

THESIS FOR THE MASTER OF DEGREE IN MATHEMATICAL STATISTICS

Optimal design of dilution experiments under volume constraints

MARYAM ZOLGHADR

CHALMERS



GÖTEBORGS UNIVERSITET

Department of Mathematical Sciences
Division of Mathematical Statistics
Chalmers University of Technology and University of Gothenburg
Gothenburg, Sweden 2013

Optimal design of dilution experiments under volume constraints
Maryam Zolghadr

Copyright © 2013 by Maryam Zolghadr, All rights reserved by the Author and Sergei Zuyev¹.

Department of Mathematical Sciences
Division of Mathematical Statistics
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg, Sweden
Phone: +46 (0)31-772 10 00

Typeset with L^AT_EX.
Department of Mathematical Sciences
Printed in Gothenburg, Sweden 2013

¹Professor Chair in Mathematical Statistics, Department of Mathematical Sciences, Chalmers University of Technology, Gothenburg, Sweden.

Optimal design of dilution experiments under volume constraints

Maryam Zolghadr

Abstract

We develop methods to construct a one-stage design of dilution experiment under the total available volume constraint typical for bio-medical applications. We consider different optimality criteria based on the Fisher information in both non-Bayesian and Bayesian settings. It turns out that the optimal design is one atomic, meaning that all the dilutions should be of the same size. Our proposed approach to solve such optimization problems is a variational analysis of functionals of a measure. The advantage of the measure optimization approach is that additional requirements like a total cost of experiment can be easily incorporated into the goal function.

Keywords: Dilution experiment, optimal design of experiments, Fisher information criteria, gradient methods, design measure, variational analysis on measures.

Acknowledgment

I would like to express my sincere gratitude to my supervisor Prof. Sergei Zuyev for the continuous support of my master study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

Besides my supervisor, I would like to thank my examiner Prof. Rebecka Jornsten, for her insightful comments.

My sincere thanks also goes to my families and friends in Sweden.

Last but not least, I would like to thank my parents for supporting me throughout my life.

Maryam Zolghadr
Gothenburg, June 5, 2013

Contents

Contents	2
1 Introduction	3
2 Dilution experiment, statistical model and optimality criteria	5
3 Optimisation of functionals of measures	10
4 Construction of optimal design	14
4.1 Optimal design for a fixed average number of HSCs	14
4.2 Optimal design with prior distribution on λ	15
4.3 Optimal design with additional costs	20
5 Discussion	23
References	32

1 Introduction

We study a new class of statistical experiments arising, in particular, in stem cells research, a very active area of experimental biology. Stem cells are the cells produced during early stages of embryonic development and they are having capacity to turn into various types of tissue cells. This potentially opens new ways to cure many diseases and explains the huge importance of the stem cells research, see, e. g., Mayhall et al. (2004) and the references therein. We aim to characterise an optimal design of dilution-type experiments under the constraint of the available solution which is typical for experiments involving counting stem cells. Specifically, this study originates in studies of *Hematopoietic* or *blood stem cells* (*HSCs*) which are the stem cells giving rise to all red and white blood cells and platelets. HSCs develop in a mammal's embryo in early days from cells. There is not much known on the biological mechanism which triggers a cell to develop into an HSC and there is no direct way so far to observe such cells, called *pre-HSCs*, which soon to become HSCs. Pre-HSCs are mostly produced in aorta-gonad-mesonephros (AGM) region of the embryonic mesoderm and also in the yolk sac, then colonise the liver. A challenging problem is how to detect the pre-HSCs and how many of them are present at a different embryo ages.

In order to estimate the number of pre-HSCs present in a given region, experiments on laboratory mice have been conducted which use the following signature property of stem cells. A mature HSC is capable to cure a mouse which receives a controlled dose of radiation if injected in its blood. A mouse recovers (*repopulates*), if and only if, it has received at least one HSC in the injected dose². Thus the number of pre-HSCs can be estimated by the so-called *limiting dilution method*: controlled doses of a substrate containing samples from the AGM are injected into irradiated mice and then the number of repopulated mice infers on the number of HSCs which developed from the initially present pre-HSCs.

The dilution method has been in the arsenal of biologists for almost a century, at least since McCrady used it for quantitative determination of *Bacillus coli* in water in 1915, McCrady (1915). Since then many studies used it to estimate the number of objects of interest in a medium without their direct count, Fisher (1922), Cochran (1950), Ridout (1995), to name a few.

So far, several studies have produced estimates for the number of pre-HSCs in AGM by using the dilution method which varies between just a few to, perhaps, as many as 200, see, e. g., Kumaravelu et al. (2002), Gekas et al. (2005), Ottersbach and Dzierzak (2005) Bonnefoix

²This assumption is rather questionable, so the biologists cautiously speak of one *repopulation unit* for this unknown minimal number of HSCs sufficient to cure an irradiated mouse. In this study we basically loosely speak of *one HSC* as of *one repopulation unit*.

and Callanan (2010) and Medvinsky et al. (2011). Most of the reports tend to focus on the modelling of experimental data and on the estimation methods. However, there has been little discussions on how to design the experiment to capture the most informative sample. Indeed, the experiment would be spoiled, if all the mice repopulate or if all do not. The aim of this paper is to find an optimal design of the dilution experiment to estimate the mean number of pre-HSCs. To this end, we are going to apply recently developed methods of *constrained optimisation* of functionals of measures and the corresponding steepest descent type algorithms for computations. It should be stressed that our methodology is generic in the sense that it can be applied to other dilution experiments and not only in the stem cells research. Moreover, additional constraints can be incorporated into the model which would allow, for instance, to take into account the cost of mice or other materials used in the experiment.

This study is organised in the following way. Section 2 introduces the dilution experiment we are dealing with, description of the corresponding statistical model followed by its assumptions, and at the end the optimality criterion functions. Section 3 lays out the theoretical basis for the optimisation methods we are using given that the goal functions are represented as functions of a *measure*. Consecutive sections provide the optimal design of the dilution experiments under various conditions and various goal functions: in non-Bayesian setting in Section 4.1, then under Uniform and Gamma Bayesian priors in Section 4.2, and finally in Section 4.3 we represent an optimal design that integrates cost to the criterion functions. We conclude by discussion of our findings and their extensions in Section 5.

2 Dilution experiment, statistical model and optimality criteria

Description of experiment The dilution experiment on estimation of the number of HSCs we address here involved the AGM region of an 11 days old mouse embryo. More exactly, in order to make a study more representable and not depending on features of a particular embryo, an engineered AGM is made from several such embryos. A substrate of volume V is then made from this engineered AGM and the content is thoroughly mixed. Next, n doses containing proportions x_1, \dots, x_n of the whole V are extracted from all or part of the substrate and left for a few days so that pre-HSCs in these doses, if any, mature into HSCs. Finally, these n doses³ are injected into n irradiated mice (the number n of mice was 30 in this experiment) so that each mouse receives its own dose and the mice are put to rest for a few weeks for the doses to take effect. After this, the number of repopulated mice, which are the ones having received a dose with at least one pre-HSC, is counted and the inference is drawn on the total number of pre-HSCs initially present in the AGM. The main question to address when designing such an experiment is doses of which volumes should be used for available number of mice in order to get best possible quality of the statistical estimator? Different criteria could be considered to quantify the quality of the estimator. We will consider below most common ones based on the Fisher information.

Note that an experiment in which all the mice repopulate or not repopulate is not informative and cannot be used to estimate the parameter of interest.

Statistical model Given the description of the experiment above, the following assumptions can be made.

- *Spatial homogeneity*: pre-HSCs were distributed uniformly throughout the substrate when the doses were taken. Thus there is no tendency for pairs or groups of pre-HSCs either to cluster or to reject one another. This is implied by the fact that the substrate was thoroughly mixed just before the doses are taken.
- *Orderliness*: the probability that there are more than one pre-HSCs in a small volume dv of substrate has order $o(dv)$. This is a natural assumption given that an 11 days old AGM contains about 300 thousand cells and only no more than 200 of these are pre-HSCs.
- *Independence*: each cell in substrate has the same (small) probability to turn into a pre-HSC independently of the other cells.

³These are further diluted to a standard volume, but this, obviously, does not change the number of HSCs present before the dilution.

- Only a pre-HSC can develop into a mature HSC. So that a dose contains at least one HSC at the time of injection to a mouse if and only if there was a pre-HSC present in the dose at the time of its extraction from the whole substrate.
- Finally, each dose when injected into an irradiated mouse is certain to exhibit a positive result (repopulated mouse), whenever the dose contains at least one HSC.

