

The Philosophy of Statistical Modeling

Sitti Masyitah Meliyana ^{1*}, Muhammad Rusli Baharuddin ²

^{1*} Universitas Negeri Makassar, Indonesia, email: sitti.masyitah@gmail.com

² Institut Teknologi Sepuluh Nopember, Indonesia, email: 7003241005@student.its.ac

Abstract

This research is a literature review study aimed at exploring the essence of the philosophy and the underlying rationale behind statistical modeling. The research method began with the identification of relevant keywords, namely “Statistical Modeling,” “Why Modeling is Needed,” “Operational Model,” “Estimation,” and “Hypothesis Testing.” The literature search was conducted through credible sources such as international journal databases (Scopus, Web of Science, PubMed, IEEE Xplore), Google Scholar, university digital libraries, as well as reference books and conference proceedings. Based on these keywords, the initial search resulted in approximately 435 articles. After an initial screening based on the titles and abstracts in accordance with the inclusion and exclusion criteria, 75 relevant articles were obtained. A further selection process through full-text reading and quality evaluation of the literature resulted in 30 articles that were reviewed in depth, with the entire process documented using a PRISMA flow diagram and supported by reference management software such as Mendeley, Zotero, or EndNote. The research findings indicate that statistical modeling plays a central role in the modern data-dominated era. Various approaches, such as regression, Bayesian methods, and probabilistic models, allow for the simplification of data complexity, analysis of relationships among variables, and reliable predictions. The Bayesian approach, in particular, stands out due to its ability to handle uncertainty and causal relationships through graphical models such as Bayesian networks, which have proven effective in artificial intelligence and sequential data analysis. Furthermore, probabilistic models provide a representation of uncertainty through probability distributions, which is crucial in the analysis of complex data. The main advantage of statistical modeling lies in its functions in simplification, representation, and prediction, which are supported by the integration of advanced technologies such as machine learning. The process of parameter estimation and hypothesis testing ensures the validity and reliability of the model, making statistical modeling an essential tool in decision-making and practical applications across various fields.

Keywords: Philosophy; Statistical Modeling; Literature Review

Introduction

In the modern era dominated by data, statistical modeling has become an essential tool for understanding, analyzing, and managing complex information. Data now serves as the primary foundation for decision-making in various fields, from healthcare and economics to technology. Statistical modeling transforms raw data into meaningful insights by simplifying complexity, illustrating relationships among variables, and providing reliable predictions.

A statistical model, as a mathematical representation of real-world phenomena, is designed to explain patterns in data and the relationships between variables. Various approaches, such as linear regression, logistic regression, and probabilistic models, have proven their usefulness in numerous applications, from predicting economic trends to diagnosing diseases. On the other hand, more modern approaches like Bayesianism offer greater flexibility in handling uncertainty, allowing for the updating of beliefs based on new data, and facilitating a more intuitive analysis of causal relationships.

The integration of statistical modeling with technological advancements, such as machine learning and artificial intelligence, further reinforces the role of statistics in processing large-scale data. These technologies enable models not only to process data with high efficiency but also to produce more accurate predictions in various contexts, such as genomic analysis, weather forecasting, and dynamic business decision-making.

More than just an analytical tool, statistical modeling also holds significant practical value. It helps address various real-world challenges, such as predicting health risks, identifying the relationship between education and income, and designing effective marketing strategies. With its ability to simplify complex realities and generate actionable insights, statistical modeling is a key asset in tackling the challenges of the data era.

This approach is not only relevant in an academic context but also serves as the foundation for various practical applications. The EAS assignment for the Philosophy of Science and Statistics will explore The Philosophy of Statistical Modeling by discussing: Statistical Modeling, Why is Modeling Needed? Why Must It Be Translated into an Operational Model? Modeling, Estimation, and Hypothesis Testing, as well as providing a case study of modeling in Education.

Method

This research is a literature review study. The research method begins with the identification of relevant keywords, namely "Statistical Modeling," "Why Modeling is Needed," "Operational Model," "Estimation," and "Hypothesis Testing," which are designed to capture the essence of the philosophy and underlying rationale behind statistical modeling. In this study, the literature search was conducted through credible sources such as international journal databases (e.g., Scopus, Web of Science, PubMed, IEEE Xplore), Google Scholar, university digital libraries, as well as related reference books and conference proceedings.

