Int. J. Nonlinear Anal. Appl. In Press, 1–9
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2024.34516.5158

Nonparametric Bayesian optimal designs for unit exponential regression model with respect to prior processes (with Polya Urn scheme as the base measure)

Anita Abdollahi Nanvapisheh, Habib Jafari*, Soleiman Khazaei

Department of Statistics, Razi University, Kermanshah, Iran

(Communicated by Zakieh Avazzadeh)

Abstract

Nonlinear regression models find extensive applications across various scientific disciplines. It is crucial to accurately fit the optimal nonlinear model while taking into account the biases inherent in the Bayesian optimal design. By utilizing the Dirichlet process as a prior, we present a Bayesian optimal design. The Dirichlet process serves as a fundamental tool in the exploration of Nonparametric Bayesian inference, offering multiple representations that are well-suited for application. This research paper introduces a novel one-parameter model, referred to as the "Unit-Exponential distribution", specifically designed for the unit interval. Additionally, we employ a representation to approximate the D-optimality criterion, considering the Dirichlet process as a functional tool. Through this approach, we aim to identify a Nonparametric Bayesian optimal design.

Keywords: D-optimal design, Nonparametric Bayesian optimal design, Unit Exponential model (UE)
2020 MSC: 60E05, 60Exx

1 Introduction

Within the realm of experimental design, the concept of optimal design refers to a specific category of designs that are classified based on certain statistical criteria. It is widely acknowledged that a well-designed experiment can significantly enhance the accuracy of statistical analyses. Consequently, numerous researchers have devoted their efforts to developing optimal designs for nonlinear regression models. Experimental design plays a pivotal role in scientific research domains, including but not limited to biomedicine and pharmacokinetics. Its application in these fields enables researchers to conduct rigorous investigations and yield valuable insights.

Optimal designs are sought using optimality criteria, typically based on the information matrix. Until 1959, research primarily focused on linear models, where the models were linear concerning the parameters. However, in nonlinear models, the presence of unknown parameters introduced complexities in the design problem, as the optimality criteria depended on these unknown parameters [3]. To address this challenge, researchers proposed various solutions, including local optimal designs [1], sequential optimal designs, minimax optimal designs, Bayesian optimal designs

*Corresponding author
Email addresses: anita.abdollahi@yahoo.com (Anita Abdollahi Nanvapisheh), h.jafari@razi.ac.ir (Habib Jafari), s.khazaei@razi.ac.ir (Soleiman Khazaei)

Received: June 2024 *Accepted:* September 2024

[15]and pseudo-Bayesian designs [14]. Chernoff [9] introduced the concept of local optimality, which involves specifying fixed values for the unknown parameters and optimizing a function of the information matrix to determine the design for these specified parameter values. This approach aimed to overcome the difficulties associated with the dependence of the design problem on unknown parameters in nonlinear models.

The selection of unknown parameters in local designs is typically obtained from previous studies or experiments specifically conducted for this purpose. The effectiveness of local designs heavily relies on the appropriate selection of these parameters. However, a significant challenge arises when the investigated problem lacks robustness concerning weak parameter estimation. To address this, an alternative approach for local optimal designs involves utilizing a prior distribution for the unknown parameters instead of relying solely on an initial guess. In the Bayesian method, the first step is to represent the available information in the form of a probability distribution for the model parameter, known as the prior distribution. A Bayesian optimal design aims to maximize the relevant optimality criterion over this prior distribution. Nevertheless, it is crucial to acknowledge that the selection of the prior distribution within the Bayesian framework can be problematic and may potentially lead to erroneous results. The choice of the prior distribution is subjective, relying on the researcher's beliefs, and it significantly influences the final outcome. Unfortunately, the Bayesian approach lacks a definitive method for selecting the prior distribution. Numerous researchers have investigated the effect of the prior distribution on determining design points in various types of optimal designs. For instance, Chaloner and Lorentz [7], Chaloner and Duncan [6], Burghaus and Dette [5], Chaloner and Vardinelli [8], Pronzato and Walter [16], Mukhopadhyay and Haines [14], Dette and Ngobauer [10], Fedorov [11], and Firth and Hinde [13] have contributed extensively to this field. Chapter 18 of Atkinson et al.'s book [2] provides further reading on this topic. Moreover, in situations where there is insufficient evidence from previous studies on the topic of interest, specifying an appropriate prior distribution becomes challenging. In such cases, subjective or noninformative prior distributions are used, incorporating all available information regarding the uncertainty of the parameter values. For more information, refer to Burghaus and Dette [5].

