Hypothesis Testing

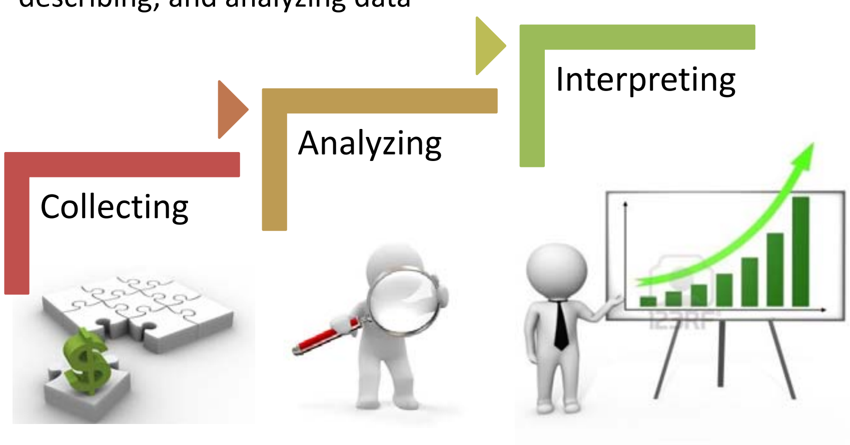
Simon Sung Jin, Ph.d

Announcement

- **Case write-up Grading**
  - Great in summarizing as well as answering the questions.
  - Basically Acceptable Vs. non-acceptable
  - Late submission penalty (50%)
- **Quiz**

What is the Statistics? …

- In general, “STATISTICS” is the science of gathering, describing, and analyzing data



Population

- In statistics, the group we wish to study is called the **population**.
- A population is:
- Defined by what the researcher is studying.
- The total number of subjects or things we are interested in studying.

Frame …

- In statistics, a **frame** is a list containing all members of the population.
- A strict definition of census is a survey that includes all elements or units in the frame.

Since a frame is the whole list which contains every members of the population, it would be much easier to develop when the population size is small or big?

Frame

- Examples of Frames:
- The frame for the population of the U.S. would be a rather long list containing over 300 million names.
- If MIS687 class were the population under consideration, the class roll would be the frame.
- For the population of registered voters, the electoral register would be the frame

Population Parameter …

- **Population parameters** are facts about the population.
- Parameters are descriptions of the population.
- A population can have many parameters.
- Parameters can be in the form of percentages, maximums, minimums, or other characteristics.

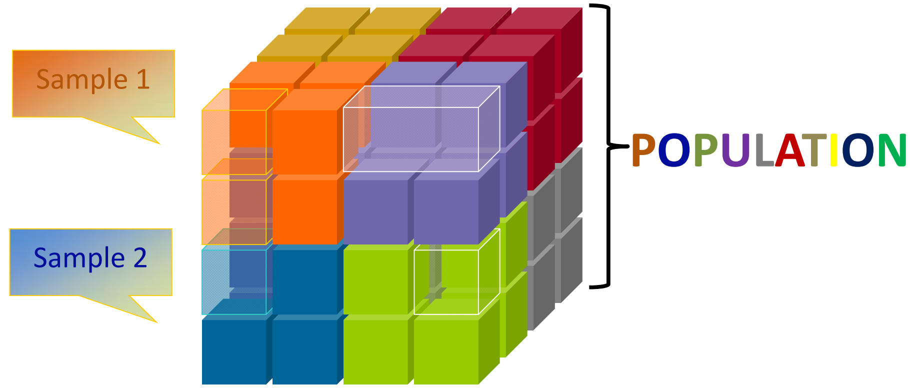
Population Parameter

- Examples of Population Parameters:
- 67% of Americans 20 and over are overweight.
- 7 out of 10 Americans do not exercise regularly.
- 70 million Americans suffer from a sleep disorder.

Sample …

- A **sample** is a subset of the population which is used to gain insight about the population. Samples are used to represent a larger group, the population.
- For example, the percentage of votes a presidential candidate received on Election Day is a parameter. Sample data is used to try to estimate this population parameter.
- Another example is sampling a college campus to represent the population of college students in the U.S.

Sample



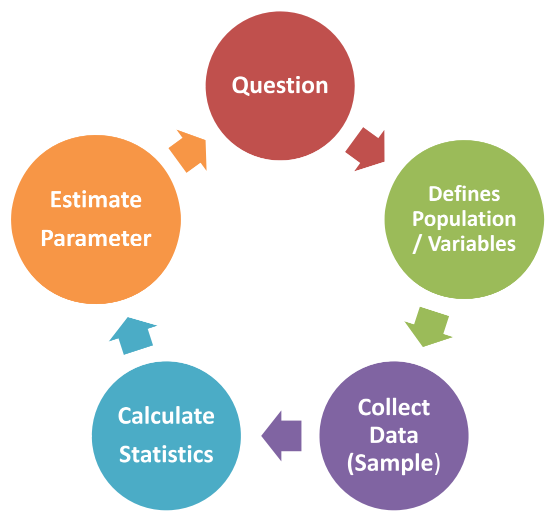
What is the Statistics? …

- A **statistic** is a fact or characteristic about a sample.

Population	Americans
Sample	All college students
Statistics	67% are sleep deprived

- For any given **sample,** a **statistic** is a fixed number.
- Because there are lots of different samples that can be drawn from the **population**, statistics vary depending on the sample collected.
- Statistics are used as estimates of **population parameters**.

What is the Statistics?



Branches of Statistics …

- **Descriptive statistics** – the collection, organization, analysis, and presentation of data.
  - The emphasis in descriptive statistics is analyzing observed measurements usually from a sample.
  - To comprehend a large set of data, it must be summarized. Descriptive techniques are the most common statistical applications.
- **Inferential statistics** – uses descriptive statistics to estimate population parameters; an educated guess about the population based on sample data.

Branches of Statistics …

- **Example 1:**
- In a survey of 100 students, 83.2% of students are happy with the food in the cafeteria.
- Identify the descriptive statistic(s). What inferences can be made?

Descriptive statistic?
83.2% of the 100 students surveyed are happy with the food in the cafeteria.

Possible inference?
83.2% of all students are happy with the food in the cafeteria.

Branches of Statistics

- **Example 2:**
- A heart researcher is interested in studying the relationship between diets which are high in calcium and blood pressure in adult females. The researcher randomly selects 20 female subjects who have high blood pressure. Ten subjects are randomly assigned to try a diet which is high in calcium. The other subjects are assigned to a diet with a standard amount of calcium. After one year the average blood pressures for subjects in both groups will be measured and compared to decide if diets high in calcium decrease the average blood pressure.
- Identify the population. Adult females
- What characteristic of the population is being measured? Average blood pressure
- Identify the sample. 20 females who have high blood pressure
- Is the purpose of the data collection to perform descriptive or inferential statistics? Inferential statistics

The Normal Distribution

Normal Distribution …

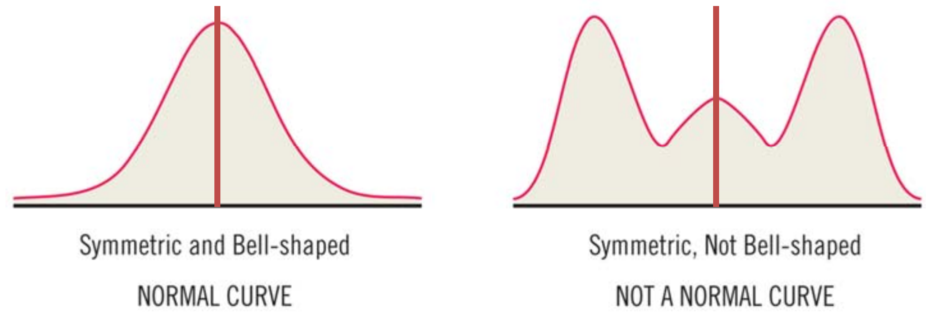
- Normal Distribution
  - A continuous probability distribution for a given random variable, X, that is completely defined by its mean and variance.

- Properties of a Normal Distribution:

1. A normal curve is symmetric and bell-shaped.
2. A normal curve is completely defined by its mean, μ, and variance, σ^2.
3. The total area under a normal curve equals 1.
4. The x-axis is a horizontal asymptote for a normal curve.

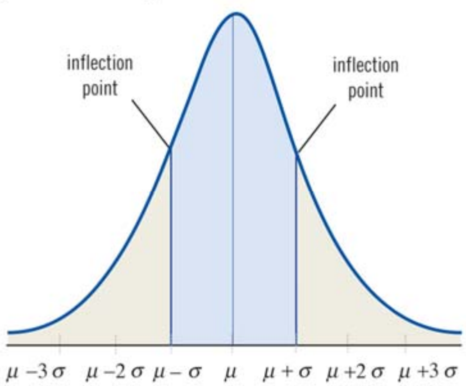
Normal Distribution …

- Symmetric and Bell-shaped?



