###### **Christiana Ogboali IDAM**

###### **Student’s ID: UM76770HHU85962**

###### **COURSE TITLE: Numbers of life: Hypothesis Testing**

###### **ATLANTIC INTERNATIONAL UNIVERSITY**

###### **JUNE 2, 2022**

**INTRODUCTION**

According to Bren 2022, Statistics is the science concerned with developing and studying methods for collecting, analysing, interpreting and presenting empirical data. Statistics is a highly interdisciplinary field; research in statistics finds applicability in virtually all scientific fields and research questions in the various scientific fields motivate the development of new statistical methods and theory. In developing methods and studying the theory that underlies the methods statisticians draw on a variety of mathematical and computational tools.

Two fundamental ideas in the field of statistics are uncertainty and variation. There are many situations that we encounter in science (or more generally in life) in which the outcome is uncertain. In some cases the uncertainty is because the outcome in question is not determined yet (e.g., we may not know whether it will rain tomorrow) while in other cases the uncertainty is because although the outcome has been determined already we are not aware of it (e.g., we may not know whether we passed a particular exam).

Probability is a mathematical language used to discuss uncertain events and probability plays a key role in statistics. Any measurement or data collection effort is subject to a number of sources of variation. By this we mean that if the same measurement were repeated, then the answer would likely change. Statisticians attempt to understand and control (where possible) the sources of variation in any situation.

A [hypothesis](https://www.thoughtco.com/definition-of-hypothesis-605234) as stated by Helmenstine 2020, (plural hypotheses) is a proposed explanation for an observation. The definition depends on the subject. In science, a hypothesis is part of the scientific method. It is a prediction or explanation that is tested by an experiment. Observations and experiments may disprove a scientific hypothesis, but can never entirely prove one while in the study of logic, a hypothesis is an if-then proposition, typically written in the form, "If X, then Y." In common usage, a hypothesis is simply a proposed explanation or prediction, which may or may not be tested.

# **1.** Hypothesis testing is a statistical process whereby the researcher tests an assumption in relation to a population parameter. The method preferred in the analysis is dependent on the nature of the data collected and the reason for the analysis. Hypothesis testing uses sample data to assess the plausibility of a hypothesis.

# Step-by-step explanation

# An analyst tests a statistical sample so as to provide evidence on the plausibility of the null hypothesis. S/he measures and examines a random sample of the population being analyzed while testing a hypothesis. In most cases, analysts use a random population sample while testing the null hypothesis and the alternative hypothesis. The hypothesis of equality between population parameters is the null hypothesis while the alternative hypothesis is the opposite of a null hypothesis. One of the hypotheses is always true.

# **Steps of hypothesis testing.**

# Stating the null and alternative hypotheses so that only one can be correct.

# Formulating an analysis plan outlining how to evaluate the collected data.

# Implementing the plan and physically analyzing the sample data.

# Analyzing the outcomes and either rejecting the null hypothesis or stating that the null hypothesis is plausible according to the data.

# Parametric testing is designed to provide the data required for the analysis in parametric statistics which assumes that some information about the population under study is already known (probability distribution). On the other hand, a non-parametric testing is used when the gathered data is not expected to fit a normal distribution curve or ordinal data. Therefore, a non-parametric test does not assume anything about the population in question, unlike a parametric test. Furthermore, a parametric test is based on probabilistic distribution whereas a non-parametric test is not based on anything i.e it is completely arbitrary making it more flexible and easier to fit the hypothesis of the obtained data.

# 2. A null hypothesis is a research statement that states that there is no significant difference between two or more specified populations.

# **Step-by-step explanation**

# A null hypothesis is a research statement that states that there is no significant difference between specified populations and therefore, any observed difference is due to sampling or experimental error.

# 3. Alternative hypotheses are used in statistical hypothesis testing to distinguish between new and old theories about what's happening in the data. In most cases, this is in line with the overall trend. Because it is based on existing literature, the research hypothesis to review, to look into the past, etc.

