Compound Optimal Designs for Percentile Estimation in Dose-Response Models with Restricted Design Intervals

Stefanie Biedermann[1], *Holger Dette*[1], *Wei Zhu*[2]

Abstract

In dose-response studies, the dose range is often restricted due to ethics concerns over drug toxicity and/or efficacy, particularly when human subjects are involved. We present locally optimal designs for the estimation of several percentiles simultaneously on restricted as well as unrestricted design intervals. Our results hold true for most of the commonly applied link functions with respect to the model under consideration.

Keywords: Dose-response model; link function; percentile estimation; compound optimal design; A-optimality.

1 Introduction

We consider the common binary response model where a subject is administered a stimulus at a certain dose level x to study the relationship between the dose level and the probability $p = p(x)$ of a response. The response Y at dose level x is modeled as a binary random variable with success probability p, i.e. $Y \sim Bin(1, p)$. In this article, we deal with the following parametrization of a two parameter binary response model,

$$p(x) = F((x-\alpha)/\beta), \qquad \vartheta = (\alpha, \beta)^T,\ \alpha \in I\!R,\ \beta \in I\!R^+, \tag{1}$$

where F denotes a known distribution function with density f. The Fisher information for the parameter ϑ of an observation at a dose level x is thus given by

$$I(z) = \frac{h^2(z)}{\beta^2} \begin{pmatrix} 1 & z \\ z & z^2 \end{pmatrix}, \qquad z = \frac{x-\alpha}{\beta}, \tag{2}$$

where $h^2(z) = f^2(z)/(F(z)(1-F(z)))$. An approximate design ξ is a probability measure with finite support on $I\!R$ such that the observations are taken at the support points of ξ with frequencies proportional to the corresponding masses. The Fisher information matrix $M(\xi)$ of a design ξ is defined as the integral of $I(z)$ over the measure ξ, i.e.

$$M(\xi) = \int_{I\!R} I(z)\, d\xi(z), \tag{3}$$

and an optimal design minimizes a real-valued function $\Phi(\xi)$ of the inverse of the Fisher information matrix, which is usually referred to as an optimality criterion.

[1]Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany
[2]State University of New York, Department of Applied Mathematics and Statistics, Stony Brook, NY 11794-3600, USA

In the framework of dose-response studies accomplished on human subjects such as clinical trials prior to the launch of new drugs, we often encounter the problem that the support points of an optimal design ξ with respect to some criterion function $\Phi(\cdot)$ lie outside a reasonable dosage range, i.e. Φ-optimal dose levels are either below zero or exceed safety levels such as the maximum tolerated dose of the drug. Practitioners are therefore in need of designs that take possibly restricted design intervals into account. In spite of an extensive amount of literature on optimal design for the binary response model on an unrestricted design space, so far there are relatively few articles concerning the topic of optimal design on restricted design spaces in this model. Extensive literature search yielded three related papers, one by Mats, Rosenberger and Flournoy (1998) where they derived the locally c- and D-optimal design for estimating the maximum tolerated dose in a Phase I clinical trial on a restricted design space, one by Haines, Perevozskaya and Rosenberger (2003) where they extend the latter approach to Bayesian c- and D-optimal designs and one by Biedermann, Dette and Zhu (2004), which deals with optimal designs with respect to a very general class of optimality criteria for the estimation of the vector of weighted parameters $(\sqrt{\lambda}\alpha, \sqrt{1-\lambda}\beta)^T$ on restricted and unrestricted design intervals.

In addition to estimating the model parameters of the underlying dose-response curve, there is also a great need to estimate other percentiles besides the median effective dose α. For example, the low percentiles are of particular interest in toxicity studies such as in virtually safe dose extrapolation studies, and the high percentiles are of interest in efficacy studies. The focus of this article is on the design situation where we try to estimate several percentiles simultaneously with different emphasis on the respective percentiles assuming that the corresponding design spaces are either unrestricted, one-side restricted or two-side restricted. The problem of optimal design for percentile estimation in dose-response experiments has first been addressed by Wu (1988) who derived designs that are optimal with respect to the estimation of one percentile at a time. This approach has been extended by several authors; see, e.g., a work of Zhu and Wong (2000) who focus on Bayesian optimal design for estimating the $ED50$ precisely, subject to the constraint that the efficiencies for estimating the other two quartiles $ED25$ and $ED75$ are not too low, or a recent work of Biedermann, Dette and Pepelyshev (2004) where model robust designs for percentile estimation in dose-response models are derived. The above authors, however, assume that the design interval comprises the entire real axis.

