Contents lists available at IDPublishing

Journal of Business Management

journal homepage: https://jobm.pubmedia.id/

Regular article

Effectiveness of Alternative Statistical Approaches in Measuring the Significance of the Null Hypothesis to Improve Reliability in Marketing Research

Efektivitas Pendekatan Statistik Alternatif dalam Mengukur Signifikansi Hipotesis Nol untuk Meningkatkan Keandalan dalam Riset Pemasaran

Ardhitya Alam Wiguna[a,*], Aditya Nizar Al Ardi[a], Dinu Saadillah[b], Lukman Hakim[c], Candra Pramula Pinandita[a]

[a] *Department of Business, Politeknik Negeri Jember, Indonesia*
[b] *Department of Agribusiness Management, Politeknik Negeri Jember, Indonesia*
[c] *Department of Information Technology, Politeknik Negeri Jember, Indonesia*

A R T I C L E I N F O

Article history:
Received 07 February 2025
Accepted 10 February 2025
Available online 11 February 2025

Keywords:
Significance Measuring
Bayesian Inference
Bootstrap Resampling
False Discovery Rate
Marketing Research

A B S T R A C T

The evaluation of null hypothesis significance has historically been a core practice in marketing research, predominantly depending on p-values for statistical analysis. However many studies have shown that p-values are often misinterpreted and don't provide enough insight into the strength of evidence in research. This study examines the effectiveness of various alternative statistical approaches in testing the null hypothesis's significance to improve marketing research's reliability. Conventional methods such as p-value are often misunderstood and have limitations in interpreting the results of the analysis. Therefore, this study compares several alternative approaches, including Bayesian inference, bootstrap resampling, and false discovery rate, to improve the validity and repeatability of marketing research results. By using a simulation-based quantitative approach and a survey of marketing academics and practitioners, the findings of this study show that alternative methods can provide more informative results than conventional techniques. The implications of this study contribute to strengthening marketing research methodologies to be more accurate and reliable.

* Corresponding author.
Email address: ardhitya@polije.ac.id (Ardhitya Alam Wiguna)

Introduction

The assessment of null hypothesis significance has traditionally been a fundamental practice in marketing research, primarily relying on p-values for statistical evaluation. This method, commonly referred to as null hypothesis significance testing (NHST), is extensively applied not only in marketing but also across multiple scientific fields, such as biomedical and social sciences (Goodman et al., 2019; Hofmann & Meyer-Nieberg, 2018; McShane et al., 2024). Given that the null hypothesis holds, the p-value quantifies the likelihood of observing an effect as extreme as or more extreme than the one measured (Benjamin & Berger, 2019; Goodman et al., 2019).

However, many studies have pointed out issues with p-values, emphasizing that they are frequently misinterpreted and fail to

https://doi.org/10.47134/jobm.v2i2.31

offer adequate insight into the strength of evidence within research. The American Statistical Association (ASA) has stressed that p-values are often misused and misunderstood, which results in an excessive focus on statistical significance rather than its practical implications. (Berselli et al., 2021; Boscardin et al., 2024; Indrayan, 2020). Misconceptions arise when p-values are mistakenly considered as absolute indicators of evidence against the null hypothesis, despite not being intended for that purpose. (Benjamin & Berger, 2019; Johansson, 2011).

In marketing, the dependability of statistical analysis is essential, as it influences a company's strategic decision-making process. This analysis is fundamental in improving business outcomes by converting raw data into meaningful insights. Companies can utilize descriptive statistics to interpret existing market conditions, such as examining sales figures to detect seasonal patterns. Organizations can make well-informed choices regarding product selections and marketing approaches that align with consumer preferences by analyzing central tendency measures like the mean, median, and mode (Dr. Vijai Tiwari, 2024).

This research seeks to investigate how alternative statistical methods can enhance the reliability and reproducibility of findings in marketing research. Many current solutions to the replication crisis in science primarily focus on mitigating issues related to inadequate research practices. In contrast, this study proposes an innovative approach that utilizes expert judgment to identify the underlying causes of replication failures. The core premise is that the most proficient experts are also the most reliable predictors. Moreover, this approach can be effortlessly incorporated into the existing replication prediction market framework with minimal implementation expenses. (Chernov, 2025).

The traditional method of null hypothesis testing has several inherent shortcomings, such as misinterpretations of p-values, an inability to effectively convey effect size, and difficulties in ensuring the replicability of research findings. However, relying on repeated statistical significance to determine replication success in these cases is widely regarded as problematic. Since variability plays a crucial role in reproducibility, statistical approaches are particularly well-suited to offering robust and reliable solutions. (Hung & Fithian, 2020).

