The Scientific Temper (2024) Vol. 15 (2): 2380-2384
E-ISSN: 2231-6396, ISSN: 0976-8653
Doi: 10.58414/SCIENTIFICTEMPER.2024.15.2.54
https://scientifictemper.com/

RESEARCH ARTICLE

Bayesian Optimization Phase I Design of Experiment Models

Maria D. Roopa, Nimitha John*

Abstract

This paper offers a concise overview of a novel model for clinical trials, focusing on measurable phase I outcomes from a Bayesian perspective. It outlines hypothetical Bayesian criteria standards and discusses utilization model techniques, including sampler Bayesian models. Bayesian methodologies are increasingly popular in clinical research for their ability to incorporate prior information and adapt trial designs based on accumulating data. Phase I trials are vital for assessing new treatment safety, making them ideal for Bayesian approaches. The model leverages Bayesian principles to guide trial decisions, like dose escalation and maximum tolerated dose determination. By merging prior knowledge with observed data, Bayesian methods provide a framework for informed decisions, especially in scenarios with small sample sizes or historical data. Additionally, the paper explores various Bayesian model techniques, including samplers for posterior inference, enhancing decision-making in clinical trials. Overall, it contributes to Bayesian methodologies by outlining a tailored model for phase I trials and offering practical implementation guidance to improve early-phase trial efficiency and reliability.

Keywords: Optimization, Clinical Data, Bayesian sampler, Outcomes, Prior information, Design model.

Introduction

Prior data might be accessible from prior tests or from guesses which conjectures the examination. The Bayesian methodology gives a sound system where earlier data and optimal designs with respect to obscure amounts of the clinical data can be consolidated to find an exploratory plan that improves the objectives of the clinical trial.

This study provides an initial examination of hierarchical Bayes models, as well as a summary of their effective implementations. An examination of the underlying assumptions is presented in the following section, which is then followed by an introduction to the Markov Chain Monte Carlo (MCMC) methods. Within the scope of this work, a case study is presented that illustrates the implementation of Bayesian approaches in a conjoint study. Additionally, additional examples of successful applications are demonstrated. In conclusion, we will now explore the difficulties faced while utilizing hierarchical Bayesian models.

Department of Statistics, Christ University Bangalore, India.

***Corresponding Author:** Nimitha John, Department of Statistics, Christ University Bangalore, India., E-Mail: john@christuniversity.in

How to cite this article: Roopa, M. D., John, N. (2024). Bayesian Optimization Phase I Design of Experiment Models. The Scientific Temper, **15**(2):2380-2384.

Doi: 10.58414/SCIENTIFICTEMPER.2024.15.2.54

Source of support: Nil

Conflict of interest: None.

Methodology

The application of Bayesian hierarchical modeling is a well-established method in the field of meta-analysis. Previous studies have explored hierarchical models across adverse side effects within body systems. However, we combine the novel approach of combining these two methods is of great significance. Typically, numerous clinical trials are carried out to evaluate a drug, with most primarily focusing on its efficacy. It becomes crucial to leverage the entire body of evidence available. It's important to note that safety pertains to the drug itself and not just the individual trial. Sometimes, a safety signal may only become evident after analyzing data from multiple trials [P Müller *et al.,* (2005); MJ BAYARRI *et al.,* (1999)].

The Bayesian framework finds particular relevance during the planning phase of clinical trials. At this stage, there is often a wealth of external information available, including historical data, insights from prior studies, and expert opinions, waiting to be effectively leveraged. As emphasized, we tend to adopt a Bayesian mindset when designing trials. Health authorities have issued significant guidelines concerning the statistical, clinical, and regulatory aspects of Bayesian clinical trials. In recent times, there has been a growing acceptance and even encouragement of innovative methods, especially adaptive designs. This shift towards Bayesian approaches is evident in the numerous proposals for trial designs with a Bayesian perspective, some of which involve virtual reinterpretations of previously published trials.

Received: 05/04/2024 **Accepted:** 03/05/2024 **Published** : 15/06/2024

The model consists of the following parameters

X is the design matrix of cohort points and theta is the doses.

Q is the priors

P is the design matrix of inverse design end points.