Based on the main keywords "Statistical Modeling," "Why Modeling is Needed," "Operational Model," "Estimation," and "Hypothesis Testing" the initial search yielded approximately 435 articles. After an initial screening based on titles and abstracts in accordance with the inclusion criteria—which encompassed articles, books, and other sources that explicitly discuss the philosophy or underlying rationale behind statistical modeling, have undergone peer review, and present an in-depth discussion regarding the translation of statistical models into operational models and their relationship with estimation and hypothesis testing—and the exclusion criteria, which eliminated sources that only provided superficial discussions or lacked academic validity, 75 relevant articles were obtained. Subsequently, through further selection by reading the full texts and evaluating the quality of the literature, 30 articles were finally selected for an in-depth review. The entire process was systematically documented, for example, using a PRISMA flow diagram, and was supported by reference management software such as Mendeley, Zotero, or EndNote.

In the literature selection process, the inclusion criteria encompassed articles, books, and other sources that explicitly discuss the philosophy or underlying rationale of statistical modeling, have undergone peer review, and present an in-depth analysis of the translation of statistical models into operational models and their relationship with estimation and hypothesis testing; whereas the exclusion criteria eliminated sources that only provided superficial discussions or lacked academic validity, such as blogs or non-reviewed publications.

The analysis and synthesis of the literature were carried out by categorizing the sources based on the main themes, namely the fundamentals of statistical modeling, the urgency of using models, the translation into operational models, and the relationship between modeling, estimation, and hypothesis testing. Then, through thematic analysis, the main patterns, debates, and arguments were identified, which were subsequently synthesized narratively to provide a

comprehensive and comparative overview of various approaches and interpretations in the literature. The evaluation of the quality of the literature was conducted by assessing the methodological strength, data reliability, and contribution of each source, as well as considering the potential bias and limitations of the studies, which were validated through cross-checking among sources to ensure the consistency and validity of the findings.

Results and Discussion

The ability to analyze and understand information is becoming increasingly important. Statistical modeling serves as the primary tool that enables humans to bridge raw data with meaningful insights. Through various approaches and methodologies, statistical modeling helps simplify complexity, capture hidden patterns, and provide a strong foundation for decision-making across various fields. With the continuous advancement of technology and the growing need for data analysis, statistical modeling is becoming increasingly relevant as a foundation for solving the challenges of the modern world

Tabel 1. Reference The Philosophy of Statistical Modeling

The Philosophy of Statistical Modeling	Reference
Statistical Modeling?	Box (1976). Gelman et al (1995). Pearl (2000). Lindley (2000). Silver (2012). Zaharatos, (2022). Cressie (1993).
Why is Modeling Needed?	Tukey (1977). Breiman, L. (2001). Cox (2006). Silver (2012). Pearl (2000). Lindley (2000) David (1962). Casella & Berger (2002) Rubin (1987). Hastie et al (2009), Gelman et al, (2013).
Why Must It Be Translated into an Operational Model?	Nilsen (2015). Tabak et al, (2012). Moullin et al, (2015). Birken et al (2017). Damschroder & Hagedorn (2011). Otaia, (2016). Sunaryo, et al. (2003)
Modeling, Estimation, and Hypothesis Testing	Nasution (2017). Yasin, et al (2022). Suárez (2020). Box & Tiao (2011). Danks (2018) Hasad (2011).

Statistical Modeling?

Data have become a central component in almost every aspect of human life. Data not only serve as a source of information but also as the basis for making better decisions across various fields—from healthcare to business, and from the social sciences to artificial intelligence. Statistical modeling is the primary tool used to analyze this data, handle uncertainty, and predict relevant patterns or outcomes. Through modeling, experts can simplify the complexity of the real world into a more organized and easily understandable structure.

A statistical model is a mathematical representation designed to explain the relationships between variables and patterns in data (Zaharatos, 2022). Common examples of statistical models include linear regression, which is used to predict the relationship between independent and dependent variables; logistic regression, which aids in classification, such as diagnosing diseases based on symptoms; and time series models, which are useful for analyzing data over time (Wulff, 2017), such as predicting trends in stock prices or rainfall. Box (1976) stated that a model is a simplification of reality, or in other words, a model is an essential tool for understanding real-world phenomena. However, the success of a statistical model depends not only on its mathematical structure but also on the approaches and methodologies used in its implementation.

