AG Publishing House

Advanced Research Techniques: Theories, Methods, And Practices (Volume-2)

Volume 2
Year: 2025

AG PH Books

A Review of Statistical Method for Hypothesis Testing

Dr. Sunita Sharma [1*]

[1]*Faculty in Delhi public world school.*

Abstract

In the context of research initiatives, hypothesis testing is an essential component of statistical analyses. It entails estimating the probability of an event occurring by coincidence. In this article review the various literature's study on statistical method for hypothesis testing. In conclusion, this review highlights the enduring popularity of Null Hypothesis Significance Testing (NHST), despite growing awareness of its limitations. While disciplines like psychology and statistics show stable use, NHST remains widely adopted. Nonparametric methods offer flexible alternatives when assumptions of parametric tests are unmet. The importance of multiple testing correction, careful model selection, and transparent reporting is emphasized to ensure validity and reproducibility. Researchers should move beyond sole reliance on p-values, incorporating confidence intervals and contextual relevance. Advancing statistical literacy and methodological rigor is essential for enhancing the quality and impact of hypothesis-driven research across diverse fields.

Keywords: Null Hypothesis Significance Testing (NHST), Statistical Method, Hypothesis Testing, Alternative Hypothesis, ANOVA, Logit/Probit models, Decision Making, etc.

1 Introduction

A statistical sample is assessed by an analyst during hypothesis testing to establish the plausibility of the null hypothesis. A statistical analyst measures and examines a random sample of the population under investigation. Every analyst employs a random population sample to test two distinct hypotheses: the null hypothesis and the alternative hypothesis [1]. Generally speaking, the null hypothesis states that

* ISBN No. - 978-93-49028-02-9

every population parameter is the same. It may claim, for instance, that the population mean return is zero. In essence, the alternative hypothesis is the opposite of the null hypothesis. Therefore, only one of them can be true, and they are mutually exclusive [2]. One of the two theories, nevertheless, will always be correct. The satirical writer John Arbuthnot, who examined male and female births in England in 1710 after noting that male births outnumbered female births by a little margin in almost every year, is credited by some statisticians with creating the first hypothesis tests. Arbuthnot determined that because there was little prospect of this occurring by accident, divine providence was responsible [3], [4].

A. Hypothesis Testing

One statistical technique for determining whether or not experiment findings are significant is hypothesis testing. It comprises the establishment of a null hypothesis as well as an alternative hypothesis. These two theories will never be compatible. This implies that the alternative hypothesis is false if the null hypothesis is true, and vice versa. A typical instance of hypothesis testing is putting up a test to see if a new drug cures a disease more successfully [5], [6].

Null Hypothesis: A concise mathematical statement known as the null hypothesis is employed to suggest that there is no distinction between two possibilities. Other words, there is no distinction between specific data characteristics. An experiment's results are presumed to be solely determined by coincidence in this hypothesis. It is designated as H0. The null hypothesis is evaluated through hypothesis testing to determine whether it can be refuted. Let us assume that an experiment is conducted to determine whether females are shorter than boys at the age of five. The null hypothesis will be that they are of the same height [7].

Alternative Hypothesis: In contrast to the null hypothesis, the alternative hypothesis is a different proposal. It serves as an indicator that the results of an investigation are attributable to a genuine effect. It signifies the existence of a statistically significant difference between two potential outcomes and is represented as H1 or Ha. The alternative hypothesis for the aforementioned example may be that females are shorter than boys at the age of five.

B. Statistical method

To turn raw data into useful insights in a variety of disciplines, statistical techniques are essential. Using these mathematical tools, researchers, analysts, and decision-makers may efficiently gather, arrange, analyse, interpret, and present data. These techniques make it easier to comprehend large, complicated data sets, find trends, and make wise choices in the social sciences, business, healthcare, and engineering domains. Statistical approaches provide a methodical approach to data analysis, ranging from using descriptive statistics to summarise data to using inferential techniques to test hypotheses and make predictions [8].

C. Descriptive and Inferential Statistics

The two primary subfields of statistics are known as descriptive statistics, which describes the properties of sample and population data, and inferential statistics, which uses these properties to test hypotheses and draw conclusions. Variance, skewness, kurtosis, and mean (average) are examples of descriptive statistics. Analysis of variance (ANOVA), logit/Probit models, linear regression analysis, and null hypothesis testing are examples of inferential statistics.