Normal Distribution …

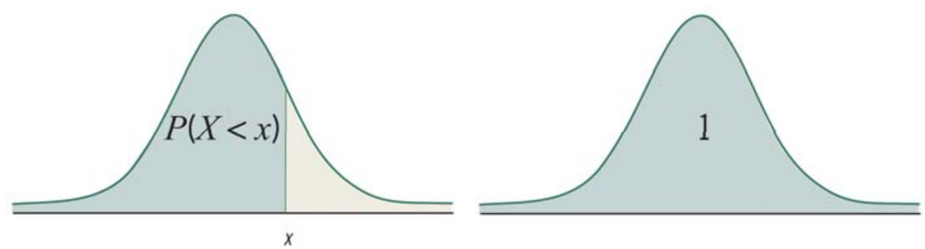
- Completely Defined by its Mean and Standard Deviation



An **inflection point** is a point on the curve where the curvature of the line changes. The inflection points are located at $\mu - \sigma$ and $\mu + \sigma$.

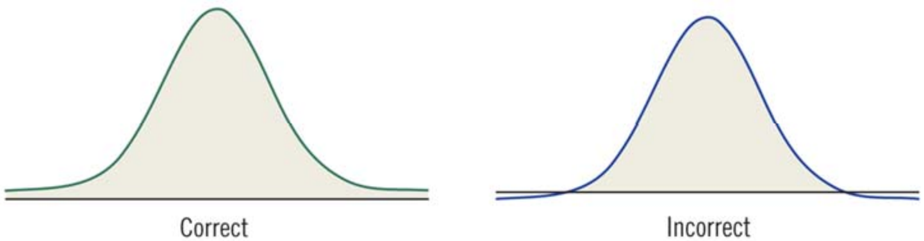
Normal Distribution …

- Total Area Under the Curve = 1



Normal Distribution …

- The x-Axis is a Horizontal Asymptote

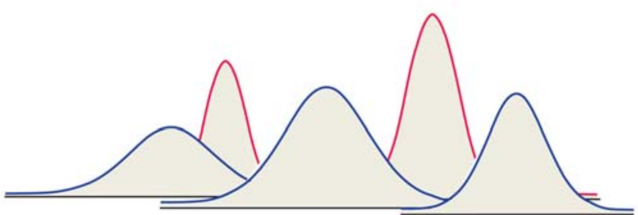


Normal Distribution …

- Example
  - Determine if the following is a normal distribution
    a. Birth weights of 75 babies.
       Normal
    b. Ages (in year) of 250 students in 10^{th} grade.
       No, this would be almost uniform
    c. Heights of 100 random adult males.
       Normal
    d. Frequency of outcomes from rolling a die.
       No, because the data is discrete
    e. Weights of 50 random fully grown tigers.
       Normal

Normal Distribution …

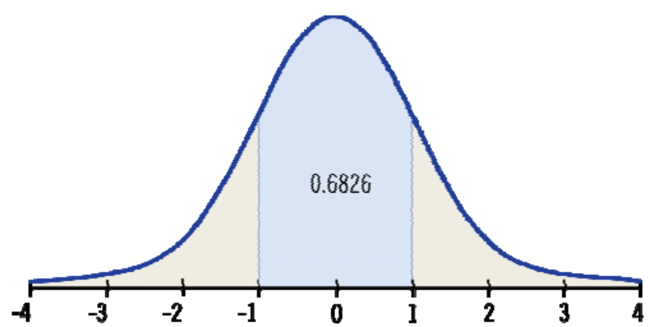
- How Many Normal Curves are there?



Because there are an infinite number of possibilities for μ and σ, there are an infinite number of normal curves.

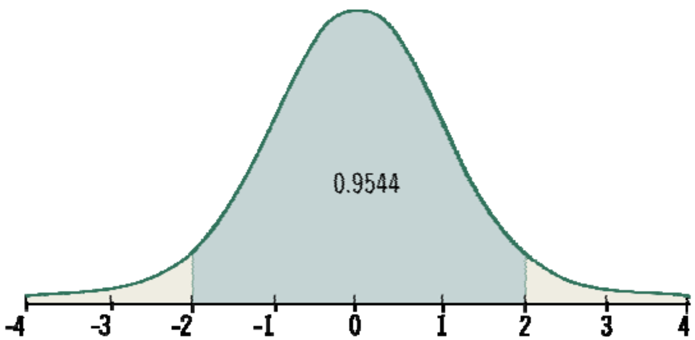
Normal Distribution …

- Area within One Standard Deviation
  - The area under the curve and the probability of being within one standard deviation $(\pm 1\sigma)$ of the mean, μ, equals 0.6826.



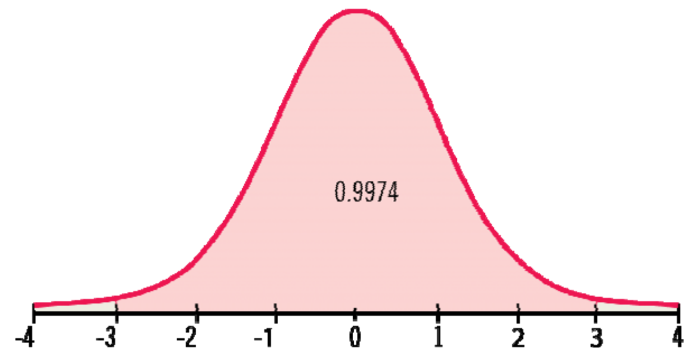
Normal Distribution …

- Area within Two Standard Deviations
  - The area under the curve and the probability of being within two standard deviations $(\pm 2\sigma)$ of the mean, μ, equals 0.9544.



Normal Distribution …

- Area within Three Standard Deviations
  - The area under the curve and the probability of being within three standard deviations (± 3σ) of the mean, μ, equals 0.9974.



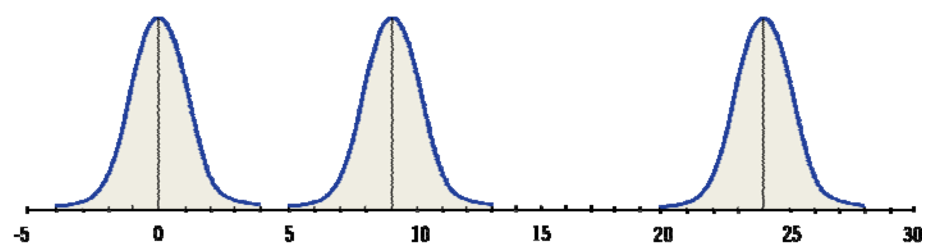
Normal Distribution …

- Definition
  - Normal distribution – a continuous probability density function completely defined by its mean and variance.

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

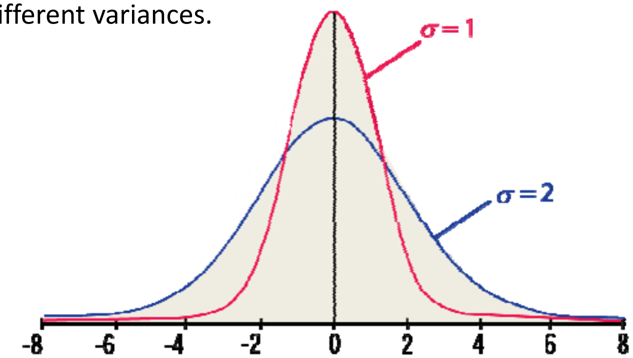
Normal Distribution …

- Normal Curves
  - The mean defines the location and the variance determines the dispersion.
  - Below are three different normal curves with different means and identical variances.



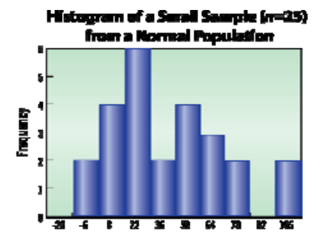
Normal Distribution …

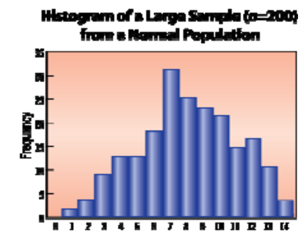
- Normal Curves
  - Changing the variance parameter can have rather significant effects on the shape of the distribution
  - Below are two different normal curves with identical means and different variances.



Normal Distribution

- Data from Normal Distributions
  - As the following three histograms demonstrate, data from a population that is assumed to come from a normal population will more closely represent a bell curve as the sample size n grows larger





The Standard Normal Distribution

Standard Normal Distribution

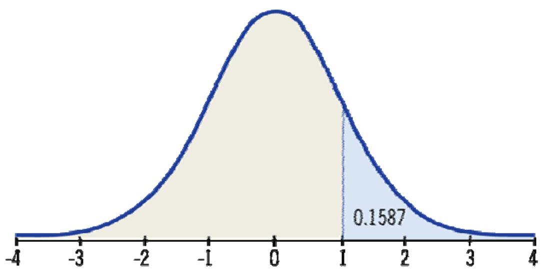
- A standard normal distribution has the same properties as the normal distribution; in addition, it has a mean of 0 and a variance of 1.

- Properties of a Standard Normal Distribution

1. The standard normal curve is symmetric and bell-shaped.
2. It is completely defined by its mean and standard deviation,
   $\mu = 0$ and $\sigma^2 = 1$.
3. The total area under a standard normal curve equals 1.
4. The *x*-axis is a horizontal asymptote for a standard normal curve.

Reading a Normal Curve Table …

- Probability of a Normal Curve
  - The probability of a random variable having a value in a given range is equal to the area under the curve in that region.