One of the most flexible and popular approaches in statistical modeling is Bayesianism. The Bayesian approach has an advantage due to its ability to handle uncertainty through probabilities. Bayesianism involves three main components: the prior, the likelihood, and the posterior. The prior reflects initial knowledge before new data are obtained, while the likelihood represents the information derived from the data. The combination of these produces the posterior, which

represents updated beliefs based on new evidence. Gelman (1995) explains, “The Bayesian paradigm provides a coherent framework for updating beliefs in light of new evidence.” This advantage makes Bayesianism highly relevant in various situations, such as medical analyses that predict the likelihood of a disease based on early symptoms, dynamic weather forecasting, and business decision-making that requires quick responses to changing market conditions.

Bayesianism not only excels at handling uncertainty but also enables the modeling of causal relationships in data through Bayesian networks. Bayesian networks are probabilistic graphical models that represent causal relationships between variables. In these models, nodes represent variables while arrows indicate causal links among them. The applications of Bayesian networks are extensive. In healthcare, for example, these networks are used to predict disease risk based on genetic factors and lifestyle. In finance, Bayesian networks help analyze the relationship between monetary policy and market conditions, while in artificial intelligence, they are used to build automated decision-making systems. Pearl (2000), a pioneer in causal theory, stated, “Bayesian networks excel in representing and reasoning about probabilistic dependencies, making them indispensable for modern AI and data science.” This quote underscores the important role of Bayesian networks in modern data analysis, particularly in situations where complex relationships among variables require a structured and intuitive approach.

In addition to Bayesian networks, probabilistic models in general have a very broad scope in statistical modeling. Probabilistic models are used to describe uncertainty in data by considering the probability distributions of variables. This approach allows for more realistic predictions by reflecting the natural variation within the data. Examples of probabilistic model applications include Gaussian Mixture Models, which are used to cluster data based on normal distributions, and Hidden Markov Models (HMMs), which are very useful for analyzing sequential data such as text or genomics. Lindley (2000) stated, “Probability is the language of uncertainty, and statistical models are its grammar.” This quote illustrates the importance of probability in statistical modeling, as probability is used not only to predict outcomes but also to mathematically explain uncertainty. Probabilistic models enable a deeper understanding of complex phenomena, such as genetic patterns associated with diseases or linguistic patterns in text analysis.

Statistical modeling is not only theoretically relevant but also has broad practical applications in the real world. In healthcare, for example, logistic regression is often used to predict the risk of diseases such as heart attacks based on risk factors like blood pressure, cholesterol, and lifestyle. In economics, time series models are used to analyze trends such as inflation, GDP growth, or stock market fluctuations. In the social sciences, regression models help identify the relationship between education and income, providing insights into how investment in education can affect a society’s economic well-being. Nate Silver stated, “The beauty of statistics is that it turns data into actionable insights, shaping decisions in every field imaginable.” This view highlights how statistical modeling provides insights that can be directly applied to solve real-world problems, while also helping decision-makers create more effective, evidence-based policies.

Furthermore, the advantages of statistical modeling lie not only in its broad application but also in its ability to adapt to new technologies and the ever-growing volume of data. With advancements in computing and artificial intelligence, statistical models such as Bayesian networks and other probabilistic models are increasingly being enhanced by machine learning algorithms. This technology enables models to process large amounts of data and produce more accurate predictions in a shorter time. For example, genomic analysis that once took weeks can now be completed in hours with the aid of advanced statistical models integrated with machine learning.

Overall, statistical modeling is a cornerstone of modern science and practical application. From linear regression to Bayesian networks, these models enable in-depth analysis, reliable predictions, and effective management of uncertainty. With their ability to explain relationships

between variables, understand causality, and provide actionable insights, statistical models continue to be essential tools for addressing the complex challenges of the modern world. Experts in statistics emphasize the importance of this approach in understanding and leveraging data for collective advancement. As technology evolves and more data become available, the relevance of statistical modeling will continue to increase, helping humanity address fundamental questions and tackle future practical challenges. In this way, statistical modeling is not only a tool for data analysis but also a key to understanding the world more deeply.

Why is Modeling Needed?

Humans are always faced with a world full of complexity, where every natural, social, or economic phenomenon consists of many interrelated and interacting variables. In such situations, models become essential tools that enable humans to understand, analyze, and make better decisions. The need for modeling is rooted in its three main functions: simplification, representation, and prediction.