This research paper presents the introduction of a novel one-parameter model, referred to as the UE distribution, specifically designed for the unit interval in section 2. In Section 3, the optimal design for nonlinear models is derived. In the fourth section, the nonparametric Bayesian D-optimal design, the Dirichlet process and the Polya Urn Scheme are introduced. Finally, Section 5 concludes the paper with some closing remarks.

2 The Unit-Exponential distribution

The exponential distribution is continuous distribution in statistics and probability theory. If $Y \sim Exp(\theta)$, then using the transformation $X=\frac{Y}{1+Y}$ we have a new distribution with support on the unit-interval that the CDF and the PDF of the resulting distribution are respectively:

$$F(x \mid \theta) = 1 - Exp(\frac{-\theta x}{1-x});\ 0 \leq x < 1,\ \theta > 0, \tag{2.1}$$

$$f(x \mid \theta) = \frac{\theta}{(1-x)^2} Exp(\frac{-\theta x}{1-x});\ 0 \leq x < 1,\ \theta > 0. \tag{2.2}$$

The Hazard Rate Function (HRF) of this distribution is as follows:

$$h(x \mid \theta) = \frac{f(x \mid \theta)}{1 - F(x \mid \theta)} = \frac{\theta}{(1-x)^2};\ 0 \leq x < 1,\ \theta > 0. \tag{2.3}$$

In the following figure, the PDF and the HRF of this distribution are plotted for different values of the parameter θ. Acording to this figurs, it can be seen that the HRF is increasing in $0 \leq x < 1$,.

3 Optimal Design for Nonlinear Models

In the context of nonlinear experimental design, a common issue arises where the relationship between the response variable y and the independent variable x is given by the equation $y = \eta(x, \boldsymbol{\theta}) + \epsilon$ where $x \in \chi \subseteq \mathrm{R}$ and y is a response variable and $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ is the unknown parameter vector and ϵ is a normally distributed residual value with mean 0 and known variance $\sigma^2 > 0$. For simplicity, we assume $\sigma^2 = 1$ in this problem. If $\eta(x, \boldsymbol{\theta})$ is differentiable with respect to $\boldsymbol{\theta}$ then, the information matrix at a given point x can be represented as follows:

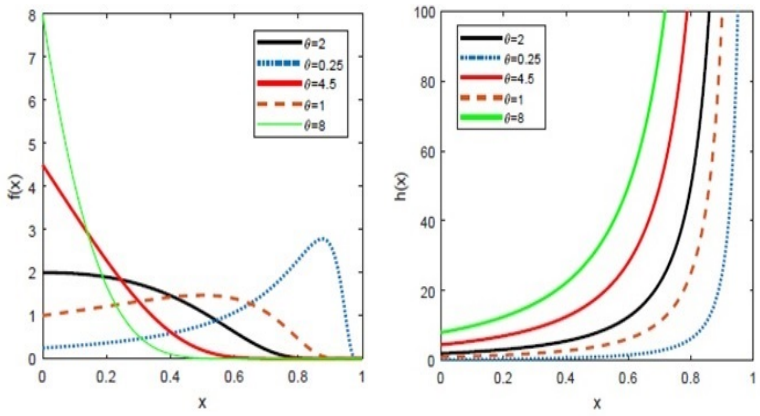

Figure 1: Plot of density function (left) and hrf (right)

$$I(\xi, \boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \eta(x, \boldsymbol{\theta}) \frac{\partial}{\partial \boldsymbol{\theta}^T} \eta(x, \boldsymbol{\theta}). \tag{3.1}$$

There exist several optimality criteria used to obtain the optimal design, including D-optimality and A-optimality. These criteria are functions of the information matrix and can be expressed as follows:

$$\Psi_D(\xi, \boldsymbol{\theta}) = -\log(\det(M(\xi, \boldsymbol{\theta}))),$$

and

$$\Psi_A(\xi, \boldsymbol{\theta}) = tr(M^{-1}(\xi; \boldsymbol{\theta})),$$

where ξ denotes a design with two components; the first component represents specific values from the design space χ and the second component corresponds to the weights assigned to these values, so that design ξ can be defined as follows:

$$\xi = \left\{ \begin{array}{cccc} x_1 & x_2 & \dots & x_\ell \\ w_1 & w_2 & \dots & w_\ell \end{array} \right\} \in \boldsymbol{\Xi}, \tag{3.2}$$

where $\boldsymbol{\Xi}$=$\{\xi \mid 0 \leq w_j \leq 1 \, ; \sum_{j=1}^{\ell} w_j = 1 \, , \, x \in \boldsymbol{\chi}\}$, [14]. When considering a discrete probability measure ξ with finite support, the information function of ξ can be expressed as follows [3]:

$$M(\xi, \boldsymbol{\theta}) = \sum_{j=1}^{\ell} w_j I(x_j, \boldsymbol{\theta}). \tag{3.3}$$