$P(X>1) = 0.1587$

Reading a Normal Curve Table …

- Standard Normal Distribution Table

Standard Normal Distribution Table from –∞ to positive z					
z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**
0.0	0.5000	0.5040	0.5080	0.5120	0.5160
0.1	0.5398	0.5438	0.5478	0.5517	0.5557
0.2	0.5793	0.5832	0.5871	0.5910	0.5948
0.3	0.6179	0.6217	0.6255	0.6293	0.6331
0.4	0.6554	0.6591	0.6628	0.6664	0.6700
0.5	0.6915	0.6950	0.6985	0.7019	0.7054
0.6	0.7257	0.7291	0.7324	0.7357	0.7389
0.7	0.7580	0.7611	0.7642	0.7673	0.7704
0.8	0.7881	0.7910	0.7939	0.7967	0.7995

Reading a Normal Curve Table …

- Standard Normal Distribution Table

1. The standard normal tables reflect a *z*-value that is rounded to two decimal places.

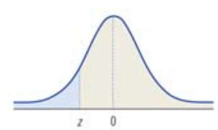
2. The first decimal place of the *z*-value is listed down the left-hand column.

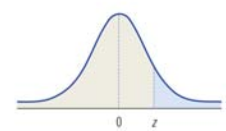
3. The second decimal place is listed across the top row.

4. Where the appropriate row and column intersect, we find the amount of area under the standard normal curve to the **left** of that particular *z*-value.

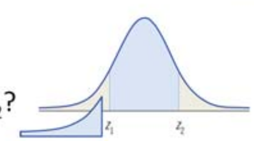
Reading a Normal Curve Table …

- Possible cases of reading the table
  - Area to the Left of z?

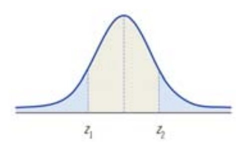




  - Area to the Right of z?
  - Area Between z_1 and z_2?

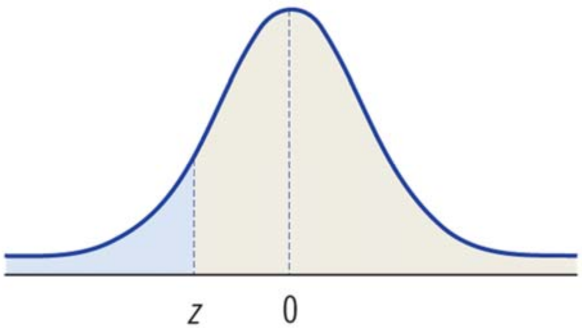


  - Area in the Tails?



Reading a Normal Curve Table …

- Area to the Left of z?



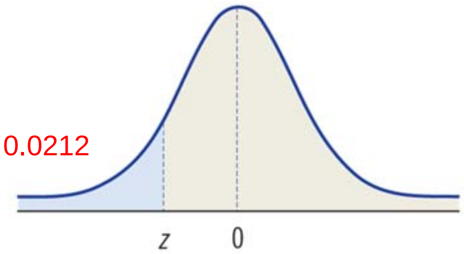
Reading a Normal Curve Table …

- Find the area to the left of z

  a. $z = 1.69$
  0.9545

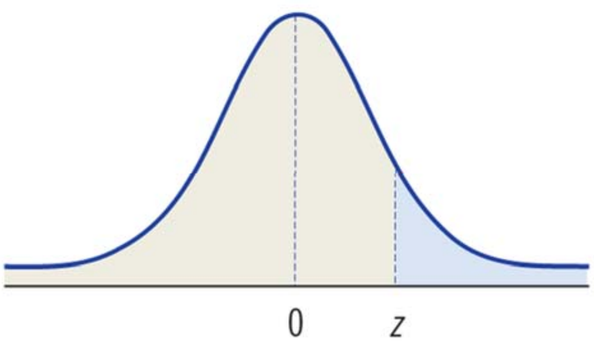
  b. $z = –2.03$
  1 – 0.9788 = 0.0212

  c. $z = 0$
  0.5000



Reading a Normal Curve Table …

- Area to the Right of z?



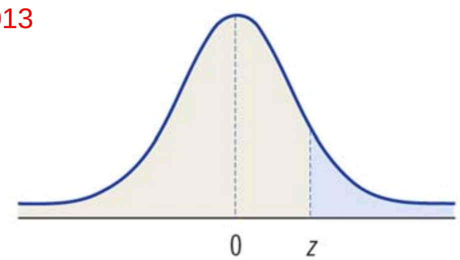
Reading a Normal Curve Table …

- Find the area to the Right of z

  a. $z = 3.02$
  1-0.9987 = 0.0013

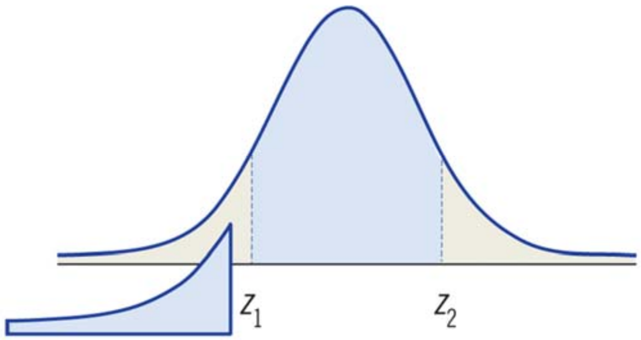
  b. $z = –1.70$
  0.9554

  c. $z = 0$
  1-0.5 = 0.5000



Reading a Normal Curve Table …

- Area Between z_1 and z_2?



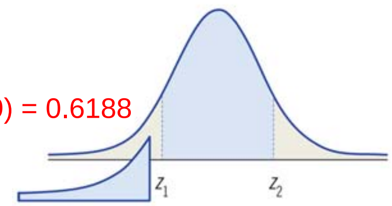
Reading a Normal Curve Table …

- Find the area Between z_1 and z_2

  a. $z_1 = 1.16$, $z_2 = 2.31$
     0.9896-0.8770 = 0.1126

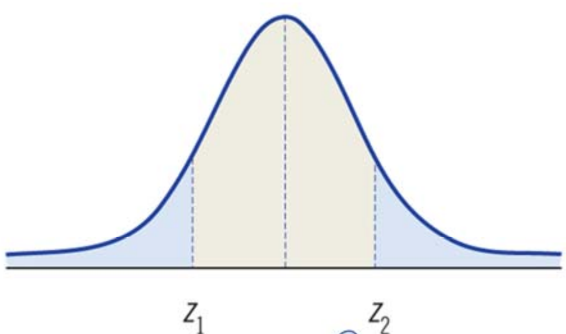
  b. $z_1 = -2.76$, $z_2 = 0.31$
  0.6217 - (1- 0.9971 = 0.0029) = 0.6188

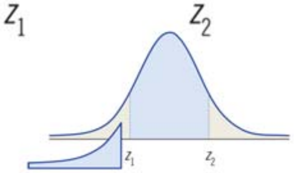
  c. $z_1 = -3.01$, $z_2 = -1.33$
     0.0918-0.0013 = 0.0905



Reading a Normal Curve Table …

- Area in the Tails?



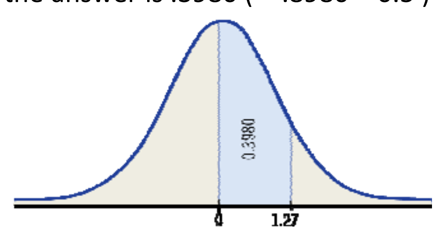
1 -  = Answer

Example 2

- Calculate the probability that a standard normal random variable is between 0 and 1.27.

  Solution

  - Look up the value of 1.27 in the table.
  - The table value of .8980 is the area under the curve left of 1.27.
  - So, the answer is .3980 (= .8980 – 0.5)

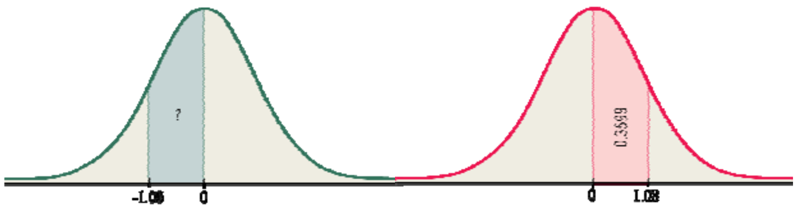


Example 2

- Calculate the probability that a standard normal random variable is between −1.08 and 0.

Solution

- The value −1.08 is not given in the table.
- Since the distribution is symmetric, the probability that the random variable is between −1.08 and 0 is equal to the probability the random variable is between 0 and 1.08.
- The table value of 0.3599 is the area under the curve between 0 and 1.08.



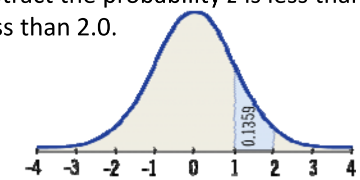
$P(-1.08 < z < 0) \quad = \quad P(0 < z < 1.08) \quad = .8599 - .5 \quad = .3599$

Example 3

- Calculate the probability that a standard normal random variable is between 1.0 and 2.0.

Solution

- First determine the probability that *z* is less than 2.0, which the table gives as .9772.
- Then determine the probability that *z* is less than 1.0, which the table gives as .8413.
- The final step is to subtract the probability *z* is less than 1.0 from the probability that *z* is less than 2.0.



