Nonparametric Bayesian optimal designs for Unit Exponential regression model with respect to prior processes(with the truncated normal as the base measure)

Anita Abdollahi Nanvapisheh[1], Soleiman Khazaei[2], Habib Jafari[3]

Abstract

Nonlinear regression models are extensively applied across various scientific disciplines. It is vital to accurately fit the optimal nonlinear model while considering the biases of the Bayesian optimal design. We present a Bayesian optimal design by utilising the Dirichlet process as a prior. 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 stick-breaking 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.

Key words: D-optimal design, Bayesian optimal design, Unit Exponential model (UE), Dirichlet process, stick-breaking prior, nonparametric Bayesian.

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 dedicated their efforts to address the challenge of constructing 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 with respect to 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, 5]. To address this challenge, researchers

[1]Department of Statistics, Razi University, Kermanshah, Iran. E-mail: anita.abdollahi@yahoo.com. ORCID: https://orcid.org/0000-0003-3248-0347.

[2]Department of Statistics, Razi University, Kermanshah, Iran. E-mail: s.khazaei@razi.ac.ir. ORCID: https://orcid.org/0000-0003-2537-9232.

[3]Department of Statistics, Razi University, Kermanshah, Iran. E-mail: h.jafari@razi.ac.ir. ORCID: https://orcid.org/0000-0001-5191-2796.

proposed various solutions, including local optimal designs [2, 7, 11, 19, 30], sequential optimal designs, minimax optimal designs, Bayesian optimal designs [28, 21-24], and pseudo-Bayesian designs [26]. Chernoff (1953) 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 in relation to 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 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 [10], Chaloner and Duncan [8], Burghaus and Dette [4], Chaloner and Vardinelli [9], Pronzato and Walter [29], Mukhopadhyay and Haines [26], Dette and Ngobauer [12, 13], Fedorov [14, 15], and Firth and Hinde [17] have contributed extensively to this field. Chapter 18 of Atkinson et al.'s book [3] 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 [4]. 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. As we know, in applied statistic, a common issue is to deal with the uncertainty phenomena observed in the interval (0, 1). For example, in real life we often encounter measures like proportion or fraction of a certain characteristic, scores of some ability tests, different index, rates, etc., which lie in the interval (0, 1). In such cases continuous distributions with domain (0, 1) are indispensable to probabilistic modeling of the phenomena. So, in regression models where the response variable is in the form of ratio, rate or percentage, we use the unit exponential regression model to model the data that are concentrated in a certain sub-interval of the range of their domains. In Section 3, the optimal design for nonlinear models is derived. Finally, Section 4 concludes the paper with some closing remarks.

2. The Unit-Exponential distribution

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

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

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

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

$$h(x \mid \theta) = \frac{f(x \mid \theta)}{1 - F(x \mid \theta)} = \frac{\theta}{(1-x)^2}; \ 0 \leq x < 1, \ \theta > 0. \tag{3}$$

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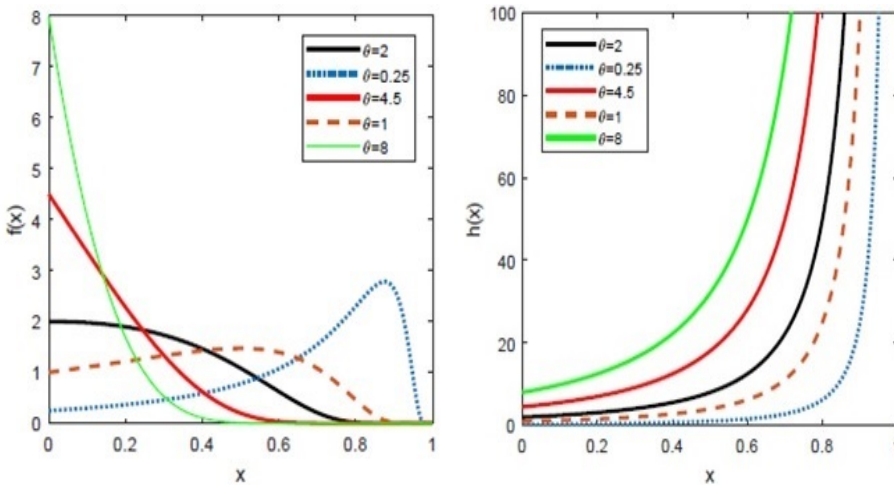