The first three assumptions above suggest that the locations of pre-HSCs in the substrate V are given by a homogeneous Poisson point process. This follows from the Poisson limit theorem for thinned point processes, see, e. g., Daley and Vere-Jones (2008, Sec. 11.3). Indeed, in every subset of volume x of the substrate, the number of cells turned into pre-HSCs is close to Poisson distribution with the parameter proportional to the mean number of pre-HSCs in the substrate, so with λx for some λ . And because of independence assumption, these numbers are independent for disjoint subsets. Changing the units if necessary, we assume from now on that the volume of the substrate is 1. The parameter λ is then the unknown density of the Poisson point process which is also the mean number of pre-HSCs in the substrate. Thus we operate with a measurable space carrying point configurations inside a set $V \subset \mathbb{R}^3$ of volume 1 (the space of finite counting measures ω on V with the minimal σ -algebra making the mappings $\omega \mapsto \omega(B)$ measurable for all Borel $B \subseteq V$) supplied with probability distribution \mathbf{P}_λ so that ω under \mathbf{P}_λ is a homogeneous Poisson point process with density λ .

The doses taken can now be associated with disjoint subsets V_1, \dots, V_n of V with volumes x_i , $i = 1, \dots, n$. The corresponding numbers of pre-HSCs $\omega(V_i)$, $i = 1, \dots, n$ in the doses are then independent Poisson distributed random variables with parameters λx_i while the total number of pre-HSCs $\omega(V)$ is Poisson distributed with parameter λ .

In the simplest case all the doses have the same volume $x \leq 1/n$. The probability that a doze is sterile, i.e. it does not contain a pre-HSC, is then

$$p = \mathbf{P}(\omega(V_i) = 0) = e^{-\lambda x}. \quad (1)$$

Thus the total number of non-repopulated mice follows Binomial distribution with parameters n, p and the maximum likelihood estimate for the average number of HSC λ is given by

$$\hat{\lambda} = -\frac{\log \hat{p}}{x}, \quad (2)$$

where \hat{p} is the proportion of non-repopulated mice provided it is not 0. However, the doses need not be necessarily all equal for an optimal design.

Let χ_i ($i = 1, \dots, n$) be an indicator that a mouse, which received the i th dose of volume x_i , has not repopulated. Thus, χ_i is a *Bernoulli*

random variable

$$\chi_i|\lambda \sim \text{Bern}(e^{-\lambda x_i}) \quad (3)$$

with the parameter equal to the probability of the i th dose to be sterile.

Hence, the *log-likelihood function* for the sequence $\boldsymbol{\chi} = (\chi_1, \dots, \chi_n)$ of non-repopulated and repopulated mice is given by

$$\ell(\lambda|\boldsymbol{\chi}, \mathbf{x}) = - \sum_{i=1}^n \chi_i \lambda x_i + \sum_{i=1}^n (1 - \chi_i) \log(1 - e^{-\lambda x_i}), \quad (4)$$

where $\mathbf{x} = (x_1, \dots, x_n)$. Maximisation of this expression over λ for an observed sample $\boldsymbol{\chi}$ provides a maximum likelihood estimator (MLE) of λ for a given design \mathbf{x} . Our goal here is to determine the optimal design in terms of the doses $\{x_i\}$, according to a suitably chosen *optimality criterion*, which we describe next.

Optimality criterion functions Recall that for a statistical model which depends on a one-dimensional parameter λ , the *Fisher information* is defined as

$$I(\mathbf{x}; \lambda) = - \mathbf{E}_\lambda \left[\frac{\partial^2 \ell(\lambda|\boldsymbol{\chi}, \mathbf{x})}{\partial \lambda^2} \right]. \quad (5)$$

The expectation \mathbf{E}_λ is taken here w.r.t the random vector $\boldsymbol{\chi}$.

Derived from (4), we have in our case

$$I(\mathbf{x}; \lambda) = \sum_{i=1}^n \frac{e^{-\lambda x_i}}{1 - e^{-\lambda x_i}} x_i^2. \quad (6)$$

The Fisher information measures the amount of information that an observable sample carries about the unknown parameter, which the likelihood function depends upon. On the other hand, it is the inverse of MLE's variance, see, e. g., Everitt and Skrondal (2010). Thus, maximising the information corresponds to minimising the variance of the MLE. Therefore, maximising the Fisher information (6) over \mathbf{x} under constraint $\sum_{i=1}^n x_i \leq 1$ is a useful design criterion.

It is typical in statistical experiment planning to describe design in term of a probabilistic *design measure*. Typically, the design measure is atomic, so it has a form $\sum_j q_j \delta_{x_j}$, where δ_x is the unit mass measure concentrated on a point $\{x\}$. The design measure reflects the (asymptotic when $n \rightarrow \infty$) frequency q_j of occurrence of the value x_j in the design, see, e. g., Atkinson and Donev (1992). By this reason, we will also describe the doses by a measure $\mu(dx)$ living on $[0, 1]$, albeit not renormalised to have mass 1. Namely, $\mu = \sum_j m_j \delta_{x_j}$ means the design when a dose of volume x_j is repeated m_j times. Obviously, we have that the total mass constraint $\sum_{j=1}^k m_j = \int \mu(dx) = n$ and that we cannot extract more doses than the total volume of the substrate: $\sum_{j=1}^k m_j x_j = \int x \mu(dx) \leq$

1. All integrals here and below are taken over $[0, 1]$, unless specified differently.

Now the basic optimisation problem for the design of our experiment is

$$G_1(\mu; \lambda) = I(\mu; \lambda) = \int \frac{e^{-\lambda x}}{1 - e^{-\lambda x}} x^2 \mu(dx) \rightarrow \sup \quad (7)$$

over measures μ with $\text{supp } \mu \subseteq [0, 1]$ satisfying

$$\mu([0, 1]) = n; \quad (8)$$

$$\int x \mu(dx) \leq 1. \quad (9)$$

A design measure describes *asymptotic* frequencies, so $q_j n = m_j$ for a given finite n may not be all integers. In this case it is reasonable to consider the nearest measure with all $m_j \in \mathbb{Z}_+$ as an approximation to the optimal solution. Or, if necessary, a choice of the optimal measure among such measures can be done by evaluation of the goal function at just a few closest approximations of this kind to the optimal design measure.

Bayesian setting Sometimes, there is an additional information available on the plausible values of the parameter λ which is given in a form of a *prior distribution* $Q(d\lambda)$, see, e. g., Marin and Robert (2007, Sec. 2.2.2). In this case the optimality criterion involves taking the expectation of the goal functions w.r.t the distribution Q .

For a single parameter, the following criterion functions are typically used to find an optimal design, see, e.g., Atkinson and Donev (1992) or Ridout (1995):

$$G_2(\mu) = \mathbf{E}_Q I(\mu; \lambda) \quad (10)$$

$$G_3(\mu) = \mathbf{E}_Q \log I(\mu; \lambda) \quad (11)$$

$$G_4(\mu) = -\mathbf{E}_Q (I(\mu; \lambda))^{-1} \quad (12)$$

Criterion function G_2 and G_3 maximize, under the same constraints (8) and (9), the expectation of the Fisher information and of its logarithm, respectively, as used in e.g., in Zacks (1977) and Chaloner and Larntz (1989). The criterion function G_4 minimises the expected asymptotic variance of the maximum likelihood estimator.

Next section will describe a general framework of optimisation of functionals of measures and the recently developed techniques for solving such optimisation problems. Apart from optimal design of statistical experiments Pukelsheim (1983), Atkinson and Donev (1992), these are frequent in different subjects, like spline approximation of curves and geometrical

bodies where the measure describes the positions of spline points Schneider (1988), maximisation the area covered by random geometric objects with the distribution determined by a measure Hall (1988), stochastic search, where a measure determines the search strategy Wynn and Zhigljavsky (1994), Zhigljavsky (1991).

3 Optimisation of functionals of measures

In this section we summarise necessary information about measures and variational analysis on them. Further details on measure theory can be found, *e. g.* in Dunford and Schwartz (2009) or Hille and Philip (1957).

Let X be a locally compact separable topological space with the Borel σ -algebra \mathcal{B} of its subsets. Let \mathbb{M} (\mathbb{M}_+) denote the set of signed (respectively, non-negative) finite measures on \mathcal{B} , *i. e.* countably additive functions from \mathcal{B} to \mathbb{R} (\mathbb{R}_+ , respectively). \mathbb{M} becomes a linear space if the sum of measures and multiplication by a number is defined by $(\eta + \nu)(B) = \eta(B) + \nu(B)$ and $(t\eta)(B) = t\eta(B)$ for any $t \in \mathbb{R}$ and $\eta, \nu \in \mathbb{M}$. \mathbb{M}_+ is a *cone* in \mathbb{M} since $\mu + \nu \in \mathbb{M}_+$ and $t\mu \in \mathbb{M}_+$ whenever $\mu, \nu \in \mathbb{M}_+$ and $t \geq 0$. The *support* $\text{supp } \mu$ of a positive measure m , is defined as the complement to the union of all open sets of zero μ -measure. Measures are *orthogonal* if their supports are disjoint. Any signed measure η can be represented as the difference $\eta^+ - \eta^-$ of two orthogonal non-negative measures $\eta^+, \eta^- \in \mathbb{M}_+$ (the Jordan decomposition). The set \mathbb{M} becomes a *Banach space* if supplied with the *total variation norm* $\|\eta\| = \eta^+(X) + \eta^-(X)$.