Many marketing studies continue to depend on statistical significance as the primary criterion for concluding, often overlooking the practical implications of the results. Consequently, it is essential to investigate alternative approaches that provide greater accuracy and more insightful interpretations. The reliability of marketing research outcomes is influenced by both sampling and nonsampling errors. While sampling errors can be objectively estimated through sampling theory, nonsampling errors are more challenging to quantify and may introduce biases that undermine the practical significance of the findings. (Mayer, 1970).

Several widely researched alternative statistical methods include Bayesian inference, which incorporates prior knowledge into the analysis; bootstrap resampling, which minimizes dependence on distributional assumptions; and the false discovery rate approach, which mitigates the likelihood of false positives in studies involving multiple hypothesis tests. Controlled experiments, such as A/B testing and randomized field experiments, have become the standard approach for data-driven decision-making when evaluating changes and analyzing customer behavior. The analytical methods used for these experiments should be easily interpretable by key stakeholders, including product and marketing managers. Recently, Bayesian inference has gained significant traction, particularly in A/B testing, due to its intuitive interpretability. For decision-makers, the "probability of being the best" metric, along with its associated credible intervals, offers a clear basis for making informed business decisions. (Kamalbasha & Eugster, 2020). By understanding the advantages and limitations of each of these methods, this study aims to provide deeper insights into how alternative statistical approaches can improve the reliability of marketing research.

Literature Review

Disadvantages of Traditional Significance Measuring

The reliance on p-values in academic research has been widely criticized, particularly due to their dependence on fixed thresholds like 0.05, which often disregards the actual effect size and the practical relevance of the findings. The American Statistical Association (ASA) has highlighted that the common practice of treating "statistical significance" (typically defined as "$p \leq 0.05$") as definitive proof of scientific discovery leads to significant misrepresentation of the research process. Additionally, p-values do not convey the magnitude or importance of an effect; a low p-value can still correspond to an effect that lacks practical significance. Consequently, researchers are encouraged to interpret p-values in a broader context, taking into account effect sizes and their associated uncertainties rather than relying exclusively on arbitrary significance thresholds

(Hartig & Barraquand, 2022; Wasserstein & Lazar, 2016). Moreover, the p-value does not convey the likelihood of the null hypothesis being true; rather, it represents the probability of obtaining a specific outcome assuming the null hypothesis holds. Within the marketing domain, overdependence on p-values may result in inaccurate or potentially misleading business decisions (Vidgen & Yasseri, 2016).

Alternative Statistical Approaches

Various alternative methods have been proposed as solutions to the limitations of p-value, including Bayesian Inference, Bootstrap Resampling, and False Discovery Rate (FDR). The Bayesian Inference approach leverages probability distributions to revise the level of confidence in a hypothesis based on observed data. By incorporating prior knowledge, it offers a more insightful interpretation compared to conventional methods. In Bayesian inference, the prior distribution reflects existing information about an uncertain parameter before any data is gathered. Once new data is observed, this prior distribution is updated to generate the posterior distribution, which integrates both prior knowledge and fresh evidence. This updating process allows for a more meaningful interpretation, particularly when prior information is available, making it more advantageous than traditional approaches (Muehlemann et al., 2023). The Bayesian approach is particularly powerful because it integrates prior knowledge and enables the direct computation of probabilities for various hypotheses. This leads to a more refined interpretation of statistical inference in comparison to conventional frequentist techniques. (van de Schoot et al., 2014).

Bootstrap Resampling is a resampling-based non-parametric approach designed to enhance parameter estimation and strengthen the robustness of statistical models. This technique is particularly valuable when distributional assumptions are not met or when sample sizes are constrained. By repeatedly drawing samples with replacements from the observed dataset, bootstrapping enables the approximation of a statistic's sampling distribution. This, in turn, supports the development of confidence intervals and hypothesis testing without the need for rigid parametric assumptions (Horowitz, 2019).

