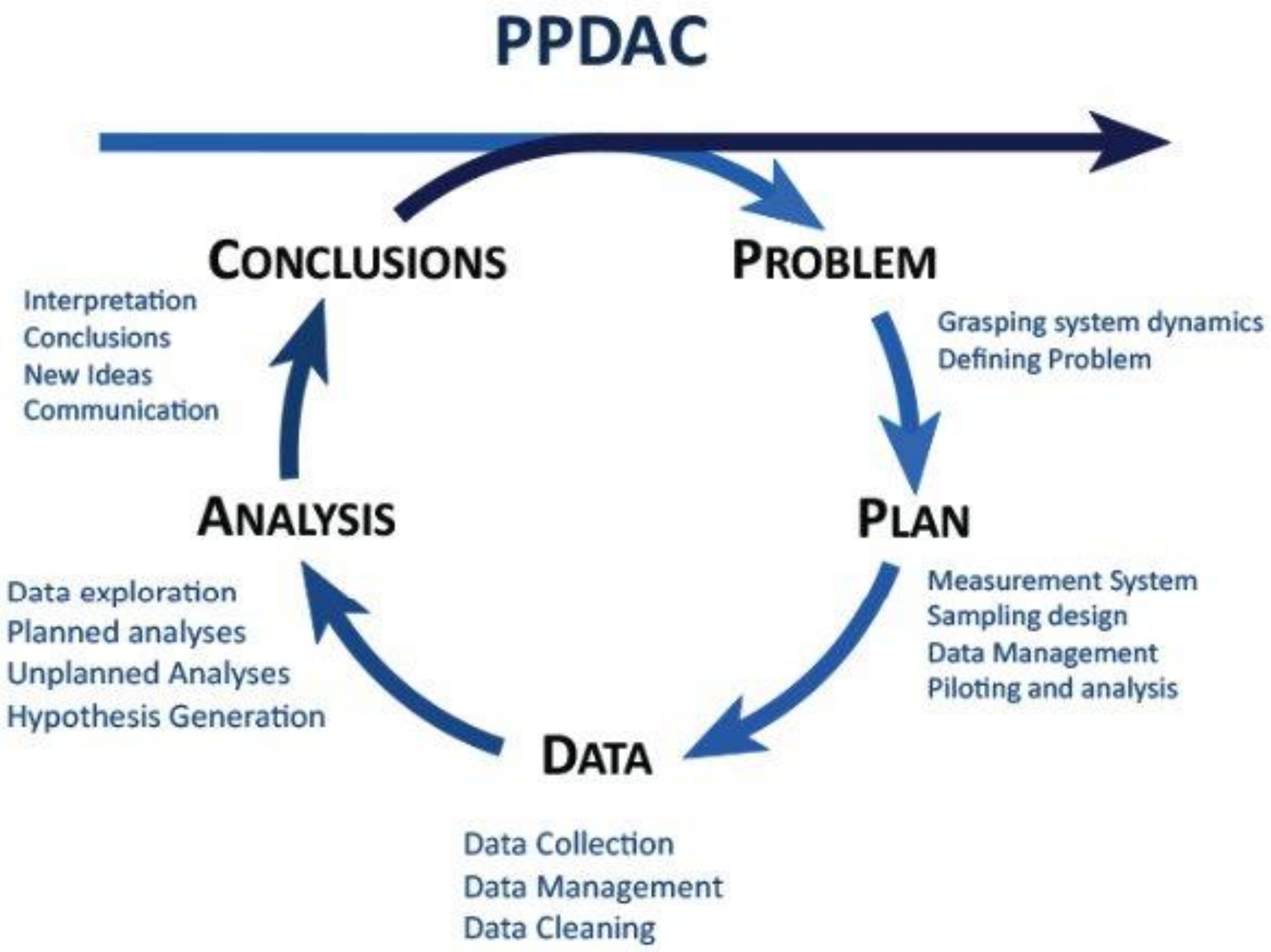
(this is what we'll study today)

Data: Say that Robi is targeting new 4GB,7 days internet packs at a value of BDT 99. They want to pick the right target group to advertise these pack to. From historical data, they know people who bought >2GB, 7 days pack are 70% likely to buy the new pack and those who bought >1GB packs are 50% likely to buy it?

Question: What should be ratio of this ad's targeted user group?

All of these can be answered with Applied Statistics.

The Investigative Statistical Cycle (PPDAC)



But first, Let’s get oriented with some terms and concepts of statistics.

Population

- The entire possible dataset (that i ***should*** base my decisions on).
- But more often than not, the entire dataset is not available.
- For example, Let’s say Samsung already manufactured and shipped 2.5M SSDs. Now, to test the customer’s claim (and to clear samsung’s name!) I can’t call back all 2.5M SSDs to test their actual speed again.
- However, in case of the internet packs, the telecom operator should have the entire dataset. So we should be able make decisions more accurately (depending on how good my model is).
- So although we ***should***, we ***can’t*** always analyze the entire dataset.