1. Simplification: Making Complexity More Understandable

One of the main reasons humans need models is to simplify the complexity of the real world. This reflects the idea that models are imperfect abstractions, yet they remain important tools for understanding a complex world. Without the simplification provided by models, available data would be difficult to manage and patterns in the data would be hard to identify. Tukey (1977), known for his contributions to exploratory data analysis, emphasized the importance of simplification in data analysis by stating, “An approximate answer to the right question is worth a good deal more than an exact answer to an approximate question.” Tukey stressed that models allow us to focus on the main relevant elements, even if not all details are represented. This concept is also supported by Breiman (2001), who emphasized that models not only simplify data but also help detect patterns that were previously hidden. Breiman believed that the simplification provided by statistical models creates accessibility to complex insights.

2. Representation: Capturing the Essence of the Real World

Models function as representations of the real world, helping humans create abstractions that capture the essence of the systems being studied. David Cox stated, “Models are not meant to be true; they are meant to be useful.” Cox (2006) explained that models are designed to depict the relevant parts of reality without claiming to be perfect mirrors of it. Gelman (2013) reinforced this philosophy with the Bayesian approach, which emphasizes the ability to update beliefs based on new evidence. In other words, the representation provided by models is not static but dynamic, allowing for adaptation to new data. Noel Cressie added that representation in statistical models enables us to visualize relationships between data that are distributed geographically or spatially (Firman, 2019).

3. Prediction: Assisting Decision-Making

Models not only simplify and represent reality but also allow for predictions that form the basis of decision-making. This philosophy is rooted in pragmatism, where models must provide direct benefits to users. Nate Silver (2012) stated, “The beauty of statistics is that it turns data into actionable insights, shaping decisions in every field imaginable.” According to Silver, the ability of models to predict is a key added value that makes statistics relevant across various fields, from politics to economics. Pearl (2000), a pioneer in causal theory and Bayesian networks, argued that prediction is not only about forecasting the future but also about understanding causality: “Bayesian networks excel in representing and reasoning about probabilistic dependencies, making them indispensable for modern AI and data science.” Additionally, Lindley (2000) highlighted that prediction is at the core of probability by stating, “Probability is the language of uncertainty, and statistical models are its grammar.” In other words, models provide a framework for making decisions in the face of uncertainty.

4. Validation and Empirical Testing

The philosophy behind models is also based on the need to test and validate theories using real data. David (1962) emphasized the importance of connecting theory with empirical evidence by stating, “Models are the bridge between theory and data, helping us test our understanding of the world.” Casella & Berger (2002) added that models enable systematic hypothesis testing, noting, “A model is only as good as its ability to stand against real-world data.” In this context, models become tools for testing the validity of scientific claims.

5. Adaptation to Technology and Changing Times

Statistical models are not only relevant in traditional contexts but also evolve with technological advances. Hastie et al. (2009) noted that “Statistical models evolve with the data, becoming more sophisticated as technology advances.” They explained that the integration of statistics with machine learning has enhanced the ability of models to analyze large volumes of data. Rubin (1987) highlighted how modern models enable the analysis of more complex causal relationships, especially in real-world experiments, stating, “Models allow us to understand not just what happened, but why it happened.”

Why Must It Be Translated into an Operational Model?

Translating the philosophy of models into operational models is a fundamental step that bridges theoretical ideas with practical application. Philosophically, this process reflects a pragmatic approach, which assesses the value of an idea based on its practical benefits in solving real-world problems. It allows abstract ideas to be implemented, tested, modified, and adapted to specific needs, making them more relevant and applicable (Nilsen, 2015).

Models such as Gaussian Mixture Models and Hidden Markov Models are concrete examples of translating philosophy into operational form (Otaia, 2016). These models, designed to recognize complex and hidden patterns in data, would have little impact if not applied within an operational framework. With support from technologies like machine learning, these models have proven highly effective in various fields. For instance, in healthcare, they are used to improve diagnostic accuracy; in finance, they enable more precise risk management; and in technology, they serve as the foundation for artificial intelligence (Tabak et al., 2012).

John Tukey, a prominent statistician, emphasized the importance of operational representation in data. He argued that data visualization provides extraordinary value by revealing previously unexpected insights. This illustrates the role of operational models in bridging theoretical abstraction with practical understanding (Sunaryo et al., 2003). By visualizing data or patterns, abstract concepts become easier to understand and implement. Research across various disciplines has demonstrated the importance of translating models into operational forms. In healthcare, for example, the use of tools based on operational models has been shown to enhance the effectiveness of evidence-based policies and clinical decision-making.