The organization of this article is as follows. In the first paragraph of section 2, the theoretical background is given and an appropriate optimality criterion for the problem of estimating several percentiles simultaneously is derived. We then apply some results of Biedermann, Dette and Zhu (2004) to obtain the structure of the support of the optimal designs with respect to unrestricted, one-side restricted and two-side restricted design intervals. The next paragraph will be devoted to the derivation of the optimal weights utilizing a result of Pukelsheim and Torsney (1991). In this article, we are taking the Frequentists' approach (Chernoff, 1953) and thus our designs are termed "locally optimal". In the following, we will omit the word "locally" for simplicity.

2 Compound optimal designs for estimating several percentiles simultaneously

With parametrization (1), the $100p^{th}$ percentile Q_p of the underlying quantal response curve is given by

$$Q_p = ED100p = \beta\, F^{-1}(p) + \alpha. \tag{4}$$

As the maximum likelihood estimate $\hat{Q}_p$ for the $100p^{th}$ percentile, we therefore obtain $\hat{Q}_p = \hat{\beta}\, F^{-1}(p) + \hat{\alpha}$ where $\hat{\alpha}$ and $\hat{\beta}$ denote the maximum likelihood estimators of α and β, respectively. If the goal is to design the experiment optimally for the estimation of one percentile Q_p at a time, it is thus reasonable to choose as optimality criterion to minimize the function

$$\varphi_p(\xi) = \mathrm{Var}(\hat{Q}_p) = \mathrm{Var}(\hat{\alpha}) + F^{-2}(p)\,\mathrm{Var}(\hat{\beta}) + 2F^{-1}(p)\,\mathrm{Cov}(\hat{\alpha}, \hat{\beta}), \tag{5}$$

i.e. to minimize the variance of the estimator $\hat{Q}_p$. If, in contrast, the experimenter's interest is in finding a good design for estimating several percentiles $Q_{p_1}, \ldots, Q_{p_k},\ k \geq 2$, simultaneously, a reasonable choice of optimality criterion is the compound criterion $\Phi(\xi)$ where

$$\Phi(\xi) = \sum_{i=1}^{k} \lambda_i\, \varphi_{p_i}(\xi), \qquad \sum_{i=1}^{k} \lambda_i = 1, \tag{6}$$

i.e. $\Phi(\xi)$ minimizes a weighted average of the variances of the maximum likelihood estimators for the respective percentiles where the weights $\lambda_i,\ i = 1, \ldots, k$ are chosen accordingly with respect to the emphasis on the particular percentile $Q_{p_i},\ i = 1, \ldots, k$. Since the variances of the percentile estimators can be of very different scale many authors [see, e.g., Dette (1997)] recommend the use of standardized optimality criteria. The above formulation of the criterion function (6) allows for this modification as follows. Assume that the aim is to minimize the standardized criterion

$$\tilde{\Phi}(\xi) = \sum_{i=1}^{k} \tilde{\lambda}_i \frac{\varphi_{p_i}(\xi)}{\varphi_{p_i}(\xi^*_{p_i})}$$

for a particular choice of weights $\tilde{\lambda}_i$, which add up to one, where $\xi^*_{p_i}$ denotes the optimal design for estimating the percentile Q_{p_i}. This is equivalent to minimizing (6) where the weights are given by

$$\lambda_i = \frac{\tilde{\lambda}_i}{\varphi_{p_i}(\xi^*_{p_i})} \Big/ \sum_{l=1}^{k} \frac{\tilde{\lambda}_l}{\varphi_{p_l}(\xi^*_{p_l})}, \qquad i = 1, \ldots, k$$

since the normalizing constant in the denominator of λ_i does not depend on ξ. The designs $\xi^*_{p_i},\ i = 1, \ldots, k$, are given in Wu (1988) so the expressions for the weights λ_i can easily be implemented in standard software such as Mathematica and standardized optimal designs can be calculated in the same way as their non standardized counterparts.

A design ξ minimizing $\Phi(\cdot)$ is called a compound optimal design. Following Cook and Wong (1994), each compound optimal design is at the same time a constrained optimal design in the sense of Lee (1987), i.e. the individual criterion function φ_{p_j} for some $j \in \{1, \ldots, k\}$ is minimized subject to the constraints that the other percentiles are estimated with certain precisions. Solving the compound optimal design problem therefore also gives a solution to the constrained optimal design problem described above.
In the model framework of (1)-(3), we can rewrite the criterion function $\Phi(\xi)$ in terms of the Fisher information matrix

$$\Phi(\xi) = \mathrm{tr}(C^{-1}(\xi)), \quad C^{-1}(\xi) = K^T M^{-1}(\xi) K, \quad K = \begin{pmatrix} 1 & 0 \\ c_1 & \sqrt{c_2 - c_1^2} \end{pmatrix} \tag{7}$$

where the expressions c_1 and c_2 are given by the first two moments of $F^{-1}(\cdot)$ with respect to the probability measure allocating weight λ_i to the point p_i, $i = 1, \ldots, k$, i.e.