Figure 1: Plot of density function (left) and hrf (right)

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 $y = \eta(x, \theta) + \varepsilon$ where $x \in \chi \subseteq \mathbb{R}$ and y is a response variable and $\theta \in \Theta$ 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 $\eta(x, \theta)$ is differentiable with respect to θ, then the information matrix at a given point x can be represented as follows:

$$I(\xi, \theta) = \frac{\partial}{\partial \theta} \eta(x, \theta) \frac{\partial}{\partial \theta^T} \eta(x, \theta). \tag{4}$$

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:

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

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:

$$\xi = \left\{ \begin{array}{cccc} x_1 & x_2 & \ldots & x_l \\ w_1 & w_2 & \ldots & w_l \end{array} \right\} \in \Xi, \tag{5}$$

where $\Xi = \{\xi \mid 0 \le w_j \le 1 ; \sum\limits_{j=1}^{l} w_j = 1 , x \in \chi\}$, [25].

When considering a discrete probability measure ξ with finite support, the information function of ξ can be expressed as follows [3]:

$$M(\xi, \theta) = \sum_{j=1}^{l} w_j I(x_j, \theta). \tag{6}$$

Because of the dependence of the information matrix $M(\xi, \theta)$ to the unknown parameter θ, 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:

$$\Psi_\Pi(\xi) = E(\psi(\xi; \theta)) = \int_\Theta \psi(\xi; \theta) d\Pi(\theta) = \int_\Theta -log(det(M(\xi, \theta))) d\Pi(\theta), \tag{7}$$

where Π represents the prior distribution for θ and the Bayesian D-optimal design is attained by minimizing (7). According to Dette and Neugebauer [11], 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 $n \frac{p(p+1)}{2}$. 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 Θ can be challenging for the experimenter. In such cases, an alternative approach is to consider an

unknown prior distribution Π for the parameter θ. In this condition, Π is treated as a parameter itself. Consequently, equation (7) 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 (1973) introduced the concept of the Dirichlet process in this context, an overview of which will be provided in Section 3.1.1.

3.1. Nonparametric Bayesian D-optimal design

In this section, we introduce the nonparametric Bayesian optimal designl. In the nonparametric Bayesian framework, it is assumed that $\theta \mid P \sim P$, where P is a random probability distribution and $P \sim \Pi$. A general method of construction of a random measure is to start with the stochastic processes. Ferguson (1973) formulated the requirements which must be imposed on a prior distribution and proposed a class of prior distributions, named the Dirichlet processes. One of the main argument in 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 give the definition of the Dirichlet process.

3.1.1 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 Θ and α be a positive real number. For any finite measurable partition $A_1, A_2, ..., A_r$ of Θ the vector $(G(A_1), G(A_2), ..., G(A_r))$ is random since G is random. We say G is the Dirichlet process distributed with base distribution H and concentration parameter α, written $G \sim \text{DP}(\alpha, H)$, if the following conditions hold:

$$(G(A_1), G(A_2), ..., G(A_r)) \sim Dir(\alpha H(A_1), ..., \alpha H(A_r)), \tag{8}$$

for every finite measurable partition $A_1, A_2, ..., A_r$ of Θ.