In the financial sector, translating theory into practice enables sharper risk analysis and better strategic decision-making. In technology, algorithms built on theoretical models have become the backbone of innovation—from facial recognition to natural language processing (Birken et al., 2017). Statistician George Box once said, “All models are wrong, but some are useful.” This statement reinforces that while models cannot perfectly represent reality, their value lies in their practical relevance for solving problems. Operational models allow us to acknowledge these limitations while still effectively addressing real-world issues (Damschroder & Hagedorn, 2011). Thus, translating the philosophy of models into operational models is not merely a technical process; it is a crucial step in integrating theory into practice with tangible impact. This process ensures that theories remain not only academically relevant but also applicable in solving real-world challenges with adaptive, sustainable, and impactful solutions. George Box emphasized the importance of translating theoretical models into usable forms despite their limitations, as this

accelerates analysis, enhances prediction accuracy, and yields insights relevant for decision-making across various fields. Turning the philosophy of models into operational models thus aids practical implementation and ensures their relevance and effectiveness in facing real-world challenges.

Modeling, Estimation, and Hypothesis Testing

Estimation and hypothesis testing are essential components of statistical modeling, serving to connect theory with empirical data and to ensure the validity and reliability of the models used (Box & Tiao, 2011). Parameter estimation enables the determination of the most likely values of model parameters based on available data, allowing the model to accurately represent the phenomena being studied (Suárez, 2020). For example, in a regression model, parameter estimation helps in understanding the relationship between independent and dependent variables, which can then be used for further prediction and inference. Hypothesis testing is used to assess the validity of certain assumptions or claims about population parameters. Through hypothesis testing, we can determine whether the observed data are consistent with the initial hypothesis or whether there is sufficient evidence to reject it. For instance, in regression analysis, hypothesis testing can be used to evaluate the significance of a particular predictor variable in the model, ensuring that the identified relationship is not merely the result of random variability.

The relationship between estimation and hypothesis testing in modeling is very close (Danks, 2018). Parameter estimation provides the quantitative basis necessary for conducting hypothesis tests, while hypothesis testing offers a framework to assess the reliability and validity of those estimates (Hasad, 2011). Without accurate estimation, hypothesis testing might not yield valid results, and without hypothesis testing, parameter estimates might not be interpreted with adequate confidence.

Statistical literature has emphasized the importance of integrating estimation and hypothesis testing in modeling. For example, a study by Nasution (2016) highlighted the use of Maximum Likelihood Estimation (MLE) for parameter estimation in a Gamma regression model, followed by hypothesis testing using the Likelihood Ratio Test (LRT) to assess the significance of the estimated parameters. This approach ensures that the constructed model not only fits the observed data but also possesses strong statistical validity. Similarly, research by Yasin et al. (2022) discussed parameter estimation in a Geographically Weighted Generalized Gamma Regression model using the Berndt-Hall-Hausman (BHHH) algorithm, along with hypothesis testing to assess the uniformity of parameters across different locations. This study emphasized the importance of hypothesis testing in ensuring that variations in estimated parameters reflect true differences among locations, rather than merely sample variability.

Conclusion

Statistical modeling plays a central role in the modern, data-dominated era. With various approaches such as regression, Bayesianism, and probabilistic models, statistical modeling enables the simplification of complexity, analysis of relationships between variables, and reliable prediction. This approach is not only theoretically relevant but also has wide practical applications, ranging from healthcare to economics. The Bayesian approach, with its ability to handle uncertainty and causal relationships, is one of the most flexible methods. The integration of Bayesianism into graphical models like Bayesian networks has proven its usefulness in various fields, including artificial intelligence and sequential data analysis. Additionally, probabilistic models provide the ability to represent uncertainty through probability distributions, which is essential in the analysis of complex data.

The advantages of statistical modeling lie in its three main functions: simplification, representation, and prediction. Simplification helps in understanding patterns within complex

data, representation allows for a relevant abstraction of the real world, and prediction provides actionable insights for decision-making. This approach is also supported by the integration of advanced technologies, such as machine learning, which enhances the efficiency of big data analysis. The modeling process also includes parameter estimation and hypothesis testing to ensure the validity and reliability of the model. Parameter estimation helps accurately represent phenomena, while hypothesis testing assesses the validity of model assumptions based on empirical data. The integration of these components ensures that the model has strong statistical validity, making it reliable for various applications.

Overall, statistical modeling is an essential tool for addressing the challenges of the modern world, helping bridge raw data with meaningful insights. With technological advances and increasing data complexity, the relevance of statistical modeling will continue to grow, making it a key element in understanding and utilizing data for progress across various fields.