$$c_1 = \sum_{i=1}^{k} \lambda_i \, F^{-1}(p_i), \quad c_2 = \sum_{i=1}^{k} \lambda_i \, F^{-2}(p_i).$$

A design ξ minimizing the criterion function $\Phi(\cdot)$ is therefore at the same time A-optimal for the estimation of the parameter vector $K^T \vartheta = (\alpha + c_1 \beta, \sqrt{c_2 - c_1^2}\, \beta)^T$.
In order to derive bounds on the number of support points of the Φ-optimal design ξ^* the following conditions on $h(\cdot)$ and thus the link function chosen to fit the binary response model (1) will be needed.
Condition (I): Let $g(z) = 1/h^2(z)$. Suppose that the function $g(\cdot)$ is twice differentiable on the entire real axis $\mathbb{R}$ and that the equation $g''(z) = c$ has at most two solutions for any real constant c.
Condition (II): $z \cdot h(z) \to 0$ as $z \to \pm\infty$.
Condition (I) is satisfied for most of the commonly applied link functions, such as the familiar logit and probit links as well as the asymmetrical complementary log-log and skewed logit link functions. The double exponential and double reciprocal links do not meet condition (I) due to their non-differentiability at the origin. Condition (II), in contrast, is complied with by all the above-mentioned link functions.
In the following lemma, we derive the number of support points of the Φ-optimal design ξ^* on any class of design intervals.

Lemma 1 *Assume that condition (I) is satisfied. Let the design interval $\mathcal{Z}$ be either unrestricted, one-side restricted or two-side restricted. Then the Φ-optimal design ξ^* with respect to any class for $\mathcal{Z}$ is supported on exactly two points, which are uniquely determined.*

We note that for any design space $\mathcal{Z}$, the Φ-optimal design ξ^* features exactly two points of support, thus leaving a three-dimensional minimization problem to solve. Theorem 1 summarizing the main results of this article gives further simplifications of this problem with respect to the position of the support.

Theorem 1 *Assume that conditions (I) and (II) are satisfied.*

(i) Let the design space be unrestricted, i.e. $\mathcal{Z} = \mathbb{R}$. *If* $h(\cdot)$ *is symmetric and there is interest in estimating a set of percentiles symmetric about the* $ED50$ *with* $\lambda_i = \lambda_j$ *for* $p_i = 1 - p_j$, *i.e.* $c_1 = 0$, *the* Φ*-optimal design* ξ^* *with respect to* $\mathcal{Z}$ *is symmetric about zero with equal weights.*
(ii) Assume that the design interval $\mathcal{Z}$ *is left-restricted, i.e.* $\mathcal{Z} = [A, \infty)$, *such that the lower support point of the* Φ*-optimal design* ξ^* *on the unrestricted design space is not included in* $\mathcal{Z}$. *Then the* Φ*-optimal design* ξ_A^* *with respect to the left-restricted design space* $[A, \infty)$ *has the boundary* A *as its lower support point. Analogously, for the right-restricted case* $\mathcal{Z} = (-\infty, B]$ *with the upper support point of the* Φ*-optimal design* ξ^* *on the unrestricted design space not included in* $\mathcal{Z}$, *we obtain that the upper support point of the* Φ*-optimal design* ξ_B^* *with respect to* $(-\infty, B]$ *is given by the boundary* B.
(iii) Let the design interval be two-side restricted, i.e. $\mathcal{Z} = [A, B]$ *with the upper support point of* ξ_A^* *and the lower support point of* ξ_B^* *not included in* $\mathcal{Z}$. *Then the support of the* Φ*-optimal design* $\xi_{A,B}^*$ *with respect to* $\mathcal{Z} = [A, B]$ *is given by the two ending points* A *and* B.

The proofs of Lemma 1 and Theorem 1 follow exactly the same lines as the corresponding proofs in Biedermann, Dette and Zhu (2004) for a more general class of optimality criteria and another matrix K and are therefore omitted.
From Theorem 1 it follows that in most cases, the three-dimensional minimization problem can be reduced to a one- or two-dimensional problem. In the subsequent paragraph, we derive a formula for the weights corresponding to the optimal design points, thus reducing the problem by a further dimension.
Denote the support points of the Φ-optimal design ξ^* with respect to some design interval $\mathcal{Z}$ by z_1 and z_2 where without loss of generality we assume that $z_1 < z_2$. The optimal weights ω_1 and ω_2 corresponding to z_1 and z_2 can then be derived from a result by Pukelsheim and Torsney (1991) as

