Title
Optimal designs for health risk assessments using fractional polynomial models.

Journal
Stochastic Environmental Research and Risk Assessment, 36(9)

ISSN
1436-3240

Authors
Casero-Alonso, Víctor
López-Fidalgo, Jesús
Wong, Weng

Publication Date
2022-09-01

DOI
10.1007/s00477-021-02155-1

HHS Public Access
Author manuscript
Stoch Environ Res Risk Assess. Author manuscript; available in PMC 2023 September 01.

Published in final edited form as:
Stoch Environ Res Risk Assess. 2022 September ; 36(9): 2695–2710. doi:10.1007/s00477-021-02155-1.

Optimal designs for health risk assessments using fractional polynomial models

Víctor Casero–Alonso,
Department of Mathematics, Castilla–La Mancha University, Spain.

Jesús López–Fidalgo,
DATAI institute, Navarra University, Spain

Weng Kee Wong
Department of Biostatistics, UCLA, USA.

Abstract

Fractional polynomials (FP) have been shown to be more flexible than polynomial models for fitting data from an univariate regression model with a continuous outcome but design issues for FP models have lagged. We focus on FPs with a single variable and construct *D*-optimal designs for estimating model parameters and *I*-optimal designs for prediction over a user-specified region of the design space. Some analytic results are given, along with a discussion on model uncertainty. In addition, we provide an applet to facilitate users find tailor made optimal designs for their problems. As applications, we construct optimal designs for three studies that used FPs to model risk assessments of (a) testosterone levels from magnesium accumulation in certain areas of the brains in songbirds, (b) rats subject to exposure of different chemicals, and (c) hormetic effects due to small toxic exposure. In each case, we elaborate the benefits of having an optimal design in terms of cost and quality of the statistical inference.

Keywords

Approximate design; *D*-optimal design; Equivalence theorem; *I*-optimal design; Mathematica applet

1 Introduction

Polynomial models of low degrees are widely used to model a continuous response using a couple of covariates. Royston and Altman (1994) noted that polynomial models tend to fit poorly at the extreme values of the covariates and fitting data using high order

victormanuel.casero@uclm.es .

[6]Declarations
Not applicable, except:

polynomials often has ill-conditioning problems. They proposed polynomials with fractional powers, called them fractional polynomials (FPs) and demonstrated their utility for fitting data from the biological sciences. The popularity of FPs can be seen in the huge number of citations of Royston and Altman (1994). There are more than 1200 at the time of writing this paper and they come from different disciplines, particularly, in the epidemiology literature and the biological sciences. For instance, in epidemiology, Royston, Ambler, and Sauerbrei (1999) showed that several continuous risk and confounding variables can be used in FP models to assess risk accurately in epidemiological studies and Shkedy, Aerts, Molenberghs, Beutels, and Van Damme (2006) proposed FPs to model the force of infection for infectious diseases using cross-sectional seroprevalence data. Other areas of applications of FPs include microbiology (Namata, Aerts, Faes, & Teunis, 2008) oncology (Atzpodien, Royston, Stoerkel, & Reitz, 2007; Royston, Reitz, & Atzpodien, 2006; Royston & Sauerbrei, 2004; Sauerbrei, Royston, Bojar, Schmoor, & Schumacher, 1999), rheumatology (Krishnan, Tugwell, & Fries, 2004; Wolfe, 2000) or toxicology (Groten et al., 1997), to name a few. Additionally, Mayer, Keller, Syrovets, and Wittau (2016) used FP models to estimate half-life periods in antibiotic tissue concentrations in visceral surgery, nephropharmacology and clinical pharmacology.

In environmental sciences, Knafl (2015) gave an example of using FPs to model mercury level (in ppt) in fish depending on weight of the fish caught in two rivers. Related with risk assessment Silke, Kellett, Rooney, Bennett, and O'riordan (2010) proposed an improved medical admissions risk system using multivariable fractional polynomial logistic regression modelling. And Austin, Park-Wyllie, and Juurlink (2014) described the relationship between cumulative duration of use of amiodarone (an antiarrhythmic drug) and the risk of thyroid dysfunction. Finally about risk assessment in toxicity studies, Geys, Molenberghs, Declerck, and Ryan (2000) and Faes, Geys, Aerts, and Molenberghs (2003) considered FP models for dose-response modelling to determine safe dose levels of an exposure on developing fetuses. We will use in this work these latter studies.

In addition, several software statistical packages provide an option for fitting FP models. For example, the commercial software package STATA has a built in command *fp* for fitting FP models and in R, there is a package, *mfp*, for fitting FPs and studying influence of the continuous covariates on the outcome in the regression model. This R package was updated by Axel Benner in 2015 and now allows for models with binary or categorical variables, which are not subject to FP transformation. And more recently, in 2021, have been uploaded to CRAN the *bfp* package which implements the Bayesian paradigm for FPs. This suggests that FP models are gaining recognition as an modelling tool in statistics.

Despite the increasing use of FPs model, design issues for such models have not been addressed. Our goal is to develop optimal designs for estimating parameters in a FP model or for estimating its mean response surface across a user-specified region S with possibly different interests in different parts in S. Additionally, we create an online interactive tool to facilitate practitioners, who may not be familiar with the design theory, to implement the optimal designs. The tool is an interactive applet to find and compare optimal designs for FP models up to degree 3. The applet, OEDforFPmodels, is available at https://github.com/victormanuelcasero/OED_FPmodels and is created from Mathematica.

Section 2 discusses FP models and briefly reviews fundamentals for finding and verifying whether a design is optimal. Section 3 constructs *D*-optimal designs for estimating all parameters in the model and *I*-optimal designs for estimating the mean response averaged over a user-specified region. In addition, Section 3 discusses design issues when there is model uncertainty in the mean function. Section 4 applies our design techniques to find a few types of optimal designs for FP models useful for biomedical, environmental and toxicity studies. Section 5 offers a summary where we also mention challenging design issues for FP models. The appendix contains proofs of theorems for our optimal designs for low degree FP models.

2 Fractional polynomial models and design fundamentals

We first give an overview of FP models, describe approximate designs and present design optimality criteria. We then discuss the theory behind the construction of the optimal designs and present properties of *D*- and *I*-optimal designs.

2.1 Fractional polynomial models

Let x be a continuous variable assumed to take on only positive values, let m be a positive integer and let $\mathbf{p} = (p_1,\ldots,p_m)$ be a real-valued vector of powers ($p_1 \le \ldots \le p_m$). A fractional polynomial (FP) is given by $\phi_m(x;\alpha,\mathbf{p}) = \alpha_0 + \sum_{j=1}^{m} \alpha_j H_j(x)$, where α_j are the real-valued coefficients and $H_j(x)$ are defined sequentially, $H_1(x) = x^{(p1)}$,

$$H_j(x) = \begin{cases} x^{(p_j)}, & \text{if } p_j \neq p_{j-1}, \\ H_{j-1}(x)\ln[x], & \text{if } p_j = p_{j-1}, \end{cases} \text{ for } j = 2,\ldots,m. \quad (1)$$

The powers are defined by $x^{(pj)} = x^{pj}$ if $p_j \neq 0$, otherwise $x^{(0)} = \ln[x]$. For applications, Royston and Altman (1994) recommended 'powers' in a FP be selected from the set $\{-2,-1,-0.5,0,0.5,1,2,\ldots,\max(3,m)\}$.

The degree of the fractional polynomial is the number of terms in the FP not counting the intercept term. The value of m is user-selected and its value suggests how large the class of FP models we are willing to consider as a plausible model for the study. A larger integer-valued of m suggests that we are willing to consider a more complex FP model with more terms for inference purposes. We focus on FP models of degree $m = 3$ or lower because FPs of degree 3 or lower are widely used in practice. We denote FP models of degree 1, 2 and 3 by FP1, FP2 and FP3, respectively, and more specific ones by their explicit powers taken from $\mathscr{P} = \{-2,-1,-0.5,0,0.5,1,2,3\}$. There are 8 FP1 models, 28 FP2 models with different values of p_1 and p_2, hereafter $p_1 = p$ and $p_2 = q$, and 36 FP2s when repeated values of p and q are allowed. Specifically, if y is the continuous outcome variable, 7 FP2 models have mean outcome given by $Ey = \alpha_0 + \alpha_1 \ln[x] + \alpha_2 x^q$ when $0 = p < q$ or $Ey = \alpha_0 + \alpha_1 x^p + \alpha_2 \ln[x]$ when $p < q = 0$, where Ey stands for the expectation of y. There are 8 FP2s with repeated values of p and q and their mean outcome is given by $Ey = \alpha_0 + \alpha_1 x^p + \alpha_2 x^p \ln[x]$ if $p \neq 0$ and $Ey = \alpha_0 + \alpha_1 \ln[x] + \alpha_2 \ln^2[x]$ if $p = 0$. Table 1 shows all possible FP2 models. Figure 1 in Royston and Altman (1994) displays shapes of various

FP2 models. It is clear that some of them cannot be modeled satisfactorily by polynomials or FP1 models.

FP3 models are more flexible than FP1 and FP2 models for approximating the shapes of the mean response. When all the powers are different and nonzero, the mean response is $Ey = \alpha_0 + \alpha_1 x^p + \alpha_2 x^q + \alpha_3 x^r$ with $p, q, r \in \mathscr{P}$ and $p < q < r$. Other mean functions for different values of p,q and r can be worked out directly using the rule near (1).

2.2 Approximate designs and design criteria

Throughout we consider approximate designs, which are probability measures ξ defined on a user-selected design space Ω. An approximate design ξ has k points at $x_1,\ldots,x_k \in \Omega$ and positive weight w_i at $x_i, i = 1,\ldots,k$ such that $w_1 + \ldots + w_k = 1$. We assume the sample size n is predetermined either by budgetary constraint or the number of subjects available in the study. Such an approximate design ξ is implemented by taking $[nw_i]$ observations at $x_i, i = 1,\ldots,k$, where each $[nw_i]$ is a positive integer rounded from nw_i subject to $[nw_1] + \ldots[nw_k] = n$. For instance in a dose response study, Ω typically represents the dose interval of interest and $[nw_i]$ represents the number of subjects assigned to dose x_i. Approximate designs were proposed by Kiefer in 1950's and while they were controversial at that time, they are now commonly used as benchmarks for designing studies when we have a parametric model (J. C. Kiefer, 1985).