Because of the dependence of the information matrix $M(\xi, \boldsymbol{\theta})$ on the unknown parameter $\boldsymbol{\theta}$, one approach to address this issue is to employ the Bayesian method and incorporate a prior distribution for the parameter vector. The Bayesian D-optimality criterion can be formulated as follows:

$$\Psi_\Pi(\xi) = E(\psi(\xi; \boldsymbol{\theta})) = \int_\Theta \psi(\xi; \boldsymbol{\theta}) d\Pi(\boldsymbol{\theta}) = \int_\Theta -\log(det(M(\xi, \boldsymbol{\theta}))) d\Pi(\boldsymbol{\theta}), \tag{3.4}$$

where Π represents the prior distribution for $\boldsymbol{\theta}$ and the Bayesian D-optimal design is attained by minimizing (3.4). According to Dette and Neugebauer [10], in the general case of optimal designs which can include designs with two and more points, if the support of the prior distribution has n points, then the maximum number of Bayesian optimal design points is given by $n \quad \frac{p(p+1)}{2}$. Hence, in the specific scenario of nonlinear models with one parameter ($p = 1$), this implies that the support of the Bayesian optimal design does not contain more points than the support of the prior distribution.

In certain situations, specifying a prior distribution on the parameter space $\boldsymbol{\Theta}$ can be challenging for the experimenter. In such cases, an alternative approach is to consider an unknown prior distribution Π for the parameter $\boldsymbol{\theta}$. In this condition, Π is treated as a parameter itself. Consequently, equation (3.4) becomes a random functional, and it becomes necessary to determine its distribution or approximation. From a Bayesian perspective, we construct a prior distribution on the space of all distribution functions to address this issue. Ferguson [12] introduced the concept of the Dirichlet process in this context, and in section 4.1, an overview of the Dirichlet process will be provided.

4 Nonparametric Bayesian D-optimal design

Now suppose we have the following regression model:

$$E(y|x) = \eta(x, \boldsymbol{\theta}) = \frac{\theta}{(1-x)^2} \exp(\frac{-\theta x}{1-x}), 0 \leq x < 1, \theta > 0. \tag{4.1}$$

In this section, we introduce the nonparametric Bayesian optimal design. In the nonparametric Bayesian framework, it is assumed that $\theta \mid P \sim P$, where P is a random probability distribution and $P \sim \Pi$. The general method of constructing a random measure is to start with the stochastic processes. Ferguson [12] formulated the requirements which must be imposed on a prior distribution and proposed a class of prior distributions, named Dirichlet processes. One of the main arguments for using the Dirichlet distribution in practical applications is based on the fact that this distribution is a good approximation of many parametric probability distributions. Below, we define the Dirichlet process.

4.1 Dirichlet Process (DP)

To have a random distribution G distributed according to a Dirichlet process (DP), its marginal distributions must follow a Dirichlet distribution. Specifically, let H be a distribution over $\boldsymbol{\Theta}$ and α be a positive real number. For any finite measurable partition $A_1, A_2, ..., A_r$ of $\boldsymbol{\Theta}$ the vector $(G(A_1), G(A_2), ..., G(A_r))$ is random since G is random. We say G is Dirichlet process distributed with base distribution H and concentration parameter α, written $G \sim \mathrm{DP}(\alpha, H)$, if the following conditions hold:

$$(G(A_1), G(A_2), ..., G(A_r)) \sim Dir(\alpha H(A_1), ..., \alpha H(A_r)), \tag{4.2}$$

