U O W

Research Methodology

Quantitative Research

UNIVERSITY
OF WOLLONGONG
AUSTRALIA

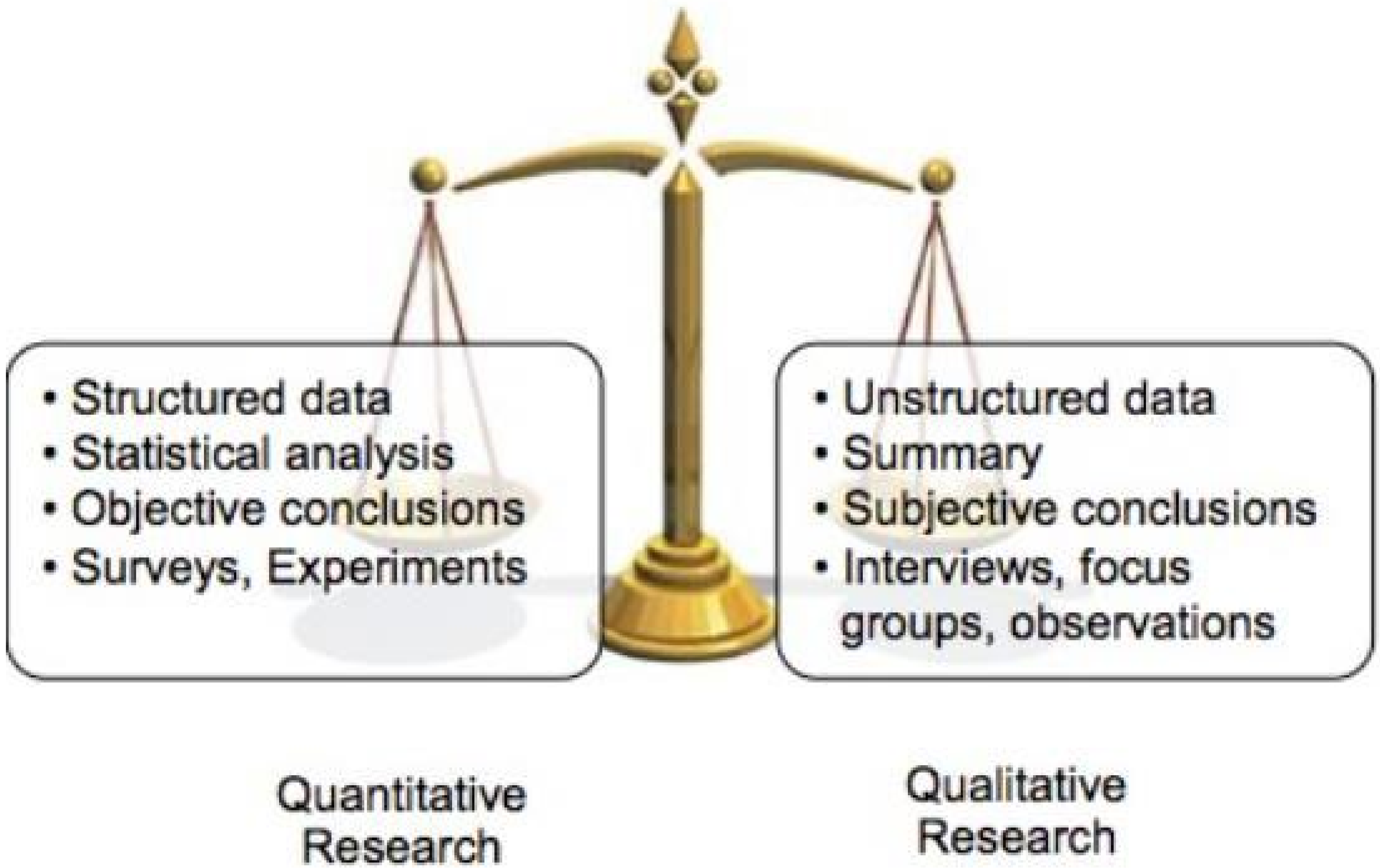
Review

- Data collection
- Basics in statistics and probability
- Sampling

Content

- Definition
- Main research methodology
  - Descriptive
  - Explanatory
  - Predictive
- General steps
  - Hypothesis testing

Quantitative Research vs Qualitative Research



UNIVERSITY OF WOLLONGONG AUSTRALIA

Definition

- Quantitative research is “Explaining phenomenon by collecting numerical data that are analysed using mathematically based methods (in particular statistics)” [Aliagaand Gunderson 2000]

Quantitative Research

- Quantitative research is defined as a systematic investigation of phenomena by gathering quantifiable data and performing statistical, mathematical, or computational techniques.
- Quantitative research methods are research methods dealing with numbers and anything that is measurable in a systematic way of investigation of phenomena and their relationships.
- An entire quantitative study usually ends with confirmation or disconfirmation of the hypothesis tested.
- Researchers using the quantitative method identify one or a few variables that they intend to use in their research work and proceed with data collection related to those variables.

UNIVERSITY OF WOLLONGONG AUSTRALIA

What is Quantitative Research

- Data gathering instrument
  - Quantitative research makes use of tools such as experiments, surveys, measurements and other equipment to collect numerical or measurable data.
- Type of data
  - What most likely are tables containing data in the forms of numbers and statistics
- Approach
  - Researchers tend to remain objectively separated from the subject matter. This is because quantitative research is objective research in the sense that it only seeks precise measurements and analysis of target concepts to answer his inquiry.

UNIVERSITY OF WOLLONGONG AUSTRALIA

When do Quantitative Research

- If your study aims to find out the answer to an inquiry through numerical evidence, then you should make use of the quantitative research.
- In general, use of qualitative research at the beginning of a design process to uncover innovations. Use quantitative research at the end of a design process to measure improvement.
- The main activity for which quantitative research is especially studied is the testing of hypothesis.

What are quantitative Data

- Quantitative data is information about quantities; that is, information that can be measured and written down with numbers.
- Some examples of quantitative data are your height, your shoe size, and attendance rate.
- Quantitative data defines whereas qualitative data describes.

Quantitative data

- Units
  - When we collect data in quantitative research, we have to collect them from someone or something. The people or things (e.g. schools) we collect data on or from are known as units or cases.
- Variables
  - The data that we are collecting from these units are known as variables. Variables are any characteristics of the units we are interested in and want to collect (e.g. height, age)

Quantitative research design

- Quantitative research design is the standard experimental method of most scientific disciplines.
  - They are most commonly used by physical scientists, (although social sciences, education and economics have been known to use this type of research). It is the opposite of qualitative research.
  - Quantitative experiments all use a standard format, with a few minor inter-discipline differences, of generating a hypothesis to be proved or disproved. This hypothesis must be provable by mathematical and statistical means, and is the basis around which the whole experiments is designed.

Quantitative research design

- Randomization of any study groups is essential, and a control group should be included, wherever possible.
- A sound quantitative research design should only manipulate one variable at a time, or statistical analysis becomes cumbersome and open to question.
- Ideally, the research should be constructed in a manner that allows to repeat the experiment and obtain similar results.

Some basic quantitative research method

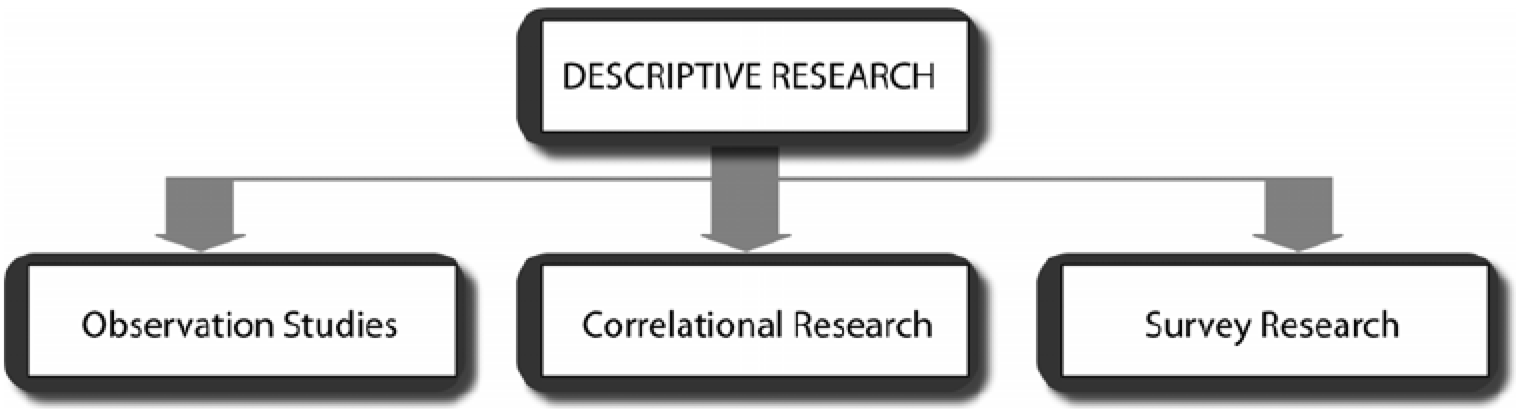
- Quantitative method typically begins with data collection based on a hypothesis or theory and it is followed with application of descriptive or inferential statistics.
- Levels of quantitative analysis
  - Descriptive
  - Explanatory
  - Predictive

Descriptive

- This type of research describes what exists and may help to uncover new facts and meaning. The purpose is to observe, describe, document aspects of a situation as it naturally occurs.
- This involves the collection of data that will provide an account or description of individuals, groups or situations. Instruments we use to obtain data in descriptive studies include questionnaires, interviews (closed question), observation (checklists, etc)
- There is no experimental manipulation or indeed any random selection to groups, as there is in experimental research.
- The characteristics of individuals and groups such as nurses, patients and families may be the focus of descriptive research. It can provide a knowledge base which can act as a springboard for other types of quantitative research methods.