Our regression function is $f(x)^T = (f_1(x),\ldots, f_t(x))$ and its components are linearly independent. For FP models of degree m, the number of parameters in the model is $t = m + 1$. We assume all errors are independent and normally distributed, each with mean 0 and common variance, and note these assumptions can be relaxed when finding optimal designs. Following convention, we measure the worth of a design by its information matrix. This $t \times t$ information matrix constructed from an observation at x is proportional to $f(x)f(x)^T$ and the normalized information matrix from an approximate design ξ is proportional to

$$M(\xi) = \int_\Omega f(x)f(x)^T \xi(dx).$$

Several commonly used design criteria are formulated as convex functions of the information matrix. For example, the popular D-optimality criterion given by $\Phi_D(\xi) = -\ln|M(\xi)|$ is a convex function of the information matrix and the design ξ_D that minimizes the criterion over all approximate designs on Ω is called D-optimal. D-optimal designs minimize the volume of the confidence ellipsoid for the model parameters when errors are normally distributed and so provide the most accurate inference for model parameters in the mean function.

Another useful convex design criterion is I-optimality for predicting the mean response at a given point x_0. The variance of the prediction of the response using design ξ is proportional to

$$d(x_0, \xi) = f(x_0)^T M(\xi)^{-1} f(x_0) = tr\, f(x_0) f(x_0)^T M(\xi)^{-1}$$

where tr is the matrix trace, and so the best design for making inference at x_0 is the design that minimizes $d(x_0,\xi)$ among all approximate designs on Ω. If x_0 is inside Ω, we have an interpolation design problem and if it is outside of Ω, we have an extrapolation design problem (J. Kiefer & Wolfowitz, 1964). Extrapolation design problems typically arise in dose response studies where there is interest to make inference on the mean response at a dose, which frequently is outside the known safety limits of the drug. As another example, in the maintenance of a nuclear plant, it is desirable to have observations at extremely low temperature settings of certain variables. However, in practice, it is frequently problematic or risky to obtain such data and they will have to be inferred by extrapolation.

For estimating the mean response over a user-selected region S with varying interest, an appropriate design criterion is I-optimality defined by

$$\Phi_I(\xi) = \int_S f(x)^T M(\xi)^{-1} f(x)\mu(dx) = \int_S d(x,\xi)\mu(dx) = tr\ AM(\xi)^{-1}, \quad (2)$$

where $A = \int_S f(x)f(x)^T \mu(dx)$ and μ is a user-selected weighting measure over S that assigns greater weights to more interesting areas in S. For example, if there is equal interest throughout S, μ is the uniform measure. For a given weight measure μ, the design ξ_I that minimizes $\Phi_I(\xi)$ over all approximate designs on Ω is called I-optimal. The I-optimality criterion is also differentiable, and algorithms are available for finding I-optimal designs.

2.3 Theoretical and computational tools

For convex design criteria, equivalence theorems are available to check optimality of a design. Equivalence theorems are based on directional derivatives and are discussed widely in design monographs (Fedorov, 1972; Pázman, 1986, among others). A convex analysis argument shows that if there are t parameters in the mean function, the design ξ_D is D-optimal if and only if for all $x \in \Omega$,

$$f(x)^T M(\xi_D)^{-1} f(x) - t \le 0. \quad (3)$$

Similarly the design ξ_I is I-optimal if and only if for all $x \in \Omega$

$$f(x)^T M(\xi_I)^{-1} A M(\xi_I)^{-1} f(x) - tr\ AM(\xi_I)^{-1} \le 0. \quad (4)$$

Further, if the design is optimal among all designs on Ω, the above inequalities become an equality at the design points. The functions on the left hand sides of the inequalities are sometimes called sensitivity functions. To verify if a design ξ is D- or I-optimal, one plots the corresponding sensitivity function and observes if the function has the desired properties. With one variable, the optimality of a design can be quickly confirmed by plotting the sensitivity function across Ω.

Algorithms are available for generating D- and I-optimal designs and equivalence theorems can also be used to confirm optimality of the designs. The theorems can also evaluate the proximity of the generated design to the optimum if the generated design is not optimal, or more general, evaluate the proximity of any design to the optimum without knowing the optimum, including designs obtained when the algorithm is prematurely terminated, see page 128 of Pázman (1986) for details when he discussed D- and I-optimality.

The proximity of a design ξ to the optimal design is measured in terms of design efficiency. For example, the D-efficiency and I-efficiency of a design ξ are, respectively, defined by

$$\left(\frac{|M(\xi)|}{|M(\xi_D)|}\right)^{1/t} \text{ and } \frac{\Phi_I(\xi_I)}{\Phi_I(\xi)}.$$

The reason for having the ratio of the determinants in the D-efficiency raised to the power of $1/t$ is to standardize interpretation of design efficiency. For example, if the D- or I-efficiency of ξ is 0.5, the practical implication is that ξ needs to be replicated twice to perform as well as the optimal design. In practice, designs with high efficiencies are sought.

We finish this section introducing the well-known uniform designs, which we will use in Sections 3.1 and 4.1 as a benchmark. Uniform designs have points equally spread out in the design space with equal number of observations at each point. They are easy to implement and intuitively appealing which explains in part their popularity. However the number of points in a uniform design has to be carefully selected; otherwise, too many points can reduce their efficiencies (Wong & Lachenbruch, 1996). We denote a uniform design with k design points by U_k and will evaluate efficiencies of uniform designs under various FP models.

3 Optimal designs for FP models

We provide D- and I-optimal designs for FP models and present justifications for them. Algorithms can be directly used to find optimal designs but it is desirable to have analytical descriptions whenever possible because they facilitate studying properties of the optimal designs under model mis-specifications. For example, Chang and Lay (2002) found analytic D-, A- and E-optimal designs for the growth curve model with regression function $f(x) = (1,x,x^p)^T$ and p is known and larger than one and used them to study properties of the optimal designs.

In the next two subsections, we present D- and I-optimal designs for FP models. Section 3.3 discusses the common case when the choice of the model is uncertain and we determine designs for the FP models that are robust to mis-specification in the mean function.

Throughout, we assume we have a known design space $\Omega = [\epsilon, a]$, where for all practical purposes, we may assume that the left extreme point of the design space is 0, unless it causes numerical problems, as for example, when one or more of the powers in the FP are negative. For such situations, we replace the left extreme point by a very small positive constant ϵ, say, $\epsilon = 0.0001$.

3.1 *D*-optimal designs

FP models like other models are used to approximate the unknown true response function, which may be linear or non-linear model. This means that the parameters in the FP models may no longer have meaningful physical interpretations. However, using an adequately fitted FP model for prediction is both practical and useful because of the celebrated equivalence theorem of J. Kiefer and Wolfowitz (1960), which established that *D*- and *G*-optimal designs are equivalent when errors in the model are homoscedastic. This is a remarkable result because the two criteria are very different with the latter aiming to find a design that minimizes the maximal variance of the fitted response across the design space. It follows that *D*-optimal designs for FP models are also *G*-optimal for estimating the response surface over the design space. *D*-optimal designs for FP models up to degree 2 can be obtained using the theory of Tchebycheff systems and the above equivalence theorems; proofs are in the appendix.

Theorem 1 Let ξ_D denote the *D*-optimal design for a given FP model on the design space $\Omega = [\epsilon, a]$.

1. For FP1 models with values of $p \in \mathscr{P}$,

$$\xi_D = \begin{Bmatrix} \epsilon & a \\ 1/2 & 1/2 \end{Bmatrix},$$

that is, it is equally supported at the extreme points of Ω, and

2. For FP2 models with values of p and $q \in \mathscr{P}$, ξ_D is equally supported at three points: ϵ, s and a, where the interior point s is given by

$$s = \left(\frac{(a^q - \epsilon^q)p}{(a^p - \epsilon^p)q}\right)^{1/(q-p)} \quad (5)$$

and p and q are both unequal and nonzero. The remaining cases are:

$$s = \begin{cases} \left(\dfrac{a^q - \epsilon^q}{(\ln[a] - \ln[\epsilon]q)}\right)^{1/q}, & p = 0 \neq q \text{ (and analogously for } p \neq 0 = q), \\ \epsilon \exp\left(\dfrac{(\ln[a] - \ln[\epsilon])a^p}{a^p - \epsilon^p} - \dfrac{1}{p}\right), & p = q \neq 0, \\ \sqrt{\epsilon\ a}, & p = q = 0. \end{cases}$$

Theorem 1 applies more generally to other real numbers outside the set $\mathscr{P}$ (see the proof). When $p = 0 \neq q$, $p = q \neq 0$ or $p = q = 0$ the formula (5) for obtaning s leads to an indeterminate value. Then the formulae for s is obtained using the corresponding FP2 model. But they can be obtained by evaluating the limit of s as p or/and q tends to zero, or p tends to q. We observe that when $p + q = 0$, the interior point s becomes $\sqrt{\epsilon\ a}$, and the *D*-optimal design is the same for 4 different FP2 models, namely FP2(−2,2), FP2(−0.5,0.5), FP2(0,0) and FP2(−1,1).

It is helpful to list FP2s in a systematic way with each FP2 identified uniquely by a model number. We propose the following system. We first arrange the value of the interior support point s of each of the 3-point D-optimal design for all 36 models from the smallest to the largest. The FP2 model with the smallest interior point is assigned model number 1, and so on (see Table 1). When the models have the same interior support point, see for example models 13–16 in Table 1, the numbering for these models start from the smallest power to the largest power starting with the power p first and then q in an ordered manner. Unlike another numbering system for FP models (Duong & Volding, 2015), there are advantages in our numbering system when we study robustness properties of optimal designs for FP models in section 3.3.

A desirable property of D-optimal designs for polynomial models is that they are invariant under linear transformation on the design space. This property does not hold for FP models as can be seen from the formula for the interior point; if $c \neq 0$ and d are constants and x is replaced by $cx + d$, the interior point of the new design space does not move to an interior point given by $cs+d$, unless $d = 0$. This suggests studying properties of D-optimal designs for FP models is more challenging than for polynomial models, which may be another reason for the lack of design work for FP models.

Are uniform designs with 3 and 4 points efficient for estimating parameters in FP2 models? Figure 1 plots the D-efficiencies of the 3-point uniform design U_3 under different FP2(p,q) models. Its maximum D-efficiency of 1 is attained at $p = 1$ and $q = 2$ when the interior point of the D-optimal design $s = 0.50005$ coincides with the middle point of the interval. The minimum D-efficiency of this 3-point uniform design is near zero and is attained when $p = q = -2$. A general observation is that when q and p decrease, the interior support point of the D-optimal design tends to 0 and the optimal design is supported at nearly 2 points. Consequently, the design U_3 becomes very inefficient and is to be avoided. Conversely, when both p and q increase, the interior support points increase and the D-efficiency increases.

