RUHR UNIVERSITÄT BOCHUM RUB

Optimal Designs for Dose Finding Studies with an Active Control

Holger Dette[1] Norbert Benda[2] Frank Bretz[3]
Katrin Kettelhake[1] Christine Kiss[1]

[1]Department of Mathematics, Ruhr-University Bochum, Germany

[2]Federal Institute for Drugs and Medical Devices

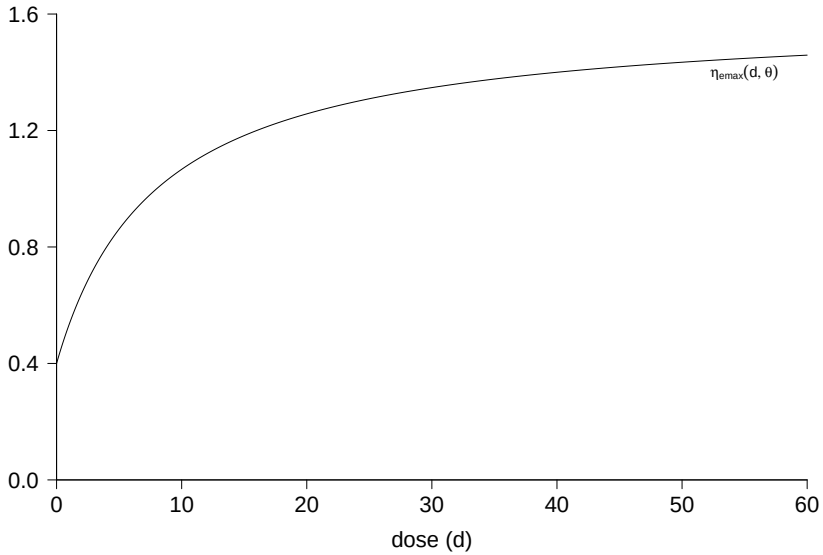
[3]Statistical Methodology, Novartis Pharma AG

Wien 2014

Outline

- Motivation
- Introduction to optimal design theory
- Dose finding studies with an active control
- Examples
- Conclusion

Motivation

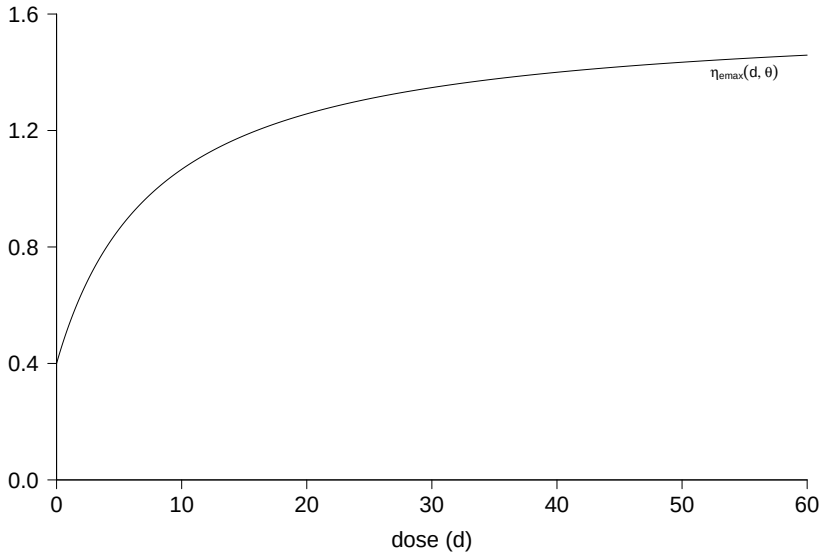


- Dose-response relationship

$$\eta(d,\theta) = E_0 + \frac{E_{max}\, d}{ED_{50}+d} = \vartheta_0 + \frac{\vartheta_1 d}{\vartheta_2+d}$$

$\rightarrow$ EMAX-model

Motivation



- Dose-response relationship

  $\eta(d, \theta) = E_0 + \frac{E_{max} d}{ED_{50}+d} = \vartheta_0 + \frac{\vartheta_1 d}{\vartheta_2+d}$

  $\rightarrow$ EMAX-model

$\rightarrow$ aim: estimate $\theta = (\vartheta_0, \vartheta_1, \vartheta_2)^t$ (or a functional of θ, such as the MED)

Mathematical background

- Dose range $\mathcal{D} \subset \mathbb{R} \rightarrow d_1, \ldots, d_k \in \mathcal{D}$
- n_i observations at dose level d_i
  - $Y_{1,1}, \ldots, Y_{1,n_1}, \ldots, Y_{k,1}, \ldots, Y_{k,n_k}$
  - $Y_{i,j} \rightarrow$ effect of the drug on patient j at dose level d_i
  - $N = \sum_{i=1}^{k} n_i$ total sample size
- Assumption:

$$Y_{i,j} = \eta(d_i, \theta) + \varepsilon_{i,j}$$

  - $\varepsilon_{i,j} \sim \mathcal{N}(0, \sigma_1^2)$
  - $\theta = (\vartheta_0, \ldots, \vartheta_s)^t$ model parameter
- Expected effect $\mathbb{E}[Y_{i,j} \mid d_i, \theta] = \eta(d_i, \theta)$

Optimal Design

Problem: Choice of the dose levels d_i and sample sizes n_i for a most efficient inference.

Approximate Designs

Probability measure with weights $\omega_1, \ldots, \omega_k \in (0, 1)$ and $\sum_{i=1}^{k} \omega_i = 1$

$$\xi = \begin{pmatrix} d_1 & \ldots & d_k \\ \omega_1 & \ldots & \omega_k \end{pmatrix} \quad d_i \in \mathcal{D}$$

Optimal Design

Problem: Choice of the dose levels d_i and sample sizes n_i for a most efficient inference.

Approximate Designs

Probability measure with weights $\omega_1, \ldots, \omega_k \in (0, 1)$ and $\sum_{i=1}^{k} \omega_i = 1$

$$\xi = \begin{pmatrix} d_1 & \ldots & d_k \\ \omega_1 & \ldots & \omega_k \end{pmatrix} \quad d_i \in \mathcal{D}$$

Example

$$\xi_{ex} = \begin{pmatrix} 15 & 30 & 45 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

- 3 dose levels
- $N =$ 9 patients $\rightarrow$ 3 observations at each dose level
- $N = 10$ patients $\rightarrow$ rounding
  $\rightarrow$ 3, 4 and 3 observations at dose levels 15, 30 and 45

Measuring the quality of estimation

Information Matrix $M(\xi, \theta)$

- $\hat{\theta}_{ML}$ maximum likelihood (ML) estimator
- $Var(\hat{\theta}_{ML}) \approx \frac{\sigma_1^2}{N} M^{-1}(\xi, \theta)$

$$M(\xi, \theta) = \sum_{j=1}^{k} \omega_j \left(\frac{\partial}{\partial \theta} \eta(d_j, \theta) \right) \left(\frac{\partial}{\partial \theta} \eta(d_j, \theta) \right)^t,$$

where $\frac{\partial}{\partial \theta} \eta(d, \theta) = (\frac{\partial}{\partial \vartheta_0} \eta(d, \theta), \ldots, \frac{\partial}{\partial \vartheta_s} \eta(d, \theta))^t$

Measuring the quality of estimation

Information Matrix $M(\xi, \theta)$

- $\hat{\theta}_{ML}$ maximum likelihood (ML) estimator
- $Var(\hat{\theta}_{ML}) \approx \frac{\sigma_1^2}{N} M^{-1}(\xi, \theta)$

$$M(\xi, \theta) = \sum_{j=1}^{k} \omega_j \left(\frac{\partial}{\partial \theta} \eta(d_j, \theta) \right) \left(\frac{\partial}{\partial \theta} \eta(d_j, \theta) \right)^t,$$

where $\frac{\partial}{\partial \theta} \eta(d, \theta) = (\frac{\partial}{\partial \vartheta_0} \eta(d, \theta), \ldots, \frac{\partial}{\partial \vartheta_s} \eta(d, \theta))^t$

Note: other estimates give different precision measures!