Optimisation of functions defined on a Banach space, which are commonly called *functionals*, relies on the notions of differentiability. In our case a functional $f : \mathbb{M} \mapsto \mathbb{R}$ is called *strongly* or *Fréchet differentiable* at $\nu \in \mathbb{M}$ if

$$f(\nu + \eta) - f(\nu) = Df(\nu)[\eta] + o(\|\eta\|) \text{ as } \|\eta\| \downarrow 0, \quad (13)$$

where $Df(\nu)$ is a bounded linear continuous functional on \mathbb{M} called the *differential*.

When a function is strongly differentiable at ν then there also exists a *weak* or *directional* or *Gateaux derivative*, *i. e.*

$$\lim_{t \downarrow 0} t^{-1}[f(\nu + t\eta) - f(\nu)] = Df(\nu)[\eta] \quad (14)$$

for any ‘direction’ $\eta \in \mathbb{M}$.

The differential $Df(\nu)$ often has an integral form:

$$Df(\nu)[\eta] = \int_X g(x; \nu) \eta(dx)$$

for some function $g(\bullet; \nu) : X \mapsto \mathbb{R}$ which is then called the *gradient function* to f at ν . Not all linear functionals are integrals, unless the space X is a finite set in which case \mathbb{M} can just be identified with an Euclidean space. In most applications, however, including the experimental design, differentiable functionals do possess a gradient function, so this assumption is not too restrictive in practice.

In this study we are interested in optimisation of functionals of positive measures. A general constrained optimisation problem can be written as follows:

$$f(\mu) \longrightarrow \inf, \quad \mu \in \mathbb{A}, \quad (15)$$

where $\mathbb{A} \subseteq \mathbb{M}_+$ is a set of measures describing the constraints. If f is strongly differentiable then the first order necessary optimality condition states that if μ^* provides a local minimum in the problem (15) then

$$Df(\mu^*)[\eta] \geq 0 \quad \text{for all } \eta \in T_{\mathbb{A}}(\mu^*), \quad (16)$$

where

$$T_{\mathbb{A}}(\mu) = \liminf_{t \downarrow 0} \frac{\mathbb{A} - \mu}{t} \quad (17)$$

is the *tangent cone* to \mathbb{A} at μ . Here the ‘+’ (respectively ‘−’) operation on sets indicates all pairwise sums of (respectively difference between) the points from the corresponding sets. The tangent cone is the closure of all admissible directions $\eta \in \mathbb{M}$ at μ meaning that $\mu + t\eta \in \mathbb{A}$ for all sufficiently small $t > 0$, see, e. g., Cominetti (1990). In other words, derivative in all admissible directions should be non-negative at a point of local minimum. For any \mathbb{A} of interest, one needs to characterise the tangent cone $T_{\mathbb{A}}(\mu)$.

General optimisation theory for functionals of measures has been developed in a series of papers Molchanov and Zuyev (2000a), Molchanov and Zuyev (2000b) and Molchanov and Zuyev (2004). For us here an optimisation with finite number of equality and inequality constraints is relevant.

Consider the following optimisation problem:

$$f(\mu) \rightarrow \inf, \quad \mu \in \mathbb{M}_+ \text{ subject to} \quad (18)$$

$$\begin{cases} H_i(m) = 0 & i = 1, \dots, k, \quad k \leq d \\ H_j(m) \leq 0 & j = k + 1, \dots, d. \end{cases} \quad (19)$$

where $f : \mathbb{M}_+ \mapsto \mathbb{R}$ and $H = (H_1, \dots, H_d) : \mathbb{M}_+ \mapsto \mathbb{R}^d$ are strongly differentiable functions. Alternatively, the constraints (19) can be written in the form $H(\mu) \in C$, where $C \subset \mathbb{R}^d$ is the cone $\{\mathbf{y} \in \mathbb{R}^d : y_1 = \dots = y_k = 0, y_{k+1} \leq 0, \dots, y_d \leq 0\}$.

Definition 1 (Robinson (1976)). A measure μ is called *regular* for optimisation problem (18) if the origin 0 of \mathbb{R}^d belongs to the interior of the set

$$H(\mu) - C + DH(\mu)[\mathbb{M}_+ - \mu] \subseteq \mathbb{R}^d. \quad (20)$$

Robinson’s regularity condition which as shown in Zowe and Kurcyusz (1979), guarantees the existence and boundedness of the *KuhnTucker*

vector appearing in the next theorem. See also Maurer and Zowe (1979) for the discussion of different forms of regularity condition.

The following theorem gives the first-order necessary conditions for a minimum in the problem (18).

Theorem 1. *Molchanov and Zuyev (2000a, Th. 4.1). Let μ^* be a regular local minimum of f over \mathbb{M}_+ , subject to (19). Assume that f and H are Fréchet differentiable at μ^* , and there exist the corresponding gradient functions $g(x, \mu^*)$ and $h_i(x, \mu^*), i = 1, \dots, d$. Then there exists Kuhn-Tucker vector $(u_1, \dots, u_d) \in \mathbb{R}^d$ such that $u_j \leq 0$ (resp. $u_j = 0$) for those $j \in \{k+1, \dots, d\}$ satisfying $H_j(\mu^*) = 0$ (resp. $H_j(\mu^*) < 0$), such that*

$$g(x, \mu^*) \begin{cases} = \sum_{i=1}^d u_i h_i(x, \mu^*) & \text{for } \mu^*\text{-almost all } x, \\ \geq \sum_{i=1}^d u_i h_i(x, \mu^*) & \text{for all } x \in X. \end{cases} \quad (21)$$

One can show that in the case of finitely many constraints (19) satisfying (21), the regularity condition (20) becomes the so-called *Mangasarian-Fromowitz constraints qualification* (see Cominetti (1990, p. 274)), that is the linear independence of the gradient functions $h_1(\cdot, \mu^*), \dots, h_k(\cdot, \mu^*)$ and the existence of a measure $\eta \in \mathbb{M}$ such that

$$\int h_i(x, \mu^*) \eta(dx) = 0 \quad \text{for all } i = 1, \dots, k; \quad (22)$$

$$\int h_j(x, \mu^*) \eta(dx) < 0 \quad \text{for all } j \in \{k+1, \dots, d\} \\ \text{verifying } H_j(\mu^*) = 0. \quad (23)$$

Without the inequality constraints, condition (23), trivially holds for η being the zero measure.

The design problems we consider naturally fall in the above described general framework of optimisation of functionals defined on finite measures. Theorem 1 provides the necessary conditions for optimality of a design. Moreover, it allows one to easily incorporate into the model other constraints on the optimal design measure, if needed. Constraints (8) and (9) correspond to linear functionals $H_1(\mu) = \int \mu(dx) - n$ and $H_2(\mu) = \int x \mu(dx) - 1$ with the corresponding gradient functions $h_1(x; \mu) \equiv 1$ and $h_2(x; \mu) = x$. These constraints are regular for any μ since (22) and (23) are satisfied, for instance, for a measure $\eta = \delta_0 - \delta_1$. We therefore have the following important corollary of Theorem 1 which we use in the next section. Note that we mostly *maximise* the goal function so that the inequalities in (21) change to opposite.

Theorem 2. *Let μ^* be a local maximum of a strongly differentiable function $f : \mathbb{M}_+ \mapsto \mathbb{R}$ possessing a gradient function $g(x; \mu^*)$, subject*

to constraints (8) and (9). Then, there exist constants u_1 and u_2 , where $u_2 \geq 0$ if $\int x\mu^*(dx) = 1$ and $u_2 = 0$ if $\int x\mu^*(dx) < 1$, such that

$$g(x, \mu^*) \begin{cases} = u_1 + u_2x & \mu^*\text{-almost everywhere,} \\ \leq u_1 + u_2x & \text{for all } x \in X. \end{cases} \quad (24)$$

4 Construction of optimal design

In this section we apply the necessary condition for extremum of a functional of measures to find optimal designs for a range of goal functions and most common prior distributions in the Bayesian settings. First we assume that the parameter λ , the mean number of HSCs in the substrate, is known from previous experiments, and obtain the optimal design, in terms of maximisation of the Fisher information, for each λ .

4.1 Optimal design for a fixed average number of HSCs

Here we are dealing with optimisation problem (7) under constraints (8) and (9). The goal function is a linear function of μ , so that its differential is the function itself with the gradient function

$$g_1(x; \lambda) = \frac{e^{-\lambda x}}{1 - e^{-\lambda x}} x^2, \quad (25)$$

independent of μ . Note that $g_1(x; \lambda) = \lambda^{-2} r(\lambda x)$, where

$$r(y) = \frac{e^{-y}}{1 - e^{-y}} y^2. \quad (26)$$

The graph of r is shown on Figure 1. It attains its unique maximum at point $y_{max} \approx 1.59362$ and it is strictly concave on $[0, y_{max}]$.