The uniform design U_3 cannot be used to diagnose model misfit in a FP2 model because it has only 3 points. U_4 is appealing because it has an additional support point and can be used to assess model adequacy. Figure 1 shows the D-efficiencies of the 4-point uniform design U_4 for various FP2 models. The maximum D-efficiency of U_4 is 0.905375, which is attained for the model FP2(0.5,3); otherwise the trends observed for U_3 apply for U_4 as well. We do not consider uniform designs with more points for FP2 models because they can be inefficient when they have many points, just like in polynomial models (Wong & Lachenbruch, 1996).

Theoretical results are not available for FP3 models but they can be directly generated using one of the standard algorithms discussed in the literature. The first row of figures in Figure 2 displays the D-optimal designs for two FP2 and two FP3 models defined on the interval $[\epsilon,a] = [0.0001,1]$, along with their sensitivity plots that confirm their optimality. The figures in the second row are examples and similar plots for I-optimal designs to be discussed next. The two plots in the right most column seem to show the two optimal designs are supported at 3 points but there are actually 4 support points for both the D- and I-optimal designs for the FP3 models. The vertical lines in both plots signify that the two smallest support points of each of the two optimal designs are very close together.

3.2 *I*-optimal designs

In contrast to *G*-optimality, which seeks to minimize the largest possible prediction variance across the design space, *I*-optimality assumes the researcher knows a priori which part or parts of the response surface are of interest to predict at the design stage. Naturally, a larger weight is assigned to parts that are deemed more important or interesting. In practice, numerical methods are used to find *I*-optimal designs unless the model is relatively simple as the below result shows. The proof is deferred to the appendix.

Theorem 2 Suppose the regression model is FP1 with values of $p \in \mathscr{P}$ and the weight measure μ is uniform on $S = \Omega = [\epsilon, a]$, the design space. The *I*-optimal design has the form

$$\xi_I = \begin{Bmatrix} \epsilon & a \\ w & 1-w \end{Bmatrix},$$

where

1. if $p \neq -1, -0.5, 0$:

$$1/w = 1 + \sqrt{\frac{(p+1)a^{2p+1} + a(2p+1)\left[(p+1)\epsilon^{2p} - 2(a\epsilon)^p\right] - 2p^2\epsilon^{2p+1}}{2p^2a^{2p+1} - (p+1)(2p+1)\epsilon a^{2p} + \epsilon\left[2(2p+1)(a\epsilon)^p - (p+1)\epsilon^{2p}\right]}};$$

2. otherwise,

$$1/w = \begin{cases} 1 + a^2\epsilon\sqrt{\dfrac{a^2 + 4a\epsilon\ln[\epsilon] - 2a\epsilon\ln[a\epsilon] - \epsilon^2}{a^3\epsilon^3\left(a^2 - 4a\epsilon\ln[a] + 2a\epsilon\ln[a\epsilon] - \epsilon^2\right)}}, & p = -1, \\ 1 + \sqrt{\dfrac{a\epsilon^{-1} - 4a(a\epsilon)^{-0.5} + \ln[a] - \ln[\epsilon] + 3.}{-\epsilon a^{-1} + 4\epsilon(a\epsilon)^{-0.5} + \ln[a] - \ln[\epsilon] - 3.}}, & p = -0.5, \\ 1 + \sqrt{\dfrac{2(a-\epsilon) + a(\ln[a] - \ln[\epsilon] - 2)(\ln[a] - \ln[\epsilon])}{2(a-\epsilon) + \epsilon(\ln[a] - \ln[\epsilon])(\ln[a] - \ln[\epsilon] + 2)}}, & p = 0. \end{cases}$$

3. When $p > 0$ and $\epsilon = 0$, $1/w = 1 + \sqrt{(1+p)/2p^2}$, independent of the value of the right end point a of the design space.

When $p = -1, -0.5$ or 0 the above 'general' formula for obtaning $1/w$, given in *1.*, leads to an indeterminate value. Then the appropriate value is obtained using the corresponding FP1 model. But in the cases $p = -1$ or -0.5 the weights can be obtained by evaluating the limit of $1/w$ when p tends to -1 or -0.5. However it is not the case for $p = 0$.

The optimal weights depend on the given weight measure and S, and both can be arbitrary. For example, suppose $S = [b, c]$ and the weight measure takes on the form $\mu = \alpha + \beta x$ or $\mu = \exp[-\text{abs}(x - \delta)]$, reflecting different prediction interests over S. A direct calculation shows that the *I*-optimal design is always supported at the extreme ends of the design space. The weights of the *I*-optimal design at the support points can be directly computed for selected values of b, c, α, β or δ. The general formulae for the weights at the support points for the I-optimal design with an arbitrary weight function μ can be derived analytically and the results are complicated and not necessarily useful. The applet we have created can assist in this task to find the optimal design analytically. For instance, from (2) one can find a more general formula for the weights of the I-optimal design. For FP1 models, the optimal designs are supported at the two ends of the design interval and knowing the matrix A is symmetric,

$$A = \begin{pmatrix} \mu_{11} & \mu_{12} \\ \mu_{12} & \mu_{22} \end{pmatrix},$$

the weight at the lower endpoint when $p \neq 0$ is

$$1/w = 1 + \sqrt{\frac{\epsilon^{2p}\mu_{11} - 2\epsilon^{p}\mu_{12} + \mu_{22}}{\alpha^{2p}\mu_{11} - 2a^{p}\mu_{12} + \mu_{22}}}.$$

As an application, suppose we have 3 possible weight measures over $S = [b,c]$: a uniform measure, say $\mu = \alpha = 1/(c-b)$, a linear measure, say $\mu = x-b$, implying greater interest near c, the right limit of S, and a linear measure, say $\mu = c - x$, implying less interest near c. Suppose the model is FP1(3), Ω = [0.0001,1] and the weight measure is uniform on S, then the mass at the smaller extreme end is $w = 0.6796$ when $S = \Omega$ and $w = 0.2281$ when $S =$ [1,1.2]. The corresponding weights for the increasing and decreasing linear measure over S described above are $w = 0.2518$ and $w = 0.1829$, respectively. Our results show that for all values of p, the linear increasing weight measure requires a larger value of w and its lowest value comes from the linear decreasing measure. We note that when we use the uniform measure over $\Omega = [\epsilon,a]$ for the model FP1(1), the I-optimal design is equally weighted at both ends of the design interval and so coincides with the D-optimal design.

Theoretical I-optimal designs for FP2 and FP3 models are not available but they are special cases of L-optimality discussed in Fedorov (1972). In particular, they can be generated directly using a modified version of Fedorov's algorithm and have their optimality verified using (4). Table 2 lists I-optimal designs for FP2 and FP3 models and different measures μ across Ω and beyond.

Tables 1 and 2 show that the interior support points of the I-optimal and D-optimal designs for FP2 models are generally different but close. There are two exceptions: (i) for model FP2(0,0), the interior support points of the D-optimal and I-optimal designs are 0.01 and 0.0334 respectively, and (ii) for model FP2(1,2), both optimal designs have the same interior support point 0.50005. In all cases, the I-optimal designs do not have equal weights, unlike D-optimal designs. We note that if $\epsilon = 0$, the above I-optimal design for the FP2(1,2) model is the D_s-optimal design for estimating the parameter α_2 of the FP2 model; see, for example, Atkinson, Donev, and Tobias (2007), p.160, Table 11.2. We also note that the interior support point of the I-optimal design increases when either value of p or q increases.

Table 2 shows selected I-optimal designs for the FP3 models on Ω = [0.0001,1]. For these designs, the middling support points of the equally weighted D-optimal designs are 0.0068 and 0.6667 for FP3(−2,2,3), 0.0004 and 0.3507 for FP3(−2,0,3), 0.0745 and 0.5861 for FP3(0.5,0.5,3) and 0.2765 and 0.7236 for FP3(1,2,3). The second row in Figure 2 shows examples and the sensitivity plots of the *I*-optimal designs for 2 FP2 and 2 FP3 models from Table 2 with weight measure $U[\epsilon,1]$.

3.3 Optimal designs when there is model uncertainty

Model assumptions can be wrong in various ways and to fix ideas, we focus on finding an efficient design when there is mis-specification in the mean function. We postulate that the true mean response function belongs to a class of user-specified functions. A traditional choice for this class is polynomial models up to a given degree, see for example, Stigler (1971), Studden (1982) or Song and Wong (1999). Here, the natural choice for this class is the class of FPs up to a user-specified degree, or a subclass thereof. The practitioner can either find a design that has high *D*- or *I*-efficiencies among all the models in the class or the subclass, that is, the most robust design to the mean specification.

If the postulated class of models is FP1 models, all *D*- and *I*-optimal designs are supported at the extreme endpoints. The *D*-optimal designs are equally supported and so designs equally supported at the extremes of the design space have a *D*-efficiency of 1 regardless which FP1 model is the true model. For *I*-optimality with the uniform weighting measure on $[\epsilon,1]$, the *I*-efficiency of the *I*-optimal design for a FP1 model varies depending which FP1 model is the true model. Our calculation shows that the *I*-optimal design for the FP1(0.5) model is the most robust across all FP1 models. Its minimum *I*-efficiency is 63.8% when the true model is FP1(−2). *I*-optimal designs with non-uniform weight measures can be found and similarly assessed.

For FP2 models, a similar investigation suggests that some *D*-optimal designs have high efficiencies among some FP2 models. For example, the 4 *D*-optimal designs for FP2 models with $p+q=0$, all have the same interior point, so their relative efficiencies are 1. When the class of interest includes all FP2 models, one can compare efficiencies of an optimal design under one model across all other models and pick the design that has high efficiencies across all models of interest. Figure 3 shows the *D*-efficiencies of the *D*-optimal design for each FP2 model relative to other *D*-optimal designs for FP2 models in two ways. The top sub-figure uses a contour plot and the bottom sub-figure displays these efficiencies using boxplots, with one per assumed model. In the top subfigure, the *D*-efficiencies of the *D*-optimal design for the assumed model on the horizontal axis under another model can be read off directly from the contour plot. Generally, we look for models with the most light patches above their model number (see Table 1) on the x-axis in the sub-figure. This suggests that the *D*-optimal designs for these models are robust to model mis-specification. For example, the sub-figure shows that the *D*-optimal design of the assumed model has high efficiencies when the true model has a number close to that of the assumed model. Conversely the top left and right bottom corners of the plot vividly shows FP2 models with very different model numbers are very incompatible in the sense that if one model is assumed and the true model has a very different model number, the *D*-optimal design for the assumed model is going to have very poor *D*-efficiency. Our proposed numbering of the FP2 models described in section 3.1 was motivated after several attempts to make such a contour plot easier to interpret for the reader.