Descriptive

- Quantitative research methods fall under the broad heading of descriptive research.
- Describe the phenomenon with statistical analysis.
- Identify the characteristics of an observed phenomenon
- Explore correlations between two or more entities.



Observation studies

- Involved in both quantitative and qualitative research methods.
- In quantitative methods, focus on a particular factor of behaviour and it is quantified.

UNIVERSITY OF WOLLONGONG AUSTRALIA

Strategies used in observation studies

No.	Strategies	Description
1	Using rating scale	Using rating scale (e.g. Likert Scale) to evaluate the behaviour in terms of specific factor or reasons.
2	Defining the behaviour	Defining the behaviour being studied in a precise and solid manner so that the behaviour is easily recognised during its occurence.
3	Rated by two or more individuals	Having two or more individual ratings the same behaviour independently, without the knowledge of one another's ratings.
4	Clustering the observation periods	Divide observation period into small clusters and then record whether the behaviour does or does not occur during each cluster or segment. Time period may be assigned with some intervals depending on the studies requirement.
5	Train the rater(s)	Train the rater(s) of the behaviour to follow some specific requirement until consistent ratings are obtained during any of the behaviour occurences.

UNIVERSITY OF WOLLONGONG AUSTRALIA

Correlational research

- Quantitative correlational research aims to systematically investigate and explain the nature of the relationships between variables in the real world. Often the quantifiable data (i.e. data that we can quantify or count) from descriptive studies are frequently analysed in this way.
- Correlational research studies go beyond simply describing what exists and are concerned with systematically investigating relationships between two or more variables of interest.
- Such studies only describe and attempt to explain the nature of relationships that exist, and do not examine causality (i.e. whether one variable causes the other).

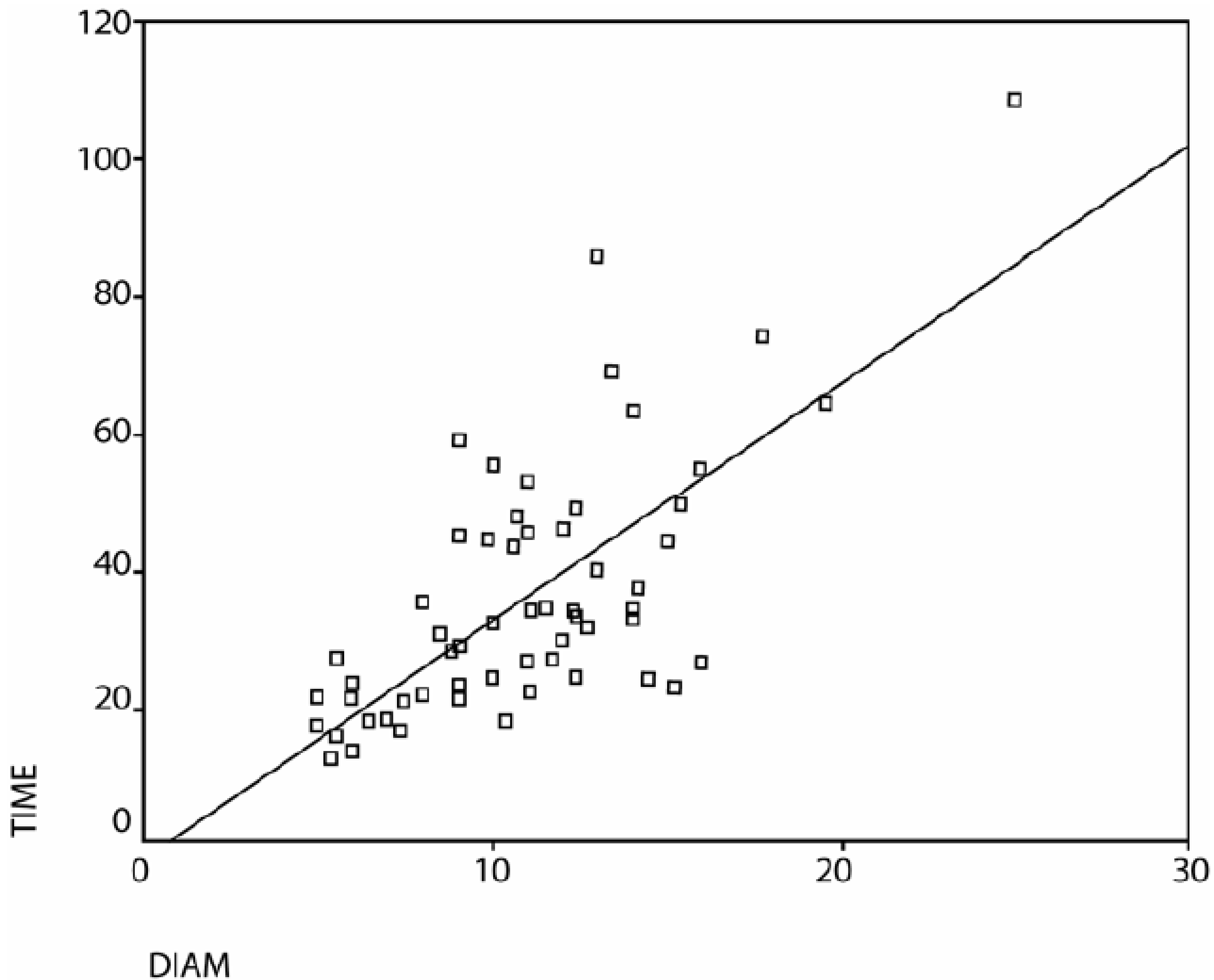
Examples

- The relation between storage size and database capacity
  - as storage becomes larger, it is easier to manage databases => there is a correlation between storage size and database capacity.
- The equation between fame and money.
- The relationship between stress and depression.

Correlational research (features)

- Examine differences of characteristics or variables of two or more entities.
- A correlation exists when one variable increases or decreases correspondingly with the other variable.
- A researcher gathers data about two or more variables in a particular group; these data are numbers that reflect measurement of the characteristics of research questions.

- Correlational results can be represented using various mean of visualisation, for example **scatterplot**



Draw a **scatterplot**

- Step 1: Decide the Two Variables
- Step 2: Collect Data
- Step 3: Map the Data
- Step 4: The Line of Best Fit
- Step 5: Get the result

Examining Scatterplot, we can

- Identify or describe the homogeneity or heterogeneity of the two variables.
- Describe the degree of which the two variables are intercorrelated or using statistical approach known as correlation coefficient.
- Interpret these data and give them meaning

Survey research

- One of the most common methodologies used in ICT
- Kerlinger 1973: survey research is a study on large and small populations by selecting samples chosen from the desired population and to discover relative incidence, distribution and interrelations.
- Ultimate goal: learn about a large population by surveying a sample of the population.

Examples

- A researcher poses a series of questions to the respondents, summaries their responses in percentages, frequency distribution and some other statistical approach.
- Employ face-to-face interviews, telephone interviews or using questionnaires.

Example of questions asked in survey research

What are the beliefs?
Should Service Oriented Architecture (SOA) be adopted in seeking human capital for Human Resource Management?

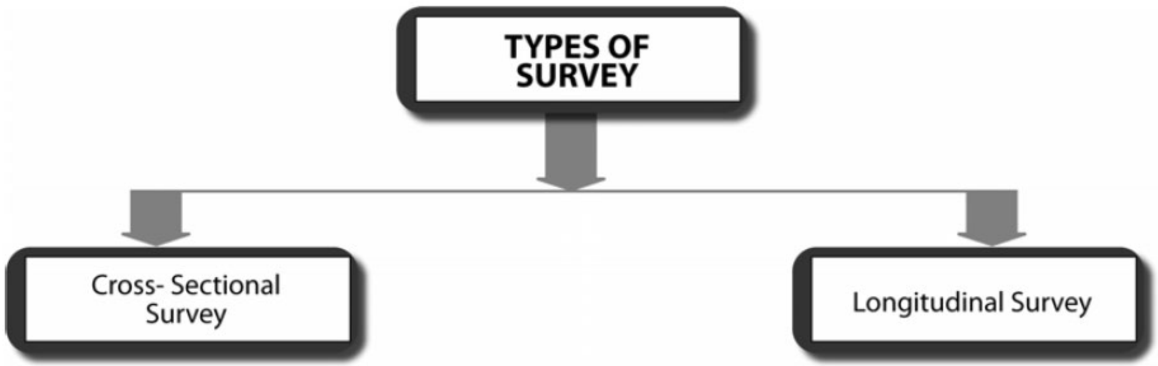
What do they know?
What was the first IT enterprise architecture that dominated Malaysian scenario?

What have they done?
How often do you perform virus scan and monitoring in your organisation?

What do they expect?
Do you think Open Source policy should be implemented in all educational agencies for wider acceptance and cost savings?