The gradient function is just scaled in both directions of the function r , attaining its maximum at point y_{max}/λ , respectively. It follows from Theorem 2, that if μ^* is a measure at which G_1 attains its maximum, then $g_1(x; \lambda) \leq u_1 + u_2 x$ for all $x \in [0, 1]$ with $u_2 \geq 0$. Moreover, $g_1(x; \lambda) = u_1 + u_2 x$ for $x \in \text{supp } \mu^*$. But this is only possible if $u_1 + u_2 x$ is a tangent line to g_1 at some $x^* \in [0, y_{max}/\lambda]$ and hence the support of optimal μ^* consists of only this point x^* . Using (8) and substituting $\mu = n\delta_x$ into (7) and (9), we come to the optimisation problem of one variable:

$$r(\lambda x) \rightarrow \sup \quad \text{over } x \in [0, 1/n], \quad (27)$$

so that $x^* = 1/n$ for $\lambda \leq y_{max} n$ and $x^* = y_{max}/\lambda$, otherwise. Thus we have proved the following theorem.

Theorem 3. *The optimal design for the problem (7) under constraints (8) and (9) consists in n equal doses of volume $1/n$ for $\lambda \leq y_{max} n$ and of volume y_{max}/λ for $\lambda > y_{max} n$, where $y_{max} \approx 1.59362$ is the maximum point of the function r (26).*

This indicates that for those $\lambda > 1.59362n$ we need to sample a proportion of the substrate $1.59362n/\lambda$, and for those $\lambda \leq 1.59362n$ we have to take all the substrate. Therefore, if a good prior point estimate of λ is available, a near optimal doses of volume x^* can be selected, see Figure 1.

4.2 Optimal design with prior distribution on λ

Typically researchers already have an idea on what are the most likely values of λ . This can be suggested by previous experiments or by analogy with other similar cases and it is given in the form of a *prior distribution* $Q(d\lambda)$. The optimal experimental design now becomes dependent not only on the choice of a criterion, but also on the parameters of Q . Formally, the locations of pre-HSCs in the substrate now conform to a Cox (or mixed Poisson) point processes with parameter measure Q .

In this section we carry out optimisation of the three goal functions introduced above in (10)–(12) for the most common prior distributions: Uniform and Gamma.

Uniform prior distribution. The Uniform prior distribution is appropriate when there is no knowledge on the mean number λ of pre-HSCs in substrate apart from its range. Certainly, λ should be greater than 1 since *pre-HSCs* do develop in the AGM. So, it is reasonable to assume that $\lambda \sim \text{Unif}(1, u)$, where $u > 1$ is a known upper bound. It was already mentioned that the number of pre-HSCs hardly exceeds 200, so one can set $u = 200$, or, keeping in mind the variance of the Poisson distribution, $u = 170$, for instance.

The goal function (10) now becomes

$$G_2(\mu; u) = \frac{1}{u-1} \int_1^u \int_0^1 \frac{e^{-\lambda x}}{1-e^{-\lambda x}} x^2 \mu(dx) d\lambda. \quad (28)$$

By *Fubini's theorem* we can change the order of integrals above to arrive at

$$G_2(\mu; u) = \frac{1}{u-1} \int_0^1 x \log \frac{1-e^{-ux}}{1-e^{-x}} \mu(dx). \quad (29)$$

Thus $G_2(\mu; u)$ is a linear functional of μ with the gradient function

$$g_2(x; u) = \frac{x}{u-1} \log \frac{1-e^{-ux}}{1-e^{-x}}. \quad (30)$$

This function varies very little for all practically interesting values of $u \geq 5$ with the maximum value attained around 0.69–0.7, see Figure 2. Reasoning the same way as we did in Section 4.1, we can conclude that

for any $u \geq 2$ the optimal design is attained on one atom design measure concentrated on the point $1/n$ (*cf.* Plot (c) in Figure 1). So the whole volume of the substrate should be divided into n equal doses for this goal function.

Consider now optimisation criterion (11) which now takes the form

$$G_3(\mu; u) = \frac{1}{u-1} \int_1^u \log \left(\int_0^1 \frac{e^{-\lambda x}}{1-e^{-\lambda x}} x^2 \mu(dx) \right) d\lambda. \quad (31)$$

This goal function is not linear, but still Fréchet differentiable with the gradient function given by

$$g_3(x, \mu; u) = \frac{1}{u-1} \int_1^u I^{-1}(\mu; \lambda) \frac{e^{-\lambda x}}{1-e^{-\lambda x}} x^2 d\lambda, \quad (32)$$

where

$$I(\mu; \lambda) = \int_0^1 \frac{e^{-\lambda x}}{1-e^{-\lambda x}} x^2 \mu(dx) \quad (33)$$

is the Fisher information written in terms of a design measure. Since the gradient now depends on μ , numeric methods have to be employed to find the optimal design for a given value of the upper bound u and the number of mice n . This was done by means of R-library `medea` which finds an optimal solution to a measure optimisation problem under linear constraints of equality type, see Molchanov and Zuyev (2002). Since `medea` does not yet allow for inequality constraints, to deal with (9), the procedure looks for an optimum measure for a given value b of the integral $\int x \mu(dx)$ and then optimises over $b \leq 1$. The R-code is freely available from one of the authors web-page⁴.

Numeric experiments conducted for various values of u show that the obtained optimal solution is always one-atom as in the previous cases. Typically, the numeric solution gives two atoms at the neighbouring points of the discretised space $[0, 1]$ for x which indicates a single atom is situated in between these grid points, see Figure 3. Although we cannot formally prove that the optimum design is one-atom, such designs (*i. e.* equal doses) are certainly of an interest.

One-atom measures under constraint (8) have a form $\mu = n\delta_x$ for some $x \in (0, 1/n]$. Making use of (26), we come to a one variable optimisation problem: maximise

$$G_3(n\delta_x) = \log n - 2 \mathbf{E}_Q \log \lambda + \mathbf{E}_Q \log r(\lambda x) \quad (34)$$

⁴<http://www.math.chalmers.se/~sergei>

subject to $x^* \leq 1/n$ and Q is $\text{Unif}[1, u]$. Only the last term depends on x , so the equivalent problem is to maximise

$$\tilde{G}_3(x; u) = \mathbf{E}_Q \log r(\lambda x) = \frac{1}{u-1} \int_1^u \log r(\lambda x) d\lambda. \quad (35)$$

As u grows, the maximum x_{max} of this function, which can easily be computed numerically, approaches 0, see the right plot on Figure 4. So when $x_{max} \geq 1/n$, the constrained maximum is attained at the point $1/n$. Otherwise, for large u , $x_{max} < 1/n$ and the solution is to take n doses of volume x_{max} . This is exemplified at Figure 4 for $n = 30$: the optimal dose is given by

$$x^* = \begin{cases} \frac{1}{n} & u \leq u^*; \\ x_{max} & u > u^*. \end{cases} \quad (36)$$

where $u^* \approx 90.66$. This indicates that for those $u \leq u^*$ one needs to take all the substrate to make n equal doses, and the volume nx_{max} otherwise.

Finally, consider objective function G_4 given by (12) which corresponds to the expected asymptotic variance of the maximum likelihood estimator. In the case of a Uniform prior $\lambda \sim \text{Unif}(1, u)$,

$$G_4(\mu; u) = -\frac{1}{u-1} \int_1^u I^{-1}(\lambda, \mu) d\lambda, \quad (37)$$

where $I(\mu; \lambda)$ is given by (33). Again, G_4 is Fréchet differentiable with the gradient function

$$g_4(x, \mu; u) = \frac{1}{u-1} \int_1^u I^{-2}(\lambda, \mu) \frac{e^{-\lambda x}}{1 - e^{-\lambda x}} x^2 d\lambda \quad (38)$$

and a numeric procedure should be used to find the optimal measure for any given values of u and n . Similarly to the case of objective function G_3 above, numeric experiments show that the optimal measure is one atomic, although we cannot show this rigorously. In the class of one-atomic measures $\mu = n\delta_x$ the goal function simplifies to

$$G_4(n\delta_x; u) = -\frac{e^{ux} - e^x - x(u-1)}{nx^3(u-1)} \quad (39)$$

Similarly to G_3 above, the point of maximum x_{max} of this function approaches 0 when u grows, see the right plot on Figure 5. So that when u is such that $x_{max} > 1/n$ the optimal dose is $1/n$ (the whole substrate is used), otherwise the optimal dose is x_{max} . For example, for $n = 30$ mice, the optimal dose is also given by (36) but with $u^* \approx 64.47$ this time.

Gamma prior distribution. Gamma prior distribution arises naturally as a posterior distribution for the Poisson parameter λ which has a Uniform prior distribution. Since the probability of having k pre-HSCs in the substrate is

$$\mathbf{P}(\omega(V) = k) = \mathbf{E}_Q \mathbf{P}(\omega(V) = k|\lambda) = \frac{1}{u-1} \int_1^u \frac{\lambda^k e^{-\lambda}}{k!} d\lambda, \quad (40)$$

The posterior p.d.f. for λ is then

$$f_\lambda(x|\omega(V) = k) = \frac{x^k e^{-x}}{\int_1^u \lambda^k e^{-\lambda} d\lambda} \propto x^k e^{-x}, \quad (41)$$

which is close to $\Gamma(k+1, 1)$ distribution for large u . So once an estimate for the total number of pre-HSCs k is available, the Gamma prior would be a reasonable for λ in the subsequent experiment. Moreover, posterior distribution for a Gamma prior $\Gamma(\alpha, 1)$ given k pre-HSCs will be $\Gamma(\alpha + k, 2)$ and so on. So generally, $\lambda \sim \Gamma(\alpha, \beta)$ with rate parameter $\beta \in \mathbb{N}$ could be a reasonable assumption.