$P(1 < z < 2)$

.1359

z-Transformations

z-Transformations …

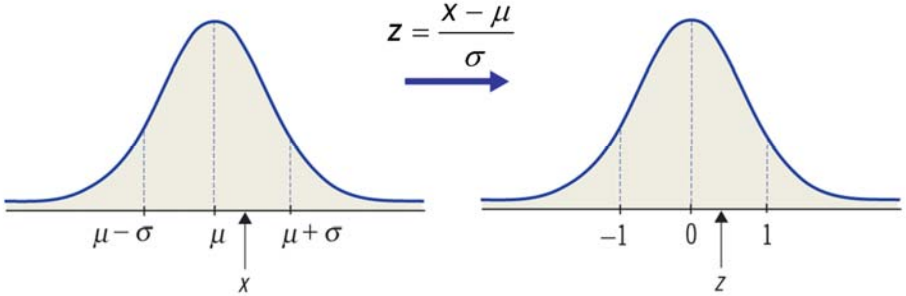
- Definition
  - z-Transformation – a transformation of any normal variable into a standard normal variable. The z-transformation is denoted by z and is given by the formula

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

z-Transformations …

- Converting to the Standard Normal Curve

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

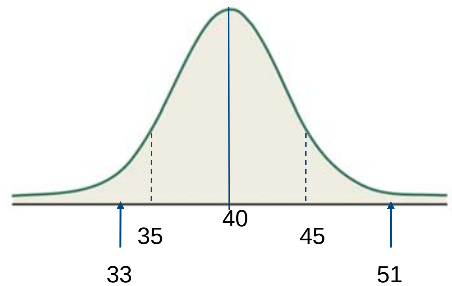


z-Transformations …

- Example 1
  - Given μ = 40 and σ = 5, indicate the mean, each of the inflections points, and where each given value of *x* will appear on the curve

    x_1 = 33 and x_2 = 51

Solution:



z-Transformations

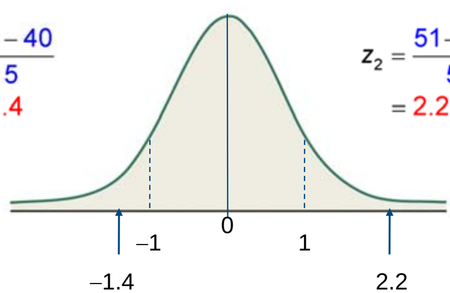
- Example 2
  - Given μ = 40 and σ = 5, calculate the standard score for each x value and indicate where each would appear on the standard normal curve

    x_1 = 33 and x_2 = 51

Solution:

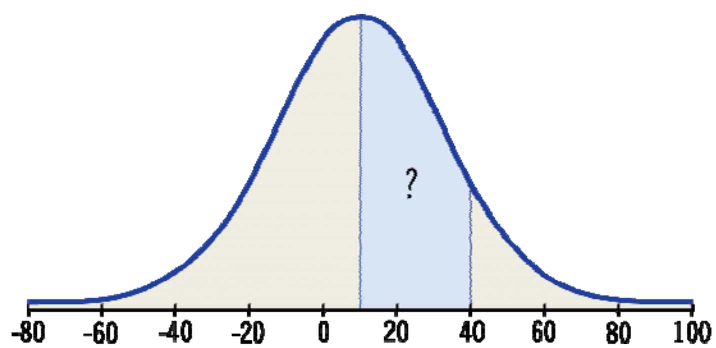
$$z_1 = \frac{33 - 40}{5} = -1.4$$

$$z_2 = \frac{51 - 40}{5} = 2.2$$



Example 1 …

- Calculate the probability that a normal random variable with a mean of 10 and a standard deviation of 20 will lie between 10 and 40.



Example 1

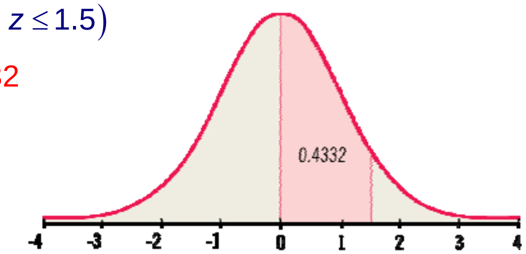
Solution

Applying the z-transformation yields

$$P(10 \le X \le 40) = P\left(\frac{(10-10)}{20} \le \frac{x-\mu}{\sigma} \le \frac{(40-10)}{20}\right)$$

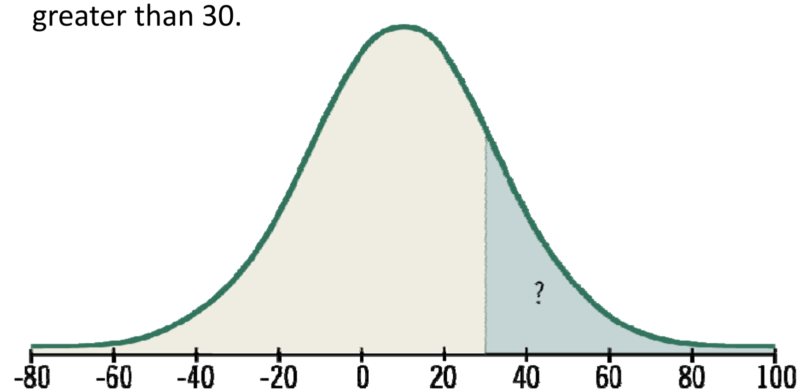
$$= P(0 \le z \le 1.5)$$

$$= 0.4332$$



Example 2 …

- Calculate the probability that a normal random variable with a mean of 10 and a standard deviation of 20 will be greater than 30.



Example 2

Solution

Applying the *z*-transformation yields

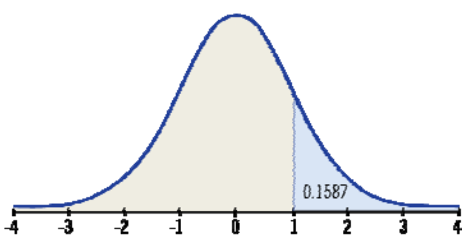
$$P(X > 30) = P\left(z > \frac{30-10}{20}\right)$$

$$= P(z > 1)$$

$$P(z > 1) = 1 - P(z<1)$$

$$= 1 - .8413$$

$$= .1587$$

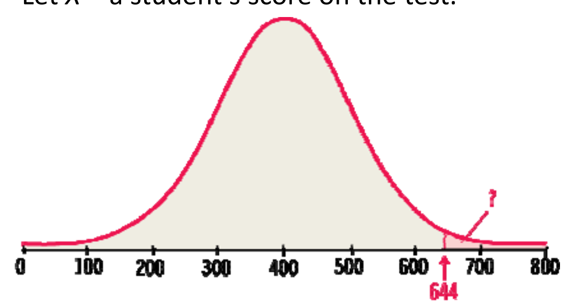


Example 3 …

- Suppose that a national testing service gives a test in which the results are normally distributed with a mean of 400 and a standard deviation of 100. If you score a 644 on the test, what fraction of the students taking the test exceeded your score?

Solution

Let *X* = a student's score on the test.



Example 3

Solution

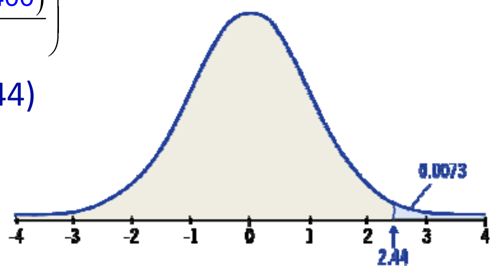
The first step is to apply the z-transformation

$$P(X > 644) = P\left(z > \frac{(644-400)}{100}\right)$$

$= 1 - P\,(z<2.44)$

$= 1 - .8927$

$= .0073$



Thus, only 0.73% of the students scored higher than your score of 644.

Developing a Hypothesis

Definitions

- **Hypothesis** – a theory or premise, usually the claim that someone is investigating.
- **Null Hypothesis, H_0** – describes the currently accepted value for the population parameter.
- **Alternative Hypothesis, H_a** – describes the claim that is being tested; the mathematical opposite of the null hypothesis.
- **Hypothesis Test** – compares the merit of the two competing hypotheses by examining the data that is collected

Hypothesis Test

- In a hypothesis test, the null hypothesis is always given the benefit of the doubt.
- That is, we assume that the null hypothesis is true unless there is overwhelming evidence from the sample that goes against the null hypothesis.
- Hypothesis Testing is formalized decision making
  - Choose between **REJECTING** & **FAILING TO REJECT** a hypothesis on the basis of a set of observations

Rejecting & Accepting – Wrong!!!

- Acceptance implies that the null hypothesis is true.
- Failure to reject implies that the data are not sufficiently persuasive for us to prefer the alternative hypothesis over the null hypothesis

Developing a Hypothesis …

- To write the null and alternative hypotheses:

1. Write down the claim
2. Write down the mathematical opposite
3. Assign the null and alternative labels

Remember that the null hypothesis is the currently accepted value for the population parameter.

Developing a Hypothesis …

- Example 1: Determine the null and alternative hypotheses
  - It is generally accepted among leading educators that the average student studies no more than 15 hours per week. In the spring newsletter, a national student organization claims that its members study more per week than the average student

Solution:

Claim: Members of the national student organization study more per week than the average student.