# 4. Hypothesis tests employ an alpha level, also referred to as the significance level, to determine the likelihood that a decision will be incorrect if the null hypothesis is correct. Most of these tests have an alpha level of 5 percent for soft sciences (society) and 1 percent for hard sciences (natural sciences). Type 1 and Type 2 errors are possible in hypothesis tests. Type 1 supports the alternative hypothesis when the null hypothesis is true, whereas type 2 does not support the alternative hypothesis when it is true. The alternative hypothesis proposes that the population parameters are larger, smaller, or otherwise different from what was assumed in the null hypothesis. Thus, the smaller the alpha level, the smaller the area to reject the null hypothesis. A type 2 error occurs when the null hypothesis is rejected in a small area despite the fact that it should not be the case. That is, the more researchers/scientists try to avoid a type 1 error, the more likely a type 2 error will occur. Scientists have discovered that an alpha level of 5% is a good compromise between these two issues. This means that there is a 95% probability that the true population value is within the confidence interval around the sample results. The best estimate of the population value is the sample value, because scientists only use a sample and not the entire population, and thus results may be inaccurate. The error's most likely size is estimated using the confidence interval.

# Assuming that there is no difference between the sample mean and population mean, the null hypothesis is supported by an alpha of 0.5, which means that there is a 5% chance that this conclusion will be incorrect. Because alpha levels were found to be a good balance between the two types of hypothesis errors, I support them. If a child's education and the teacher assigned to them would be successful 95 times out of 100, the alpha level must be subtracted from 1. For a 95 percent confidence level in an analysis, the alpha level will be 1 -.95 = 5% if only one tail was tested (Frost, n. d). Thus, there is only a 5% chance of failure/risks for the hypothesis tests to be deemed successful.

# 5. A population value's confidence interval (CI) is a range of possible values in which the value is likely to fall. It is commonly expressed as a percentage where the population mean falls somewhere between the upper and lower bounds.

# 6. When conducting hypothesis testing, a critical value is a point on the test distribution that is compared to the test statistic to determine whether or not to reject the null hypothesis. If your test statistic's absolute value is greater than the critical value, you can declare statistical significance and reject the null hypothesis. Critical values correspond to α, so their values become fixed when you choose the test's α.

# 7. The test statistic is a number derived from a statistical test of a hypothesis, and it is used as a measure of validity. That statistical test's null hypothesis shows how closely your observed data matches the predicted distribution. Use the p-value of your results as a guide to whether or not to reject your null hypothesis. Null hypothesis assumes correctness until evidence points in the opposite direction in the end, there are only two possible outcomes following a hypothesis test. If your p-value is less than or equal to your significance level, the null hypothesis is rejected as a possibility. Alternative hypothesis supported by data.

# 8. If the sample data is in the critical region, the values are unlikely to happen if the null is true; the sample is not consistent with the null, and we reject the null hypothesis

# 9. In statistical analysis, a type I error is the rejection of a true null hypothesis, while a type II error describes the error that occurs when one fails to reject a null hypothesis that is actually false. Despite the fact that it is not a random occurrence, the error rules out the alternative theory.

# While

# A type II error is a statistical term used in hypothesis testing to describe the error that occurs when one fails to reject a null hypothesis that is in fact false. A type II error, also known as an error of omission, results in a false negative. Some tests may report negative results, even though the patient is actually infected with the disease. This is a type II error because we accept the test results as negative, despite the fact that they are incorrect.

**CONCLUSION**

# The hypothesis of equality between population parameters is the null hypothesis while the alternative hypothesis is the opposite of a null hypothesis. One of the hypotheses is always true. A parametric test is based on probabilistic distribution whereas a non-parametric test is not based on anything i.e it is completely arbitrary making it more flexible and easier to fit the hypothesis of the obtained data. A null hypothesis can only be rejected and not proven. The alternative hypothesis is a statement used in**statistical inference experiment.** It is contradictory to the null hypothesis and denoted by Ha or H1. We can also say that it is simply an alternative to the null. The significance level, also denoted as alpha or α, is the**probability of rejecting the null hypothesis when it is true** anda type I error is the rejection of a true null hypothesis, while a type II error describes the error that occurs when one fails to reject a null hypothesis that is actually false.

# **References.**

Bren D. (2022), UCI Department of Statistics: School of Information & Computer Sciences

<https://www.stat.uci.edu/>

Helmenstine, Anne Marie, Ph.D. "What Is a Hypothesis? (Science)." ThoughtCo, Aug. 25, 2020, thoughtco.com/what-is-a-hypothesis-609092.Frost, J. (n. d). Significance Level. <https://statisticsbyjim.com/glossary/significance-level/>

# How Hypothesis Testing Works. (2020). Retrieved 16 February 2020, from <https://www.investopedia.com/terms/h/hypothesistesting.asp>

# Type II Error: https://www.investopedia.com/terms/t/type-ii-error.asp#:~:text=In%20statistical%20analysis%2C%20a%20type,not%20occur%20due%20to%20chance.