Sub-Population

- Classification of the entire possible dataset based on some criteria.
- For example, in the first scenario, if i was calculating the average height of CSE 17 students instead of marks, it might be a good idea to divide the population of 84 into two sub-population, namely Male and Female, since height are more closely matched within that sub group.
- however, for marks, no such distinction (hopefully) will be there and no need to divide into sub-populations.

Sample

- A subset of the population.
- For example, since we can't call back all 2.5M SSDs. We'll do the speed testing on 1000 SSD cards ready to be shipped. We'll reject the customer's claim (or revise our claim/tweak production) based on the result from this "sample" of 1000 SSDs.
- In most production cases, this is more often the statistical scenario.

Population vs Sample Mean

- Mean is simply the average value of a particular parameter for a sample or population.
- We divide the sum of all values of a variable by the total number of available data.

Population vs Sample Mean

Population Mean	Sample Mean
$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$	$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$
N = number of items in the population	*n* = number of items in the sample

Note: notice the different symbols used to denote the different means.

Population vs Sample Variance and Std Deviation

Population Variance

$$\sigma^2 = \frac{\sum_{i=1}^{n}(X_i - X_{avg})^2}{n}$$

Population Standard Deviation

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(X_i - X_{avg})^2}{n}}$$

Sample Variance

$$s^2 = \frac{\sum_{i=1}^{n}(X_i - X_{avg})^2}{n-1}$$

Sample Standard Deviation

$$s = \sqrt{\frac{\sum_{i=1}^{n}(X_i - X_{avg})^2}{n-1}}$$

Note: Variance is a measure of how varied the values are from the calculated mean.

Hypothesis Testing

Hypothesis Testing

- Let's go back to the Samsung SSD example.
- We have contradictory claims by each group.
- How to resolve? The concept of hypothesis testing.

Testing A Hypothesis - Scenario

Data: You're working as maintenance engineer at Samsung. The new SSDs that they are making have an advertised read speed of 550mb/s. But after purchasing some users are reporting much lower read speeds with different values.

To verify their claims, you decided to gather 100 SSDs ready to be shipped to stores from your production line. After careful testing, the read speed of these test SSDs were found to be 548.9 mb/s. You know from previous tests that the population variance of your SSDs is 6.5 mb/s.

Question: Test the claim by the customers with 95% confidence and establish if the advertised read speed of 550 mb/s is acceptable or not.

Null and Alt Hypothesis

1. Null hypothesis (H_0):
   a. It's the currently claimed or accepted value of a parameter.
   b. In our example, 550mb/s is the Null hypothesis.

2. Alternate Hypothesis (H_1 or H_a):
   a. It's the claim different than the current claim.
   b. May be the exact opposite of a range. but never equal to the H_0 value.

Null and Alt Hypothesis

One-tailed

Two-tailed	Left-tailed	Right-tailed
$H_0: \mu = 550$	$H_0: \mu \geq 550$	$H_0: \mu \leq 550$
$H_1: \mu \neq 550$	$H_1: \mu < 550$	$H_1: \mu > 550$

This is meaningless for our case.

Possible Outcomes for H_0

1. Accepted or "fail to reject": We couldn't gather enough evidence from our sample dataset to reject the Null hypothesis. Hence, we accepted it.
2. Rejected: We concluded from our sample dataset that the original claim (H_0) was wrong.

(correction from class)

Level of Confidence

- However, we are testing on a sample dataset.
- So We can’t be that rigid.
- We need to give some leeway (instead of a value, we’ll test for a range).
- Hence we have the concept of “level of confidence” (C).
- The smaller the sample dataset the higher we should set the val for C.
- α (next slide) should be smaller and C should be higher for a smaller sample size. Since a higher val for C gives us a wider acceptance region and that is what we want for smaller sample sizes (we shouldn’t be too rigid if n is small).

Level of Significance

- The complement of C.
- Denoted by $\boldsymbol{\alpha}$
- $\boldsymbol{\alpha} = 1 - C$
- This defines the critical values or the "area of rejection".