$\mu > 15$

Mathematical opposite: $\mu \leq 15$

H_0: $\mu \leq 15$ Current accepted belief

H_a: $\mu > 15$ Testing hypothesis

Developing a Hypothesis …

- Example 2: Determine the null and alternative hypotheses
  - A leading news authority claims that the President's job approval rating has dropped over the past 3 months. Previous polls put the President's approval rating at a minimum rate of 56%. The President's chief of staff is concerned about this claim since it is an election year, and he wants to run a test on the claim

Solution:

Claim: The President's approval rating has dropped lower than 56%.

$p < 0.56$

Mathematical opposite: $p \geq 0.56$

H_0: $p \geq 0.56$ Current accepted belief

H_a: $p < 0.56$ Testing hypothesis

Developing a Hypothesis

- Example 3: Determine the null and alternative hypotheses
  - Leading authorities have stated that approximately 1% of dogs attack people even if unprovoked. The animal rights organization believes that 1% is not accurate.

Solution:

Claim: 1% is not accurate for the percentage of dogs who attack people.

$p \neq 0.01$

Mathematical opposite: $p = 0.01$

H_0: $p = 0.01$ Current accepted belief

H_a: $p \neq 0.01$ Testing hypothesis

Definitions

- **Test Statistic** – a component of the criteria used to evaluate the hypothesis which is calculated from the sample data gathered
- **Statistically Significant** – Data is said to be statistically significant if it is unlikely that a sample similar to the one chosen would occur by chance if the null hypothesis is true
- **Level of Significance, α** – the probability of rejecting a true null hypothesis. It is also the complement to the level of confidence, α, where $\alpha = 1 - c$

Steps for Hypothesis Testing

1. State the null and alternative hypotheses
2. Set up the hypothesis test by choosing the test statistic and determining the values of the test statistic that would lead to rejecting the null hypothesis
3. Gather data and calculate the necessary sample statistics
4. Draw a conclusion

Conclusions for a Hypothesis Test …

1. Reject the null hypothesis.
2. Fail to reject the null hypothesis.

Once you have decided on a conclusion, a discussion of the meaning of this conclusion in terms of the original claim is appropriate.

Conclusions for a Hypothesis Test …

- Example 1 – Determine the conclusion
  - It is generally accepted among leading educators that the average student studies no more than 15 hours per week. In their spring newsletter, a national student organization claims that its members study more per week than the national average. After performing a hypothesis test at the 95% level of confidence to evaluate the claim of the organization, the researchers' conclusion is to reject the null hypothesis. Does this support the organization's claim?

Claim: Members of the national student organization study more per week than the average student.
$\mu > 15$

Conclusions for a Hypothesis Test …

- Solution 1

Claim: Members of the national student organization study more per week than the average student.
$\mu > 15$

Mathematical opposite: $\mu \leq 15$

H_0: $\mu \leq 15$ Current accepted belief
H_a: $\mu > 15$ Testing hypothesis

After performing a hypothesis test at the 95% level of confidence to evaluate the claim of the organization, the researchers' conclusion is to reject the null hypothesis.

Since the null hypothesis was rejected, the organization can be 95% confident in its claim that its members study more per week than the national average.

Conclusions for a Hypothesis Test …

- Example 2 – Determine the conclusion
  - The Board of Education for one large school district uses at least 10% as the percentage of high school sophomores considering dropping out of school. A high school counselor in this district claims that this percentage is too high. A hypothesis test with $\alpha = 0.02$ is performed on the counselor's claim. The result is to fail to reject the null hypothesis. Do the findings support the counselor's claim?

Claim: The drop out rate of at least 10% is too high.
$p < 0.10$

Mathematical opposite: $p \geq 0.10$

Conclusions for a Hypothesis Test

- Solution 2

Claim: The drop out rate of at least 10% is too high.
$p < 0.10$

Mathematical opposite: $p \geq 0.10$

H_0: $p \geq 0.10$ Current accepted belief
H_a: $p < 0.10$ Testing hypothesis

Since the null hypothesis was not rejected, the evidence is not strong enough at this level of significance to support the counselor's claim.

Reaching a Conclusion

Definitions

- **Type I Error** – rejecting a true null hypothesis; the probability of making a Type I error is denoted by $\boldsymbol{\alpha}$.

Inversely related

- **Type II Error** – failing to reject a false null hypothesis; the probability of making a Type II error is denoted by $\boldsymbol{\beta}$.

		Reality	
		H_0 is true	**H_0 is false**
Decision	**H_0 is rejected**	Type I error	Correct Decision
	H_0 is not rejected	Correct Decision	Type II error

Reaching a Conclusion …

- Example 1 - Determine if an error was made
  - A television executive believes that at least 99% of households own a television set. An intern at her company is given the task of testing the claim that the percentage is actually less than 99%. The hypothesis test is completed, and based on the sample collected the intern decides to fail to reject the null hypothesis.
  - If, in reality, 97.5% of households own a television set, was an error made? If so, what type?

First state the hypotheses.

H_0: $p \geq 0.99$
H_a: $p < 0.99$

Reaching a Conclusion …

- Solution 1

First state the hypotheses.

H_0: $p \geq 0.99$
H_a: $p < 0.99$

The decision was to fail to reject the null hypothesis.
In reality, the null hypothesis is false since less than 99% of households own a TV set.
Therefore, the intern failed to reject a false null hypothesis.
This is a Type II error.

		Reality	
		H_0 is true	**H_0 is false**
Decision	**H_0 is rejected**	Type I error	Correct Decision
	H_0 is not rejected	Correct Decision	**Type II error**

Reaching a Conclusion …

- Example 2 - Determine if an error was made
  - Insurance companies commonly use 1000 miles as the average number of miles a car is driven per month. One insurance company claims that due to our more mobile society, the average is more than 1000 miles per month. The insurance company tests their claim with a hypothesis test and decides to reject the null hypothesis.
  - Suppose that in reality, the average number of miles a car is driven per month is 1500 miles. Was an error made? If so, what type?

First state the hypotheses.

H_0: $\mu \leq 1000$
H_a: $\mu > 1000$

Reaching a Conclusion …

- Solution 2

First state the hypotheses.

H_0: $\mu \leq 1000$
H_a: $\mu > 1000$

The decision was to reject the null hypothesis.
In reality, the null hypothesis is false since the average number of miles is more than 1000.
Therefore, the decision was to reject a false null hypothesis.
This is a correct decision.

		Reality	
		H_0 is true	H_0 is false
Decision	H_0 is rejected	Type I error	Correct Decision
	H_0 is not rejected	Correct Decision	Type II error

Reaching a Conclusion …

- Example 3 - Determine if an error was made
  - A study of the effects of television viewing on children reports that children watch an average of 4 hours of television per night. A researcher believes that the average number of hours of television watched per night by children in her neighborhood is not really 4. She performs a hypothesis test and rejects the null hypothesis.
  - In reality, children in her neighborhood do watch an average of 4 hours of television per night. Was an error made? If so, what type?

First state the hypotheses.

H_0: $\mu = 4$
H_a: $\mu \neq 4$

Reaching a Conclusion …

- Solution 3

First state the hypotheses.

H_0: $\mu = 4$
H_a: $\mu \neq 4$

The decision was to reject the null hypothesis.
In reality, the null hypothesis is true since the children in her neighborhood watch an average of 4 hours of television per night.
Therefore, the researcher rejected a true null hypothesis.
This is a Type I error.

		Reality	
		H_0 is true	H_0 is false
Decision	H_0 is rejected	Type I error	Correct Decision
	H_0 is not rejected	Correct Decision	Type II error

Testing a Hypothesis About a Population Mean

In the case of n > 30

Definition

- Rules for a z-test for a population mean
  - If $n > 30$, or
  - the standard deviation of the population, σ, is known and the sample is drawn from a normal population,
  - then by the Central Limit Theorem the test statistic is given by

$$z = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}}, \text{ where } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}},$$

$\bar{x}$ is the sample mean,

μ_0 is the hypothesized value of the population mean,

and n is the sample size.

- If σ is unknown and $n > 30$, the sample standard deviation, s, can be used as an approximation of σ.
- The test statistic, z, has a standard normal distribution

Steps for Hypothesis Testing …

1) Define the hypotheses in plain English.
2) Select the appropriate statistical measure.
3) Determine whether the alternative hypothesis should be one-sided (> or <) or two-sided (= or ≠).
4) State the hypotheses using the statistical measure that reflects the hypotheses under consideration.
5) Specify α, the level of the test.
6) Select the appropriate test statistic $z = \frac{\bar{x} \quad \mu_0}{\sigma_{\bar{x}}}$, where $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$
   (A test statistic is a quantity calculated from our sample of data)

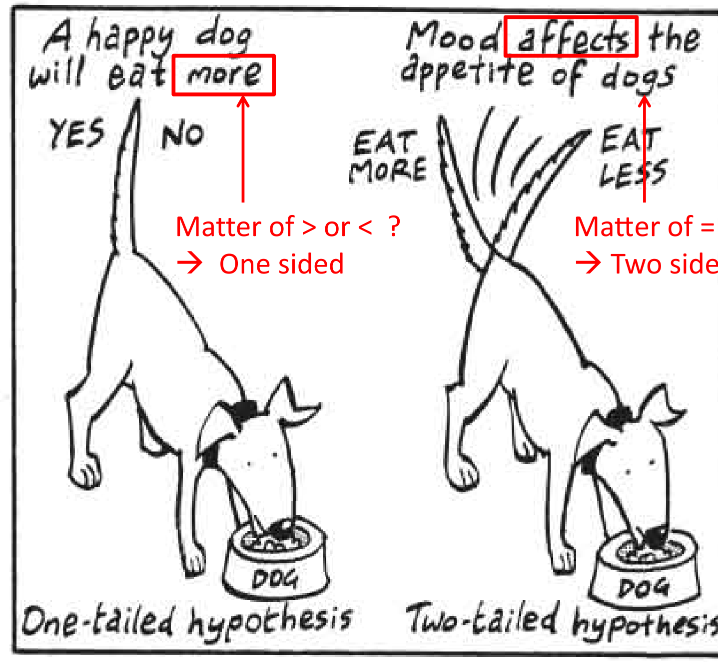
Steps for Hypothesis Testing

7) Determine the critical value of the test statistic.
(Critical value is a threshold to which the value of the test statistic in a sample is compared to determine whether or not the null hypothesis is rejected. Critical value corresponds to a given significance level)

8) Compute the test statistic using the sample data.

