# Statistics Bootcamp

# Statistics Bootcamp

## An overview of statistical concepts and basics

- Assumption of some level of prior experience
- A lot of ground to be covered!
- Practical examples / exercises  will focus on general approach rather than calculation

**Topics:**

- Descriptive / Inferential Statistics
- Data Visualisation,
- Probability
- Hypothesis Testing
- Probability Distributions
- Confidence Intervals
- One and Two Tailed Tests, Single Samples
- Comparing Samples, Multiple samples, ANOVA
- Correlation

Part 1.

# Statistics Bootcamp

## An overview of statistical concepts and basics

- Assumption of some level of prior experience
- A lot of ground to be covered!
- Practical examples / exercises will focus on general approach rather than calculation

**Topics:**

- Descriptive / Inferential Statistics
- Data Visualisation,
- Probability
- Hypothesis Testing
- Probability Distributions
- Confidence Intervals
- One and Two Tailed Tests, Single Samples
- Comparing Samples, Multiple samples, ANOVA
- Correlation

Part 1.

Part 2. Starts at slide 66

# The Science of Statistics

**Statistics** is the science of data. This involves collecting, classifying, summarizing, organizing, analyzing, presenting, and interpreting numerical information.

# Types of Statistical Applications

**Descriptive statistics** utilizes numerical and graphical methods to look for patterns in a data set, to summarize the information revealed in a data set, and to present that information in a convenient form.

**Inferential statistics** utilizes sample data to make estimates, decisions, predictions, or other generalizations about a larger set of data.

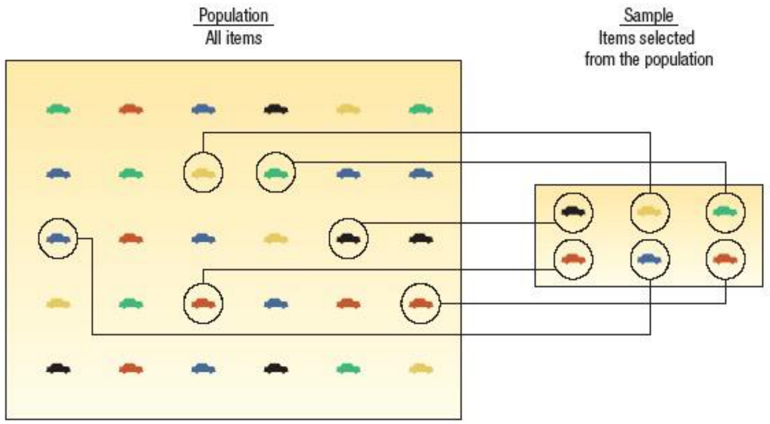
# Fundamental Elements of Statistics

An **experimental** (or **observational**) **unit** is an object (e.g., person, thing, transaction, or event) about which we collect data.

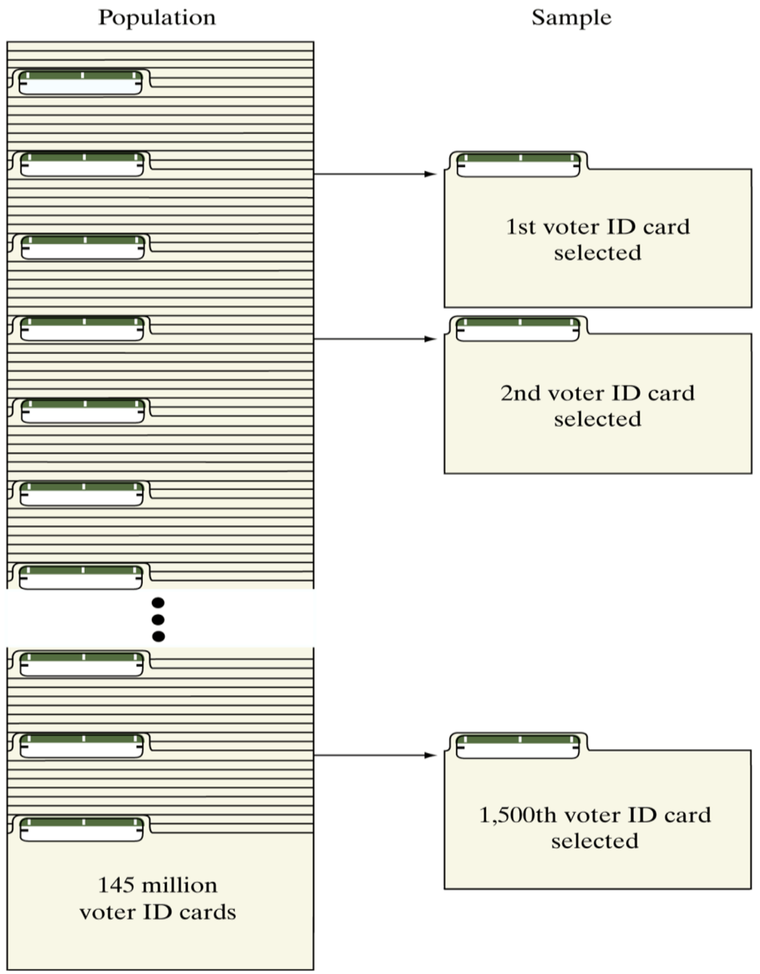
A **population** is a set of units (usually people, objects, transactions, or events) that we are interested in studying.

A **variable** is a characteristic or property of an individual experimental (or observational) unit in the population.

A **sample** is a subset of the units of a population.



# A sample of voter registration cards for all registered voters




A **statistical inference** is an estimate, prediction, or some other generalization about a population based on information contained in a sample.

# Reliability

A **measure of reliability** is a statement (usually quantitative) about the degree of uncertainty associated with a statistical inference.

**Four Elements of Descriptive Statistical Problems**

1. The population or sample of interest
2. One or more variables (characteristics of the population or sample units) that are to be investigated
3. Tables, graphs, or numerical summary tools
4. Identification of patterns in the data

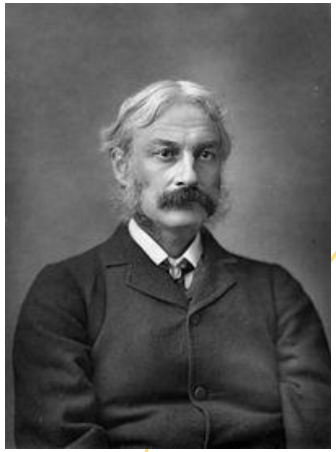
**Five Elements of Inferential Statistical Problems**

1. The population of interest
2. One or more variables (characteristics of the population units) that are to be investigated
3. The sample of population units
4. The inference about the population based on information contained in the sample
5. A measure of the reliability of the inference

# Fundamental Importance of Source

*There are three kinds of lies: lies, damned lies, and statistics.*

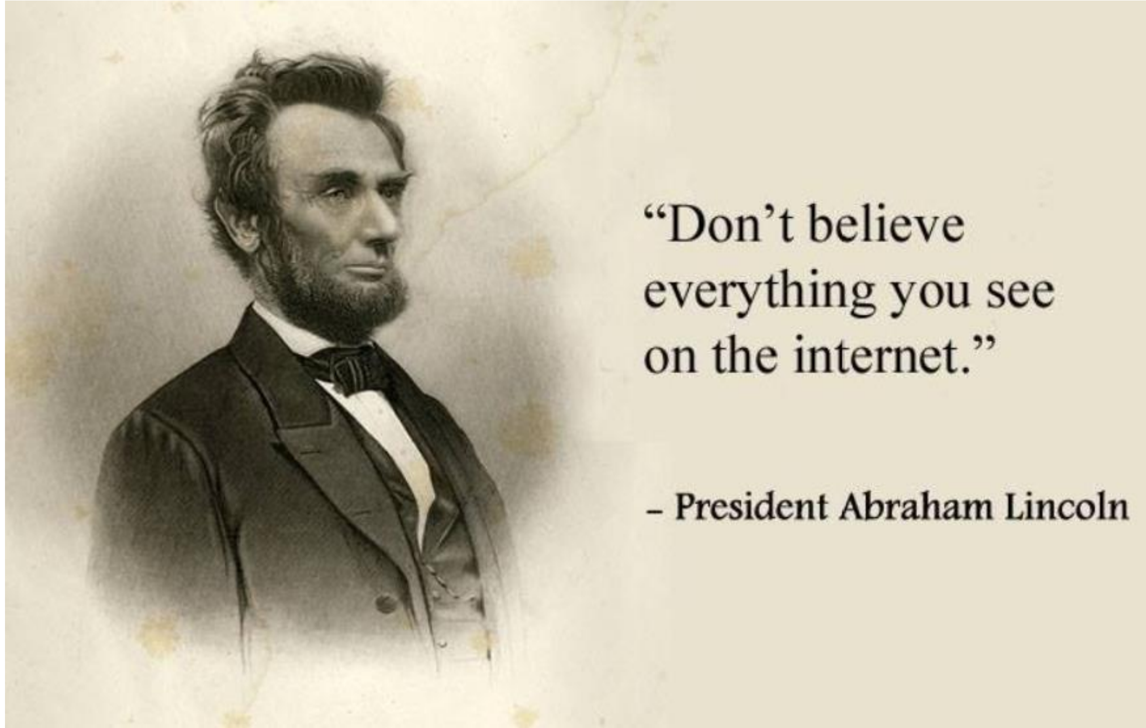
Benjamin Disraeli



*An unsophisticated forecaster uses statistics as a drunken man uses lamp-posts - for support rather than for illumination*

Andrew Lang

(Image source: Wikipedia)




*"I only believe in statistics that I doctored myself"*

Winston Churchill

(Image source: Wikipedia)

# "Cola Wars"


- Suppose...
  - Sample 1,000 tasters
  - 56% Preferred Pepsi
- How "*accurate*" is this?
- How "*confident*" would you be that this reflects entire population's preference?

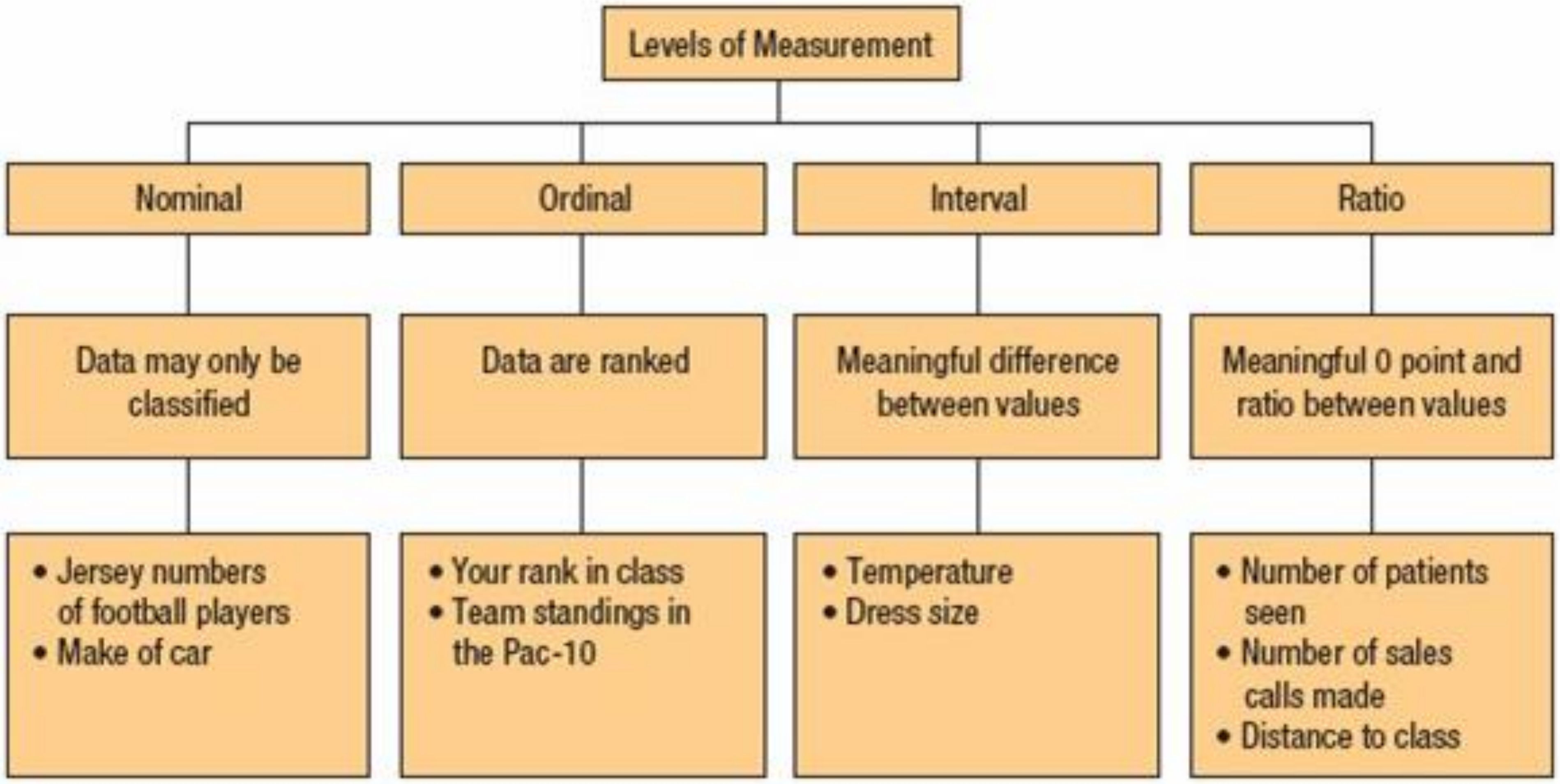
# Types of Data

**Quantitative data** are measurements that are recorded on a naturally occurring numerical scale.

- Temperature
- Unemployment rate
- Test scores
- Number of car accidents per year

**Qualitative** (or **categorical**) **data** are measurements that cannot be measured on a natural numerical scale; they can only be classified into one of a group of categories.

- Political party affiliation (FF/FG/Lab/SF...)
- Defect status (defective or not)
- Model of car – Ford: Ka, Focus, Mondeo, Kuga....
- Ratings – Very good, Good, Neutral Poor, Very Poor
- Gender Male, Female…




Summary of the Characteristics for Levels of Measurement

# Collecting Data

A **designed experiment** is a data collection method where the researcher exerts full control over the characteristics of the experimental units sampled. These experiments typically involve a group of experimental units that are assigned the *treatment* and an untreated (or *control*) group.

An **observational study** is a data collection method where the experimental units sampled are observed in their natural setting. No attempt is made to control the characteristics of the experimental units sampled. (Examples include *opinion polls* and *surveys*.)

# Samples

A **representative sample** exhibits characteristics typical of those possessed by the target population.

- Literary Digest 1936 U.S. Presidential Poll
  - 25% response rate (2.4m responses)

- Forecast Landon 57% : Roosevelt 43%
- Actual Landon 37% : Roosevelt 62%

- Based on phone numbers, drivers' registrations, and country club memberships

*Source: http://www.oxfordmathcenter.com/drupal7/node/251*

# Bias

**Selection bias** results when a subset of the experimental units in the population is excluded so that these units have no chance of being selected in the sample.

**Nonresponse bias** results when the researchers conducting a survey or study are unable to obtain data on all experimental units selected for the sample.

**Measurement error** refers to inaccuracies in the values of the data recorded. In surveys, this kind of error may be due to ambiguous or leading questions and the interviewer's effect on the respondent.

# The Role of Statistics in Critical Thinking and Ethics

- *"Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write"*

**Statistical thinking** involves applying rational thought and the science of statistics to critically assess data and inferences. Fundamental to the thought process is that variation exists in populations of data.

spurious correlations

# Avocado Prices and Bitcoin Values


https://twitter.com/tracyalloway/status/1187167482877865984?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1187167482877865984%7Ctwgr%5Eshare_3&ref_url=https%3A%2F%2Fboingboing.net%2F2019%2F11%2F04%2Fthe-mysterious-correlation-bet.html

# Leaving Cert. exam upgrades double

The number of upgrades awarded, following appeals, to 2019 Leaving Certificate candidates more than doubled in 2019 compared to 2018.

Almost 3,000 (2,916) individual grades awarded in the 2019 exams were revised upwards compared to 1,450 in 2018.

38% of the Japanese exam students who appealed their results received an upgrade.

16% of papers appealed in 2018 resulted in an upgrade,  17% - were revised upwards in 2019.

Five students (of a total of thirteen) Japanese exam students who appealed their results received an upgrade in 2019

https://www.irishtimes.com/news/education/almost-3-000-leaving-cert-results-upgraded-in-record-year-1.4021574

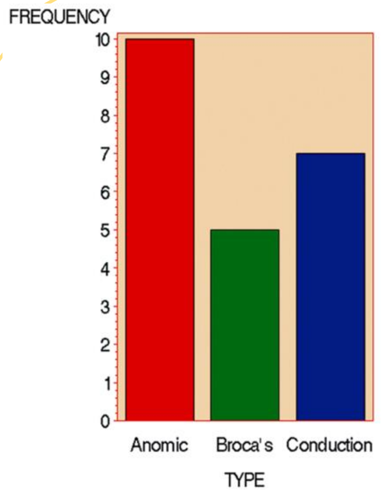
https://www.rte.ie/news/education/2019/0918/1076597-exam-appeals/

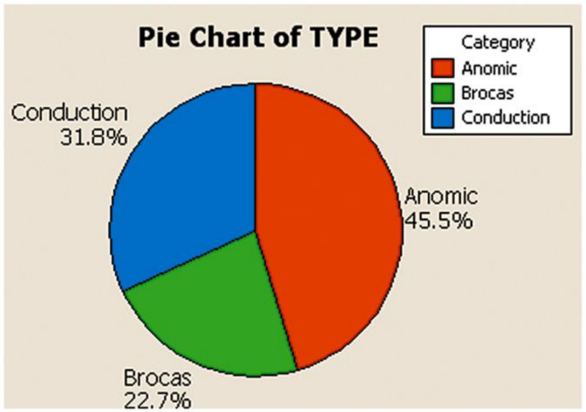
University of Reading meteorologist Dr Rob Thompson's experiment

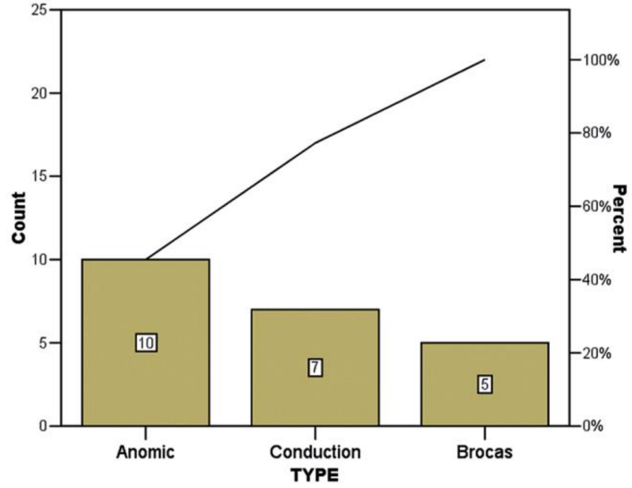

**Experiment Shows Risk of Flash Flooding After Drought**

# Methods for Describing Sets of Data

# Representing Data - Graphs




**Summary of Graphical Descriptive Methods for Qualitative Data**

**Bar Graph:** The categories (classes) of the qualitative variable are represented by bars, where the height of each bar is either the class frequency, class relative frequency, or class percentage.

**Pie Chart:** The categories (classes) of the qualitative variable are represented by slices of a pie (circle). The size of each slice is proportional to the class relative frequency.

**Pareto Diagram:** A bar graph with the categories (classes) of the qualitative variable (i.e., the bars) arranged by height in descending order from left to right.

# Graphical Methods for Describing Quantitative Data

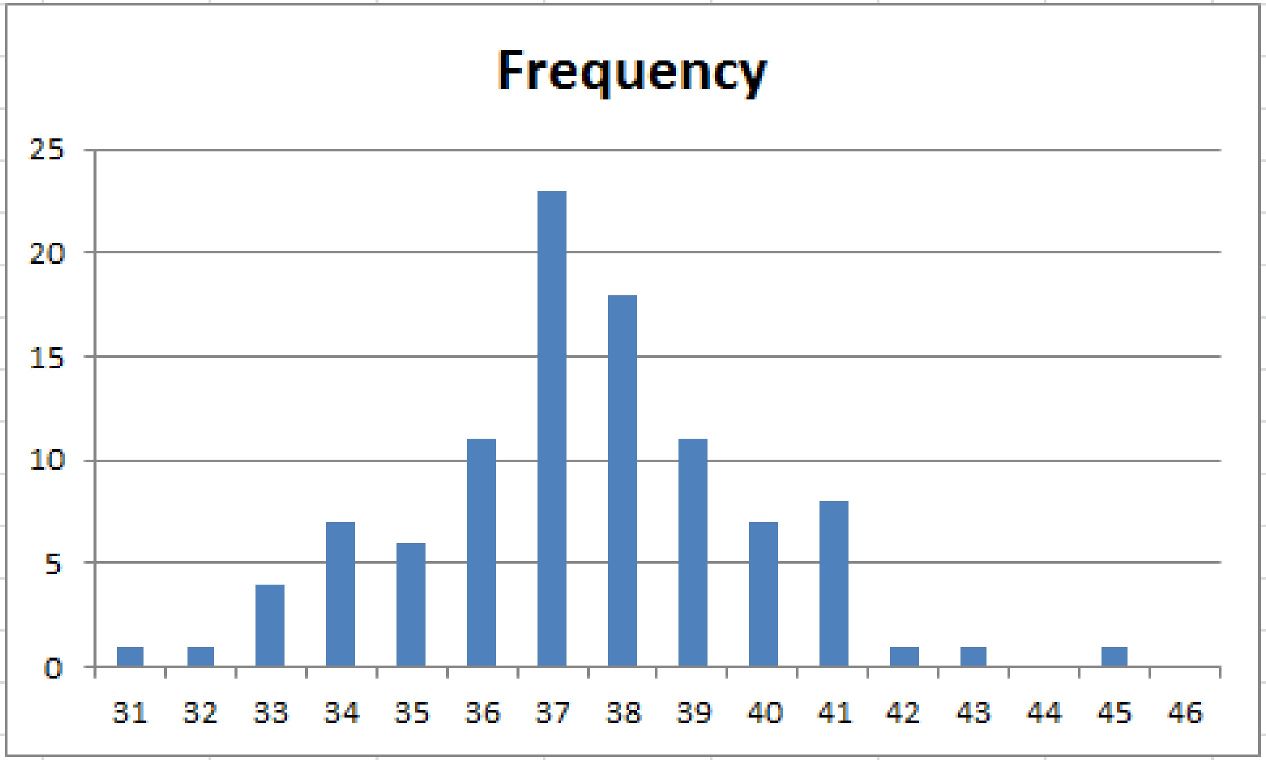
**Table 2.2** EPA Mileage Ratings on 100 Cars

| | | | | |
|---|---|---|---|---|
| 36.3 | 41.0 | 36.9 | 37.1 | 44.9 |
| 32.7 | 37.3 | 41.2 | 36.6 | 32.9 |
| 40.5 | 36.5 | 37.6 | 33.9 | 40.2 |
| 36.2 | 37.9 | 36.0 | 37.9 | 35.9 |
| 38.5 | 39.0 | 35.5 | 34.8 | 38.6 |
| 36.3 | 36.8 | 32.5 | 36.4 | 40.5 |
| 41.0 | 31.8 | 37.3 | 33.1 | 37.0 |
| 37.0 | 37.2 | 40.7 | 37.4 | 37.1 |
| 37.1 | 40.3 | 36.7 | 37.0 | 33.9 |
| 39.9 | 36.9 | 32.9 | 33.8 | 39.8 |
| 36.8 | 30.0 | 37.2 | 42.1 | 36.7 |
| 36.5 | 33.2 | 37.4 | 37.5 | 33.6 |
| 36.4 | 37.7 | 37.7 | 40.0 | 34.2 |
| 38.2 | 38.3 | 35.7 | 35.6 | 35.1 |
| 39.4 | 35.3 | 34.4 | 38.8 | 39.7 |
| 36.6 | 36.1 | 38.2 | 38.4 | 39.3 |
| 37.6 | 37.0 | 38.7 | 39.0 | 35.8 |
| 37.8 | 35.9 | 35.6 | 36.7 | 34.5 |
| 40.1 | 38.0 | 35.2 | 34.8 | 39.5 |
| 34.0 | 36.8 | 35.0 | 38.1 | 36.9 |

Draw A Normal Frequency Distribution Histogram for these data, usings intervals starting at 30-31, 31-32, etc...

# Frequency Distribution

| Bins | Frequency |
|---|---|
| 31 | 1 |
| 32 | 1 |
| 33 | 4 |
| 34 | 7 |
| 35 | 6 |
| 36 | 11 |
| 37 | 23 |
| 38 | 18 |
| 39 | 11 |
| 40 | 7 |
| 41 | 8 |
| 42 | 1 |
| 43 | 1 |
| 44 | 0 |
| 45 | 1 |
| 46 | 0 |



# Numerical Measures of Central Tendency - Definitions

The **mean** of a set of quantitative data is the sum of the measurements, divided by the number of measurements contained in the data set.

The **median** of a quantitative data set is the middle number when the measurements are arranged in ascending (or descending) order.

The **mode** is the measurement that occurs most frequently in the data set.