- $\boldsymbol{\alpha} = 1 - 0.95 = 0.05$

Null and Alt Hypothesis

One-tailed

Two-tailed

$H_0: \mu = 550$

$H_1: \mu \neq 550$

Left-tailed

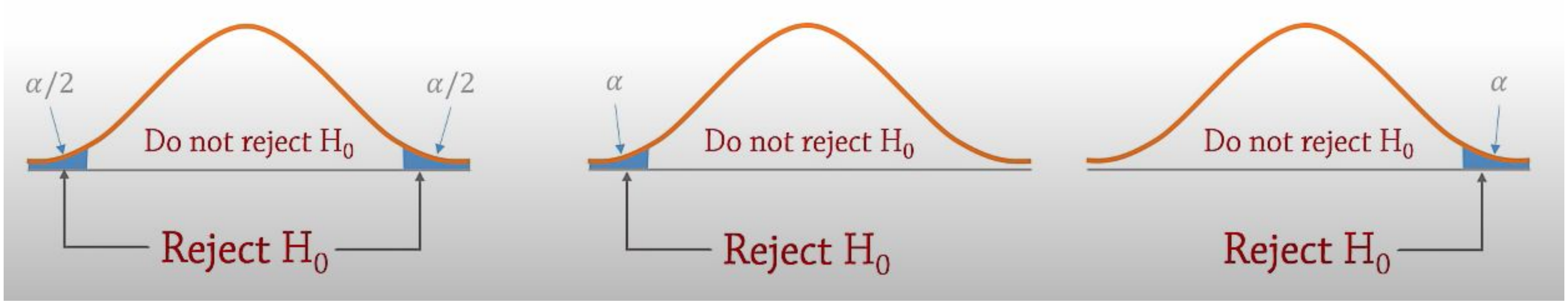
$H_0: \mu \geq 550$

$H_1: \mu < 550$

Right-tailed

$H_0: \mu \leq 550$

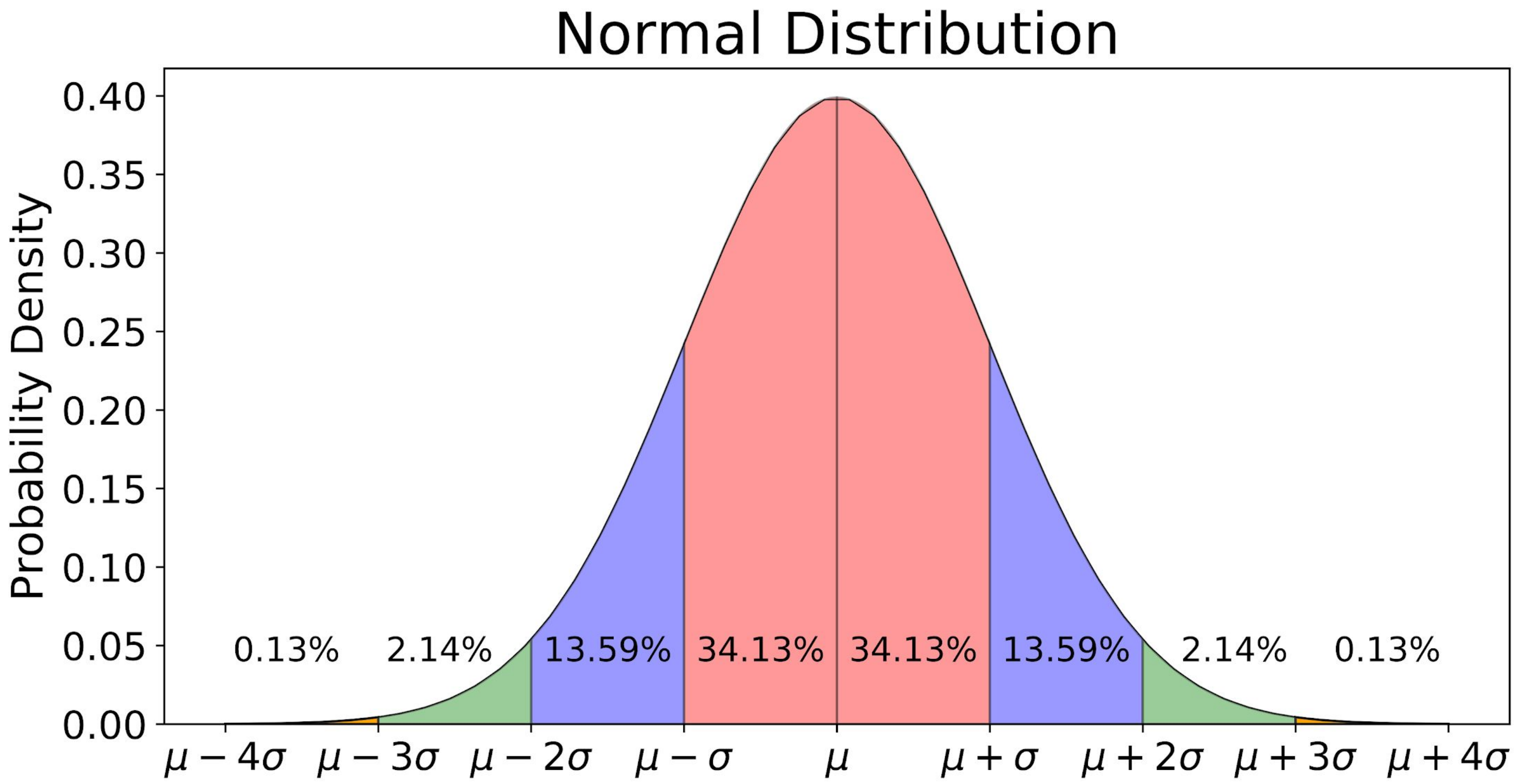
$H_1: \mu > 550$



Prelude: Normal Distribution

- Most real life means are distributed this way.
- We have a very specific graph for standard normal distribution (avg=0 and std deviance=1).
- We’ll use this to test the hypothesis assuming the data is normally distributed.
- Let’s see what it looks like.