for every finite measurable partition $A_1, A_2, ..., A_r$ of $\boldsymbol{\Theta}$. The parameters of H and α play intuitive roles in the definition of the DP. The base distribution H represents the mean of the Dirichlet process, such that for any measurable set $A \subset \boldsymbol{\Theta}$ we have $E[\mathrm{G(A)}] = H(A)$. On the other hand, the concentration parameter α can be viewed as an inverse variance: $V[G(A)] = H(A)(1 - H(A))/(\alpha + 1)$. The larger α is, the smaller the variance, and the DP will concentrate more of its mass around the mean. The concentration parameter is also referred to as the strength parameter, referring to the strength of the prior when using the DP as a nonparametric prior in Bayesian nonparametric models. It can be interpreted as the amount of mass or sample size associated with the observations. It is worth noting that α and H only appear as their product in the definition of the Dirichlet process (equation 4.2). Consequently, some authors treat $\tilde{\mathrm{H}} = \alpha H$, as the same as the single (positive measure) parameter of the DP, writing $\mathrm{DP}(\tilde{\mathrm{H}})$ instead of $\mathrm{DP}(\alpha, H)$. This parametrization can be notationally convenient, but it loses the distinct roles α and H play in describing the DP.

As the concentration parameter α increases, the mass of the DP becomes more concentrated around its mean. Consequently, when α approaches infinity ($\alpha \to \infty$), $G(A)$ approaches $H(A)$ for any measurable set A, indicating weak or pointwise convergence of G to H. However, it's important to note that this does not imply a direct convergence of G to H as a whole. In fact, as we will explore later, samples drawn from a DP will typically be discrete distributions with probability one, even if the base distribution H is smooth. Therefore, G and H may not be absolutely continuous with respect to each other. Despite this, some authors still utilize the DP as a nonparametric extension of a parametric model represented by H. However, if the desire is to maintain smoothness, it is possible to extend the DP by convolving G with kernels, resulting in a random distribution with a density function.

An alternative definition of the Dirichlet process is proposed by Ferguson [12] that defines a random probability measure which is a Dirichlet process on $(\boldsymbol{\Theta}, \mathrm{B}(\boldsymbol{\Theta}))$, as:

$$P(.) = \sum_{i=1}^{\infty} p_i \delta_{\theta_i}(.), \tag{4.3}$$

where θ_i $(i > 1)$ be a sequence of *i.i.d.* random variables with common distribution Q, δ_{θ_i} represents a probability measure that is degenerate at θ where δ_{θ_i}=1 if $\theta_i \in A$ and 0 otherwise, and $p_i^{,}$s are the random weights satisfying p_i¿0 and $\sum_{i=1}^{\infty} p_i$=1. The random distribution P is discrete with probability one. Several authors have proposed alternative series representations of the Dirichlet process. Bondesson [4], Sethuraman [17], and Zarepour and Al Labadi [18] are among those who have contributed to this area. A method of producing samples from the Dirichlet process is to use the Polya urn process that in the upcoming section, we will discuss about it. Then the nonparametric Bayesian D-optimal design for the UE model is discussed.

4.2 Polya Urn Scheme

The Polya Urn Scheme was used by Blackwell and McQueen (1973) to demonstrate the existence of the Dirichlet process. The method of producing a sample of the Dirichlet process is to use a Polya Urn Scheme [5]. Consider a Polya urn with $a(\chi)$ balls of which $a(i)$ are of color i ; $i = 1, 2, ..., k$. For the moment, assume that $a(i)$'s are whole numbers or 0. Draw balls at random from the urn, replacing each ball drawn by two balls of the same colour. Let $X_i = j$ if the i The ball is of colour j. Then:

$$P(X_1 = j) = \frac{a(j)}{a(\chi)}, \tag{4.4}$$

$$P(X_2 = j \mid X_1) = \frac{a(j) + \delta_{X_1}(j)}{a(\chi) + 1}, \tag{4.5}$$

and in general

$$P(X_{n+1} = j \mid X_1, X_2, ..., X_n) = \frac{a(j) + \sum_{1}^{n} \delta_{X_i}(j)}{a(\chi) + n}. \tag{4.6}$$

That n is the number of extracted balls and $\delta_{X_i}(j)$ is equal to one if $X_i = j$, otherwise it is equal to zero.