Types of survey

- There are two types of survey, depending on the scope of the research work



- Cross-sectional survey: if researchers need a pool of opinions and practices
- Longitudinal survey: if researchers want to compare differences in opinion and practices over time

UNIVERSITY OF WOLLONGONG AUSTRALIA

Cross-sectional survey

- A researcher collects information from a sample drawn from a population
- It involves collecting data at one point of time. The period of data collection can vary and depends on the study weightage.
- Eg: you administer a questionnaire on broadband usage among 500 UOW students for information dissemination using research network. The UOW students may comprise 20-23 year old students. The students could be males and females from different course backgrounds. In this case, the data you obtain is derived from a cross-section of the population at one point of time.

Longitudinal surveys

- Data collection is done at different points of time to observe the changes. Two common types:
  - **Cohort studies**: identify a category of people of interest then randomly select individuals from within that category to survey over time.
  - E.g: a research specifies population and lists the names of all members of this population. At each data collection point, a researcher will select a sample of respondents from the population to see the trends and changes.

  - **Panel studies**: focuses on the same people each time the survey is administered
  - E.g: a researcher can identify a sample from the beginning and follow the respondents over a specified period of time to observe changes in specific respondents and highlight the reasons why these respondents have changed.

Examples

- A survey to study the effects of exercise on 100 kids over a 10-year period beginning at age 12 and following them until they are 22.
- A survey to know how Americans' views on healthcare have changed over the past 10 years.

Comparison between the two surveys

- It depends on the nature of your research project and the questions you're trying to answer.
- longitudinal surveys give us more information about trends, tendencies, opinions or ideas over a long period of time.
- Longitudinal surveys can be costly and difficult to administer, and your research work may not require a temporal data analysis
- Cross-sectional surveys have limited explanatory power because they only capture one moment, but easier to implement.

Explanatory

- Understand why things happen as they do, how reliable is that understanding, or are there other explanations?
  - For example, if the unemployment rate has risen, we want to understand why – is it because of young people entering the labour market, is it because restaurants are sacking dishwashers and replacing them with machines?
  - Statistical models can be invaluable in this situation.

Predictive

- Need a model of future outcomes, how well does it work?
  - For example, if we are looking at admissions to universities, can we predict which students will receive admission offers before they get their public exam results, if we know their performance in school exams? This is not hard for the group of all school students, but is very hard for individual students.

Basic Steps for Quantitative Research

- It uses the deductive reasoning, where the researcher forms a hypothesis, collects data in an investigation of the problem, and then uses the data from the investigation, after analysis is made and conclusions are shared, to prove the hypothesis not false or false

1. Make you observations about something that is unknown, unexplained, or new. Investigate current theory surrounding your problem or issue.
2. Hypothesise an explanation for these observations.
3. Make a prediction of outcomes based on your hypotheses. Formulate a plan to test your prediction.
4. Collect and process your data. If your prediction was correct, go to step 5. If not, the hypothesis has been proven false. Return to step 2 to form a new hypothesis based on your own knowledge.
5. Verify your findings. Make your final conclusions. Present your findings

Quantitative Research (statistics)

- Key Elements:
  - **Counting**
    Example: how many boys in our sample?
  - **Testing Hypotheses**
    Example: are there more boys than girls in the population?
  - **Sample size and power**
    Example: how big a sample do I need to have a good chance of reaching a useful conclusion from my experiment?
  - **Modelling Variability**
    Example: how much do Age & Gender determine academic results?
  - **Prediction**
    Example: can I predict whether a boy aged 16 from a band 1 school will pass an exam?

What is Hypothesis Testing?

- A statistical hypothesis is an assumption about a population parameter. This assumption may or may not be true
- Hypothesis testing refers to the formal procedures used by statisticians to accept or reject statistical hypotheses.
- The best way to determine whether a statistical hypothesis is true would be to examine the entire population.
  - Since that is often impractical, researchers typically examine a random sample from the population.
  - If sample data are not consistent with the statistical hypothesis, the hypothesis is rejected.

What is Hypothesis Testing?

There are two types of statistical hypotheses : null hypothesis and alternative hypothesis.

- **Null hypothesis**. The null hypothesis, denoted by H_0, is usually the hypothesis we wish to be true. Researchers work to reject, nullify or disprove the null hypothesis.
- **Alternative hypothesis**. The alternative hypothesis, denoted by H_1 or H_a. Set up in opposition to null hypothesis.

What is Hypothesis Testing?

- For example, to determine whether a coin was fair and balanced. A null hypothesis is that half the flips would result in Heads and half in Tails. The alternative hypothesis is that the number of Heads and Tails would be very different.

$$H_0: P = 0.5$$
$$H_a: P \neq 0.5$$

- Suppose we flipped the coin 50 times, resulting in 40 Heads and 10 Tails. Given this result, we would be inclined to reject the null hypothesis. We would conclude, based on the evidence, that the coin was probably not fair and balanced.

Hypothesis Testing for a Proportion

Examples: Research Questions

The following are research questions that could be answered using a hypothesis test for one proportion. In each case, we would test the hypothesis by comparing data from a sample to the hypothesized population parameter.

Babies. Is the proportion of babies born male different from .50?

Handedness. Are more than 80% of American's right handed?

Ice cream. Is the percentage of Creamery customers who prefer chocolate ice cream over vanilla less than 80%?

Simplified methods to test a hypothesis

- Set up the hypothesis
- Random variable
- Distribution for the test (mean, standard deviation…)
- Interpretation of the result (p value, confidence interval)
- Make a decision

Bernoulli Distribution

- Variables with boolean numbers
- The Bernoulli distribution is the probability distribution of a random variable which takes the value 1 with probability p and the value 0 with probability q = 1 − p

- $p \qquad x=1$
- $1\text{-}p \qquad x=0$

UNIVERSITY OF WOLLONGONG AUSTRALIA

- Then, we use

$$\bar{p} = \frac{\sum_{i=1}^{n} x_i}{n}$$

to estimate *p, which is the population mean.*

UNIVERSITY OF WOLLONGONG AUSTRALIA

Bernoulli Distribution

- What is the distribution for the population mean
- Central Limit Theorem: In probability theory, the central limit theorem establishes that, in some situations, when independent random variables are added, their properly normalized sum tends toward a normal distribution even if the original variables themselves are not normally distributed.
- Based on the Central Limit Theorem,

$$z = \frac{\sum_{i=1}^{n} x_i/n - p}{\sqrt{Var(\sum_{i=1}^{n} x_i/n)}} = \frac{\bar{p} - p}{\sqrt{p(1-p)/n}}$$

$$\sigma = \sqrt{\frac{p(1-p)}{n}}$$ (Standard Deviation)

- has asymptotically normal distribution, i.e. z~N(0,1)

UNIVERSITY OF WOLLONGONG AUSTRALIA

Example

- Consider estimating the proportion p of the current UOW graduating class who plan to go to graduate school. Suppose we take a sample of 40 graduating students, and suppose that 6 out of the 40 are planning to go to graduate school. Then our estimate is

$$\bar{p} = 6/40 = 0.15$$

of the graduating class plan to go to graduate school

- $\bar{p}$ is based on the sample. Unless we are lucky, it might not be 0.15.

Example

- On average, a random variable misses the mean by one standard deviation

$$\sigma = \sqrt{\frac{p(1-p)}{n}}$$

- That is, the expected size of the miss is σ

This is called the <u>standard error </u>of estimation of the sample proportion or simply standard error of the proportion

UNIVERSITY OF WOLLONGONG AUSTRALIA

Five Steps Hypothesis Testing Procedure

1. Check any necessary assumptions and write null and alternative hypotheses. The null and alternative hypotheses will also be written in terms of population parameters; the null hypothesis will always contain the equality (i.e., =).
2. Calculate an appropriate test statistic. This will vary depending on the test, but it will typically be the difference observed in the sample divided by a standard error. We will see z test statistics.
3. Determine a p-value associated with the test statistic.
4. Decide between the null and alternative hypotheses. If $p \leq \alpha$ reject the null hypothesis. If $p > \alpha$ fail to reject the null hypothesis.
5. State a "real world" conclusion. Based on your decision in step 4, write a conclusion in terms of the original research question.

1. **Check Any Necessary Assumptions and Write Null and Alternative Hypotheses**
   – In terms of the hypotheses, the null hypothesis will always contain the equality, the alternative hypothesis will never contain an equality.

Research Question	**Is the proportion different from p_0?**	**Is the proportion greater than p_0?**	**Is the proportion less than p_0?**
Null Hypothesis, H_0	$p = p_0$	$p = p_0$	$p = p_0$
Alternative Hypothesis, H_a	$p \neq p_0$	$p > p_0$	$p < p_0$
Type of Hypothesis Test	Two-tailed, non-directional	Right-tailed, directional	Left-tailed, directional

p_0 is the hypothesized value of the population proportion.

UNIVERSITY OF WOLLONGONG AUSTRALIA

2. Calculate an Appropriate Test Statistic.

When testing on proportion, a z test statistic using the following formula:

Test statistic: One Group Proportion

$$z = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

$\bar{p}$ = sample proportion

p_0 = hypothesize population proportion

n = sample size

This formula is actually the difference between the sample proportion and hypothesized population proportion divided by the standard error of $\bar{p}$. In doing so, this formula is finding the *z* score for the observed sample in terms of the hypothesized distribution of sample proportions.