Prelude: Normal Distribution



Testing our Hypothesis (Exercise 1)

Data: You're working as maintenance engineer at Samsung. The new SSDs that they are making have an advertised read speed of 550mb/s. But after purchasing some users are reporting much lower read speeds with different values.

To verify their claims, you decided to gather 100 SSDs ready to be shipped to stores from your production line. After careful testing, the read speed of these test SSDs were found to be 548.9 mb/s. You know from previous tests that the population std deviation of your SSDs is 6.5 mb/s.

Question: Test the claim by the customers with 95% level of confidence and establish if the advertised read speed of 550 mb/s is acceptable or not.

Testing our Hypothesis

Step 1: Figure out the Null and Alternate Hypothesis.

Step 2: Gather and if needed, calculate necessary test data. **n,** $\bar{X}$**,** σ**, α**.

Step 3: Decide which test to do (one or two tailed) and/or which table to use (z or t table).

Step 4: Calculate critical values from level of confidence or significance.

Step 5: Find test statistic. The “z” value.

Step 6: Interpret the data and decide whether to accept (fail to reject) or reject the H_0

Now, let’s see the question again

Data: You’re working as maintenance engineer at Samsung. The new SSDs that they are making have an advertised read speed of 550mb/s. But after purchasing some users are reporting much lower read speeds with different values.

To verify their claims, you decided to gather 100 SSDs ready to be shipped to stores from your production line. After careful testing, the read speed of these test SSDs were found to be 548.9 mb/s. You know from previous tests that the population std deviation of your SSDs is 6.5 mb/s.

Question: Test the claim by the customers with 95% confidence and establish if the advertised read speed is less atha 550 mb/s.

Step 1

Null Hypothesis: $H_0 : \mu = 550$

Alternate Hypothesis: $H_1 : \mu \neq 550$

Step 2

$n = 100$

$\bar{x} = 548.9$

$\sigma = 6.5$

$\alpha = 1 - C$.

Here, $C = 0.95$. So, $\alpha = 1 - 0.95 = 0.05$.

Step 3

Decision 1: Which test to use

Look at the Alt Hypothesis H_1 to decide.
Here, we see it'll be a two tailed test.

Decision 2: Which table to use.

If population's std deviance, σ is known, use z table.
And/or if n>100, use z table. Some tests may use z table with n >30!
Otherwise, use t table.

Here, Population's std deviance, σ (5.5 given) is known. Also, n = 100. So we'll use z table.