$y \mid s, \theta, \upsilon \sim N(X\theta + Us, \upsilon^{-1}I)$, $s|\theta, \upsilon \sim N(0, \tau^2 I)$

$\theta|\upsilon \sim N(\mu_0, (\upsilon Q_0)^{-1})$, $\upsilon \sim Ga(\alpha_0, \beta_0)$

$\theta|\upsilon \sim N(\mu, (\upsilon Q)^{-1})$, $\upsilon \sim Ga(\alpha, \beta)$

where

$\alpha = \alpha 0 + p/2$,

$\beta = \beta_0 + (y'Py + \mu'_0 Q_0 \mu_0 - \mu' Q \mu)/2$,

$\mu = (Q_0 + X'PX)^{-1}(Q'_0 \mu_0 + X'Py)$ *and*

$Q = (Q_0 + X'PX)$

Utilizing Bayesian theory to incorporate informed conjectures about the likelihood of events. Bayesian approach aimed to enhance predictionresults from meticulously controlled experiments. Bayesian techniques those acknowledging the inherent biases within scientific inquiry recognized that formalizing the role of informed conjectures could lead to improved replication of existing effects.

Bayesian analysis is conceptually simple. In marketing, it is implemented using Markov chain Monte Carlo (MCMC) methods to handle informed conjectures and integrated models. The specific details of MCMC methods will be explained below.

Hierarchical Bayes models are a combination of two components: i) a model organized in a hierarchical structure and ii) estimation performed using Bayesian approaches. Hierarchical models are composed of separate modules, with one module dedicated to analyzing behavior inside a unit (such as individual responder behavior) and another module dedicated to analyzing behavior across units. The Bayes' theorem is formed by combining these sub-models, and a hierarchical model is used to incorporate these components and manage the related uncertainties. Hierarchical models harmonize well with MCMC methods, serving as the driving force behind the advancement and application of Bayes' theorem [DA Berry *et al.,* (2018); R Etzioni *et al.,*(1993)].

In addition to its application in meta-analysis, Bayesian hierarchical modeling plays a crucial role in assessing adverse side effects within body systems. The integration of these two methods holds significant promise for advancing our understanding of drug safety across diverse patient populations and clinical settings. As clinical trials primarily focus on efficacy, leveraging the entire body of evidence becomes imperative for comprehensive safety evaluations, especially considering that safety signals may only emerge from analyzing data across multiple trials. During the planning phase of clinical trials, the Bayesian framework offers a powerful tool for incorporating a wealth of external information, including historical data and expert insights. This Bayesian mindset in trial design is further supported by guidelines from health authorities, which increasingly recognize the value of innovative methods like adaptive designs. This paradigm shift towards Bayesian approaches is evident in the growing number of trial proposals adopting this perspective, including virtual reinterpretations of past trials. The model parameters encompass various elements such as the design matrix of cohort points (X), doses (theta), priors (Q), and design matrix of inverse design endpoints (P). Utilizing Bayesian theory facilitates the incorporation of informed conjectures about event likelihoods, enhancing predictions from meticulously controlled experiments and addressing inherent biases in scientific inquiry. This formalization of informed conjectures can lead to improved replication of existing effects and better interpretation of study findings. In marketing, Bayesian analysis, often implemented using Markov chain Monte Carlo (MCMC) methods, provides a conceptually simple yet powerful approach for handling informed conjectures and integrating models. Hierarchical Bayes models, combining hierarchical structures with Bayesian estimation, offer a comprehensive framework for analyzing behavior both within and across units. These models, in conjunction with MCMC methods, serve as the driving force behind the advancement and application of Bayes' theorem across various disciplines, including clinical trials and marketing research.

Results

Quantitative analysis involves studying the parameters in linear models, which are also the main focus in hierarchical Bayes models. These models go beyond just determining if a random effect is present or absent. Probability distributions are used to quantify prior opinions about certain properties.

For the purpose of obtaining a posterior distribution of the parameter, Bayes' theorem had to be implemented by multiplying probability densities for the prior and the likelihood prior to the advent of Markov chain Monte Carlo (MCMC).

An analogous process is employed to calculate the likelihood, which represents the information pertaining to the parameter encompassed in the data, by combining the prior distribution with the variance term. Multiplying the likelihood by the prior distribution yields the posterior distribution of the parameter β. Bayes' theorem, despite its conceptual elegance, faced initial resistance in marketing and other applied fields due to the computational complexity of analytical computations, which limited its applicability to just the most basic issues [AC Atkinson *et al.,* (1996)].

