Mathematics is the art of giving the same name to different things.

Jules Henri Poincarré

Table of Contents

Model and inference (likelihood)

Using likelihood and letting

$$\Pr(Y_i = 1 | X_i = d_j) = (\alpha_j)^a$$

then the models

α_i	.81	.85	.89	.92	.95	.98
α_i	.01	.03	.09	.16	.35	.59

behave identically,

whereas the models

α_i	.05	.10	.20	.30	.50	.70
α_i	.05	.10	.20	.30	.40	.50

behave differently.

Model and inference (Bayes)

For distance measure use;

1. O'Quigley, Pepe, Fisher (1990) suggest $E\,\psi(d_j, a)$
2. O'Quigley, Pepe, Fisher (1990) suggest $\psi(d_j, E(a))$
3. Chu, Lin, Shih (2009) suggest $\psi_{1-\gamma}(d_j, a)$
4. Shih (1999) suggest $\gamma = 0.5$ corresponding to median.
5. Babb, Rogatko, Zacks (1998) suggest $\gamma = 0.75$
   This is known as EWOC.

Table of Contents

1-parameter versus 2-parameter models

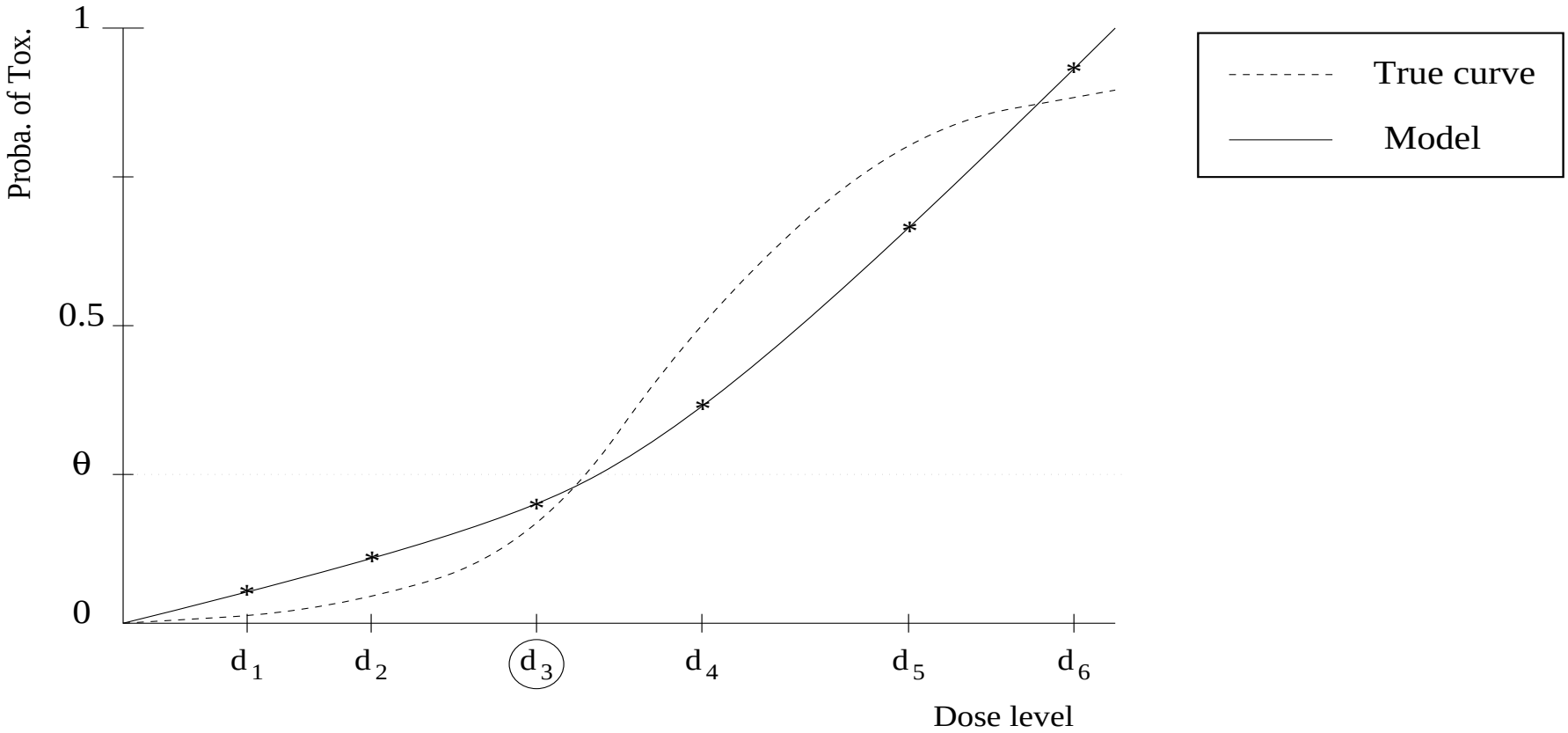
O'Quigley, Pepe and Fisher (1990) show that;