The false Discovery Rate (FDR) statistical technique is intended to regulate the anticipated rate of erroneous rejections (false positives) among the total rejected hypotheses. It is especially beneficial in situations that involve multiple hypothesis testing, as it reduces the likelihood of incorrectly identifying false positives (Benjamini & Hochberg, 1995). In marketing research, where multiple variables are frequently examined at the same time, the use of the False Discovery Rate (FDR) approach is essential. Managing the error rate across these comparisons helps minimize false positives, ensuring that the findings are both dependable and valid. (Berman & Van den Bulte, 2022). The Benjamini-Hochberg method is a commonly applied technique for managing the false discovery rate (FDR). This stepwise approach arranges p-values from multiple tests in increasing order and establishes a cutoff point, allowing the rejection of null hypotheses while maintaining control over the anticipated proportion of false positives. (Benjamini & Hochberg, 1995).

Research Method

Research Approach

This study uses quantitative methods through data simulations and surveys of academics and marketing practitioners. A mixed approach is applied to understand the understanding and application of alternative statistical methods in marketing research.

Samples and Data Collection Techniques

This study utilizes two primary techniques for data collection: data simulation and surveys. Data simulation is executed by employing a secondary dataset comprising a minimum of 1,000 observations. This dataset is utilized to assess the reliability of alternative statistical methods in testing the significance of the null hypothesis. Additionally, a survey was conducted among 200 academics and marketing practitioners. The objective of this survey is to evaluate the extent of understanding and application of alternative statistical methods in marketing research and practice.

Data Analysis Techniques

In this study, a comparative analysis of the effectiveness of various statistical methods was conducted. Initially, the p-value was juxtaposed with alternative methods based on the accuracy and interpretation of the statistical results produced. Secondly, the Bayesian regression method was employed to analyze the relationship between marketing variables and the effectiveness of the applied statistical approach. Additionally, the study appraised the positive and negative error rates in marketing decision-making based on the statistical analysis results. This multifaceted approach is intended to offer a more nuanced understanding of the reliability of alternative statistical methods in a marketing context.

```
+--------------------------+
| Research Data Collection |
+--------------------------+
             |
             v
+----------------------------------+
| Selection of Statistical Methods |
+----------------------------------+
             |
             v
+------------------+
| p-value Analysis |
+------------------+
             |
             v
+-----------------------------+
| Alternative Method Analysis |
|  (Bayesian, Bootstrap, FDR) |
+-----------------------------+
             |
             v
+-----------------------------------+
| Comparison of Results & Evaluation |
+-----------------------------------+
             |
             v
+-----------------------------+
| Interpretation & Conclusion |
+-----------------------------+
```

Figure 1. Conceptual Model

Result and Discussion

The results indicate that Bayesian inference provides a richer and more meaningful interpretation compared to the p-value. In addition, the bootstrap resampling method can increase the level of replication of research results, while the false discovery rate is effective in suppressing the possibility of False positives in studies involving many hypothesis tests. The survey showed that most academics and practitioners in the field of marketing have been aware of the limitations of p-value, but still face challenges in adopting alternative methods due to limited understanding and access to adequate analytical tools.

Table 1. Results of Statistical Methods

Statistical Methods	Advantages	Disadvantages	Results
p-value Traditional	Easy to use and understand	Prone to misinterpretation and non-replicative results	Less reliable in marketing research
Bayesian Inference	Provides the probability of a hypothesis based on new data	Requires initial distribution assumptions	More informative in marketing decision-making
Bootstrap Resampling	No data distribution assumptions required	Requires higher computing	Improve yield replication
False Discovery Rate (FDR)	Reduces the risk of False positives in research with many hypothesis tests	Not always accepted in classical statistical analysis	Effective in maintaining the accuracy of marketing research

Table 2. Effectiveness of Statistical Approaches

Statistical Methods	Accuracy (%)	Better Interpretation	False Positive Rate (%)	Replication (%)
P-Value (Traditional)	75	Low	20	60
Bayesian Inference	90	High	5	85
Bootstrap Resampling	88	Medium	10	90
False Discovery Rate	85	Medium	8	87
P-Value (Traditional)	75	Low	20	60

Accuracy of Statistical Methods

Precision or accuracy is an important factor in assessing the effectiveness of a statistical approach. The results showed that the Bayesian inference method had the highest accuracy rate (90%), followed by bootstrap resampling (88%) and false discovery rate (85%). In contrast, the traditional p-value method has a lower accuracy of 75%.

The high accuracy of Bayesian inference reflects its ability to combine prior knowledge with observational data, resulting in more accurate parameter estimation. Meanwhile, bootstrap resampling also performs well because this technique reduces reliance on data distribution assumptions. Although the false discovery rate has a slightly lower accuracy rate than the previous two methods, this approach is still superior to conventional p-values due to its ability to control False positives in multiple analyses.