3. Determine the p-value Associated with the Test Statistic.

Now, we use the test statistic that we computed in step 2 to determine the probability of obtaining a sample that deviates from the hypothesized population as much as or more than the sample that we have.

$$
\begin{aligned}
H_0 : p = p_0, &\quad H_1 : p \neq p_0 \quad \rightarrow \quad p\text{-value} = 2 \times P(z < -|z_0|) = 2 \times P(z > |z_0|) \\
H_0 : p = p_0, &\quad H_1 : p > p_0 \quad \rightarrow \quad p\text{-value} = P(z > z_0) \\
H_0 : p = p_0, &\quad H_1 : p < p_0 \quad \rightarrow \quad p\text{-value} = P(z < z_0)
\end{aligned}
$$

z_0 is calculated from the z formula

4. Decide Between the Null and Alternative Hypotheses.

We can decide between the null and alternative hypotheses by examining our *p*-values.

If p-value $\leq \alpha$, reject the null hypothesis.

If p-value $> \alpha$, fail to reject the null hypothesis. Assume that $\alpha = .05$.

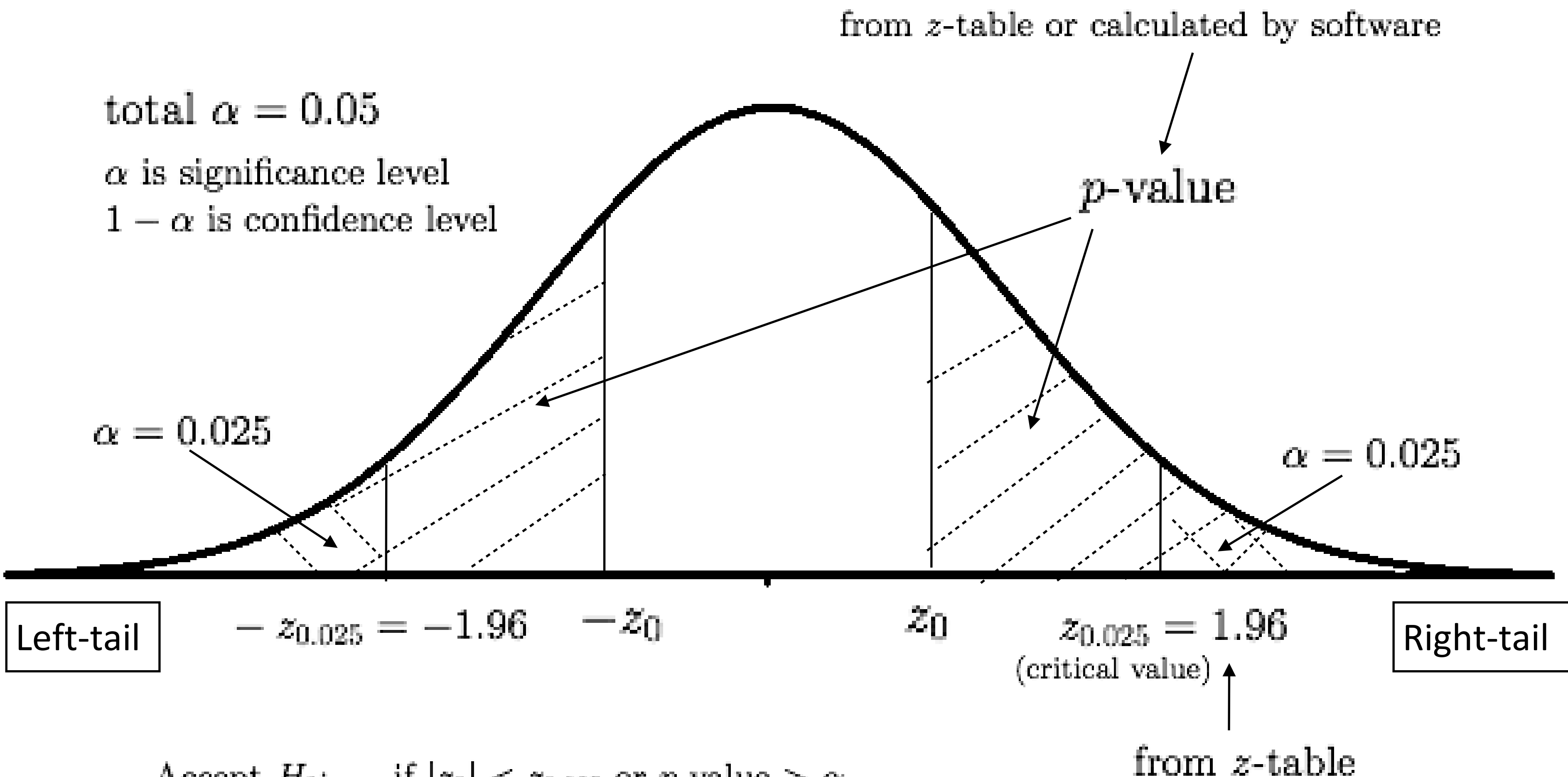
When we reject the null hypothesis, results are said to be statistically significant.

5. State a "Real World" Conclusion.

- Based on our decision in step 4, we will write a sentence or two concerning our decision in relation to the original research question.
- If the sample statistic falls within the rejection region, the null hypothesis will be rejected, otherwise accepted.
- Only one of two decisions is possible in hypothesis testing, either accept or reject the null hypothesis. Instead of "accepting" the null hypothesis, some researchers prefer to phrase the decision as "Do not reject H_0" or "We fail to reject H_0" or "The sample results do not allow us to reject H_0".