The bottom sub-figure helps us to select a design robust to mean mis-specification. There are 36 boxplots, each showing the distribution of the D-efficiencies of the D-optimal design for an assumed FP2 model when one of the other models is the true model. If the user is uncertain which FP2 model is appropriate at the onset and is satisfied if the design has at least 50% D-efficiency across all FP2 models, the sub-figure suggests that the D-optimal design for the FP2 model numbered 21 is most appropriate because its boxplot has the largest median and is larger than 84%. If the first quartile was used as a robustness criteria instead of the median, the most robust design is the one numbered 29, which is also the design with the minimum interquartile range of D-–efficiencies across all models. In contrast, the D-optimal design for the FP2 model numbered 1 is to be avoided because its 25%, 50% and 75% D-efficiencies are consistently the smallest among all other D-optimal designs.

4 Optimal designs for FP models with applications to risk assessments

The design methodology described in the previous sections is general and can be applied to any regression problems where the response can be adequately modeled by a FP. This section describes three applications to detect various health risk assessments using a well designed study and the model is a fractional polynomial. In what is to follow, we use real studies that were conducted seemingly with minimal design considerations because no rationale was provided for the number of doses, the choices of the doses, and the number of replicates at each of the doses. For each of the three applications, we determine the D and I-optimal designs and show the benefits of using optimal designs.

4.1 Optimal designs for assessing risk of brain damage in songbirds

Serroyen, Molenberghs, Verhoye, Van Meir, and Van der Linden (2005) conducted an experiment to study the impact of testosterone on the dynamics of Mn^{2+} accumulation using data measured by magnetic resonance imaging in three songbird brain areas. We are interested in data from the high vocal center (HVC) area because they applied FP models to fit the data. Serroyen et al. (2005) determined the best fitting model was FP2(0,0.5) given by $Ey = \alpha_0 + \alpha_1 ln(t) + \alpha_2 t^{0.5}$, where t is the time. In this study, the range of the *time* variable is from 0.25 to 7 hours.

From Theorem 1, assuming independent observations, the equally weighted D-optimal design for FP2(0,0.5) model has support at the two extreme ends, 0.25 and 7, and an interior point at $s = 1.66$ from formula (5). The I-optimal design for the same model has support at 0.25, 1.90 and 7, with weights at the left and right extremes equal to 0.1389 and 0.3663, respectively, when the weight measure μ is the uniform distribution U[0.25,7]. Both optimal designs have different weights but are supported at about the same points. The D-efficiency of the I-optimal design is 87.5% and the I-efficiency of the D-optimal design is 83.1% suggesting these optimal designs are relative robust under a change of criterion variation. The implemented design is not well described in the study and our best guess is that it is a uniform design U_{28} with every 15-minute measurements between 0.25 and 7 hours. A direct calculation shows the D-efficiency of this design is 54.3% ant its I-efficiency is 75.2%. In either case, a substantial amount of cost and effort could have been saved if the I-optimal or D-optimal design was used.

If there is interest to predict the mean response with different interests in a region S outside the design space Ω, an I-optimal design is appropriate. Table 3 shows the I-optimal design for two different regions S and 3 different measures μ that reflects the varying interest in predicting the outcome over different regions in S. One measure has uniform interest across S, the second measure signifies increasingly more interest in region in S further away from Ω, and the third measure signifies increasingly more interest in region in S nearer to the design space Ω. The two regions for S are similar except one is a slightly longer interval. All the I-optimal designs have three design points with two at the end-points of the design interval, 0.25 and 7 and an interior point around 1.7 hours. The main differences among these optimal designs are in their weight distributions. The I-optimal designs over the extrapolated regions S at least 75% of observations to be taken at 7 hours whereas the corresponding proportion of observations required by the I-optimal design with the uniform weighting measure on the interval [0.25,7] is only about 36%.

The take home message from our proposed optimal designs is that the implemented design that take observations from the HVC area every 15 minutes can be time consuming and laborious without a substantial gain in accuracy in the statistical inference. Our results show that when we believe the postulated models based on these data are adequate, then future studies should collect the data from the HVC area at 3 time periods only. Such a design provides maximal statistical efficiency for the desired inference, saves time and labor in terms of not having to collect data from 28 time periods as in the original experiment.

4.2 Optimal Designs for risk assessment of rats exposed to various chemicals

Geys et al. (2000) and Faes et al. (2003) investigated the use of FP models for modelling a dose-response problems, as part of a Quantitative Risk Assesment (QRA). Specifically, the goal was to ascertain the effects in mice exposed to three chemicals: ethylene glycol (EG), di(2-ethyhexyl)-phthalate (DEHP) and diethylene glycol dimethyl ether (DYME). We discuss each of them sequentially and demonstrate there are benefits of using optimal designs versus the implemented design.

Using data from the EG study, the authors concluded that the FP2 model with the best fitting was the FP2(0.5, 1) model given by $Ey = \alpha_0 + \alpha_1 d^{0.5} + \alpha_2 d$, where d is the dose. From our earlier results, we observe that (i) the D-optimal design for the FP2 model on the dose space Ω = [0,3000] (mg/kg/day) is equally supported at the doses 0,750 and 3000, i.e. each of these dose has the same number of mice, and (ii) the I-optimal design with a uniform weight function for predicting across Ω is at doses: 0,853.65 and 3000 and the corresponding weights at these points are 13.2%,52.7% and 34.1%, respectively. In contrast, the original experiment was designed with approximately equal number of observations at the following four doses: 0,750,1500 and 3000, i.e. 25,24,22, and 23 pregnant dams at these doses. A direct calculation shows that this design has a D-efficiency of 91.8% and an I-efficiency of 87.4%. That means that practitioners have a very good initial design, but the D- and I-optimal designs could have helped the investigators obtained the same inference accuracy, saving an additional 8% or 13% of efficiency, depending on the criterion. We note that for this application, the *D*-efficiency of the *I*-optimal design is 86% and the *I*-efficiency of the *D*-optimal design is 80.6%, implying that the two criteria are not competitive and either one of the two optimal designs performs roughly the same under the two criteria. Even though the implemented design was a reasonably good design for inference purposes, the optimal designs are less costly to use because fewer observations are required for the same level of statistical efficiency. Thus the two optimal designs satisfy more adequately than the implemented design a federal guideline to always use the least number of animals for experimental studies.

For the other two studies there can be substantial gains applying optimal designs rather than the original designs. A noticeable percentage of experiments/mice, and related costs, can be saved.

For the DEHP study, the FP model selected was FP2(−1,−2). Specifically, the model is $Ey = \alpha_0 + \alpha_1(d + 1)^{-1} + \alpha_2(d + 1)^{-2}$, where a transformation on the dose is carried out to fit the data and to avoid numerical problems, since such a FP model needs strictly positive values of the explanatory variable due to its negative powers. The doses of the *D*-optimal design were at 0,0.0698 and 0.150 with the same number of mice at each dose. The *I*-optimal has doses at 0,0.0704 and 0.150 with weights, respectively, equal to 23.8%,49.9% and 26.3%. The implemented design in the study had doses at 0,0.025,0.050,0.100, and 0.150 (% DEHP) and the number of dams at these doses were 25,26,26,17, and 9. For this example the *D*(*I*)-efficiency of the implemented design is 70.7% (70.6%), suggesting that the implemented design does equally well for estimating model parameters and estimating the response surface. A direct calculation shows the *D*-efficiency of the *I*-optimal design is 94.4% and the *I*-efficiency of the *D*-optimal design is 88.9%.

In the study using the chemical DYME, the doses selected for the study were 0,62.5,125,250, and 500 mg/kg/day and the number of pregnant dams were 21,20,24,23, and 23, respectively. The best fitting FP model for the data was a FP1 given by $Ey = \alpha_0 + \alpha_1 d^{-1}$. There was no mention about working with transformed doses, but to avoid numerical problems we assume a small strictly positive minimum dose of 0.1. By Theorem 1, the *D*-optimal design for the FP1(−1) is equally supported at the control and maximum doses, i.e. 0.1 and 500. The *I*-optimal design has the same doses but requires very unequal number of mice at these doses: 1.4% at control dose and 98.6% at the other dose. For this example the implemented design has a *D*-efficiency of 78.3% and an *I*-efficiency of 83.2%.

The *I*-optimal designs found in this subsection assumes a uniform weight function for prediction in Ω. If there is different interest in predicting over different parts of the design space Ω, another suitable weight function will have to be used and the above efficiency results may be different. The same applies if we have an extrapolation problem where interest is in predicting outside the design space. For such situations, the Mathematica online tools we provide in this work can be readily applied to obtain the designs and produce the figures similar to what we show here.

4.3 Optimal Designs for detecting hormesis

Hormesis is a characteristic of many biological processes, and sometimes informally referred to as instances when a little poison can be healthy. Such observances occur frequently in the biological and health sciences, see for example, Calabrese (2004); Calabrese and Baldwin (2001).

We focus on an aquatic toxicological experiment that was conducted by the US Environmental Protection Agency (https://www3.epa.gov/) to ascertain possible existence of hormesis, when organisms were subject to low exposures of toxins. A species of water flea, *Ceriodaphnia dubia*, was exposed in a static renewal system, in brood cups, to different concentrations of effluent. Test results were based on their survival and reproduction abilities, modeled by an extended Gompertz model and a linear Logistic model. Casero-Alonso, Pepelyshev, and Wong (2018) considered design issues and provided different optimal designs for detecting different aspects of hormesis as efficiently as possible.

We have fitted a FP model to the available data and Figure 4 shows the best fitting FP2 model FP2(0.5, 2), along with the fitted Gompertz and Logistic models used in the paper. To obtain that FP2 model, we used the R-package *mfp* mentioned in the introduction. The R output gives an AIC 321.7, close to the AIC obtained for linear Logistic model (321.4) and slightly better than the AIC for the extended Gompertz model (323.3). For the fit of the FP2 model, the dose (d) is transformed to match those from the eGompertz and linear logistic models: $Ey = \alpha_0 + \alpha_1[(d + 1.6)/10]^{0.5} + \alpha_2[(d + 1.6)/10]^2$.

Theorem 1 provides the D-optimal design for the fitted model on the dose interval Ω = [0,12.5]: the doses are 0,4.96 and 12.5 with equal number of brood cups at the three doses. The I-optimal design on Ω with a uniform weight function was obtained numerically: the doses are 0,5.08 and 12.5 with 17.5%,54.3% and 28.2% of the experiment units respectively. The implemented design with 10 brood cups at doses 0,1.56,3.12,6.25 and 12.5, has a D-efficiency of 82.9% and an I-efficiency of 85.2%. This means that with that the implemented is about equally efficient for estimating model parameters in the FP model and for predicting when there is uniform interest across the design space. However, if an optimal design were used, substantial gain can still be realized in terms of cost, without loss of accuracy for inference.