The implementation of Markov chain Monte Carlo (MCMC) techniques resolved this computational limitation. MCMC approaches utilize iterative computations to simulate samples from the posterior distribution instead of obtaining its analytical form. Subsequently, these Monte Carlo samples are utilized to calculate pertinent statistics, such

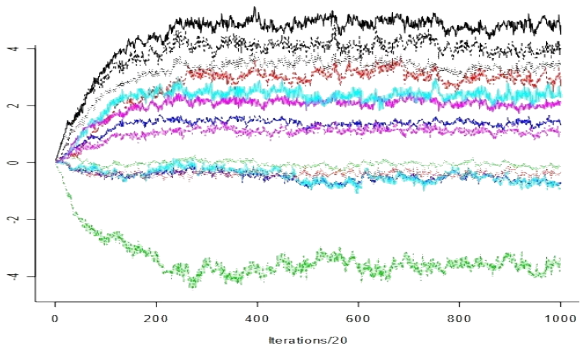

Figure 1: Conversion graph of utility data function data points

as estimations of parameters and intervals of confidence. The MCMC sampler, which is responsible for the hierarchical Bayes revolution, establishes a Markov chain that produces samples from the posterior distribution of model parameters [M Clyde et al., (1996); DV Lindley et al., (1972)].

The Markov chain's steps depend on the specific model being used in the hierarchical set up. Although the notion of MCMC methods is easy to understand, its implementation requires the derivation of suitable conditional distributions for generating draws, which are obtained using Bayes' theorem. Thankfully, various tools are available to assist researchers in generating draws from more intricate models, making this approach broadly applicable (Figure 1).

We employ a hierarchical Bayes random-effects logit model, which allows us to consolidate data while retaining the ability to examine the preferences and characteristics of individual customers. We present and demonstrate the model within the context of a conjoint analysis. Our illustration showcases that the covariance matrix of random effects holds significant implications for product design. Furthermore, the model aids in targeting and identifying individuals who exhibit strong preferences for specific data for priors. Additionally, we highlight that the proposed model provides a more accurate means of integrating demographic variables into the analysis and offers a solution for handling large datasets [K Chaloner *et al.,* (1995)].

Assuming we want to provide an example, let's define the target dose as the group of persons whose net utility is higher than 10.0. The low dosage distribution has around 4.5% of its values exceeding this threshold, while the low gaussian distribution has approximately 7.5%. Despite the fact that the average utility for tumor response is 0.8 units higher, the low dose distribution contains nearly double the amount of mass in the region of the tumor distribution that corresponds to individuals with strong previous beliefs. These individuals are more likely to accurately approximate the ideal tumor size as indicated by MCMC priors.

Dose-response experiments

Assume that the k groups represent the administration of a "dose" at progressively higher levels. Based on Smith's references from 1973 b and 1977, it is plausible to expect that the θ's will approximately reside on a response curve that can be represented by a low degree polynomial. The design is used with reference to Goerke(2007) for the dose levels0.60,1.20,2,3,4,5.33,7.11,9.48,12.64,16.86 mg [M Clyde *et al.,* (1996); DA Berry *et al.,* (2018)].

Results

Numerical Optimization in Bayesian Experimental Design

In Bayesian experimental design, the process of finding optimal designs, denoted as e∗, often involves numerical optimization. In some cases, this procedure may be as challenging as solving complex optimization problems like the TSP. One way to think about designs is as sets of support points. These points represent distinct treatment levels, and the weights assigned to them suggest how many observations should be devoted to each point [RE Kass *et al.,* (1995)]. To ensure that the count of observations at each support point is a whole number, a "exact" design is required. However, in practice, this exact design is not always feasible, and "continuous" designs are explored, where weights can take real-number values. Although the process of obtaining the most optimal exact design can be complex, the relaxed problem, which incorporates continuous designs, is mathematically easier to handle. Interestingly, when the answer aligns perfectly with a specific design, it is also the most ideal design on a global scale. Continuous designs can be estimated to obtain precise solutions that closely approach optimality [Etzioni *et al.,*(1993); JB Kadane et al.,(2011)].