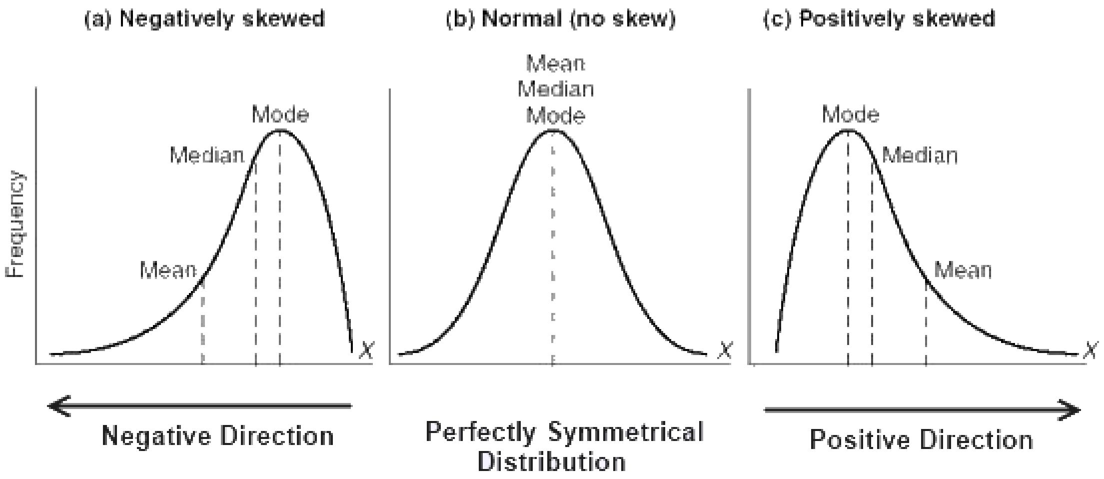



Image source: http://www.durofy.com/tutorial-8-mode-merits-demerits

# Numerical Measures of Variability

Other Measures:

- Range
- Variance
- Standard Deviation

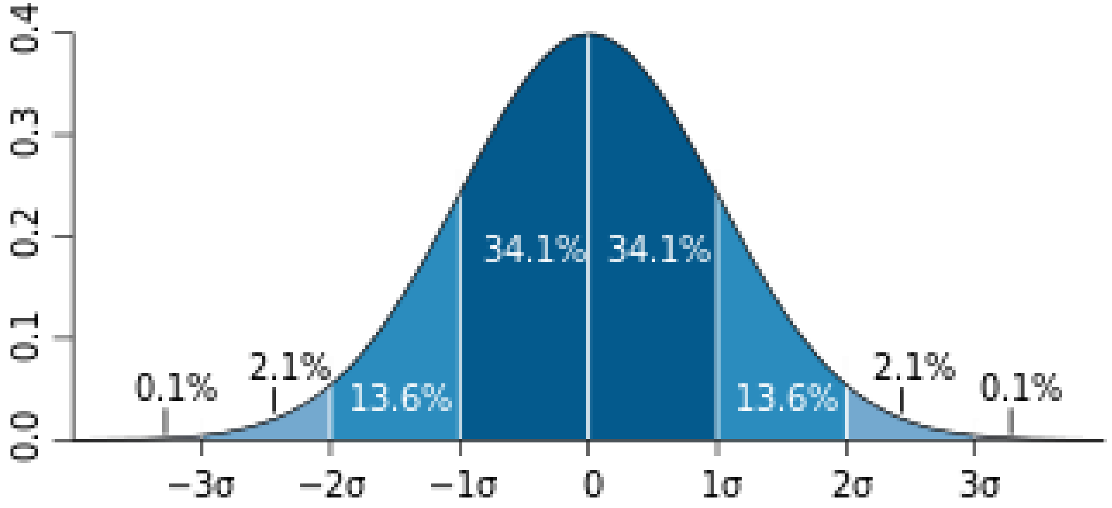
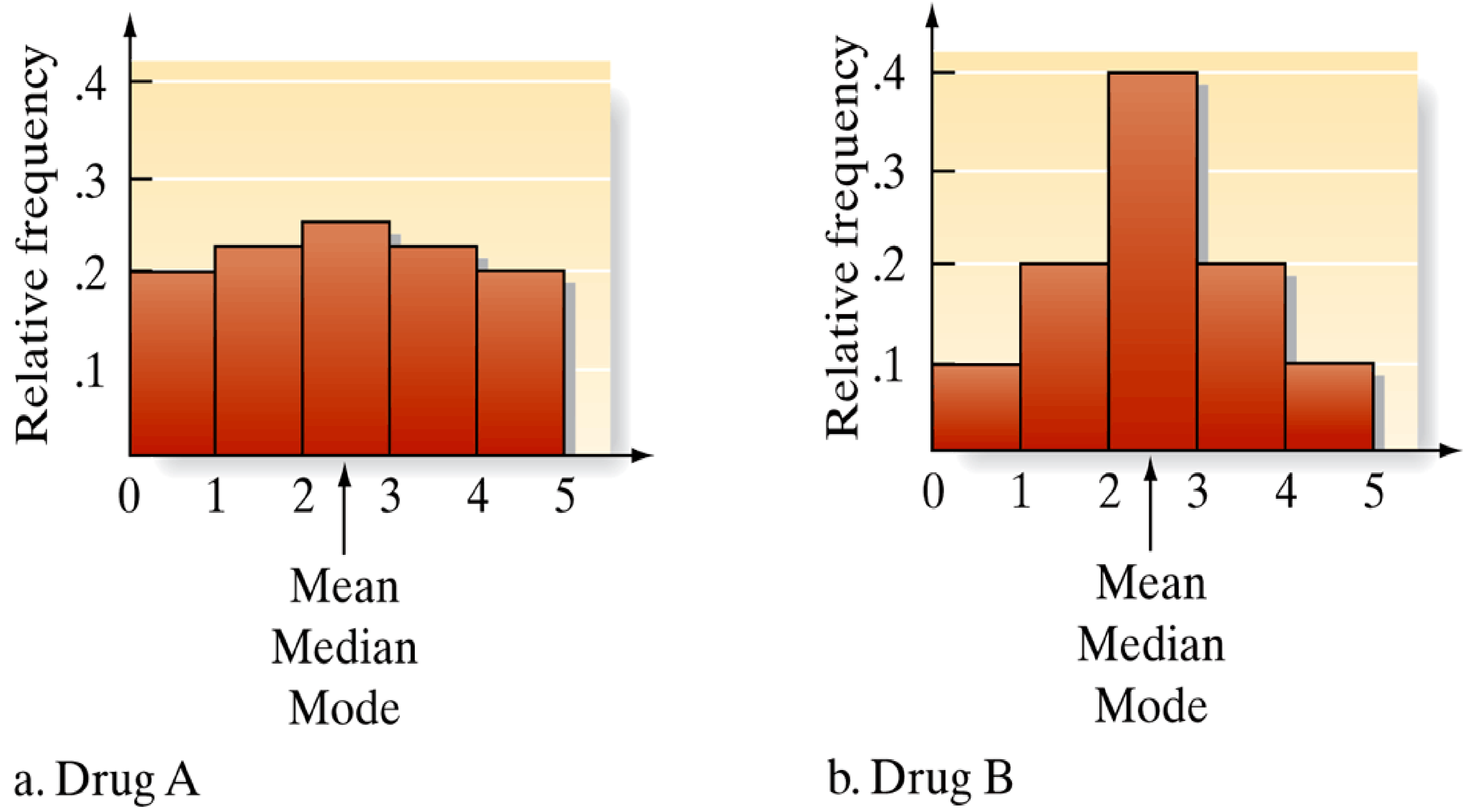



Image source: http://en.wikipedia.org/wiki/Standard_deviation

# Response time histograms for two drugs



# Range

The **range** of a quantitative data set is equal to the largest measurement minus the smallest measurement.

Calculate Range in following sets of data:

```
Data set X: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
Data set Y: 3, 3, 5, 4, 7, 8, 3, 2, 9, 7
Data set Z: 3, 3, 4, 6, 4, 5, 7, 4, 3, 4
```

Which is most variable?

# Variance ($s^2$)

The **sample variance** for a sample of $n$ measurements is equal to the sum of the squared deviations from the mean, divided by $(n - 1)$. The symbol $s^2$ is used to represent the sample variance.

$$s^2 = \frac{\sum(x - \bar{x})^2}{n - 1}$$

# Variance

$$s^2 = \frac{\sum(x - \bar{x})^2}{n-1}$$

***Sample data set A*: 1, 2, 3, 4, 5**

***Sample data set B*: 2, 3, 3, 3, 4**

$$s^2 = \frac{(1-3)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 + (5-3)^2}{5-1} = \frac{10}{4} = 2.5$$

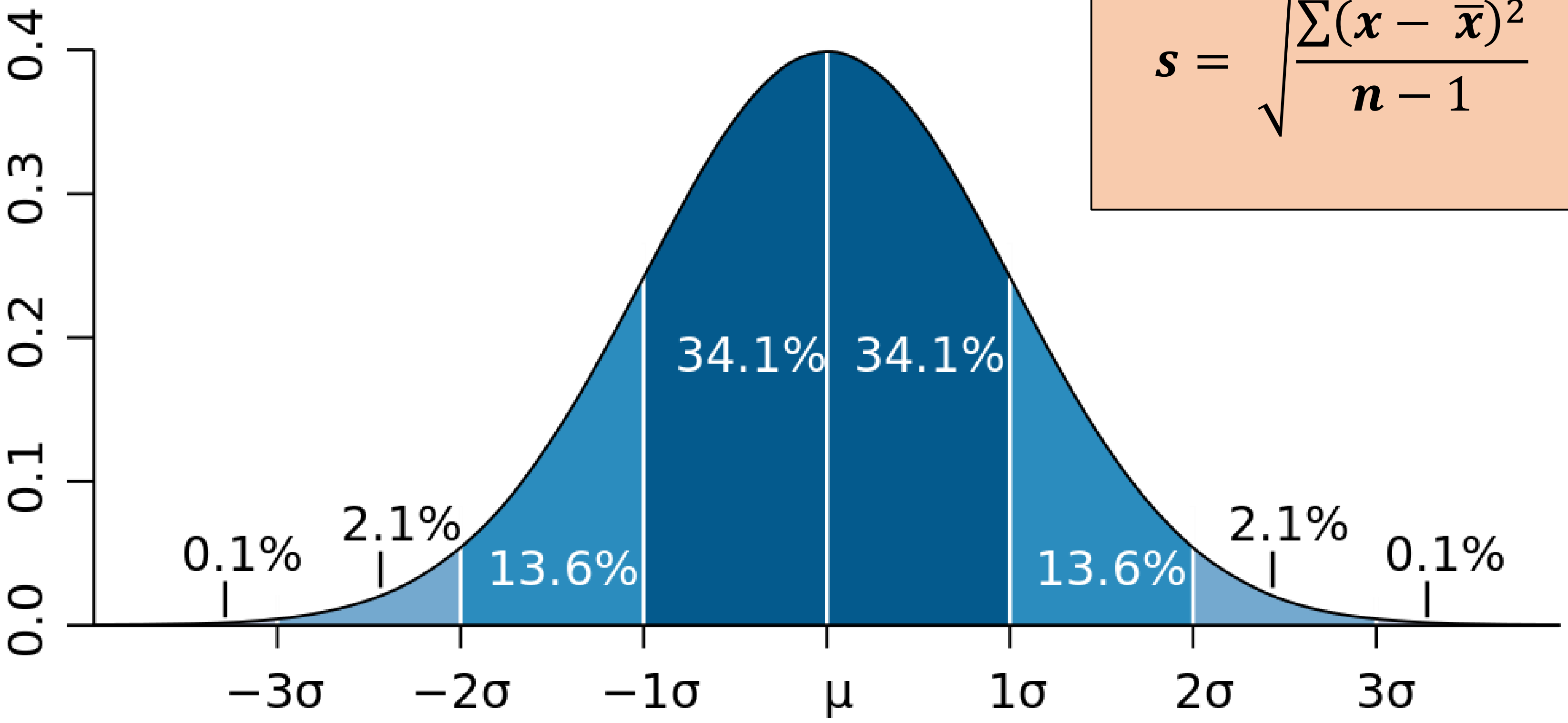
$$s^2 = \frac{(2-3)^2 + (3-3)^2 + (3-3)^2 + (3-3)^2 + (4-3)^2}{5-1} = \frac{2}{4} = 0.5$$

# Standard Deviation

The **sample standard deviation**, $s$, is defined as the positive square root of the sample variance, $s^2$, or, mathematically,

$$s = \sqrt{s^2}$$

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$



# Descriptive Data – Population vs Sample Formulas

| | Population | Sample |
|---|---|---|
| **Mean** | $\mu = \frac{\sum x}{N}$ | $\bar{x} = \frac{\sum x}{n}$ |
| **Variance** | $\sigma^2 = \frac{\sum(x - \mu)^2}{N}$ | $s^2 = \frac{\sum(x - \bar{x})^2}{n-1}$ |
| **Standard Deviation** | $\sigma = \sqrt{\frac{\sum(x - \mu)^2}{N}}$ | $s = \sqrt{\frac{\sum(x - \bar{x})^2}{n-1}}$ |

# Variance and Standard Deviation Symbols

**Symbols for Variance and Standard Deviation**

$s^2$ = Sample variance

$s$ = Sample standard deviation

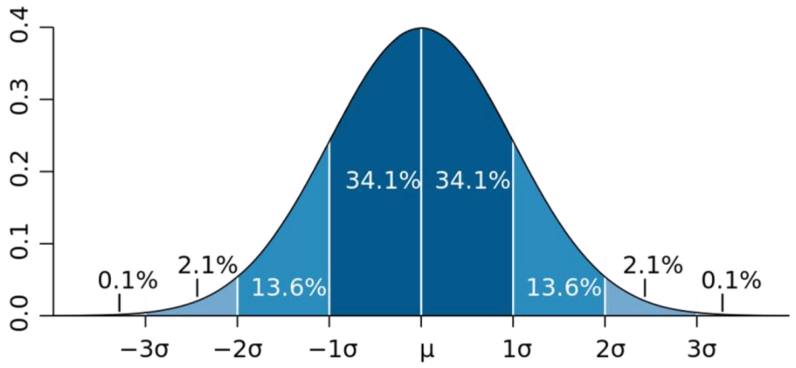
$\sigma^2$ = Population variance

$\sigma$ = Population standard deviation

# Empirical Rule

**Table 2.7 Interpreting the Standard Deviation: The Empirical Rule**

The **empirical rule** is a rule of thumb that applies to data sets with frequency distributions that are mound shaped and symmetric, as follows:




**a.** Approximately 68% of the measurements will fall within one standard deviation of the mean [i.e., within the interval $(\bar{x} - s, \bar{x} + s)$ for samples and $(\mu - \sigma, \mu + \sigma)$ for populations].

**b.** Approximately 95% of the measurements will fall within two standard deviations of the mean [i.e., within the interval $(\bar{x} - 2s, \bar{x} + 2s)$ for samples and $(\mu - 2\sigma, \mu + 2\sigma)$ for populations].

**c.** Approximately 99.7% (essentially all) of the measurements will fall within three standard deviations of the mean [i.e., within the interval $(\bar{x} - 3s, \bar{x} + 3s)$ for samples and $(\mu - 3\sigma, \mu + 3\sigma)$ for populations].

# Interpreting the Standard Deviation

**Table 2.6 Interpreting the Standard Deviation: Chebyshev's Rule**

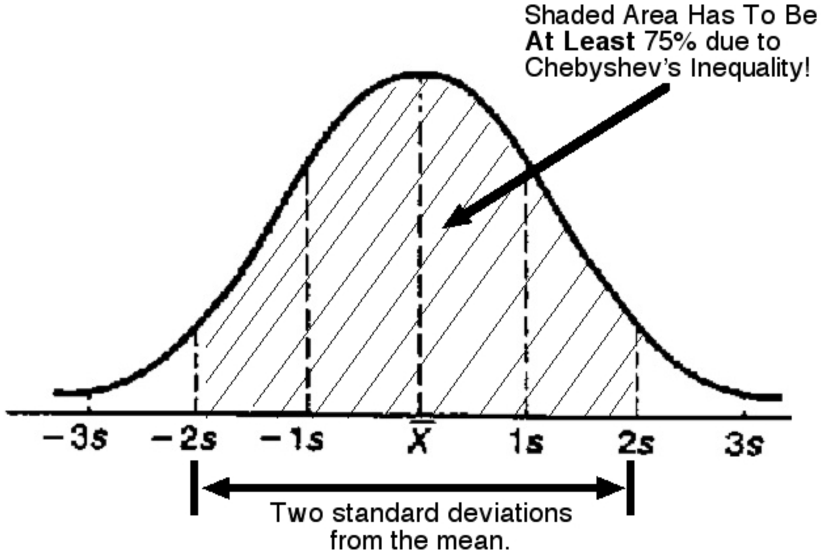
**Chebyshev's rule** applies to any data set, regardless of the shape of the frequency distribution of the data.

**a.** It is possible that very few of the measurements will fall within one standard deviation of the mean [i.e., within the interval $(\bar{x} - s, \bar{x} + s)$ for samples and $(\mu - \sigma, \mu + \sigma)$ for populations].

**b.** At least $\frac{3}{4}$ of the measurements will fall within two standard deviations of the mean [i.e., within the interval $(\bar{x} - 2s, \bar{x} + 2s)$ for samples and $(\mu - 2\sigma, \mu + 2\sigma)$ for populations].

**c.** At least $\frac{8}{9}$ of the measurements will fall within three standard deviations of the mean [i.e., within the interval $(\bar{x} - 3s, \bar{x} + 3s)$ for samples and $(\mu - 3\sigma, \mu + 3\sigma)$ for populations].

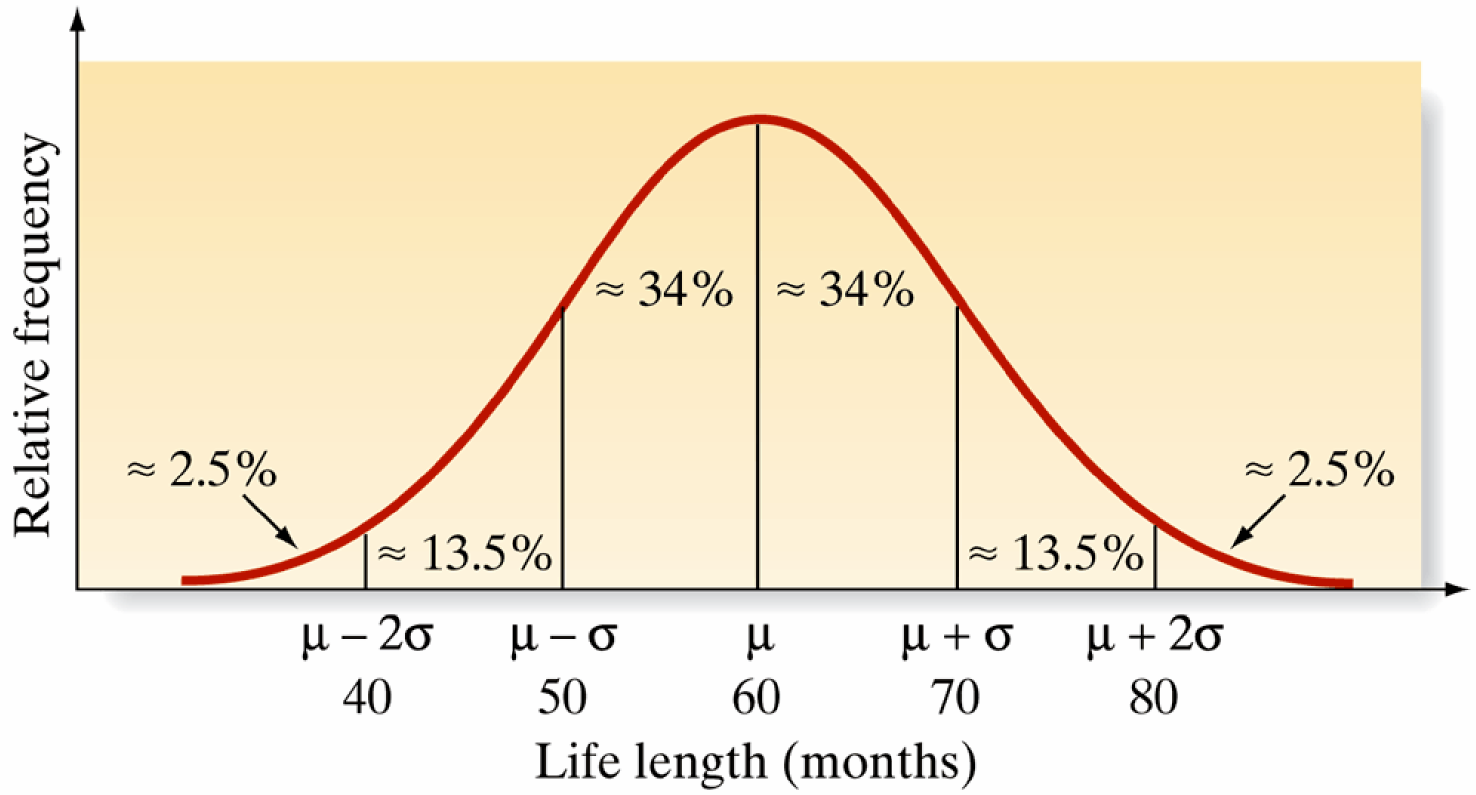
**d.** Generally, for any number $k$ greater than 1, at least $(1 - 1/k^2)$ of the measurements will fall within $k$ standard deviations of the mean [i.e., within the interval $(\bar{x} - ks, \bar{x} + ks)$ for samples and $(\mu - k\sigma, \mu + k\sigma)$ for populations].



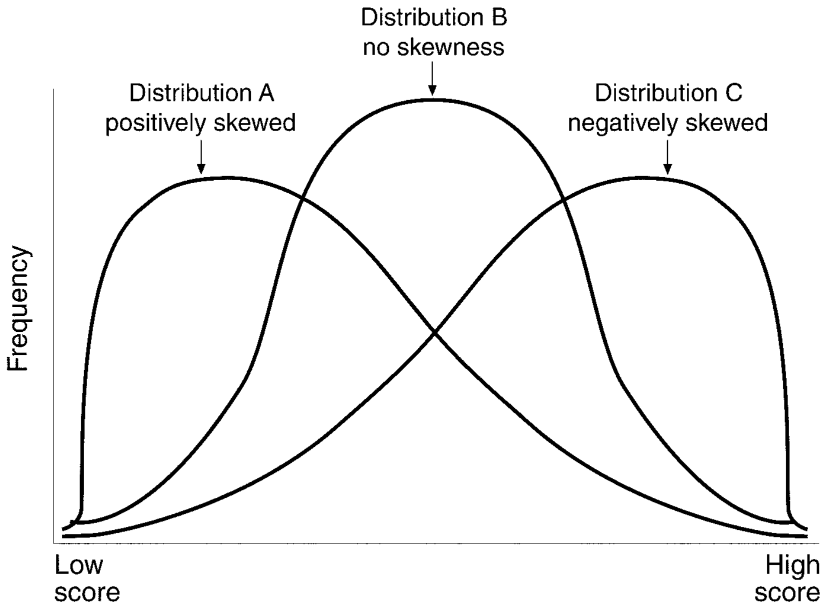
# Example - Battery life-length distribution: manufacturer's claim assumed true

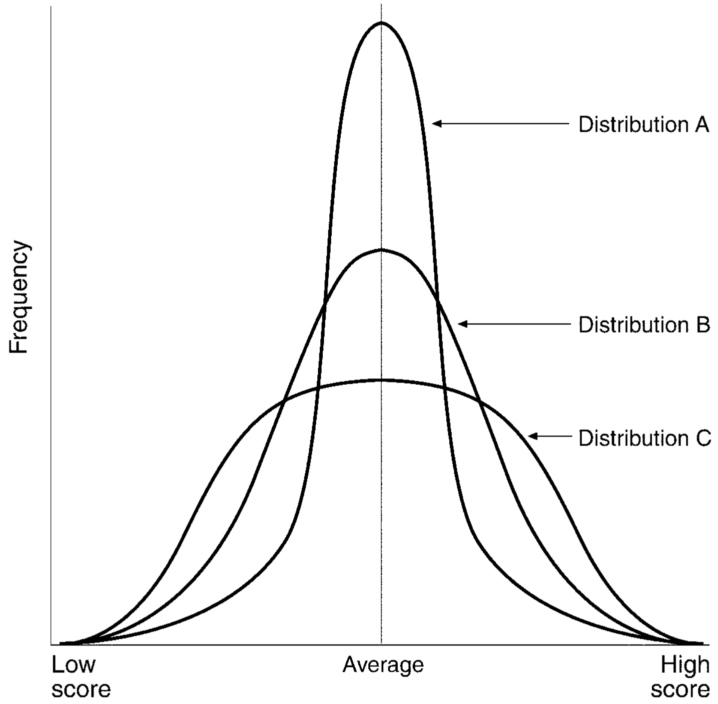
## Claim:

- Battery life = 60 months
- Guarantee = 36 months
- Standard deviation = 10 months



# How Distributions Can Differ in Variability

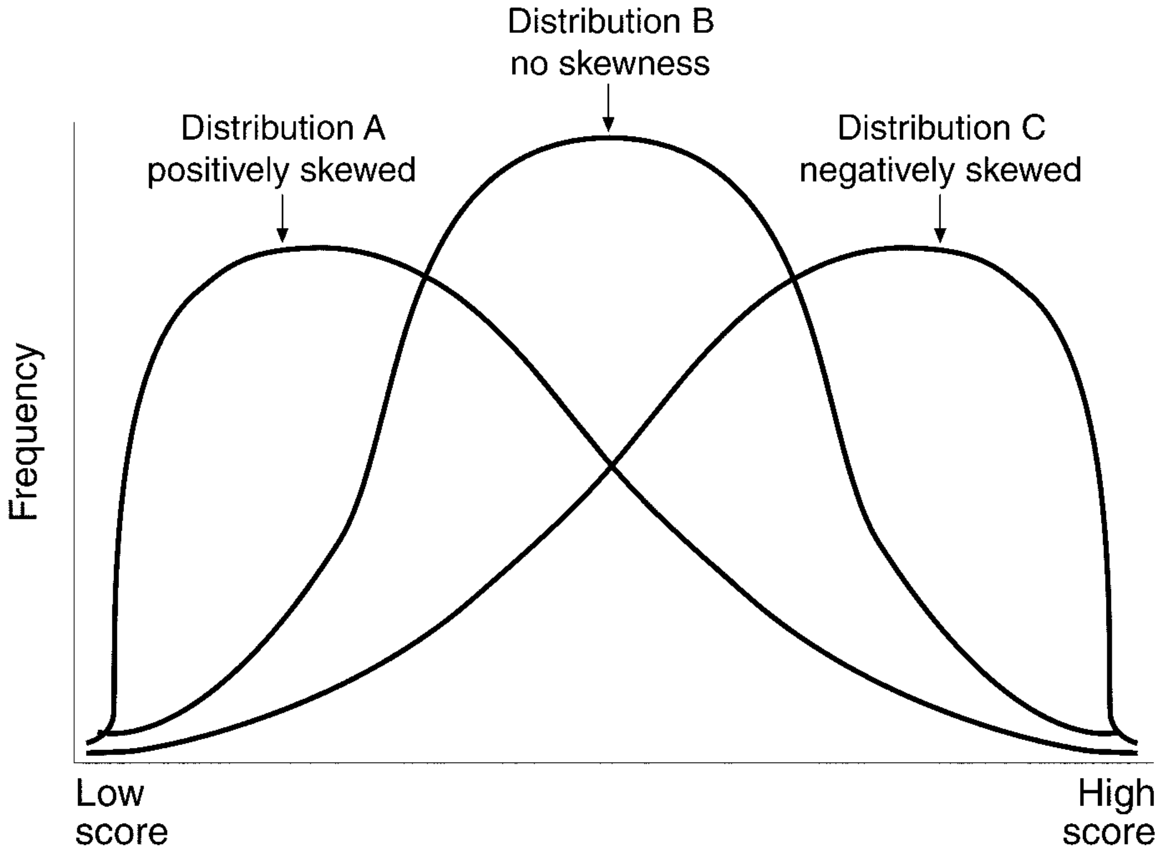



Images sourced from: Salkind (2014) p66.

# Skewness

- Measure of the lack of symmetry, or the lobsidedness, of a distribution



# Skewness – Formula

- Skewness:

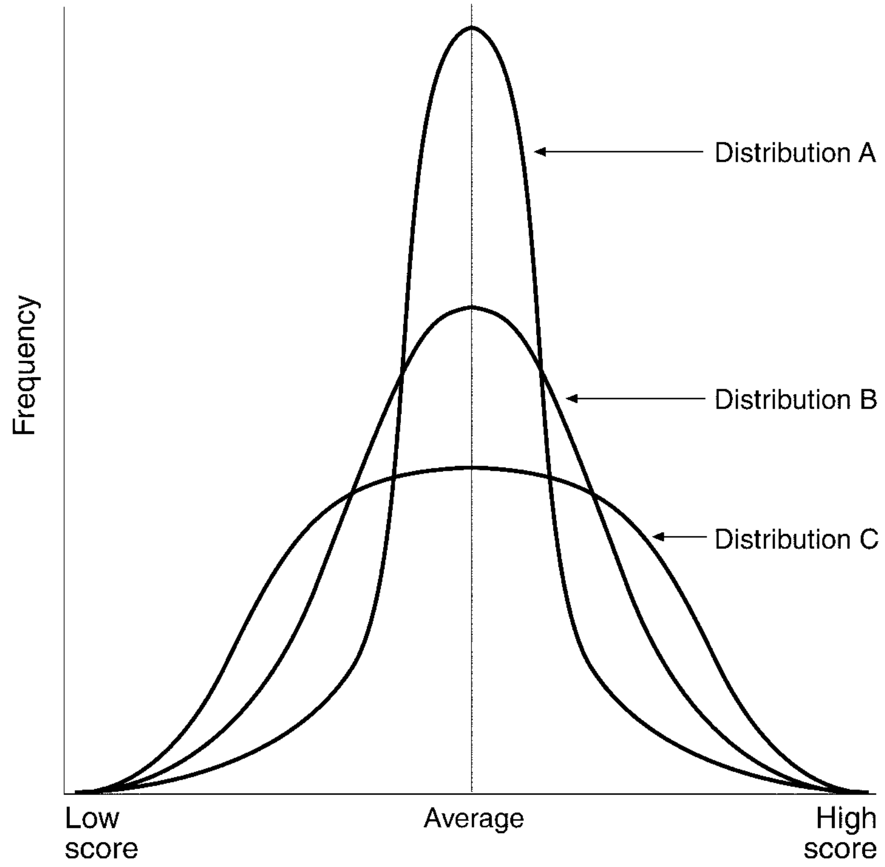
$$Sk = \frac{3(\bar{X} - M)}{s}$$

- EG negative Skewness
  - Mean ($\bar{X}$) = 100
  - Median (M) = 105
  - Standard Deviation (s) = 10
  - Skewness = -1.5

- EG positive Skewness
  - Mean ($\bar{X}$) = 120
  - Median (...) = 116
  - Standard Deviation (s) = 10
  - Skewness = 1.2

# Kurtosis

- How flat or peaked a distribution appears
- (platykurtic/flat, leptokurtic/peaked)

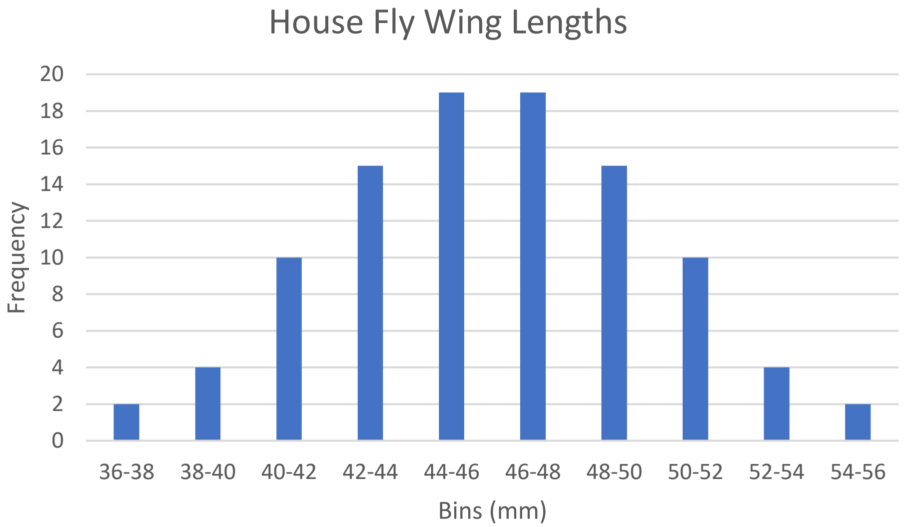


# Normally Distributed Data

- Normally Distributed Housefly Wing Lengths

*Descriptive Statistics*

| length (x.1mm) | |
|---|---|
| | |
| Mean | 45.5 |
| Standard Error | 0.391964748 |
| Median | 45.5 |
| Mode | 45.0 |
| Standard Deviation | 3.91964748 |
| Sample Variance | 15.36363636 |
| Kurtosis | -0.292931078 |
| Skewness | 1.83092E-17 |
| Range | 19 |
| Minimum | 36 |
| Maximum | 55 |
| Sum | 4550 |
| Count | 100 |




Sokal, R.R., and P.E. Hunter. (1955).