4.3 Nonparametric Bayesian D-optimal design for UE model

Now let's consider the regression model (4.1), Therefore, the Bayesian D-optimality criterion, denoted as $\Psi_\Pi(\xi)$ can be expressed as follows:

$$\Psi_\Pi(\xi) = E(\psi(\xi;\boldsymbol{\theta})) = \int_\Theta \psi(\xi;\theta) d\Pi(\boldsymbol{\theta}) = \int_\Theta -\log(\sum_{j=1}^{\ell} w_j [\exp(\frac{-\theta x_j}{1-x_j})(\frac{1}{(1-x_j)^2} - \frac{\theta x_j}{(1-x_j)^3})]^2) d\Pi(\theta) \tag{4.7}$$

where Π is the prior distribution for θ. The Bayesian D-optimal design is attained by minimizing equation (4.7). In the nonparametric Bayesian framework, we consider $P \sim \mathrm{DP}(\alpha, P_0)$ and its collective representation as $P(.) = \sum_{i=1}^{\infty} p_i \, \delta_{\theta_i}(.)$. In this context, the optimality criterion can be expressed as follows:

$$\Psi_\Pi(\xi) = \sum_{i=1}^{\infty} p_i (-\log(\sum_{j=1}^{\ell} w_j [\exp(\frac{-\theta_i x_j}{1-x_j})(\frac{1}{(1-x_j)^2} - \frac{\theta_i x_j}{(1-x_j)^3})]^2)). \tag{4.8}$$

Chernoff [9] demonstrated that when searching for a local optimal design, there exists an optimal design where all the mass is concentrated at a single point within the design's support. Caratheodory's theorem also confirms the existence of a one-point optimal design. However, when employing the Bayesian optimality criterion, a more complex situation arises. Brice and Dette showed that with a uniform prior distribution, as the support of the prior distribution increases, the number of optimal design points for the single-parameter model also increases. Challoner suggested that if the researcher aims to obtain a one-point optimal design, it is advisable to consider a small support for the uniform prior distribution. The same principle applies to nonparametric Bayesian designs. In this case, assuming a uniform distribution over the interval [1, B] as the basic distribution, the one-point optimal design can be achieved.

Equation (4.7) is a stochastic function of the Dirichlet process. According to Ferguson's definition of the Dirichlet process, the calculation (4.8) is not easily possible, so to solve this problem in obtaining the optimal nonparametric Bayesian criterion, methods such as the stick breaking process are used to approximate this criterion. Another method

has been presented by Zarepour and Ellabadi [18] whose simulation speed and accuracy is much higher than the stick breaking process.

Since the weights produced by the stick breaking process don't follow a decreasing trend, therefore, the Dirichlet process can be simulated in a way where the weights are produced in a decreasing manner. The reason for this is that the speed of reaching the cutting point increases. Zarepour and Ellabadi presented a finite collective representation of the Dirichlet process in order to generate data from the Dirichlet process, which almost certainly converges to the Ferguson collective representation that we present below, and the weights produced from this method are uniformly descending, while the weights produced by the stick-breaking method are randomly descending. Ferguson showed that the Dirichlet process with parameters (α, P_0) can be presented using the following series representation:

$$P_n^{Ferg.}(.) = \sum_{i=1}^{\infty} \frac{N^{-1}(\Gamma_i)}{\sum_{i=1}^{\infty} N^{-1}(\Gamma_i)} \delta_{\theta_i}(.),$$

where

$$N(x) = \alpha \int_x^{\infty} \frac{\exp(-t)}{t} dt, x > 0. \tag{4.9}$$

is the Levy measure of a Gamma(α, 1) random variable and $\delta_{\theta_i}(.)$ denotes the Dirac measure. Now, in this section, we present the finite sum representation of the Dirichlet process presented by Zarepour and Ellabadi [18]. Let X_n be a random variable with distribution Gamma($\frac{\alpha}{n}$,1) and with survival and quantile function, respectively as follows:

$$G_n(x) = P(X_n > x) = \int_x^{\infty} \frac{1}{\Gamma(\frac{\alpha}{n})} \exp(-t) t^{\frac{\alpha}{n}-1} dt,$$

$$G_n^{-1}(y) = \inf\{x : G_n(x) \leq y\}, \quad 0 < y < 1.$$

According to the dominated convergence theorem, n $\to \infty$, we have:

$$nG_n(x) \to N(x).$$

Notice that the left hand side of the above quantitative is a sequence of monotone functions converging to a monotone function. We have:

$$G_n^{-1}(x/n) \to N^{-1}(y).$$

Zarepour and Ellabadi showed that for each $E_i \sim \exp(1)$, $i = 1, 2, ..., n$ and for each $\theta_i \sim P_0$, that Γ_i=$E_1 + E_2 + \cdots + E_i$ the obtained approximation almost certainly converges to Ferguson's collective representation; that's mean:

