INDEX

1. Hypothesis Statement

If you are going to propose a hypothesis, it's customary to write a statement. Your statement will look like this:

"If I…(do this to an independent variable)….then (this will happen to the dependent variable)."

For example

If I (decrease the amount of water given to herbs) then (the herbs will increase in size)

If I (give patients counseling in addition to medication) then (their overall depression scale will decrease)

If I (give exams at noon instead of 7) then (student test scores will improve)

If I (look in this certain location) then (I am more likely to find new species)

Hypothesis Statement

A good hypothesis statement should

Include an “if” and “then” statement (according to the University of California).

- Include both the independent and dependent variables
- Be testable by experiment, survey or other scientifically sound technique
- Be based on information in prior research (either yours or someone else’s)
- Have design criteria (for engineering or programming projects)

2. Points to be considered while formulating Hypothesis

- **Hypothesis should be clear and precise**
- **Hypothesis should be capable of being tested**
- **Hypothesis should state relationship between variables**
- **Hypothesis should be limited in scope and must be specific**
- **Hypothesis should be stated as far as possible in most simple terms so that the same is easily understandable by all concerned**
- **Hypothesis should be amenable to testing within a reasonable time**
- **Hypothesis must explain empirical reference**

3. Hypothesis testing

Hypotheses are theoretical guesses based on limited knowledge; they need to be tested. Hypothesis testing is a decision-making process to evaluate claims about a population.

We use various statistical analysis to test hypotheses and answer research questions. In formal hypothesis testing, we test the null hypothesis and usually want to reject the null because rejection of the null indirectly supports the alternative hypothesis to the null, the one we deduce from theory as a tentative explanation. Thus, a hypothesis test mutually exclusive statements about a population to determine which statement is best supported by the sample data.

Hypothesis Testing

The main purpose of statistics is to test a hypothesis. For example, you might run an experiment and find that a certain drug is effective at treating headaches. But if you can't repeat that experiment, no one will take your results seriously.
A hypothesis is an educated guess about something in the world around you. It should be testable, either by experiment or observation. For example

A new medicine you think might work

A way of teaching you think might be better

A possible location of new species

A fairer way to administer standardized tests

Hypothesis Testing

$$z = \frac{\hat{p} - p}{\sqrt{p\,q/n}}$$

Hypothesis testing in statistics is a way for you to test the results of a survey or experiment to see if you have meaningful results. You're basically testing whether your results are valid by figuring out the odds that your results have happened by chance. If your results may have happened by chance, the experiment won't be repeatable and so has little use.

Hypothesis testing can be one of the most confusing aspects for students, mostly because before you can even perform a test, you have to know what your null hypothesis is. Often, those tricky word problems that you are faced with, can be difficult to decipher.

- **Figure out your null hypothesis**
- **State your null hypothesis**
- **Choose what kind of test you need to perform**
- **Either support or reject the null hypothesis**

4. Procedure of Hypothesis Testing

a. **Making a formal statement:** The step consists in making a formal statement of the null hypothesis (H0) and also of the alternative hypothesis (Ha or H1). This means that hypotheses should be clearly stated, considering the nature of the research problem.

b. **Selecting a significance level:** The hypotheses are tested on a pre-determined level of significance and as such the same should be specified. Generally, in practice, either 5% level or 1% level is adopted for the purpose.

c. **Deciding the distribution to use:** After deciding the level of significance, the next step in hypothesis testing is to determine the appropriate sampling distribution. The choice generally remains between normal distribution and the t-distribution.