# SPSS Output (Frequencies)

**Statistics**

Wing Length (x.1mm)

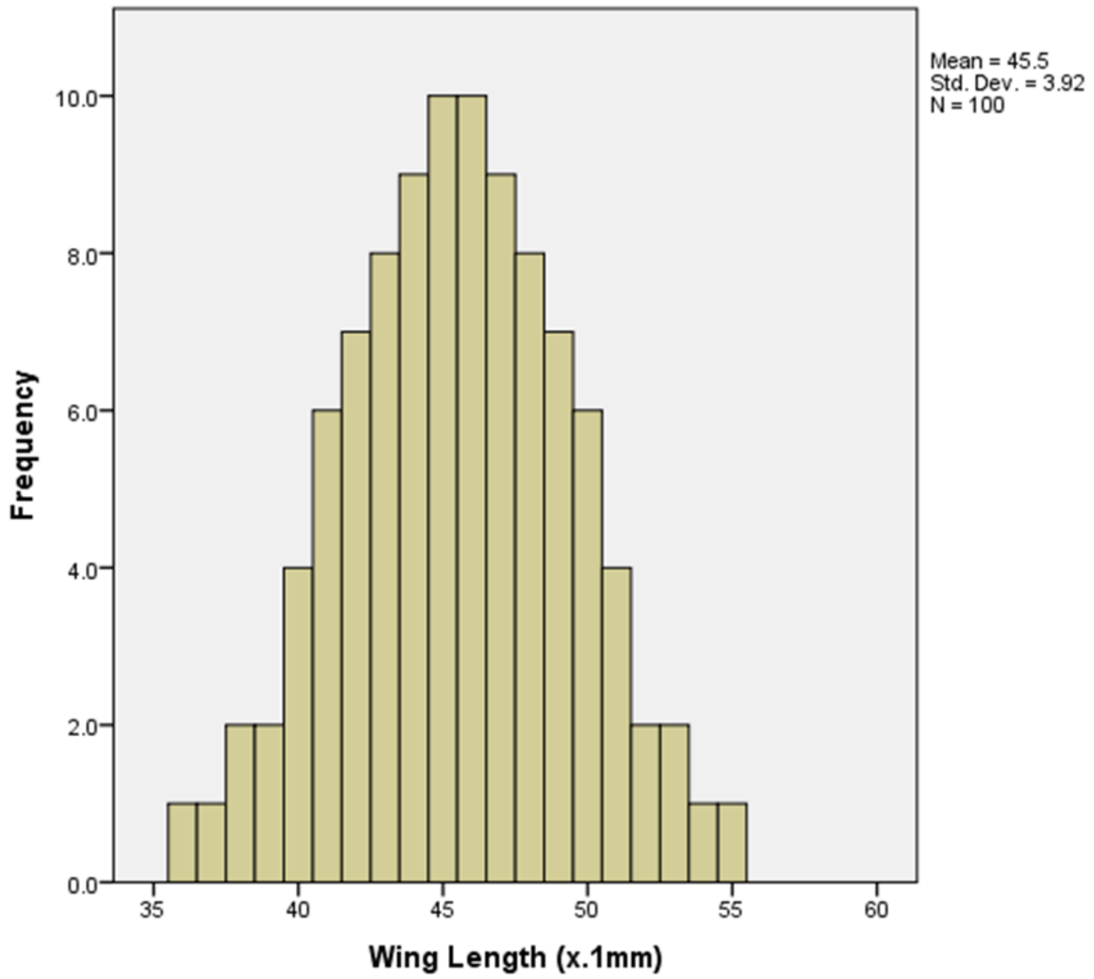
| | | |
|---|---|---|
| N | Valid | 100 |
| | Missing | 0 |
| Mean | | 45.50 |
| Median | | 45.50 |
| Mode | | 45[a] |
| Std. Deviation | | 3.920 |
| Variance | | 15.364 |
| Skewness | | .000 |
| Std. Error of Skewness | | .241 |
| Kurtosis | | -.293 |
| Std. Error of Kurtosis | | .478 |
| Range | | 19 |
| Minimum | | 36 |
| Maximum | | 55 |

a. Multiple modes exist. The smallest value is shown

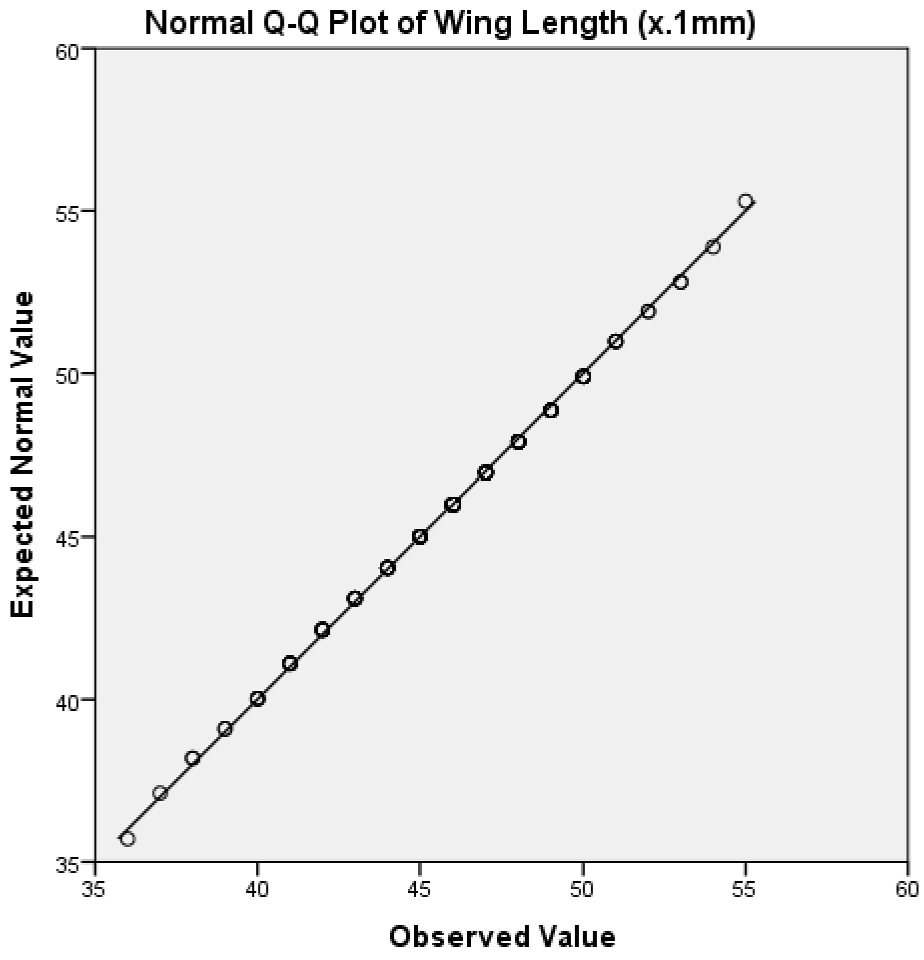
**Wing Length (x.1mm)**

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | 36 | 1 | 1.0 | 1.0 | 1.0 |
| | 37 | 1 | 1.0 | 1.0 | 2.0 |
| | 38 | 2 | 2.0 | 2.0 | 4.0 |
| | 39 | 2 | 2.0 | 2.0 | 6.0 |
| | 40 | 4 | 4.0 | 4.0 | 10.0 |
| | 41 | 6 | 6.0 | 6.0 | 16.0 |
| | 42 | 7 | 7.0 | 7.0 | 23.0 |
| | 43 | 8 | 8.0 | 8.0 | 31.0 |
| | 44 | 9 | 9.0 | 9.0 | 40.0 |
| | 45 | 10 | 10.0 | 10.0 | 50.0 |
| | 46 | 10 | 10.0 | 10.0 | 60.0 |
| | 47 | 9 | 9.0 | 9.0 | 69.0 |
| | 48 | 8 | 8.0 | 8.0 | 77.0 |
| | 49 | 7 | 7.0 | 7.0 | 84.0 |
| | 50 | 6 | 6.0 | 6.0 | 90.0 |
| | 51 | 4 | 4.0 | 4.0 | 94.0 |
| | 52 | 2 | 2.0 | 2.0 | 96.0 |
| | 53 | 2 | 2.0 | 2.0 | 98.0 |
| | 54 | 1 | 1.0 | 1.0 | 99.0 |
| | 55 | 1 | 1.0 | 1.0 | 100.0 |
| | Total | 100 | 100.0 | 100.0 | |

# SPSS Output (Frequency Chart)

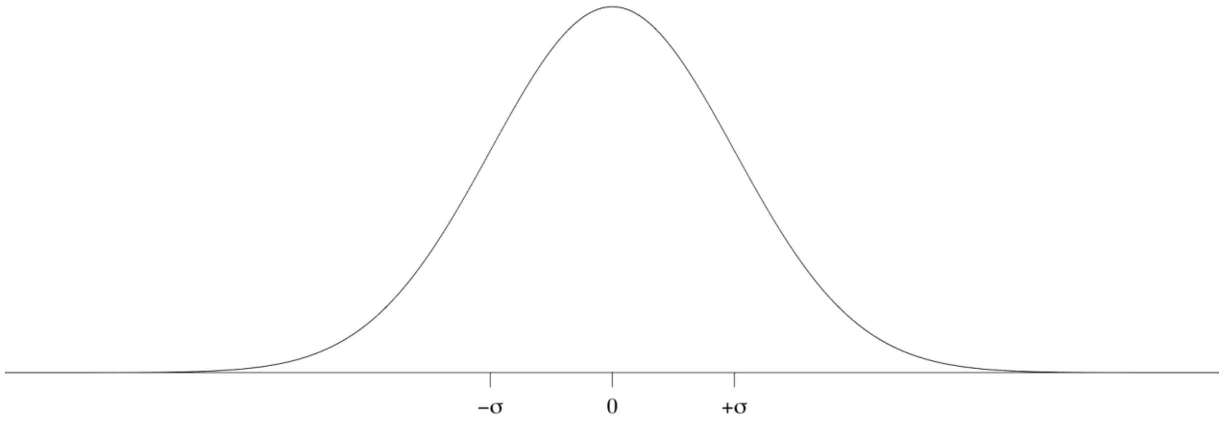


# SPSS Output (Q-Q Plot)



# Tests for Normality

- D'Agostino's K-squared test,
- Jarque–Bera test
- Cramér–von Mises criterion,
- Shapiro-Wilk test
- Kolmogorov-Smirnov test
- Lilliefors test
- Anderson-Darling test




Normal Distribution

# Probability

# Simple Experiment

Link: 

- One Coin:
  - Toss a coin 20 times, record number of heads or tails
- Using two coins:
  - Toss two separate coins ten times, record number of
    - Heads and Heads
    - Heads and Tails
    - Tails and Heads
    - Tails and Tails

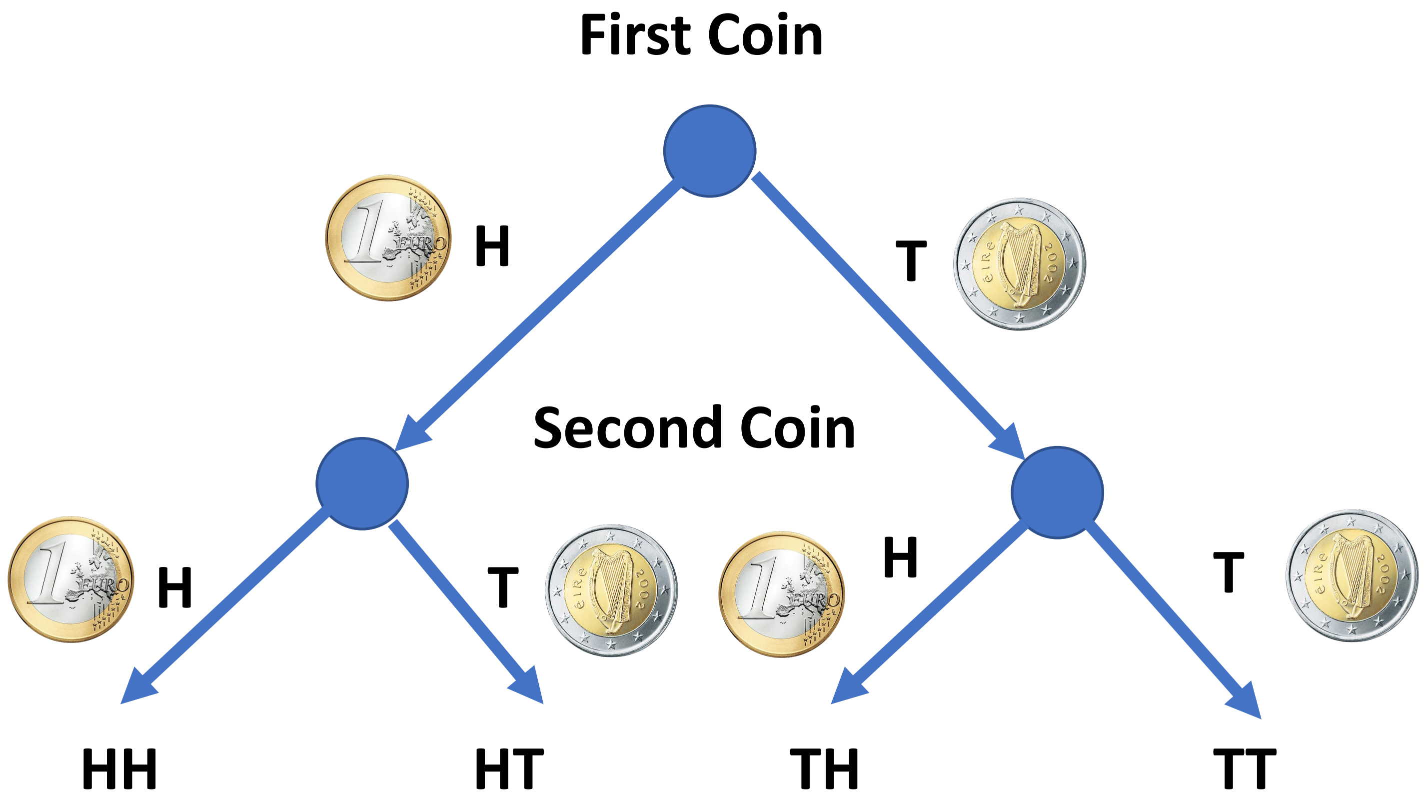
Euro: Heads = Common side
Tails = National side

# Definitions

An **experiment** is an act or process of observation that leads to a single outcome that cannot be predicted with certainty.

A **sample point** is the most basic outcome of an experiment.


# Tree Diagram for Coin Tossing experiment



# Sample spaces

**Experiments and Their Sample Spaces**

*Experiment:* Observe the up face on a coin.
*Sample Space:*
1. Observe a head.
2. Observe a tail.

This sample space can be represented in set notation as a set containing two sample points:

$$S: \{H, T\}$$

Here, $H$ represents the sample point Observe a head and $T$ represents the sample point Observe a tail.

---

*Experiment:* Observe the up face on a die.
*Sample Space:*
1. Observe a 1.
2. Observe a 2.
3. Observe a 3.
4. Observe a 4.
5. Observe a 5.
6. Observe a 6.

This sample space can be represented in set notation as a set of six sample points:

$$S: \{1, 2, 3, 4, 5, 6\}$$

---

*Experiment:* Observe the up faces on two coins.
*Sample Space:*
1. Observe $HH$.
2. Observe $HT$.
3. Observe $TH$.
4. Observe $TT$.

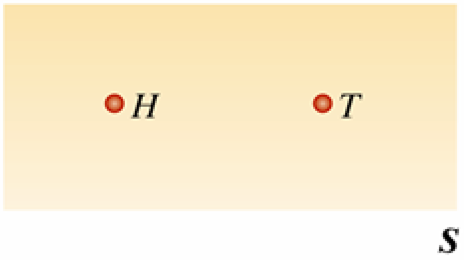
This sample space can be represented in set notation as a set of four sample points:

$$S: \{HH, HT, TH, TT\}$$

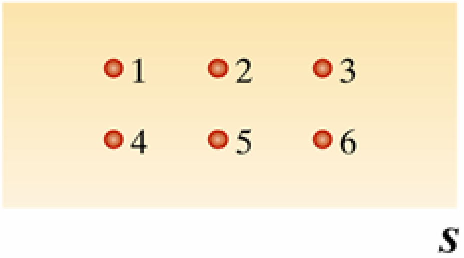
The **sample space** of an experiment is the collection of all its sample points.

# Sample spaces

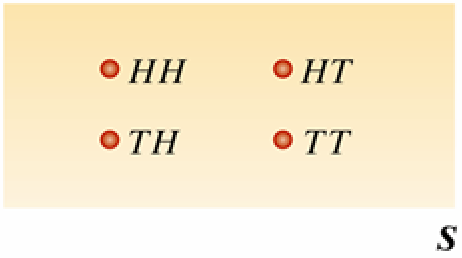
- Each sample represented by a dot
- Labelled accordingly




a. Experiment: Observe the up face on a coin




b. Experiment: Observe the up face on a die




c. Experiment: Observe the up faces on two coins

# Probability

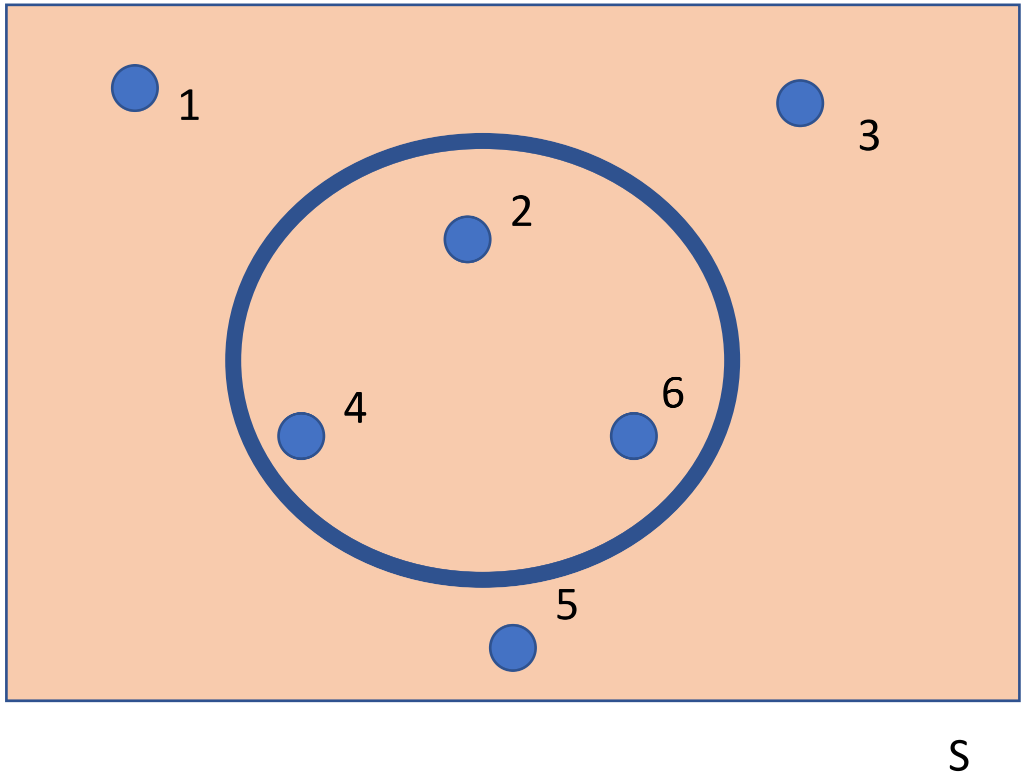
**Probability Rules for Sample Points**

Let $p_i$ represent the probability of sample point $i$. Then

1. All sample point probabilities *must* lie between 0 and 1 (i.e., $0 \leq p_i \leq 1$).
2. The probabilities of all the sample points within a sample space *must* sum to 1 (i.e., $\Sigma p_i = 1$).

# Die Toss

- Die-toss experiment with event A, observe an even number



# Events

An **event** is a specific collection of sample points.

**Probability of an Event**

The probability of an event $A$ is calculated by summing the probabilities of the sample points in the sample space for $A$.

**Steps for Calculating Probabilities of Events**

1. Define the experiment; that is, describe the process used to make an observation and the type of observation that will be recorded.
2. List the sample points.
3. Assign probabilities to the sample points.
4. Determine the collection of sample points contained in the event of interest.
5. Sum the sample point probabilities to get the probability of the event.

# Combinatorial Mathematics

- Determine the number of sample points in large experiments

**Combinations Rule**

Suppose a sample of $n$ elements is to be drawn without replacement from a set of $N$ elements. Then the number of different samples possible is denoted by $\binom{N}{n}$ and is equal to

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}$$

where

$$n! = n(n-1)(n-2)\cdots(3)(2)(1)$$

and similarly for $N!$ and $(N-n)!$ For example, $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$. [*Note:* The quantity $0!$ is defined to be equal to 1.]

# Simple Example

- Consider the task of choosing two students from a group of four to make a presentation
- How many different selections can be made?

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}$$

$$n! = n(n-1)(n-2)\cdots(3)(2)(1)$$

**Solution:**

N = 4, n = 2

$$\binom{4}{2} = \frac{4!}{2!2!} = \frac{4 \times 3 \times 2 \times 1}{(2 \times 1)(2 \times 1)} = 6$$

**There are 6 possible selections.**

# More complex example

- Suppose a  movie reviewer for a newspaper reviews 5 movies each month. This month, the reviewer has 20 new movies from which to make a selection

- How many different samples of the 5 movies can be selected from the 20?

N = 20, n = 5

Answer = 15,504

Use “FACT” function in Excel for calculations.

# Random Sampling

- How a sample is selected from a population is of vital importance in statistical inference
- The probability of an observed sample will be used to infer the characteristics of a sampled population

# Random Sampling

- **What is probability of drawing four aces from a shuffled deck of 52 cards?**

(i.e. first 4 cards drawn without replacement)

- Chances of picking one ace = 4/52
- A second ace = 3/51
- Third ace = 2/50
- Fourth ace = 1/49

- **Overall** 4/52 x 3/51 x 2/50 x 1/49

= 0.0000003693785

=> **1 in 270,725**

# Hypothesis Testing

# Test of Hypothesis

- How many tissues in a box of Kleenex?
- Does a driver's blood alcohol level exceed the legal limit after two drinks?
- Do the majority of registered voters approve of the President's performance?

Image source: http://openreflections.wordpress.com/2011/08/03/on-crowd-funding-open-access-scholarly-books/

# Hypotheses

A statistical **hypothesis** is a statement about the numerical value of a population parameter.

The **null hypothesis,** denoted $H_0$, represents the hypothesis that will be accepted unless the data provide convincing evidence that it is false. This usually represents the "status quo" or some claim about the population parameter that the researcher wants to test.

The **alternative (research) hypothesis,** denoted $H_a$, represents the hypothesis that will be accepted only if the data provide convincing evidence of its truth. This usually represents the values of a population parameter for which the researcher wants to gather evidence to support.

# Making a Judgement

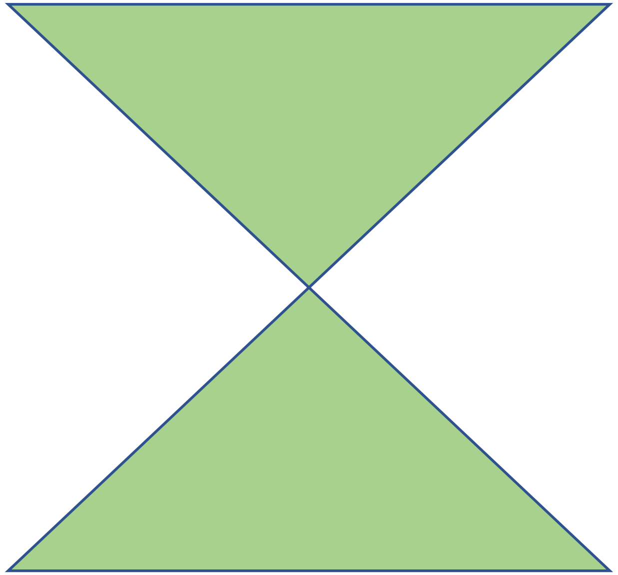
- *"Innocent"…*

  *…until proven…*

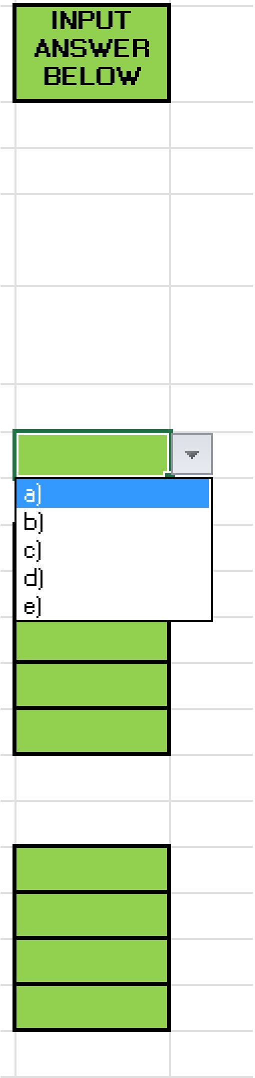
  *"Guilty"*

- Basis of Justice system
- Burden of proof is on Prosecution

# Complete BC1 Quiz  - on Moodle Page



# Questions?




| | |
|---|---|
| Quiz Start Date | 7/10/2024 |
| End Date | 21/10/2024 |

(record any incorrect answers before End Date)

**Only input cells are available - other cells are locked**

# Hypothesis

A statistical **hypothesis** is a statement about the numerical value of a population parameter.

- Suppose:

  Building Regulations state that the crushing strength of domestic concrete blocks should be more than 25 $kN/m^2$

# Null Hypothesis

The **null hypothesis,** denoted $H_0$, represents the hypothesis that will be accepted unless the data provide convincing evidence that it is false. This usually represents the "status quo" or some claim about the population parameter that the researcher wants to test.

- Null Hypothesis ($H_0$): $\mu \leq 25$

(i.e. the manufacturer's blocks do not meet the required specification)

*$\mu$ = population mean*

# Alternate (Research) Hypothesis

The **alternative (research) hypothesis,** denoted $H_a$, represents the hypothesis that will be accepted only if the data provide convincing evidence of its truth. This usually represents the values of a population parameter for which the researcher wants to gather evidence to support.

- Alternative Hypothesis ($H_a$): $\mu > 25$

(i.e. the manufacturer's blocks meet requirements)

*$\mu$ = population mean*

# Population Vs. Sample

- Hypothesis concerns population mean ($\mu$)
- Use sample mean ( $\overline{X}$ ) to make the inference
- We conclude that the blocks meet specifications <u>only</u> when the sample mean <u>convincingly</u> indicates that the population mean exceeds 25 KiloNewtons per square metre.