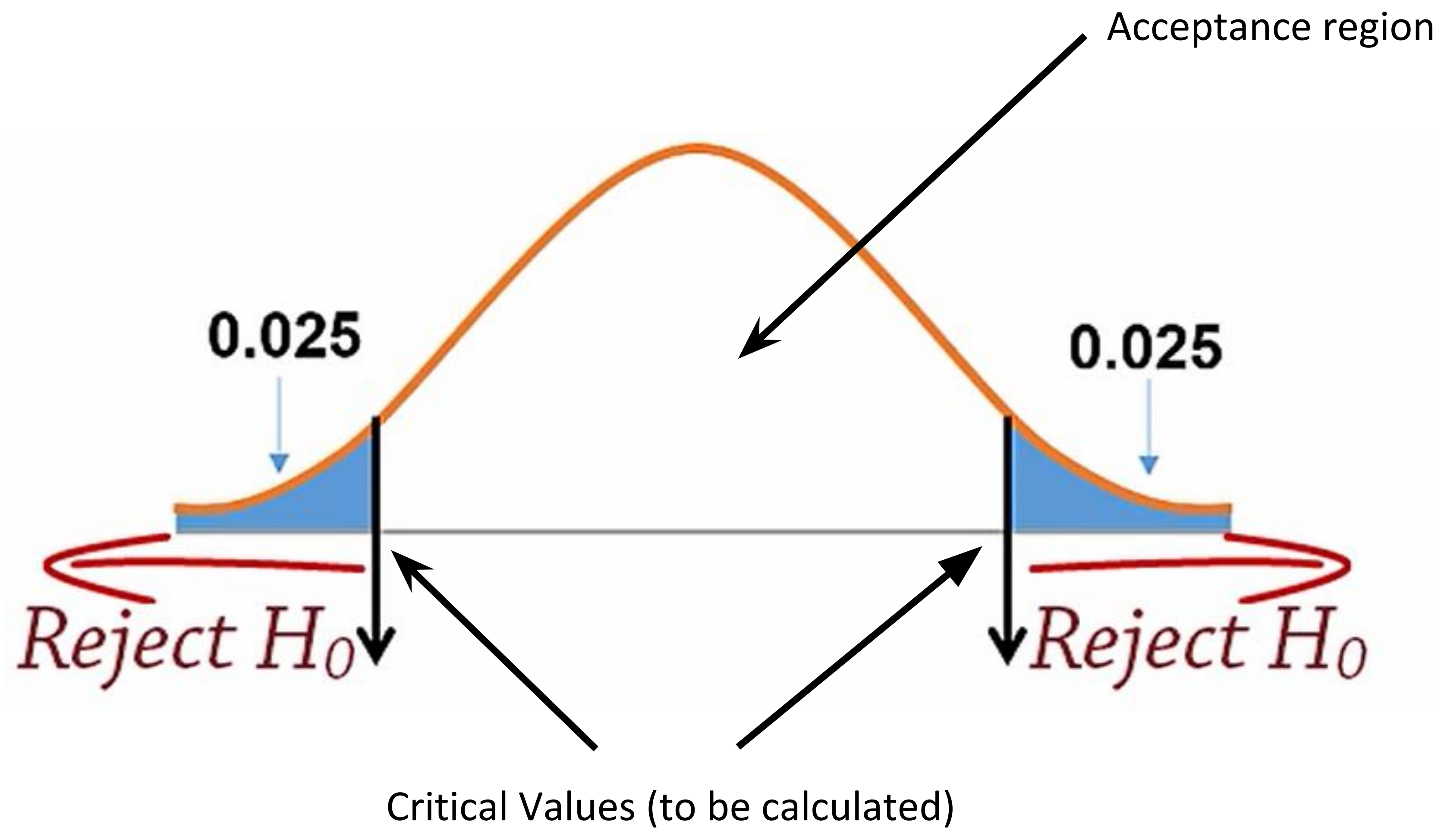
Step 4

We'll calculate the critical value based on the level of significance (**α** value), the type of test we're doing and the table we're using.

Here, **α** = 0.05

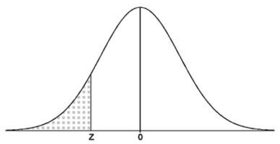
We're using a two-tailed test. So as per symmetry, **α**/2 = 0.025 will be the corresponding value for each side.

Step 4


Acceptance region
0.025
0.025
Reject H_0
Reject H_0
Critical Values (to be calculated)

Step 4 Let's Look at the z table

Cumulative Standard Normal Distribution Table



Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-0.00	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
-0.10	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.20	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	[illegible]	0.3936	0.3897	0.3859
-0.30	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.40	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.50	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.60	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.70	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.80	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.90	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-1.00	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-1.10	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.20	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.30	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.40	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.50	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.60	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.70	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.80	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.90	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-2.00	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-2.10	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.20	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.30	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.40	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.50	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.60	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.70	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.80	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.90	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-3.00	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-3.10	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.20	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.30	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.40	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.50	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
-3.60	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.70	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.80	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001

*Note: z-values less than -3.89 produce a probability of zero.

Department of Mathematics, Sinclair Community College, Dayton, OH

Download the full table here.

Here, the left critical value is -1.96 for 95% confidence. Since the graph is symmetric the right critical value is +1.96.