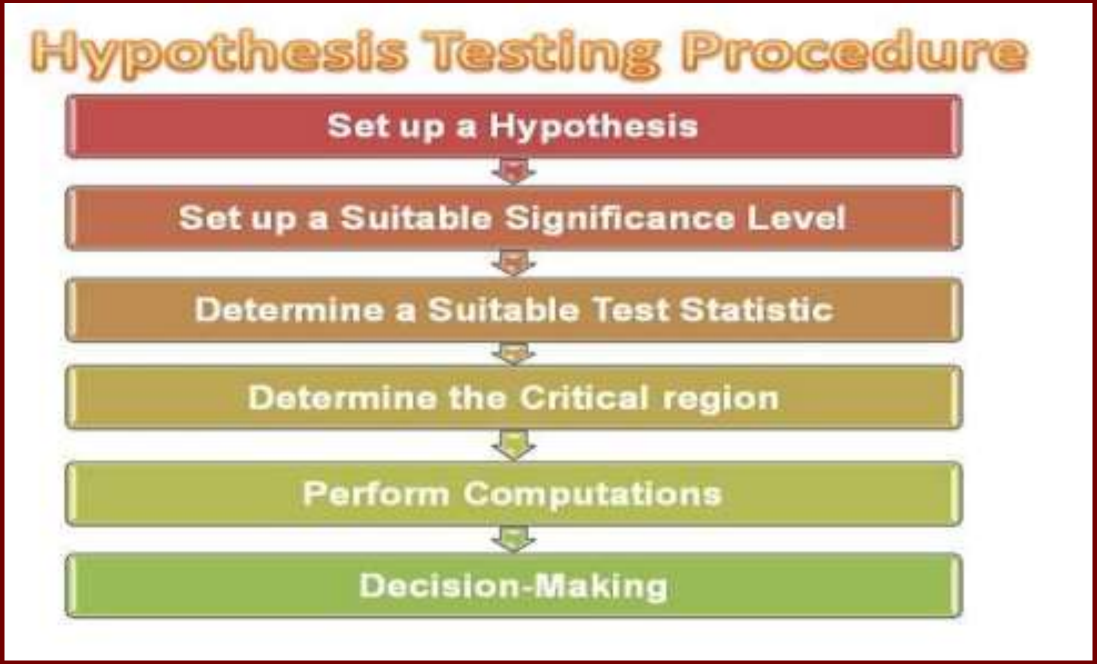
Procedure of Hypothesis Testing

d. Selecting a random sample and computing an appropriate value: Another step is to select a random sample(s) and compute an appropriate value from the sample data concerning the test statistic, utilizing the relevant distribution. In other words, draw a sample to furnish empirical data.

e. Calculation of the probability: One has then to calculate the probability that the sample result would diverge as widely as it has from expectations, if the null hypothesis were in fact true.

Procedure of Hypothesis Testing

f. Comparing the probability and Decision making: Yet another step consists in comparing the probability thus calculated with the specified value for α, the significance level. If the calculated probability is equal to or smaller than the α value in case of one-tailed test (and α /2 in case of two-tailed test), then reject the null hypothesis (i.e., accept the alternative hypothesis), but if the calculated probability is greater, then accept the null hypothesis.



5. There are 2 types of Hypothesis Testing

Different types of hypothesis are involved in finding whether the tested samples test positive for a hypothesis or not. In this segment, we shall discover the different types of hypotheses and understand the role they play in hypothesis testing:

a) Alternative Hypothesis

b) Null Hypothesis

a) Alternative Hypothesis

Alternative Hypothesis (H1) or the research hypothesis states that there is a relationship between two variables, where one variable affects the other. The alternative hypothesis is the main driving force for hypothesis testing.

It implies that the two variables are related to each other and the relationship that exists between them is not due to chance or coincidence.

When the process of hypothesis testing is carried out, the alternative hypothesis is the main subject of the testing process. The analyst intends to test the alternative hypothesis and verifies its plausibility.

A statistical hypothesis that states the existence of a difference between a parameter and a specific value, or states that there is a difference between two parameters. Alternative hypothesis is created in a negative meaning of the null hypothesis. Suppose we want to test the hypothesis that the population mean (μ) is equal to the hypothesized mean (μH0) = 100. Then we would say that the null hypothesis is that the population mean is equal to the hypothesized mean 100 and symbolically we can

express as:

H0: μ = μ H0 = 100

If our sample results do not support this null hypothesis, we should conclude that something else is true. What we conclude rejecting the null hypothesis is known as alternative hypothesis. The set of alternatives to the null hypothesis is referred to as the alternative hypothesis. If we accept H0, then we are rejecting H1 and if we reject H0, then we are accepting H1.