# Test Statistic

- “Convincing” evidence
  - When $\overline{X}$ cannot be readily attributed to sampling variability

The **test statistic** is a sample statistic, computed from information provided in the sample, that the researcher uses to decide between the null and alternative hypotheses.

- z-value
  - Measures the distance between the value of $\overline{X}$ and the value of $\mu$

# Calculate Z

***Formula for Z Statistic***

$$Z = \frac{\bar{X} - \mu}{\underbrace{\frac{\sigma}{\sqrt{n}}}_{\text{Standard Error of the Mean (SEM)}}}$$

$\bar{X}$ = Sample Mean

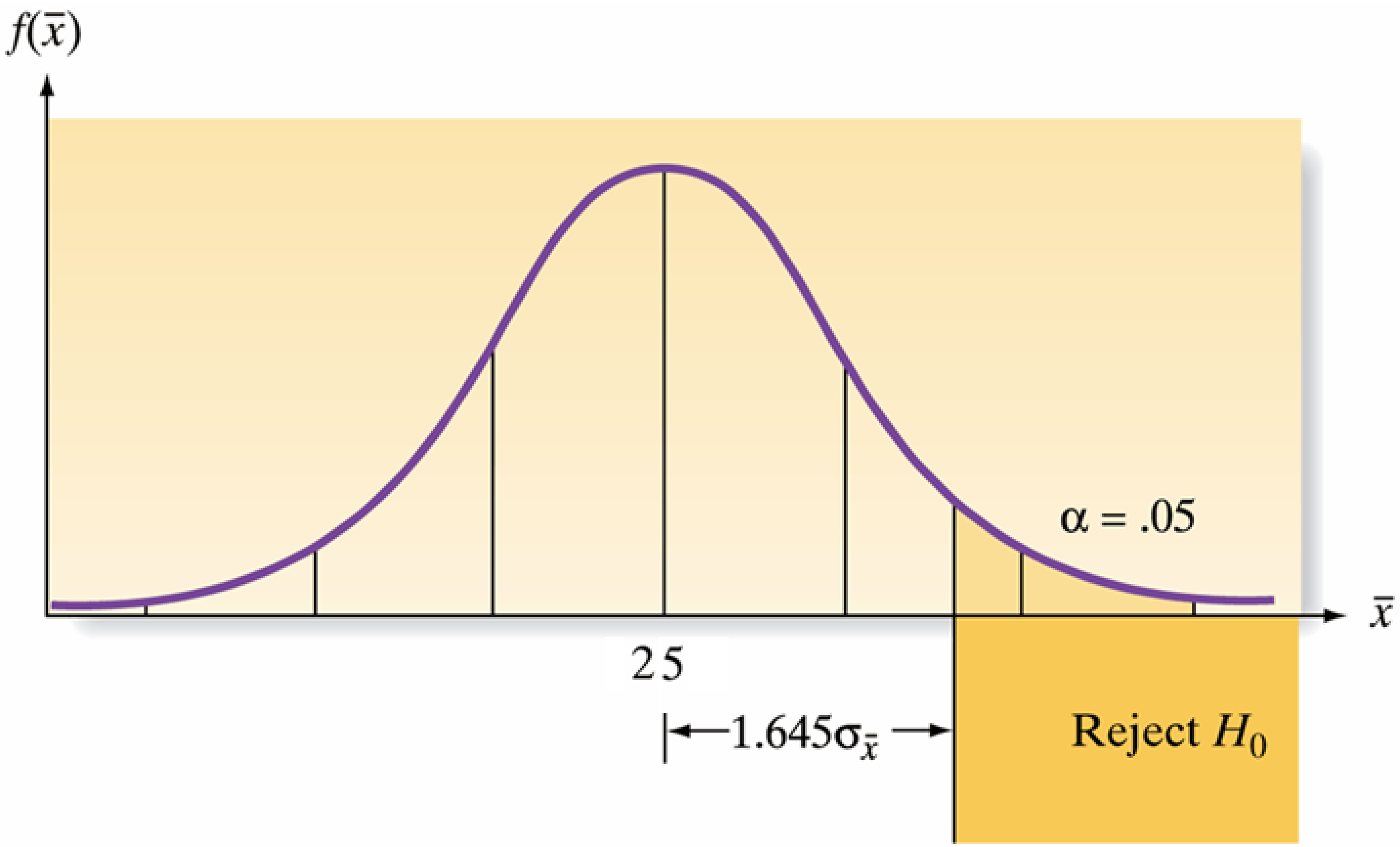
$\mu$ = Population Mean

σ = Population Standard Deviation

n = Size of Sample

# The sampling distribution of $\bar{x}$, Assuming $\mu$ = 25

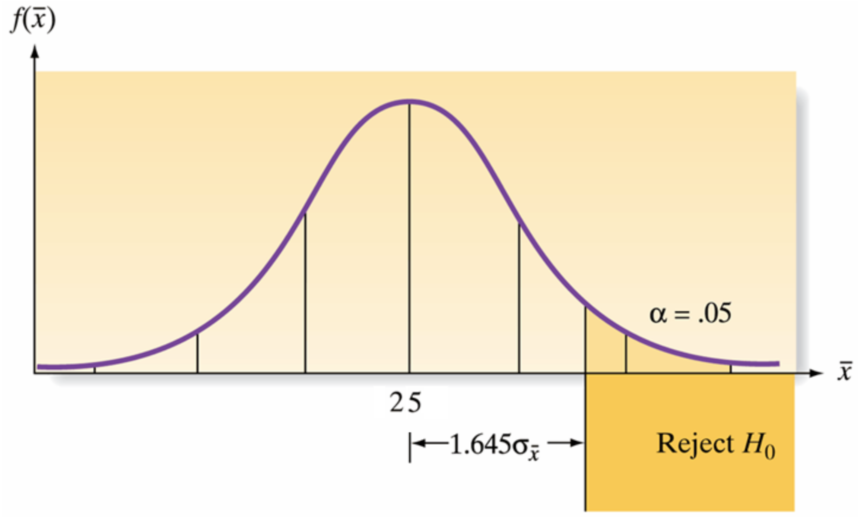
The Central Limit Theorem states that **the distribution of sample means approximates a normal distribution as the sample size gets larger, regardless of the population's distribution**.




| Rejection Regions for Common Values of $\alpha$ | | | |
|---|---|---|---|
| | Alternative Hypotheses | | |
| | Lower-Tailed | Upper-Tailed | Two-Tailed |
| $\alpha = .10$ | $z < -1.28$ | $z > 1.28$ | $z < -1.645$ or $z > 1.645$ |
| $\alpha = .05$ | $z < -1.645$ | $z > 1.645$ | $z < -1.96$ or $z > 1.96$ |
| $\alpha = .01$ | $z < -2.33$ | $z > 2.33$ | $z < -2.575$ or $z > 2.575$ |

# Accept/Reject Hypothesis?

- If the sample mean is more than 1.645 standard deviations above 25,



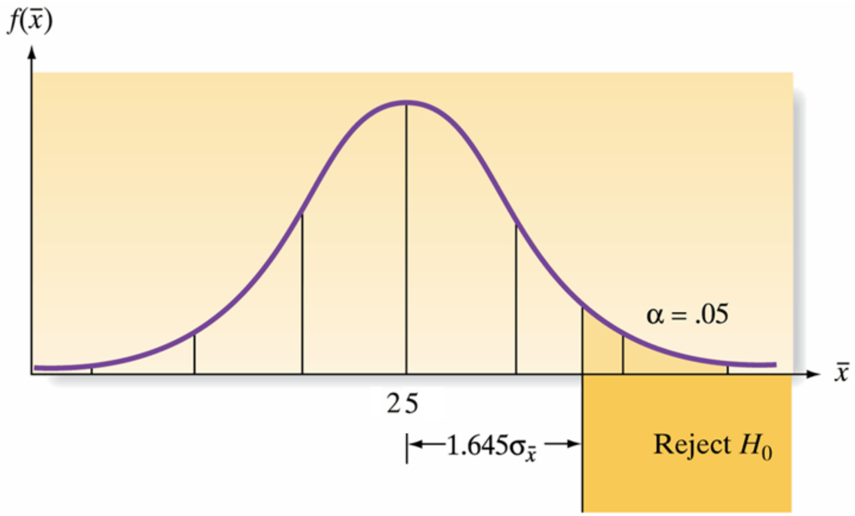

<u>either</u>:

- $H_0$ is true and a relatively rare event has occurred (0.05 probability,) <u>or</u>
- $H_a$ is true and the population mean exceeds 25

# Accept/Reject Hypothesis?

- What is the probability that this procedure will lead us to an incorrect decision?

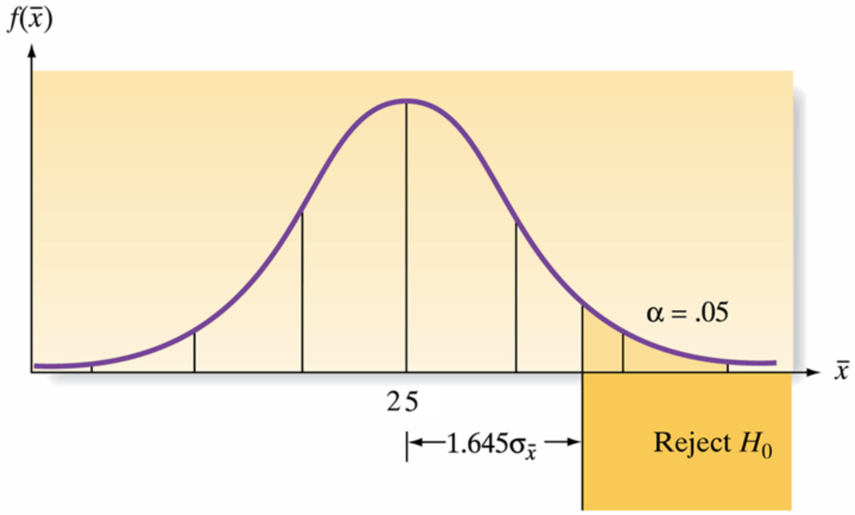
A **Type I error** occurs if the researcher rejects the null hypothesis in favor of the alternative hypothesis when, in fact, $H_0$ is true. The probability of committing a Type I error is denoted by $\alpha$.



# Type I Error

- The risk of making a Type I error is denoted by the symbol α – that is:
  - α = *P* (Type I error)

    = *P* (rejecting the null hypothesis when in fact the null hypothesis is true)



# Example

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

In our example:

$\alpha = P\,(z > 1.645 \text{ when in fact } \mu = 25) = 0.05$

We now summarize the elements of the test:

$H_0$: $\mu \leq 25$ (Blocks do not meet specification)

$H_a$: $\mu > 25$ (Blocks do meet specification)

$$z \text{ value} = \frac{\bar{X} - 25}{\sigma/\sqrt{N}}$$

Rejection region: $z > 1.645$, which corresponds to $\alpha = 0.05$

The **rejection region** of a statistical test is the set of possible values of the test statistic for which the researcher will reject $H_0$ in favor of $H_a$.

# Illustration

- To illustrate use of the test, suppose we test 50 blocks and find the mean and standard deviation to be:
  - $\bar{X}$ = 25.60 kN/m$^2$
  - s = 2.00 kN/m$^2$
  - The test statistic is:

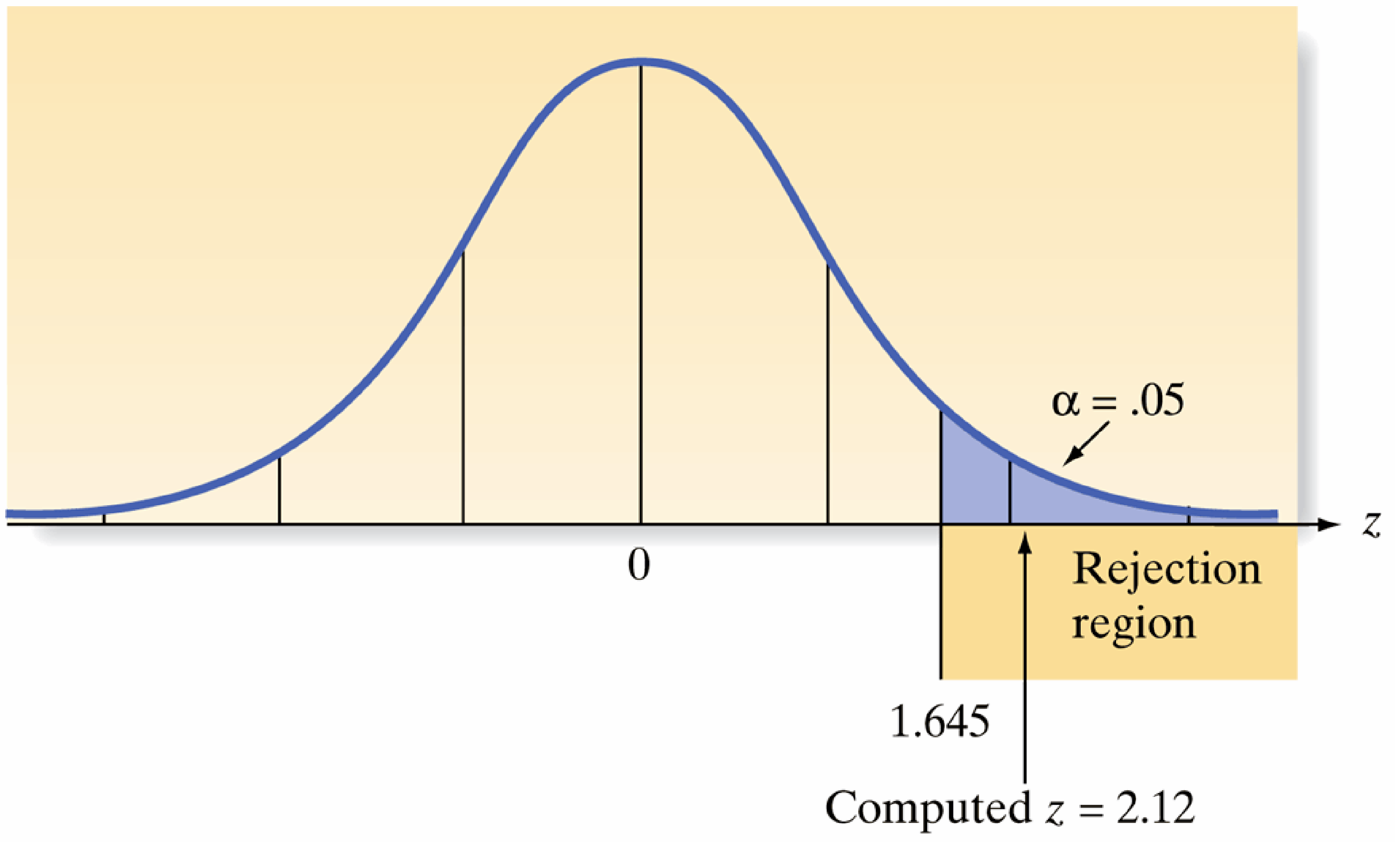
Use *s* to approximate σ when *s* is calculated from a large set of sample measurements.

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$z = \frac{25.60 - 25}{2.00 / \sqrt{50}} = \frac{0.60}{28.28} = 2.121$$

Now What?

# Location of the test statistic for a test of the hypothesis $H_0$: $\mu = 25$




The sample mean lies 2.121 $\sigma_{\bar{x}}$ above the hypothesized value of $\mu$ (25).

# What if…

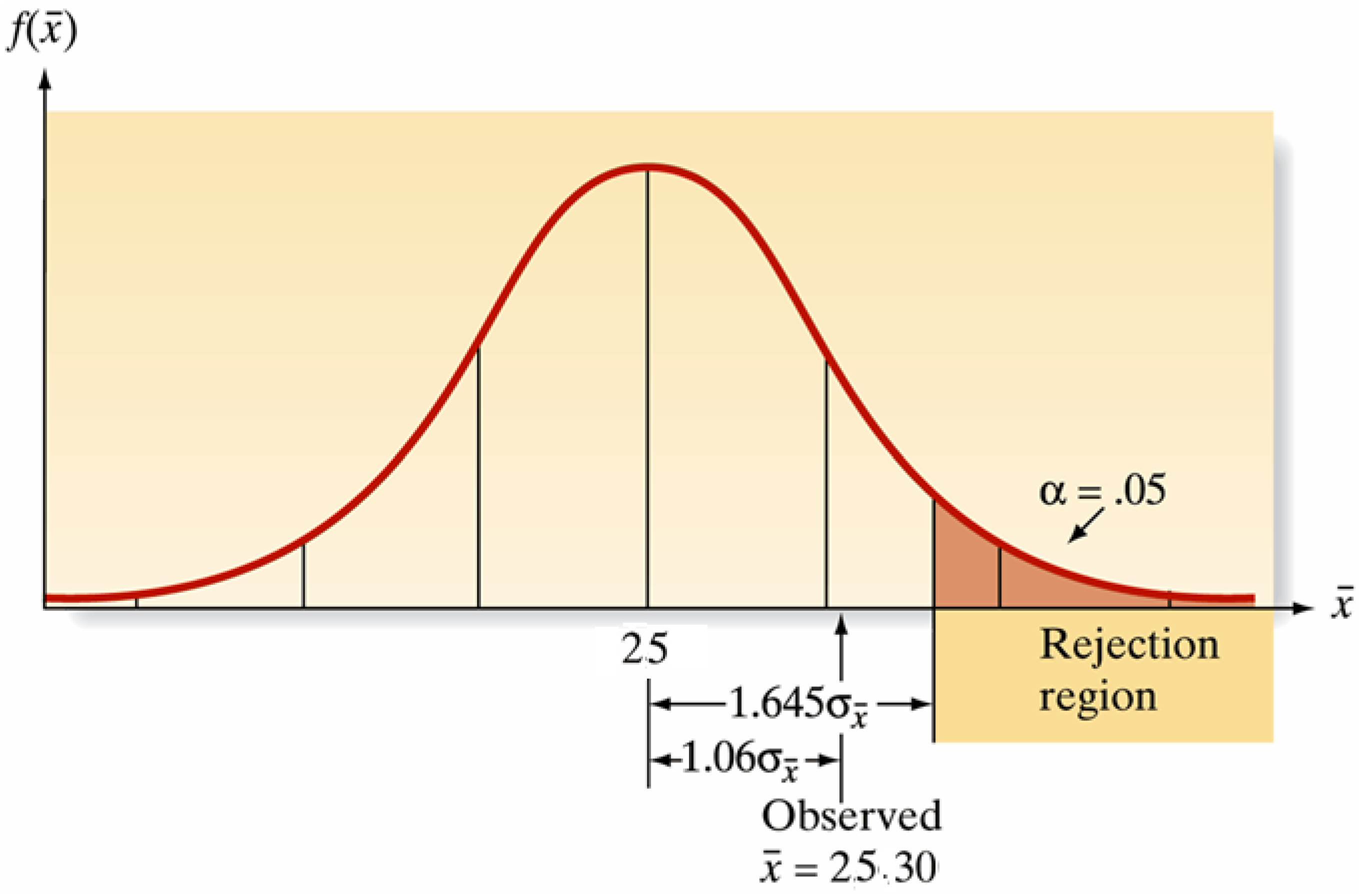
- Suppose we test 50 blocks and find the mean and standard deviation to be:
  - $\bar{X}$ = 25.3 kN/m2
  - s = 2.00 kN/m2
  - The test statistic is:

Use *s* to approximate σ when *s* is calculated from a large set of sample measurements.

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$z = \frac{25.30 - 25}{200 / \sqrt{50}} = \frac{0.30}{28.28} = 1.06$$

# Location of test statistic when $\bar{x} = 25$



# "Convincing" evidence

- We cannot reject $H_0$ using $\alpha = 0.05$

- Even though the sample mean ($\bar{x}$) of 25.3 exceeds specification (25) by 0.30

- It does not exceed the specification by enough to provide **convincing** evidence that the **population mean** ($\mu$) exceeds 25.

A **Type II error** occurs if the researcher accepts the null hypothesis when, in fact, $H_0$ is false. The probability of committing a Type II error is denoted by $\beta$.

Conclusions

# Errors in Hypothesis Testing

- Type I Error
  - Defined as the probability of rejecting the null hypothesis when it is actually true.
  - This is denoted by the Greek letter $\alpha$.
  - Also known as the *significance level* of a test.
- Type II Error
  - Defined as the probability of failing to reject the null hypothesis when it is actually false.
  - This is denoted by the Greek letter $\beta$.

| Null Hypothesis | Researcher: Does Not Reject $H_0$ | Researcher: Rejects $H_0$ |
|---|---|---|
| $H_0$ is true | Correct decision | Type I error |
| $H_0$ is false | Type II error | Correct decision |

- Both a small $\alpha$ and a small $\beta$ are desirable.
- For a given type of test and fixed sample size, there is a trade-off between $\alpha$ and $\beta$.
- The larger critical value needed to reduce $\alpha$ risk makes it harder to reject $H_0$, thereby increasing $\beta$ risk. Both $\alpha$ and $\beta$ can be reduced simultaneously only by increasing the sample size.

# Formulating Hypotheses

- Recall…
  - The null hypothesis and alternative hypothesis form the basis for inference using a test of hypothesis
- We tested…

$(H_0)$: $\mu \leq 25$ (Concrete Blocks do not meet specification)

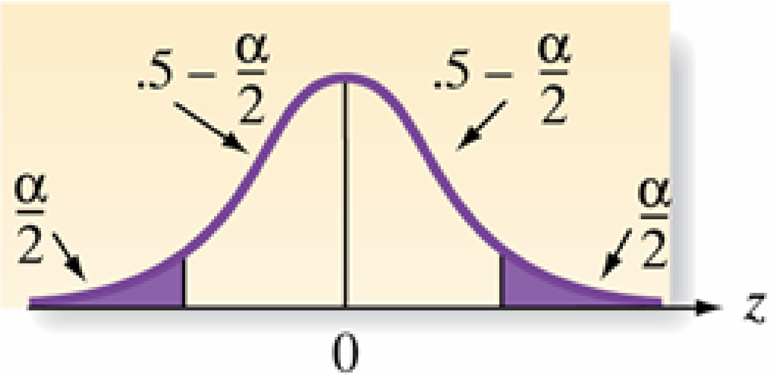
$(H_a)$: $\mu > 25$ (Concrete Blocks do meet specifications)

- This is a **one-tailed** (or one-sided) statistical test

# Suppose…

- Want to show that the population parameter is **either larger or smaller** than a specified value

- Such an alternative hypothesis is called a **two-tailed** hypothesis



c. Form of $H_a$: $\neq$

# One/Two tail tests

**Steps for Selecting the Null and Alternative Hypotheses**

1. Select the *alternative hypothesis* as that which the sampling experiment is intended to establish. The alternative hypothesis will assume one of three forms:
   - **a.** One-tailed, **upper-tailed** (e.g., $H_a$: $\mu > 2{,}400$)
   - **b.** One-tailed, **lower-tailed** (e.g., $H_a$: $\mu < 2{,}400$)
   - **c.** Two-tailed (e.g., $H_a$: $\mu \neq 2{,}400$)

2. Select the *null hypothesis* as the status quo, that which will be presumed true unless the sampling experiment conclusively establishes the alternative hypothesis. The null hypothesis will be specified as that parameter value closest to the alternative in one-tailed tests and as the complementary (or only unspecified) value in two-tailed tests.

   (e.g., $H_0$: $\mu = 2{,}400$)

A **one-tailed test** of hypothesis is one in which the alternative hypothesis is directional and includes the symbol "<" or " > ."

A **two-tailed test** of hypothesis is one in which the alternative hypothesis does not specify departure from $H_0$ in a particular direction and is written with the symbol "$\neq$."

# Example

- A metal lathe is checked periodically by quality control inspectors to determine whether it is producing machine bearings with a mean diameter of 2cms
- If the mean is larger or smaller than 2cms, then the process is out of control
- Formulate the null hypothesis and alternative hypotheses for a test to determine whether the bearing production process is out of control


Image source: http://metal.baileighindustrial.com/metal-lathe-pl-1440e


Image source: http://faqb2b.blog.com/ball-bearing-faq/

# Solution

- The hypothesis must be stated in terms of a population parameter
- $\mu$ is the true mean diameter (in cms)

- If either $\mu > 2$ or $\mu < 2$ then the process is out of control
  - These are the **alternative** hypotheses ($H_a$)

- $\mu = 2$ represents an in-control process (the *status quo*)
  - this is the **Null** hypothesis

$(H_0)$: $\mu = 2$ (the process is in control)

$(H_a)$: $\mu \neq 2$ (the process is <u>out</u> in control)

# Rejection regions

## One-tailed vs. Two-tailed Test

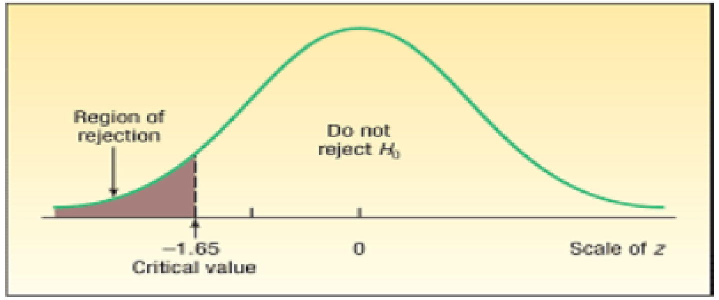
The *direction of the test* is indicated by which way the inequality symbol points in $H_1$:

> indicates a right-tailed test

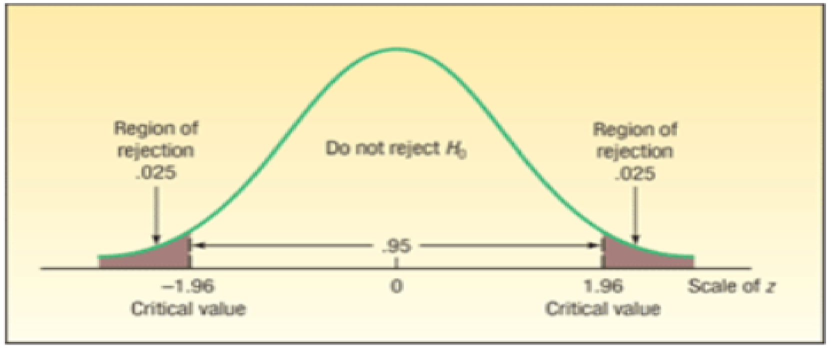
< indicates a left-tailed test

≠ indicates a two-tailed test

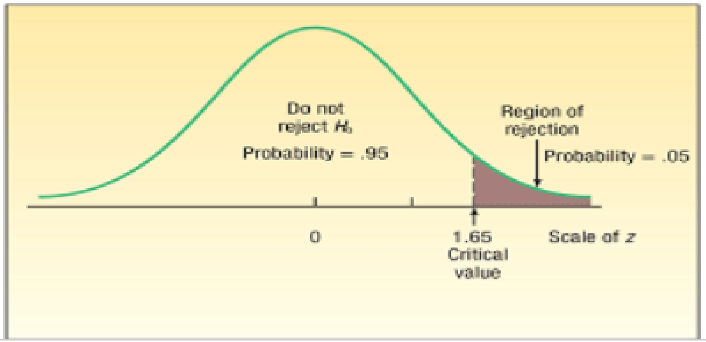
One-tailed or Left-tailed Test




Two-tailed or Non-directional Test




One-tailed or Right-tailed Test



# Typical values

| Rejection Regions for Common Values of $\alpha$ | | | |
|---|---|---|---|
| | Alternative Hypotheses | | |
| | Lower-Tailed | Upper-Tailed | Two-Tailed |
| $\alpha = .10$ | $z < -1.28$ | $z > 1.28$ | $z < -1.645$ or $z > 1.645$ |
| $\alpha = .05$ | $z < -1.645$ | $z > 1.645$ | $z < -1.96$ or $z > 1.96$ |
| $\alpha = .01$ | $z < -2.33$ | $z > 2.33$ | $z < -2.575$ or $z > 2.575$ |

- Note that the smaller α you select, the more evidence (i.e. larger Z) you will need before you can reject $H_0$

# Normal Z Statistic

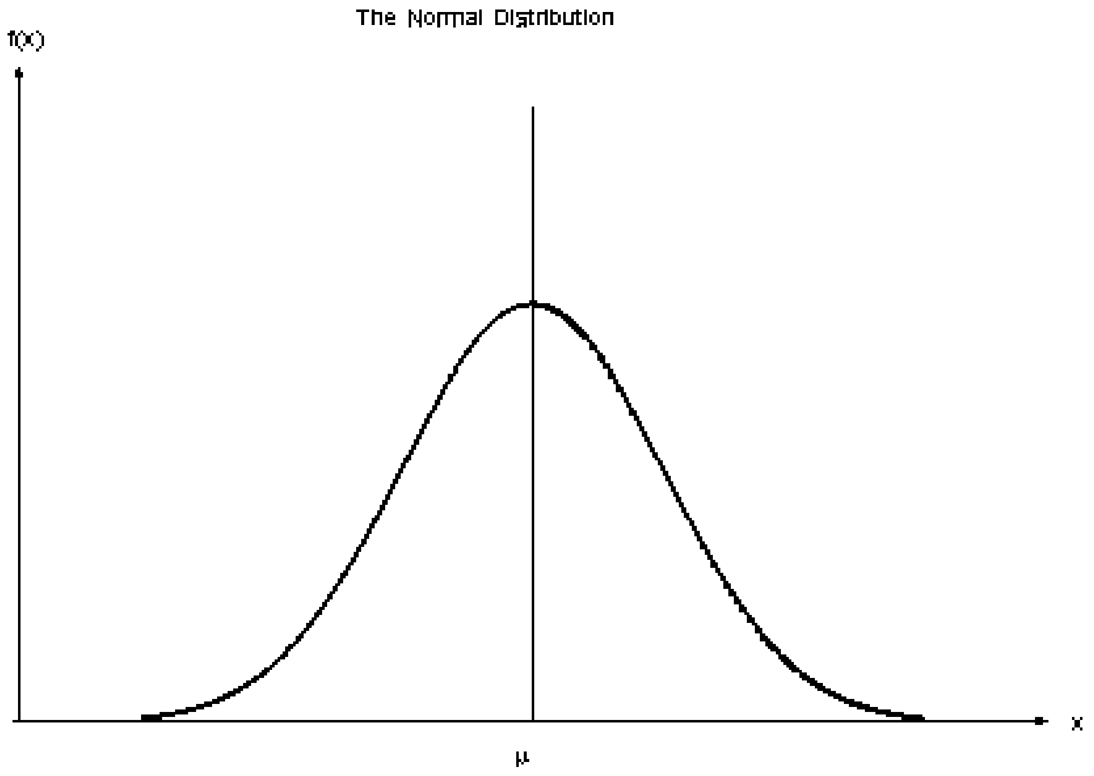
- When *n* is large, the sample standard deviation (*s*) provides a good approximation to population standard deviation (σ)
- Z statistic can be approximated as follows:

$$\text{z value} = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \approx \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

# Large Samples

- When testing a hypothesis about a **Population Mean** ($\mu$), the test statistic (z) we use will depend on whether the sample size is large (e.g. $n \geq 30$) or small, and whether or not we know the value of the population standard deviation ($\sigma$)

*In large samples the distribution* of $\overline{X}$ is approximately normal




The Central Limit Theorem states that **the distribution of sample means approximates a normal distribution as the sample size gets larger, regardless of the population's distribution**.