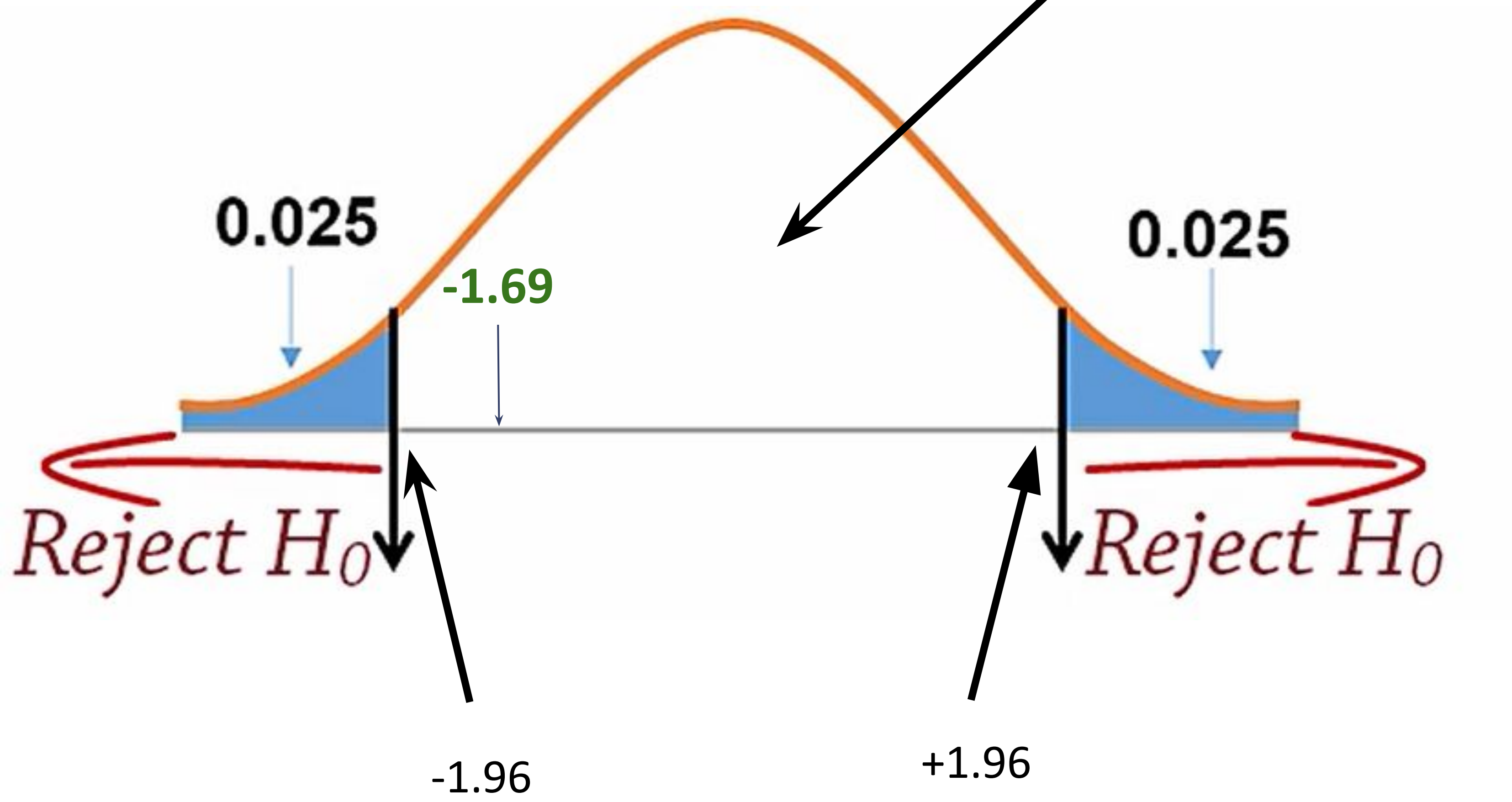
Step 5

Calculate the test stat “z”.

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{548.9 - 550}{6.5 / \sqrt{100}} = -1.69$$

Step 6 - Interpret the data

Acceptance region. The z value is inside. so we “accept” H_o.



Step 6 - Interpret the data

Judging from our test data of 100 ready to be shipped SSDs, as per the level of significance taken in our test and population variance, the advertised read speed of 550 mb/s is correct.

So we accept or fail to reject the Null hypothesis that the SSDs indeed have a read speed of 550 mb/s.

Exercise 2

Data: You’re working as maintenance engineer at Samsung. The new SSDs that they are making have an advertised read speed of 550mb/s. But after purchasing some users are reporting much lower read speeds with different values.

To verify their claims, you decided to gather 100 SSDs ready to be shipped to stores from your production line. After careful testing, the read speed of these test SSDs were found to be **548.7 mb/s.** You know from previous tests that the std deviation of your SSD model is **6.6 mb/s.**

Question: Test the claim by the customers with **90% confidence** and establish if the advertised read speed of 550 mb/s is acceptable or not.

Change in Step 2

n = 100

$\bar{x}$ **= 548.7**

σ **= 6.6**

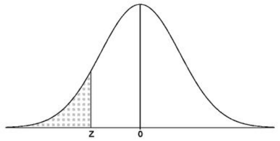
α = 1 - C.