Information Matrix - Example

$k = 3$ dose levels and

$$\xi_{ex} = \begin{pmatrix} 15 & 30 & 45 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

- Linear-model $\eta(d, \theta) = \vartheta_0 + \vartheta_1 d$
- Gradient

$$\frac{\partial}{\partial \theta} \eta(d, \theta) = (1, d)^t$$

Information Matrix - Example

$k = 3$ dose levels and

$$\xi_{ex} = \begin{pmatrix} 15 & 30 & 45 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

- Linear-model $\eta(d, \theta) = \vartheta_0 + \vartheta_1 d$
- Gradient

$$\frac{\partial}{\partial \theta} \eta(d, \theta) = (1, d)^t$$

$$\begin{aligned} M(\xi_{ex}, \theta) &= \sum_{j=1}^{3} \omega_j \left(\frac{\partial}{\partial \theta} \eta(d_j, \theta) \right) \left(\frac{\partial}{\partial \theta} \eta(d_j, \theta) \right)^t \\ &= \frac{1}{3} \begin{pmatrix} 1 \\ 15 \end{pmatrix} (1, 15) + \frac{1}{3} \begin{pmatrix} 1 \\ 30 \end{pmatrix} (1, 30) + \frac{1}{3} \begin{pmatrix} 1 \\ 45 \end{pmatrix} (1, 45) \\ &= \begin{pmatrix} 1 & 30 \\ 30 & 1050 \end{pmatrix} \end{aligned}$$

Comparing different designs

An "optimal design" minimizes $Var(\hat{\theta}_{ML})$, that is

$$M^{-1}(\xi, \theta) \rightarrow min$$

Note: $M^{-1}(\xi, \theta)$ is a matrix.

Comparing different designs

An "optimal design" minimizes $Var(\hat{\theta}_{ML})$, that is

$$M^{-1}(\xi, \theta) \rightarrow min$$

Note: $M^{-1}(\xi, \theta)$ is a matrix.

$\rightarrow$ Solution: minimize (statistical meaningful) real valued functionals of the matrix $M^{-1}(\xi, \theta) \rightarrow$ "**optimality criteria**"

Comparing different designs

An "optimal design" minimizes $Var(\hat{\theta}_{ML})$, that is

$$M^{-1}(\xi, \theta) \rightarrow min$$

Note: $M^{-1}(\xi, \theta)$ is a matrix.

$\rightarrow$ Solution: minimize (statistical meaningful) real valued functionals of the matrix $M^{-1}(\xi, \theta) \rightarrow$ "**optimality criteria**"

Optimality criteria

$\phi(\xi, \theta) = (\det(M^{-1}(\xi, \theta)))^{\frac{1}{s+1}}$	D-optimality
$\phi(\xi, \theta) = \lambda_{max}(M^{-1}(\xi, \theta))$	E-optimality
$\phi(\xi, \theta) = c^t M^{-1}(\xi, \theta) c$	c-optimality

Locally optimal designs

Note: Information matrix depends on (unknown) parameters.

Fix $\theta_0 \in \Theta$ (guess) $\quad \rightarrow \quad M^{-1}(\xi, \theta_0) \quad \rightarrow \quad \phi(\xi, \theta_0)$

A design $\xi^*_{\theta_0}$ is called **locally** ϕ-optimal if it minimizes $\phi(\xi, \theta_0)$.

Locally optimal designs

Note: Information matrix depends on (unknown) parameters.

Fix $\theta_0 \in \Theta$ (guess) $\quad \rightarrow \quad M^{-1}(\xi, \theta_0) \quad \rightarrow \quad \phi(\xi, \theta_0)$

A design $\xi^*_{\theta_0}$ is called **locally** ϕ-optimal if it minimizes $\phi(\xi, \theta_0)$.

Note: $\xi^*_{\theta_0}$ depends on the (unknown) parameter θ_0

$\rightarrow$ "locally" optimal designs

- locally optimal designs serve as benchmark for commonly used designs
- locally optimal designs are used in more advanced design strategies (e.g. Bayesian- or adaptive designs)

ϕ-Efficiency

Comparison of the performance of a given design ξ with respect to the "best" design $\xi^*_{\theta_0}$

$$\mathit{eff}(\xi, \theta_0) = \frac{\phi(\xi^*_{\theta_0}, \theta_0)}{\phi(\xi, \theta_0)} \quad \in [1, \infty)$$

ϕ-Efficiency

Comparison of the performance of a given design ξ with respect to the "best" design $\xi^*_{\theta_0}$

$$\mathit{eff}(\xi, \theta_0) = \frac{\phi(\xi^*_{\theta_0}, \theta_0)}{\phi(\xi, \theta_0)} \quad \in [1, \infty)$$

Example

- $\mathcal{D} = [0, 60]$, EMAX-model,

  $\theta_{ex} = (0.4, 1.2, 8)^t \rightarrow \eta(d, \theta) = 0.4 + \frac{1.2d}{d+8}$

- $\xi_D^* = \begin{pmatrix} 0 & 6.32 & 60 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$ locally D-optimal design

  $\xi_{ex} = \begin{pmatrix} 15 & 30 & 45 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$

Example

- $\mathcal{D} = [0, 60]$, EMAX-model,

  $\theta_{ex} = (0.4, 1.2, 8)^t \rightarrow \eta(d, \theta) = 0.4 + \frac{1.2d}{d+8}$

- $\xi_D^* = \begin{pmatrix} 0 & 6.32 & 60 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$ locally D-optimal design

  $\xi_{ex} = \begin{pmatrix} 15 & 30 & 45 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$

$$eff(\xi_{ex}, \theta_{ex}) = \left(\frac{\det(M(\xi_D^*, \theta_{ex}))}{\det(M(\xi_{ex}, \theta_{ex}))} \right)^{\frac{1}{3}} = 22.486$$

Example

- $\mathcal{D} = [0, 60]$, EMAX-model,

  $\theta_{ex} = (0.4, 1.2, 8)^t \rightarrow \eta(d, \theta) = 0.4 + \frac{1.2d}{d+8}$

- $\xi_D^* = \begin{pmatrix} 0 & 6.32 & 60 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$ locally D-optimal design

$$\xi_{ex} = \begin{pmatrix} 15 & 30 & 45 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}, \quad \xi_{ex_2} = \begin{pmatrix} 0 & 10 & 60 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

$$eff(\xi_{ex}, \theta_{ex}) = \left(\frac{\det(M(\xi_D^*, \theta_{ex}))}{\det(M(\xi_{ex}, \theta_{ex}))} \right)^{\frac{1}{3}} = 22.486$$

$$eff(\xi_{ex_2}, \theta_{ex}) = \left(\frac{\det(M(\xi_D^*, \theta_{ex}))}{\det(M(\xi_{ex_2}, \theta_{ex}))} \right)^{\frac{1}{3}} = 1.04748$$

c-optimal designs

- A design ξ_c^* is called locally c-optimal if it minimizes

$$c^t M^-(\xi, \theta) c$$

where $c \in \mathbb{R}^{s+1}$ is a given vector.

c-optimal designs

- A design ξ_c^* is called locally c-optimal if it minimizes

$$c^t M^-(\xi, \theta) c$$

  where $c \in \mathbb{R}^{s+1}$ is a given vector.