# Procedure

**Large-Sample Test of Hypothesis about $\mu$ Based on a Normal ($z$) Statistic**

| One-Tailed Test | Two-Tailed Test |
|---|---|
| $H_0$: $\mu = \mu_0$ | $H_0$: $\mu = \mu_0$ |
| $H_a$: $\mu < \mu_0$ (or $H_a$: $\mu > \mu_0$) | $H_a$: $\mu \neq \mu_0$ |
| *Test Statistic* ($\sigma$ Known): $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ | *Test statistic* ($\sigma$ Known): $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ |
| *Test Statistic* ($\sigma$ unknown): $z \approx \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$ | *Test statistic* ($\sigma$ unknown): $z \approx \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$ |
| *Rejection region*: $z < -z_\alpha$ (or $z > z_\alpha$ when $H_a$: $\mu > \mu_0$) where $z_\alpha$ is chosen so that $P(z > z_\alpha) = \alpha$ | *Rejection region*: $\lvert z \rvert > z_{\alpha/2}$ where $z_{\alpha/2}$ is chosen so that $P(\lvert z \rvert > z_{\alpha/2}) = \alpha/2$ |

*Note:* The symbol for the numerical value assigned to $\mu$ under the null hypothesis is $\mu_0$.

# The Procedure of Hypothesis Testing

## critical value approach

1. State the null hypothesis ($H_0$)
2. State the alternative hypothesis ($H_1$)
3. State the significance level ($\alpha$) of the test
4. Find a critical value for the test statistic from a table
5. Calculate an appropriate test statistic from the data
6. Reach a conclusion by comparing the test statistic with the critical value.

# Conditions

**Conditions Required for a Valid Large-Sample Hypothesis Test for $\mu$**

1. A random sample is selected from the target population.
2. The sample size $n$ is large (i.e., $n \geq 30$). (Due to the Central Limit Theorem, this condition guarantees that the test statistic will be approximately normal regardless of the shape of the underlying probability distribution of the population.)

- Once the test has been set up, the sampling experiment is performed and the test statistic calculated

# Possible conclusions

**Possible Conclusions for a Test of Hypothesis**

1. If the calculated test statistic falls in the rejection region, reject $H_0$ and conclude that the alternative hypothesis $H_a$ is true. State that you are rejecting $H_0$ at the $\alpha$ level of significance. Remember that the confidence is in the testing *process*, not the particular result of a single test.
2. If the test statistic does not fall in the rejection region, conclude that the sampling experiment does not provide sufficient evidence to reject $H_0$ at the $\alpha$ level of significance. [Generally, we will not "accept" the null hypothesis unless the probability $\beta$ of a Type II error has been calculated

# Example

- Carrying out a Hypothesis Test

for $\mu$ - Mean drug response Time

- Suppose a neurologist is testing the effect of a drug response time by:
  - injecting 100 rats with a unit dose of the drug
  - subject each rat to a neurological stimulus
  - record the response time
- Response time for rats not injected with the drug (control group) is 1.2 seconds
- Test whether the mean response time for drug-injected rats differs from 1.2 secs

**Drug Response Times for 100 Rats**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1.90 | 2.17 | 0.61 | 1.17 | 0.66 | 1.86 | 1.41 | 1.30 | 0.70 | 0.56 |
| 2.00 | 1.27 | 0.98 | 1.55 | 0.64 | 0.60 | 1.55 | 0.93 | 0.48 | 0.39 |
| 0.86 | 1.19 | 0.79 | 1.37 | 1.31 | 0.85 | 0.71 | 1.21 | 1.23 | 0.89 |
| 1.84 | 0.80 | 0.64 | 1.08 | 0.74 | 0.93 | 1.71 | 1.05 | 1.44 | 0.42 |
| 0.70 | 0.54 | 1.40 | 1.06 | 0.54 | 0.17 | 0.98 | 0.89 | 1.28 | 0.68 |
| 0.98 | 1.14 | 1.16 | 1.64 | 1.16 | 1.01 | 1.09 | 0.77 | 1.58 | 0.99 |
| 0.57 | 0.27 | 0.51 | 1.27 | 1.81 | 0.88 | 0.31 | 0.92 | 0.93 | 1.66 |
| 0.21 | 0.79 | 0.94 | 0.45 | 1.19 | 1.60 | 0.14 | 0.99 | 1.08 | 1.57 |
| 0.55 | 1.65 | 0.81 | 1.00 | 2.55 | 1.96 | 1.31 | 1.88 | 1.51 | 1.48 |
| 0.61 | 0.05 | 1.21 | 0.48 | 1.63 | 1.45 | 0.22 | 0.49 | 1.29 | 1.40 |

# Solution

- To carry out the test, we need to find values of $\overline{X}$ and $s$

**Table 8.3** **Drug Response Times for 100 Rats, Example 8.4**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1.90 | 2.17 | 0.61 | 1.17 | 0.66 | 1.86 | 1.41 | 1.30 | 0.70 | 0.56 |
| 2.00 | 1.27 | 0.98 | 1.55 | 0.64 | 0.60 | 1.55 | 0.93 | 0.48 | 0.39 |
| 0.86 | 1.19 | 0.79 | 1.37 | 1.31 | 0.85 | 0.71 | 1.21 | 1.23 | 0.89 |
| 1.84 | 0.80 | 0.64 | 1.08 | 0.74 | 0.93 | 1.71 | 1.05 | 1.44 | 0.42 |
| 0.70 | 0.54 | 1.40 | 1.06 | 0.54 | 0.17 | 0.98 | 0.89 | 1.28 | 0.68 |
| 0.98 | 1.14 | 1.16 | 1.64 | 1.16 | 1.01 | 1.09 | 0.77 | 1.58 | 0.99 |
| 0.57 | 0.27 | 0.51 | 1.27 | 1.81 | 0.88 | 0.31 | 0.92 | 0.93 | 1.66 |
| 0.21 | 0.79 | 0.94 | 0.45 | 1.19 | 1.60 | 0.14 | 0.99 | 1.08 | 1.57 |
| 0.55 | 1.65 | 0.81 | 1.00 | 2.55 | 1.96 | 1.31 | 1.88 | 1.51 | 1.48 |
| 0.61 | 0.05 | 1.21 | 0.48 | 1.63 | 1.45 | 0.22 | 0.49 | 1.29 | 1.40 |

$\overline{X} = 1.05$

$s = 0.5$

# Substitute $\bar{x}$ and $s$ (for σ) in formula

$$\text{z value} = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \approx \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

$$\text{z value} = \frac{\bar{x} - 1.2}{\sigma_{\bar{x}}} = \frac{\bar{x} - 1.2}{\sigma/\sqrt{n}}$$

$$\approx \frac{1.05 - 1.2}{0.5/\sqrt{100}} = -3.0$$

# What does this mean?

$\overline{X}$ = 1.05

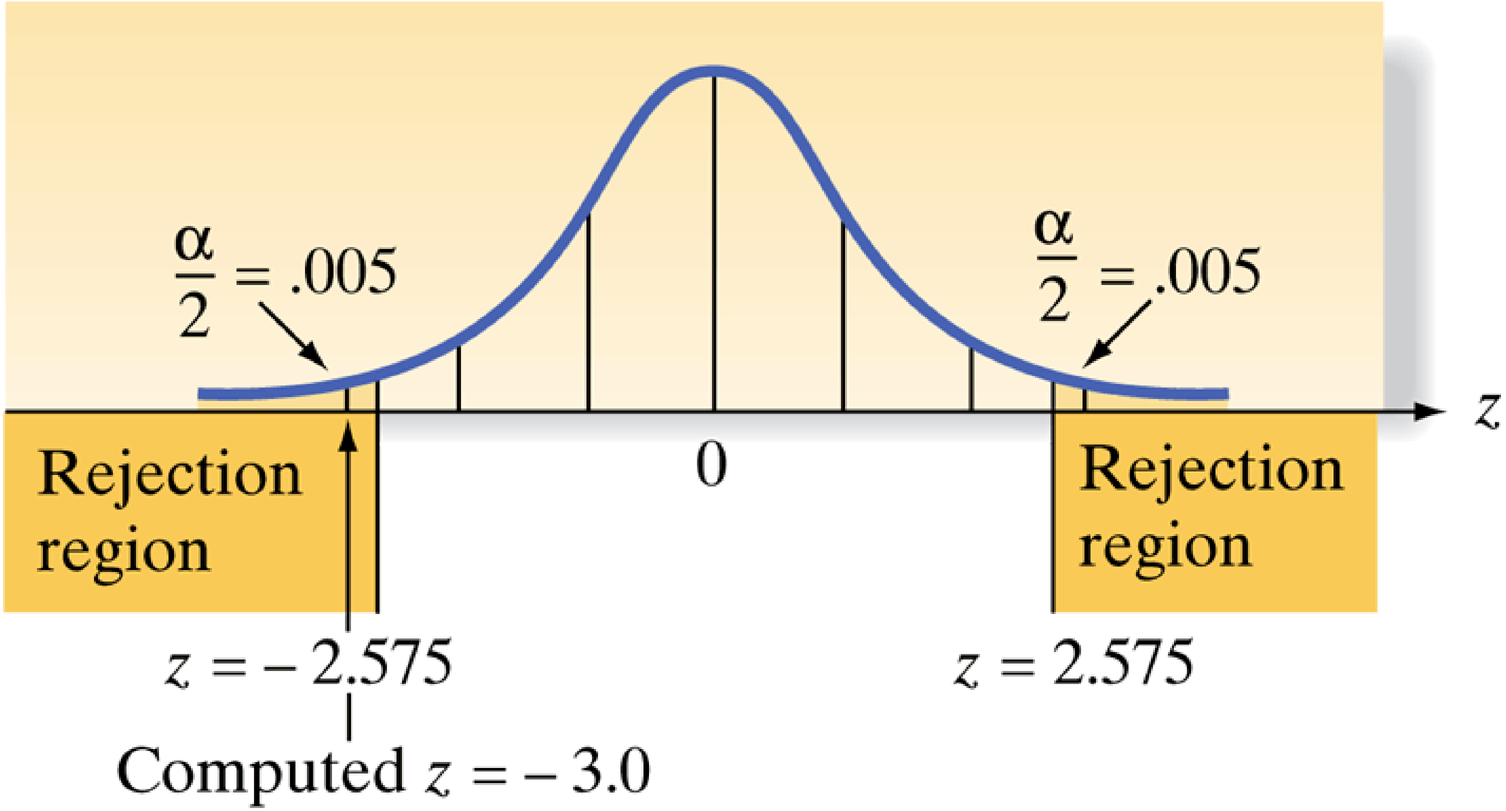
$s$ = 0.5

z = -3.0

($\alpha$ = 0.01)

# Recall…

- Two-tailed rejection region: $\alpha = 0.01$



# Implication of results

$\overline{X}$ = 1.05

$s$ = 0.5

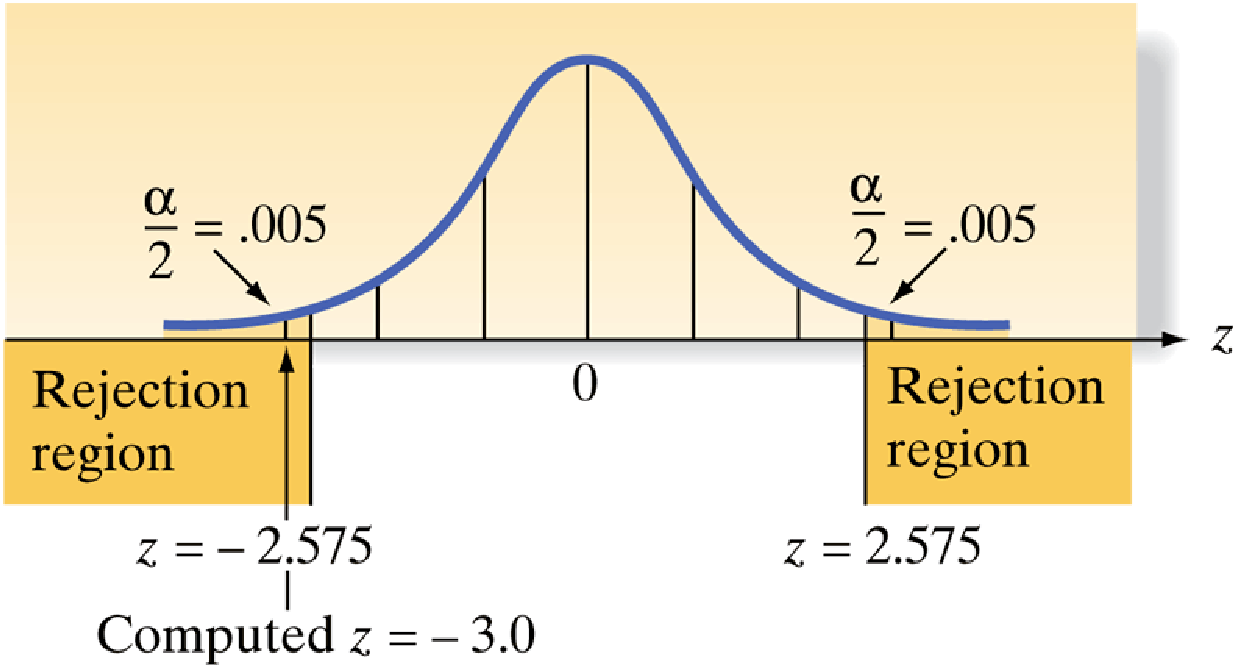
z = -3.0

($\alpha$ = 0.01)

The sample mean (1.05) is approx. three standard deviations below the null-hypothesized value of 1.2 in the sampling distribution of $\overline{X}$

# Reject $H_0$

- Z = -3.0 falls in the lower-tail rejection region, which consists of all values of z < -2.575



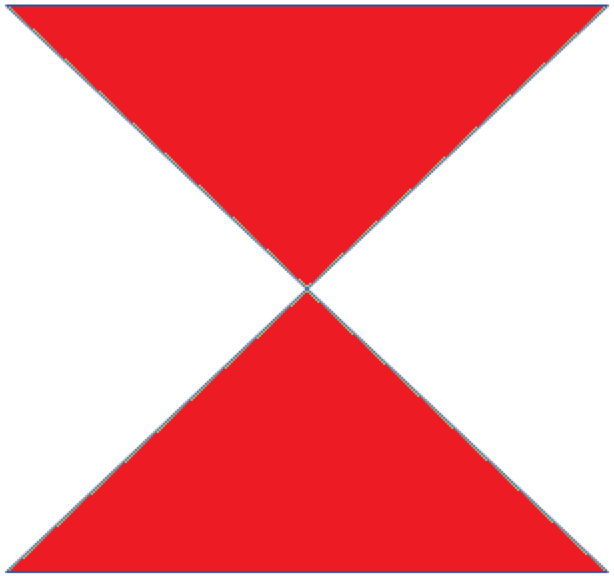

- Enough evidence to reject $H_0$

# Conclusion

- At the $\alpha = 0.01$ level of significance, the mean response time for drug-injected rats differs from the control mean of 1.2 seconds
- It appears that the rats receiving an injection have a mean response time that is not equal to 1.2 seconds

# Complete BC2 Quiz

- Available on Moodle



| | |
|---|---|
| Quiz Start Date | 11/10/2024 |
| End Date | 25/10/2024 |

(record any incorrect answers before End Date)

# More than a single sample….

## Student’s t-Test

# Objective

- Be able to compute and interpret test for difference between the means of two groups

# Student’s t test

- Student's t-test is used to compare the means of two samples
- Developed by William Gosset of Guinness's Brewery

Image source: http://en.wikipedia.org/wiki/William_Sealy_Gosset


Image source: http://tastingvideos.blogspot.ie/2011/11/welcome-to-ireland-with-pint-of.html

# Null hypothesis

- The statistical null hypothesis is that the means of the measurement variable are equal for the two categories.

# Put simply…

- A t-Test is used to compare two sample means to make an inference about a population

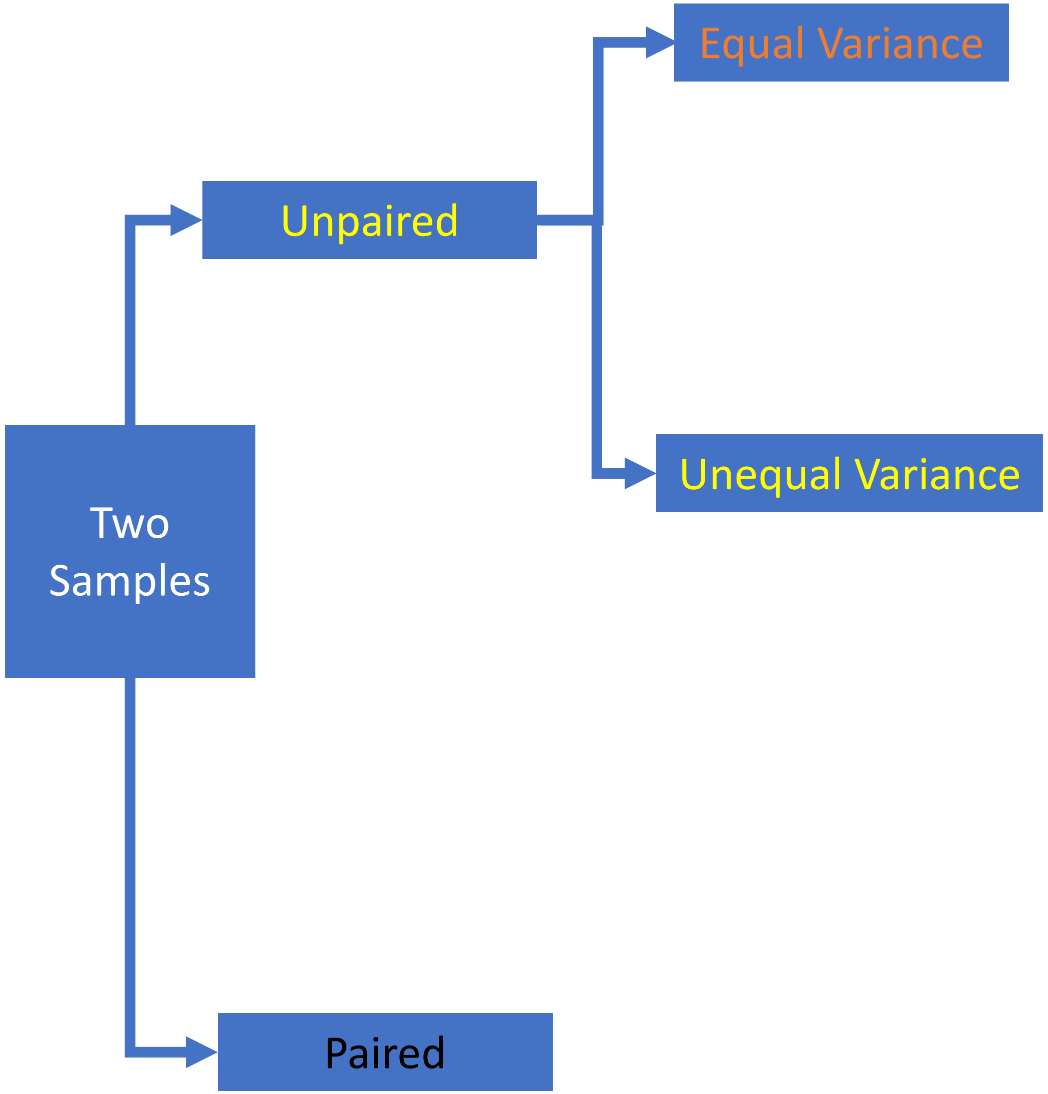
($H_0$) There ***isn't*** a difference between the populations

($H_1$) There ***is*** a difference between the populations

# Two Types of t-Test

- Unpaired (independent)
  - e.g. - evaluating the effect of a medical treatment, and we enrol 100 subjects into our study, then randomize 50 subjects to the treatment group and 50 subjects to the control group
  - Equal vs Unequal Variance
- Paired (dependent)
  - e.g. - subjects are tested prior to a treatment, say for high blood pressure, and the same subjects are tested again after treatment with a blood-pressure lowering medication

# Deciding on the correct t-Test



# t-Test for Unpaired (Independent) Samples

- First step – check are:
  - Variances Equal or Unequal ?

  Applies to **Unpaired** t-Test only

- Equal vs Unequal Variance
  - Slightly different formulas
  - Different calculation for degrees of freedom

- Check using F-test

# Degrees of Freedom

- Degrees of freedom (df) are defined as the number of scores in a sample that are free to vary

Visual explanation:

- There are five balloons: one blue, one red, one yellow, one pink, & one green
- If 5 students (n=5) are each to select one balloon only 4 will have a choice of colour (df=4)
- The last person will get whatever colour is left.

# Assumptions

- The different options of the *t*-test revolve around the assumption of equal variances or unequal variances

$$F = \frac{\text{larger sample variance}}{\text{smaller sample variance}}$$

- Values of F close to "1" indicate equal Variance

# Example (F Test)

- Two samples:

$$n_1 = 25$$
$$n_2 = 46$$
$$s_1 = 13.2$$
$$s_2 = 15.3$$

$$F = \frac{s_2^2}{s_1^2} = \frac{15.3^2}{13.2^2} = \frac{234.09}{174.24} = 1.34$$

# Degrees of Freedom

- Two samples:

$$n_1 = 25$$

$$n_2 = 46$$

- Degrees of Freedom = n-1

$$DF_1 = 25-1 = 24$$

$$DF_2 = 46-1 = 45$$

- F tables

# Data Summary

| | Sample 1 | Sample 2 |
|---|---|---|
| Count (n) | 25 | 46 |
| Variance ($s^2$) | 174.24 | 234.09 |
| Degrees of Freedom | 24 | 45 |

- F Distribution Table:
  - Column – sample with higher variance
  - Row – sample with lower variance

**Table of the F-Distribution**

Critical values for right-hand tail area equal to **0.05**

| **df1:** | **16** | **17** | **18** | **19** | **20** | **21** | **22** | **23** | **24** | **25** | **30** | **40** | **50** | **60** | **120** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **df2: 1** | 246 | 247 | 247 | 248 | 248 | 248 | 249 | 249 | 249 | 249 | 250 | 251 | 252 | 252 | 253 |
| **2** | 19.4 | 19.4 | 19.4 | 19.4 | 19.4 | 19.4 | 19.5 | 19.5 | 19.5 | 19.5 | 19.5 | 19.5 | 19.5 | 19.5 | 19.5 |
| **3** | 8.69 | 8.68 | 8.67 | 8.67 | 8.66 | 8.65 | 8.65 | 8.64 | 8.64 | 8.63 | 8.62 | 8.59 | 8.58 | 8.57 | 8.55 |
| **4** | 5.84 | 5.83 | 5.82 | 5.81 | 5.80 | 5.79 | 5.79 | 5.78 | 5.77 | 5.77 | 5.75 | 5.72 | 5.70 | 5.69 | 5.66 |
| **5** | 4.60 | 4.59 | 4.58 | 4.57 | 4.56 | 4.55 | 4.54 | 4.53 | 4.53 | 4.52 | 4.50 | 4.46 | 4.44 | 4.43 | 4.40 |
| **6** | 3.92 | 3.91 | 3.90 | 3.88 | 3.87 | 3.86 | 3.86 | 3.85 | 3.84 | 3.83 | 3.81 | 3.77 | 3.75 | 3.74 | 3.70 |
| **7** | 3.49 | 3.48 | 3.47 | 3.46 | 3.44 | 3.43 | 3.43 | 3.42 | 3.41 | 3.40 | 3.38 | 3.34 | 3.32 | 3.30 | 3.27 |
| **8** | 3.20 | 3.19 | 3.17 | 3.16 | 3.15 | 3.14 | 3.13 | 3.12 | 3.12 | 3.11 | 3.08 | 3.04 | 3.02 | 3.01 | 2.97 |
| **9** | 2.99 | 2.97 | 2.96 | 2.95 | 2.94 | 2.93 | 2.92 | 2.91 | 2.90 | 2.89 | 2.86 | 2.83 | 2.80 | 2.79 | 2.75 |
| **10** | 2.83 | 2.81 | 2.80 | 2.79 | 2.77 | 2.76 | 2.75 | 2.75 | 2.74 | 2.73 | 2.70 | 2.66 | 2.64 | 2.62 | 2.58 |
| **11** | 2.70 | 2.69 | 2.67 | 2.66 | 2.65 | 2.64 | 2.63 | 2.62 | 2.61 | 2.60 | 2.57 | 2.53 | 2.51 | 2.49 | 2.45 |
| **12** | 2.60 | 2.58 | 2.57 | 2.56 | 2.54 | 2.53 | 2.52 | 2.51 | 2.51 | 2.50 | 2.47 | 2.43 | 2.40 | 2.38 | 2.34 |
| **13** | 2.51 | 2.50 | 2.48 | 2.47 | 2.46 | 2.45 | 2.44 | 2.43 | 2.42 | 2.41 | 2.38 | 2.34 | 2.31 | 2.30 | 2.25 |
| **14** | 2.44 | 2.43 | 2.41 | 2.40 | 2.39 | 2.38 | 2.37 | 2.36 | 2.35 | 2.34 | 2.31 | 2.27 | 2.24 | 2.22 | 2.18 |
| **15** | 2.38 | 2.37 | 2.35 | 2.34 | 2.33 | 2.32 | 2.31 | 2.30 | 2.29 | 2.28 | 2.25 | 2.20 | 2.18 | 2.16 | 2.11 |
| **16** | 2.33 | 2.32 | 2.30 | 2.29 | 2.28 | 2.26 | 2.25 | 2.24 | 2.24 | 2.23 | 2.19 | 2.15 | 2.12 | 2.11 | 2.06 |
| **17** | 2.29 | 2.27 | 2.26 | 2.24 | 2.23 | 2.22 | 2.21 | 2.20 | 2.19 | 2.18 | 2.15 | 2.10 | 2.08 | 2.06 | 2.01 |
| **18** | 2.25 | 2.23 | 2.22 | 2.20 | 2.19 | 2.18 | 2.17 | 2.16 | 2.15 | 2.14 | 2.11 | 2.06 | 2.04 | 2.02 | 1.97 |
| **19** | 2.21 | 2.20 | 2.18 | 2.17 | 2.16 | 2.14 | 2.13 | 2.12 | 2.11 | 2.11 | 2.07 | 2.03 | 2.00 | 1.98 | 1.93 |
| **20** | 2.18 | 2.17 | 2.15 | 2.14 | 2.12 | 2.11 | 2.10 | 2.09 | 2.08 | 2.07 | 2.04 | 1.99 | 1.97 | 1.95 | 1.90 |
| **21** | 2.16 | 2.14 | 2.12 | 2.11 | 2.10 | 2.08 | 2.07 | 2.06 | 2.05 | 2.05 | 2.01 | 1.96 | 1.94 | 1.92 | 1.87 |
| **22** | 2.13 | 2.11 | 2.10 | 2.08 | 2.07 | 2.06 | 2.05 | 2.04 | 2.03 | 2.02 | 1.98 | 1.94 | 1.91 | 1.89 | 1.84 |
| **23** | 2.11 | 2.09 | 2.08 | 2.06 | 2.05 | 2.04 | 2.02 | 2.01 | 2.01 | 2.00 | 1.96 | 1.91 | 1.88 | 1.86 | 1.81 |
| **24** | 2.09 | 2.07 | 2.05 | 2.04 | 2.03 | 2.01 | 2.00 | 1.99 | 1.98 | 1.97 | 1.94 | 1.89 | 1.86 | 1.84 | 1.79 |
| **25** | 2.07 | 2.05 | 2.04 | 2.02 | 2.01 | 2.00 | 1.98 | 1.97 | 1.96 | 1.96 | 1.92 | 1.87 | 1.84 | 1.82 | 1.77 |
| **30** | 1.99 | 1.98 | 1.96 | 1.95 | 1.93 | 1.92 | 1.91 | 1.90 | 1.89 | 1.88 | 1.84 | 1.79 | 1.76 | 1.74 | 1.68 |
| **40** | 1.90 | 1.89 | 1.87 | 1.85 | 1.84 | 1.83 | 1.81 | 1.80 | 1.79 | 1.78 | 1.74 | 1.69 | 1.66 | 1.64 | 1.58 |
| **60** | 1.82 | 1.80 | 1.78 | 1.76 | 1.75 | 1.73 | 1.72 | 1.71 | 1.70 | 1.69 | 1.65 | 1.59 | 1.56 | 1.53 | 1.47 |
| **120** | 1.73 | 1.71 | 1.69 | 1.67 | 1.66 | 1.64 | 1.63 | 1.62 | 1.61 | 1.60 | 1.55 | 1.50 | 1.46 | 1.43 | 1.35 |