$$P_n^{New.}(.) = \sum_{i=1}^{n} \frac{G_n^{-1}(\frac{\Gamma_i}{\Gamma_{n+1}})}{\sum_{i=1}^{n} G_n^{-1}(\frac{\Gamma_i}{\Gamma_{n+1}})} \delta_{\theta_i}(.) \to P^{Ferg.}(.) = \sum_{i=1}^{\infty} \frac{N^{-1}(\Gamma_i)}{\sum_{i=1}^{\infty} N^{-1}(\Gamma_i)} \delta_{\theta_i}(.), \tag{4.10}$$

where n is as follows:

$$n = \inf \left\{ m : \frac{G_m^{-1}(\frac{\Gamma_m}{\Gamma_{m+1}})}{\sum_{i=1}^{m} G_m^{-1}(\frac{\Gamma_i}{\Gamma_{m+1}})} < \epsilon \right\}. \tag{4.11}$$

It is important to emphasize that unlike in the previously discussed truncation approximations, the weights:

$$p_i = \frac{G_n^{-1}(\frac{\Gamma_i}{\Gamma_{n+1}})}{\sum_{i=1}^{n} G_n^{-1}(\frac{\Gamma_i}{\Gamma_{n+1}})}, \tag{4.12}$$

decrease monotonically for any fixed positive integer n, which leads to the fact that the speed of simulation and its accuracy are much higher than the stick breaking process. In the following, a nonparametric Bayesian optimal design is obtained for different selections of Dirichlet process parameters. For this purpose, at first, we generate p_i from (4.12). We obtain n from (4.11) and generate $E_i \sim Exp(1)$, $i = 1, 2, \ldots, n$, and let Γ_i=$E_1 + E_2 + ... + E_i$. We calculate $G_n^{-1}(\frac{\Gamma_i}{\Gamma_{n+1}})$, $i = 1, 2, \ldots, n$ from the equation $G_n^{-1}(y)$= $\inf\{x : G_n(x) < y\}$, and generate θ_i from Base measure P_0. Finally, we evaluate the functional:

$$\Psi_P(\xi) = \sum_{i=1}^{\infty} p_i[-\log(\det(M(\xi, \theta_i)))],$$

and obtain ξ^* from the following equation:

$$\xi^* = \arg\min \Psi_P(\xi).$$

Now, in this section, we consider the Polya Urn Scheme as the base measure in DP. We get the results by using a nonlinear optimization programming with R package *Rsolnp*. For a better understanding of the effect of the α parameter, we tabulate the results for four different values of α=1, 5, 10, 50, in Tables 1-4. We also fixed ϵ=10^{-10}. Without loss of generality, we consider a bounded design space χ=[0, 1].

Tables 1-4 represent the results when the concentration parameter and uncertainty in the base measure increase. According to the results, when the value of α increases, the support points in the points design do not significantly change. The weight of the minimum point increases rapidly, and the smallest point will have the most weight. This weight almost increases or remains fixed by increasing the concentration parameter. Also, for a three-point design, the minimum support point has the greatest weight. In addition, in the range under investigation, the results show that we don't have a three-point design for $\mu = 5, \sigma = 2$, and in fact, it converts to a design with fewer points. This observation is clearer for a larger concentration parameter. But, by increasing the parameter space, optimal two three-point designs are obtained.

Table 1: Nonparametric Bayesian D-optimal designs with truncated normal base distribution and concentration parameter when α=1. First row: support points; second row: weights.

Parameters	*Design*	*Two points*		*Three points*		
$\mu = 5, \sigma = 2$	x	0.00039	0.23748	–	–	–
	w	0.99738	0.00262	–	–	–
$\mu = 50, \sigma = 30$	x	0.03358	0.18865	0.038036	0.18630	0.29563
	w	0.97085	0.02915	0.949231	0.050768	0.0000001
$\mu = 150, \sigma = 90$	x	0.01520	0.19838	0.01595	0.19625	0.29908
	w	0.99393	0.00607	0.98983	0.00813	0.00204
$\mu = 1000, \sigma = 500$	x	0.002302	0.19991	0.00275	0.20004	0.29995
	w	0.999998	0.000002	0.999999	0.0000006	0.0000005

Table 2: Nonparametric Bayesian D-optimal designs with truncated normal base distribution and concentration parameter when α=5. First row: support points; second row: weights.