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 $A \subset \Theta$ we have $E[G(A)] = H(A)$. 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 modelsl, 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 8). Consequently, some authors treat $\tilde{H} = \alpha H$, as the single (positive measure) parameter of the DP, writing DP($\tilde{H}$) instead of DP(α, H). This parametrization can be notationally convenient, but 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 is 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 the 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 [16], who, defined a random probability measure, which is a Dirichlet process on $(\Theta, B(\Theta))$, as:

$$P(.) = \sum_{i=1}^{\infty} p_i \delta_{\theta_i}(.), \tag{9}$$

where θ_i $(i > 1)$ is 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 $\delta_{\theta_i}=1$ 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. Sethuraman [31], and Zarepour and Al Labadi [32] are among those who have contributed to this area. In the upcoming section, we will discuss the nonparametric Bayesian D-optimal design for the UE model.

3.1.2   Nonparametric Bayesian D-optimal design for UE model

Now, let us consider the following regression model:

$$E(y|x) = \eta(x, \theta) = \frac{\theta}{(1-x)^2} \exp\left(\frac{-\theta x}{1-x}\right), \theta > 0. \tag{10}$$

Therefore, the Bayesian D-optimality criterion, denoted as $\Psi_{\Pi}(\xi)$, can be expressed as follows:

$$\Psi_{\Pi}(\xi) = E(\psi(\xi; \theta)) = \int_{\Theta} \psi(\xi; \theta) d\Pi(l\theta) \tag{11}$$

$$= \int_{\Theta} -log\left(\sum_{j=1}^{l} w_j \left[\exp\left(\frac{-\theta x_j}{1-x_j}\right)\left(\frac{1}{(1-x_j)^2} - \frac{\theta x_j}{(1-x_j)^3}\right)\right]^2\right) d\Pi(\theta) \tag{12}$$

where Π is the prior distribution for θ. The Bayesian D-optimal design is attained by minimizing equation (11). 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:

$$\Psi_{\Pi}(\xi) = \sum_{i=1}^{\infty} p_i \left(-log\left(\sum_{j=1}^{l} w_j \left[\exp\left(\frac{-\theta_i x_j}{1-x_j}\right)\left(\frac{1}{(1-x_j)^2} - \frac{\theta_i x_j}{(1-x_j)^3}\right)\right]^2\right)\right). \qquad (13)$$

Chernoff [7] 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. Braess 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 (11) represents a stochastic function of the Dirichlet process. According to Ferguson's definition of the Dirichlet process, the direct calculation of (12) is not straightforward. To address this challenge and obtain an approximation of the optimal nonparametric Bayesian criterion, methods such as the stick-breaking process is employed [31]. Saturaman (1994) introduced this method as a significant approach for generating realizations of the Dirichlet process, which we will explain below. Additionally, we will highlight the discreteness of the Dirichlet process within the framework of the stick-breaking process. To generate a realization of the Dirichlet process P with a concentration parameter α and base distribution H we can follow the stick-breaking process.

First, we generate a sequence of random samples $\theta_1, \theta_2, \ldots$ from the base distribution H. Additionally, we generate a sequence of random samples $V_1, V_2, \ldots$ from the $Beta(1, \alpha)$ distribution. We define a sequence of probabilities $p_1, p_2, \ldots, p_k, \ldots$ as follows. We start by choosing a point called V_1 on a unit-length piece of wood and set p_1 equal to V_1. In other words, $p_1 = V_1$. Then, we divide the remaining part of the wood into two parts, $V_2(1 - V_1)$ and $(1 - V_1)(1 - V_2)$. We consider the first part as p_2. To calculate p_3, we divide the remaining part of the wood into two parts in the same manner as in step 2. We continue this process, dividing each remaining part into two parts and assigning the first part as the next weight in the sequence. By following these steps, we obtain a sequence of weights $p_1, p_2, \ldots, p_k, \ldots$ that represents the probabilities associated with the generated samples $\theta_1, \theta_2, \ldots$. This sequence of weights reflects the stick-breaking process used to approximate the Dirichlet process. So:

$$p_1 = V_1,$$
$$p_i = V_i \prod_{j=1}^{i-1} (1 - V_j), \quad i \geq 2$$

According to the structure described, it can be proven that $\sum_{i=1}^{\infty} p_i = 1$. For this purpose we have:

$$1 - \sum_{i=1}^{\infty} p_i = 1 - V_1 - V_2(1 - V_1) - V_3(1 - V_2)(1 - V_1) - \ldots$$

$$= (1 - V_1)(1 - V_2 - V_3(1 - V_2) - \ldots)$$

$$\vdots$$

$$= \prod_{i=1}^{\infty} (1 - V_i) \tag{14}$$

By problem 32 in Chapter 1 of Folland (1999), we have [18]:

$$\prod_{i=1}^{\infty} (1 - V_i) = 0 \Leftrightarrow \sum_{i=1}^{\infty} V_i = \infty$$

So, for every $\varepsilon \in (0,1)$ we can write the following relation:

$$\sum_{i=1}^{\infty} Pr(V_i > \varepsilon) = \infty$$

And using Borel-Cantelli's Lemma, we will have:

$$Pr(V_i > \varepsilon, i.o) = 1 \Rightarrow \sum_{i=1}^{\infty} V_i = \infty \quad a.s \tag{15}$$

Therefore, by setting the relation (3.10) equal to zero, we will have $\sum_{i=1}^{\infty} p_i = 1$.

In this section, we focus on the use of a truncated normal distribution as the base measure in the DP. To obtain the results, we employ nonlinear optimization programming using the R package Rsolnp. The nonparametric Bayesian optimal designs are examined using the stick-breaking method, and tables presenting the results are provided. To better understand the influence of the α parameter, we present the results for four different values of $\alpha=1, 5, 10, 50$. It is important to note that we consider a bounded design space $\chi=[0,1]$ without any loss of generality. Tables 4-7 display the results obtained when the concentration parameter (α) and uncertainty in the base measure increase. Based on these results, we can observe in the class of two-point design, that largest weight corresponds to the smallest point. This pattern is consistent across the investigated range of α values. According to the results, when the value of α increases, the support points in two-point design do not significantly change. The smallest point will have the most weight that this weight almost increases or remains fixed by increasing the concentration parameter. Also, for three-point design, minimum support point has the greatest weight. In addition, in the range under investigation, the results show that we do not have a three-point design for $\mu = 5, \sigma = 2$, and in fact, it converts to the design by less points. This observation is more clear for larger concentration parameter. But, by increasing the parameter space, optimal two and three-point design are obtained.

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

Parameters	Design	Two − point		Three − point		
$\mu = 5, \sigma = 2$	x	0.0000000	0.3419671	−	−	−
	w	0.9999995	0.0000005	−	−	−
$\mu = 50, \sigma = 30$	x	0.0245942	0.2728781	0.0341488	0.2776520	0.5086153
	w	0.9696928	0.0303072	0.9789477	0.0210526	0.0000007
$\mu = 150, \sigma = 90$	x	0.0065494	0.2995342	0.01578005	0.2997138	0.5013164
	w	0.9999903	0.0000097	0.9999994	0.0000003	0.0000003

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

Parameters	Design	Two − point		Three − point		
$\mu = 5, \sigma = 2$	x	0.0000000	0.3233669	−	−	−
	w	0.9999995	0.0000005	−	−	−
$\mu = 50, \sigma = 30$	x	0.0204877	0.2772758	0.03387816	0.2638516	0.5001318
	w	0.9799968	0.0200032	0.9494947	0.0505048	0.0000005
$\mu = 150, \sigma = 90$	x	0.0009694	0.2993877	0.01462319	0.3000076	0.4991483
	w	0.9999854	0.0000146	0.9999999	0.0000004	0.0000004

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

Parameters	Design	Two − point		Three − point		
$\mu = 5, \sigma = 2$	x	0.0000000	0.3021963	−	−	−
	w	0.9999995	0.0000005	−	−	−
$\mu = 50, \sigma = 30$	x	0.0156330	0.2706337	0.0257019	0.2071970	0.5050722
	w	0.9898957	0.0101043	0.9265122	0.0734868	0.0000010
$\mu = 150, \sigma = 90$	x	0.0006769	0.2990424	0.0126487	0.2992510	0.5007835
	w	0.9863551	0.0136449	0.9999868	0.0000135	0.0000007