Under this assumption, Criterion (10) becomes

$$G_2(\mu; \alpha, \beta) = \int_0^\infty \int_0^1 f_{\alpha, \beta}(\lambda) \frac{e^{-\lambda x}}{1 - e^{-\lambda x}} x^2 \mu(dx) d\lambda, \quad (42)$$

where

$$f_{\alpha, \beta}(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta \lambda}. \quad (43)$$

So G_2 is a linear functional with the gradient function

$$\begin{aligned} g_2(x, \mu; \alpha, \beta) &= \int_0^\infty \frac{e^{-\lambda x}}{1 - e^{-\lambda x}} x^2 f_{\alpha, \beta}(\lambda) d\lambda = \int_0^\infty r(\lambda x) \lambda^{-2} f_{\alpha, \beta}(\lambda) d\lambda \\ &= \frac{\beta^2}{(\alpha-1)(\alpha-2)} \int_0^\infty r(\lambda x) f_{\alpha-2, \beta}(\lambda) d\lambda \\ &= \frac{\beta^2}{(\alpha-1)(\alpha-2)} \int_0^\infty r(\lambda) f_{\alpha-2, \beta/x}(\lambda) d\lambda. \end{aligned} \quad (44)$$

Thus the situation here is similar to the case of the Uniform distribution: depending on whether the maximum of this function is below or above $1/n$ one needs to take equal doses of volumes x_{max} the point where g_2 attains its maximum or $1/n$, respectively.

Consider as an example the first-iteration case when $\beta = 1$. The gradient function and the point of maximum are shown in Figure 6. The optimal design is given by

$$x^* = \begin{cases} \frac{1}{n} & \alpha \leq \alpha^*; \\ x_{max} & \alpha > \alpha^*. \end{cases} \quad (45)$$

where $\alpha^* \approx 49.68$.

Under Gamma prior (43), criterion function (11) takes the following form:

$$G_3(\mu; \alpha, \beta) = \int_0^{\infty} f_{\alpha, \beta}(\lambda) \log I(\mu; \lambda) d\lambda, \quad (46)$$

where $I(\mu; \lambda)$ is given by (33). Note that G_3 is not a linear functional of a measure, however it is Fréchet differentiable with the gradient function

$$g_3(x, \mu; \alpha, \beta) = \int_0^{\infty} I^{-1}(\mu; \lambda) \frac{e^{-\lambda x}}{1 - e^{-\lambda x}} x^2 f_{\alpha, \beta}(\lambda) d\lambda. \quad (47)$$

Thus, as before numeric methods for $\beta = 1$ and a given values of α and n should be employed. Our experiments show that the optimum measure given by the steepest descent algorithm still contains only one atom. So, optimising (46) for one atomic measures $n\delta_x$ over $x \in (0, 1/n]$ leads to the design given by (45), with $\alpha^* \approx 47.70$. Figure 7 shows $G_3(n\delta_x; \alpha)$ and the optimal equal doses for different values of α .

Finally, consider G_4 in (12) with Gamma prior distribution on λ :

$$G_4(\mu; \alpha, \beta) = - \int_0^{\infty} I^{-1}(\mu; \lambda) f_{\alpha, \beta}(\lambda) d\lambda, \quad (48)$$

which is Fréchet differentiable with the gradient function

$$g_4(x, \mu; \alpha, \beta) = \int_0^{\infty} I^{-2}(\mu; \lambda) \frac{e^{-\lambda x}}{1 - e^{-\lambda x}} x^2 f_{\alpha, \beta}(\lambda) d\lambda. \quad (49)$$

Here also, our numeric experiments for $\beta = 1$ and any given values of α and n , show that the optimal measure contains a single atom. Thus, in the class of one atomic measures, (48) simplifies to

$$G_4(n\delta_x; \alpha) = - \frac{1}{nx^2} \int_0^{\infty} \frac{1 - e^{-\lambda x}}{e^{-\lambda x}} f_{\alpha}(\lambda) d\lambda. \quad (50)$$

which leads to the one-atom optimal design (45) with $\alpha^* \approx 45.74$, see Figure 8.

4.3 Optimal design with additional costs

The measure optimisation formalism we used above allows for seamless inclusion of different costs associated with the experiment either as additional terms in the goal function or as additional constraints. This can be done in both non-Bayesian and Bayesian settings. Typical loss incurred in the above described experiment are the non-repopulated laboratory mice which are quite expensive. Added to this, the cost of all the experiment when all mice die or all mice survive so it does not yield meaningful results. In this section we demonstrate how these costs affect the optimal experimental design.

Non-Bayesian setting. To take into account the cost of non-repopulated mice, note that for a fixed λ , the indicator χ_i that the mouse which received the i th dose of volume x_i has not repopulated is a Bernoulli random variable with a parameter $e^{-\lambda x_i}$, see (3). Thus, the mean number of non-repopulated mice $\mathbf{E} \sum_{i=1}^n \chi_i$ is

$$T_1(\mathbf{x}; \lambda) = \sum_{i=1}^n e^{-\lambda x_i}, \quad (51)$$

which can be represented as a functional of the design measure μ as follows:

$$T_1(\mu; \lambda) = \int e^{-\lambda x} \mu(dx). \quad (52)$$

Recall that as in the previous sections, all the integrals are taken over the range $(0, 1]$. To include the cost associated with a spoilt experiment, compute the probability that all the mice do not repopulate which can be written as follows:

$$\begin{aligned} T_{21}(\mathbf{x}; \lambda) &= \prod_i^n \mathbf{P}(\chi_i = 1) \\ &= \prod_{i=1}^n e^{-\lambda x_i} = e^{-\lambda \sum_{i=1}^n x_i}. \end{aligned} \quad (53)$$

Consequently, the probability that all the mice repopulate is

$$T_{20}(\mathbf{x}; \lambda) = \prod_{i=1}^n \mathbf{P}(\chi_i = 0) = \prod_{i=1}^n (1 - e^{-\lambda x_i}) = e^{\sum_{i=1}^n \log(1 - e^{-\lambda x_i})}. \quad (54)$$

In terms of the design measure, (53) and (54) are the following functionals:

$$T_{21}(\mu; \lambda) = e^{-\lambda \int x \mu(dx)} \quad (55)$$

$$T_{20}(\mu; \lambda) = e^{\int \log(1 - e^{-\lambda x}) \mu(dx)}. \quad (56)$$

This gives rise to the expression for the probability of a spoilt experiment due to either all the mice repopulating or all the mice non-repopulating:

$$T_2(\mu; \lambda) = T_{21}(\mu; \lambda) + T_{20}(\mu; \lambda). \quad (57)$$

We note that T_1 and T_2 are Fréchet differentiable with the gradient functions

$$t_1(x; \lambda) = e^{-\lambda x} \quad (58)$$

and

$$t_2(\mu, x; \lambda) = -\lambda x T_{21}(\mu; \lambda) + \log(1 - e^{-\lambda x}) T_{20}(\mu; \lambda). \quad (59)$$

Now consider maximisation of a new goal function

$$\tilde{G}_1(\mu; \lambda) = G_1(\mu; \lambda) - c_1 T_1(\mu; \lambda) - c_2 T_2(\mu; \lambda), \quad (60)$$

where G_1 is given by (7), under constraints (8) and (9). The positive constants c_1, c_2 should be set by the experimentator to reflect the cost of mice and of a spoilt experiment which should be offset against the usefulness of the results reflected in the original goal function G_1 involving the Fisher information.

The function \tilde{G}_1 is strongly differentiable with the gradient function given by

$$\tilde{g}_1(\mu, x; \lambda) = g_1(x; \lambda) - c_1 t_1(x; \lambda) - c_2 t_2(\mu, x; \lambda), \quad (61)$$

where g_1 as in (25).

Figure 9 shows the optimal dose as a function of λ , for $n = 30$, $c_1 = 10^{-4}$ and $c_2 = 1$. Compared to the optimal design for G_1 shown on Figure 1 one has to dilute more and also to start dilution earlier at $\lambda = 41.8$ compared to 47.8 for G_1 which reflects the caution not to allow all the mice to survive as this would mean a spoilt experiment.