9) Make the decision (Compare Test Statistic and the critical value)
   - If Test Statistic > Critical Value → Reject H_0
   - If Test Statistic < Critical Value → Fail to reject H_0

10) State the conclusion in terms of the original question.



Hypothesis Testing …

- Example 1
  - Suppose the national average reading speed for high school sophomores is 150 words per minute with a standard deviation of 15. A local school board member wants to know if sophomore students at Lincoln High School read at a level different from the national average for tenth graders. The level of the test is to be set at 0.05. A random sample of 100 tenth graders has been drawn, and the sample mean is 157 words per minute.

Hypothesis Testing …

- Solution 1

1) H_0: Lincoln High School tenth graders are reading at the national average.
   H_a: Lincoln High School tenth graders are *not* reading at the national average.

2) μ = average number of words read per minute by Lincoln High School sophomores.

3) The key word here is "different." Since the board member is interested in whether there is any difference, the test is two-sided.

Hypothesis Testing …

- Solution 1 – Contd

4) $H_0: \mu = 150 \quad H_a: \mu \neq 150$

5) $\alpha = 0.05$

6) Since $n > 30$, we can assume that the sampling distribution of $\bar{x}$ is approximately normally distributed and therefore we can use the z-test.

7) Since $\alpha = 0.05$ and the test is two-sided, $z_{\alpha/2} = z_{0.025}$

Note that 1 - 0.025 = 0.975
You need to find 0.975 in the table → 1.96.

Hypothesis Testing …

- Solution 1 – Contd

8) $z = \dfrac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \dfrac{157 - 150}{\frac{15}{\sqrt{100}}} = \dfrac{7}{\frac{15}{10}} \approx 4.67$

9) Since $|4.67| > 1.96 = z_{\alpha/2}$, reject the null hypothesis.

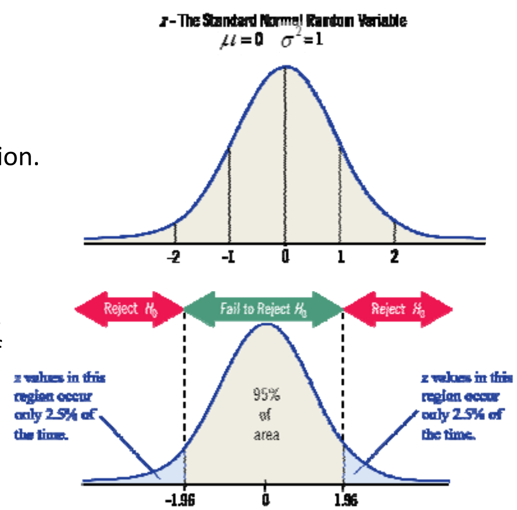
10) There is significant evidence at the 0.05 level that tenth graders at Lincoln High School do not read at the national average.

Hypothesis Testing …

- Solution 1 – Contd

The z-test statistic has a standard normal distribution.

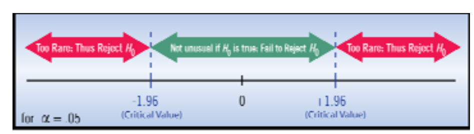
If the null hypothesis is true, 95% of the time the value of the z-test statistic will be between -1.96 and 1.96.



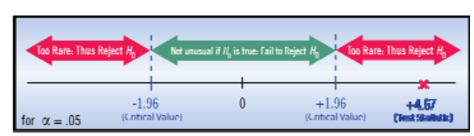
Hypothesis Testing …

- Solution 1 – Contd

The rejection regions can be graphed on the real number line, as follows.



Then the test statistic can be compared against the rejection regions on the real number line.



Hypothesis Testing …

- Example 2
  - A microprocessor designer has developed a new fabrication process which he believes will increase the usable life of a chip. Currently the usable life is 16,000 hours with a standard deviation of 2500. Test the hypothesis that the process increases the usable life of a chip, at the 0.01 level. A sample of 1000 microprocessors is tested and the mean is found to be 16,500. Assume the standard deviation of the life of the new chips will be equal to the standard deviation of the current chips.

Hypothesis Testing …

- Solution 2

1) H_0: The new process does *not* increase the life of the chip.
   H_a: The new process does increase the life of the chip.

2) μ = the mean life of the newly fabricated chips.

3) The key word here is "increase." Since the goal is to determine if there sufficient evidence that the new fabrication process increases the usable life of a chip, the test is one-sided.

4) $H_0 : \mu \leq 16{,}000$ $\quad$ $H_a : \mu > 16{,}000$

5) $\alpha = 0.01$

Hypothesis Testing …

- Solution 2 – Contd

6) Since $n > 30$, we can assume that the sampling distribution of $\overline{x}$ is approximately normally distributed and therefore we can use the z-test.

7) Since $\alpha = 0.01$ and the test is one-sided, $z_\alpha = z_{0.01} = 2.33$.

8) $z = \dfrac{\overline{x} - \mu_0}{\dfrac{\sigma}{\sqrt{n}}} = \dfrac{16{,}500 - 16{,}000}{\dfrac{2500}{\sqrt{1000}}} \approx \dfrac{500}{\dfrac{2500}{31.622777}} \approx 6.32$

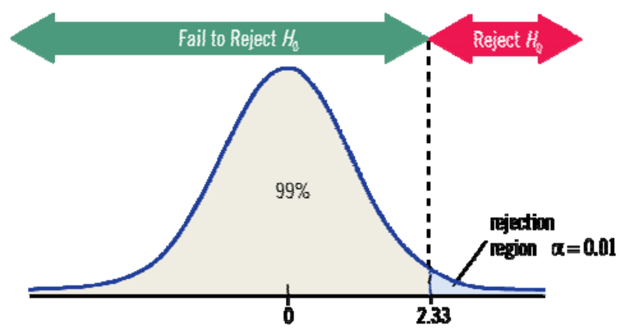
9) Since $6.32 > 2.33 = z_\alpha$, reject the null hypothesis.

10) There is significant evidence at the 0.01 level that the new fabrication process increases the usable life of a chip.

Hypothesis Testing

- Solution 2 – Contd

If the null hypothesis is true, 99% of the time the value of the z-test statistic will be less than 2.33.



Critical Values of the z-Test Statistic …

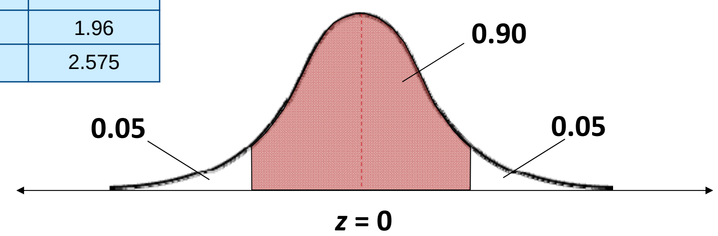
- Two sided (= Two tailed)

Critical Values of the *z*-Test Statistic for Two-Sided Alternatives		
Level of Test	**Definition of Ordinary Variability**	***z***
0.20	80% interval around the hypothesized mean	1.28
0.10	90% interval around the hypothesized mean	1.645
0.05	95% interval around the hypothesized mean	1.96
0.02	98% interval around the hypothesized mean	2.33
0.01	99% interval around the hypothesized mean	2.575

Critical Value for Z-Distribution …

- If the level of significance is 0.1

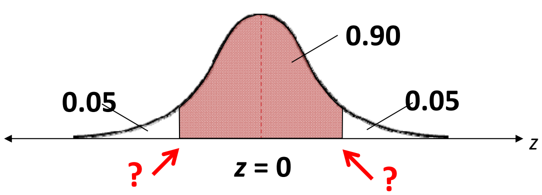
Level of Test	z
0.20	1.28
0.10	**1.645**
0.05	1.96
0.01	2.575



Critical Value for Z-Distribution …

- If the level of confidence is 0.1,

Level of Test	z
0.20	1.28
0.10	**1.645**
0.05	1.96
0.01	2.575



1. Look up the z-table and find the value of 0.95 (= 1 – 0.05)
Why? Since the Table C of appendix shows the area left of critical value
2. It is not shown in the table. What to do next?
P(z<1.64) = 0.9495, and P(z<1.65) = 0.9505
0.95 is between these two z-value (exactly middle of these two value)
3. Make average of 1.64 and 1.65 → 1.645

Critical Values of the z-Test Statistic …

- One sided (= One tailed)

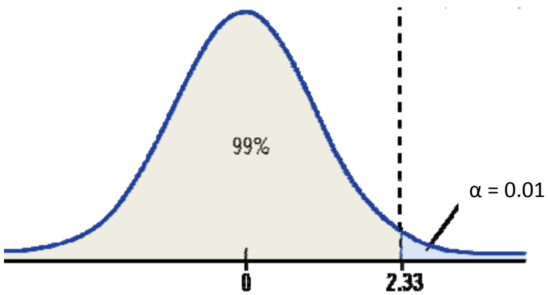
Critical Values of the *z*-Test Statistic for One-Sided Alternatives		
Level of Test	**Definition of Ordinary Variability**	***z***
0.20	lower/upper 80% of the distribution	±0.84
0.10	lower/upper 90% of the distribution	±1.28
0.05	lower/upper 95% of the distribution	±1.645
0.02	lower/upper 98% of the distribution	±2.05
0.01	lower/upper 99% of the distribution	±2.33

For a test with a one-sided "less than" alternative hypothesis (H_a), the critical value will be negative. For a test with a one-sided "greater than" alternative hypothesis (H_a), the critical value will be positive.