Acknowledgment

References

- Birken, S. A., Powell, B. J., Shea, C. M., Haines, E. R., Kirk, M. A., Leeman, J., Rohweder, C., Damschroder, L., & Presseau, J. (2017). Criteria for selecting implementation science theories and frameworks: results from an international survey. *Implementation Science*, 12(1), 124. <https://doi.org/10.1186/s13012-017-0656-y>
- Box, G. E. (1976). Science and statistics. *Journal of the American Statistical Association*, 71(356), 791-799.
- Box, G. E., & Tiao, G. C. (2011). *Bayesian inference in statistical analysis*. John Wiley & Sons.
- Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science*, 16(3), 199–231. <https://doi.org/10.1214/ss/1009213726>
- Casella, G., & Berger, R. L. (2002). *Statistical inference* (2nd ed.). Duxbury Press.
- Cox, D. R. (2006). *Principles of statistical inference*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511813559>
- Cressie, N. (1993). *Statistics for spatial data*. Wiley. <https://doi.org/10.1002/9781119115151>
- Damschroder, L. J., & Hagedorn, H. J. (2011). A guiding framework and approach for implementation research in substance use disorders treatment. *Psychology of Addictive Behaviors*, 25(2), 194–205. <https://doi.org/10.1037/a0022284>
- Danks, D. (2018). Probabilistic models. In *The Routledge handbook of the computational mind* (pp. 149-158). Routledge.
- David, F. N. (1962). *Games, gods, and gambling: The origins and history of probability and statistical ideas*. Hafner Publishing.
- Firman, H. (2019). Kepastian dan ketidakpastian dalam sains. *Jurnal Filsafat Indonesia*, 2(1), 33-36.
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. Chapman and Hall/CRC.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). CRC Press. <https://doi.org/10.1201/b16018>
- Hasad, A. (2011). Verifikasi dan validasi dalam simulasi model. *Departemen Ilmu Komputer*, 5-6.

- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- Lindley, D. V. (2000). The philosophy of statistics. *Journal of the Royal Statistical Society Series D: The Statistician*, 49(3), 293-337.
- Moullin, J. C., Sabater-Hernández, D., Fernandez-Llimos, F., & Benrimoj, S. I. (2015). A systematic review of implementation frameworks of innovations in healthcare and resulting generic implementation framework. *Health Research Policy and Systems*, 13(1), 16. <https://doi.org/10.1186/s12961-015-0005-z>
- Nasution, A. S. (2017). Estimasi Parameter dan Pengujian Hipotesis Pada Model Regresi Gamma (Studi Kasus: Pemodelan Pencemaran Sungai Di Surabaya). *Jurnal Paidagogeo*, 2(2), 17-26.
- Nilsen, P. (2015). Making sense of implementation theories, models and frameworks. *Implementation Science*, 10(1), Article 53. <https://doi.org/10.1186/s13012-015-0242-0>
- Otaya, L. G. (2016). Probabilitas Bersyarat, Independensi dan Teorema Bayes Dalam Menentukan Peluang Terjadinya Suatu Peristiwa. *Tadbir: Jurnal Manajemen Pendidikan Islam*, 4(1), 68-78.
- Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511803161>
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Wiley.
- Silver, N. (2012). *The signal and the noise: Why so many predictions fail—but some don't*. Penguin Press.
- Suárez, M. (2020). *Philosophy of probability and statistical modelling*. Cambridge University Press.
- Sunaryo, S., Setiawan, S., Djuraidah, A., & Saefuddin, A. (2003). sejarah perkembangan statistika dan aplikasinya. In *Forum Statistika Dan Komputasi*, 8(1).
- Tabak, R. G., Khoong, E. C., Chambers, D. A., & Brownson, R. C. (2012). Bridging research and practice: Models for dissemination and implementation research. *American Journal of Preventive Medicine*, 43(3), 337–350. <https://doi.org/10.1016/j.amepre.2012.05.024>
- Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.
- Wulff, S. S. (2017). *Time series analysis: Forecasting and control*.
- Yasin, A., Purnadi, & Choiruddin, A. (2022). Estimasi Parameter dalam Model Geographically Weighted Generalized Gamma Regression Menggunakan Algoritma Berndt-Hall-Hausman (BHHH). *Gaussian: Jurnal Matematika dan Pendidikan Matematika*.
- Zaharatos, B. (2022). *The Philosophy of Statistics*