A further and important advantage of implementing an optimal design for a fractional polynomial is that, if the fractional polynomial provides an adequate fit, its optimal designs (D- or I-) are independent on the nominal values of the model parameters. In contrast, optimal designs for nonlinear models depend on the nominal parameters and they can be sensitive to mis-specifications in the nominal values, as can be seen for the two nonlinear models used to fit the data displayed in Tables 1 and 2 in Casero-Alonso et al. (2018). We also note that the number of support points of the optimal designs for these models are different, 3 for the FP and 4 for the fitted nonlinear models. Depending on the problem at hand, cost can vary substantially when we take observations at a new dose; so design considerations can guide our choice for the implemented design.

5 Conclusions and Discussion

FP models are increasingly used in many areas of biomedical and environmental studies because they are more flexible than polynomials. Different types of optimal designs for polynomial models have been found analytically many decades ago and there is no corresponding work for FP models. The mathematical derivation of optimal designs for FP models seems challenging and they are hard to study them analytically. We constructed D- and I-optimal designs up to degree 3 and when analytical formulae of the optimal designs are not available, we provide a user-friendly applet for finding the optimal designs for FPs up to degree 3. Figure 5 displays a screen shot from the applet. The applet, OEDforFPmodels, is freely available at https://github.com/victormanuelcasero/OED_FPmodels and is created from Mathematica. To use the applet, the user first downloads a free 'Wolfram Player' from Wolfram.com. In the second step, the user selects one of the two design criteria and inputs the degree and powers of the FP model. In the third step, the applet is run to obtain the D- or I-optimal design, whereupon the sensitivity plot is also automatically shown. Our Mathematica codes can also be directly modified to find other types of optimal designs.

We also discussed robustness properties of optimal designs for FPs to mis-specifications in the model assumptions. Here we follow traditional design work in the literature and assume the model is known (apart from the parameters' values). This may not be a realistic assumption in practice, which is why we have included Section 3.3 that addresses model uncertainty issues. Using some illustrative examples, our recommendation there is to implement a design that is most robust to model assumptions, to the extent possible. In addition, we provided illustrative applications of how such optimal designs can be constructed and implemented for different type of studies. Our examples are in risk assessments but the design methodology can be applied generally to other problems. As mentioned in the introduction, multivariate FP models may be of great interest, even the case of correlated observations. That deserves further research, especially for additive and multiplicative models formed from marginal FP functions, or nonlinear models with random effects.

Since a particular FP model has to be chosen at the end, criteria for discriminating among all possible models should be considered in future work (López-Fidalgo, Tommasi, & Trandafir, 2007). Actually robust designs for both selecting the model and then estimating the parameters will be of great interest.

We close by offering a mathematical challenge for FP models. Royston and Altman (1994) recommended selected powers for FP models without theoretical justifications and it is not clear to date whether their collection of power indices has desirable mathematical properties. For example, given compact design space, can we show that their recommended restricted class of FPs can always approximate any continuous function 'adequately' in some sense? If not, can the set of the recommended powers be expanded to have desirable properties?

Acknowledgements

All authors were sponsored by Ministerio de Economía y Competitividad and fondos FEDER MTM2016–80539–C2–1–R and by Junta de Comunidades de Castilla–La Mancha SBPLY/17/180501/000380. Dr. Wong wishes to acknowledge the support from the University of Castilla–La Mancha and the program FEDER of Castilla–La Mancha 2014–2020 that made his visit possible. He also wishes to acknowledge the hospitality of the department during the visit. We also thank Diego Urruchi-Mohíno for his helpful assistance in the development of the Mathematica codes to generate optimal designs for FP models. Dr. Wong was also partially supported by a grant from the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R01GM107639. The contents in this paper are solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We also want to acknowledge the valuable comments from reviewers.

Appendix

This appendix provides proofs for Theorem 1: D-optimal designs for FP1 and FP2 models and Theorem 2: I-optimal designs for the FP1 models.

Proof (Theorem 1: construction of D-optimal designs)

Suppose that a design ξ for a FP1 model defined on $[\epsilon, a]$ with 3 points at $\epsilon \le s_1 < s_2 < s_3 \le a$. From the General Equivalence Theorem of J. Kiefer and Wolfowitz (1960), if ξ were D-optimal, its sensitivity function $c(x)$ must satisfy $c(x) = f^T(x)M^{-1}(\xi)f(x) - 2 \le 0$ for all $x \in [\epsilon, a]$ with equality at the support points. If $p \ne 0$, the component functions in $c(x)$ are $1, x^p, x^{2p}$. These 3 component functions form a Tchebycheff system on the interval $[\epsilon, a]$ because the determinant of the matrix

$$\begin{vmatrix} 1 & x_1^p & x_1^{2p} \\ 1 & x_2^p & x_2^{2p} \\ 1 & x_3^p & x_3^{2p} \end{vmatrix} = -\left(x_1^p - x_2^p\right)\left(x_1^p - x_3^p\right)\left(x_2^p - x_3^p\right),$$

has the same sign for any $\epsilon \le x_1 < x_2 < x_3 \le a$. It follows that for each $0 \ne p \in \mathscr{P}$, the sensitivity function has at most 2 zeros and since the D-optimal design for a FP1 model has at least 2 points for it to have a nonsingular information matrix, the D-optimal design is equally supported at 2 points (Fedorov, 1972; Pukelsheim, 1993). A similar argument shows that the same conclusion applies when $p = 0$ and the component functions are $\{1, \ln[x], \ln[x]^2\}$. Direct calculus shows that the sensitivity function of the design equally supported at the two end-points is $c(x) = 4(a^p - x^p)(\epsilon^p - x^p)/(a^p - \epsilon^p)^2$ when $p \ne 0$. Clearly, this function satisfies the conditions required in the equivalence theorem for D-optimality in (3) and so the design is D-optimal for FP1 models. A similar argument applies when $p = 0$.

Note that the previous reasoning about $p \ne 0$ is valid for any real number value of p. Thus the result proved here is more general than for $p \in \mathscr{P}$. The same applies for the following results.

For the FP2 models, we first establish that the D-optimal design ξ_D has three points. For such models, the sensitivity function $c(x)$ has at most 6 component functions, i.e. $1, x^p, x^{2p}, x^q, x^{2q}$ and x^{p+q}. In selected cases, such as when $p = -0.5$ and $q = -1$, the sensitivity function has only 5 component functions. A direct calculation shows that the associated Wronskians for this system of functions are positive for any values of p and q and so the component functions form a Tchebycheff system (Gasull, Lázaro, & Torregrosa, 2012). It follows that there are at most 5 zeroes (counting multiplicities). The interior support points have multiplicity two, because the maximum value of the sensitivity function of the D-optimal design has to be less than or equal to zero in the interval with the maximum value attained at the support points. This implies only three support points are possible, either two interior points and one extreme point of the design interval or one interior support point and the two extreme points of the interval $[\epsilon, a]$. Because the number of support points is the same as the number of parameters, the D-optimal design is equally weighted (Pukelsheim, 1993). We next argue that the optimal designs have to include the two extreme points.

Suppose the equally weighted design ξ is supported at $s_1 < s_2 < a$, where $\epsilon < s_1$. The determinant of the information matrix of this design for model FP2(p,q) with $0 \neq p \neq q \neq 0$ is

$$|M(\xi)| = \frac{1}{27}\begin{vmatrix} 1 & s_1^p & s_1^q \\ 1 & s_2^p & s_2^q \\ 1 & a^p & a^q \end{vmatrix} = \frac{1}{27}D^2 > 0$$

since D is always either positive or negative for any values of s_1, s_2 and a because we have a Chebyshev system. Further,

$$\frac{\partial D}{\partial s_1} = q s_1^{q-1}\left(a^p - s_2^p\right) - p s_1^{p-1}\left(a^q - s_2^q\right) \neq 0$$

implies that

$$\partial|M(\xi)|/\partial s_1 = (2/27) D\, \partial D/\partial s_1 \neq 0$$

and so D is a decreasing function of s_1. Consequently, ϵ is a support point of the D-optimal design. We note that the last equation holds if and only if $s_1^{q-p}\dfrac{q\left(a^p - s_2^p\right)}{p\left(a^q - s_2^q\right)} \neq 1$, i.e. $\left(\frac{s_1}{c}\right)^{q-p} \neq 1$. For the first equivalence we note that $p s_1^{p-1}\left(a^q - s_2^q\right) \neq 0$. The last equivalence obtains because by the mean value theorem, there exists a $c \in [s_2, a]$ such that $\left(a^p - s_2^p\right)/\left(a^q - s_2^q\right) = c^{p-q} p/q$. Consequently, since $s_1 < c$ and $q > p$ without loss of generality, $(s_1/c)^{q-p} < 1$ proving the result. It follows that $\partial|M(\xi)|/\partial s_1$ is negative, otherwise the optimal design is singular. It is easy to verify that the determinant of the information matrix is a decreasing function of s_1 and its maximum is obtained at $s_1 = \epsilon$, the left end-point of the design space. Similar reasoning leads to the optimal design as being supported at $\epsilon < s_2 < s_3$ and its determinant is maximized when $s_3 = a$. The upshot is that the D-optimal design is equally supported at the two end-points of the design space and at an interior point. The above arguments apply to other cases of p and q (p and q unequal and one of them zero, p

and q equal and $p = q = 0$). In either case, the interior support point s is the unique root of the derivative of the sensitivity function.

□

Proof (Theorem 2: construction for I-optimal designs)

The previous reasoning can be used to directly prove that *I*-optimal designs for FP1 models are unequally supported at the end-points of the design space. We note that the component functions in the sensitivity function of (4) are $1, x^p, x^{2p}$ when $p \neq 0$ and $\{1, \ln[x], \ln[x]^2\}$ when $p = 0$, and they form a Tchebycheff system on $[\epsilon, a]$. Similarly, the weights in section 3.2 are found by finding the roots of the sensitivity function of the *I*-optimal design evaluated at $x = \epsilon$ and $x = a$.

In addition, a direct calculation, from $1/w$ given in *1.*, shows the result *3.* for $p > 0$ and $\epsilon = 0$.