1. 2-parameter logistic model more noisy
2. Final recommendations less accurate

Table: 2-param logistic (O'Quigley, Pepe, Fisher 1990)

	Dose					
	1	2	3	4	5	6
$R(d_i)$	.06	.08	.12	.18	.40	.71
1-CRM	.00	.04	.23	.57	.15	.00
2-CRM	.01	.11	.16	.48	.19	.05

One parameter CRM models



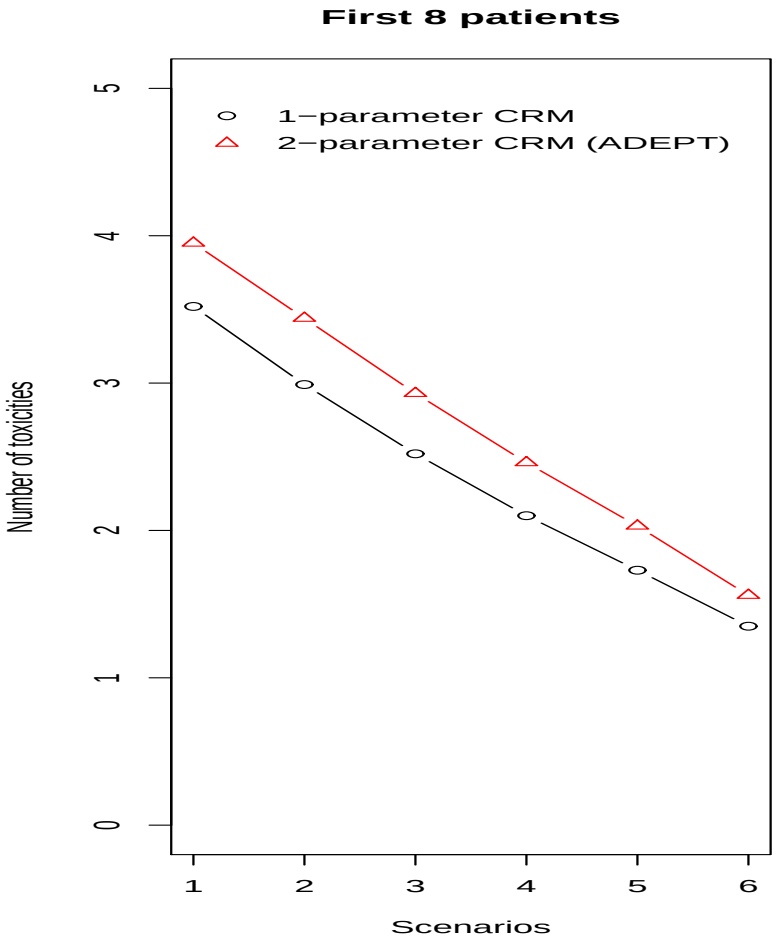
Two parameter CRM (ADEPT, BLR)

- 2-CRM has weaker theoretical foundation
- 2-CRM can be erratic, eg., first patient treated at level 1, suffers DLT, the recommendation is treat at level 6 (Shu 2008).
- 2-CRM gets stuck, eg, 1 our of 2 DLT at level i. More than one million non-DLTs needed in order to return to level i (Cheung 2008).
- ADEPT is 2-CRM, using patient benefit as metric.
- BLR (Neuenschwander et al 2007) is also 2-CRM

Two versus one parameter CRM models

1. $\hat{R}_j = \psi(d_j, \hat{a})$ may be too inflexible to work well for all j.
2. $\hat{R}_j \approx \sum Y_{ij}/n_j$ at recommended level.
3. $\hat{R}_j \xrightarrow{\mathrm{P}} R_j$ and is fully efficient (Shen & O'Quigley, *Biometrika 96*)
4. $R_j = \psi(d_j, a, b)$ is over-parameterized, cannot identify a and b.