Two-tailed z-test

$$H_0 : p = p_0, \quad H_1 : p \neq p_0$$

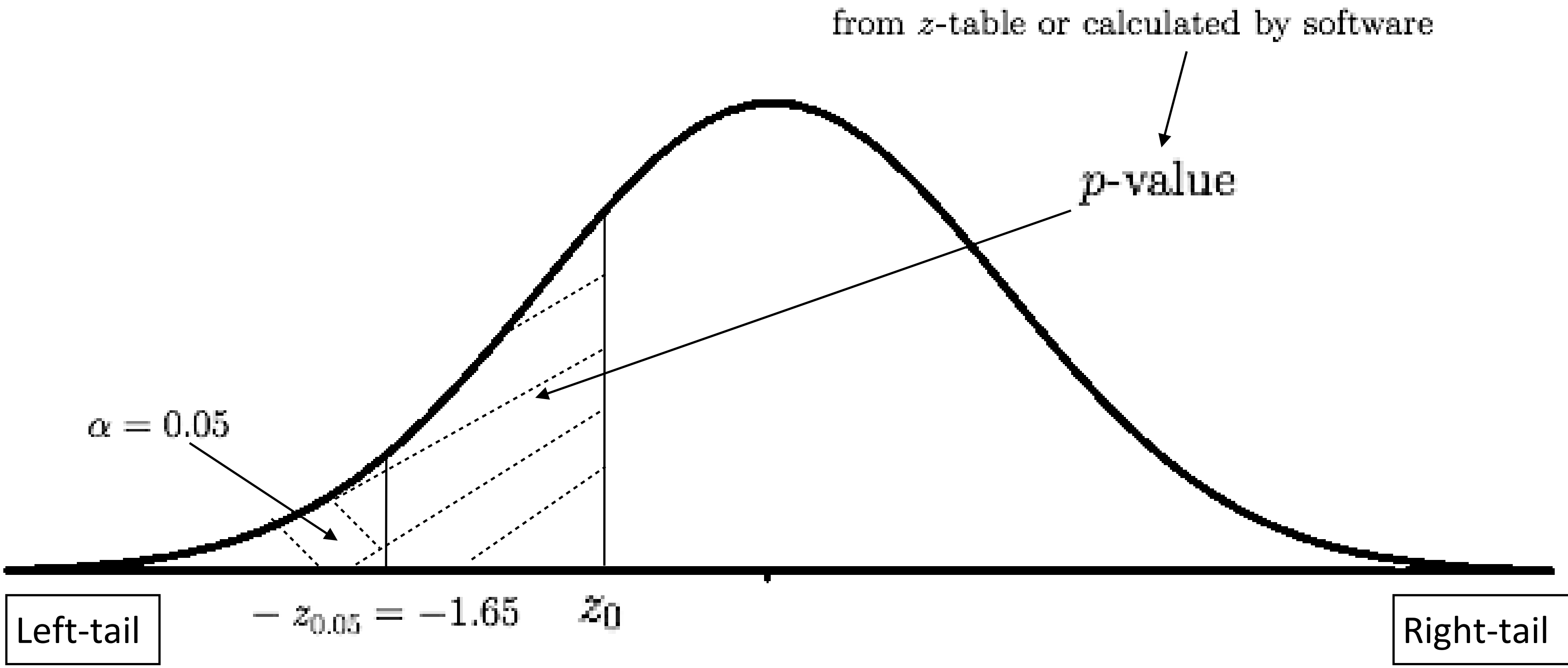


Accept H_0: if $|z_0| < z_{0.025}$ or p-value $> \alpha$

Reject H_0: if $|z_0| \geq z_{0.025}$ or p-value $\leq \alpha$

Left-tailed z-test

$$H_0 : p = p_0, \quad H_1 : p < p_0$$

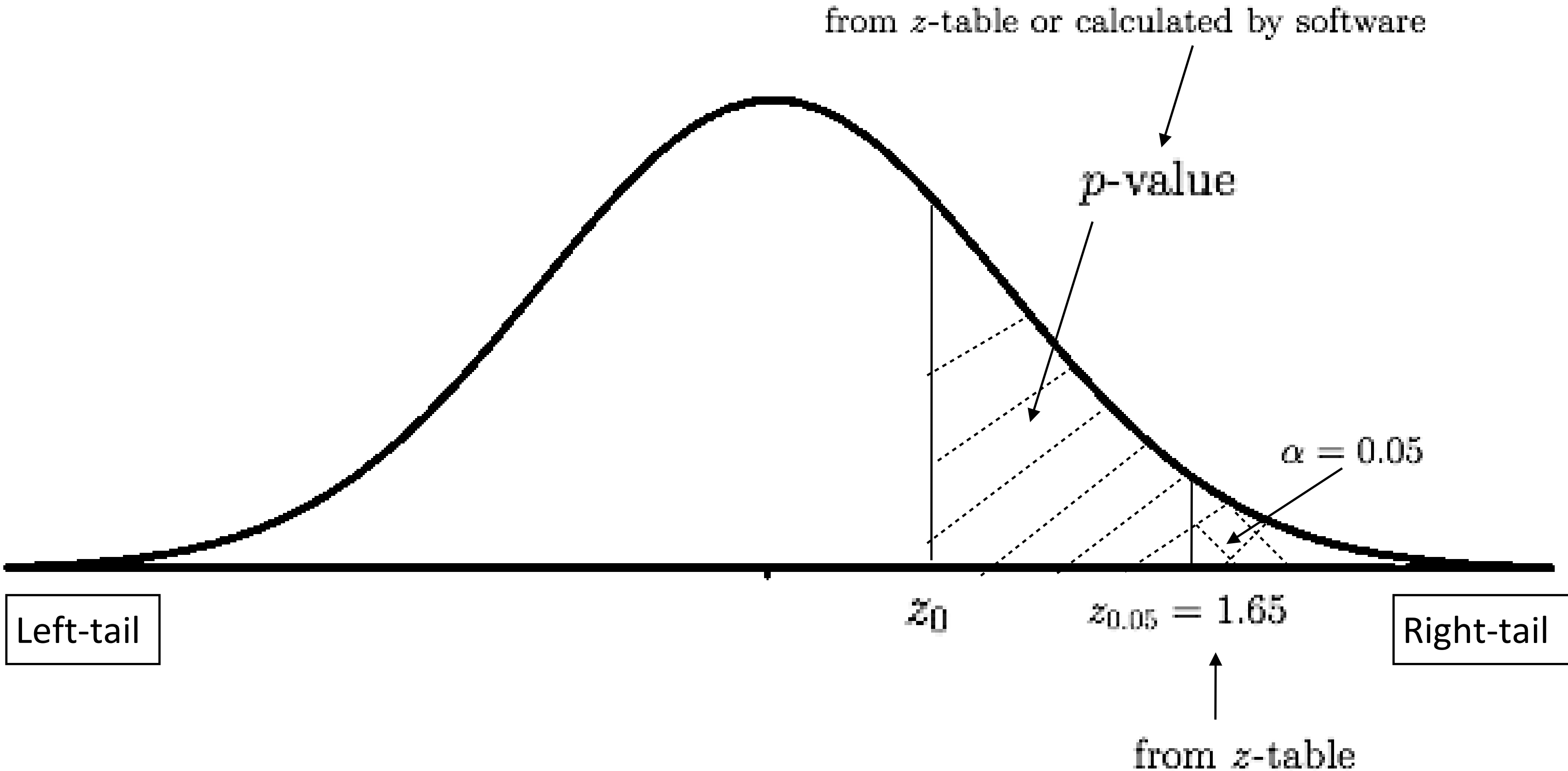


Accept H_0: if $z_0 > z_{0.05}$ or p-value $> \alpha$

Reject H_0: if $z_0 \le z_{0.05}$ or p-value $\le \alpha$

Right-tailed z-test

$$\boxed{H_0 : p = p_0, \quad H_1 : p > p_0}$$



Accept H_0: if $z_0 < z_{0.05}$ or p-value $> \alpha$

Reject H_0: if $z_0 \geq z_{0.05}$ or p-value $\leq \alpha$

Babies Example

Research Question: Is the proportion of babies born male different from .50?

In a sample of 200 babies, 96 were male.

1. Check Any Necessary Assumptions and Write Null and Alternative Hypotheses.

- Is the proportion of babies born male different from 0.50?

$H_0: \; p = p_0 = 0.50$

$H_a: \; p \neq 0.50$

This is a two-tailed test because our research question does not state if the proportion of males should be less than or more than 0.50, it just states that it is different.

2. Calculate an Appropriate Test Statistic.

$$z = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Our z test statistic is -0.566. Given that the population proportion is 0.50, a sample of n = 200 translates to a z score of -0.566.

3. Determine the p-value Associated with the Test Statistic.

Using the standard normal distribution, we want to find the probability of obtaining a z score of -0.566 or more extreme (i.e., less than -0.566).

Table A Standard Normal Cumulative Probabilities

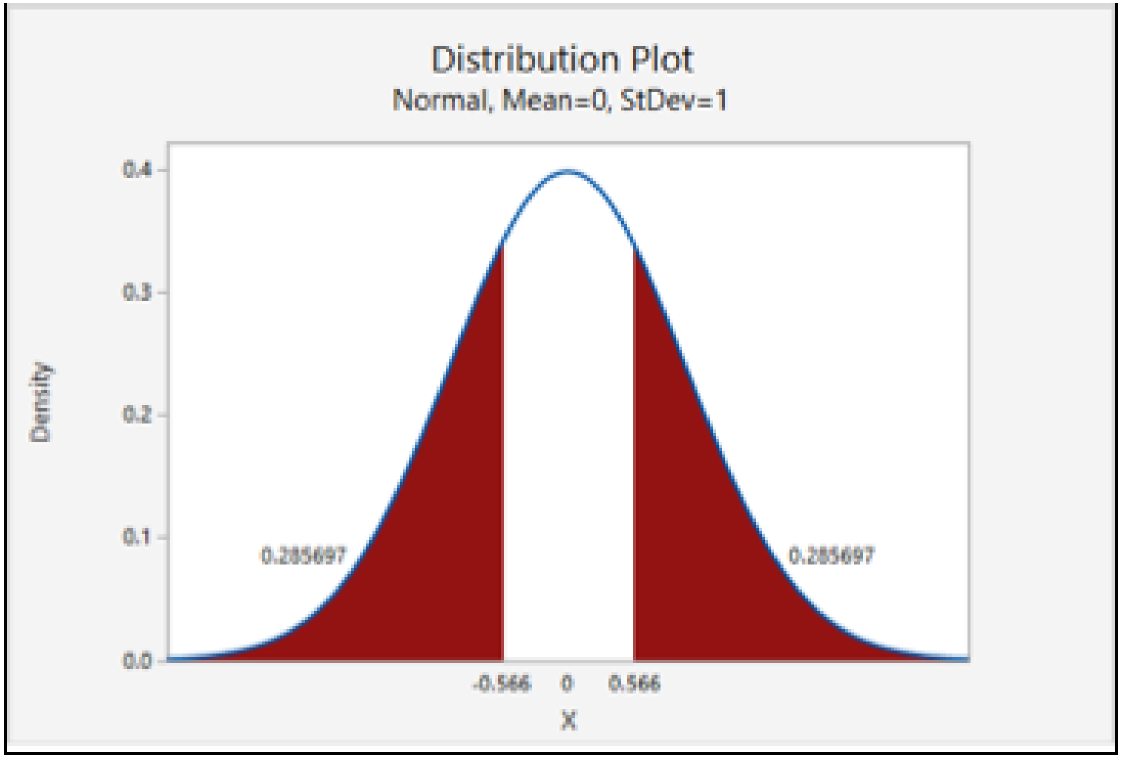
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	[illegible]	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	[illegible]	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	[illegible]	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	[illegible]	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	[illegible]	.0010	.0010
⋮										
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	[illegible]	.2483	.2451
−0.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	$\boxed{.2843}$	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	[illegible]	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

p-value($z < -0.566$) = .2843

Because this is a two-tailed test we must take into account both the left and right tails. To do so, we

multiply the value above by two (p-value=.2843+.2843=.5686). Our p-value is .5686

We could also find this probability in R, we still need to add the proportions in the two tails:



4. Decide Between the Null and Alternative Hypotheses.

p-value > 0.5, therefore our decision is to fail to reject the null hypothesis

5. State a "Real World" Conclusion.

We do not have sufficient evidence to state that in the population the proportion of babies born male is different from .50

- In the above, we carry out the hypothesis test on the population proportion (i.e. the mean of x) when x has Bernoulli distribution.
- If x does not have the Bernoulli distribution, how can we carry out the hypothesis test on the mean of x?
- There are two types of tests can be used for the hypothesis test.
  - One is called z-test
  - The other is call t-test.