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 − point		Three − point		
$\mu = 5, \sigma = 2$	x	0.0000000	0.3030561	−	−	−
	w	0.9999995	0.0000005	−	−	−
$\mu = 50, \sigma = 30$	x	0.0132530	0.2859840	0.0236265	0.2357064	0.5016003
	w	0.9999973	0.0000027	0.9361683	0.06383056	0.0000001
$\mu = 150, \sigma = 90$	x	0.0000608	0.2990344	0.0107339	0.2991683	0.5012125
	w	0.9999865	0.0000135	0.9999959	0.0000020	0.0000021

Table 5 presents the results when assuming a constant mean of the base distribution and increasing the variance. Specifically, in the two-point designs, it can be observed that the smallest point has the highest weight. This table provides insights into the distribution of weights in this scenario.

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 − point		
$\mu = 50, \sigma = 2$	x	0.0000000	0.3000000	−	−	−
	w	0.9999942	0.0000059	−	−	−
$\mu = 50, \sigma = 30$	x	0.0237781	0.2842176	0.0384189	0.2794133	0.5005586
	w	0.9795880	0.0204120	0.9587626	0.04123712	0.0000002
$\mu = 50, \sigma = 90$	x	0.0108601	0.2875706	0.0257537	0.2810997	0.4938304
	w	0.9899937	0.0100063	0.9791666	0.02083332	0.0000002

4. 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 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 10) 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 two-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. These designs can be considered as one-point designs, as the weight of the additional point becomes negligible.

In three-point designs, similar observations can be made. In some cases, two of the obtained optimal points are very similar, leading to a design with fewer points. This indicates that the additional point does not significantly contribute to the design in such cases.

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.

Acknowledgements

References

Abdollahi, A., H. Jafari and S.Khazaei, (2024). Locally, Bayesian and Nonparametric Bayesian optimal designs for Unit Exponential regression model. *Journal of Communications in Statistics – Theory and Methods*, doi.org/10.1080/03610926.2024.2328182.

Aminnejad, M., Jafari H., (2017). Bayesian A and D-optimal designs for gamma regression model with inverse link function. *Communications in Statistics-Simulation and Computation*, 46 (10), pp. 8166–89, doi:10.1080/03610918.2016.1271888.

Atkinson, A. C., A. N. Donev and R. D. Tobias, (2007). Optimum experimental design, with SAS. *Oxford*, UK: Oxford University Press.

Burghaus, I., Dette H., (2014). Optimal designs for nonlinear regression models with respect to non-informative priors. *Journal of Statistical Planning and Inference*, 154, pp. 12–25, doi:10.1016/ j.jspi.2014.05.009.

Burkner P. C., Schwabe R., Holling H., (2019). Optimal designs for the generalized partial credit model. *British Journal of Mathematical and Statistical Psychology*, doi:10.1111/ bmsp. 12148.

Braess, D., Dette, H., (2007). On the number of support points of maximin and Bayesian optimal designs. *The Annals of Statistics*, 35(2), pp. 772–792.

Chernoff, H., (1953). Locally optimal designs for estimating parameters. *The Annals of Mathematical Statistics*, 24 (4), pp. 586–602, doi:10.1214/aoms/1177728915.

Chaloner, K. M., Duncan, G. T., (1983). Assessment of a beta prior distribution:PM elicitation. *The Statistician*, pp. 174–180.

Chaloner, K., Verdinelli, I., (1995). Bayesian experimental design: a review. *Statistical Science*, 10, pp. 273–304.

Chaloner, K., Larntz, K., (1989). Optimal Bayesian design applied to logistic regression experiments. *Journal of Statistical Planning and Inference*, 21, pp. 191–208.

Dette, H., Melas, V. B. and Wong, W. K., (2006). *Locally D-optimal designs for exponential regression Statistica Sinica*, 16, pp. 789–803.