□

References

Atkinson A, Donev A, & Tobias R (2007). Optimum experimental designs, with SAS (Vol. 34). Oxford University Press. doi: 10.1080/10543406.2010.481801

Atzpodien J, Royston P, Stoerkel S, & Reitz M (2007). Fractional polynomials in a new metastatic renal carcinoma continuous prognostic index involving histology, laboratory, and clinical predictors. Cancer biotherapy & radiopharmaceuticals, 22(6), 812–818. doi: 10.1089/cbr.2007.375 [PubMed: 18158772]

Austin PC, Park-Wyllie LY, & Juurlink DN (2014). Using fractional polynomials to model the effect of cumulative duration of exposure on outcomes: applications to cohort and nested case-control designs. Pharmacoepidemiology and Drug Safety, 23(8), 819–829. doi: 10.1002/pds.3607 [PubMed: 24664670]

Calabrese EJ (2004). Hormesis: a revolution in toxicology, risk assessment and medicine. EMBO reports, 5(S1), S37–S40. doi: 10.1038/sj.embor.7400222 [PubMed: 15459733]

Calabrese EJ, & Baldwin LA (2001). The frequency of U-shaped dose responses in the toxicological literature. Toxicological Sciences, 62(2), 330–338. doi: 10.1093/toxsci/62.2.330 [PubMed: 11452146]

Casero-Alonso V, Pepelyshev A, & Wong WK (2018). A web-based tool for designing experimental studies to detect hormesis and estimate the threshold dose. Statistical Papers, 59(4), 1307–1324. doi: 10.1007/s00362-018-1038-5 [PubMed: 30930546]

Chang F-C, & Lay C-F (2002). Optimal designs for a growth curve model. J. Stat. Plan. Inference, 104(2), 427–438. doi: 10.1016/s0378-3758(01)00255-5

Duong H, & Volding D (2015). Modelling continuous risk variables: Introduction to fractional polynomial regression. Vietnam J. Sci, 2, 19–26. doi: 10.1080/17476933.2019.1631287

Faes C, Geys H, Aerts M, & Molenberghs G (2003). Use of fractional polynomials for dose-response modelling and quantitative risk assessment in developmental toxicity studies. Statistical Modelling, 3(2), 109–125. doi: 10.1191/1471082×03st051oa

Fedorov VV (1972). Theory of optimal experiments Elsevier.

Gasull A, Lázaro JT, & Torregrosa J (2012). On the chebyshev property for a new family of functions. J Math Anal Appl, 387(2), 631–644. doi: 10.1016/j.jmaa.2011.09.019

Geys H, Molenberghs G, Declerck L, & Ryan L (2000). Flexible quantitative risk assessment for developmental toxicity based on fractional polynomial predictors. Biometrical Journal: Journal of Mathematical Methods in Biosciences, 42(3), 279–302. doi: 10.1002/1521-4036(200007)42:3<279::aid-bimj279>3.0.co;2-f

Groten JP, Schoen ED, Van Bladeren PJ, Kuper CF, Van Zorge JA, & Feron VJ (1997). Subacute toxicity of a mixture of nine chemicals in rats: detecting interactive effects with a fractionated two-level factorial design. Fundamental and Applied Toxicology, 36(1), 15–29. doi: 10.1093/toxsci/36.1.15 [PubMed: 9073463]

Kiefer J, & Wolfowitz J (1960). The equivalence of two extremum problems. Canadian J Math, 12, 363–366. doi: 10.1007/978-1-4615-6660-1_5

Kiefer J, & Wolfowitz J (1964). Optimum extrapolation and interpolation designs, I. Annals of the Institute of Statistical Mathematics, 16(1), 79–108. doi: 10.1007/BF02868564

Kiefer JC (1985). Collected papers III: design of experiments Springer.

Knafl GJ (2015). Adaptive fractional polynomial modeling in SAS®

Krishnan E, Tugwell P, & Fries JF (2004). Percentile benchmarks in patients with rheumatoid arthritis: Health Assessment Questionnaire as a quality indicator (QI). Arthritis Res Ther, 6(6), 1–9. doi: 10.1186/ar1220

López-Fidalgo J, Tommasi C, & Trandafir PC (2007). An optimal experimental design criterion for discriminating between non-normal models. J. R. Stat. Soc. B, 69(2), 231–242. doi: 10.1111/j.1467-9868.2007.00586.x

Mayer B, Keller F, Syrovets T, & Wittau M (2016). Estimation of half-life periods in nonlinear data with fractional polynomials. Statistical Methods in Medical Research, 25(5), 1791–1803. doi: 10.1177/0962280213502403 [PubMed: 23999890]

Namata H, Aerts M, Faes C, & Teunis P (2008). Model averaging in microbial risk assessment using fractional polynomials. Risk Analysis: An International Journal, 28(4), 891–905. doi: 10.1111/j.1539-6924.2008.01063.x

Pázman A (1986). Foundations of optimum experimental design (Vol. 14). Springer.

Pukelsheim F (1993). Optimal design of experiments (Vol. 50). SIAM. doi: 10.1137/1.9780898719109

Royston P, & Altman DG (1994). Regression using fractional polynomials of continuous covariates: Parsimonious parametric modelling. Journal of the Royal Statistical Society. Series C (Applied Statistics), 43(3), 429–467. doi: 10.2307/2986270

Royston P, Ambler G, & Sauerbrei W (1999). The use of fractional polynomials to model continuous risk variables in epidemiology. Int J Epidemiol, 28(5), 964–974. doi: 10.1093/ije/28.5.964 [PubMed: 10597998]

Royston P, Reitz M, & Atzpodien J (2006). An approach to estimating prognosis using fractional polynomials in metastatic renal carcinoma. Br. J. Cancer, 94(12), 1785–1788. doi: 10.1038/sj.bjc.6603192 [PubMed: 16736003]

Royston P, & Sauerbrei W (2004). A new approach to modelling interactions between treatment and continuous covariates in clinical trials by using fractional polynomials. Statistics in Medicine, 23(16), 2509–2525. doi: 10.1002/sim.1815 [PubMed: 15287081]

Sauerbrei W, Royston P, Bojar H, Schmoor C, & Schumacher M (1999). Modelling the effects of standard prognostic factors in node-positive breast cancer. Br. J. Cancer, 79(11), 1752–1760. doi: 10.1038/sj.bjc.6690279 [PubMed: 10206288]

Serroyen J, Molenberghs G, Verhoye M, Van Meir V, & Van der Linden A (2005). Dynamic manganese-enhanced MRI signal intensity processing based on nonlinear mixed modeling to study changes in neuronal activity. J. Agric. Biol. Environ. Stat, 10(2), 170–183. doi:10.1198/108571105×46426

Shkedy Z, Aerts M, Molenberghs G, Beutels P, & Van Damme P (2006). Modelling age-dependent force of infection from prevalence data using fractional polynomials. Statistics in Medicine, 25(9), 1577–1591. doi: 10.1002/sim.2291 [PubMed: 16252265]

Silke B, Kellett J, Rooney T, Bennett K, & O'riordan D (2010). An improved medical admissions risk system using multivariable fractional polynomial logistic regression modelling. QJM: An International Journal of Medicine, 103(1), 23–32. doi: 10.1093/qjmed/hcp149 [PubMed: 19846579]

Song D, & Wong WK (1999). On the construction of G_{rm}-optimal designs. Statistica Sinica, 263–272.

Stigler SM (1971). Optimal experimental design for polynomial regression. J. Am. Stat. Assoc, 66(334), 311–318. doi: 10.1080/01621459.1971.10482260

Author Manuscript
Author Manuscript
Author Manuscript

Studden W (1982). Some robust-type *D*-optimal designs in polynomial regression. J. Am. Stat. Assoc, 77(380), 916–921. doi: 10.2307/2287327

Wolfe F (2000). A reappraisal of HAQ disability in rheumatoid arthritis. Arthritis & Rheumatism, 43(12), 2751–2761. doi: 10.1002/1529-0131(200012)43:12<2751::aid-anr15>3.0.co;2-6 [PubMed: 11145033]

Wong WK, & Lachenbruch PA (1996). Designing studies for dose response. Statistics in Medicine, 15(4), 343–359. doi: 10.1002/0470023678.ch3a [PubMed: 8668866]

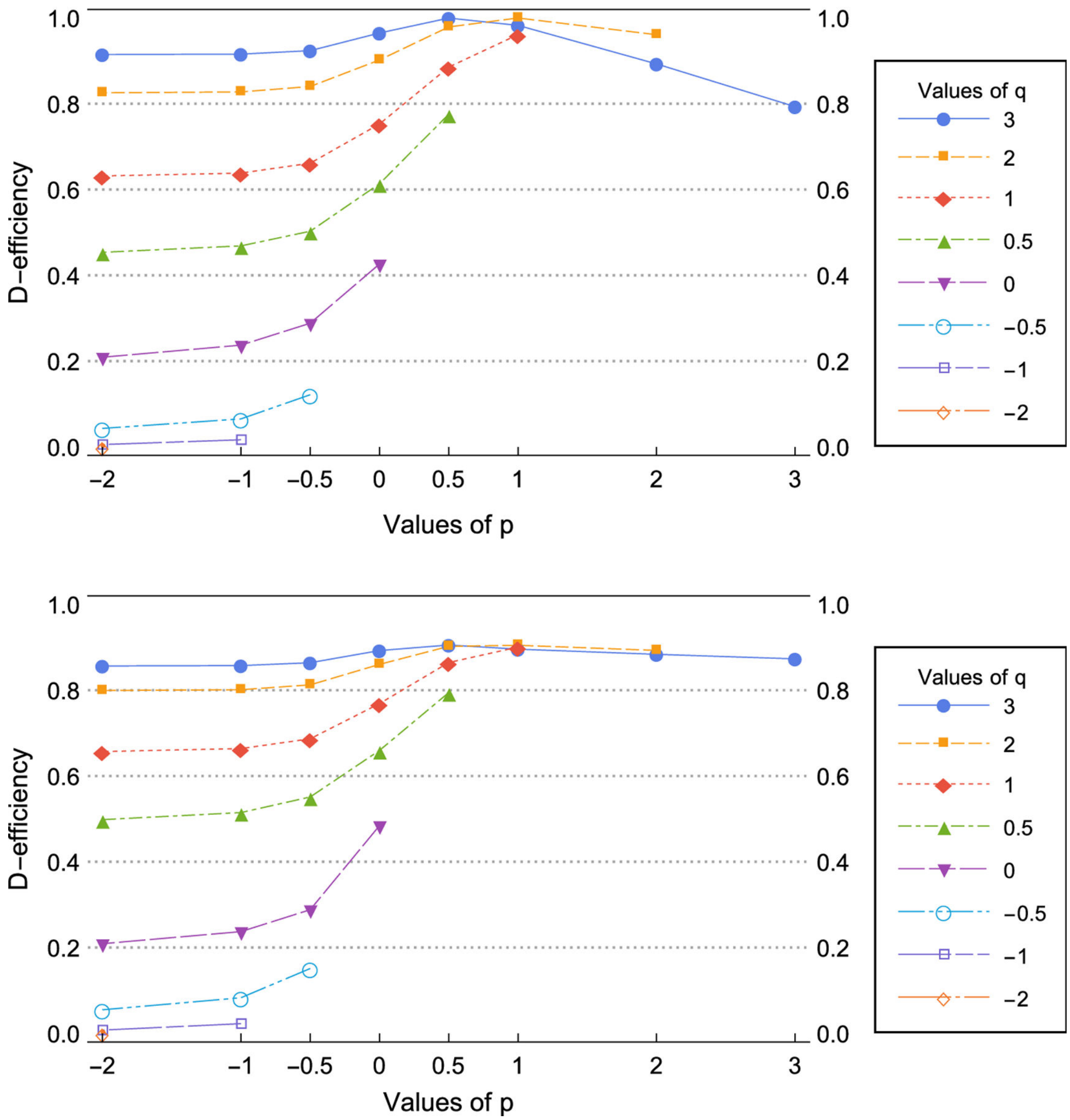

Fig. 1.