Bayesian setting. Under $\Gamma(\alpha, 1)$ prior distribution for λ , the average number of non-repopulated mice becomes

$$\mathbf{E}_Q T_1(\mu; \lambda) = \int \frac{1}{(1+x)^\alpha} \mu(dx), \quad (62)$$

with the gradient function

$$\mathbf{E}_Q t_1(x; \lambda) = \frac{1}{(1+x)^\alpha}. \quad (63)$$

The probability of a spoilt experiment is

$$\begin{aligned} \mathbf{E}_Q T_2(\mu; \lambda) &= \mathbf{E}_Q T_{21}(\mu; \lambda) + \mathbf{E}_Q T_{20}(\mu; \lambda) \\ &= \frac{1}{(1+H(\mu))^\alpha} + \mathbf{E}_Q T_{20}(\mu; \lambda), \end{aligned} \quad (64)$$

with the corresponding gradient function

$$\mathbf{E}_Q t_2(\mu, x; \lambda) = -\frac{\alpha x}{(1 + H(\mu))^{\alpha+1}} + \mathbf{E}_Q \log(1 - e^{-\lambda x}) T_{20}(\mu; \lambda), \quad (65)$$

where $H(\mu) = \int x \mu(dx)$.

Take G_4 as in (48) with $\beta = 1$ and consider a new goal function

$$\tilde{G}_4(\mu; \alpha) = G_4(\mu; \alpha) - c_1 \mathbf{E}_Q T_1(\mu; \lambda) - c_2 \mathbf{E}_Q T_2(\mu; \lambda). \quad (66)$$

It is Fréchet differentiable and possesses a gradient function

$$\tilde{g}_4(\mu, x; \alpha) = g_4(\mu, x; \alpha) - c_1 \mathbf{E}_Q t_1(x; \lambda) - c_2 \mathbf{E}_Q t_2(\mu, x; \lambda). \quad (67)$$

We are aiming to maximise \tilde{G}_4 under the same constraints (8) and (9).

Again, numeric solutions are still one-atomic for different tried values of the parameters α, c_1 and c_2 we tried. Figure 10 shows a typical picture for $-\log |\tilde{G}_4(n\delta_x; \alpha)|$ for the class of one-atomic design measures and the optimal doses for three different values of α .

Mixture prior distribution Finally, consider the setting typical for testing two alternative hypotheses about possible values of the mean number of pre-HSCs. Suppose that λ can take two different values, our prior belief is that it is rather λ_1 than λ_2 with probability $0 < p < 1$. This gives rise to optimisation of the following goal function

$$G_{1M}(\mu; p, \lambda_1, \lambda_2) = p G_1(\mu; \lambda_1) + (1 - p) G_1(\mu; \lambda_2). \quad (68)$$

under constraints (8) and (9).

The gradient function for (68) is

$$g_{1M}(x; p, \lambda_1, \lambda_2) = p g_1(x; \lambda_1) + (1 - p) g_1(x; \lambda_2). \quad (69)$$

and numerical optimisation can be employed to find the optimal design for any particular values of λ_1, λ_2 and p . For example of $\lambda_1 = 25, \lambda_2 = 150$ and $p = 0.05$, Figure 11 shows a numerically obtained optimal measure which is *two-atomic*: 19 mice should receive equal doses of 0.019 and the rest 11 of them should receive the same doses of 0.058.

5 Discussion

We have considered dilution experiments with volume constraints typical in biological and medical research. An important particularity of the experiments we consider here is that the time delay necessary for pre-HSCs to develop and then for the injected doses to take effect in mice prevents from planning multiple stage experiments: the estimation should necessarily be done from the first and only stage. As we have seen, in all the considered cases of the goal functions, the optimal design is attained on a one-point measure, meaning that all the doses should have equal volume. This parallels the well known result about the D-optimal design measure for a linear regression model: The Kiefer-Wolfowitz theorem assures that such measure is atomic with the number of atoms to be at most the number of constraints plus one, see, e. g., Kiefer and Wolfowitz (1960). In our case we have one volume constraint (in addition to the measure to have a fixed total mass) so the number of atoms is at most two. Indeed, the gradient functions we observed are convex on the interval from 0 to the point of maximum, so the only possibility to satisfy the necessary optimality condition given in Theorem 2, is for the optimal measure to have only one atom.

In practical terms, if the number of pre-HSC expected to be not very large (a few dozens or less), the whole substrate should be used to derive the doses, otherwise only part of it. We have characterised above what is ‘not very large’ and how much the doses should be diluted. We have fixed some of the parameters here (the number of mice $n = 30$, $\beta = 1$), driven by practical applications to stem cell research. For other applications and other values of these parameters one make use the computer codes the authors make freely available for download.

Advantage of measure optimisation approach is that additional requirements can be easily incorporated into the goal function or added as further constraints, *e. g.* limited cost associated with non-repopulated mice and/or the cost of the whole experiment if all mice are repopulated or all are not repopulated.

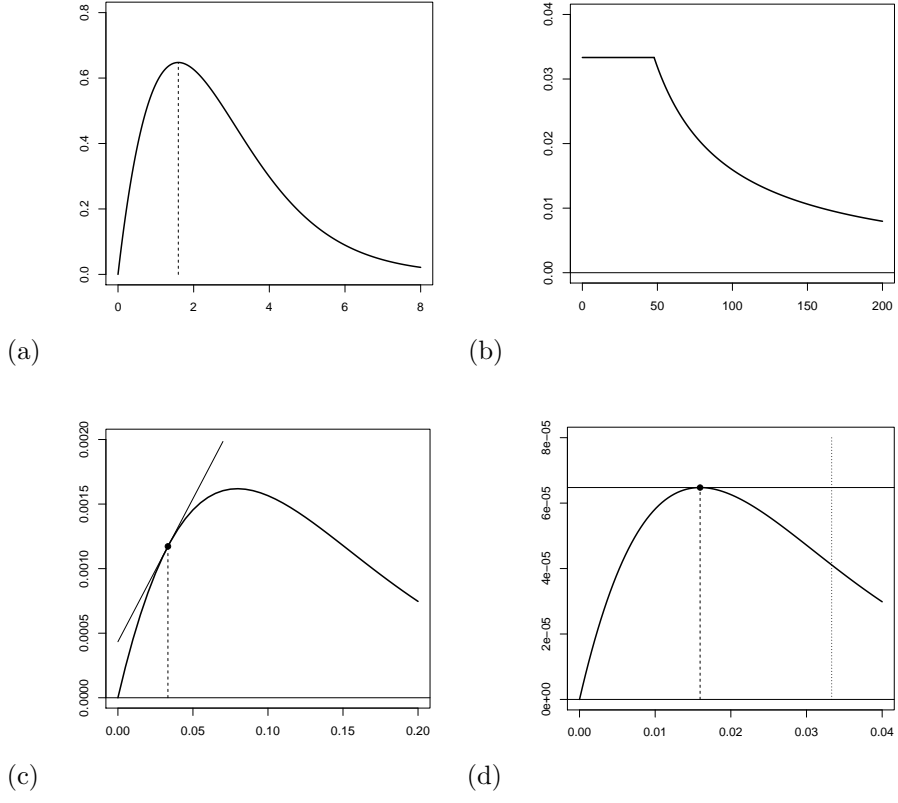


Figure 1: Plot (a): function r , the maximum is attained at the point $y_{max} \approx 1.59362$. (b): the optimal dose volume for $n = 30$ mice as a function of λ : $1/30$ for $\lambda \leq 47.8$ and $1.59362/\lambda$ otherwise. The line $u_1 + u_2x$ and the gradient function $g_1(x; \lambda)$ satisfying conditions (24) for $\lambda = 20$ (Plot (c), the tangent line at point $1/30$) and for $\lambda = 100$ (Plot (d), the tangent line at the point of maximum $y_{max}/\lambda < 1/30$).

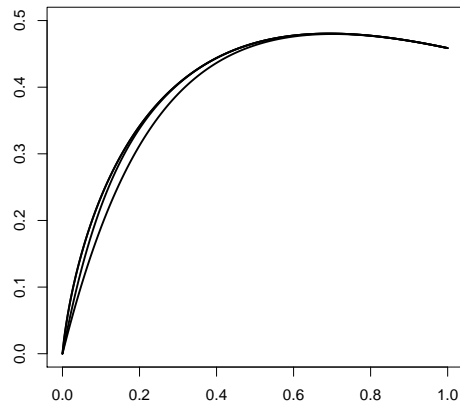


Figure 2: Function $(u - 1)g_2(x; u)$ for the values $u = 10, 20, 200, 1000$ (in the order of increase). The last two curves are almost indistinguishable.