Parameters	*Design*	*Two points*		*Three points*		
$\mu = 5, \sigma = 2$	x	0.00068	0.22791	–	–	–
	w	0.99734	0.00266	–	–	–
$\mu = 50, \sigma = 30$	x	0.03373	0.18516	0.03688	0.17411	0.28843
	w	0.96606	0.03394	0.91814	0.07302	0.00884
$\mu = 150, \sigma = 90$	x	0.01457	0.19748	0.14979	0.19533	0.29895
	w	0.99192	0.00808	0.9979605	0.0020304	0.0000001
$\mu = 1000, \sigma = 500$	x	0.002169	0.19994	0.00245	0.19965	0.29999
	w	0.999999	0.000001	0.9999908	0.0000001	0.0000001

Now, if we assume the mean of the base distribution to be constant and increase the variance, it can be seen that in the two-point designs, the smallest point has the most weight. The results related to this case have been presented in Table 5.

5 Concluding Remarks And Future Works

Nonlinear regression models are widely used in various scientific fields, and the Bayesian method is commonly employed to obtain optimal designs in such models. However, one of the challenges in the Bayesian framework is the

Table 3: Nonparametric Bayesian D-optimal designs with truncated normal base distribution and concentration parameter when α=10. First row: support points; second row: weights.

Parameters	*Design*	*Two points*		*Three points*		
$\mu = 5, \sigma = 2$	x	0.00031	0.23268	–	–	–
	w	0.99735	0.00265	–	–	–
$\mu = 50, \sigma = 30$	x	0.03269	0.17711	0.03922	0.17145	0.29027
	w	0.94385	0.05615	0.92004	0.07552	0.00444
$\mu = 150, \sigma = 90$	x	0.01361	0.19819	0.01601	0.19434	0.29891
	w	0.99798	0.00202	0.9879202	0.0120707	0.0000001
$\mu = 1000, \sigma = 500$	x	0.00226	0.20000	0.00251	0.19948	0.29997
	w	0.999992	0.0000008	0.9979901	0.0020004	0.0000005

Table 4: Nonparametric Bayesian D-optimal designs with truncated normal base distribution and concentration parameter when α=50. First row: support points; second row: weights.

Parameters	*Design*	*Two points*		*Three points*		
$\mu = 5, \sigma = 2$	x	0.00119	0.21958	–	–	–
	w	0.98904	0.01096	–	–	–
$\mu = 50, \sigma = 30$	x	0.03164	0.18099	0.03504	0.16948	0.29022
	w	0.96095	0.03905	0.91685	0.07641	0.00674
$\mu = 150, \sigma = 90$	x	0.01236	0.19683	0.14969	0.19641	0.29876
	w	0.99596	0.00404	0.9979801	0.0020107	0.0000002
$\mu = 1000, \sigma = 500$	x	0.00221	0.19988	0.002462	0.19979	0.29999
	w	0.9999999	0.00000001	0.9999908	0.0000006	0.0000006

Table 5: Nonparametric Bayesian D-optimal designs with truncated normal base distribution and concentration parameter when α=1. First row: support points; second row: weights.

Parameters	*Design*	*Two points*		*Three points*		
$\mu = 50, \sigma = 30$	x	0.03342	0.18889	0.03342	0.19076	0.29333
	w	0.97926	0.02074	0.96545	0.03227	0.00228
$\mu = 50, \sigma = 90$	x	0.02209	0.19534	0.02491	0.19477	0.29463
	w	0.98720	0.01280	0.98060	0.1939609	0.0000001
$\mu = 50, \sigma = 500$	x	0.00567	0.20004	0.008246	0.19944	0.29863
	w	0.99797	0.00203	0.9979821	0.0020104	0.0000005

subjective selection of the prior distribution, which can potentially lead to incorrect results. The choice of the prior distribution is often based on the researcher's beliefs, and it strongly influences the final outcome. Unfortunately, the Bayesian approach lacks a systematic method for selecting the prior distribution. To overcome these limitations and reduce reliance on restrictive parametric assumptions, nonparametric Bayesian methods are pursued. In this study, we consider the prior distribution as an unknown parameter and utilize the Dirichlet process to derive nonparametric Bayesian D-optimal designs. Specifically, we focus on a nonlinear model with one parameter, namely the Unit-Exponential distribution. We investigate the Bayesian D-optimal design for the unit exponential regression model (equation 4.1) using a truncated normal prior distribution, examining various parameter values. By adopting a nonparametric Bayesian approach and utilizing the Dirichlet process, we aim to address the challenges associated with selecting the prior distribution in Bayesian optimal design construction. This allows us to account for uncertainty and mitigate the impact of restrictive parametric assumptions, providing more flexible and robust designs for nonlinear regression models.