For H0: μ = μHo =100, we may consider three possible alternative hypotheses as follows:

Alternative hypothesis	To be read as follows
$H_1: \mu \neq \mu_{Ho}$	(The alternative hypothesis is that the population mean is not equal to 100 i.e., it may be more or less than 100)
$H_1: \mu > \mu_{Ho}$	(The alternative hypothesis is that the population mean is greater than 100)
$H_1: \mu < \mu_{Ho}$	(The alternative hypothesis is that the population mean is less than 100)

The null hypothesis and the alternative hypothesis are chosen before the sample is drawn. The researcher must avoid the error of deriving hypotheses from the data that he/she collects and then testing the hypotheses from the same data.

In the choice of null hypothesis, the following considerations are usually kept in view

I. Alternative hypothesis is usually the one which one wishes to prove and the null hypothesis is the one which one wishes to disprove. Thus, a null hypothesis represents the hypothesis we are trying to reject, and alternative hypothesis represents all other possibilities.

II. Null hypotheses should always be specific hypothesis i.e., it should not state about or approximately a certain value.

III. In testing hypothesis, there are two possible outcomes

- **Reject H0 and accept H1 because of sufficient evidence in the sample in favor of H1;**
- **Do not reject H0 because of insufficient evidence to support H1.**

6. Null Hypothesis

If you trace back the history of science, the null hypothesis is always the accepted fact. Simple examples of null hypotheses that are generally accepted as being true are:

- DNA is shaped like a double helix.
- There are 8 planets in the solar system (excluding Pluto).
- Taking Vioxx can increase your risk of heart problems (a drug now taken off the market).

You won't be required to actually perform a real experiment or survey in elementary statistics (or even disprove a fact like "Pluto is a planet"!), so you'll be given word problems from real-life situations. You'll need to figure out what your hypothesis is from the problem. This can be a little trickier than just figuring out what the accepted fact is. With word problems, you are looking to find a fact that is nullifiable (i.e. something you can reject).

7. Basic concepts concerning testing of hypotheses

A. The level of significance: This is a very important concept in the context of hypothesis testing. It is always some percentage (usually 5%) which should be chosen with great care, thought and reason. In case we take the significance level at 5 %, then this implies that H0 will be rejected when the sampling result (i.e., observed evidence) has a less than 0.05 probability of occurring if H0 is true. In other words, the 5 % level of significance means that researcher is willing to take as much as a 5 per cent risk of rejecting the null hypothesis when it (H0) happens to be true. **The significance level is the maximum value of the probability of rejecting H0 when it is true & is usually determined in advance before hypothesis testing.**

Basic concepts concerning testing of hypotheses

B. Decision rule or Test of Hypothesis: A decision rule is a procedure that the researcher uses to decide whether to accept or reject the null hypothesis. The decision rule is a statement that tells under what circumstances to reject the null hypothesis. The decision rule is based on specific values of the test statistic (e.g., reject H0 if Calculated value > table value at the same level of significance).

C. Types of Error: In the context of testing of hypotheses, there are basically two types of errors we can make.

a. Type 1 error: To reject the null hypothesis when it is true is to make what is known as a type I error. The level at which a result is declared significant is known as the type I error rate, often denoted by α.

b. Type II error: If we do not reject the null hypothesis when in fact there is a difference between the groups, we make what is known as a type II error. The type II error rate is often denoted as β.