D-efficiencies of the uniform designs U_3 (top) and U_4 (bottom) for the FP2(p,q) models with $p \leq q$.

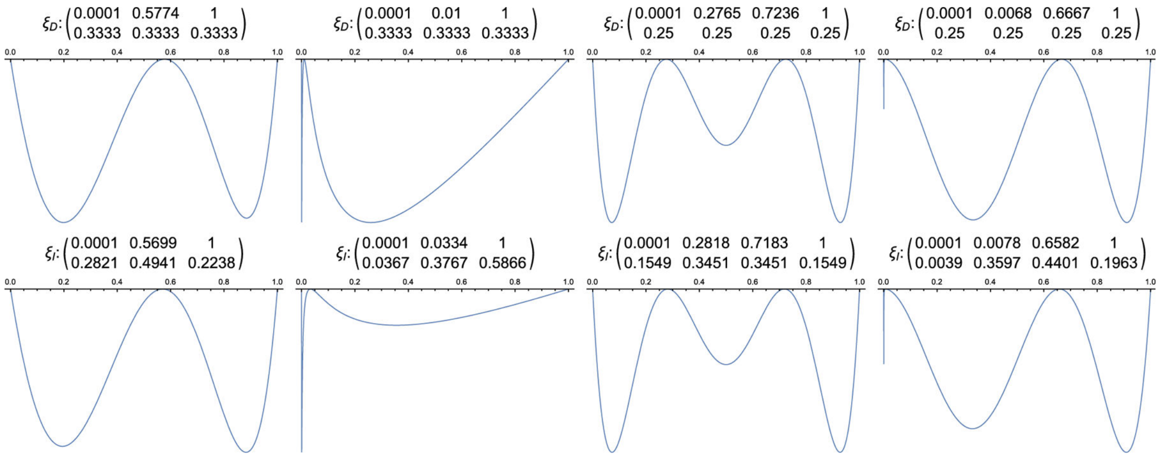

Fig. 2.

D- (first row) and *I*-optimal (second row) designs and their sensitivity plots on the design space $[\epsilon, a] = [0.0001, 1]$ for selected FP models starting from left: FP2(1,3), FP2(0,0), FP3(1,2,3) and FP3(−2,2,3). The *I*-optimal designs were computed using the uniform weight measure on $[\epsilon, 1]$.

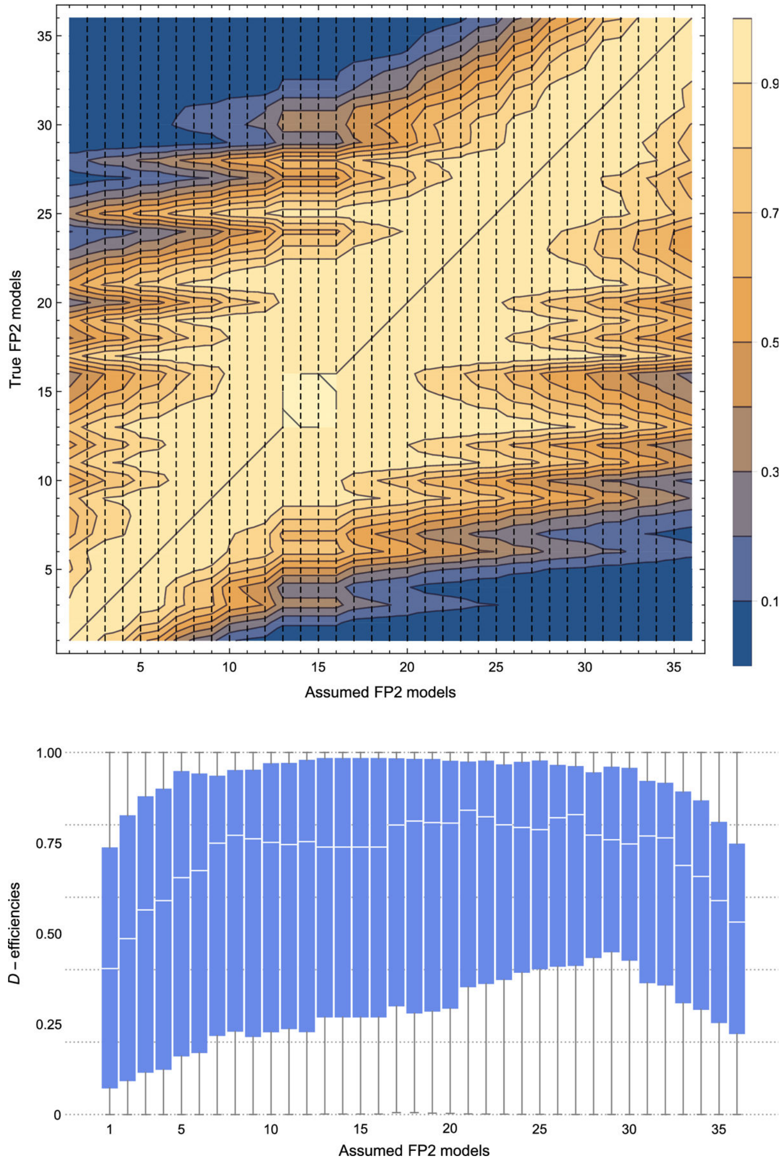

Fig. 3.

D-efficiencies of the *D*-optimal design for an assumed FP2 model when one of the 36 FP2 model is the true model in a contour plot (top) and in boxplots (bottom).

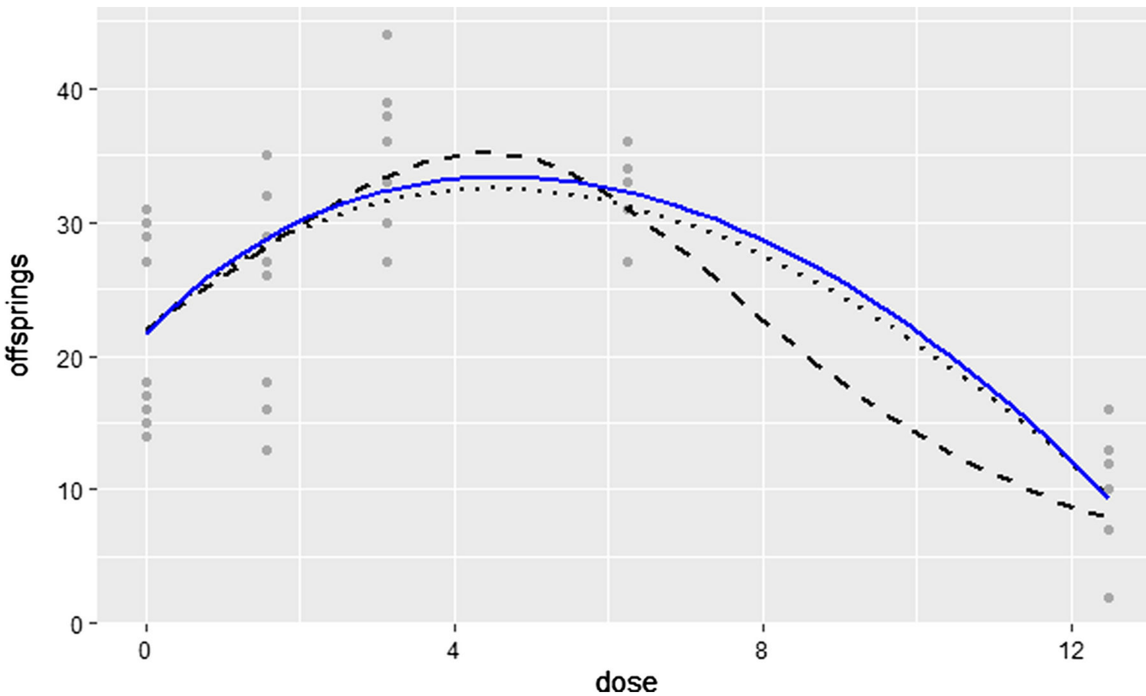

Fig. 4.
Plot of the response (number of offsprings Ceriodaphnia dubia) versus dose concentration from a whole effluent toxicity test [circles]. Models fitted: extended Gompertz (dotted line), Linear Logistic (dashed line), FP2(0.5,2) (blue solid line)