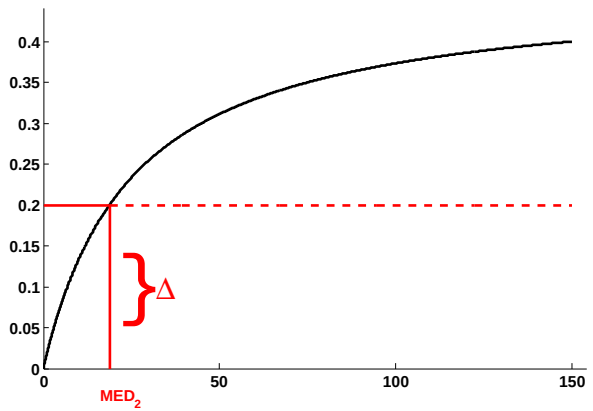
- Many other statistical problems yield to c-optimal design problems:
  - quantile estimation (MED)
  - estimation of individual parameters
  - extrapolation
  - estimation of the area under the curve
  - estimation of extrema
  - etc.

Example: Michaelis Menten model

- Michaelis Menten model

$$\eta(d, \theta) = \frac{\vartheta_1 d}{\vartheta_2 + d}$$

- Typical application in dose response studies. Estimation of the minimum effective dose.
- Example: $\vartheta_1 = 0.467$, $\vartheta_2 = 25$; $\mathcal{X} = [0\mu g, 150\mu g]$



- $\frac{\partial}{\partial \theta}\eta(d, \theta) = \left(\frac{d}{\vartheta_2 + d}, -\frac{\vartheta_1 d}{(\vartheta_2 + d)^2}\right)^t$

Optimal designs for MED-estimation

- Variance of the ML-estimate for the MED is approximately given by

$$\frac{\sigma_1^2}{N} c^t(\theta) M^-(\xi, \theta) c(\theta)$$

where the vector $c(\theta)$ is given by

$$c(\theta) = \left(-\frac{\vartheta_2}{\vartheta_1 - \Delta}, 1 \right)^t$$

Optimal designs for MED-estimation

- Variance of the ML-estimate for the MED is approximately given by

$$\frac{\sigma_1^2}{N} c^t(\theta) M^-(\xi, \theta) c(\theta)$$

  where the vector $c(\theta)$ is given by

$$c(\theta) = \left(-\frac{\vartheta_2}{\vartheta_1 - \Delta}, 1 \right)^t$$

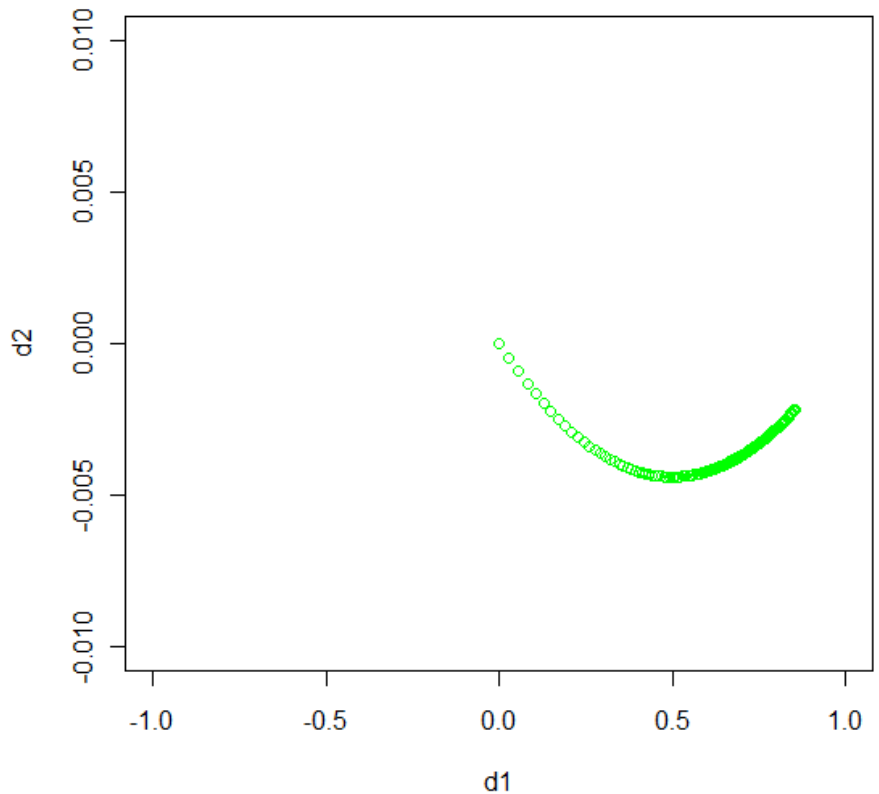
- The optimal design for MED-estimation minimizes the variance of the ML-estimate, i.e.

$$\xi^*_{MED} = \arg\min_{\xi} c^t(\theta) M^-(\xi, \theta) c(\theta)$$

- Locally c-optimal designs can be found geometrically.

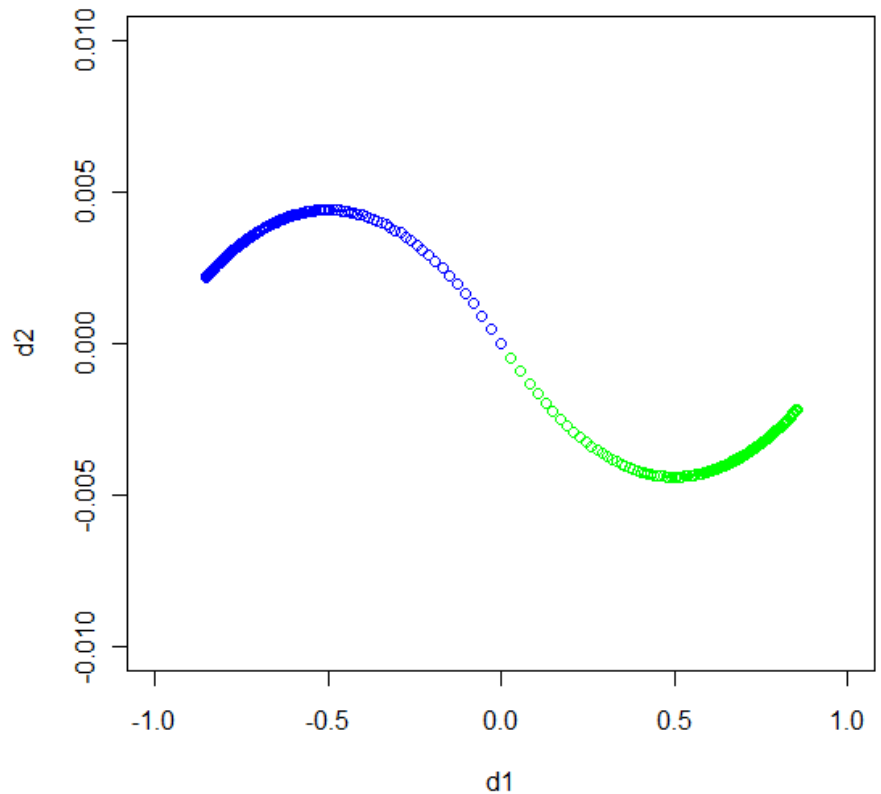
Construction of the Elfving set - Step 1