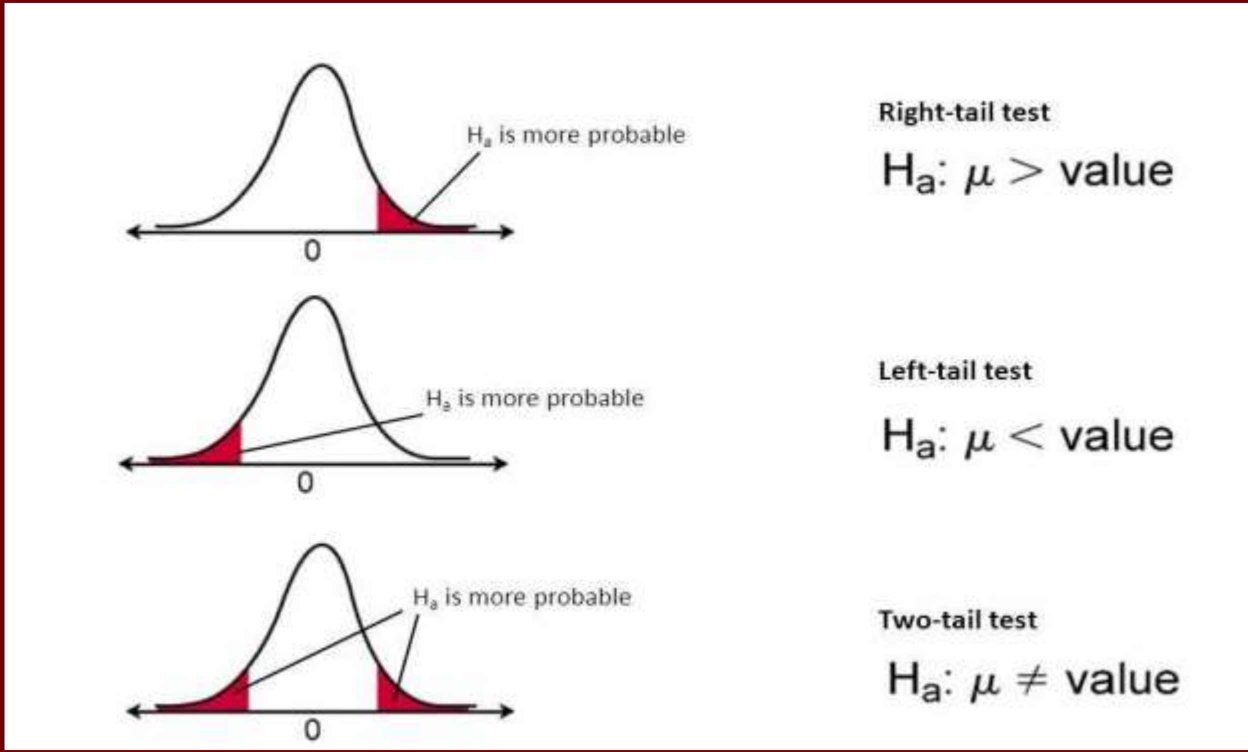
Basic concepts concerning testing of hypotheses

In a tabular form the said two errors can be presented as follows:

Particulars	Decision	
	Accept H_0	Reject H_0
H_0 (True)	Correct Decision	Type I error (α error)
H_0 (False)	Type II error (β error)	Correct decision

One- tailed and Two-tailed Tests: A test of statistical hypothesis, where the region of rejection is on only one side of the sampling distribution, is called a one tailed test. For example, suppose the null hypothesis states that the mean is less than or equal to 10. The alternative hypothesis would be that the mean is greater than 10. The region of rejection would consist of a range of numbers located on the right side of sampling distribution i.e., a set of numbers greater than 10.

Basic concepts concerning testing of hypotheses



A test of statistical hypothesis, where the region of rejection is on both sides of the sampling distribution, is called a two-tailed test. For example, suppose the null hypothesis states that the mean is equal to 10. The alternative hypothesis would be that the mean is less than 10 or greater than 10. The region of rejection would consist of a range of numbers located on both sides of sampling distribution; i.e., the region of rejection would consist partly of numbers that were less than 10 and partly of numbers that were greater than 10.

8. Tests of Hypotheses

Hypothesis testing determines the validity of the assumption (technically described as null hypothesis) with a view to choose between two conflicting hypotheses about the value of a population parameter. Hypothesis testing helps to decide on the basis of a sample data, whether a hypothesis about the population is likely to be true or false. Statisticians have developed several tests of hypotheses (also known as the tests of significance) for the purpose of testing of hypotheses which can be classified as:

a) Parametric tests or standard tests of hypotheses

b) Non-parametric tests or distribution-free test of hypotheses

Parametric tests usually assume certain properties of the parent population from which we draw samples. Assumptions like observations come from a normal population, sample size is large, assumptions about the population parameters like mean, variance, etc., must hold good before parametric tests can be used.