1. **Descriptive Statistics**

In general, descriptive statistics focus on sample data's central tendency, variability, and distribution. The term central tendency refers to the evaluation of a characteristic or typical feature of a population or sample. Descriptive statistics like the mean, median, and mode are included. A group of statistics known as variability shows how much a population or sample's constituent parts vary from one another along the qualities under evaluation. Included are metrics like range, variance, and standard deviation [9]. The term "distribution" refers to the general "shape" of the data, which may be shown on a chart like a dot plot or histogram and include characteristics like kurtosis, skewness, and the probability distribution function. Additionally, disparities between observed attributes of a data set's components may be described using descriptive statistics. In addition to aiding in the understanding of the collective qualities of the components of a data sample, they may be employed to test hypotheses and make predictions using inferential statistics [10].

2. **Inferential Statistics**

Through the application of inferential statistics, statisticians may make inferences about the features of a population based on observations of a sample. Additionally, it is utilised to assess the degree of certainty they can have on the validity of those judgements. Based on the sample size and distribution, statisticians can calculate the likelihood that "statistics, which quantify the central tendency, variability, distribution, and relationships" among characteristics within a data sample, will accurately represent the corresponding parameters of the entire population from which the sample is drawn [11]. Generalisations about large groups are made using inferential statistics, such as trying to forecast future occurrences or predicting typical demand for a product by surveying the purchasing patterns of a sample of users. A securities or asset class's future return may be projected using returns from a sample period [12].

Regression analysis is a frequently employed mathematical method for statistical inference. The correlation between a dependent variable and one or more explanatory (independent) variables is assessed by examining the strength and nature of the relationship. A regression model's output is frequently subjected to statistical significance analysis, which implies that the results of testing or experimentation are unlikely to have occurred by coincidence or at random. Or, in other words, statistical significance implies that the results are due to a particular cause that is revealed by the data [13].

D. Importance of Hypothesis Testing in Research

Several critical purposes are served by hypothesis testing, which is a critical component of the scientific research process:

Guiding Research: Through the formulation of precise predictions regarding the relationship between variables, hypothesis testing offers a distinct direction for research. In this way, researchers can concentrate their resources and efforts on the investigation of particular hypotheses.

Testing Predictions: By doing so, researchers can evaluate the accuracy of their predictions or assumptions regarding the parameters of the population. Through the comparison of sample data to the hypothesised population parameters, researchers can ascertain whether their predictions are corroborated by the evidence.

Decision Making: Hypothesis testing offers a methodical framework for data-driven decision-making. The acceptance or rejection of a null hypothesis can be determined by researchers through the use of statistical tests, which in turn informs conclusions regarding the research query.

Generalizability: The generalisability of researchers' findings to the broader population can be evaluated through hypothesis testing. Based on the sample data, researchers can draw conclusions about the population by investigating hypotheses.

E. Application of hypothesis testing

Medicine and Healthcare: In clinical trials and medical research, hypothesis testing is essential. To ascertain if novel medications or therapies are successful, researchers use hypothesis testing. Data is gathered from a sample population, and they develop a null hypothesis (the medicine has no impact) and an alternative hypothesis (the drug has a substantial effect). They may determine if the evidence supports rejecting or failing to reject the null hypothesis by examining the data and doing statistical tests.

Quality Control and Manufacturing: To guarantee product quality and process efficiency in manufacturing sectors, hypothesis testing is used. For instance, a vehicle manufacturer may test the hypothesis that a new model's mean fuel economy is higher than that of the old model. Through hypothesis testing and the collection of fuel efficiency data from a sample of automobiles, they may ascertain if the claim is sufficiently supported.

Economics and Finance: In order to assess hypotheses and get well-informed conclusions, economic and financial research use hypothesis testing. An economist may, for example, examine a theory on how a certain policy affects economic growth. They may evaluate the importance of the connection between the policy and economic results by looking at relevant data and doing statistical tests.

Social Sciences: Political science, sociology, psychology, and other social sciences all make extensive use of hypothesis testing. In order to test their theories, researchers in these domains gather data on social events, human behaviour, and the impacts of policies. After then, the data is analysed using statistical tests to see whether the evidence backs up the hypotheses put forward.

Market Research: Market research uses hypothesis testing to verify presumptions and create data-driven choices. Businesses may test theories on customer preferences, market trends, or the efficacy of marketing initiatives via surveys or experiments. They may make wise business judgements and learn more about customer behaviour by using hypothesis testing to the data analysis.