# Table of the F Distribution (p27)

$F_{stat} = 1.34$

$DF_1 = 24$

$DF_2 = 45$

$\alpha = 0.05$

$F_{crit} = 1.89$

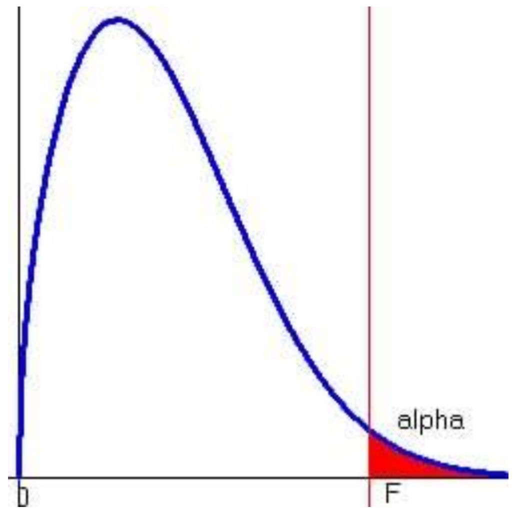
**Table of the F-Distribution**

Critical values for right-hand tail area equal to **0.05**

| **df1:** | **16** | **17** | **18** | **19** | **20** | **21** | **22** | **23** | **24** | **25** | **30** | **40** | **50** | **60** | **120** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **df2: 1** | 246 | 247 | 247 | 248 | 248 | 248 | 249 | 249 | 249 | 249 | 250 | 251 | 252 | 252 | 253 |
| **2** | 19.4 | 19.4 | 19.4 | 19.4 | 19.4 | 19.4 | 19.5 | 19.5 | 19.5 | 19.5 | 19.5 | 19.5 | 19.5 | 19.5 | 19.5 |
| **3** | 8.69 | 8.68 | 8.67 | 8.67 | 8.66 | 8.65 | 8.65 | 8.64 | 8.64 | 8.63 | 8.62 | 8.59 | 8.58 | 8.57 | 8.55 |
| **4** | 5.84 | 5.83 | 5.82 | 5.81 | 5.80 | 5.79 | 5.79 | 5.78 | 5.77 | 5.77 | 5.75 | 5.72 | 5.70 | 5.69 | 5.66 |
| **5** | 4.60 | 4.59 | 4.58 | 4.57 | 4.56 | 4.55 | 4.54 | 4.53 | 4.53 | 4.52 | 4.50 | 4.46 | 4.44 | 4.43 | 4.40 |
| **6** | 3.92 | 3.91 | 3.90 | 3.88 | 3.87 | 3.86 | 3.86 | 3.85 | 3.84 | 3.83 | 3.81 | 3.77 | 3.75 | 3.74 | 3.70 |
| **7** | 3.49 | 3.48 | 3.47 | 3.46 | 3.44 | 3.43 | 3.43 | 3.42 | 3.41 | 3.40 | 3.38 | 3.34 | 3.32 | 3.30 | 3.27 |
| **8** | 3.20 | 3.19 | 3.17 | 3.16 | 3.15 | 3.14 | 3.13 | 3.12 | 3.12 | 3.11 | 3.08 | 3.04 | 3.02 | 3.01 | 2.97 |
| **9** | 2.99 | 2.97 | 2.96 | 2.95 | 2.94 | 2.93 | 2.92 | 2.91 | 2.90 | 2.89 | 2.86 | 2.83 | 2.80 | 2.79 | 2.75 |
| **10** | 2.83 | 2.81 | 2.80 | 2.79 | 2.77 | 2.76 | 2.75 | 2.75 | 2.74 | 2.73 | 2.70 | 2.66 | 2.64 | 2.62 | 2.58 |
| **11** | 2.70 | 2.69 | 2.67 | 2.66 | 2.65 | 2.64 | 2.63 | 2.62 | 2.61 | 2.60 | 2.57 | 2.53 | 2.51 | 2.49 | 2.45 |
| **12** | 2.60 | 2.58 | 2.57 | 2.56 | 2.54 | 2.53 | 2.52 | 2.51 | 2.51 | 2.50 | 2.47 | 2.43 | 2.40 | 2.38 | 2.34 |
| **13** | 2.51 | 2.50 | 2.48 | 2.47 | 2.46 | 2.45 | 2.44 | 2.43 | 2.42 | 2.41 | 2.38 | 2.34 | 2.31 | 2.30 | 2.25 |
| **14** | 2.44 | 2.43 | 2.41 | 2.40 | 2.39 | 2.38 | 2.37 | 2.36 | 2.35 | 2.34 | 2.31 | 2.27 | 2.24 | 2.22 | 2.18 |
| **15** | 2.38 | 2.37 | 2.35 | 2.34 | 2.33 | 2.32 | 2.31 | 2.30 | 2.29 | 2.28 | 2.25 | 2.20 | 2.18 | 2.16 | 2.11 |
| **16** | 2.33 | 2.32 | 2.30 | 2.29 | 2.28 | 2.26 | 2.25 | 2.24 | 2.24 | 2.23 | 2.19 | 2.15 | 2.12 | 2.11 | 2.06 |
| **17** | 2.29 | 2.27 | 2.26 | 2.24 | 2.23 | 2.22 | 2.21 | 2.20 | 2.19 | 2.18 | 2.15 | 2.10 | 2.08 | 2.06 | 2.01 |
| **18** | 2.25 | 2.23 | 2.22 | 2.20 | 2.19 | 2.18 | 2.17 | 2.16 | 2.15 | 2.14 | 2.11 | 2.06 | 2.04 | 2.02 | 1.97 |
| **19** | 2.21 | 2.20 | 2.18 | 2.17 | 2.16 | 2.14 | 2.13 | 2.12 | 2.11 | 2.11 | 2.07 | 2.03 | 2.00 | 1.98 | 1.93 |
| **20** | 2.18 | 2.17 | 2.15 | 2.14 | 2.12 | 2.11 | 2.10 | 2.09 | 2.08 | 2.07 | 2.04 | 1.99 | 1.97 | 1.95 | 1.90 |
| **21** | 2.16 | 2.14 | 2.12 | 2.11 | 2.10 | 2.08 | 2.07 | 2.06 | 2.05 | 2.05 | 2.01 | 1.96 | 1.94 | 1.92 | 1.87 |
| **22** | 2.13 | 2.11 | 2.10 | 2.08 | 2.07 | 2.06 | 2.05 | 2.04 | 2.03 | 2.02 | 1.98 | 1.94 | 1.91 | 1.89 | 1.84 |
| **23** | 2.11 | 2.09 | 2.08 | 2.06 | 2.05 | 2.04 | 2.02 | 2.01 | 2.01 | 2.00 | 1.96 | 1.91 | 1.88 | 1.86 | 1.81 |
| **24** | 2.09 | 2.07 | 2.05 | 2.04 | 2.03 | 2.01 | 2.00 | 1.99 | 1.98 | 1.97 | 1.94 | 1.89 | 1.86 | 1.84 | 1.79 |
| **25** | 2.07 | 2.05 | 2.04 | 2.02 | 2.01 | 2.00 | 1.98 | 1.97 | 1.96 | 1.96 | 1.92 | 1.87 | 1.84 | 1.82 | 1.77 |
| **30** | 1.99 | 1.98 | 1.96 | 1.95 | 1.93 | 1.92 | 1.91 | 1.90 | 1.89 | 1.88 | 1.84 | 1.79 | 1.76 | 1.74 | 1.68 |
| **40** | 1.90 | 1.89 | 1.87 | 1.85 | 1.84 | 1.83 | 1.81 | 1.80 | 1.79 | 1.78 | 1.74 | 1.69 | 1.66 | 1.64 | 1.58 |
| **60** | 1.82 | 1.80 | 1.78 | 1.76 | 1.75 | 1.73 | 1.72 | 1.71 | 1.70 | 1.69 | 1.65 | 1.59 | 1.56 | 1.53 | 1.47 |
| **120** | 1.73 | 1.71 | 1.69 | 1.67 | 1.66 | 1.64 | 1.63 | 1.62 | 1.61 | 1.60 | 1.55 | 1.50 | 1.46 | 1.43 | 1.35 |

# Equal or Unequal?

- Decision:
  - $F_{stat} = 1.34$
  - $DF_1 = 24$
  - $DF_2 = 45$
  - $\alpha = 0.05$
  - $F_{crit} = 1.89$




- $F_{stat} < F_{crit}$
- Therefore sample variances are “equal”

# Unpaired t-Test formulas

- Equal variance:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}\right)}}$$

- Unequal variance:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

**Note differences!**

# Unpaired t-Test (Equal Variance)

- Formula:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}\right)}}$$

- Pooled Variance:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

- Degrees of Freedom:

$$df = n_1 + n_2 - 2$$

# Unpaired t-Test (Unequal Variance)

- Formula:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

- Degrees of Freedom

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$$

# So Far....

- Two Samples
  - Unpaired
    - $F = \dfrac{\text{larger sample variance}}{\text{smaller sample variance}}$
    - Equal Variance
      - $$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}\right)}}$$
      - $$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$
      - $$df = n_1 + n_2 - 2$$
    - Unequal Variance
      - $$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$
      - $$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$$
  - Paired

# Example
# t-Test for Unpaired (Independent)

- Examine the differences between two groups of Alzheimer's patients

  - Group 1
    - Taught using visuals
  - Group 2
    - Taught using visuals and intense verbal rehearsal

- Participants are being tested only once

# Hypotheses

- The Null Hypothesis is:

  $H_0$: $\mu_1 = \mu_2$ (No difference between populations)

- The Alternate Hypothesis is:

  $H_1$: $\bar{\boldsymbol{X}}_1 \neq \bar{\boldsymbol{X}}_2$ (There is a difference between samples)

# Alzheimer’s Data

- Memory scores
- Two groups
- 30 participants per group

| Group 1 | Group 2 |
|---|---|
| 7 | 5 |
| 3 | 4 |
| 3 | 4 |
| 2 | 5 |
| 3 | 5 |
| 8 | 7 |
| 8 | 8 |
| 5 | 8 |
| 8 | 9 |
| 5 | 8 |
| 5 | 3 |
| 4 | 2 |
| 6 | 5 |
| 10 | 4 |
| 10 | 4 |
| 5 | 6 |
| 1 | 7 |
| 1 | 7 |
| 4 | 5 |
| 3 | 6 |
| 5 | 4 |
| 7 | 3 |
| 1 | 2 |
| 9 | 7 |
| 2 | 6 |
| 5 | 2 |
| 2 | 8 |
| 12 | 9 |
| 15 | 7 |
| 4 | 6 |

# Equal Variance?

- First – test for equality of variance:

$$F_{stat} = \frac{s_1^2}{s_2^2} = \frac{11.70}{4.36} = 2.75$$

$DF_1 = 30-1 = 29$

$DF_2 = 30-1 = 29$

$F_{crit} = 1.84$

- $F_{stat} > F_{crit}$
- Therefore sample variances are "unequal"

# Formula for Unpaired (Independent) t-Test (Unequal Variance)

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

- Where
  - $\bar{X}_1$ is the mean for sample Group 1
  - $\bar{X}_2$ is the mean for sample Group 2
  - $\mu_1$ is the mean for Population 1
  - $\mu_2$ is the mean for Population 2
  - $s_1^2$ is the variance of Group 1
  - $s_2^2$ is the variance of Group 2
  - $n_1$ is the number of participants in Group 1
  - $n_2$ is the number of participants in Group 2

# Amended formula

- Assume population means are the same
- Therefore can omit $(\boldsymbol{\mu_1} - \boldsymbol{\mu_2})$ from formula

$$\boldsymbol{t} = \frac{(\bar{\boldsymbol{X}}_1 - \bar{\boldsymbol{X}}_2)}{\sqrt{\left(\frac{\boldsymbol{s}_1^2}{\boldsymbol{n}_1} + \frac{\boldsymbol{s}_2^2}{\boldsymbol{n}_2}\right)}}$$

→ **Difference between means**

→ **Variances**

→ **Sample sizes**

# Alzheimer’s Data

- Memory scores
- Two groups
- 30 participants per group

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

| Group 1 | Group 2 |
|---|---|
| 7 | 5 |
| 3 | 4 |
| 3 | 4 |
| 2 | 5 |
| 3 | 5 |
| 8 | 7 |
| 8 | 8 |
| 5 | 8 |
| 8 | 9 |
| 5 | 8 |
| 5 | 3 |
| 4 | 2 |
| 6 | 5 |
| 10 | 4 |
| 10 | 4 |
| 5 | 6 |
| 1 | 7 |
| 1 | 7 |
| 4 | 5 |
| 3 | 6 |
| 5 | 4 |
| 7 | 3 |
| 1 | 2 |
| 9 | 7 |
| 2 | 6 |
| 5 | 2 |
| 2 | 8 |
| 12 | 9 |
| 15 | 7 |
| 4 | 6 |

# Substitute values

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

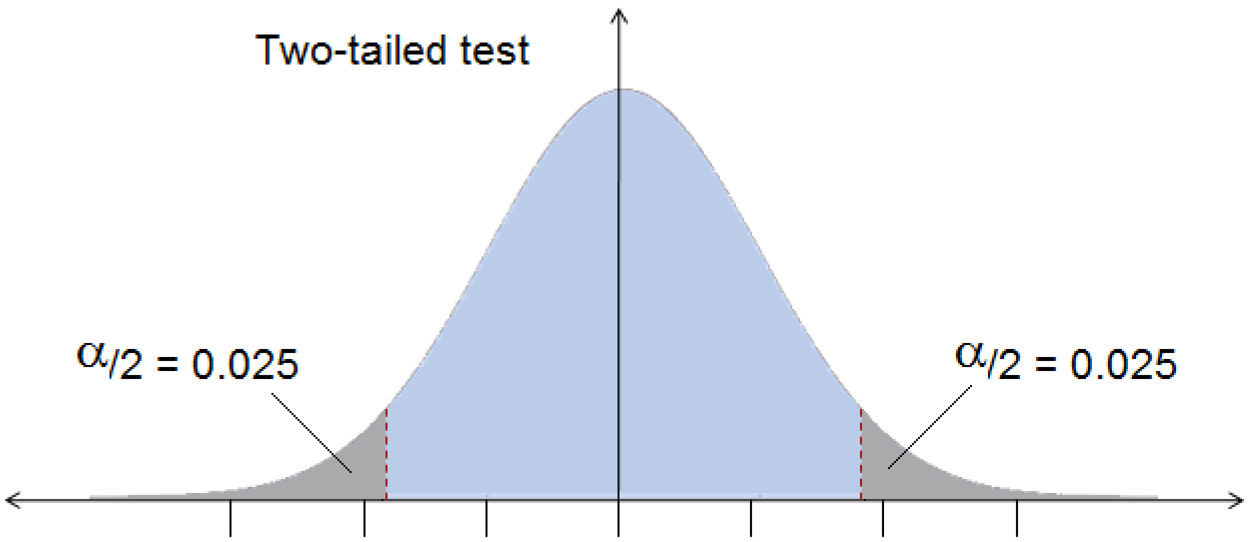
$$t = \frac{5.43 - 5.53}{\sqrt{\left(\frac{11.70}{30} + \frac{4.26}{30}\right)}}$$

$$t = -0.137$$

Now What?

# Determine α value

- $\alpha = 0.05$



# Degrees of Freedom

- In this sample there are two groups
  - $n_1 = 30$, $n_2 = 30$
  - $\boldsymbol{s}_1^2 = 11.7$, $\boldsymbol{s}_2^2 = 4.26$

- In this test, the Degrees of Freedom are:

$$\boldsymbol{df} = \frac{\left(\frac{\boldsymbol{s}_1^2}{\boldsymbol{n}_1} + \frac{\boldsymbol{s}_2^2}{\boldsymbol{n}_2}\right)^2}{\frac{\left(\boldsymbol{s}_1^2/\boldsymbol{n}_1\right)^2}{\boldsymbol{n}_1 - 1} + \frac{\left(\boldsymbol{s}_2^2/\boldsymbol{n}_2\right)^2}{\boldsymbol{n}_2 - 1}} \quad = 47.6$$

# Report result of t-Test

- Results:
  - t = -0.137
  - α = 0.05
  - DF = 48

- Report as:

$$t_{(48)} = -0.137, \, p < 0.05$$

t = -0.137

α = 0.05

DF = 48

## *t* Table

| 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 |
| | | | | | | 1.833 | 2.262 | 2.821 | 3.250 | 4.297 | 4.781 |
| | | | | | | 1.812 | 2.228 | 2.764 | 3.169 | 4.144 | 4.587 |
| | | | | | | 1.796 | 2.201 | 2.718 | 3.106 | 4.025 | 4.437 |
| | | | | | | 1.782 | 2.179 | 2.681 | 3.055 | 3.930 | 4.318 |
| | | | | | | 1.771 | 2.160 | 2.650 | 3.012 | 3.852 | 4.221 |
| | | | | | | 1.761 | 2.145 | 2.624 | 2.977 | 3.787 | 4.140 |
| | | | | | | 1.753 | 2.131 | 2.602 | 2.947 | 3.733 | 4.073 |
| | | | | | | 1.746 | 2.120 | 2.583 | 2.921 | 3.686 | 4.015 |
| | | | | | | 1.740 | 2.110 | 2.567 | 2.898 | 3.646 | 3.965 |
| | | | | | | 1.734 | 2.101 | 2.552 | 2.878 | 3.610 | 3.922 |
| | | | | | | 1.729 | 2.093 | 2.539 | 2.861 | 3.579 | 3.883 |
| | | | | | | 1.725 | 2.086 | 2.528 | 2.845 | 3.552 | 3.850 |
| | | | | | | 1.721 | 2.080 | 2.518 | 2.831 | 3.527 | 3.819 |
| | | | | | | 1.717 | 2.074 | 2.508 | 2.819 | 3.505 | 3.792 |
| | | | | | | 1.714 | 2.069 | 2.500 | 2.807 | 3.485 | 3.768 |
| | | | | | | 1.711 | 2.064 | 2.492 | 2.797 | 3.467 | 3.745 |
| | | | | | | 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** | | | | | | | | | | |

t-value (-0.137) is less than critical value (2.021) therefore *t* is not significant and you have not found a difference

# SPSS Result

**T-Test**

[DataSet3]

**Group Statistics**

| | Group | N | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| MemoryTest | 1 | 30 | 5.43 | 3.421 | .625 |
| | 2 | 30 | 5.53 | 2.063 | .377 |

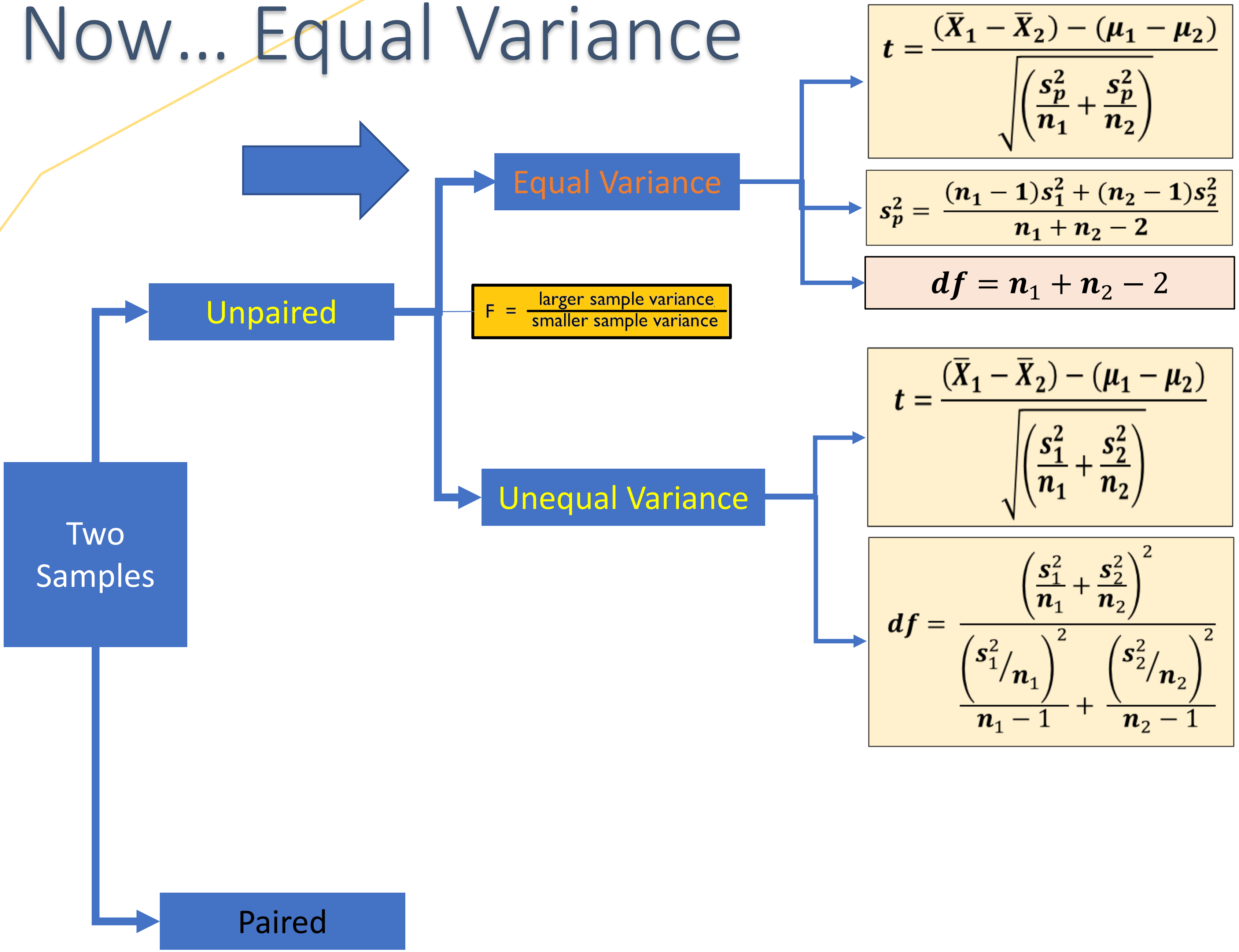
**Independent Samples Test**

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | 95% Confidence Interval of the Difference | |
| | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | Lower | Upper |
| MemoryTest | Equal variances assumed | 4.994 | .029 | -.137 | 58 | .891 | -.100 | .729 | -1.560 | 1.360 |
| | Equal variances not assumed | | | -.137 | 47.635 | .892 | -.100 | .729 | -1.567 | 1.367 |

# Excel Result

| | | |
|---|---|---|
| t-Test: Two-Sample Assuming Unequal Variances | | |
| | | |
| | *Group 1* | *Group 2* |
| Mean | 5.433333 | 5.533333 |
| Variance | 11.7023 | 4.257471 |
| Observations | 30 | 30 |
| Hypothesized Mean Difference | 0 | |
| df | 48 | |
| t Stat | -0.1371 | |
| P(T<=t) one-tail | 0.445761 | |
| t Critical one-tail | 1.677224 | |
| P(T<=t) two-tail | 0.891523 | |
| t Critical two-tail | 2.010635 | |

# Now… Equal Variance



# Sample Data (Hay Fever/Driving)

- To investigate the effect of a new hay fever drug on driving skills, a researcher studies 24 individuals with hay fever:
  - 12 who have been taking the drug
  - 12 who have not

- All participants then entered a simulator and were given a driving test which assigned a score to each driver as summarized

| Control | Drug |
|---|---|
| 23 | 16 |
| 15 | 21 |
| 16 | 16 |
| 25 | 11 |
| 20 | 24 |
| 17 | 21 |
| 18 | 18 |
| 14 | 15 |
| 12 | 19 |
| 19 | 22 |
| 21 | 13 |
| 22 | 24 |

# Example (F Test)

- Two samples:

$$n_1 = 12$$
$$n_2 = 12$$
$$s_1^2 = 15.18$$
$$s_2^2 = 17.88$$

$$F = \frac{s_2^2}{s_1^2} = \frac{17.88}{15.18} = 1.18$$

$$F_{stat}(1.18) < F_{crit}(2.82)$$

Therefore variances are “equal”

# Unpaired t-Test (Equal Variance)

- Formula:

$$\boldsymbol{t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}\right)}}}$$

- Pooled Variance:

$$\boldsymbol{s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Degrees of Freedom:

$$df = n_1 + n_2 - 2$$

# Pooled Variance

- In this sample there are two groups
  - $n_1 = 12$, $n_2 = 12$
  - $s_1^2 = 15.18$, $s_2^2 = 17.88$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$= 16.56$

# Formula for Unpaired (Independent) t-Test (Equal Variance)

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}\right)}}$$

- Where
  - $\bar{X}_1$ is the mean for sample Group 1
  - $\bar{X}_2$ is the mean for sample Group 2
  - $\mu_2$ is the mean for Population 1
  - $\mu_1$ is the mean for Population 2
  - $s_p^2$ is the pooled variance of Group 1
  - $s_p^2$ is the pooled variance of Group 2
  - $n_1$ is the number of participants in Group 1
  - $n_2$ is the number of participants in Group 2

# Calculate t statistic

- Formula:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}\right)}}$$

| Control | Drug |
|---|---|
| 23 | 16 |
| 15 | 21 |
| 16 | 16 |
| 25 | 11 |
| 20 | 24 |
| 17 | 21 |
| 18 | 18 |
| 14 | 15 |
| 12 | 19 |
| 19 | 22 |
| 21 | 13 |
| 22 | 24 |

t = 0.10

- Report as:

$t_{(22)} = 0.10$, $p < 0.05$

t-value (0.10) is less than critical value (2.074) therefore *t* is not significant, and you have not found a difference