In this study, we focus on utilizing the Polya Urn Scheme as the base distribution in the Dirichlet process. To better understand the influence of the concentration parameter α, we present the results in tables for four different values of α=1, 5, 10, 50. These tables provide valuable insights into the nonparametric Bayesian optimal designs, showcasing the distribution of weights and support points. By analyzing the results for different values of α, we can better understand the impact of this parameter on the design outcomes. This approach allows us to explore and evaluate the performance of the nonparametric Bayesian optimal designs under varying levels of concentration parameter α.

In the investigated range, the results reveal interesting findings. For small parameter values, there are no three-point designs observed. However, by increasing uncertainty in the base measure, another optimal point is obtained with a very small weight, resulting in a design where the smallest point has the highest weight.

Moreover, as the uncertainty in the base measure and the concentration parameter in the Dirichlet process increase,

the support points in the two-point designs do not undergo significant changes. The weight of the smallest point increases rapidly, and it becomes the point with the highest weight. This weight tends to either increase or remain relatively stable with an increase in the concentration parameter.

It is important to note that this approach can be applied to other optimality criteria and various models with two or more parameters. For example, nonparametric Bayesian optimal designs using the A- or E-optimality criterion for the nonlinear model discussed in this paper, along with a Dirichlet process prior, hold potential for further research. We hope to report new results in this area in the near future.

References

[1] M. Aminnejad and H. Jafari, *Bayesian A and D-optimal designs for gamma regression model with inverse link function*, Commun. Statist. Simul. Comput. **46** (2017), 8166–8189.

[2] A.C. Atkinson, A.N. Donev, and R.D. Tobias, *Optimum Experimental Design, With SAS*, Oxford University Press, Oxford, 2007.

[3] D. Blackwell, *Discreteness of Ferguson selections*, Ann. Statist. **1** (1973), 356–358.

[4] L. Bondesson, *On simulation from infinitely divisible distributions*, Adv. Appl. Probab. **14** (1982), no. 4, 855–869.

[5] I. Burghaus and H. Dette, *Optimal designs for nonlinear regression models with respect to non-informative priors*, J. Statist. Plann. Infer. **154** (2014), 12–25.

[6] K.M. Chaloner and G.T. Duncan, *Assessment of a beta prior distribution: PM elicitation*, J. Royal Statist. Soc.: Ser. D (The Statistician) **32** (1983), no. 1-2, 174–180.

[7] K. Chaloner and K. Larntz, *Optimal Bayesian design applied to logistic regression experiments*, J. Statist. Plann. Infer. **21** (1989), 191–208.

[8] K. Chaloner and I. Verdinelli, *Bayesian experimental design: A review*, Statist. Sci. **10** (1995), 273–304.

[9] H. Chernoff, *Locally optimal designs for estimating parameters*, Ann. Math. Statist. **24** (1953), no. 4, 586–602.

[10] H. Dette and H.M. Neugebauer, *Bayesian D-optimal designs for exponential regression models*, J. Statist. Plann. Infer. **60** (1997), no. 2, 331–349.

[11] V.V. Fedorov and S.L. Leonov, *Optimal Design for Nonlinear Response Models*, CRC Press, 2013.

[12] T.S. Ferguson, *A Bayesian analysis of some nonparametric problems*, Ann. Statist. **1** (1973), no. 1, 209–230.

[13] D. Firth and J. Hinde, *On Bayesian D-optimum design criteria and the equivalence theorem in nonlinear models*, J. Royal Statist. Soc. B **59** (1997), no. 4, 793–797.

[14] S. Mukhopadhyay and L.M. Haines, *Bayesian D-optimal designs for the exponential growth model*, J. Statist. Plann. Infer. **44** (1995), no. 3, 385–397.

[15] P. Parsamaram and H. Jafari, *Bayesian D-optimal Design for the logistic regression model with exponential distribution for random intercept*, J. Statist. Comput. Simul. **86** (2016), no. 10, 1856–1868.

[16] L. Pronzato and E. Walter, *Robust experiment design via stochastic approximation*, Math. Biosci. **75** (1985), no. 1, 103–120.

[17] J. Sethuraman, *A constructive definition of Dirichlet priors*, Statist. Sinica **4** (1994), 639–650.

[18] M. Zarepour and L. Al Labadi, *On a rapid simulation of the Dirichlet process*, Statist. Probab. Lett. **82** (2012), no. 5, 916–924.