Critical Value for Z-Distribution …

- One sided (= One tailed)

Critical Values of the *z*-Test Statistic for One-Sided Alternatives		
Level of Test	**Definition of Ordinary Variability**	***z***
0.01	**lower/upper 99% of the distribution**	**±2.33**



1. Look up the z-table and find the value of 0.99 (= 1 – 0.01)
2. It is not shown in the table. P(Z = 2.33) = 0.9901

3.1 We can approximate z value with 2.33 (for the right tail)

3.2 We can approximate z value with – 2.33 (for the left tail)

In the case of n ≤ 30

Definition

- Rules for a *t*-test for a population mean
  - If $n \leq 30$, *and*
  - the standard deviation of the population, σ, is *unknown* and the sample is drawn from a normal population,
  - then the test statistic is given by

$$t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}}, \text{ where } s_{\bar{x}} = \frac{s}{\sqrt{n}},$$

$\bar{x}$ is the sample mean,

μ_0 is the hypothesized value of the population mean,

s is the sample standard deviation,

and n is the sample size.

The test statistic has a *t*-distribution with n - 1 degrees of freedom.

P value

Definition

- P-value
  - The probability of observing a value of the test statistic as extreme or more extreme than the observed one, assuming the null hypothesis is true.
    → The p-value is the probability that your null hypothesis is actually correct
- Decision making with the P-value
  - If the computed *P*-value is less than α, reject the null hypothesis in favor of the alternative.
  - If the computed *P*-value is greater than or equal to α, fail to reject the null hypothesis.

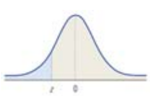
P-value …

- P-values allow us to express the strength of our conclusion to reject or fail to reject the null hypothesis. **A smaller P-value** indicates a rarer test statistic, which implies that the **null hypothesis is less likely to be true**.
- The further a test statistic penetrates a rejection region, the smaller the P-value, and the more confidence a researcher can place in the decision to reject the null hypothesis.
- On the other hand, if a P-value is larger than (or equal to) α, the test statistic is not sufficiently rare to reject the null hypothesis. It could have been caused by ordinary sampling variation.

P-value …

- Less than tests (one-sided, left-tailed) – If you are testing whether an observed value is significantly less than the hypothesized mean, the P-value is the probability of observing a test statistic less than or equal to the calculated value.
  - Example: If the test statistic for a left-tailed test is z = -2.45, then, what is the P-value?

  Solution

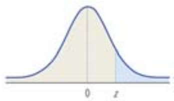


  Since it is left-tailed test, the p-value for this situation is the probability that z is less than –2.45 = P(z < -2.45) = 0.0071.

P-value …

- Greater than tests (one-sided, right-tailed) – If you are testing whether an observed value is significantly greater than the hypothesized mean, the P-value is the probability of observing a test statistic greater than or equal to the calculated value.
  - Example: If the test statistic for a right-tailed test is z = 2.12, then, what is the P-value?

  Solution

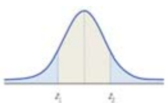


  Since it is right-tailed test, the p-value for this situation is the probability that z is greater than 2.12 = P(z > 2.12) = 0.0170

P-value …

- Not equal tests (two-sided) – If you are testing whether an observed value is significantly different from the hypothesized mean, the P-value is the probability of observing a test statistic whose absolute value is greater than or equal to the absolute value of the calculated value. Thus, to compute the P-value for a two-sided test, simply double the tail probability of the test statistic.
  - Example: If the test statistic for a two-tailed test is z = -2.56, then

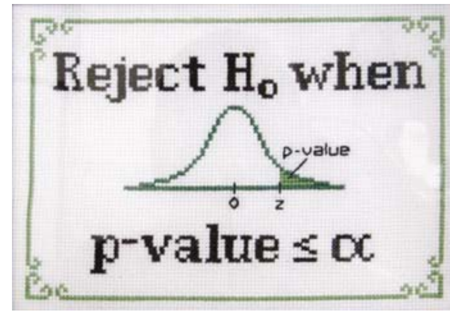
$$\begin{aligned} P\text{-value} &= P\left(|z| \geq |-2.56|\right) \\ &= 2 \cdot P\left(z \leq -2.56\right) \\ &= 2 \cdot 0.0052 \\ &= 0.0104. \end{aligned}$$



The *P*-value of 0.0104 is the probability of observing a value of the test statistic that is either greater than or equal to 2.56 or less than or equal to -2.56 assuming the null hypothesis is true.

P-value …

- Conclusions for a Hypothesis Testing Using p-Values
  1. If $p \leq \alpha$, then *reject* the null hypothesis.
  2. If $p > \alpha$, then *fail to reject* the null hypothesis.



P-value …

- Example 1
  - Suppose a hypothesis test about a population mean was conducted and the P-value was calculated to be 0.0156.

a) If the level of significance is 0.01, what is the conclusion of the test?

b) If the level of confidence is 90%, what is the conclusion of the test?

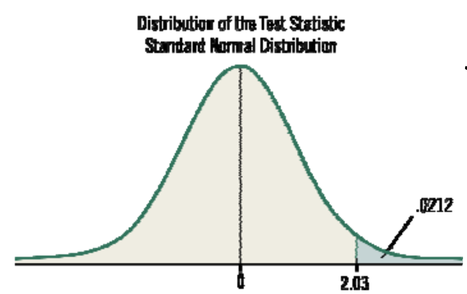
Solutions:

a. Since *P*-value = 0.0156 $\geq$ 0.01 = α, fail to reject the null hypothesis.

b. Since α = 1 - level of confidence = 1 - 0.9 = 0.1 and *P*-value = 0.0156 < 0.1 = α, reject the null hypothesis.

P-value …

- Example 2
  - The following graph displays a z-test statistic of 2.03 and the corresponding P-value of 0.0212. At a significance level of 0.05, would you reject or fail to reject the null hypothesis?

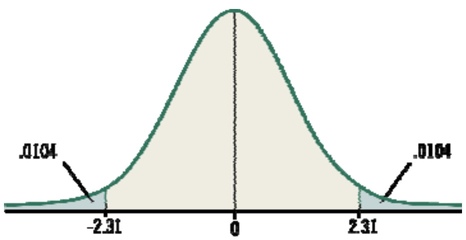


Solution:

Since *P*-value = 0.0212 < 0.05 = α (the significance level), we would reject the null hypothesis.

P-value

- Example 3
  - The following graph illustrates the P-value for a two-tailed test where the computed test statistic was z = 2.31. At a significance level of 0.05, would you reject or fail to reject the null hypothesis?



Solution:

To calculate the *P*-value for a two-sided test, double the tail probability of the test statistic.

Thus, *P*-value = 0.0104 + 0.0104
= 0.0208.

Since *P*-value = 0.0208 < 0.05 = α , we reject the null hypothesis.

Hypothesis Testing …

- Example 1
  - Suppose the national average reading speed for high school sophomores is 150 words per minute with a standard deviation of 15.
  - A local school board member wants to know if sophomore students at Lincoln High School read at a level different from the national average for tenth graders. The level of the test is to be set at 0.05.
  - A random sample of 100 tenth graders has been drawn, and the sample mean is 155 words per minute.

Hypothesis Testing …

- Solution 1

1) H_0: Lincoln High School tenth graders are reading at the national average.
   H_a: Lincoln High School tenth graders are *not* reading at the national average.
2) μ = average number of words read per minute by Lincoln High School sophomores.
3) The key word here is "different." Since the board member is interested in whether there is any difference, the test is two-sided.
4) $H_0: \ \mu = 150 \quad H_a: \mu \neq 150$
5) $\alpha = 0.05$
6) Since $n > 30$, we can assume that the sampling distribution of $\overline{x}$ is approximately normally distributed and therefore we can use the *z*-test.

Hypothesis Testing …

- Solution 1 – Contd

7) $z = \dfrac{\overline{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \dfrac{155 - 150}{\frac{15}{\sqrt{100}}} = \dfrac{5}{\frac{15}{10}} \approx 3.33$

   Note that *s* is used as an approximation of σ.