One/two parameter CRM models: first patients



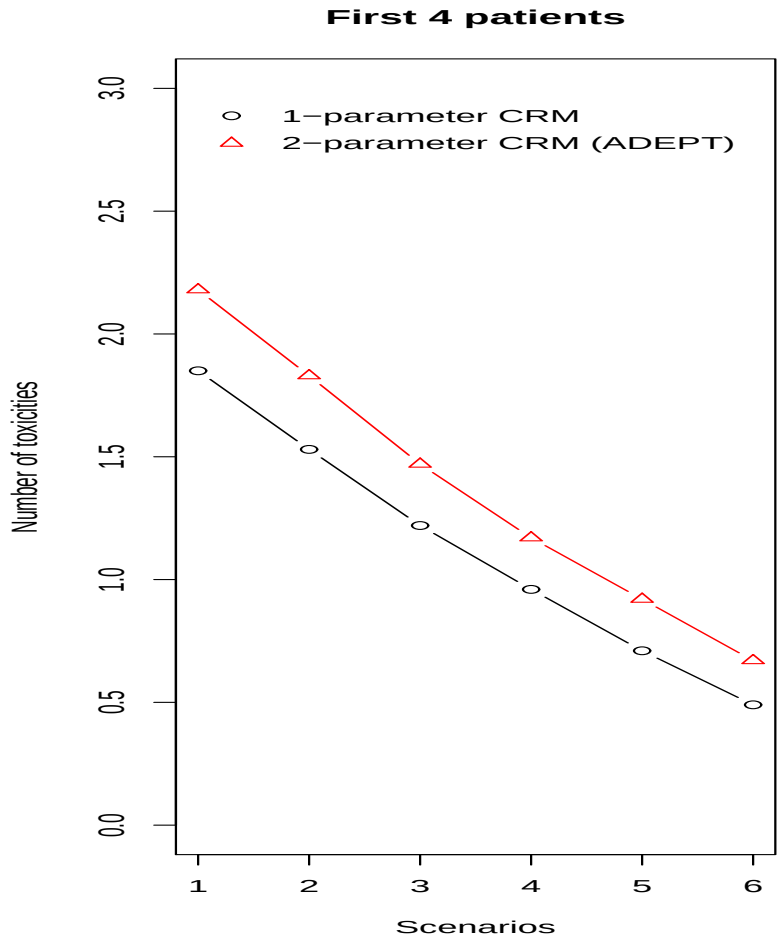


Table of Contents

Curve free designs (Gasparini, Eisele 2000)

1. $\theta_1 = 1 - R(d_1)\,,$

$$\theta_i = \frac{1 - R(d_i)}{1 - R(d_{i-1})}\,,\ i = 2, \ldots, k.$$

2. For each $\theta_i\,, (i = 1, \ldots, k)$,

$$f(\theta_i) = B^{-1}(a_i, b_i)\theta_i^{a_i - 1}(1 - \theta_i)^{b_i - 1}$$

for parameters a_i and b_i and where $B(a, b)$ is the beta function. with parameters a and b.

3. $R(d_i) = 1 - \theta_1\theta_2 \ldots \theta_i$

4. Curve free designs are rigid (Cheung 2003)

Simple take home message

2-param CRM / EWOC / BLRM are all equivalent

2-param CRM / EWOC / BLRM are all rigid

2-param CRM / EWOC / BLRM are not proven coherent

2-param CRM / EWOC / BLRM are all inefficient

1-param CRM $\equiv$ curve free under restrictions on curve free

Table of Contents

Novartis case study (Neuenschwander et al, Bailey and Neuenschwander 2008

doses	1.0	2.5	5	10	15	20	25	30	40	50
# pats	3	4	5	4	0	0	2	-	-	-
# DLTs	0	0	0	0	-	-	2	-	-	-

- 1-param CRM produces unintuitive dose recommendations (Jaki, T. and Hampson, L., Schmidl, Vlachos,)
- 2-param CRM / BLRM produces good recommendations (Jaki, T. and Hampson, L.)
- Ananysis of above incorrect (Iasonos et al 2016)

Logical errors in Novartis case study

doses	1.0	2.5	5	10	15	20	25	30	40	50
prior	.17	.02	.01	.01	.01	.02	.04	.07	.09	**.56**
post	.00	.00	.00	.01	.02	.03	.10	.28	**.37**	.19

- CRM recommends **decrease** and **not** an increase.
- CRM is coherent (Cheung 2003).
- Bayesian methods require care.
- MCMC = My Clueless Magical Calculator. Use caution!

Table of Contents

How good can any design be?