$$\left\{ \frac{\partial}{\partial \theta} \eta(d, \theta) \mid d \in \mathcal{D} \right\} = \left\{ \frac{d}{\vartheta_2 + d} \begin{pmatrix} 1 \\ -\frac{\vartheta_1}{\vartheta_2 + d} \end{pmatrix} \mid d \in \mathcal{D} \right\}$$



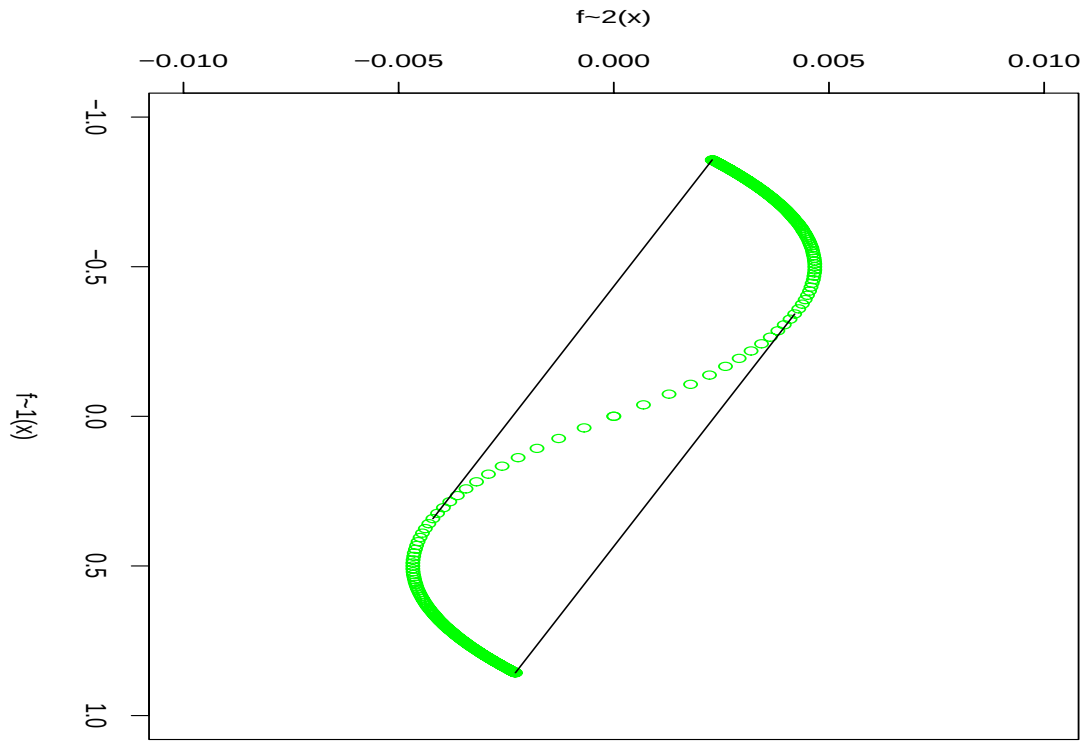
Construction of the Elfving set - Step 2

$$\left\{ \frac{d}{\vartheta_2 + d} \begin{pmatrix} 1 \\ -\frac{\vartheta_1}{\vartheta_2 + d} \end{pmatrix} \mid d \in \mathcal{D} \right\} \cup \left\{ -\frac{d}{\vartheta_2 + d} \begin{pmatrix} 1 \\ -\frac{\vartheta_1}{\vartheta_2 + d} \end{pmatrix} \mid d \in \mathcal{D} \right\}$$



Elfving set for the Michaelis Menten model

$$\mathcal{R}_1 = conv\left(\left\{\frac{d}{\vartheta_2+d}\begin{pmatrix}1\\ -\frac{\vartheta_1}{\vartheta_2+d}\end{pmatrix} \mid d \in \mathcal{D}\right\} \cup \left\{-\frac{d}{\vartheta_2+d}\begin{pmatrix}1\\ -\frac{\vartheta_1}{\vartheta_2+d}\end{pmatrix} \mid d \in \mathcal{D}\right\}\right)$$



Elfving's Theorem (1952)

- A design ξ with weights ω_i at the points d_i, $i = 1, \ldots, k$, is c-optimal if and only if there exist constants $\gamma > 0$ and $\varepsilon_1, \ldots, \varepsilon_k \in \{-1, 1\}$ such that:
  (a) The point γc can be represented as

$$\gamma c = \sum_{i=1}^{k} \omega_i \varepsilon_i \frac{\partial}{\partial \theta} \eta(d_i, \theta).$$

Elfving's Theorem (1952)

- A design ξ with weights ω_i at the points d_i, $i = 1, \ldots, k$, is c-optimal if and only if there exist constants $\gamma > 0$ and $\varepsilon_1, \ldots, \varepsilon_k \in \{-1, 1\}$ such that:
  (a) The point γc can be represented as

$$\gamma c = \sum_{i=1}^{k} \omega_i \varepsilon_i \frac{\partial}{\partial \theta} \eta(d_i, \theta).$$

  (b) The point γc is a boundary point of the **Elfving set**

$$\mathcal{R}_1 = conv\left(\left\{\frac{\partial}{\partial \theta}\eta(d, \theta) \mid d \in \mathcal{D}\right\} \cup \left\{-\frac{\partial}{\partial \theta}\eta(d, \theta) \mid d \in \mathcal{D}\right\}\right).$$

Elfving's Theorem (1952)

- A design ξ with weights ω_i at the points d_i, $i = 1, \ldots, k$, is c-optimal if and only if there exist constants $\gamma > 0$ and $\varepsilon_1, \ldots, \varepsilon_k \in \{-1, 1\}$ such that:
  (a) The point γc can be represented as

$$\gamma c = \sum_{i=1}^{k} \omega_i \varepsilon_i \frac{\partial}{\partial \theta} \eta(d_i, \theta).$$