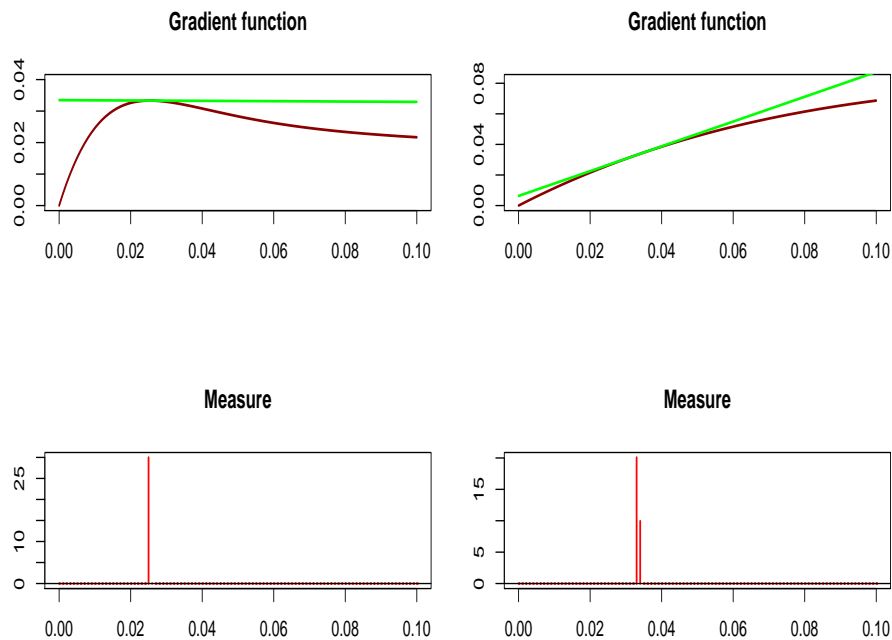


Figure 3: Numeric solution obtained by `medea` for $u = 120$ (left pair of plots) and for $u = 20$ (right pair) for $n = 30$ mice. For $u = 120$, one atomic optimal measure is $30 \delta_{x^*}$, where $x^* = 0.02522$. The numeric solution is two atom measure $29.97 \delta_{0.025} + 0.03 \delta_{0.026}$. In a similar way, we have for $u = 20$, one atomic optimal measure $30 \delta_{x^*}$ with $x^* = 1/30$. Moreover, the numeric solution is two atom measure $20.07 \delta_{0.033} + 9.93 \delta_{0.034}$. So, the total mass of 30 was distributed between the grid points surrounding the true atom position.

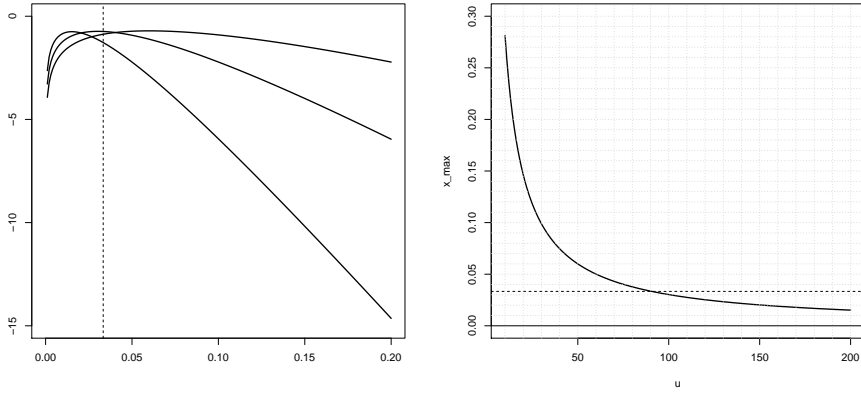


Figure 4: Left plot: function \tilde{G}_3 for $u = 50, 100$ and 200 . As u grows, the point of maximum x_{max} approaches 0. The vertical dashed line is through the point $1/n$ with $n = 30$. Right plot: x_{max} as a function of u (horizontal dashed line is through $1/n$ for $n = 30$).

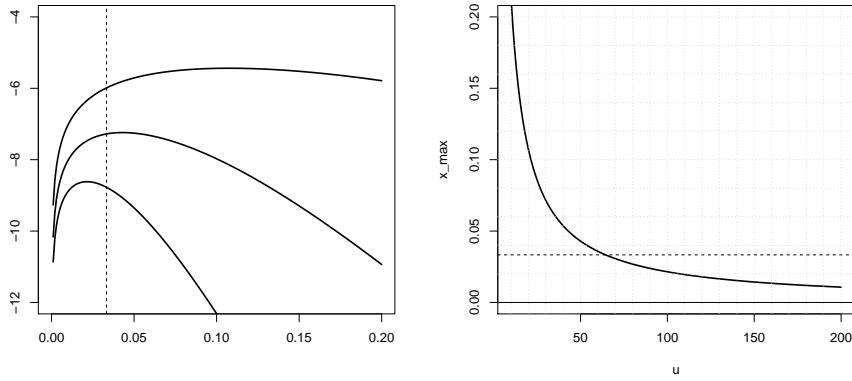


Figure 5: Left plot: function $-\log(|nG_4(n\delta_x; u)|)$ for $u = 20, 50$ and 100 (from upper to lower curve). As u grows, the point of maximum x_{max} approaches 0. The vertical dashed line is through the point $1/n$ with $n = 30$. Right plot: x_{max} as a function of u (horizontal dashed line is through $1/n$ for $n = 30$).

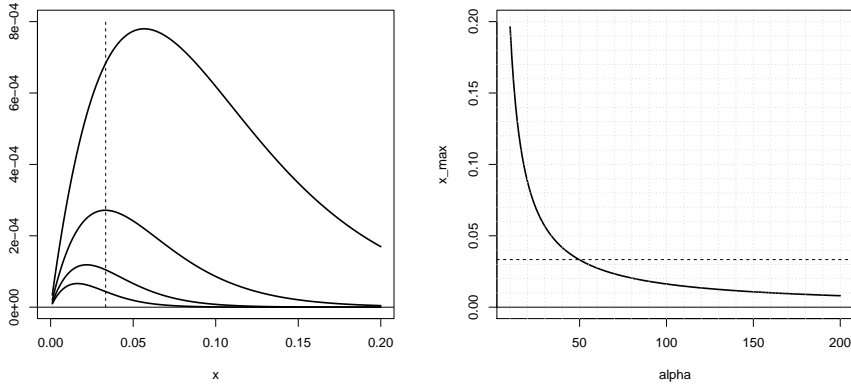


Figure 6: Left plot: function $g_2(x)$ for $\beta = 1$ and $\alpha = 30, 50, 75, 100$ (from top to bottom). The vertical dashed line is through the point $1/n$ with $n = 30$. Right plot: x_{max} as a function of α (horizontal dashed line is through $1/n$ for $n = 30$ intersection the curve at point $u^* \approx 49.68$).

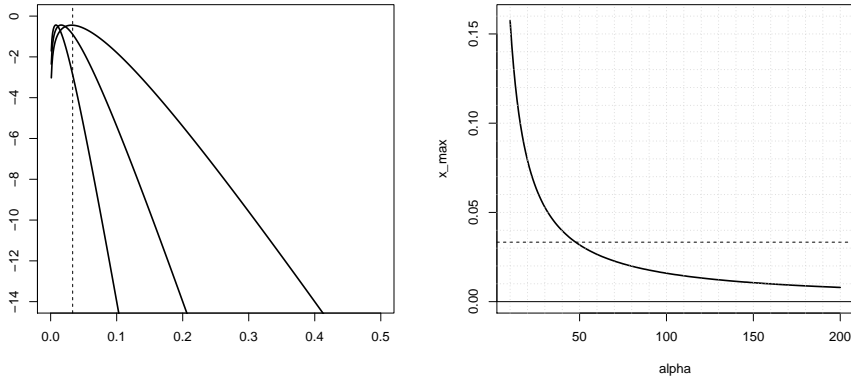


Figure 7: Left plot: function $G_3(n\delta_x; \alpha)$ for $\beta = 1$ and $\alpha = 50, 100$ and 200. Similarly to the Uniform prior case as α grows, the point of maximum x_{max} approaches 0. The vertical dashed line is through the point $1/n$ with $n = 30$. Right plot: x_{max} as a function of α (horizontal dashed line is through $1/n$ for $n = 30$).

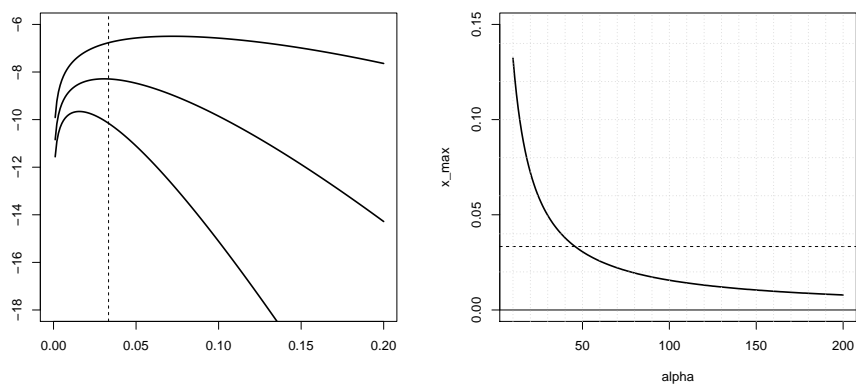


Figure 8: Left plot: function $-\log(|nG_4(n\delta_x; \alpha)|)$ for $\beta = 1$ and $\alpha = 20, 50$ and 100 . Similar to the Uniform case for G_4 , as α grows, the point of maximum x_{max} approaches 0. The vertical dashed line is through the point $1/n$ with $n = 30$. Right plot: x_{max} as a function of α (horizontal dashed line is through $1/n$ for $n = 30$).