Computing expected utility for generalized linear models, nonlinear models, and other "nonlinear" design challenges is typically not possible in a straightforward manner and requires the use of approximations. In nonlinear design, asymptotic normal approximations are frequently employed for expectation estimation. Numerical quadrature, Laplace integration, and Monte Carlo integration are all methods that can be used to approximate integrals [CF Powell et al., (1995)]. With the rise of powerful computers, simulation-based optimal design is now within reach, but tailored approaches are usually necessary for every specific use case. Simulation-based design notably enables the use of more practical utility functions that are better suited to fit with the goals of the experiment. Simulation-based optimal design techniques now allow for the incorporation of hierarchical models into experimental design. Not only are these models essential for modeling random and latent effects, but they are also essential for accounting for subject-to-subject variability [M Clyde *et al.,* (1996)] (Tables 1 and 2).

Applications

Bayesian experimental design is utilized in diverse domains. It has been applied in various linear models, including

Table 1: : Clinical Trial Design Points for Linear Regression Model

v	*P1*	*P2*	*P3*	*P4*	*P5*	*P6*	*P7*
0	0.145	0.145	0.145	145	145	0.145	0.145
0.25	0.15	0.3	0.04	0.08	0.08	0.2	0.15
0.5	0.25	0.1	0.1	0.01	0.04	0.25	0.25
1	0.35	0.1	0.02	0.05	0.03	0.1	
3	0.40	0.04	0.07	0.01	0.05	0.03	0.40
6	0.5	0	0	0	0	0	0.5

Table 2: Clinical Trial Design Points for Logistics Regression Model

v	*P1*	*P2*	*P3*	*P4*	*P5*	*P6*	*P7*	*P8*	*P9*
0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
0.25	0.18	0.15	0.05	0.05	0.08	0.01	0.15	0.15	0.18
0.5	0.24	0.08	0.09	0.05	0.16	0.03	0.07	0.04	0.24
1	0.36	0.12	0.04	0.01	0.01	0.02	0.03	0.05	0.36
3	0.41	0.03	0.01	0.03	0.01	0.02	0.08	0.06	0.35
6	0.49	0	0	0.2	0	0	0	0	0.49

analysis and regression of variance models, variance component models, factorial and fractional factorial trials, mixtures of linear, binary regression, hierarchical, and nonlinear regression models. It is also utilized in the process of developing experiments for clinical trials and sequential experimentation. Moreover, the field of Bayesian analysis and design for clinical trials, whether sequential or non-sequential, is a constantly evolving and applicable domain, with several practical advancements in real-world settings. Nevertheless, the growing utilization of Bayesian design in various applications frequently necessitates the usage of specialized software to enhance the accessibility and broad applicability of these methods [JB Kadane et al.,(2011)].

Example

Consider a design problem related to confirming results from a previous study. This study focuses on patients who have been diagnosed with breast cancer. The objective is to examine the likelihood of tumor regression in relation to the levels of protein expression. A logistic regression model relates expression levels to clinical outcomes. The design problem aims to find the optimal sample size for a new study to investigate whether a previous observed decline in regression probability was a chance occurrence. An approach to balancing costs and information is to utilize the predicted utility for model discrimination. Combining costs and model discrimination, a utility function is constructed, and the optimal sample size is determined to maximize the combined utility [RE Kass *et al.,* (1995); DV Lindley et al., (1972)].

Choice of Utility Functions

The selection of appropriate utility functions is of the utmost importance, and these responsibilities must be in accordance with the specific goals of a given circumstance. Designing for prediction can be very different from designing for discriminating between two different models, which can result in significant differences. When contemplating a one-way analysis of variance model, it is essential to keep in mind that the design that is most appropriate for comparing k treatments to a control group could be different from the design that is perfect for estimating the effects of k plus one treatment. This is because both studies have different objectives. It is important to consider the differential costs of treatments, which may result in varying sample sizes or selections of experiments.The assignment of treatments may be influenced by ethical considerations as well (Kadane, 1996) [P Müller *et al.,* (2005); MJ BAYARRI *et al.,* (1999)].

$$\theta \mid \nu 2 \sim N\left((X' PX)^{-1}(X' PY), (\nu_2 X' PX)^{-1}\right),$$

$$\nu 2 \sim Ga\left(\alpha_2 \frac{n}{1}, \beta 2 + \frac{1}{2}\{Y'PY - (Y' PX)(X' PX)^{-1}(X' PY\}\right)$$

The Bayesian approach is particularly suitable during the initial phase of our clinical research, where there is typically access to external information, such as historical data, findings from prior studies, and expert opinions, which can be effectively utilized. According to Donald Berry and his colleagues, as mentioned in [K Chaloner *et al.,* (1995)], throughout the design phase, we assess and supervise the treatment effect controls [J Pilz et al.,(1991)].