2 Literature Review

(Henríquez-roldán et al., 2025) [14] In order to elucidate the importance of the scientific method in statistics, this research is the first to propose an experiential learning technique in statistical education. It does this by using an actual experimental design. Several critical phases comprise the experiment, including the design of the study, the formulation and testing of hypotheses, the participation of experimental units, the analysis of the collected data, and the interpretation of the results to derive statistical inferences. Along with making statistical ideas easier to understand practically, this approach also highlights the importance of "experimental design, hypothesis testing, data analysis, and probability interpretation in real-world situations". By participating in this immersive experience, students are anticipated to develop a more profound understanding of statistics as a scientific discipline and its practical applications in daily life.

(Emmert-Streib, 2024) [15] Examine the community's response to these discussions. To be more precise, we perform a scientometric analysis of bibliographic records to examine the publication behaviour regarding the utilisation of Null hypothesis significance testing (NHST). A trend analysis is performed for the general community, specific subject areas, and individual journals. Additionally, we perform a change-point analysis to determine whether there are genuine changes or repeated movements. Therefore, we have discovered that NHST is more popular than ever among the general public. However, there is a distinct heterogeneity in the publication behaviour of specific subject areas and journals, and no uniformity is apparent.

(Jangir et al., 2024) [16] A fundamental procedure in statistical analyses for research initiatives is hypothesis testing. It entails assessment of the probability of an event transpiring by coincidence. The subjects are divided into two categories in experimental studies: a control group and a treatment group. The results of these groups are contrasted to determine whether any observed differences are indicative of the treatment effect or are merely the result of random sampling error. Conventional testing is conducted to evaluate the null hypothesis, which asserts that there is no distinction between the categories. In a study that investigates the impact of treatment on mortality rates, the null hypothesis would assert that the treatment has no effect. This methodology is the foundation of the majority of statistical analyses conducted in individual research initiatives.

(Rogers, 2022) [17] Aims to establish a paradigm for comprehending the issue by examining the history of NHST and exploring alternative methodologies. In order to contextualise the issues, a concise overview of the history, dissemination, and criticisms of NHST is presented. The spectrum of available options to NHST is illustrated through the description of supplementary and alternative methodologies. Simulated data sets are employed to illustrate a variety of viable alternative methodologies and supplements to NHST. The Monte Carlo method simulations are employed to compare the methods that provide decision rules with NHST.

(Suvorov et al., 2022) [18] One of the vital procedures in contemporary medical research is statistical hypothesis testing. Scientific researchers initially develop a research hypothesis, which is subsequently used to develop and evaluate the statistical hypothesis. This review delineates the general algorithm for evaluating the null and alternative hypotheses using a t-test, as well as the examples that are compiled for various research questions. In addition, the authors discuss type I errors, which are essential for interpreting p-values estimated from statistical tests, and type II errors, which are employed to evaluate the power of a study. The article concentrates on the distinction between statistically significant and clinically significant effects, as well as the calculation methods and measure of effect size. The connection between type II error, effect magnitude, and sample size is also examined.

(Van Witteloostuijn & van Hugten, 2022) [19] Over the past decade, there has been an additional increase in this critique, which is linked to recommendations for conducting quantitative empirical research in an alternative manner. In the present research note, we conduct a comprehensive examination of 148 articles that present the results of quantitative empirical studies published in recent issues of 18 prestigious journals in six prominent social science disciplines: business, economics, political science, psychology, public administration, and sociology. Although numerous economics studies emphasise "estimation precision" rather than "statistical significance," we conclude that NHST remains excessively prevalent in all six disciplines. In addition, we identify a few instances of innovative practices that may serve as a catalyst for change.

(Mata & Milner, 2021) [20] With a focus on their relevance for current pathology research and treatment, the statistical tests of the publications were summarised. Ninety-three percent of the 195 identified publications mentioned using one or more statistical tests. Several significant discoveries were identified through a retrospective statistical analysis of the articles. Initial reports of tests for normality were rare, and "parametric hypothesis tests" were overused. Second, research that used multisample hypothesis tests (such as analysis of variance) to look at differences between study groups seldom used post hoc testing. Third, the survival analysis, regression, and correlation methods were not used to their full potential. A primer on pertinent statistical concepts and tests is presented in light of these findings. "Regression, correlation, survival, and genetic data analysis" are among the subjects covered in this primer, along with descriptive and comparative statistics and the best research design.