Dette, H., Neugebauer, H. M., (1996). Bayesian optimal one point designs for one parameter nonlinear models. *Journal of Statistical Planning and Inference*, 52(1), pp. 17–31.

Dette, H., Neugebauer, H. M., (1997). Bayesian D-optimal designs for exponential regression models. *Journal of Statistical Planning and Inference*, 60(2), pp. 331–349.

Fedorov V, Hackl P., (2012). Model-Oriented Design of Experiments. *Springer Science and Business Media*, Vol. 125, doi:10.1007/978-1-4612-0703-0.

Fedorov V. V., Leonov S. L., (2013). Optimal design for nonlinear response models. *CRC Press. bability Letters*, 82 (5), pp. 916–24, doi:10.1016/j.spl.2012.01.020.

Ferguson, T. S., (1973). A bayesian analysis of some nonparametric problems. *The Annals of Statistics*, 1 (2), pp. 209–30, doi:10.1214/aos/1176342360.

Firth D, Hinde J., (1997). On Bayesian D-optimum Design Criteria and the Equivalence Theorem in Non-linear Models. *Journal of the Royal Statistical Society B*, 59(4), pp. 793–797, doi:10.1111/1467-9868.00096.

Folland, G. B., (1999). Real analysis: modern techniques and their applications. John Wiley and Sons.

Ford, I., Torsney, B. and Wu, C. F. J., (1992). The use of a canonical form in theconstruction of locally optimal designs for non-linear problems. *J. Roy. Statist. Soc. (Ser.) B*, 54, pp. 569–583.

Ghitany M. E. , Atieh B., Nadarajah S., (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, 78, pp. 493–506.

Goudarzi, M., H. Jafari and Khazaei S., (2019). Nonparametric Bayesian optimal designs for exponential regression model. *Communications in Statistics-Simulation and Computation*, doi:10.1080/03610918.2019.1593454.

Grashoff U., Holling H., Schwabe R., (2012). Optimal designs for the Rasch model. *Psychometrika*, 77(4), pp. 710–723, doi:10.1007/s11336-012-9276-2.

Kiefer, J., (1959). Optimum experimental designs. *J. R. Statist. Soc. B.*, 21, pp. 272–319.

Kiefer, J., Wolfowitz, J., (1959). Optimum designs in regression problems. *The Annals of Mathematical Statistics*, 30(2), pp. 271–294.

Kiefer, J., (1974). General equivalence theory for optimum designs (approximate theory). *The Annals of Statistics*, 2 (5), pp. 849–79, doi:10.1214/aos/1176342810.

Mukhopadhyay, S., Haines, L. M., (1995). Bayesian D-optimal designs for the exponential growth model. *Journal of Statistical Planning and Inference*, 44(3), 385–397.

Nadar M., Papadopoulos A., Kızılaslan F., (2013). Statistical analysis for Kumaraswamy's distribution based on record data. *Stat Pap*, 54, pp. 355–369.

Parsamaram, P., Jafari H., (2016). Bayesian D-optimal Design for the logistic regression model with exponential distribution for random intercept. textitJournal of Statistical Computation and Simulation, 86, pp. 1856–68, doi:10.1080/00949655.2015.1087525.

Pronzato, L, E. walter, (1985). Robust experiment design via stochastic approximation. *Mathematical BIOSCIENCES*, doi.org/10.1016/0025-5564(85)90068-9.

Rodriguez-Torreblanca C., Rodríguez-Díaz J. M., (2007). Locally D- and c-optimal designs for Poisson and negative binomial regression models. *Metrika*, Vol. 66, pp 161–172.

Sethuraman, J., (1994). A Constructive De nition of Dirichlet Priors. *Statistica Sinica 4*, p. 639, p. 650.

Zarepour, Mahmoud and Al Labadi, Luai, (2012). On a rapid simulation of the Dirichlet process. *Statistics and Probability Letters 82.5*, p. 916, p. 924.