Summary

In summary, Bayesian experimental design stands at the forefront of innovative methodologies with broad applications across diverse domains. Its ability to seamlessly integrate prior information into the design process holds the key to achieving more efficient and informative experimental designs while optimizing resource utilization. Moreover, the customization of utility functions allows for the tailoring of designs to address specific objectives, ensuring maximum effectiveness. Recent advancements in simulation-based methods have further bolstered the versatility and adaptability of design specifications, opening up new avenues for experimentation. However, challenges persist, particularly concerning model uncertainty. Yet, the implementation of Bayesian model averaging presents a promising solution, mitigating the impact of uncertain model assumptions and enhancing the robustness of experimental designs. As Bayesian experimental design continues to evolve, it promises to revolutionize the way experiments are conceived and executed, driving innovation and discovery across scientific disciplines.

Conclusions

The Bayesian approach, particularly with the implementation of our MCMC model, has garnered considerable attention and acceptance within the research community and among organizations alike. As we delve deeper into our model analysis, several key points emerge that underscore the significance of Bayesian methodology in clinical trial planning and design.

First and foremost, the Bayesian paradigm is firmly established and recognized as an invaluable tool in the realm of clinical trials. Its unique capability to seamlessly integrate external information into the trial design process, adapt in real-time to evolving data, and ultimately enhance trial efficiency sets it apart from traditional approaches. This adaptability and flexibility have played a pivotal role in driving the increased adoption and acknowledgment of Bayesian methods among both scientific researchers and regulatory agencies.

Furthermore, our analysis underscores the tangible benefits that Bayesian methodology offers, not only in terms of optimizing trial design but also in facilitating more informed decision-making throughout the trial lifecycle. By harnessing the power of Bayesian inference, researchers can navigate uncertainties with greater confidence, leading to more robust and reliable trial outcomes.

In essence, the growing recognition and adoption of Bayesian approaches underscore their transformative potential in shaping the future of clinical trial research. As we continue to refine and expand our understanding of Bayesian methodology, its role in revolutionizing the landscape of clinical trials is set to become even more pronounced, paving the way for more efficient, effective, and patient-centric trial designs.

References

Atkinson, A. C. (1996). The usefulness of optimum experimental designs. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *58*(1), 59-76.

BAYARRI, M. (1999, August). BAYESIAN STATISTICS 6, pp. 53-82 JM Bernardo, JO Berger, AP Dawid and AFM Smith (Eds.) Oxford University Press, 1998. In *Bayesian Statistics 6: Proceedings of the Sixth Valencia International Meeting* (Vol. 6, p. 53). Oxford University Press.

Berry, D. A., & Stangl, D. (2018). *Bayesian biostatistics*. CRC Press.

Chaloner, K., & Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical science*, 273-304.

Clyde, M., & Chaloner, K. (1996). The equivalence of constrained and weighted designs in multiple objective design problems. *Journal of the American Statistical Association*, *91*(435), 1236-1244.

Clyde, M., Mueller, P., & Parmigiani, G. (1996). Inference and design strategies for a hierarchical logistic regression model. *STATISTICS TEXTBOOKS AND MONOGRAPHS*, *151*, 297-320.

Etzioni, R., & Kadane, J. B. (1993). Optimal experimental design for another's analysis. *Journal of the American Statistical Association*, *88*(424), 1404-1411.

Kadane, J. B. (Ed.). (2011). *Bayesian methods and ethics in a clinical trial design*. John Wiley & Sons.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the american statistical association*, *90*(430), 773-795.

Lindley, D. V. (1972). *Bayesian statistics: A review*. Society for industrial and applied mathematics.

Müller, P. (2005). Simulation based optimal design. *Handbook of Statistics*, *25*, 509-518.

Pilz, J. (1991). Bayesian estimation and experimental design in linear regression models. *(No Title)*.

Powell, C. F., Campbell, I. E., & Vapor-Plating, B. G. (1955). John Wiley and Sons, Inc.