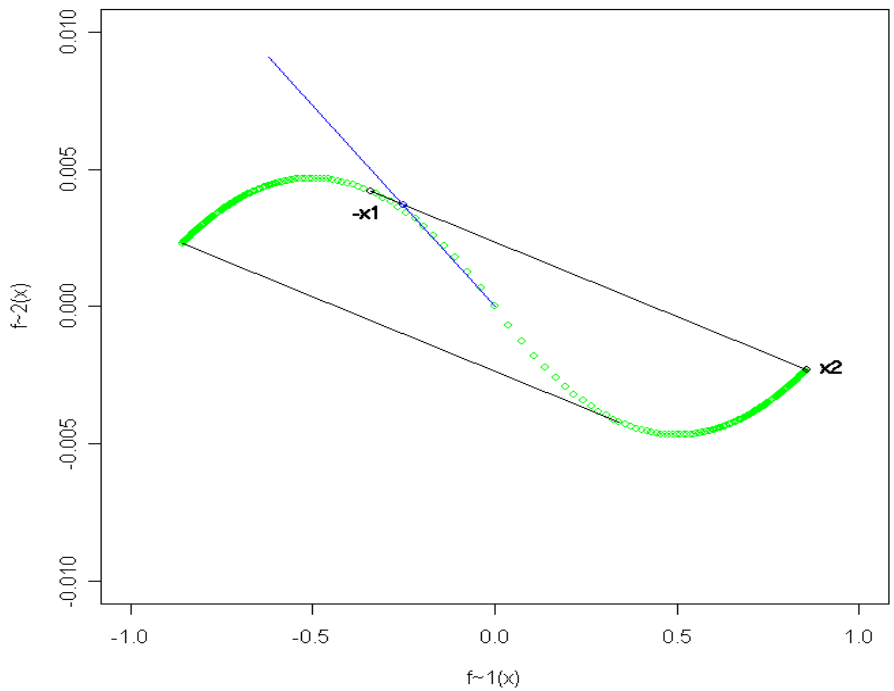
  (b) The point γc is a boundary point of the **Elfving set**

$$\mathcal{R}_1 = conv\left(\left\{\frac{\partial}{\partial \theta}\eta(d, \theta) \mid d \in \mathcal{D}\right\} \cup \left\{-\frac{\partial}{\partial \theta}\eta(d, \theta) \mid d \in \mathcal{D}\right\}\right).$$

- **Note**: c-optimal design problem reduces to a (convex) geometric problem.

MED-optimal designs for the Michaelis Menten model

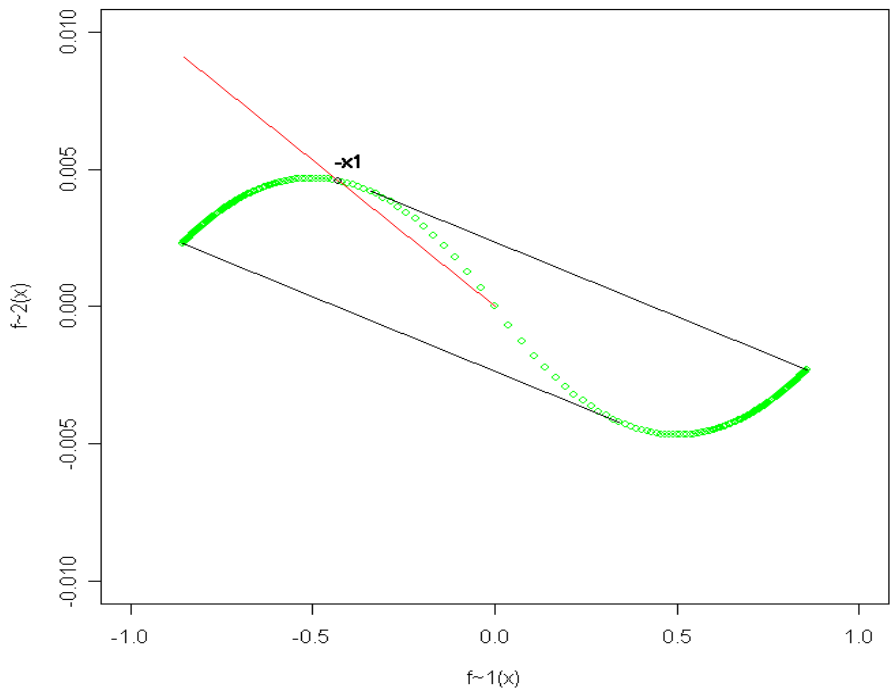
Locally MED-optimal design for $\Delta = 0.1 \Rightarrow c = (-68.12, 1)$



$\Rightarrow$ 92.5% and 7.5% of the observations at $13\mu g$ and $150\mu g$

MED-optimal designs for the Michaelis Menten model

Locally MED-optimal design for $\Delta = 0.2 \Rightarrow c = (-93.633, 1)$



$\Rightarrow$ 100% of the observations at $19\mu g$

Active-controlled dose finding studies

- $\sum_{j=1}^{k} n_j = N_1$ observations as realisations of random variables
  - $\underbrace{Y_{neu_1,1}, \ldots, Y_{neu_1,n_1}}_{d_1}, \underbrace{Y_{neu_2,1}, \ldots, Y_{neu_2,n_2}}_{d_2}, \ldots, \underbrace{Y_{neu_k,1}, \ldots, Y_{neu_k,n_k}}_{d_k}$

Active-controlled dose finding studies

- $\sum_{j=1}^{k} n_j = N_1$ observations as realisations of random variables
  - $\underbrace{Y_{neu_1,1}, \ldots, Y_{neu_1,n_1}}_{d_1}, \underbrace{Y_{neu_2,1}, \ldots, Y_{neu_2,n_2}}_{d_2}, \ldots, \underbrace{Y_{neu_k,1}, \ldots, Y_{neu_k,n_k}}_{d_k}$

    $Y_{neu_i,j} \rightarrow$ effect of the new drug on patient j at dose level d_i

Assumption: $Y_{neu_i,j} = \eta(d_i, \theta) + \varepsilon_{i,j}, \quad \varepsilon_{i,j} \sim \mathcal{N}(0, \sigma_1^2)$

Active-controlled dose finding studies

- $\sum_{j=1}^{k} n_j = N_1$ observations as realisations of random variables
  - $\underbrace{Y_{neu_1,1}, \ldots, Y_{neu_1,n_1}}_{d_1}, \underbrace{Y_{neu_2,1}, \ldots, Y_{neu_2,n_2}}_{d_2}, \ldots, \underbrace{Y_{neu_k,1}, \ldots, Y_{neu_k,n_k}}_{d_k}$

    $Y_{neu_i,j} \rightarrow$ effect of the new drug on patient j at dose level d_i

  Assumption: $Y_{neu_i,j} = \eta(d_i, \theta) + \varepsilon_{i,j}, \quad \varepsilon_{i,j} \sim \mathcal{N}(0, \sigma_1^2)$

- $N - N_1 = N_2$ observations as realisations of random variables
  - $\underbrace{Y_{ac,1}, \ldots, Y_{ac,N_2}}_{C}$

    $Y_{ac,l} \rightarrow$ effect of the active control on patient l at a fixed dose level C

  Assumption: $Y_{ac,l} = \mu + \varepsilon_l, \quad \varepsilon_l \sim \mathcal{N}(0, \sigma_2^2)$