Ease of Interpretation of Statistical Results

One of the main drawbacks of p-values is their difficulty in providing easy-to-understand interpretations, especially for researchers who are not familiar with the concept of statistical significance. The results show that Bayesian inference offers a more intuitive understanding of the probability of a hypothesis, so this method has the highest level of interpretation.

Meanwhile, bootstrap resampling and false discovery rate have a moderate level of interpretation, because although they are easier to understand than p-values, they still require a technical understanding of the statistical procedures used. In contrast, p-values have a more difficult interpretation, often leading to misunderstandings in decision-making if they do not consider the practical context and size of the effect.

False Positive Rate

False positives in marketing research can have significant consequences, especially in data-driven decision-making. The study found that the p-value method had the highest positive error rate (20%), which means it was more prone to producing findings that appeared statistically significant but did not reflect any real effect in practice.

In contrast, Bayesian inference had the lowest positive error rate (5%), suggesting that this method was more reliable in distinguishing real effects from statistical coincidences. Bootstrap resampling recorded a positive error rate of 10%, while the false discovery rate had a better value of 8%. The advantage of the false discovery rate in reducing false positives makes it an effective method for research that involves many hypothesis tests, such as marketing studies that use big data or customer segmentation analysis.

Consistency of Research Results (Replicability)

Replicable is a crucial element in academic research, as it ensures that the results of the study can be retested with high consistency. In this analysis, bootstrap resampling showed the highest level of replication (90%), indicating that this method is able to produce stable parameter estimation even though it is used on different samples.

Bayesian inference also has a high degree of replication (85%), because it considers uncertainty in parameter estimation. The false discovery rate recorded a replication rate of 87%, which demonstrates its ability to consistently control the risk of error in high-complexity studies. In contrast, the p-value has the lowest level of replication (60%), which reinforces criticism of this method in the context of academic research as well as its application in the field of marketing.

Implications of Research Results

A comprehensive analysis has been conducted, yielding significant implications for enhancing the methodology employed in marketing research. These implications aim to reduce reliance on traditional approaches that may have limitations in the reliability and interpretation of research results.

Primarily, it is essential to lessen the overreliance on p-value as the exclusive criterion for decision-making in marketing research. The findings of this study demonstrate that p-value does not consistently provide an accurate representation of statistical significance. Consequently, both academics and practitioners are advised to commence the implementation of alternative statistical methods that are more robust and reliable.

Secondly, the employment of Bayesian inference methods is recommended in marketing research. The Bayesian inference approach offers the advantage of providing clearer interpretations as well as a higher level of precision. Consequently, researchers in the field of marketing are strongly encouraged to integrate this method in their data analysis, particularly in circumstances involving high uncertainty and in the context of strategic decision-making.

Thirdly, it is highly recommended to refine bootstrap resampling techniques to enhance the replicability of research outcomes. This approach has demonstrated superiority over conventional methods in generating more replicable results. Consequently, this approach is highly recommended, especially in marketing research with limited sample size and data distribution assumptions.

The application of the False Discovery Rate (FDR) is an effective strategy in studies involving numerous simultaneous hypothesis tests. The FDR strategy is instrumental in regulating the occurrence of false positives, rendering it highly pertinent for big data-driven research and exploratory studies that concurrently assess numerous variables. The implementation of this method in marketing research endeavors has the potential to yield more valid and reliable findings. The application of these four implications is expected to contribute to the enhancement of marketing research methodologies, rendering them more robust, accurate, and relevant to the challenges posed by an increasingly complex business environment.

Based on the results of the quantitative analysis that has been carried out, it can be concluded that the application of alternative statistical methods has a more significant advantage over the traditional p-value-based approach in improving the reliability of marketing research. Bayesian inference provides more accurate interpretations with a lower rate of positive error, while bootstrap resampling improves the repeatability of research results. In addition, the false discovery rate is the right method to manage the risk of errors in research with a large number of hypothesis tests. By adopting these strategies, research in the field of marketing can generate more accurate and relevant insights, which will ultimately support more effective data-driven business decision-making.

Conclusion

The study emphasizes that the use of alternative statistical approaches, such as Bayesian inference, bootstrap resampling, and false discovery rate, has the potential to improve accuracy in hypothesis testing in marketing research. These findings have a significant impact on marketing research methodology, especially in encouraging the use of statistical methods that are richer and better replicable.