(Dyckman & Zeff, 2019) [21] Suggests enhancements to the method and quality of this type of research. We address the following limitations in current research that we believe are being disregarded or not utilised appropriately: (1) situational effects that remain unaddressed due to model limitations and what has been previously termed "data carpentry," (2) the information loss that results from undervaluing what can and cannot be learnt from replications, (3) the necessary enhancements to relying on a study's computed "p-values" instead of the economic or behavioural significance of the results, and (4) the limitations and alternatives to winsorizing.

(Emmert-Streib & Dehmer, 2019) [22] From the perspective of data science, statistical hypothesis testing is one of the most misconstrued quantitative analysis methodologies. The procedural components of the system are intricately interconnected, despite their apparent simplistic nature. In this paper, we explore the formal meaning of the components and their connections, as well as the underlying logic of statistical hypothesis testing. As a generic backbone, our presentation is applicable to all statistical hypothesis tests, rendering it useful in all application domains of artificial intelligence and data science.

(Kruschke & Liddell, 2018) [23] There is a conceptual distinction between hypothesis testing and estimation with quantified uncertainty in the context of data analysis. The New Statistics is a trend among frequentists in psychology that emphasises estimation over hypothesis testing. Bayesian methods and frequentist methods are also conceptually distinct. Our primary objective in this article is to elucidate the superiority of Bayesian methods over frequentist methods in respect to the objectives of the New Statistics. The paper examines the frequentist and Bayesian methodologies for hypothesis testing and estimation with credible or confidence intervals. Bayesian methodologies for meta-analysis, randomised controlled trials, and power analysis are additionally elaborated upon in the article.

3 Conclusion

This review highlights the enduring prominence of Null Hypothesis Significance Testing (NHST) within the research community, despite growing awareness of its limitations. While its popularity continues to rise overall, subject-specific trends—particularly in psychology, biology, and statistics—reveal a more stable or heterogeneous usage. Nonparametric methods offer flexible alternatives to traditional parametric tests, especially when assumptions of normality and homoscedasticity are violated. The widespread use of hypothesis testing in diverse domains such as medicine, economics, and data science underscores the need for careful model selection, transparency, and reproducibility. As multiple testing becomes increasingly common in fields like genomics and finance, appropriate correction techniques must be applied to mitigate errors. Researchers should move beyond an overreliance on p-values and incorporate confidence intervals, effect sizes, and the practical implications of findings to enhance interpretability and impact. Ethical and methodological rigor, combined with improved statistical literacy and communication, will be essential for advancing hypothesis testing methodologies and fostering meaningful scientific progress. The path forward involves both educating the community and adopting more nuanced, context-sensitive approaches to data analysis.

References

[1] D. P. Turner, H. Deng, and T. T. Houle, "Statistical Hypothesis Testing: Overview and Application," Headache, vol. 60, no. 2, pp. 302–308, 2019, doi: 10.1111/head.13706.

[2] P. J. Veazie, "Understanding statistical testing," SAGE Open, vol. 5, no. 1, 2015, doi: 10.1177/2158244014567685.

[3] S. Pednekar, R. D. Kulkarni, and P. Mahanwar, "STATISTICAL METHODS AND TESTING OF HYPOTHESIS," 2022.

[4] Beeranna and N. G. Basappa, "The Role of Education In Shaping Awareness and Attitudes Towards Politics In Special Reference To Yadagiri District, Karnataka," Int. J. Innov. Sci. Eng. Manag., pp. 55–62, 2024, doi: 10.69968//ijisem.2024v3i455-62.

[5] R. McNulty, "A logical analysis of null hypothesis significance testing using popular terminology," BMC Med. Res. Methodol., vol. 22, no. 1, pp. 1–9, 2022, doi: 10.1186/s12874-022-01696-5.

[6] N. Ben and S. K. Yadav, "Women Empowerment - A Study of SHGS in Satna & Maihar District of Madhya Pradesh," Int. J. Innov. Sci. Eng. Manag., pp. 324–332, 2025, doi: 10.69968/ijisem.2025v4i1324-332.

[7] D. Szucs and J. P. A. Ioannidis, "When null hypothesis significance testing is unsuitable for research: A reassessment," Front. Hum. Neurosci., vol. 11, no. August, 2017, doi: 10.3389/fnhum.2017.00390.