Here, C = 0.90. So, α = 1 - 0.90 = 0.1

Since it's a two tailed test, α/2 = 0.1/2 = 0.05

Change in step 4

Cumulative Standard Normal Distribution Table



Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-0.00	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
-0.10	0.4602	0.4562	0.4522	0.4483	0.4443	[illegible]	0.4364	0.4325	0.4286	0.4247
-0.20	0.4207	0.4168	0.4129	0.4090	0.4052	[illegible]	0.3974	0.3936	0.3897	0.3859
-0.30	0.3821	0.3783	0.3745	0.3707	0.3669	[illegible]	0.3594	0.3557	0.3520	0.3483
-0.40	0.3446	0.3409	0.3372	0.3336	0.3300	[illegible]	0.3228	0.3192	0.3156	0.3121
-0.50	0.3085	0.3050	0.3015	0.2981	0.2946	[illegible]	0.2877	0.2843	0.2810	0.2776
-0.60	0.2743	0.2709	0.2676	0.2643	0.2611	[illegible]	0.2546	0.2514	0.2483	0.2451
-0.70	0.2420	0.2389	0.2358	0.2327	0.2296	[illegible]	0.2236	0.2206	0.2177	0.2148
-0.80	0.2119	0.2090	0.2061	0.2033	0.2005	[illegible]	0.1949	0.1922	0.1894	0.1867
-0.90	0.1841	0.1814	0.1788	0.1762	0.1736	[illegible]	0.1685	0.1660	0.1635	0.1611
-1.00	0.1587	0.1562	0.1539	0.1515	0.1492	[illegible]	0.1446	0.1423	0.1401	0.1379
-1.10	0.1357	0.1335	0.1314	0.1292	0.1271	[illegible]	0.1230	0.1210	0.1190	0.1170
-1.20	0.1151	0.1131	0.1112	0.1093	0.1075	[illegible]	0.1038	0.1020	0.1003	0.0985
-1.30	0.0968	0.0951	0.0934	0.0918	0.0901	[illegible]	0.0869	0.0853	0.0838	0.0823
-1.40	0.0808	0.0793	0.0778	0.0764	0.0749	[illegible]	0.0721	0.0708	0.0694	0.0681
-1.50	0.0668	0.0655	0.0643	0.0630	0.0618	[illegible]	0.0594	0.0582	0.0571	0.0559
-1.60	[illegible]	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.70	0.0446	0.0436	0.0427	0.0418	0.0409	[illegible]	0.0392	0.0384	0.0375	0.0367
-1.80	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.90	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-2.00	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-2.10	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.20	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.30	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.40	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.50	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.60	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.70	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.80	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.90	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-3.00	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-3.10	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.20	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.30	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.40	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.50	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
-3.60	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.70	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.80	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001

*Note: z-values less than -3.89 produce a probability of zero.

Department of Mathematics, Sinclair Community College, Dayton, OH

Download the full table here.

Here, the left and right critical values are 1.65 (since 0.0495 is closest to 0.05, rounding to 2 decimal points)

Exercise 3

Data: You’re working as maintenance engineer at Samsung. The new SSDs that they are making have an advertised read speed of 550mb/s. But after purchasing some users are reporting much lower read speeds with different values.

To verify their claims, you decided to gather only **20 SSDs** ready to be shipped to stores from your production line. After careful testing, the read speed of these test SSDs were found to be 548.75 mb/s. You know from previous tests that the sample standard deviation of your SSDs is 5.9 mb/s.

Question: Test the claim by the customers with **80% confidence** and establish if the advertised read speed of 550 mb/s is acceptable or not.

Change in Step 2

n = 20

$\bar{x}$ = 548.75

s = 5.9 (sample standard deviation given, σ not given). Just put s in place of σ in step 5.
And of course, this means we'll have to use the t table.

α = 1 - C.

Here, C = 0.80. So, α = 1 - 0.80 = 0.20

We can directly find corresponding critical value for $t_{0.80}$ from the t table.

Change in step 4

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

Download the full t table here.

Here, the critical value is 1.328 for 80% confidence.

Check out the following exercises for practise.

Remember to choose which test to use and which table to use carefully.

Interpret the test result properly.

Exercise 4

Say you're working as a production system engineer at Cadbury International at their Dairy Milk Silk production line. Their 150g chocolate bars are really famous. However, recently some retailers are reporting that customers are complaining that the bars are allegedly getting smaller!

You decided to test whether the fault was with your production line or was the customer's were just too fond of the chocolate bars (and the size issue was merely psychological!).

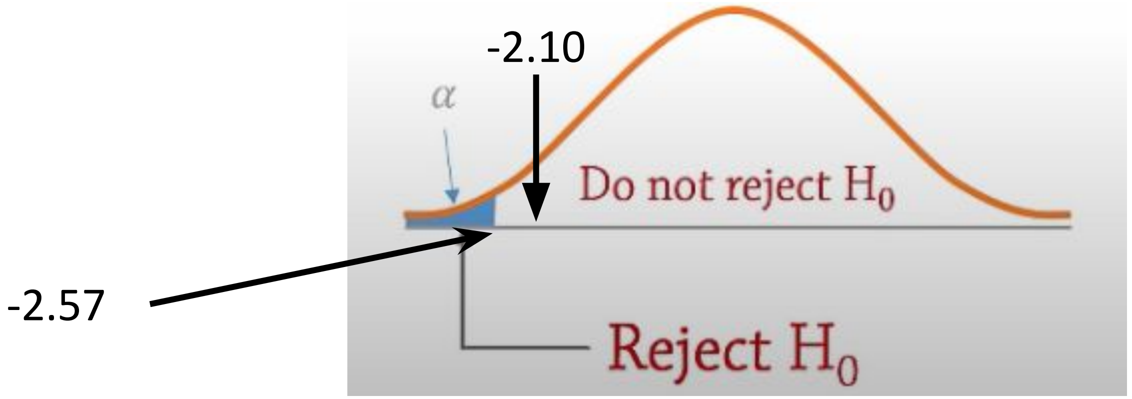
To test, you pulled 250 chocolate bars from the production line. After careful measurement, the average weight of them were 149.8 g. The standard deviation of the 250 sample chocolate bars was 1.5g.