Super-optimal designs:

1. Include zero patients in study: recommend level 2.
2. Include 5 patients at level 3. Recommend according to table:

Outcome	0/5	1/5	2/5	3/5	4/5	5/5
Recommendation	5	4	3	2	1	1

3. Above designs are competitive/very competitive depending on scenarios

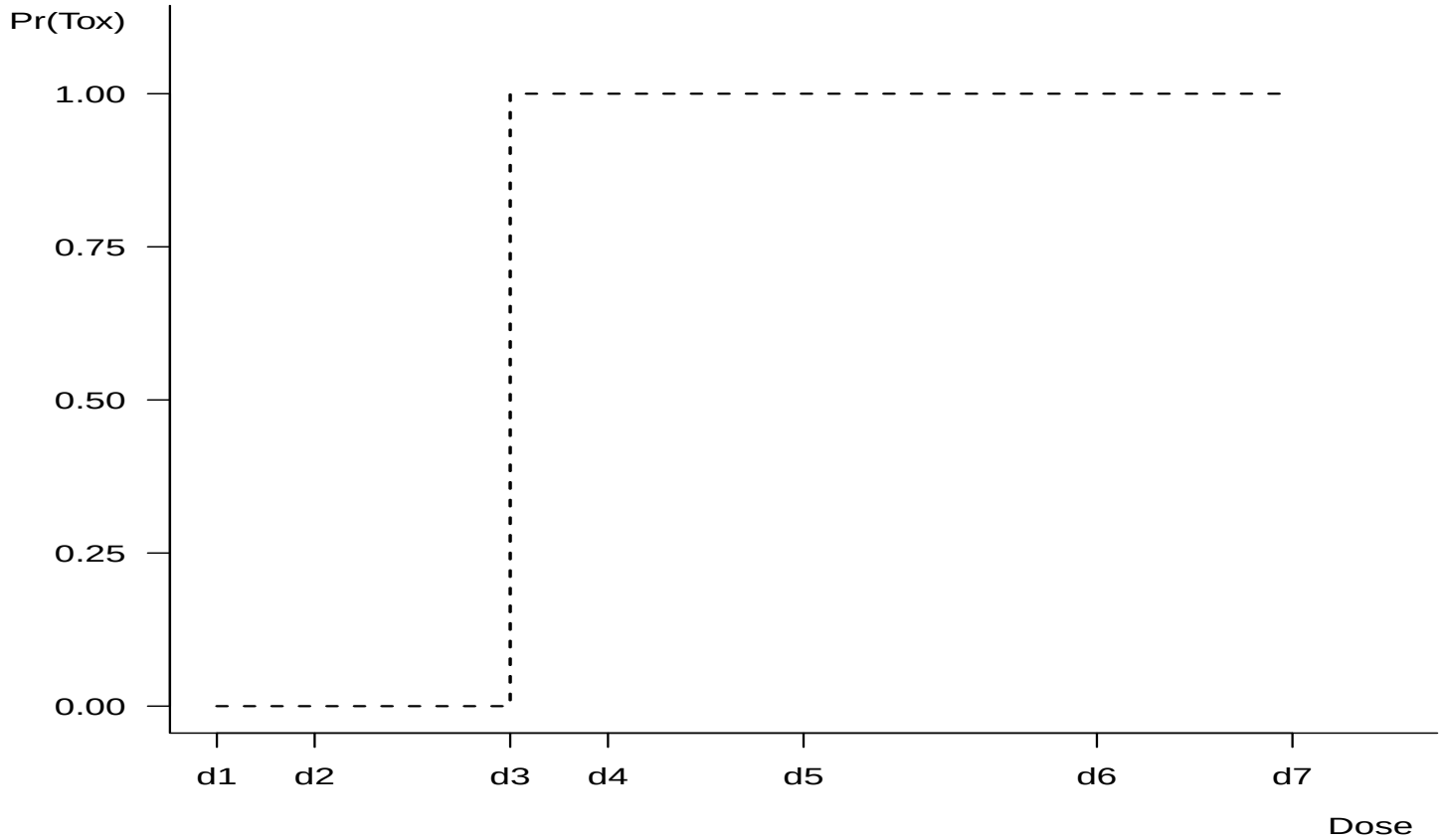
Super optimality

Super-optimality is common in the statistical literature, in particular for Bayesian designs.