There are situations when the researcher cannot or does not want to make such assumptions. In such situations we use statistical methods for testing hypotheses which are called non-parametric tests because such tests do not depend on any assumption about the parameters of the parent population. Besides, most non-parametric tests assume only nominal or ordinal data, whereas parametric tests require measurement equivalent to at least an interval scale. As a result, non-parametric tests need more observations than parametric tests to achieve the same size of Type I and Type II errors.

9. Important parametric tests

The important parametric tests are:

a. **z-test**

b. **t-test**

c. **F-test**

All these tests are based on the assumption of normality i.e., the source of data is considered to be normally distributed.

Important parametric tests

a. z- test: It is based on the normal probability distribution and is used for judging the significance of several statistical measures, particularly the mean. This test is used even when binomial distribution or t-distribution is applicable on the presumption that such a distribution tends to approximate normal distribution as 'n' becomes larger. z-test is generally used for comparing the mean of a sample to some hypothesized mean for the population in case of large sample, or when population variance is known. z-test is used for judging he significance of difference between means of two independent samples in case of large samples, or when population variance is known. z-test is also used for comparing the sample proportion to a theoretical value of population proportion or for judging the difference in proportions of two independent samples when 'n' happens to be large. Besides, this test may be used for judging the significance of median, mode, coefficient of correlation and several other measures.

b. t- test: It is based on t-distribution and is considered an appropriate test for judging the significance of a sample mean or for judging the significance of difference between the means of two samples in case of small sample(s) when population variance is not known (in which case we use variance of the sample as an estimate of the population variance). In case two samples are related, we use paired t-test (or what is known as difference test) for judging the significance of the mean of difference between the two related samples. It can also be used for judging the significance of the coefficients of simple & partial correlations.

c. F-test: It is based on F-distribution and is used to compare the variance of the two independent samples. This test is used in the context of analysis of variance (ANOVA) for judging the significance of more than two sample means at one and the same time. It is also used for judging the significance of multiple correlation coefficients.

10. Non parametric tests

Non parametric tests are used when the data isn't normal. Therefore, the key is to figure out if you have normally distributed data. The only non-parametric test you are likely to come across in elementary stats is the chi-square test. However, there are several others. For example: the Kruskal Willis test is the non-parametric alternative to the One-way ANOVA and the Mann Whitney is the non- parametric alternative to the two-sample t test.

Hypothesis testing examples #1: basic example

A researcher thinks that if knee surgery patients go to physical therapy twice a week (instead of 3 times), their recovery period will be longer. Average recovery times for knee surgery patients is 8.2 weeks.

The hypothesis statement in this question is that the researcher believes the average recovery time is more than 8.2 weeks. It can be written in mathematical terms as:

$$H1: \mu > 8.2$$

Next, you'll need to state the null hypothesis (See: How to state the null hypothesis). That's what will happen if the researcher is wrong. In the above example, if the researcher is wrong then the recovery time is less than or equal to 8.2 weeks. In math, that's:

$$H0\ \mu \leq 8.2$$

11. Rejecting the null hypothesis

Ten or so years ago, we believed that there were 9 planets in the solar system. Pluto was demoted as a planet in 2006. The null hypothesis of "Pluto is a planet" was replaced by "Pluto is not a planet." Of course, rejecting the null hypothesis isn't always that easy — the hard part is usually figuring out what your null hypothesis is in the first place.

Hypothesis testing examples (one sample Z test)

The test isn't used very often (because we rarely know the actual population standard deviation). However, it's a good idea to understand how it works as it's one of the simplest tests you can perform in hypothesis testing. In English class you got to learn the basics (like grammar and spelling) before you could write a story; think of one sample z tests as the foundation for understanding more complex hypothesis testing. This page contains two hypothesis testing examples for one sample z-tests.

Hypothesis testing examples: Mean (Using TI 83)

You can use the TI 83 calculator for hypothesis testing, but the calculator won't figure out the null and alternate hypotheses; that's up to you to read the question and input it into the calculator.

Example problem: A sample of 200 people has a mean age of 21 with a population standard deviation (σ) of 5. Test the hypothesis that the population mean is 18.9 at $\alpha = 0.05$.