Which test should we select

- z-test
  - Test of Population Mean with Known Variance
- t-test
  - Test of Population Mean with Unknown Variance

Z-Test

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

$$H_1 : \mu < \mu_0$$

$$H_1 : \mu > \mu_0$$

To use Z-test to carry out hypothesis test on H_0 against H_1 we need to know the population variance, denoted σ^2

Used to determined whether two samples' means are different when variances are known and sample is large (greater than 30)

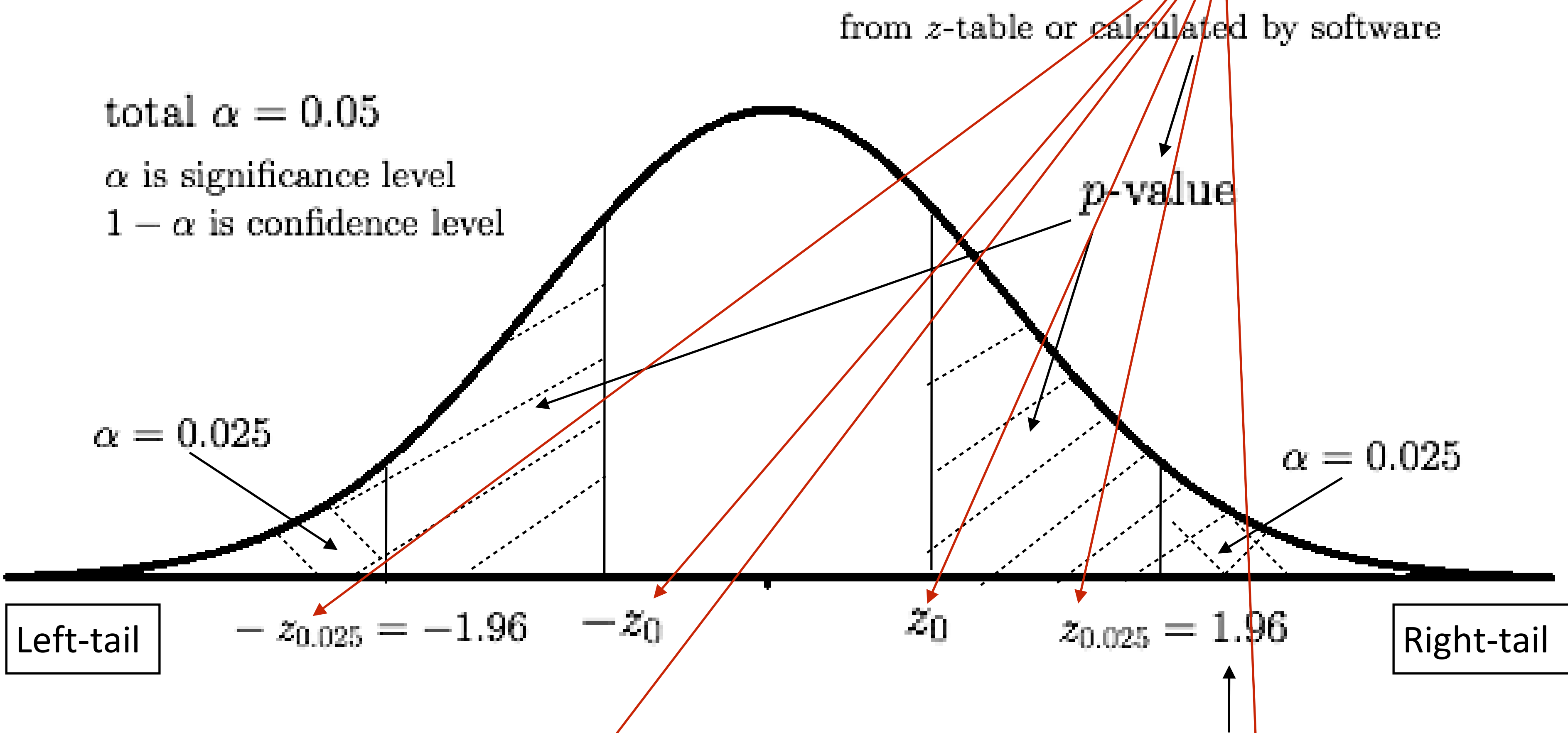
T-Test

A ***t*-test** is a statistical hypothesis test in which the test statistic follows a Student's *t*-distribution under the null hypothesis. It can be used to determine if a population mean can be accepted as a specified value or to determine if the means of two sets of data are significantly different from each other.

Two-tailed

$$H_0 : p = p_0, \quad H_1 : p \neq p_0$$

t-test: change z to t



from z-table

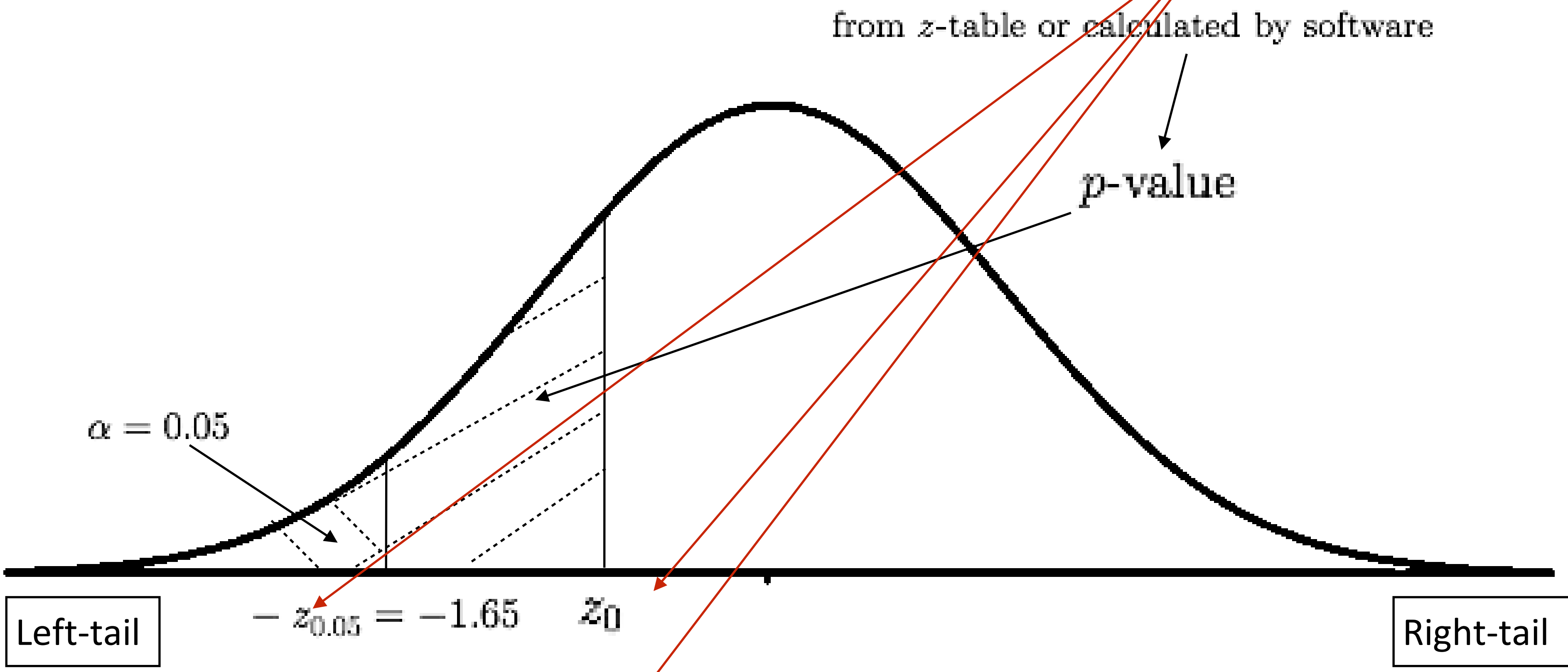
Accept H_0: if $|z_0| < z_{0.025}$ or p-value $> \alpha$