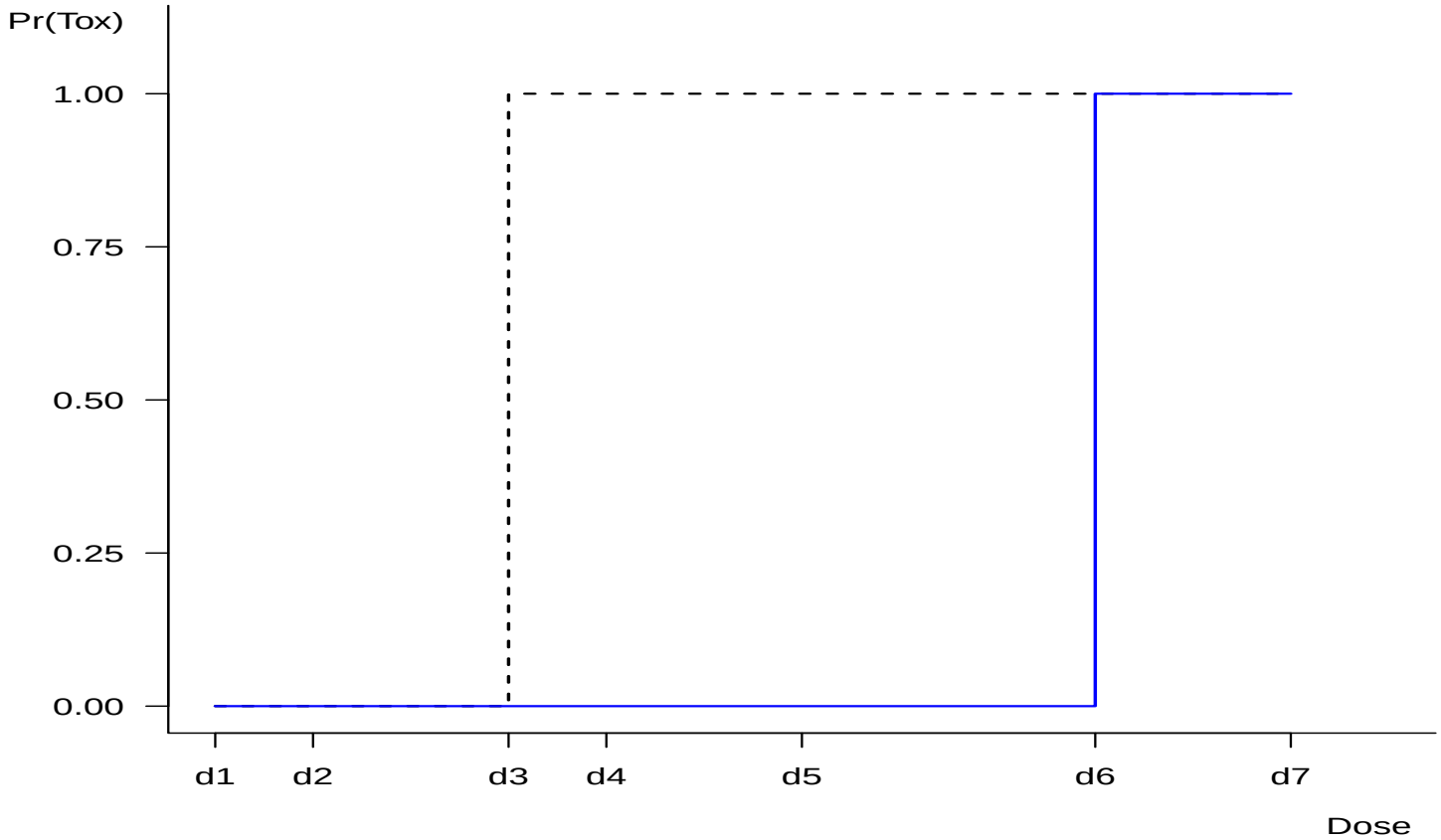
Example for combinations, using partial orderings;

1. Yin and Yuan (2009) *Appl. Statist*, 211 - 224, show for 4×4 combinations, copula design finds MTD 52%.
2. PO-CRM (Wages et al, *Biometrics* 2011) finds MTD in 45%.
3. When ordering is known, CRM finds MTD 48%.
4. When ordering is known Optimal Design finds MTD 49%.