$t = 0.10$

$\alpha = 0.05$

DF = 22

## *t* Table

| 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** | | | | | |

# Two Types of t-Test

- Unpaired (independent)
  - EG - evaluating the effect of a medical treatment, and we enrol 100 subjects into our study, then randomize 50 subjects to the treatment group and 50 subjects to the control group

- Paired (dependent)
  - EG - subjects are tested prior to a treatment, say for high blood pressure, and the same subjects are tested again after treatment with a blood-pressure lowering medication

# t-Test for Paired (Dependent) Samples

- Examine the difference between students' scores on a pre-test and on a post-test following a training intervention
- Participants are being tested more than once
- There are two groups of scores
- The appropriate test statistic is a t-Test for ***paired (dependent)*** means

# Hypotheses

- The Null Hypothesis is:

$$H_0: \mu_{\text{post-test}} = \mu_{\text{pre-test}}$$

- The Alternate Hypothesis is:

$$H_1: \bar{\boldsymbol{X}}_{\text{post-test}} \neq \bar{\boldsymbol{X}}_{\text{pre-test}}$$

# Formula for Paired (dependent) t-Test

$$\boxed{t = \frac{\Sigma D}{\sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n - 1}}}}$$

- Where
  - $\Sigma D$ is the sum of all the differences between groups of scores
  - $\Sigma D^2$ is the sum of the differences squared between groups of scores
  - $n$ is the number of (paired) observations

- Degrees of Freedom: $\boxed{\boldsymbol{df = n - 1}}$

# Choice of formula for t Test

Two Samples

- Unpaired
  - Equal Variance

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}\right)}}$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$df = n_1 + n_2 - 2$$

  - Unequal Variance

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$$

- Paired

$$t = \frac{\Sigma D}{\sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n - 1}}}$$

$$df = n - 1$$

# Data

- Test scores
- 25 students

$$t = \frac{\Sigma D}{\sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n-1}}}$$

| Pretest | Posttest |
|---|---|
| 3 | 7 |
| 5 | 8 |
| 4 | 6 |
| 6 | 7 |
| 5 | 8 |
| 5 | 9 |
| 4 | 6 |
| 5 | 6 |
| 3 | 7 |
| 6 | 8 |
| 7 | 8 |
| 8 | 7 |
| 7 | 9 |
| 6 | 10 |
| 7 | 9 |
| 8 | 9 |
| 8 | 8 |
| 9 | 8 |
| 9 | 4 |
| 8 | 4 |
| 7 | 5 |
| 7 | 6 |
| 6 | 9 |
| 7 | 8 |
| 8 | 12 |

# Calculations ($n$ = 25)

| | Pre-test | Post-test | Difference (D) | $D^2$ |
|---|---|---|---|---|
| | 3 | 7 | 4 | 16 |
| | 5 | 8 | 3 | 9 |
| | 4 | 6 | 2 | 4 |
| | 6 | 7 | 1 | 1 |
| | 5 | 8 | 3 | 9 |
| | 5 | 9 | 4 | 16 |
| | 4 | 6 | 2 | 4 |
| | 5 | 6 | 1 | 1 |
| | 3 | 7 | 4 | 16 |
| | 6 | 8 | 2 | 4 |
| | 7 | 8 | 1 | 1 |
| | 8 | 7 | -1 | 1 |
| | 7 | 9 | 2 | 4 |
| | 6 | 10 | 4 | 16 |
| | 7 | 9 | 2 | 4 |
| | 8 | 9 | 1 | 1 |
| | 8 | 8 | 0 | 0 |
| | 9 | 8 | -1 | 1 |
| | 9 | 4 | -5 | 25 |
| | 8 | 4 | -4 | 16 |
| | 7 | 5 | -2 | 4 |
| | 7 | 6 | -1 | 1 |
| | 6 | 9 | 3 | 9 |
| | 7 | 8 | 1 | 1 |
| | 8 | 12 | 4 | 16 |
| **Sum** | **158** | **188** | **30** | **180** |

$\Sigma D^2$

$\Sigma D$

# Plug values into formula

$$t = \frac{\Sigma D}{\sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n \ - 1}}}$$

$$t = \frac{30}{\sqrt{\frac{(25\ x\ 180) - 30^2}{25 \ - 1}}}$$

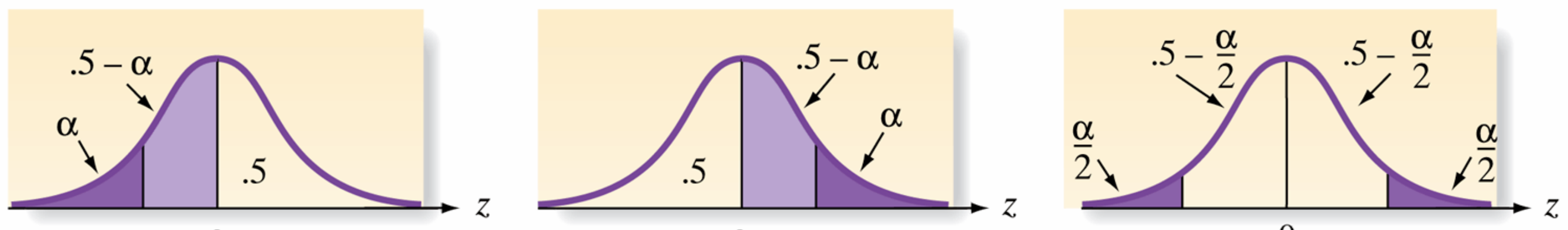
$$t = \frac{30}{\sqrt{150}} = 2.45$$

Now What?

# Determine α value

$\alpha = 0.05$

- Rejection regions corresponding to one- and two tailed tests




a. Form of $H_a$: <

b. Form of $H_a$: >

c. Form of $H_a$: ≠

(a) and (b) - one-tailed rejection regions for lower and upper tailed tests

(c) – two tailed rejection region

# Degree of Freedom

- In this sample there are 25 pairs
  - $n = 25$

- In this test, the Degrees of Freedom are:
  - $n - 1$
  - $25 - 1 = 24$

# Report result of t-Test

- Results:
  - t = 2.45
  - α = 0.05
  - DF = 24
- Report as:

$$t_{(24)} = 2.45,\ p < 0.05$$

t = 2.45

α = 0.05

DF = 24

## *t* Table

| 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 |
| [illegible] | [illegible] | [illegible] | 54 | 1.055 | 1.310 | 1.697 | 2.042 | 2.457 | 2.750 | 3.385 | 3.646 |
| [illegible] | [illegible] | [illegible] | 51 | 1.050 | 1.303 | 1.684 | 2.021 | 2.423 | 2.704 | 3.307 | 3.551 |
| [illegible] | [illegible] | [illegible] | 48 | 1.045 | 1.296 | 1.671 | 2.000 | 2.390 | 2.660 | 3.232 | 3.460 |
| [illegible] | [illegible] | [illegible] | 46 | 1.043 | 1.292 | 1.664 | 1.990 | 2.374 | 2.639 | 3.195 | 3.416 |
| [illegible] | [illegible] | [illegible] | 45 | 1.042 | 1.290 | 1.660 | 1.984 | 2.364 | 2.626 | 3.174 | 3.390 |
| [illegible] | [illegible] | [illegible] | 42 | 1.037 | 1.282 | 1.646 | 1.962 | 2.330 | 2.581 | 3.098 | 3.300 |
| [illegible] | [illegible] | [illegible] | 42 | 1.036 | 1.282 | 1.645 | 1.960 | 2.326 | 2.576 | 3.090 | 3.291 |
| | [illegible] | [illegible] | % | 70% | 80% | 90% | 95% | 98% | 99% | 99.8% | 99.9% |
| | | | | | **Confidence Level** | | | | | | |

t-value (2.45) is greater than critical value (2.064) therefore *t* is significant and you have found a difference

# SPSS Result

**T-Test**

[DataSet1]

**Paired Samples Statistics**

| | | Mean | N | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| Pair 1 | Pretest | 6.32 | 25 | 1.725 | .345 |
| | Posttest | 7.52 | 25 | 1.828 | .366 |

**Paired Samples Test**

| | | Paired Differences | | | | | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | 95% Confidence Interval of the Difference | | | | |
| | | Mean | Std. Deviation | Std. Error Mean | Lower | Upper | | | |
| Pair 1 | Pretest - Posttest | -1.200 | 2.449 | .490 | -2.211 | -.189 | -2.449 | 24 | .022 |

**t Table**

| 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 | | | | | | | | | | |

# Reading a t Table

- Once you have chosen your significance level, *p*, calculated a value for *t*, and worked out how many degrees of freedom you have, you can find the entry in the t-table that you need as follows:
  - Look down the column that corresponds to your chosen value for *p*
  - Find the row that corresponds to your degrees of freedom, and where they meet, you will find the value you need.
- This value is called the **critical value**.

# Using the Critical Value

- The final thing to do is compare the critical value with your value of *t:*
  - If your t-value is greater than or equal to this value, then *t* is significant and you have found a difference
  - If your t-value is less than this value is then *t* is not significant.

$p < 0.05$, t = 3.143, *df* = 4, tails = 1

What is critical value?

**t Table**

| 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 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 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 | [illegible] | [illegible] | [illegible] | [illegible] |
| 10 | 0.000 | 0.700 | 0.879 | 1.093 | 1.372 | 1.812 | 2.228 | [illegible] | [illegible] | [illegible] | [illegible] |
| 11 | 0.000 | 0.697 | 0.876 | 1.088 | 1.363 | 1.796 | 2.201 | [illegible] | [illegible] | [illegible] | [illegible] |
| 12 | 0.000 | 0.695 | 0.873 | 1.083 | 1.356 | 1.782 | 2.179 | [illegible] | [illegible] | [illegible] | [illegible] |
| 13 | 0.000 | 0.694 | 0.870 | 1.079 | 1.350 | 1.771 | 2.160 | [illegible] | [illegible] | [illegible] | [illegible] |
| 14 | 0.000 | 0.692 | 0.868 | 1.076 | 1.345 | 1.761 | 2.145 | [illegible] | [illegible] | [illegible] | [illegible] |
| 15 | 0.000 | 0.691 | 0.866 | 1.074 | 1.341 | 1.753 | 2.131 | [illegible] | [illegible] | [illegible] | [illegible] |
| 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**

Significant?

t-value is greater than critical value, therefore *t* is significant, and you have found a difference

# Summary

- Two Samples
  - Unpaired
    - $F = \dfrac{\text{larger sample variance}}{\text{smaller sample variance}}$
    - Equal Variance
      - $$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}\right)}}$$
      - $$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$
      - $$df = n_1 + n_2 - 2$$
    - Unequal Variance
      - $$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$
      - $$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$$
  - Paired
    - $$t = \frac{\Sigma D}{\sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n - 1}}}$$
    - $$df = n - 1$$

# More than two samples….

Analysis of Variance (ANOVA)

# Comparing Means

- t-Test
  - Compare means of two samples
- ANOVA
  - Analysis of Variance
  - Compare means of more than two samples

# Sample Problem

| Consumer | Brand A | Brand B | Brand C |
|---|---|---|---|
| 1 | 2 | 15 | 6 |
| 2 | 11 | 14 | 5 |
| 3 | 1 | 7 | 10 |
| 4 | 13 | 10 | 9 |
| 5 | 3 | 8 | 4 |

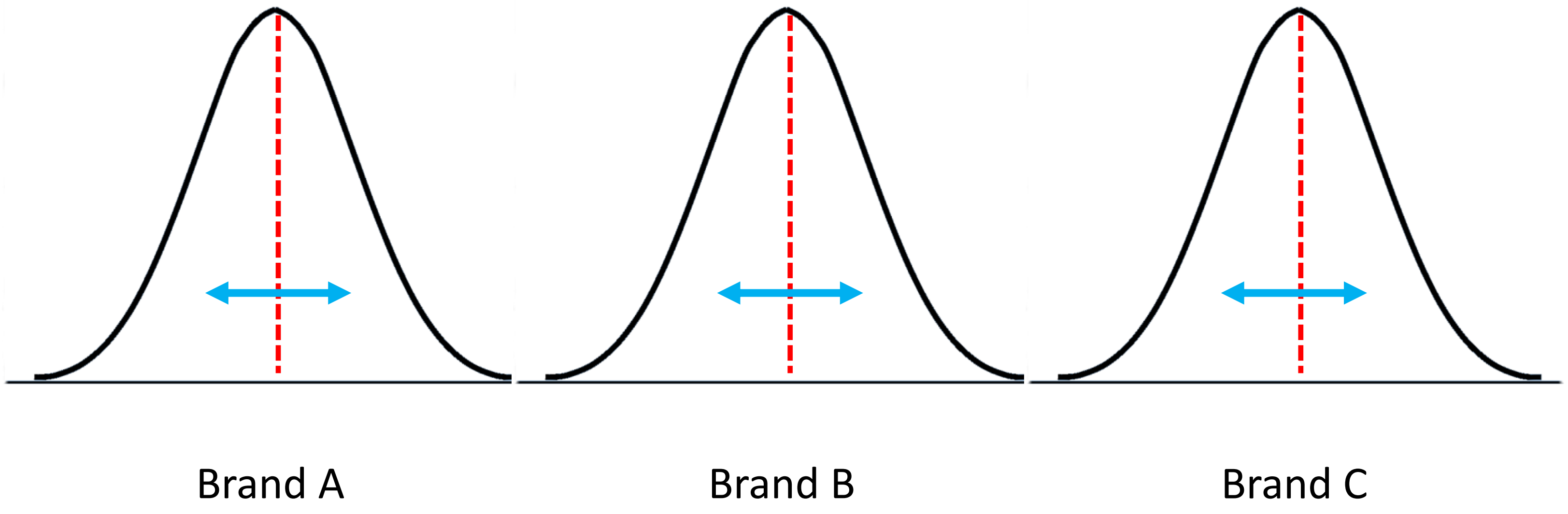
- Compare the taste preferences of consumers for three different brands of bottled water

# Why not use t-Tests?

- More than two samples
- “Alpha inflation”
- Margin of error increases
- Too many calculations!

ANOVA

# ANOVA - Distributions




- Variation ***within*** each group
- Variation ***between*** each group

# ANOVA – $H_0$ and $H_1$

- Objective is usually to compare means of more than two samples
- Null Hypothesis is that the means are all equal
  - $H_0$: $\mu_1 = \mu_2 \ \dots \ = \mu_k$
- Alternative hypothesis is that at least two of the means differ
  - $H_1$: At least two of *k* means differ
  - $H_1$: $\mu_1 \neq \mu_2 \ \dots \ \neq \mu_k$

# How the test works

- Basic idea…
  - Calculate the mean of observations within each group
  - Compare the variance among these means to the average variance within each group

- F statistic
  - Ratio of the variance among the means divided by the average variance within groups

R. A. Fischer
Image source: Wikipedia

# NCI Statistics Table / Booklet

| | |
|---|---|
| **Within Groups Mean Sum of Squares** | $MSS_W = \dfrac{\sum_{g \in G} (X - \bar{X}_g)^2}{n - k}$ |
| **Between Groups Mean Sum of Squares** | $MSS_B = \dfrac{\sum_{g \in G} n_g(\bar{X}_g - \bar{X}_G)^2}{k - 1}$ |
| **Test Statistic for Tests Concerning the Differences between the Variances of Two Populations (Normally Distributed Populations)** | $F = \dfrac{MSS_B}{MSS_W}$ <br> $df_B = k - 1$ <br> $df_W = n - k$ |

# Sample Data

| Group - Language Scores | | |
|---|---|---|
| *Group 1 - 5 Hours* | *Group 2 - 10 Hours* | *Group 3 - 20 Hours* |
| 87 | 87 | 89 |
| 86 | 85 | 91 |
| 76 | 99 | 96 |
| 56 | 85 | 87 |
| 78 | 79 | 89 |
| 98 | 81 | 90 |
| 77 | 82 | 89 |
| 66 | 78 | 96 |
| 75 | 85 | 96 |
| 67 | 91 | 93 |

Data source: Three groups of pre-schoolers and their language scores (Salkind (2014)

# Null and Alternate Hypotheses

| Group - Language Scores | | |
|---|---|---|
| *Group 1 - 5 Hours* | *Group 2 - 10 Hours* | *Group 3 - 20 Hours* |

- $H_0: \mu_1 = \mu_2 = \mu_3$
- $H_1: \mu_1 \neq \mu_2 \neq \mu_3$

# Manual Walkthrough – Excel

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Group - Language Scores | | | | | | |
| | Group 1<br>5 Hours | Group 2<br>10 Hours | Group 3<br>20 Hours | | $(x_1 - \bar{x}_1)^2$ | $(x_2 - \bar{x}_2)^2$ | $(x_3 - \bar{x}_3)^2$ |
| | 87 | 87 | 89 | | 108.16 | 3.24 | 6.76 |
| | 86 | 85 | 91 | | 88.36 | 0.04 | 0.36 |
| | 76 | 99 | 96 | | 0.36 | 190.44 | 19.36 |
| | 56 | 85 | 87 | | 424.36 | 0.04 | 21.16 |
| | 78 | 79 | 89 | | 1.96 | 38.44 | 6.76 |
| | 98 | 81 | 90 | | 457.96 | 17.64 | 2.56 |
| | 77 | 82 | 89 | | 0.16 | 10.24 | 6.76 |
| | 66 | 78 | 96 | | 112.36 | 51.84 | 19.36 |
| | 75 | 85 | 96 | | 2.56 | 0.04 | 19.36 |
| | 67 | 91 | 93 | | 92.16 | 33.64 | 1.96 |
| n | 10 | 10 | 10 | ∑ | 1288.4 | 345.6 | 104.4 |
| $\bar{x}_g$ | 76.6 | 85.2 | 91.6 | | | | |
| $\bar{x}_G$ | | 84.46666667 | | | | | |
| k | | 3 | | $SS_w$ | 1738.4 | | |
| | | | | n - k | 27 | | |
| | | | | **$MSS_W$** | **64.38518519** | | |
| | | | | | | | |
| | | | | $n_1(\bar{x}_1 - \bar{x}_G)^2$ | 618.8444444 | | |
| | | | | $n_2(\bar{x}_2 - \bar{x}_G)^2$ | 5.377777778 | | |
| | | | | $n_3(\bar{x}_3 - \bar{x}_G)^2$ | 508.8444444 | | |
| | | | | ∑ | 1133.066667 | | |
| | | | | | | | |
| | | | | **$MSS_B$** | **566.5333333** | | |
| | | | | | | | |
| | | | | **F** | **8.799125633** | | |

# SPSS Result

**Oneway**

```
[DataSet1] C:\Users\eugeneol\SkyDrive\My Documents\2014-2015 Lecture Notes\BSHBIS4-BAHTM-HDSDA
```

**ANOVA**

Language_Score

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 1133.067 | 2 | 566.533 | 8.799 | .001 |
| Within Groups | 1738.400 | 27 | 64.385 | | |
| Total | 2871.467 | 29 | | | |

# Excel Result

| Anova: Single Factor | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | | |
| SUMMARY | | | | | | |
| *Groups* | *Count* | *Sum* | *Average* | *Variance* | | |
| Group 1 - 5 Hours | 10 | 766 | 76.6 | 143.1555556 | | |
| Group 2 - 10 Hours | 10 | 852 | 85.2 | 38.4 | | |
| Group 3 - 20 Hours | 10 | 916 | 91.6 | 11.6 | | |
| | | | | | | |
| | | | | | | |
| ANOVA | | | | | | |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Between Groups | 1133.066667 | 2 | 566.5333333 | **8.799125633** | 0.001141879 | 3.35413083 |
| Within Groups | 1738.4 | 27 | 64.38518519 | | | |
| | | | | | | |
| Total | 2871.466667 | 29 | | | | |

# ANOVA Table

- The ANOVA Table (or Analysis Of Variance) table gives us the following information:
  1. Degrees Of Freedom
  2. The Sum Of The Squares
  3. The Mean Square
  4. The F ratio
  5. The p-value

# ANOVA Table

- A typical ANOVA table will look as follows:

**Oneway**

(SPSS Output)

[DataSet2]

**ANOVA**

Language_Score

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 1133.067 | 2 | 566.533 | 8.799 | .001 |
| Within Groups | 1738.400 | 27 | 64.385 | | |
| Total | 2871.467 | 29 | | | |

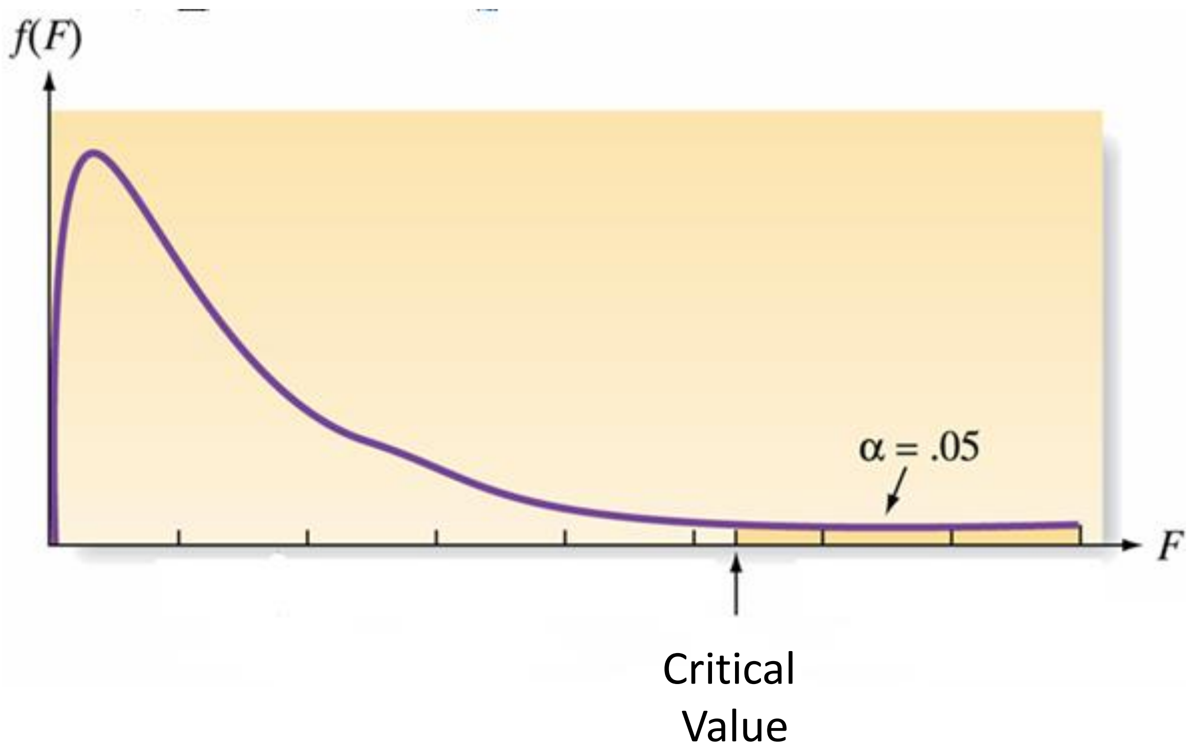
(Excel Output)

| ANOVA | | | | | | |
|---|---|---|---|---|---|---|
| Source of Variation | SS | df | MS | F | P-value | F crit |
| Between Groups | 1133.067 | 2 | 566.533 | 8.799126 | 0.0011 | 3.3541308 |
| Within Groups | 1738.4 | 27 | 64.3852 | | | |
| | | | | | | |
| Total | 2871.467 | 29 | | | | |

# Reporting ANOVA and Critical Value

- Report:
  - $F_{(2, 27)} = 8.799$, $p < 0.05$

- Determine Critical Value for *p*



# Table of the F Distribution

F = 8.799

$df_{Between}$ = 2

$df_{Within}$ = 27

α = 0.05

**Table of the F-Distribution**

Critical values for right-hand tail area equal to **0.05**

| df1: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **df2: 1** | 161 | 200 | 216 | 225 | 230 | 234 | 237 | 239 | 241 | 242 | 243 | 244 | 245 | 245 | 246 |
| **2** | 18.5 | 19.0 | 19.2 | 19.2 | 19.3 | 19.3 | 19.4 | 19.4 | 19.4 | 19.4 | 19.4 | 19.4 | 19.4 | 19.4 | 19.4 |
| **3** | 10.1 | 9.55 | 9.28 | 9.12 | 9.01 | 8.94 | 8.89 | 8.85 | 8.81 | 8.79 | 8.76 | 8.74 | 8.73 | 8.71 | 8.70 |
| **4** | 7.71 | 6.94 | 6.59 | 6.39 | 6.26 | 6.16 | 6.09 | 6.04 | 6.00 | 5.96 | 5.94 | 5.91 | 5.89 | 5.87 | 5.86 |
| **5** | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.88 | 4.82 | 4.77 | 4.74 | 4.70 | 4.68 | 4.66 | 4.64 | 4.62 |
| **6** | 5.99 | 5.14 | 4.76 | 4.53 | 4.39 | 4.28 | 4.21 | 4.15 | 4.10 | 4.06 | 4.03 | 4.00 | 3.98 | 3.96 | 3.94 |
| **7** | 5.59 | 4.74 | 4.35 | 4.12 | 3.97 | 3.87 | 3.79 | 3.73 | 3.68 | 3.64 | 3.60 | 3.57 | 3.55 | 3.53 | 3.51 |
| **8** | 5.32 | 4.46 | 4.07 | 3.84 | 3.69 | 3.58 | 3.50 | 3.44 | 3.39 | 3.35 | 3.31 | 3.28 | 3.26 | 3.24 | 3.22 |
| **9** | 5.12 | 4.26 | 3.86 | 3.63 | 3.48 | 3.37 | 3.29 | 3.23 | 3.18 | 3.14 | 3.10 | 3.07 | 3.05 | 3.03 | 3.01 |
| **10** | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.14 | 3.07 | 3.02 | 2.98 | 2.94 | 2.91 | 2.89 | 2.86 | 2.85 |
| **11** | 4.84 | 3.98 | 3.59 | 3.36 | 3.20 | 3.09 | 3.01 | 2.95 | 2.90 | 2.85 | 2.82 | 2.79 | 2.76 | 2.74 | 2.72 |
| **12** | 4.75 | 3.89 | 3.49 | 3.26 | 3.11 | 3.00 | 2.91 | 2.85 | 2.80 | 2.75 | 2.72 | 2.69 | 2.66 | 2.64 | 2.62 |
| **13** | 4.67 | 3.81 | 3.41 | 3.18 | 3.03 | 2.92 | 2.83 | 2.77 | 2.71 | 2.67 | 2.63 | 2.60 | 2.58 | 2.55 | 2.53 |
| **14** | 4.60 | 3.74 | 3.34 | 3.11 | 2.96 | 2.85 | 2.76 | 2.70 | 2.65 | 2.60 | 2.57 | 2.53 | 2.51 | 2.48 | 2.46 |
| **15** | 4.54 | 3.68 | 3.29 | 3.06 | 2.90 | 2.79 | 2.71 | 2.64 | 2.59 | 2.54 | 2.51 | 2.48 | 2.45 | 2.42 | 2.40 |
| **16** | 4.49 | 3.63 | 3.24 | 3.01 | 2.85 | 2.74 | 2.66 | 2.59 | 2.54 | 2.49 | 2.46 | 2.42 | 2.40 | 2.37 | 2.35 |
| **17** | 4.45 | 3.59 | 3.20 | 2.96 | 2.81 | 2.70 | 2.61 | 2.55 | 2.49 | 2.45 | 2.41 | 2.38 | 2.35 | 2.33 | 2.31 |
| **18** | 4.41 | 3.55 | 3.16 | 2.93 | 2.77 | 2.66 | 2.58 | 2.51 | 2.46 | 2.41 | 2.37 | 2.34 | 2.31 | 2.29 | 2.27 |
| **19** | 4.38 | 3.52 | 3.13 | 2.90 | 2.74 | 2.63 | 2.54 | 2.48 | 2.42 | 2.38 | 2.34 | 2.31 | 2.28 | 2.26 | 2.23 |
| **20** | 4.35 | 3.49 | 3.10 | 2.87 | 2.71 | 2.60 | 2.51 | 2.45 | 2.39 | 2.35 | 2.31 | 2.28 | 2.25 | 2.22 | 2.20 |
| **21** | 4.32 | 3.47 | 3.07 | 2.84 | 2.68 | 2.57 | 2.49 | 2.42 | 2.37 | 2.32 | 2.28 | 2.25 | 2.22 | 2.20 | 2.18 |
| **22** | 4.30 | 3.44 | 3.05 | 2.82 | 2.66 | 2.55 | 2.46 | 2.40 | 2.34 | 2.30 | 2.26 | 2.23 | 2.20 | 2.17 | 2.15 |
| **23** | 4.28 | 3.42 | 3.03 | 2.80 | 2.64 | 2.53 | 2.44 | 2.37 | 2.32 | 2.27 | 2.24 | 2.20 | 2.18 | 2.15 | 2.13 |
| **24** | 4.26 | 3.40 | 3.01 | 2.78 | 2.62 | 2.51 | 2.42 | 2.36 | 2.30 | 2.25 | 2.22 | 2.18 | 2.15 | 2.13 | 2.11 |
| **25** | 4.24 | 3.39 | 2.99 | 2.76 | 2.60 | 2.49 | 2.40 | 2.34 | 2.28 | 2.24 | 2.20 | 2.16 | 2.14 | 2.11 | 2.09 |
| **30** | 4.17 | 3.32 | 2.92 | 2.69 | 2.53 | 2.42 | 2.33 | 2.27 | 2.21 | 2.16 | 2.13 | 2.09 | 2.06 | 2.04 | 2.01 |
| **40** | 4.08 | 3.23 | 2.84 | 2.61 | 2.45 | 2.34 | 2.25 | 2.18 | 2.12 | 2.08 | 2.04 | 2.00 | 1.97 | 1.95 | 1.92 |
| **60** | 4.00 | 3.15 | 2.76 | 2.53 | 2.37 | 2.25 | 2.17 | 2.10 | 2.04 | 1.99 | 1.95 | 1.92 | 1.89 | 1.86 | 1.84 |
| **120** | 3.92 | 3.07 | 2.68 | 2.45 | 2.29 | 2.18 | 2.09 | 2.02 | 1.96 | 1.91 | 1.87 | 1.83 | 1.80 | 1.78 | 1.75 |