Reject H_0: if $|z_0| \geq z_{0.025}$ or p-value $\leq \alpha$

Left-tailed

$H_0 : p = p_0, \quad H_1 : p < p_0$

t-test: change z to t



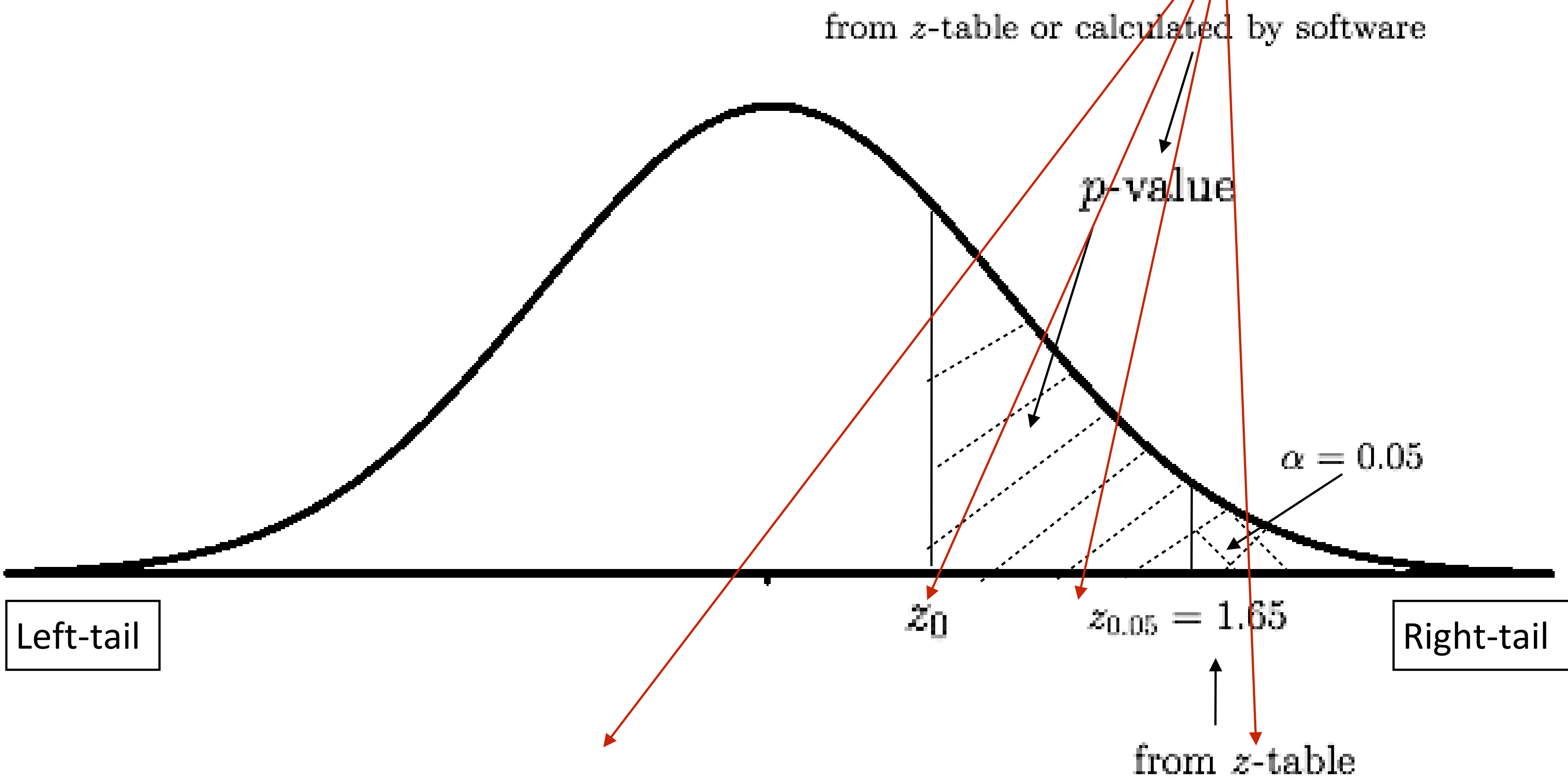
Accept H_0: if $z_0 > z_{0.05}$ or p-value $> \alpha$

Reject H_0: if $z_0 \leq z_{0.05}$ or p-value $\leq \alpha$

Right-tailed

$$H_0 : p = p_0, \quad H_1 : p > p_0$$

t-test: change z to t



Accept H_0: if $z_0 < z_{0.05}$ or p-value $> \alpha$

Reject H_0: if $z_0 \geq z_{0.05}$ or p-value $\leq \alpha$

One Sample T-Test

- The one sample t-test is a statistical procedure used to determine whether a sample of observations could have been generated by a population with a specific mean.

- Suppose you are interested in determining whether an assembly line produces laptop computers that weigh 2 kilograms. To test this hypothesis, you could collect a sample of laptop computers from the assembly line, measure their weights, and use one-sample t-test to carry out the test.

One Sample T-Test

- The procedure for a one sample t-test can be summed up in four steps. The symbols to be used are defined below:

Y = Random sample
y_i = The i^{th} observation in Y
n = The sample size
m_0 = The hypothesized value
$\overline{y}$ = The sample mean
$\hat{\sigma}$ = The sample standard deviation
T =The critical value of a t-distribution with $(n-1)$ degrees of freedom
t = The t-statistic (t-test statistic) for a one sample t-test
p = The p-value (probability value) for the t-statistic.

UNIVERSITY OF WOLLONGONG AUSTRALIA

One Sample T-Test

The four steps are listed below:

1. Calculate the sample mean.

$$\bar{y} = \frac{y_1 + y_2 + \cdots + y_n}{n}$$

2. Calculate the sample standard deviation.

$$\hat{\sigma} = \sqrt{\frac{(y_1 - \bar{y})^2 + (y_2 - \bar{y})^2 + \cdots + (y_n - \bar{y})^2}{n-1}}$$

3. Calculate the test statistic.

$$t = \frac{\bar{y} - m_0}{\hat{\sigma}/\sqrt{n}}$$

One Sample T-Test

4. Calculate the probability of observing the test statistic under the null hypothesis.

   - This value is obtained by comparing t to a t-distribution with $(n-1)$ degrees of freedom. This can be done by looking up the value in a table, such as those found in many statistical textbooks, or with statistical software for more accurate results.

$$p = 2 \cdot \Pr(T > |t|) \text{ (two-tailed)}$$
$$p = \Pr(T > t) \text{ (upper-tailed)}$$
$$p = \Pr(T < t) \text{ (lower-tailed)}$$

Once the assumptions have been verified and the calculations are complete, all that remains is to determine whether the results provide sufficient evidence to favour the alternative hypothesis over the null hypothesis.

Example

Moon illusion example: how large must moon be at zenith to appear equivalent in size to moon at horizon?

- $x = \text{size}_{zenith} / \text{size}_{horizon}$

- Null Hypothesis H_0: $\mu = 1$

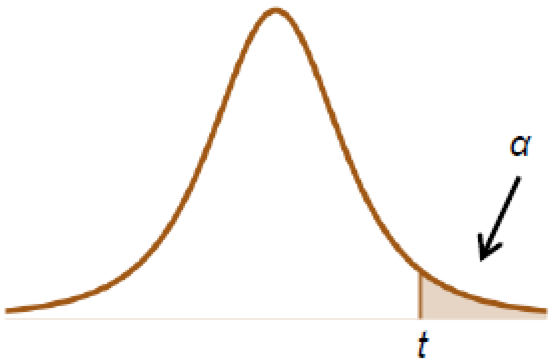
- Research Hypothesis H_1: $\mu \neq 1$

- $x = \{1.73, 1.06, 2.03, 1.40, 0.95, 1.13, 1.41, 1.73, 1.63, 1.56\}$

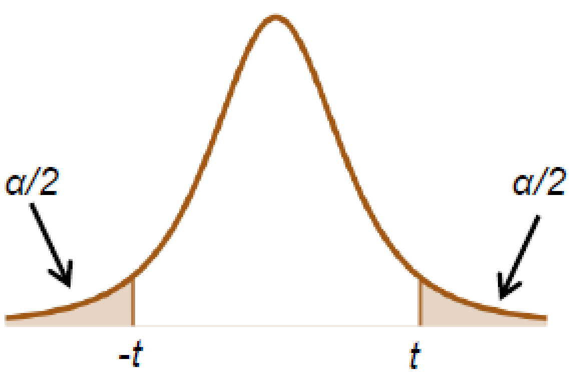
- Do we accept or reject the null hypothesis?
  - Assume a two-tailed test, with $\alpha = 0.05$

UNIVERSITY OF WOLLONGONG AUSTRALIA

t-Distribution Table



One-tailed test



Two-tailed test

	Level of significance for one-tailed test								
	0.25	0.2	0.15	0.1	0.05	0.025	0.01	0.005	0.0005
	Level of significance for two-tailed test								
df	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**	**0.05**	**0.02**	**0.01**	**0.001**
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	636.619
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.599
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	12.924
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.403	2.678	3.496
100	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.390

1. Compute sample mean and SD

	x	x^2
	1.73	2.99
	1.06	1.12
	2.03	4.12
	1.40	1.96
	0.95	0.90
	1.13	1.28
	1.41	1.99
	1.73	2.99
	1.63	2.66
	1.56	2.43
sum	14.63	22.45

$$\bar{x} = \frac{\sum_i x}{n} = \frac{14.63}{10} = 1.463$$

$$SS = \sum x^2 - \frac{\left(\sum x\right)^2}{n} = 22.45 - \frac{14.63^2}{10} = 1.046$$

$$\sigma = \sqrt{\frac{SS}{n-1}} = \sqrt{\frac{1.046}{9}} = \sqrt{0.116} = 0.341$$

2. Use these values to compute *t*-statistic

(Remember, μ=1 under H_0)

$$t(df) = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$t(n-1) = \frac{1.463 - 1}{\frac{0.341}{\sqrt{10}}}$$

$$t(9) = \frac{0.463}{0.108} = 4.29$$

$$t_{0.05} = 2.262$$

$$4.29 > 2.262; \text{ reject } H_0$$

UNIVERSITY OF WOLLONGONG AUSTRALIA

Two Sample T-Test

- The two-sample t-test is used to determine if two population means are equal.
- A common application is to test if a new process or treatment is superior to a current process or treatment.

There are several variations on this test.

1. The data may either be paired or not paired. By paired, we mean that there is a one-to-one correspondence between the values in the two samples.

   That is, if $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ are the two samples, then X_i corresponds to Y_i. For paired samples, the difference $X_i - Y_i$ is usually calculated. For unpaired samples, the sample sizes for the two samples may or may not be equal. The formulas for paired data are somewhat simpler than the formulas for unpaired data.