8) $P\text{-value} = P(|z| \geq 3.33) = 2 \cdot P(z \geq 3.33) = 2(0.0004) = 0.0008$

9) Since *P*-value = 0.0008 < 0.05 = α, reject the null hypothesis.

10) There is significant evidence at the 0.05 level that tenth graders at Lincoln High School do not read at the national average.

Hypothesis Testing …

- Example 2
  - A reporter is investigating a local resident's claim that gas pumps typically dispense less gas than the amount purchased.
  - To test the man's claim, the reporter randomly selects 46 gas pumps around town and measures the amount of gas actually dispensed into a can when exactly one gallon is purchased. The sample mean amount of gas dispensed is 0.97 gallons with a standard deviation of 0.18 gallons.
  - Does the data support the man's claim at the 0.05 significance level?

Hypothesis Testing …

- Solution 2

1) H_0: The average amount of gas dispensed from each pump is at least one gallon.
   H_a: The average amount of gas dispensed from each pump is less than one gallon.
2) μ = the average amount of gas dispensed when one gallon is purchased.
3) The key words here are "less than." Since the reporter is interested in whether the pump is dispensing less than one gallon, the test is one-sided.
4) $H_0 : \mu \geq 1 \qquad H_a : \mu < 1$
5) $\alpha = 0.05$

Hypothesis Testing

- Solution 2

6) Since $n > 30$, we can assume that the sampling distribution of $\bar{x}$ is approximately normally distributed and therefore we can use the z-test.
7) $z = \dfrac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \dfrac{0.97 - 1}{\frac{0.18}{\sqrt{46}}} \approx \dfrac{-0.03}{\frac{0.18}{6.782330}} \approx -1.13$

   Note that s is used as an approximation of s.
8) $P\text{-value} = P(z < -1.13) = 0.1292$
9) Since P-value = 0.1292 $\geq$ 0.05 = α, fail to reject the null hypothesis.
10) There is not significant evidence at the 0.05 level that the average amount of gas dispensed from each pump is less than one gallon.

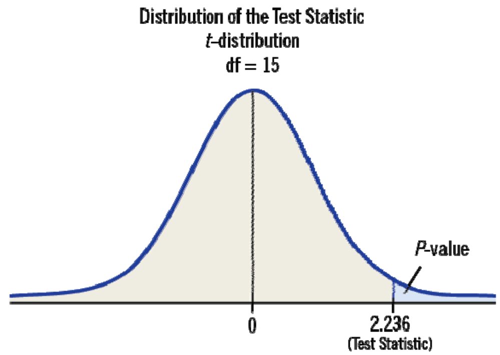
P-values for t-test statistics …

- Since there is a different t distribution for each number of degrees of freedom (*df*), the *t* table is constructed differently from the z tables.
- The t table only provides *t*-values for frequently used tail probabilities, so in most cases the exact value of the *t*-test statistic will NOT be in the table.
- When this situation arises, we can find the closest *t*-values in the table with the appropriate degrees of freedom which surround the test statistic and determine an interval estimation of the *P*-value.
- Exact *P*-values for t-test statistics can be found using a graphing calculator or statistical software.

P-values for t-test statistics …

- Example

The graph displays a t-test statistic of 2.236 for a right-tailed test with 15 *df*.
Determine the P-value.



P-values for t-test statistics

- Solution

Degree of freedom? 15

The value of the test statistic, $t = 2.236$,
→ falls between $t_{0.025,\,15} = 2.131$ and $t_{0.010,\,15} = 2.602$,
so $P(t > 2.236)$ is between 0.010 and 0.025.
In interval notation, the P-value is in the interval (0.010, 0.025).

Using a calculator or statistical software, the exact P-value can be found to be 0.0205.

Degrees of Freedom	Area in One Tail				
	$t_{0.100}$	$t_{0.050}$	$t_{0.025}$	$t_{0.010}$	$t_{0.005}$
1	3.078	6.314	12.706	31.821	63.657
⋮	⋮	⋮	⋮	⋮	⋮
15	1.341	1.753	2.131	2.602	2.947
⋮	⋮	⋮	⋮	⋮	⋮

Hypothesis Testing …

- Example 1
  - The Alexander Bolt Company produces half-inch A-class stainless steel bolts. The specified standard is that they have a mean tensile strength of more than 4000 pounds. The company is very concerned about quality and wants to be sure that its product does not fall below the standard.
  - A sample of 25 bolts is randomly selected and the mean tensile strength is found to be 4014 pounds, with a standard deviation of 20 pounds.
  - Conduct a hypothesis test to determine if there is overwhelming evidence at the 0.01 level of significance that the bolts meet the specified quality standards.

Hypothesis Testing …

- Solution 1

1) H_0: The steel bolts have a mean tensile strength of *no more* than 4000 pounds.
   H_a: The steel bolts have a mean tensile strength of *more* than 4000 pounds.
2) μ = the mean tensile strength of the steel bolts.
3) The key word here is “more.” Since the company is interested if the bolts have a tensile strength of more than a specified value, the test is one-sided.
4) $H_0: \mu \leq 4000$ $\qquad$ $H_a: \mu > 4000$
5) $\alpha = 0.01$
6) Since $n \leq 30$, we will use the t-test. We assume that the tensile strength levels of the bolts are normally distributed.

Hypothesis Testing …

- Solution 1 – contd

7) $t = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \dfrac{4014 - 4000}{\frac{20}{\sqrt{25}}} = \dfrac{14}{\frac{20}{5}} = 3.5$

8) Using the row of the t table for df = 24, we can determine that the test statistic, $t = 3.5$, is greater than $t_{0.005,\,24} = 2.797$, so $P(t \geq 3.5)$ is less than 0.005. In interval notation, this means the P-value is in the interval (0, 0.005). Using a calculator or statistical software, the exact P-value can be found to be 0.0009.

9) Since P-value = 0.0009 < 0.01 = α, reject the null hypothesis.

10) There is significant evidence at the 0.01 level that the mean tensile strength of the steel bolts is greater than 4000 pounds, so the bolts meet the specified quality standards.

Hypothesis Testing …

- Example 2
  - A sheriff claims that the mean age of prison inmates in his county is not the same as the mean age of prison inmates in the state. The mean age of current inmates in the state is 33.7 years. The sheriff randomly selects the records of 25 inmates in his county and finds that they have a mean age of 31.3 years with a standard deviation of 8.8 years. Does the data support the sheriff's claim at a 0.10 significance level? Assume that the age of prison inmates in the county has an approximately normal distribution.

Solution:

1) H_0: The inmates in the county have a mean age of 33.7 years.
   H_a: The inmates in the county *do not* have a mean age of 33.7 years.

Hypothesis Testing …

- Solution 2 – Contd

2) μ = the mean age of all inmates in the county.

3) The key words here are "not the same." Since the sheriff is testing whether the mean age of county inmates is different from the state average, the test is two-sided.

4) $H_0 : \mu = 33.7$ $\qquad$ $H_a : \mu \neq 33.7$

5) $\alpha = 0.10$

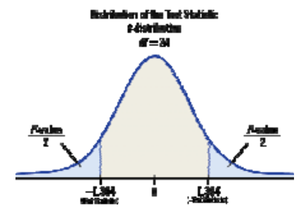
6) Since $n \leq 30$, we will use the t-test. As stated in the problem, we assume that the ages of inmates in the county are normally distributed.

Hypothesis Testing …

- Solution 2 – Contd

7) $t = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \dfrac{31.3 - 33.7}{\frac{8.8}{\sqrt{25}}} = \dfrac{-2.4}{\frac{8.8}{5}} \approx -1.364$

8) The test statistic has a t-distribution with n - 1 = 24 degrees of freedom. Since the test is two-tailed, the P-value is the probability of getting a t-test statistic that is either less than or equal to -1.364 or greater than or equal to 1.364 as shown in the graph.



Hypothesis Testing …

- Solution 2 – Contd

8) Using the row of the t table for df = 24, the absolute value of the test statistic, $|t| = 1.364$, falls between $t_{0.100,\,24} = 1.318$ and $t_{0.050,\,24} = 1.711$, in other words,

$$t_{\,0.100\,,\,24} = 1.318 < t_{\,\text{p-value/2}\,,\,24} = 1.364 < t_{\,0.050\,,\,24} = 1.711$$

$$0.05 \quad < \quad \frac{P-VALUE}{2} \quad < \quad 0.1$$

→ 0.05 * 2 < p-value < 0.1 * 2 → 0.1 < p-value < 0.2

(so the P-value, $P(|t| \geq 1.364)$, is between 2(0.050) = 0.100 and 2(0.100) = 0.200.)

Note that we multiply the area in one tail by 2 since we are conducting a two-tailed test.

In interval notation, the P-value is in the interval (0.100, 0.200). Using a calculator or statistical software, the exact P-value can be found to be 0.1852.

Hypothesis Testing …

- Solution 2 – Contd

9) The P-value is in the interval (0.100, 0.200), which means P-value > 0.100 and P-value < 0.200. Thus, P-value $\geq \alpha = 0.10$, so we fail to reject the null hypothesis.

10) At the 0.10 level of significance, the data does not support the sheriff's claim that the mean age of inmates in his county differs from the mean age of inmates in the state.

Assignment

- Review & Understand the Lecture Note
- This is critical since Regression (next week's content) needs a good understanding of all the concepts we learned today.