# How to calculate an ANOVA table

- Putting it all together in an ANOVA table:

**ANOVA**

Language_Score

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 1133.067 | 2 | 566.533 | 8.799 | .001 |
| Within Groups | 1738.400 | 27 | 64.385 | | |
| Total | 2871.467 | 29 | | | |

$$F_{(2,\ 27)} = 8.799,\ p < 0.05$$

**ANOVA Table**

| **Source of Variation** | **Df** | **SS** | **MS** $= SS/df$ | $\mathbf{F_{stat}}$ | $\mathbf{F_{crit}}$ |
|---|---|---|---|---|---|
| Between Groups ($SS_B$) | $k-1$ | $\sum_{g \in G} n_g(\bar{X}_g - \bar{X}_G)^2$ | $MSS_B = \frac{\sum_{g \in G} n_g(\bar{X}_g - \bar{X}_G)^2}{k-1}$ | $= \frac{MSS_B}{MSS_W}$ | |
| Within Groups ($SS_W$) | $n-k$ | $\sum_{g \in G} (X - \bar{X}_g)^2$ | $MSS_W = \frac{\sum_{g \in G} (X - \bar{X}_g)^2}{n-k}$ | | |
| Total ($SS_{Tot}$) | $n-1$ | | | | |

# $F_{(2,\ 27)} = 8.799,\ p < 0.05$

- F represents the test statistic that was used
- 2 and 27 are the numbers of degrees of freedom for the between-group and within-groups respectively
- 8.799 is the obtained value for F
- *P* < 0.05 indicates that the probability is less than 5% that the difference between the groups is due to chance alone

# Conclusion

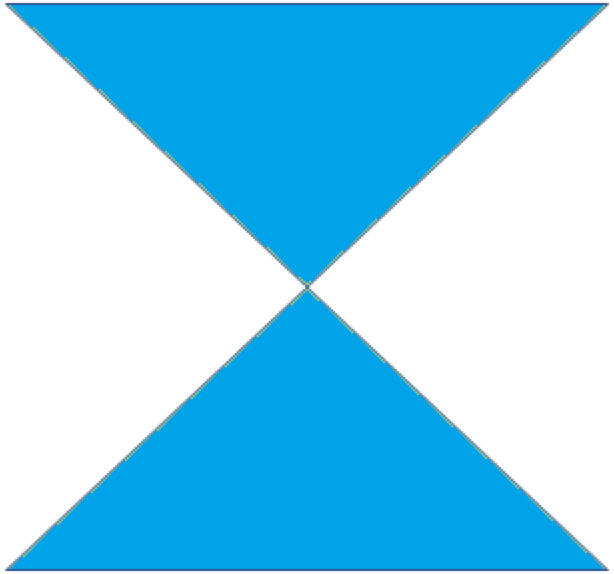
$H_0$: $\mu_1 = \mu_2 = \mu_3$
$H_1$: $\mu_1 \neq \mu_2 \neq \mu_3$

$F_{(2, 27)} = 8.799$, $p < 0.05$
Critical Value = 3.39

Reject the null hypothesis

- It appears there is a significant difference among the three sets of language scores

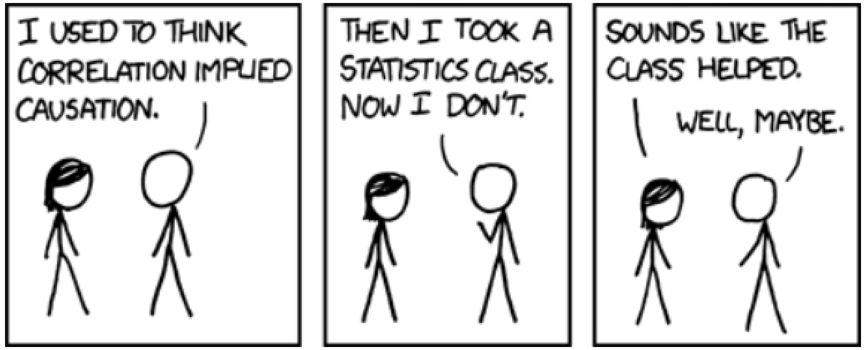
# Complete BC3 Quiz



| | |
|---|---|
| Quiz Start Date | 11/10/2024 |
| End Date | 25/10/2024 |

(record any incorrect answers before End Date)

# Correlation

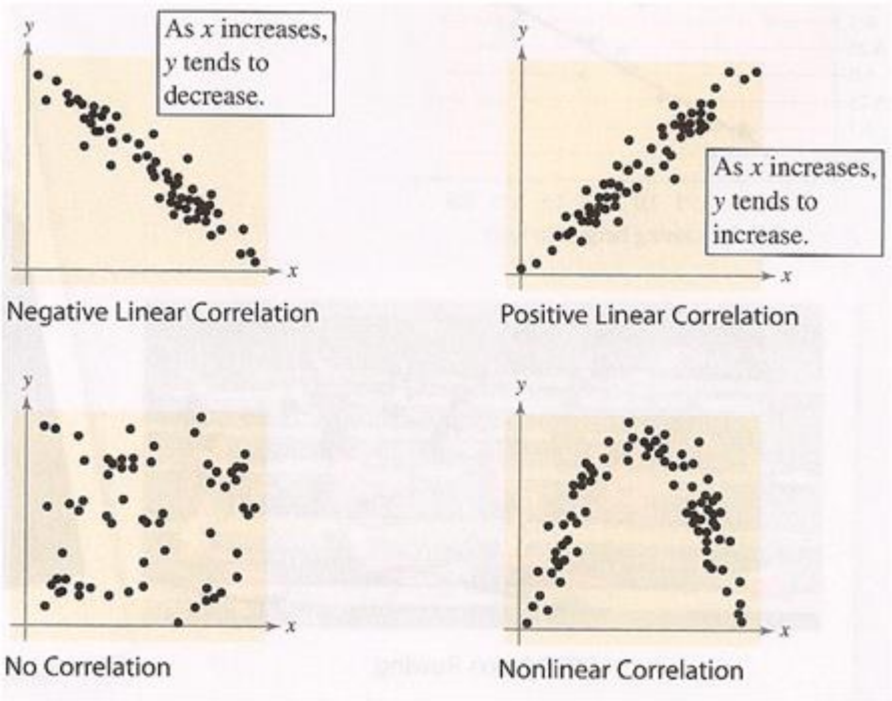



https://medium.com/@plotlygraphs/spurious-correlations-56752fcffb69

# Correlation

- Purpose:

  a way to measure how associated or related two variables are



# Direction of the Correlation

**Positive relationship** – Variables change in the same direction.

- As X is increasing, Y is increasing
- As X is decreasing, Y is decreasing

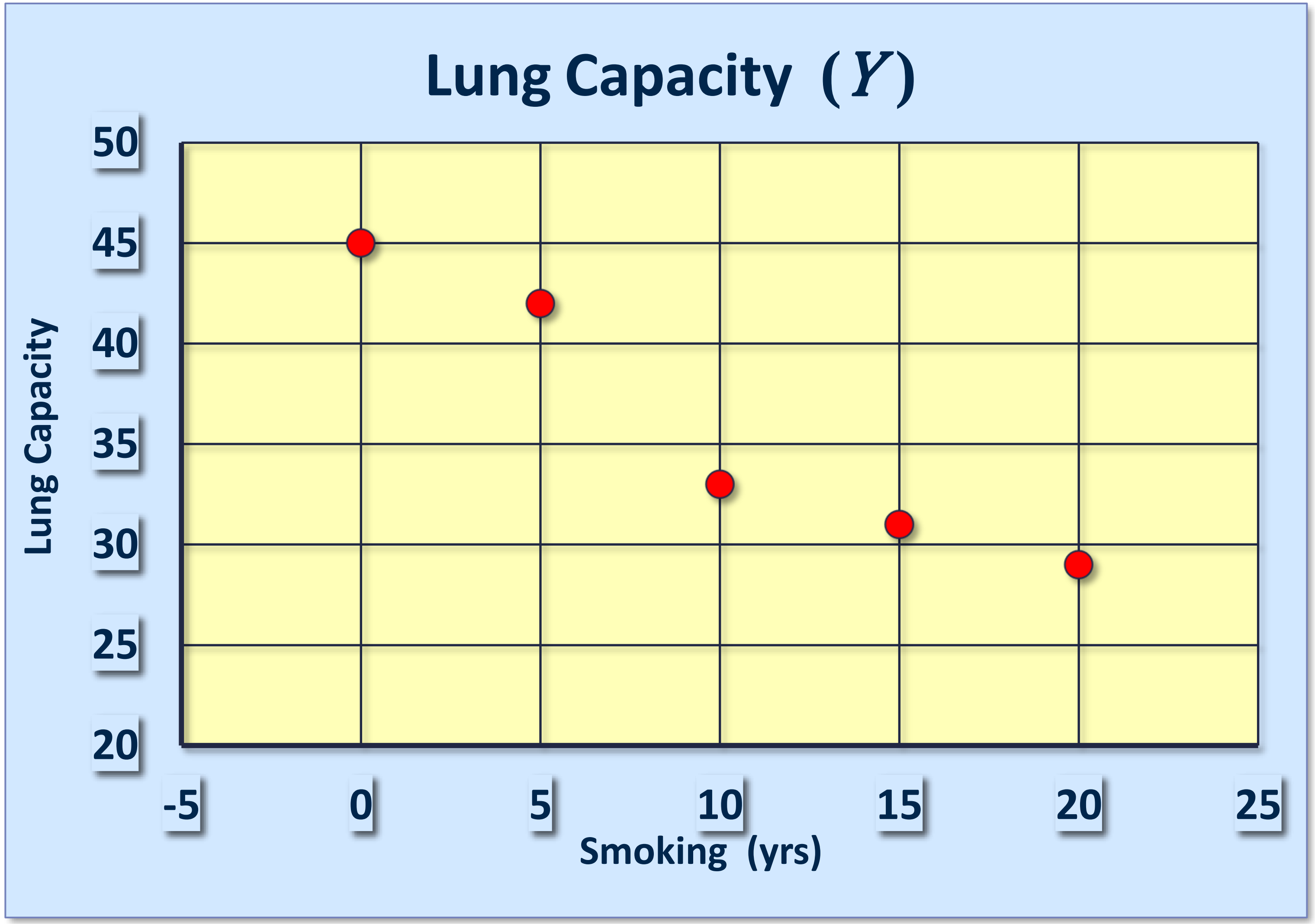
e.g. as height increases, so does weight.

**Negative relationship** – Variables change in opposite directions.

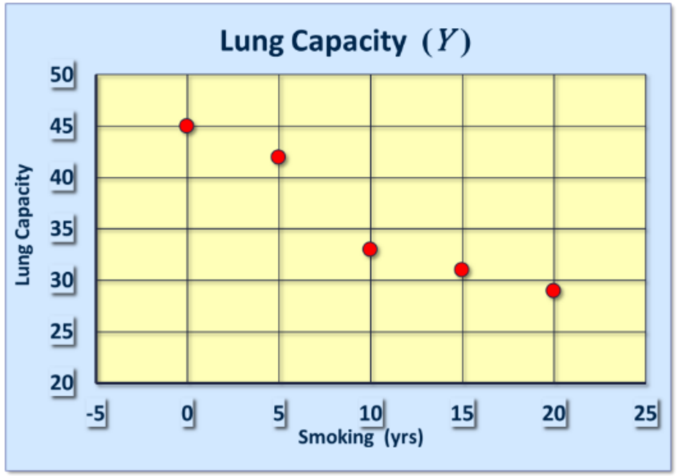
- As X is increasing, Y is decreasing
- As X is decreasing, Y is increasing

e.g. as TV time increases, grades decrease

# Smoking and Lung Capacity



# Correlation coefficient




- Interpreting correlation using a scatter plot can be subjective.
- A more precise way to measure the type and strength of a linear correlation between two variables is to calculate the correlation coefficient.
- The correlation coefficient (r) is a measure of the strength and the direction of a linear relationship between two variables.

# Linear Equation

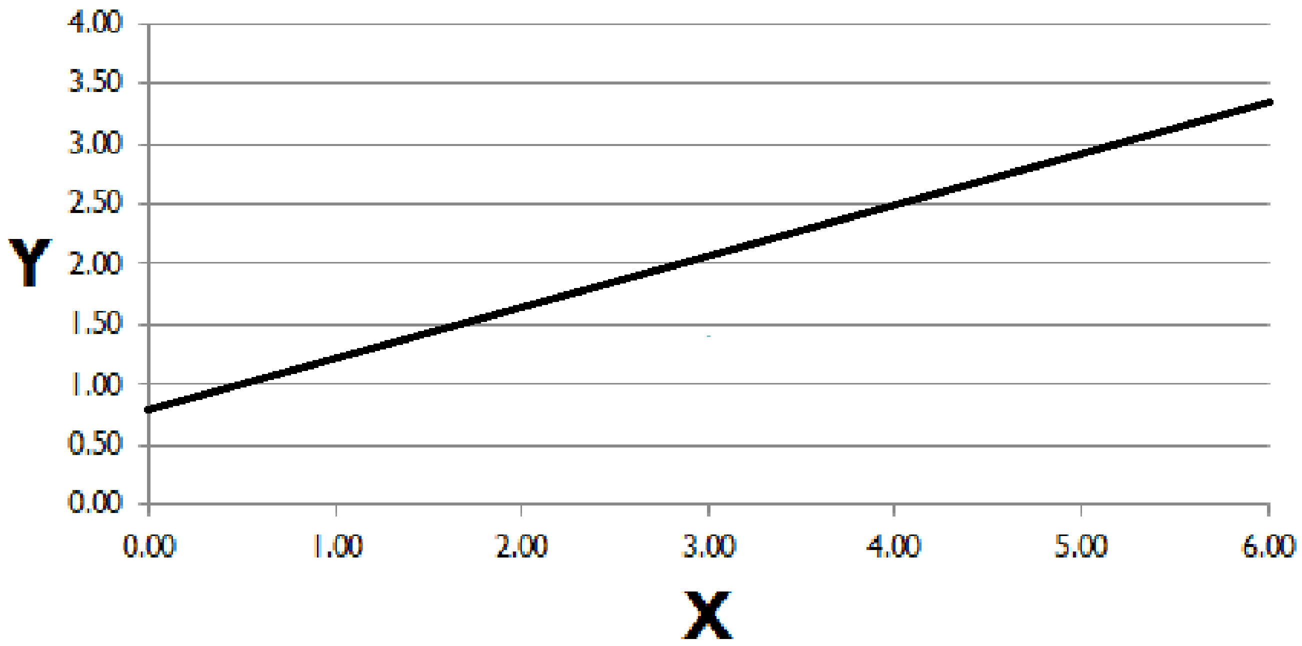
Slope of Line

$$y = a + bx$$

Dependent Variable

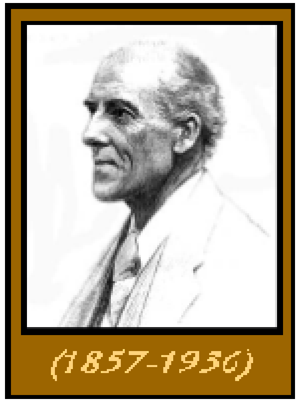
Y intercept

Independent Variable



# Pearson Correlation coefficient (r)

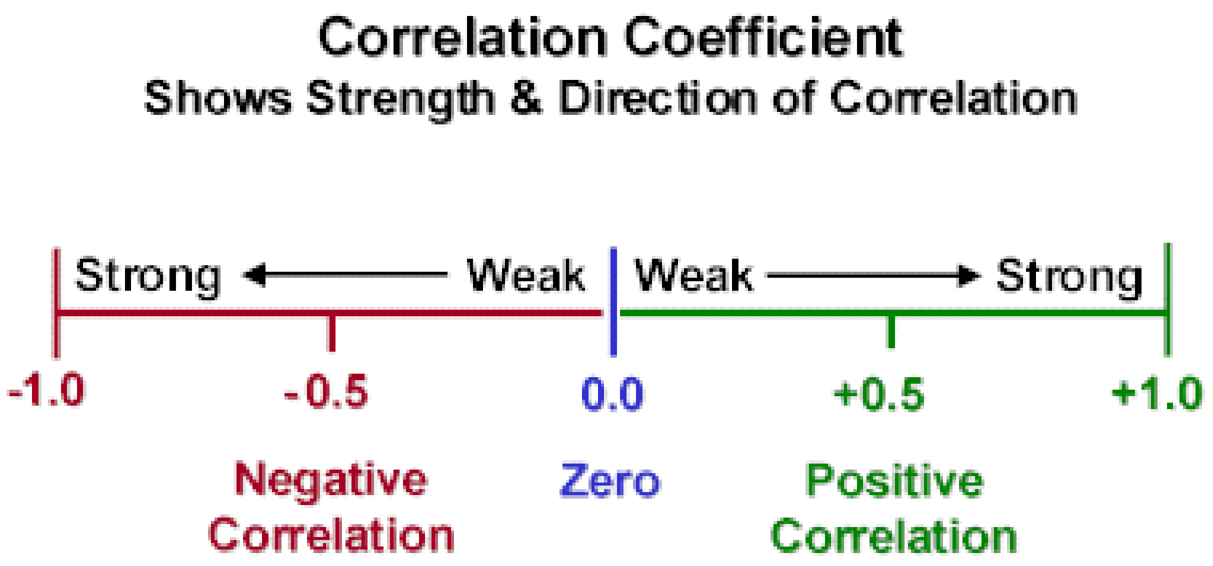
$$r = \frac{\sum_{i=1}^{n}\left(\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}}$$




Karl Pearson

Image source: http://www.alea.pt/html/nomesEdatas/swf/biografias.asp?art=13

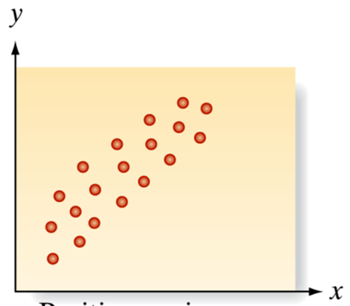
# Values of r



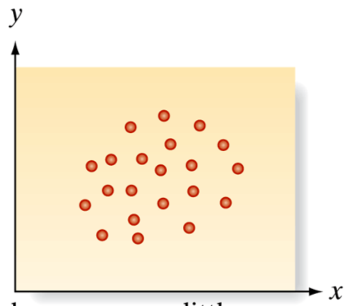

"Rule of Thumb" (Salkind)

| Size of the Correlation | Coefficient General Interpretation |
|---|---|
| **.8 to 1.0** | Very strong relationship |
| **.6 to .8** | Strong relationship |
| **.4 to .6** | Moderate relationship |
| **.2 to .4** | Weak relationship |
| **.0 to .2** | Weak or no relationship |

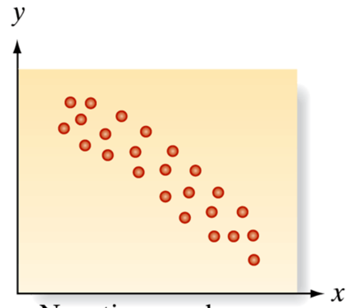
# Values of *r* and their implications



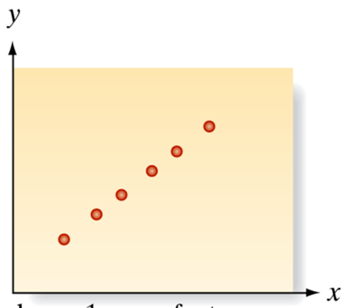

a. Positive *r*: *y* increases as *x* increases



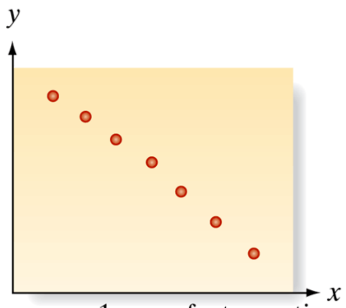

b. *r* near zero: little or no relationship between *y* and *x*



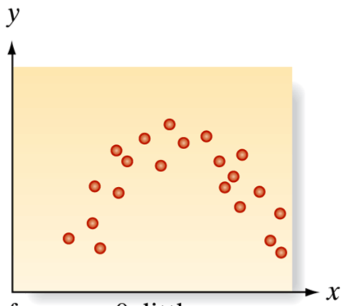

c. Negative *r*: *y* decreases as *x* increases




d. $r = 1$: a perfect positive relationship between *y* and *x*



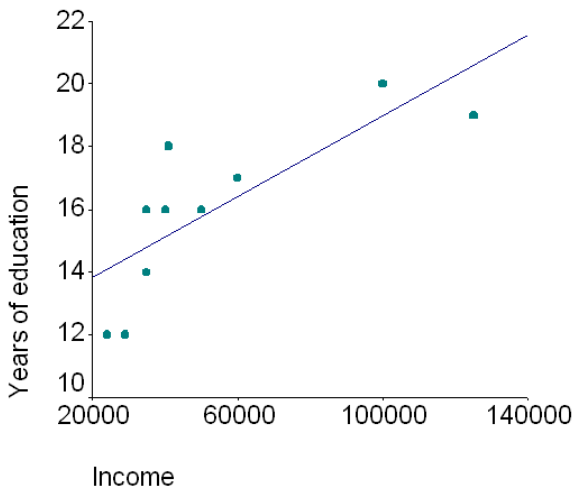

e. $r = -1$: a perfect negative relationship between *y* and *x*



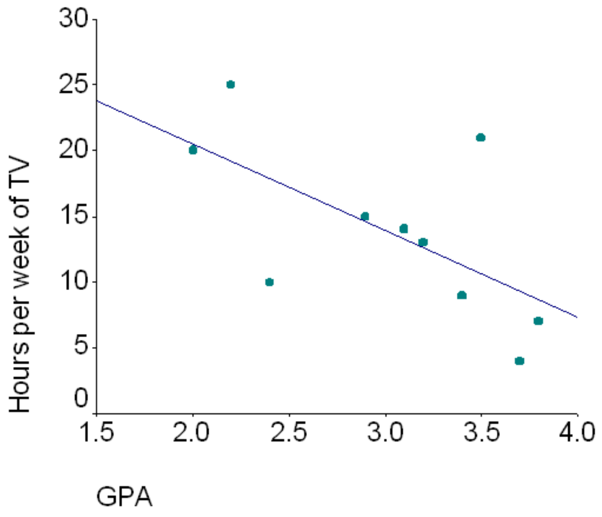

f. *r* near 0: little or no linear relationship between *y* and *x*

# *Strength*

- Correlations, whether positive or negative, range in their strength from weak to strong




In this sample, the correlation is 0.79.




In this sample, the correlation is -0.63.

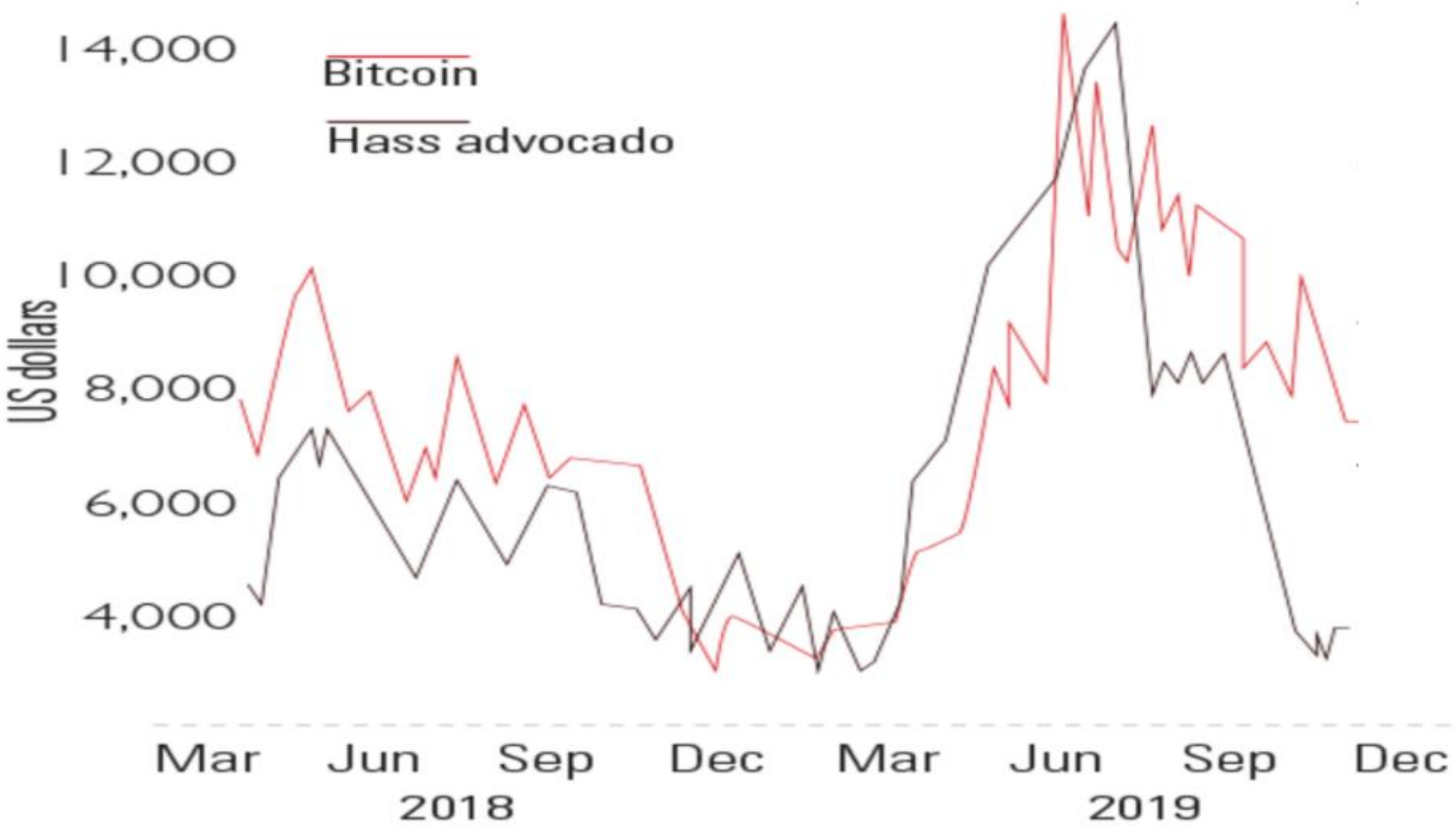
# Use

- An advantage of the correlation method is that we can make predictions based on what we know about correlations i.e.  if two variables are correlated, we can predict one based on the other

- A correlation tells us that the two variables are related, but we cannot say anything about whether one caused the other.
- This method does not allow us to come to any conclusions about cause and effect.

**Correlation is not Causation!**

## *The price of Mexican Hass avocados affects the price of bitcoin?*




https://moneyweek.com/

# Causality

- For any two correlated events A and B, the following relationships are possible:
  - A causes B
  - B causes A
  - A and B are consequences of a common cause, but do not cause each other
  - There is no connection between A and B, the correlation is coincidental

# Correlation is not Causation!