- Model parameter $\theta = (\vartheta_0, \ldots, \vartheta_s)^t$ and μ

Active-controlled dose finding studies

- Indicator $\kappa \in \{0, 1\}$

$$\kappa \to \begin{cases} 0 & \text{new drug} \\ 1 & \text{AC} \end{cases}$$

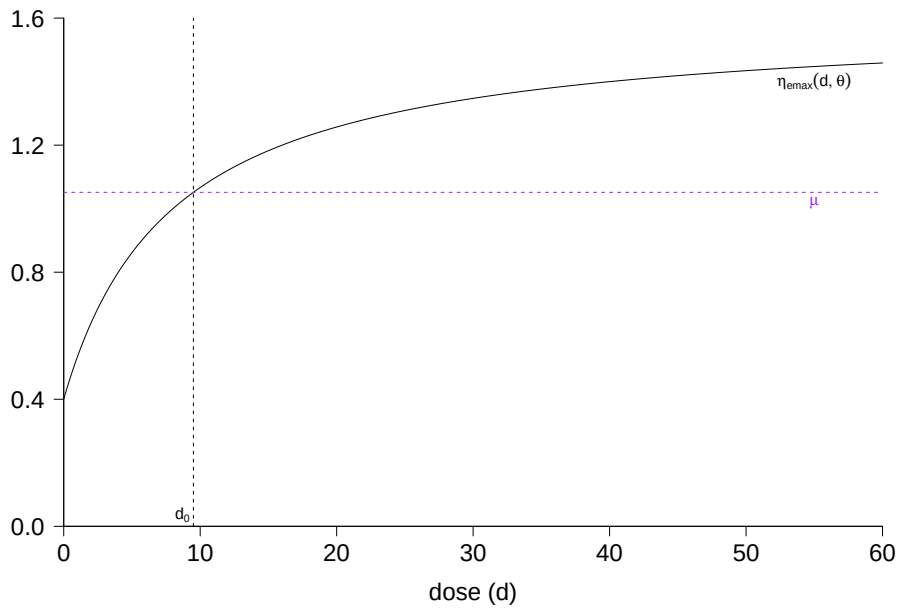
Active-controlled dose finding studies

- Indicator $\kappa \in \{0, 1\}$

$$\kappa \to \begin{cases} 0 & \text{new drug} \\ 1 & \text{AC} \end{cases}$$

- $\mathbb{E}[Y_{neu_i,j} \mid d_i, \theta] = \eta(d_i, \theta)$
  $\to$ expected effect of the new drug at dose level d_i

- $\mathbb{E}[Y_{ac,l} \mid C, \mu] = \mu$
  $\to$ expected effect of the active control



Active-controlled dose finding studies

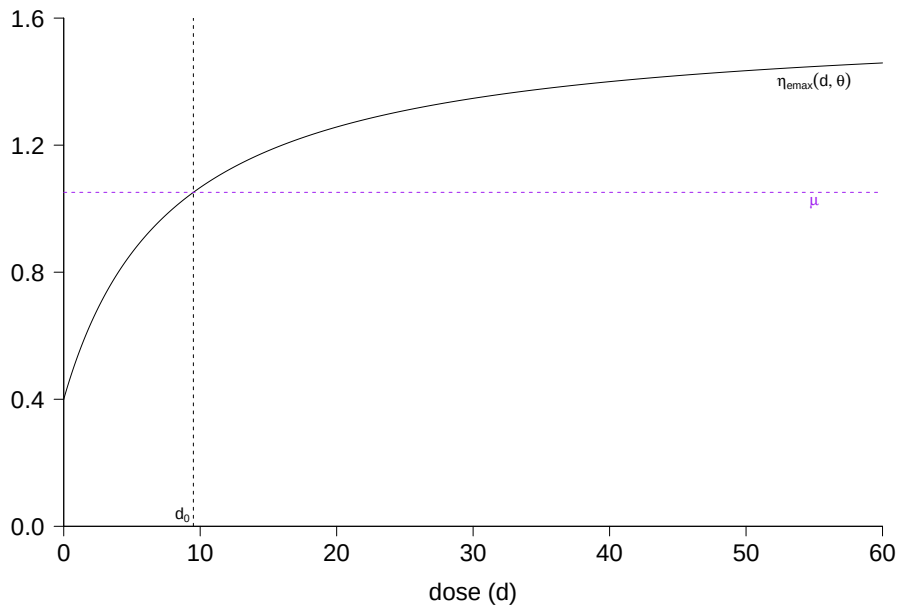
- Indicator $\kappa \in \{0, 1\}$

$$\kappa \rightarrow \begin{cases} 0 & \text{new drug} \\ 1 & \text{AC} \end{cases}$$

- $\mathbb{E}[Y_{neu_i,j} \mid d_i, \theta] = \eta(d_i, \theta)$
  $\rightarrow$ expected effect of the new drug at dose level d_i

- $\mathbb{E}[Y_{ac,l} \mid C, \mu] = \mu$
  $\rightarrow$ expected effect of the active control

- $d_0 \rightarrow$ dose level of the new drug providing the same effect as the active control



Active-controlled dose finding studies

- In an active controlled dose finding study one has to specify (d, κ)

$$\kappa = \begin{cases} 0 & \text{(new drug)}, \quad d \text{ dose level} \\ 1 & \text{(active control)}, \quad d = C \end{cases}$$

- Approximate design

$$\xi = \begin{pmatrix} (d_1, 0) & \dots & (d_k, 0) & (C, 1) \\ \omega_1 & \dots & \omega_k & \omega_{k+1} \end{pmatrix}$$

.

Active-controlled dose finding studies

- In an active controlled dose finding study one has to specify (d, κ)

$$\kappa = \begin{cases} 0 & \text{(new drug)}, \quad d \text{ dose level} \\ 1 & \text{(active control)}, \quad d = C \end{cases}$$

- Approximate design

$$\xi = \begin{pmatrix} (d_1, 0) & \dots & (d_k, 0) & (C, 1) \\ \omega_1 & \dots & \omega_k & \omega_{k+1} \end{pmatrix}$$

- Approximate design for the new drug

$$\tilde{\xi} = \begin{pmatrix} d_1 & \dots & d_k \\ \tilde{\omega}_1 & \dots & \tilde{\omega}_k \end{pmatrix},$$

where $\tilde{\omega}_i = \frac{\omega_i}{1-\omega_{k+1}}$.

Active-controlled dose finding studies - Example

Consider

$$\xi = \begin{pmatrix} (0,0) & (40,0) & (100,0) & (C,1) \\ \frac{1}{4} & \frac{1}{8} & \frac{1}{4} & \frac{3}{8} \end{pmatrix}.$$

Then

$$\tilde{\xi} = \begin{pmatrix} 0 & 40 & 100 \\ \frac{2}{5} & \frac{1}{5} & \frac{2}{5} \end{pmatrix}.$$

Active-controlled dose finding studies

- Information matrix

$$M(\xi,\theta) = \frac{1}{\sigma_1^2}\begin{pmatrix}(1-\omega_{k+1})\widetilde{M}(\tilde{\xi},\theta) & 0\\ 0 & r^2\omega_{k+1}\end{pmatrix}$$

where $r^2 = \frac{\sigma_1^2}{\sigma_2^2}$ and

$$\widetilde{M}(\tilde{\xi},\theta) = \sum_{j=1}^{k}\tilde{\omega}_j\left(\frac{\partial}{\partial\theta}\eta(d_j,\theta)\right)\left(\frac{\partial}{\partial\theta}\eta(d_j,\theta)\right)^t$$

(Information matrix of the design $\tilde{\xi}$ for the new drug).