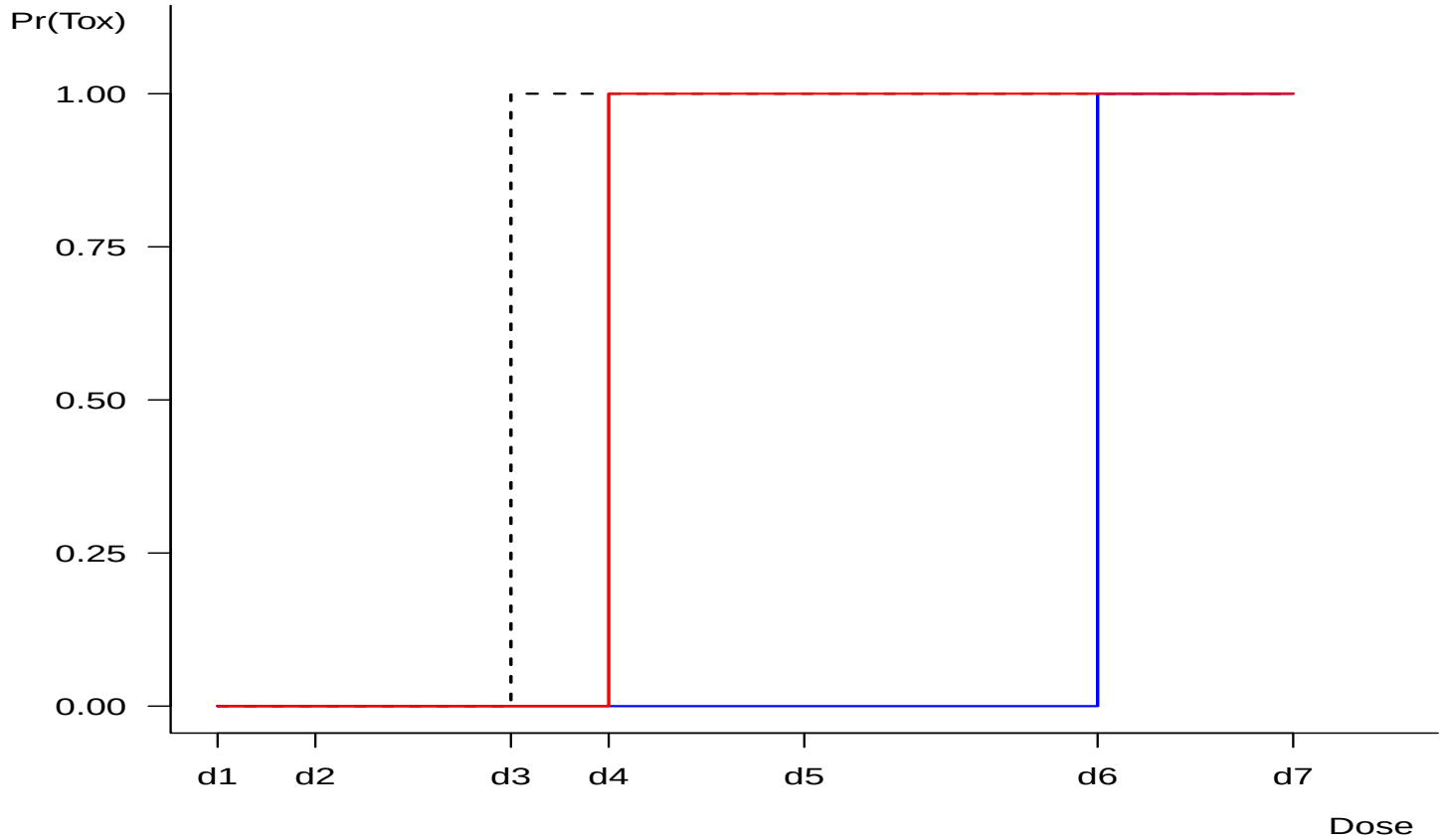
Recall basic model for cytotoxics (3 patients)



Recall basic model for cytotoxics (3 patients)



Recall basic model for cytotoxics (3 patients)



Optimal design benchmark

- Subject h experiences a toxicity at d_5.
- Subject j a non-toxicity at level d_3.

Doses	d_1	d_2	d_3	d_4	d_5	d_6
Observed Y_{hk}	X	X	X	X	1	1
Unobserved Y_{hk}	0	0	1	1	1	1
Observed Y_{jk}	0	0	0	X	X	X
Unobserved Y_{jk}	0	0	0	0	0	1

Consider;

Dose	d_1	d_2	d_3	d_4	d_5	d_6
$R_k = \Pr(Y_k = 1)$	0.05	0.11	0.22	0.35	0.45	0.60

Subject j	v_j	s_j	Toxicity at dose level					
			1	2	3	4	5	6
1	.53	6	0	0	0	0	0	1
2	.08	2	0	1	1	1	1	1
3	.29	4	0	0	0	1	1	1
4	.41	5	0	0	0	0	1	1
5	.79	-	0	0	0	0	0	0
6	.04	1	1	1	1	1	1	1
7	.87	-	0	0	0	0	0	0
8	.15	3	0	0	1	1	1	1
9	.63	-	0	0	0	0	0	0
10	.56	6	0	0	0	0	0	1
11	.32	4	0	0	0	1	1	1
12	.72	-	0	0	0	0	0	0
13	.20	3	0	0	1	1	1	1
14	.97	-	0	0	0	0	0	0
15	.52	6	0	0	0	0	0	1
16	.24	4	0	0	0	1	1	1
Frequencies		$\hat{R}_k$	0.06	0.13	0.25	0.44	0.50	0.69
		R_k	0.05	0.11	0.22	0.35	0.45	0.60