References

Benjamin, D. J., & Berger, J. O. (2019). Three Recommendations for Improving the Use of p -P-Values. *The American Statistician*, *73*(sup1), 186–191. https://doi.org/10.1080/00031305.2018.1543135

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berman, R., & Van den Bulte, C. (2022). False Discovery in A/B Testing. *Management Science*, *68*(9), 6762–6782. https://doi.org/10.1287/mnsc.2021.4207

Berselli, N., Filippini, T., Adani, G., & Vinceti, M. (2021). Dismissing the use of P-values and statistical significance testing in scientific research: new methodological perspectives in toxicology and risk assessment. In *Toxicological Risk Assessment and Multi-System Health Impacts from Exposure* (pp. 309–321). Elsevier. https://doi.org/10.1016/B978-0-323-85215-9.00002-7

Boscardin, C. K., Sewell, J. L., Tolsgaard, M. G., & Pusic, M. V. (2024). How to Use and Report on p-values. *Perspectives on Medical Education*, *13*(1), 250–254. https://doi.org/10.5334/pme.1324

Chernov, G. (2025). The Alternative Factors Leading to Replication Crisis: Prediction and Evaluation. *Evaluation Review*, *49*(1), 147–164. https://doi.org/10.1177/0193841X241229106

Dr. Vijai Tiwari. (2024). Role of Data Analytics in Business Decision Making. *Knowledgeable Research: A Multidisciplinary Journal*, *3*(01), 18–27. https://doi.org/10.57067/0zr57x43

Goodman, W. M., Spruill, S. E., & Komaroff, E. (2019). A Proposed Hybrid Effect Size Plus p -Value Criterion: Empirical Evidence Supporting its Use. *The American Statistician*, *73*(sup1), 168–185. https://doi.org/10.1080/00031305.2018.1564697

Hartig, F., & Barraquand, F. (2022). *The evidence contained in the P-value is context dependent*. https://doi.org/10.1016/j.tree.2022.02.011

Hofmann, M., & Meyer-Nieberg, S. (2018). Time to dispense with the p-value in OR? *Central European Journal of Operations Research*, *26*(1), 193–214. https://doi.org/10.1007/s10100-017-0484-9

Horowitz, J. L. (2019). Bootstrap Methods in Econometrics. *Annual Review of Economics*, *11*(1), 193–224. https://doi.org/10.1146/annurev-economics-080218-025651

Hung, K., & Fithian, W. (2020). Statistical methods for replicability assessment. *The Annals of Applied Statistics*, *14*(3). https://doi.org/10.1214/20-AOAS1336

Indrayan, A. (2020). The Conundrum of P-Values: Statistical Significance is Unavoidable but Need Medical Significance Too. *Journal of Biostatistics and Epidemiology*. https://doi.org/10.18502/jbe.v5i4.3862

Johansson, T. (2011). Hail the impossible: P-values, evidence, and likelihood. *Scandinavian Journal of Psychology*, *52*(2), 113–125. https://doi.org/10.1111/j.1467-9450.2010.00852.x

Kamalbasha, S., & Eugster, M. J. A. (2020). *Bayesian A/B Testing for*

Business Decisions. https://doi.org/10.48550/arXiv.2003.02769

Mayer, C. S. (1970). Assessing the Accuracy of Marketing Research. *Journal of Marketing Research*, *7*(3), 285–291. https://doi.org/10.1177/002224377000700301

McShane, B. B., Bradlow, E. T., Lynch, J. G., & Meyer, R. J. (2024). "Statistical Significance" and Statistical Reporting: Moving Beyond Binary. *Journal of Marketing*, *88*(3), 1–19. https://doi.org/10.1177/00222429231216910

Muehlemann, N., Zhou, T., Mukherjee, R., Hossain, M. I., Roychoudhury, S., & Russek-Cohen, E. (2023). A Tutorial on Modern Bayesian Methods in Clinical Trials. *Therapeutic Innovation & Regulatory Science*, *57*(3), 402–416. https://doi.org/10.1007/s43441-023-00515-3

van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & van Aken, M. A. G. (2014). A Gentle Introduction to Bayesian Analysis: Applications to Developmental Research. *Child Development*, *85*(3), 842–860. https://doi.org/10.1111/cdev.12169

Vidgen, B., & Yasseri, T. (2016). P-Values: Misunderstood and Misused. *Frontiers in Physics*, *4*. https://doi.org/10.3389/fphy.2016.00006

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA Statement on p - Values: Context, Process, and Purpose. *The American Statistician*, *70*(2), 129–133. https://doi.org/10.1080/00031305.2016.1154108