Now, can you **defend** with 99.5% level of confidence that the mean weight of your chocolate bars are **at least 150g?**

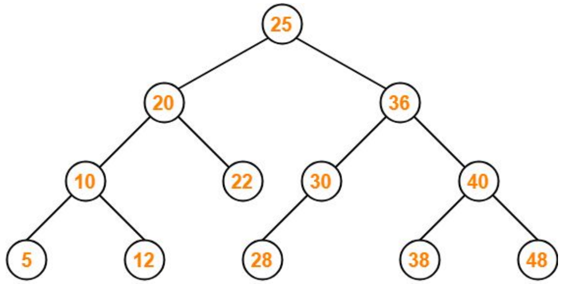
Try yourself first. Hints and final results are given on next slide.

Exercise 4

1. Use z table, since the sample size is pretty large.
2. Use a left tailed test. Since the claim by cadbury (H_0) is μ >=150 (at least 150g). You'll alternately try to prove the claim wrong by proving (H_1) μ <150.
3. Calculate critical value. One tailed test. So, **α** = 1 - 0.995 = 0.005. From z table critical value is -2.57.
4. Calculate test statistics "z". It's z = -2.10.
5. Since -2.10>-2.57. So. the value is in acceptance region. So, the candy bars are indeed acceptable as at least 150g.



Exercise 5



Binary Search Tree

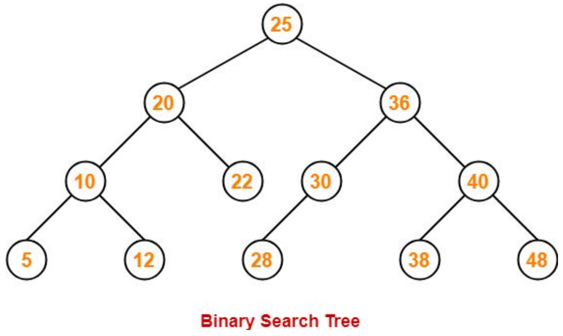
Suppose that, in your undergraduate thesis, you designed an algorithm that searches a BST with a depth of at most 20 **within** an avg of 2.25 ms (max time). However, in your pre-defence the defence team ran your algorithm and found the search time to be 2.51 ms for a particular BST.

You, very confident in your algorithm, claimed that the problem might be with the particular machine's specification. To prove that your algorithm will search a node within 2.25ms, irrespective of machine specifications, you decided to test the code again using the same tree and same search node in 5 different machines of varied specs.

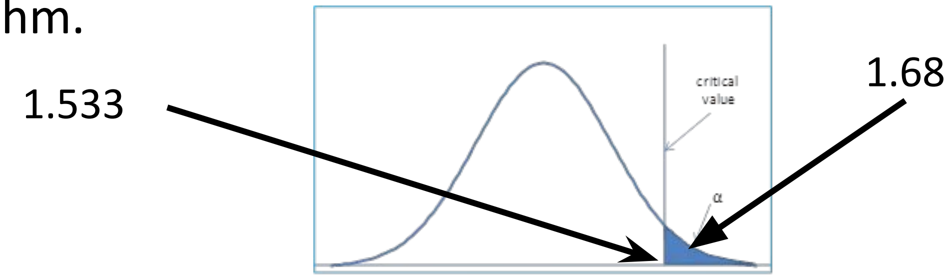
The search time for each of the 5 machines are 2.23ms, 2.21ms, 2.36ms, 2.44ms and 2.55ms. Now, based on this test data, with 90% level of confidence (or $\alpha = 0.1$), will you be able to defend your claimed 2.25ms search time or would you need to revise it in your final defence?

Try yourself first. Hints and final results are given on next slide.

Exercise 5



1. Use t table, since the sample size is very low and σ is also unknown.
2. Use a **right tailed** test. Since your claim is that the search time is at most 2.25ms i.e. H_0: μ <=2.25 ms (at most). You'll alternately try to prove the claim wrong by proving (or in this case hoping your test data doesn't prove!) μ > 2.25 (H_1).
3. Calculate critical value. One tailed test. So, α = 1 - 0.90 = 0.01. From t table, critical value is 1.533 (for df =4 and $t_{0.90}$)
4. Find $\bar{x}$ and s from respective eqns (check slide 13). $\bar{x}$ =2.358. Variance s^2 =.0247. So, s = 0.143.
5. Now, Calculate test statistics "z". z = 1.68.
6. Since 1.68>1.533 the value is in rejection region just by a bit. So, sadly, looks like you'll have to revise your claimed time or improve your algorithm.



Thank You!