Summarizing results

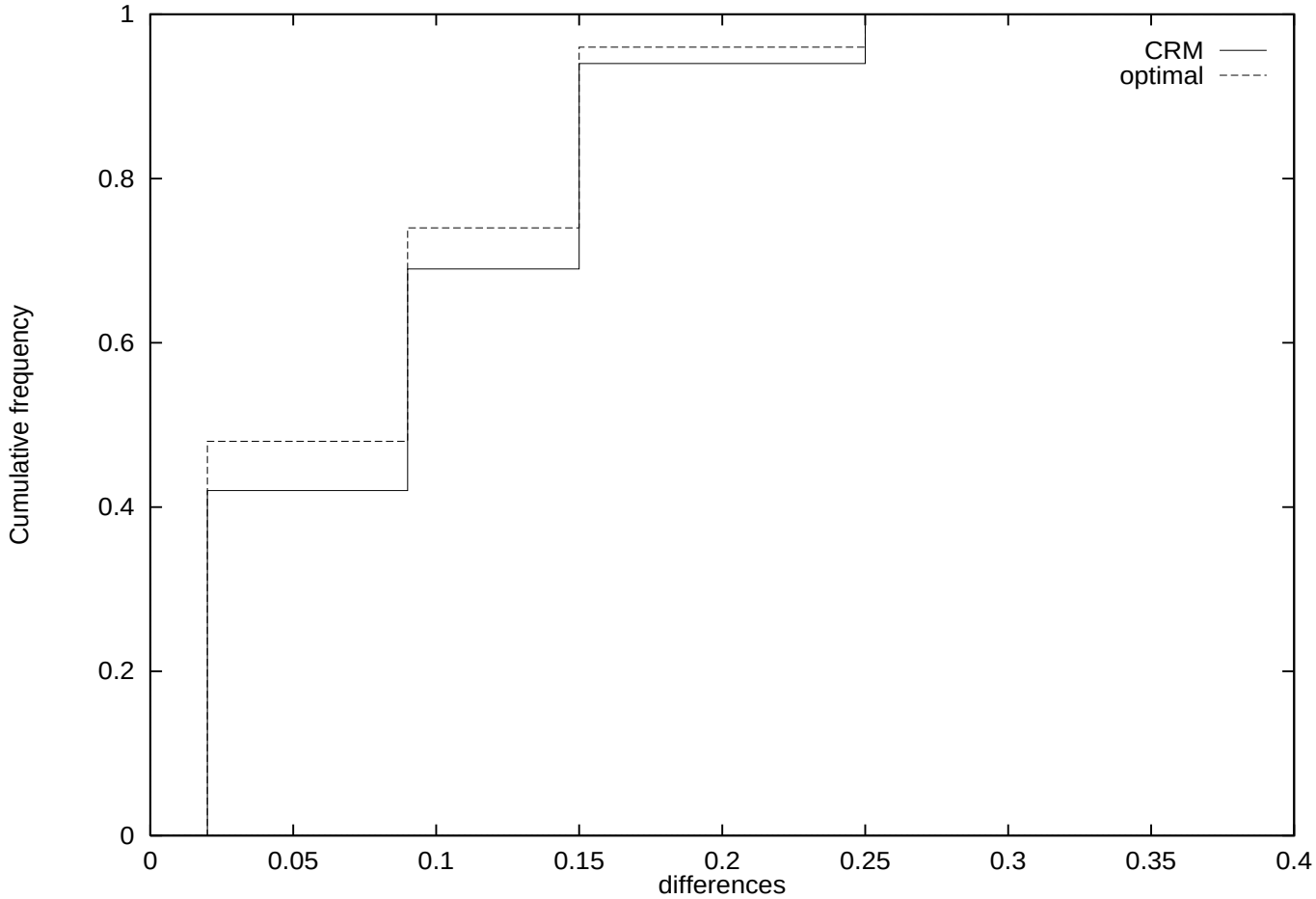
Relative performance by levels;

d_k	1	2	3	4	5	6
R_k	0.05	0.11	0.22	0.35	0.45	0.60
$p_k(16)$	0.05	0.26	0.42	0.21	0.06	0.0
$q_k(16)$	0.04	0.27	0.48	0.17	0.04	0.0