Step 1: State the null hypothesis. In this case, the null hypothesis is that the population mean is 18.9, so we write:

H0: $\mu = 18.9$

Step 2: State the alternative hypothesis. We want to know if our sample, which has a mean of 21 instead of 18.9, really is different from the population, therefore our alternate hypothesis:

H1: $\mu \neq 18.9$

Hypothesis testing examples: Mean (Using TI 83)

Step 3: Press Stat then press the right arrow twice to select TESTS.

Step 4: Press 1 to select 1:Z-Test…. Press ENTER.

Step 5: Use the right arrow to select Stats.

Step 6: Enter the data from the problem:

μ0: 18.9

σ: 5

x: 21

n: 200

μ: ≠μ0

Hypothesis testing examples: Mean (Using TI 83)

Step 7: Arrow down to Calculate and press ENTER. The calculator shows the p-value:

p = 2.87 × 10-9

This is smaller than our alpha value of .05. That means we should reject the null hypothesis.

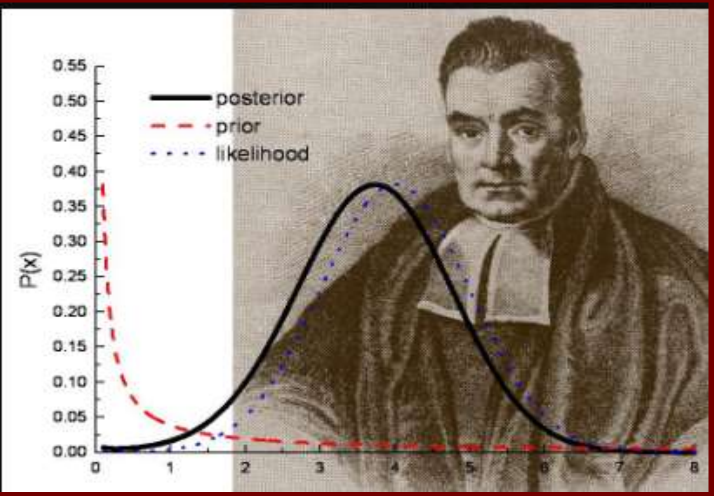
Hypothesis testing examples: Mean (Using TI 83)

Step 7: Arrow down to Calculate and press ENTER. The calculator shows the p-value:

p = 2.87 × 10-9

This is smaller than our alpha value of .05. That means we should reject the null hypothesis.

12. Bayesian Hypothesis Testing:



Bayesian hypothesis testing helps to answer the question: Can the results from a test or survey be repeated?

Why do we care if a test can be repeated? Let's say twenty people in the same village came down with leukemia. A group of researchers find that cell-phone towers are to blame. However, a second study found that cell-phone towers had nothing to do with the cancer cluster in the village.

In fact, they found that the cancers were completely random. If that sounds impossible, it actually can happen! Clusters of cancer can happen simply by chance. There could be many reasons why the first study was faulty. One of the main reasons could be that they just didn't take into account that sometimes things happen randomly and we just don't know why.

13. P Values

It’s good science to let people know if your study results are solid, or if they could have happened by chance. The usual way of doing this is to test your results with a p-value. A p value is a number that you get by running a hypothesis test on your data. A P value of 0.05 (5%) or less is usually enough to claim that your results are repeatable. However, there’s another way to test the validity of your results: Bayesian Hypothesis testing. This type of testing gives you another way to test the strength of your results.

Bayesian Hypothesis Testing: Traditional testing (the type you probably came across in elementary stats or AP stats) is called Non-Bayesian. It is how often an outcome happens over repeated runs of the experiment. It’s an objective view of whether an experiment is repeatable.

14. Differences between traditional and Bayesian Hypothesis testing

Traditional testing (Non Bayesian) requires you to repeat sampling over and over, while Bayesian testing does not. The main different between the two is in the first step of testing: stating a probability model. In Bayesian testing you add prior knowledge to this step. It also requires use of a posterior probability, which is the conditional probability given to a random event after all the evidence is considered.

15. Arguments for Bayesian testing

Many researchers think that it is a better alternative to traditional testing, because it:

- **Includes prior knowledge about the data.**
- **Takes into account personal beliefs about the results.**

Arguments against

- **Including prior data or knowledge isn't justifiable.**
- **It is difficult to calculate compared to non-Bayesian testing.**