Note:

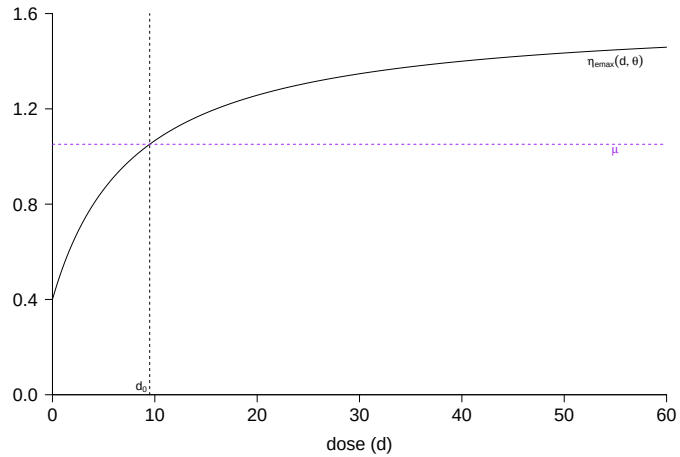
- $\mathrm{Var}(\hat{\theta},\hat{\mu}) \approx \frac{\sigma_1^2}{N}M^{-1}(\xi,\theta)$
- Block structure of the Information matrix $M(\xi,\theta)$

AC-optimality

Assumption (model)

$$\eta(d, \theta) = \vartheta_0 + \vartheta_1 \eta_{\theta_2}(d)$$

η strictly increasing, e.g.

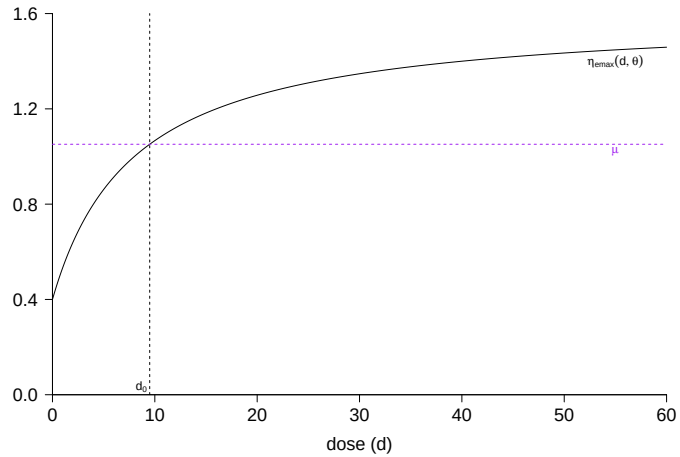


AC-optimality

Assumption (model)

$$\eta(d, \theta) = \vartheta_0 + \vartheta_1 \eta_{\theta_2}(d)$$

η strictly increasing, e.g.

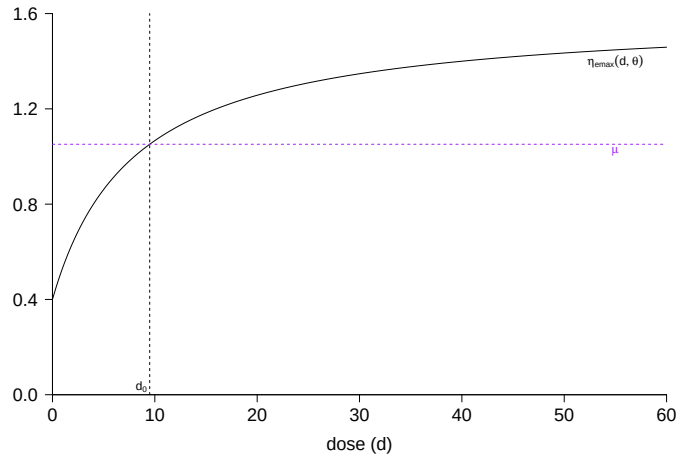


AC-optimality

Assumption (model)

$$\eta(d, \theta) = \vartheta_0 + \vartheta_1 \eta_{\theta_2}(d)$$

η strictly increasing, e.g.



Note: dose level of the new drug providing the same effect as the active control

$$d_0 = \eta^{-1}(\mu, \theta)$$

Estimate

$$\hat{d}_0 = d_0(\hat{\mu}, \hat{\theta}) = \eta^{-1}(\hat{\mu}, \hat{\theta})$$

AC-optimality

$$Var(\hat{d}_0) \approx \frac{\sigma_1^2}{N} c^t M^-(\xi, \theta) c,$$

where $c = \nabla d_0(\mu, \theta)$ the gradient of the function d_0 with respect to (θ, μ)

Note: The asymptotic variances depend on the design ξ and the parameters θ and μ!

AC-optimality

$$Var(\hat{d}_0) \approx \frac{\sigma_1^2}{N} c^t M^-(\xi, \theta) c,$$

where $c = \nabla d_0(\mu, \theta)$ the gradient of the function d_0 with respect to (θ, μ)

Note: The asymptotic variances depend on the design ξ and the parameters θ and μ!

For fixed θ_0 and μ_0 a design $\xi^*_{AC_0}$ is called **locally** AC-optimal if it minimizes

$$c_0^t M^-(\xi, \theta_0) c_0,$$

where $c_0 = \nabla d_0(\mu_0, \theta_0)$.

Note: $\xi^*_{AC_0}$ minimizes the (asymptotic) variance of the MED-estimation!

Locally AC-optimal Design I

The locally AC-optimal design $\xi^*_{AC_0}$ is given by

$$\begin{pmatrix} (d_0, 0) & (C, 1) \\ \frac{r}{r+1} & \frac{1}{r+1} \end{pmatrix},$$

where $r = \frac{\sigma_1}{\sigma_2}$.

- **Note**: The optimal design for the new drug allocates observations at d_0 (intuitively obvious).
- The proof is not obvious ($\rightarrow$ implicit function theorem + Elfving theorem).
- **However**: This observation is a consequence of the assumption of a normal distribution.

Locally AC-optimal Design II

- Locally optimal designs are often sensitive with respect to the misspecification of the initial parameters.

Locally AC-optimal Design II

- Locally optimal designs are often sensitive with respect to the misspecification of the initial parameters.

- Alternative design strategies have been developed:
  $\rightarrow$ Bayesian optimal designs
  $\rightarrow$ Minimax optimal designs
  $\rightarrow$ Adaptive designs

Standardized Bayesian AC-optimal designs

Recall

- $\frac{\sigma_1^2}{N} c^t M^-(\xi, \theta) c \approx Var(\hat{d}_0)$
- Efficiency

$$eff(\xi, \theta_0, \mu_0) = \frac{c_0^t M^-(\xi, \theta_0) c_0}{c_0^t M^-(\xi^*_{AC_0}, \theta_0) c_0} \in [1, \infty)$$

Standardized Bayesian AC-optimal designs

Recall

- $\frac{\sigma_1^2}{N} c^t M^-(\xi, \theta) c \approx Var(\hat{d}_0)$
- Efficiency

$$eff(\xi, \theta_0, \mu_0) = \frac{c_0^t M^-(\xi, \theta_0) c_0}{c_0^t M^-(\xi^*_{AC_0}, \theta_0) c_0} \in [1, \infty)$$

A design ξ_B is called standardized Bayesian AC-optimal if it minimizes

$$\int eff(\xi, \theta, \mu) d\pi(\theta, \mu),$$

where π is a prior distribution for (θ, μ).