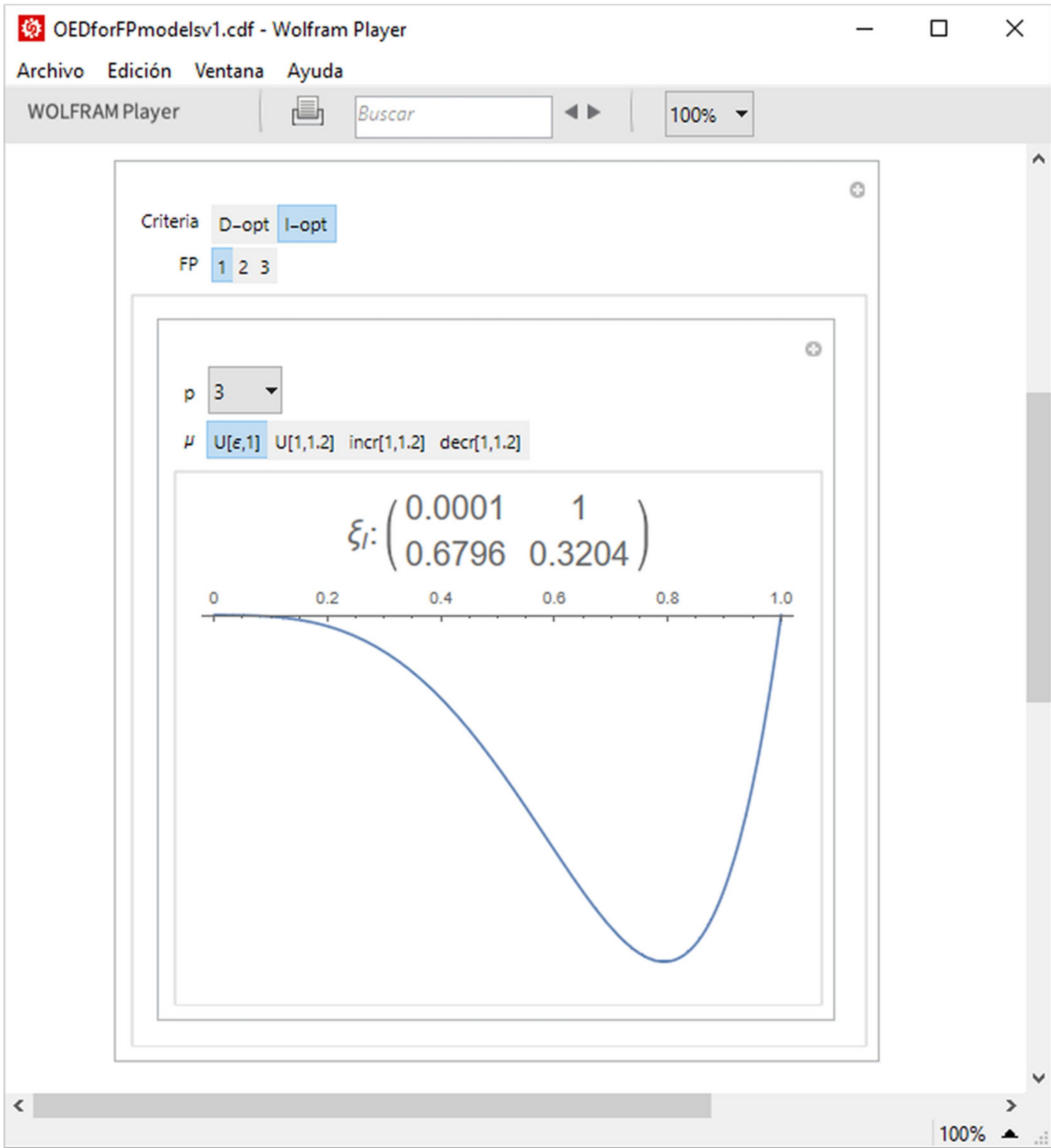

Fig. 5.

Applet showing *I*-optimal design with the uniform weight measure on the design space [0.0001,1] for model FP1(3).

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

FP2 models and their assigned unique model numbers based on the sorted values of their interior support point *s* of the *D*-optimal design for design space Ω = [0.0001,1]

Model number	FP2 (p,q)	mean response $Ey = \alpha_0 + \ldots$	interior point *s*
1	(−2,−2)	$\alpha_1 x^{-2} + \alpha_2 x^{-2} \ln[x]$	0.00016
2	(−2,−1)	$\alpha_1 x^{-2} + \alpha_2 x^{-1}$	0.00019
3	(−2,−0.5)	$\alpha_1 x^{-2} + \alpha_2 x^{-0.5}$	0.00025
4	(−1,−1)	$\alpha_1 x^{-1} + \alpha_2 x^{-1} \ln[x]$	0.00027
5	(−1,−0.5)	$\alpha_1 x^{-1} + \alpha_2 x^{-0.5}$	0.00039
6	(−2,0)	$\alpha_1 x^{-2} + \alpha_2 \ln[x]$	0.00042
7	(−0.5,−0.5)	$\alpha_1 x^{-0.5} + \alpha_2 x^{-0.5} \ln[x]$	0.00067
8	(−1,0)	$\alpha_1 x^{-1} + \alpha_2 \ln[x]$	0.00092
9	(−2,0.5)	$\alpha_1 x^{-2} + \alpha_2 x^{0.5}$	0.00109
10	(−0.5,0)	$\alpha_1 x^{-0.5} + \alpha_2 \ln[x]$	0.00216
11	(−2,1)	$\alpha_1 x^{-2} + \alpha_2 x$	0.00271
12	(−1,0.5)	$\alpha_1 x^{-1} + \alpha_2 x^{0.5}$	0.00339
13	(−2,2)	$\alpha_1 x^{-2} + \alpha_2 x^2$	0.01
14	(−1,1)	$\alpha_1 x^{-1} + \alpha_2 x$	0.01
15	(−0.5,0.5)	$\alpha_1 x^{-0.5} + \alpha_2 x^{0.5}$	0.01
16	(0,0)	$\alpha_1 \ln[x] + \alpha_2 \ln^2[x]$	0.01
17	(−2,3)	$\alpha_1 x^{-2} + \alpha_2 x^3$	0.02316
18	(−0.5,1)	$\alpha_1 x^{-0.5} + \alpha_2 x$	0.02943
19	(−1,2)	$\alpha_1 x^{-1} + \alpha_2 x^2$	0.03684
20	(0,0.5)	$\alpha_1 \ln[x] + \alpha_2 x^{0.5}$	0.04621
21	(−1,3)	$\alpha_1 x^{-1} + \alpha_2 x^3$	0.07598
22	(−0.5,2)	$\alpha_1 x^{-0.5} + \alpha_2 x^2$	0.09139
23	(0,1)	$\alpha_1 \ln[x] + \alpha_2 x$	0.10856
24	(0.5,0.5)	$\alpha_1 x^{0.5} + \alpha_2 x^{0.5} \ln[x]$	0.14853
25	(−0.5,3)	$\alpha_1 x^{-0.5} + \alpha_2 x^3$	0.16124
26	(0,2)	$\alpha_1 \ln[x] + \alpha_2 x^2$	0.23299
27	(0.5,1)	$\alpha_1 x^{0.5} + \alpha_2 x$	0.25502
28	(0,3)	$\alpha_1 \ln[x] + \alpha_2 x^3$	0.33077
29	(1,1)	$\alpha_1 x + \alpha_2 x \ln[x]$	0.36821
30	(0.5,2)	$\alpha_1 x^{0.5} + \alpha_2 x^2$	0.39951
31	(0.5,3)	$\alpha_1 x^{0.5} + \alpha_2 x^3$	0.49032
32	(1,2)	$\alpha_1 x + \alpha_2 x^2$	0.50005
33	(1,3)	$\alpha_1 x + \alpha_2 x^3$	0.57737

Model number	FP2 (p,q)	mean response $Ey = \alpha_0 + \ldots$	interior point s
34	(2,2)	$\alpha_1 x^2 + \alpha_2 x^2 \ln[x]$	0.60653
35	(2,3)	$\alpha_1 x^2 + \alpha_2 x^3$	0.66666
36	(3,3)	$\alpha_1 x^3 + \alpha_2 x^3 \ln[x]$	0.71653

Table 2

I-optimal designs for selected FP2 and FP3 models for three weighting measures on $\Omega = [0.0001,1]$ or $S = [1,1.2]$ with weight w_i at x_i. For all the cases, the extreme right endpoint of the design space is a support point and not shown.

		Weighting measures								
		***U*[Ω]**			***x* − 1(*x* ∈ *S*)**			**1:2 − *x*(*x* ∈ *S*)**		
FP2(−2,3)	x_i	0.0001	0.0252		0.0001	0.0232		0.0001	0.0235	
	w_i	0.0049	0.6763		0.	0.2518		0.	0.1830	
FP2(0,0)	x_i	0.0001	0.0334		0.0001	0.0103		0.0001	0.0111	
	w_i	0.0367	0.3767		0.0130	0.0520		0.0077	0.0319	
FP2(0,3)	x_i	0.0001	0.3441		0.0001	0.3322		0.0001	0.3368	
	w_i	0.0810	0.6319		0.0255	0.2695		0.0183	0.1996	
FP2(0.5,3)	x_i	0.0001	0.4866		0.0001	0.4919		0.0001	0.4973	
	w_i	0.2014	0.5500		0.0628	0.3010		0.0461	0.2308	
FP2(1,2)	x_i	0.0001	0.5001		0.0001	0.5020		0.0001	0.5083	
	w_i	0.25	0.5		0.0810	0.2863		0.0583	0.2166	
FP2(1,3)	x_i	0.0001	0.5699		0.0001	0.5788		0.0001	0.5839	
	w_i	0.2821	0.4941		0.0877	0.3273		0.0659	0.2584	
FP3(−2,2,3)	x_i	0.0001	0.0078	0.6582	0.0001	0.0062	0.6678	0.0001	0.0062	0.6719
	w_i	0.0039	0.3597	0.4401	0.00003	0.1153	0.3635	0.00002	0.0897	0.2992
FP3(−2,0,3)	x_i	0.0001	0.0005	0.3648	0.0001	0.0004	0.3583	0.0001	0.0004	0.3631
	w_i	0.0054	0.1046	0.6118	0.0019	0.0300	0.2711	0.0014	0.0216	0.2016
FP3(0.5,0.5,3)	x_i	0.0001	0.0962	0.5884	0.0001	0.0645	0.6201	0.0001	0.0651	0.6248
	w_i	0.0615	0.3121	0.4286	0.0343	0.1022	0.3048	0.0259	0.0777	0.2456
FP3(1,2,3)	x_i	0.0001	0.2818	0.7183	0.0001	0.2505	0.7507	0.0001	0.2517	0.7536
	w_i	0.1549	0.3451	0.3451	0.0598	0.1526	0.3384	0.0479	0.1237	0.2945

Table 3

I-optimal designs for the FP model in the testosterone example for extrapolation over two regions *S* outside the design space [0.25,7] and three weighting measures μ. Each of the *I*-optimal designs has weight w_i at x_i, i = 1,2,3.

μ	*S*		***I*-optimal designs**		
Uniform	[7,8]	x_i	0.25	1.6974	7
		w_i	0.0254	0.104	0.8706
	[7,9]	x_i	0.25	1.6914	7
		w_i	0.0437	0.1729	0.7834
Increasing	[7,8]	x_i	0.25	1.6741	7
		w_i	0.0301	0.1211	0.8488
	[7,8]	x_i	0.25	1.6713	7
		w_i	0.0502	0.1951	0.7547
Decreasing	[7,8]	x_i	0.25	1.7156	7
		w_i	0.0187	0.0784	0.9029
	[7,8]	x_i	0.25	1.709	7
		w_i	0.0336	0.1363	0.8301