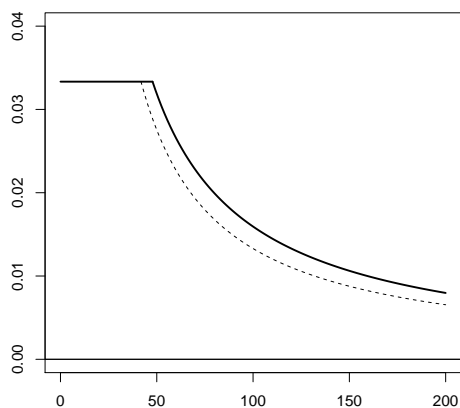


Figure 9: The optimal doses for different values of λ for the goal function \tilde{G}_1 when $n = 30$, $c_1 = 10^{-4}$ and $c_2 = 1$ (dashed) and for G_1 (solid line) as on Plot (c) in Figure 1.

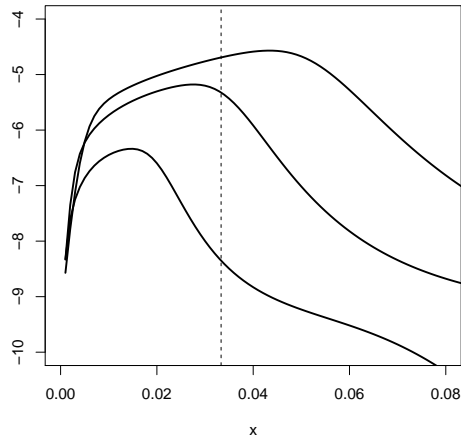


Figure 10: Function $-\log |\tilde{G}_4(n\delta_x; \alpha)|$ for the costs $c_1 = 5$, $c_2 = 10^4$ and values $\alpha = 30, 50, 100$ (in the order of a decrease). The vertical dashed line is through the point $1/n$ with $n = 30$. The extremal points are $x_{max} = 0.043, 0.028$ and 0.015 leading to optimal designs $1/30, 0.028$ and 0.015 , respectively.

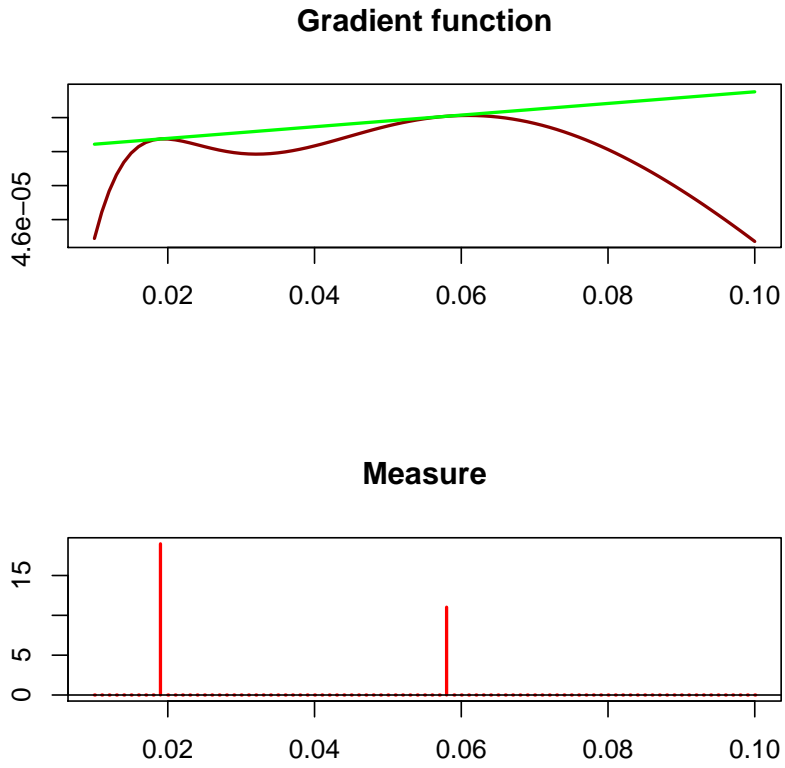


Figure 11: Solution obtained by `medea` for $p = 0.05$, $\lambda_1 = 25$, and $\lambda_2 = 150$ shows that the numeric solution is a two atomic measure $18.98 \delta_{0.019} + 11.02 \delta_{0.058}$ and all the substrate should be used.

References

- Atkinson, A. C. and A. N. Donev (1992). *Optimum Experimental Designs*. Oxford: Clarendon Press.
- Bonnefoix, T. and M. Callanan (2010). Accurate hematopoietic stem cell frequency estimates by fitting multicell poisson models substituting to the single-hit poisson model in limiting dilution transplantation assays. *Blood* 116(14), 2472–2475.
- Chaloner, K. and K. Larntz (1989). Optimal bayesian design applied to logistic regression experiments. *J. Statis. Plannng Inf.* 21, 191–208.
- Cochran, W. G. (1950). Estimation of bacterial densities by means of the "Most Probable Number". *Biometrics* 6, 105–116.
- Cominetti, R. (1990). Metric regularity, tangent sets, and second-order optimality conditions. *Appl. Math. Optim.* 21, 265–287.
- Daley, D. J. and D. Vere-Jones (2008). *An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure* (2nd ed.). New York: Springer.
- Dunford, N. and J. T. Schwartz (2009). *Linear Operators: Part 1, General Theory*. New Jersey: John Wiley & Sons.
- Everitt, B. S. and A. Skrondal (2010). *The Cambridge dictionary of statistics, 4th edition*. New York: Cambridge university press.
- Fisher, R. A. (1922). On the mathematical foundation of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character.* 222, 309–368.
- Gekas, C., F. Dieterlen-Lievre, S. H. Orkin, and H. K. A. Mikkola (2005). The placenta is a niche for hematopoietic stem cells. *Developmental* 8, 365–375.
- Hall, P. (1988). *Introduction to the theory of Coverage Processes*. New York: Wiley.
- Hille, E. and R. S. Philip (1957). *Functional Analysis and Semigroups*. New York: American Mathematical Society.
- Kiefer, J. and J. Wolfowitz (1960). The equivalence of two extremum problems. *Canadian Journal of Mathematics.* 12, 363–366.

- Kumaravelu, R., L. Hook, A. M. Morrison, J. Ure, S. Zhao, S. Zuyev, J. Ansell, and A. Medvinsky (2002). Quantitative developmental anatomy of definitive haematopoietic stem cells/long-term repopulating units (hsc/rus): role of the aorta-gonad-mesonephros (agm) region and the yolk sac in colonisation of the mouse embryonic liver. *Developmental* 129, 4891–4899.
- Marin, J. and C. P. Robert (2007). *Bayesian Core: A Practical Approach to Computational Bayesian Statistics*. New York: Springer.
- Maurer, H. and J. Zowe (1979). First and second order necessary and sufficient optimality conditions for infinite-dimensional programming problems. *Math. Programming* 16, 98–110.
- Mayhall, E. A., N. Paffett-Lugassy, and L. I. Zon (2004). The clinical potential of stem cells. *Current Opinion in Cell Biology*. 16(6), 713–720.
- McCrary, M. H. (1915). The numerical interpretation of fermentation-tube results. *The Journal of Infectious Diseases* 17(1), 183–212.
- Medvinsky, A., S. Rybtsov, and S. Taoudi (2011). Embryonic origin of the adult hematopoietic system: advances and questions. *Development*. 138, 1017–1031.
- Molchanov, I. and S. Zuyev (2000a). Tangent sets in the space of measures: with applications to variational calculus. *J. Math. Anal. Appl.* 249, 539–552.
- Molchanov, I. and S. Zuyev (2000b). Variational analysis of functionals of a poisson process. *Math. Oper. Res.* 25, 485–508.
- Molchanov, I. and S. Zuyev (2002). Steepest decent algorithm in a space of measures. *Statistics and Computing* 12(2), 115–123.
- Molchanov, I. and S. Zuyev (2004). Optimization in space of measures and optimal design. *ESAIM* 8, 12–24.
- Ottersbach, K. and E. Dzierzak (2005). The murine placenta contains hematopoietic stem cells within the vascular labyrinth region. *Dev. Cell.* 8, 377–387.
- Pukelsheim, F. (1983). *Optimal Design of Experiments*. New York: Wiley.
- Ridout, M. S. (1995). Three-stage design for seed testing experiments. *Appl. Statist.* 44, 153–162.
- Robinson, S. M. (1976). First order conditions for general nonlinear optimization. *SIAM J. Appl. Math.* 30, 597–607.

Schneider, R. (1988). Random approximations of convex sets. *J. Microscopy* 151, 211–227.

Wynn, H. P. and A. A. Zhigljavsky (1994). The theory of search from a statistical viewpoint. *Test* 3, 1–45.

Zacks, S. (1977). Problems and approaches in design of experiments for estimation and testing in non-linear models. *In Multivariate Analysis 4*, 209–223.

Zhigljavsky, A. A. (1991). *Theory of Global Random Search*. Dordrecht: Kluwer Academic.

Zowe, J. and S. Kurcyusz (1979). Regularity and stability for the mathematical programming problem in banach spaces. *Appl. Math. Optim.* 5, 49–62.