[8] A. J. Harrison et al., "Recommendations for statistical analysis involving null hypothesis significance testing," Sport. Biomech., vol. 19, no. 5, pp. 561–568, 2020, doi: 10.1080/14763141.2020.1782555.

[9] S. L. Pugh and A. Molinaro, "The nuts and bolts of hypothesis testing," Neuro-Oncology Pract., vol. 3, no. 3, pp. 139–144, 2015, doi: 10.1093/nop/npv052.

[10] P. Ranganathan and C. S. Pramesh, "An introduction to statistics: Understanding hypothesis testing and statistical errors," Indian J. Crit. Care Med., vol. 23, pp. S230–S231, 2019, doi: 10.5005/jp-journals-10071-23259.

[11] R. A. Cooper, "Making Decisions with Data: Understanding Hypothesis Testing & Statistical Significance," Am. Biol. Teach., vol. 81, no. 8, pp. 535–542, 2019, doi: 10.1525/abt.2019.81.8.535.

[12] S. Bahariniya and F. Madadizadeh, "Review of the statistical methods used in original articles published in iranian journal of public health from 2015-2019: A review article," Iran. J. Public Health, vol. 50, no. 8, pp. 1577–1585, 2021, doi: 10.18502/ijph.v50i8.6803.

[13] T. Hamasaki, F. Bretz, L. M. LaVange, P. Müller, G. Pennello, and J. C. Pinheiro, "Editorial: Roles of Hypothesis Testing, p-Values and Decision Making in Biopharmaceutical Research," Stat. Biopharm. Res., vol. 13, no. 1, pp. 1–5, 2021, doi: 10.1080/19466315.2021.1874803.

[14] C. F. Henríquez-roldán, S. I. Bangdiwala, R. I. Barrera-, Á. S. Bustos-rubilar, S. I. Bangdiwala, and R. I. Barrera-, "From hypothesis to conclusion: the essential role of statistics in science," Res. Stat., vol. 3, no. 1, p., 2025, doi: 10.1080/27684520.2025.2450532.

[15] F. Emmert-Streib, "Trends in null hypothesis significance testing: Still going strong," Heliyon, vol. 10, no. 21, p. e40133, 2024, doi: 10.1016/j.heliyon.2024.e40133.

[16] M. Jangir, D. Singh Chahar, R. Lal, and S. Sharma, "an Overview of Statistics: Comprehending Hypothesis Testing and Statistical Imperfections (Errors and Bias)," World J. Pharm. Med. Res. www.wjpmr.com │, vol. 10, no. 4, p. 90, 2024, [Online]. Available: www.wjpmr.com

[17] B. D. Rogers, "Null Hypothesis Statistical Testing: A Survey of the History, Critiques, and Alternative Methodologies," (Doctoral Diss., 2022.

[18] A. Y. Suvorov et al., "Statistical hypothesis testing: general approach in medical research," Sechenov Med. J., vol. 13, no. 1, pp. 4–13, 2022, doi: 10.47093/2218-7332.2022.426.08.

[19] A. van Witteloostuijn and J. van Hugten, “The state of the art of hypothesis testing in the social sciences,” Soc. Sci. Humanit. Open, vol. 6, no. 1, p. 100314, 2022, doi: 10.1016/j.ssaho.2022.100314.

[20] D. A. Mata and D. A. Milner, “Statistical Methods in Experimental Pathology: A Review and Primer,” Am. J. Pathol., vol. 191, no. 5, pp. 784–794, 2021, doi: 10.1016/j.ajpath.2021.02.009.

[21] T. R. Dyckman and S. A. Zeff, “Important issues in statistical testing and recommended improvements in accounting research,” Econometrics, vol. 7, no. 2, 2019, doi: 10.3390/econometrics7020018.

[22] F. Emmert-Streib and M. Dehmer, “Understanding Statistical Hypothesis Testing: The Logic of Statistical Inference,” Mach. Learn. Knowl. Extr., vol. 1, no. 3, pp. 945–962, 2019, doi: 10.3390/make1030054.

[23] J. K. Kruschke and T. M. Liddell, “The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective,” Psychon. Bull. Rev., vol. 25, no. 1, pp. 178–206, 2018, doi: 10.3758/s13423-016-1221-4.