Standardized Bayesian AC-optimal designs

- (Reparameterized) EMAX-model

$$\eta(d,\theta) = \vartheta_0 + \frac{\vartheta_1 d}{1+\vartheta_2 d}, \quad d \in [L,R]$$

Theorem

$$\xi^*_{B_{AC}} = \begin{pmatrix} (L,0) & (\frac{L+R+2\vartheta_2 LR}{2+\vartheta_2(L+R)},0) & (R,0) & (C,1) \\ p_B \frac{\sqrt{\rho_B}}{1+\sqrt{\rho_B}} & (1-2p_B)\frac{\sqrt{\rho_B}}{1+\sqrt{\rho_B}} & p_B \frac{\sqrt{\rho_B}}{1+\sqrt{\rho_B}} & \frac{1}{1+\sqrt{\rho_B}} \end{pmatrix}$$

is the standardized Bayesian AC-optimal design, where

$$\begin{aligned} p_B &= \frac{\sqrt{\mu_2+\mu_4}}{2(\sqrt{1+2\mu_2+\mu_4}+\sqrt{\mu_2+\mu_4})} \\ \rho_B &= r^2(\sqrt{1+2\mu_2+\mu_4}+\sqrt{\mu_2+\mu_4})^2 \end{aligned}$$

and μ_2 and μ_4 moments of the prior, (ϑ_2 is fixed).

Examples - locally AC-optimal designs

- EMAX-model $\rightarrow \eta(d, \theta) = \vartheta_0 + \frac{\vartheta_1 d}{1+\vartheta_2 d}$ (reparameterized)
- $\mathcal{D} = [10, 150]$, $\theta_0 = (2.5, 1.125, 0.025)^t$ and $\mu_0 = 22.5$
- $r = 1$ that is $\sigma_1^2 = \sigma_2^2$

Examples - locally AC-optimal designs

- EMAX-model $\rightarrow \eta(d, \theta) = \vartheta_0 + \frac{\vartheta_1 d}{1+\vartheta_2 d}$ (reparameterized)
- $\mathcal{D} = [10, 150]$, $\theta_0 = (2.5, 1.125, 0.025)^t$ and $\mu_0 = 22.5$
- $r = 1$ that is $\sigma_1^2 = \sigma_2^2$
- locally AC-optimal design $\xi^*_{AC_0}$

$$\xi^*_{AC_0} = \begin{pmatrix} (32, 0) & (C, 1) \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Examples - Bayesian AC-optimal designs

- $\vartheta_0 \in [1, 2]$, $\vartheta_1 \in [0.92, 1.38]$, $\vartheta_2 = 0.025$, $\mu \in [20, 23]$
- $\vartheta_0 \in [1, 2]$, $\vartheta_1 \in [0.92, 1.38]$, $\vartheta_2 \in [0.016, 0.025]$, $\mu \in [20, 23]$
- π uniform distribution
- Bayesian AC-optimal design $\xi^*_{B_1}$ and $\xi^*_{B_2}$

$$\xi^*_{B_1} = \begin{pmatrix} (10, 0) & (34, 0) & (150, 0) & (C, 1) \\ 0.07 & 0.44 & 0.03 & 0.46 \end{pmatrix}$$

$$\xi^*_{B_2} = \begin{pmatrix} (10, 0) & (31.9, 0) & (150, 0) & (C, 1) \\ 0.09 & 0.42 & 0.02 & 0.47 \end{pmatrix}$$

Examples - Efficiencies

Standard designs (from Novartis)

$$\xi_{S_1} = \begin{pmatrix} (10,0) & (45,0) & (80,0) & (115,0) & (150,0) & (C,1) \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$$

$$\xi_{S_2} = \begin{pmatrix} (10,0) & (20,0) & (39,0) & (76,0) & (150,0) & (C,1) \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$$

- Efficiencies for estimating the target dose $d_0 = 32$

	ξ_{S_1}	ξ_{S_2}		$\xi^*_{B_1}$	$\xi^*_{B_2}$
$eff(\xi, \theta, \mu)$	2.94	2.40		1.58	1.66

Example - Logistic model

- $\eta_{log}(d, \theta) = \vartheta_0 + \vartheta_1(1 + \exp(\frac{\vartheta_2 - d}{\vartheta_3}))^{-1}$, $d \in [10, 150]$
- $\vartheta_0 \in [1, 4]$, $\vartheta_1 \in [32, 37]$, $\vartheta_2 \in [45, 55]$, $\vartheta_3 \in [9, 11]$ and $\mu \in [20, 25]$
- π uniform distribution

- Bayesian AC-optimal designs $\xi^*_{B_3}$

$$\xi^*_{B_3} = \begin{pmatrix} (10, 0) & (46.2, 0) & (58, 0) & (150, 0) & (C, 1) \\ 0.02 & 0.25 & 0.29 & 0.01 & 0.43 \end{pmatrix}$$

Example - Logistic model

- $\eta_{log}(d, \theta) = \vartheta_0 + \vartheta_1(1 + \exp(\frac{\vartheta_2 - d}{\vartheta_3}))^{-1}$, $d \in [10, 150]$
- $\vartheta_0 \in [1, 4]$, $\vartheta_1 \in [32, 37]$, $\vartheta_2 \in [45, 55]$, $\vartheta_3 \in [9, 11]$ and $\mu \in [20, 25]$
- π uniform distribution

- Bayesian AC-optimal designs $\xi^*_{B_3}$

$$\xi^*_{B_3} = \begin{pmatrix} (10,0) & (46.2,0) & (58,0) & (150,0) & (C,1) \\ 0.02 & 0.25 & 0.29 & 0.01 & 0.43 \end{pmatrix}$$

- Efficiencies for estimating the target dose $d_0 = 32$

	ξ_{S_1}	ξ_{S_2}		$\xi^*_{B_3}$
$eff(\xi, \theta, \mu)$	5.81	7.53		1.70

Some conclusions

- Optimal designs improve accuracy of statistical inference.
  - Estimation of the model parameters.
  - Estimation of the MED.
  - Estimation of functionals of the parameter (AUC).
- Locally optimal designs can be used as a benchmark for commonly used designs.
- Locally optimal designs depend on
  - parameters of the model.
  - model under consideration.
- Robustification is possible (as indicated here for the parameters).

Future Research

- Other optimality criteria.

Future Research

- Other optimality criteria.
- Other distributional assumptions (exponential family).
  - Discrete data.
  - Block structure of the Information matrix remains.
  - Designs have a different structure.

References

- H. Dette, C. Kiss, N. Benda, F. Bretz (2014). *Optimal designs for dose finding studies with an active control*. Journal of the Royal Statistical Society Ser. B, 76, 265-295.

- H. Dette, C. Kiss, M. Bevanda, F. Bretz (2010). *Optimal designs for the emax, log-linear and exponential models*. Biomtrika, 97, 2, 513-518.

- H. Dette, (1997). *Designing Experiments with Respect to 'Standardized' Optimality Criteria*. Journal of the Royal Statistical Societ Ser. B, 59, 97-110.

- F. Pukelsheim, (2006). *Optimal Design of Experiments*. Society for Industrial and Applied Mathematics.