2. The variances of the two samples may be assumed to be equal or unequal. Equal variances yields somewhat simpler formulas, although with computers this is no longer a significant issue.

3. In some applications, you may want to adopt a new process or treatment only if it exceeds the current treatment by some threshold. In this case, we can state the null hypothesis in the form that the difference between the two populations means is equal to some constant $\mu_1 - \mu_2 = d_0$ where the constant is the desired threshold.

Definition

The two-sample t-test for unpaired data is defined as:

H$_0$: $\mu_1 = \mu_2$

H$_a$: $\mu_1 \neq \mu_2$

Test Statistic: $T = \frac{\bar{Y_1} - \bar{Y_2}}{\sqrt{s_1^2/N_1 + s_2^2/N_2}}$

where N_1 and N_2 are the sample sizes, $\bar{Y}_1$ and $\bar{Y}_2$ are the sample means, and s_1^2 and s_2^2 are the sample variances.

If equal variances are assumed, then the formula reduces to:

$$T = \frac{\bar{Y_1} - \bar{Y_2}}{s_p\sqrt{1/N_1 + 1/N_2}}$$

where

$$s_p^2 = \frac{(N_1-1)s_1^2 + (N_2-1)s_2^2}{N_1 + N_2 - 2}$$

Significance Level:	α.
Critical Region:	Reject the null hypothesis that the two means are equal if $$\lvert T \rvert > t_{1-\alpha/2,\nu}$$ where $t_{1-\alpha/2,\nu}$ is the critical value of the *t* distribution with ν degrees of freedom where $$\upsilon = \frac{(s_1^2/N_1 + s_2^2/N_2)^2}{(s_1^2/N_1)^2/(N_1-1) + (s_2^2/N_2)^2/(N_2-1)}$$ If equal variances are assumed, then $\nu = N_1 + N_2 - 2$

Example of Two Sample T-Test

The following two-sample *t*-test was generated for the data.txt data set. The data set contains miles per gallon for U.S. cars (sample 1) and for Japanese cars (sample 2); the summary statistics for each sample are shown below.

```
SAMPLE 1:
    NUMBER OF OBSERVATIONS      = 249
    MEAN                        =  20.14458
    STANDARD DEVIATION          =   6.41470
    STANDARD ERROR OF THE MEAN  =   0.40652

SAMPLE 2:
    NUMBER OF OBSERVATIONS      = 79
    MEAN                        = 30.48101
    STANDARD DEVIATION          =  6.10771
    STANDARD ERROR OF THE MEAN  =  0.68717
```

http://www.itl.nist.gov/div898/handbook/eda/section3/eda353.htm

- We are testing the hypothesis that the population means are equal for the two samples. We assume that the variances for the two samples are equal.

H_0: $\mu_1 = \mu_2$
H_a: $\mu_1 \neq \mu_2$

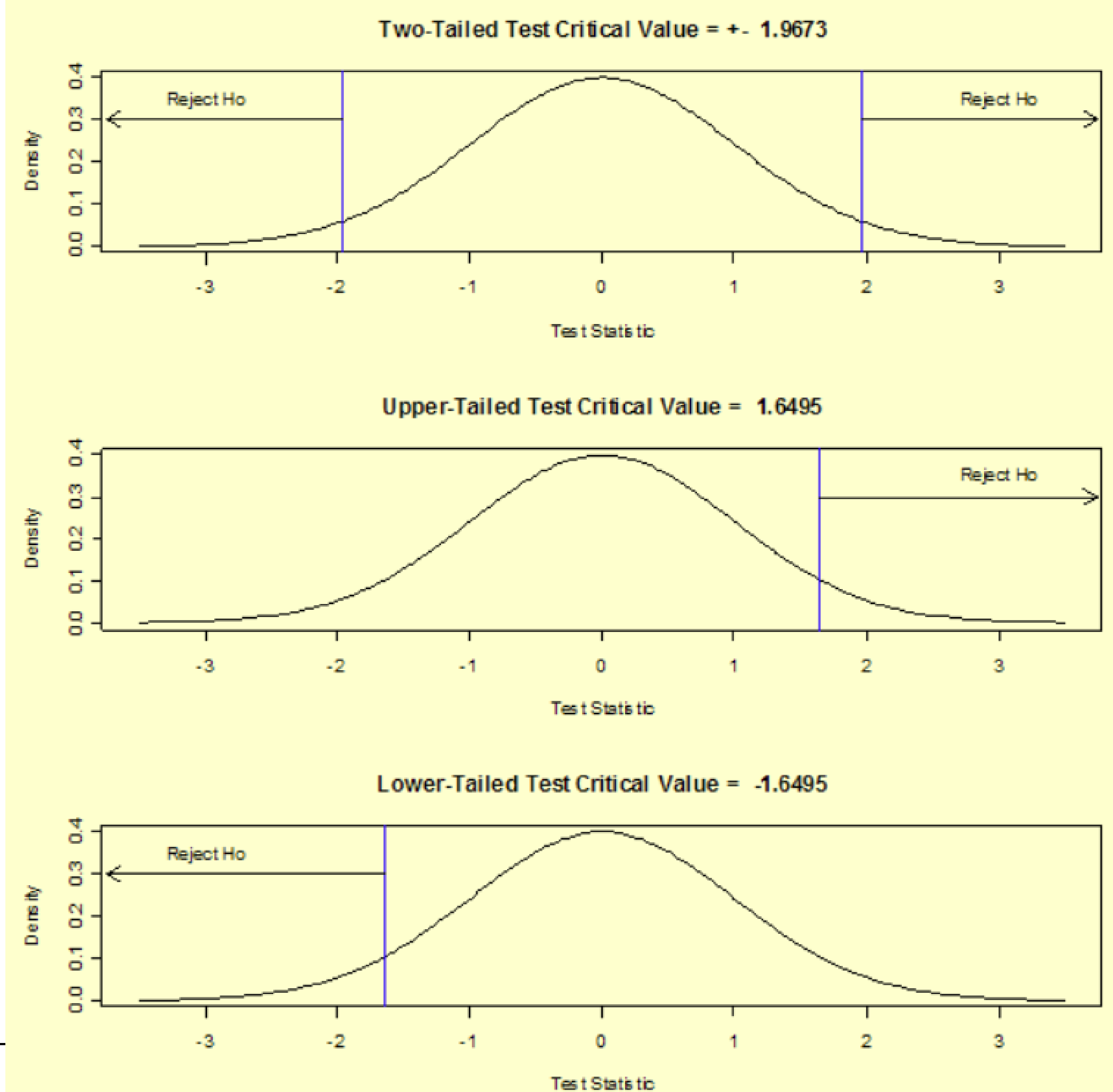
Test statistic: $T = -12.62059$
Pooled standard deviation: $s_p = 6.34260$
Degrees of freedom: $v = 326$
Significance level: $\alpha = 0.05$
Critical value (upper tail): $t_{1-\alpha/2,v} = 1.9673$
Critical region: Reject H_0 if $|T| > 1.9673$

The absolute value of the test statistic for our example, 12.62059, is greater than the critical value of 1.9673, so we reject the null hypothesis and conclude that the two population means are different at the 0.05 significance level.

In general, there are three possible alternative hypotheses and rejection regions for the one- sample *t*-test:

Alternative Hypothesis	**Rejection Region**
H_a: $\mu_1 \neq \mu_2$	$\|T\| > t_{1-\alpha/2,\nu}$
H_a: $\mu_1 > \mu_2$	$T > t_{1-\alpha,\nu}$
H_a: $\mu_1 < \mu_2$	$T < t_{\alpha,\nu}$

For our two-tailed t-test, the critical value is $t_{1\text{-}\alpha/2,v} = 1.9673$, where $\alpha = 0.05$ and $v = 326$. If we were to perform an upper, one-tailed test, the critical value would be $t_{1\text{-}\alpha,v} = 1.6495$. The rejection regions for three possible alternative hypotheses using our example data are shown on the next slide


Two-Tailed Test Critical Value = +- 1.9673
Reject Ho
Reject Ho
Density
Test Statistic
Upper-Tailed Test Critical Value = 1.6495
Reject Ho
Density
Test Statistic
Lower-Tailed Test Critical Value = -1.6495
Reject Ho
Density
Test Statistic

Paired T-Test

- The same object with two test results

$$x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6,\ x_7,\ x_8,\ x_9$$
$$y_1,\ y_2,\ y_3,\ y_4,\ y_5,\ y_6,\ y_7,\ y_8,\ y_9$$

$$x_i - y_i = z_1,\ z_2,\ z_3,\ z_4,\ z_5,\ z_6,\ z_7,\ z_8,\ z_9$$

Use one sample t-test:

- Determine whether the mean of the differences between two paired samples differs from 0 (or a target value)
- Calculate a range of values that is likely to include the population mean of the differences

Paired T-Test

For example, suppose managers at a fitness facility want to determine whether their weight-loss program is effective. Because the "before" and "after" samples measure the same subjects, a paired t-test is the most appropriate analysis.

The paired t-test calculates the difference within each before-and-after pair of measurements, determines the mean of these changes, and reports whether this mean of the differences is statistically significant.

A paired t-test can be more powerful than a 2-sample t-test because the latter includes additional variation occurring from the independence of the observations. A paired t-test is not subject to this variation because the paired observations are dependent.