$$\omega_1 = \sqrt{L_{11}}/(\sqrt{L_{11}} + \sqrt{L_{22}}) \quad \text{and} \quad \omega_2 = 1 - \omega_1 \tag{8}$$

where $L_{ii},\ i = 1, 2$ are the diagonal elements of the non-negative definite 2×2 matrix $L = VV^T$ and $V = (XX^T)^{-1}XK$ with $X^T = (\phi(z_1), \phi(z_2)) \in \mathbb{R}^{2\times 2}$ and $\phi_1(z) = h(z)/\beta$, $\phi_2(z) = h(z)\ (c_1 - z)/(\beta\sqrt{c_2 - c_1^2})$. From

$$V = \frac{\beta}{z_2 - z_1} \begin{pmatrix} \frac{z_2 - c_1}{h(z_1)} & -\frac{\sqrt{c_2 - c_1^2}}{h(z_1)} \\ -\frac{z_1 - c_1}{h(z_2)} & \frac{\sqrt{c_2 - c_1^2}}{h(z_2)} \end{pmatrix} \tag{9}$$

and (8), it then follows that the optimal weight corresponding to the lower support point z_1 is given by

$$\omega_1 = \sqrt{\frac{z_2^2 - 2z_2c_1 + c_2}{(z_2 - z_1)^2 h^2(z_1)}} \Bigg/ \left(\sqrt{\frac{z_2^2 - 2z_2c_1 + c_2}{(z_2 - z_1)^2 h^2(z_1)}} + \sqrt{\frac{z_1^2 - 2z_1c_1 + c_2}{(z_2 - z_1)^2 h^2(z_2)}} \right). \tag{10}$$

In the two-side restricted case we thus obtain the optimal design ξ^* directly by plugging the boundary values A and B of the design interval into formula (10). If the design interval is one-side restricted or unrestricted, i.e. the support points

of the optimal design ξ^* are not known in advance, plugging the weight formula (10) into the criterion function $\Phi(\xi)$ reduces the minimization problem by one dimension. With the assertions of Lemma 1, Theorem 1 and (10), the design problem can easily be implemented in standard software such as Mathematica or Matlab so that the Φ-optimal design ξ^* with respect to any design interval $\mathcal{Z}$ can be calculated.

Acknowledgements: The support of the Deutsche Forschungsgemeinschaft (SFB 475, Komplexitätsreduktion in multivariaten Datenstrukturen, Teilprojekt A2) and the U.S. National Institute of Health (NIH 3 M01 RR010710-06S1 and NIH 2 P30 AG08051-11) are gratefully acknowledged. The authors are also grateful to M. Marchlewski and S. Titoff for helpful computational assistance.

References

[1] Biedermann, S., Dette, H. and Pepelyshev, A. (2004). Some robust design strategies for percentile estimation in binary response models. preprint. http://www.ruhr-uni-bochum.de/mathematik3/preprint.htm

[2] Biedermann, S., Dette, H. and Zhu, W. (2004). Optimal designs for dose-response models with restricted design spaces. preprint. http://www.ruhr-uni-bochum.de/mathematik3/preprint.htm

[3] Chernoff, H. (1953). Locally optimal designs for estimating parameters. *Ann. Math. Statist.* **24**, 586-602.

[4] Cook, R.D. and Wong, W.K. (1994). On the equivalence of constrained and compound optimal designs. *Journal of the American Statistical Association* **89**, 687-692.

[5] Dette, H. (1997). Designing experiments with respect to standardized optimality criteria. *Journal of the Royal Statistical Society, Ser. B*, **59 (1)**, 97-110.

[6] Haines, L., Perevozskaya, I. and Rosenberger, W. (2003). Bayesian optimal designs for phase I clinical trials. *Biometrics* **59**, 591-600.

[7] Lee, C.M.S. (1987). Constrained optimal designs for regression models. *Comm. Statist. Part A - Theory and Methods* **16**, 765-783.

[8] Mats, V.A., Rosenberger, W.F. and Flournoy, N. (1997). Multiple-Objective Designs in Dose-Response Experiments. Institute of Mathematical Statistics Lecture Notes – Monograph Series: New Developments and Applications in Experimental Designs, 50-61.

[9] Pukelsheim, F. and Torsney, B. (1991). Optimal weights for experimental designs on linearly independent support points. *Annals of Statistics* **19 (3)**, 1614-1625.

[10] Wu, C.F.J. (1988). Optimal design for percentile estimation of a quantal response curve. In: Optimal Design and Analysis of Experiments, Editors: J. Dodge, V.V. Fedorov and H.P. Wynn. North Hollander, Amsterdam, 213-233.

[11] Zhu, W. and Wong, W.K. (2000). Multiple-Objective Designs in Dose-Response Experiments. *Journal of Biopharmaceutical Statistics* **10**, 1-14.