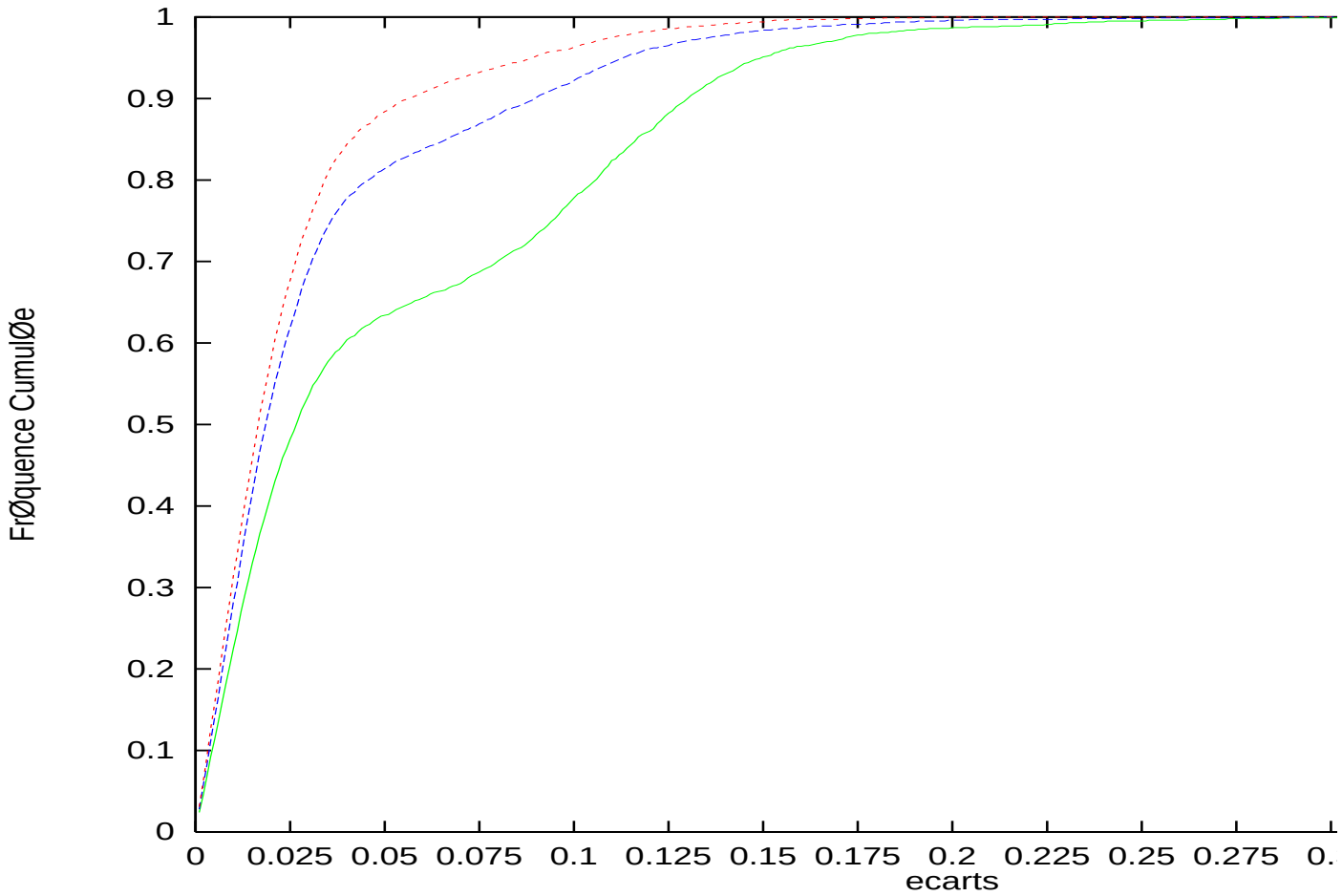
Relative performance by cumulative errors; Let $\alpha = 0.1$ be % simulations where $\Pr(Y = 1) \in (0.10, 0.30)$. This is 0.69 for CRM and 0.74 for optimal.

α	0.02	0.05	0.10	0.15	0.20
p_α	0.42	0.42	0.69	0.94	1.0
q_α	0.48	0.48	0.74	0.96	1.0

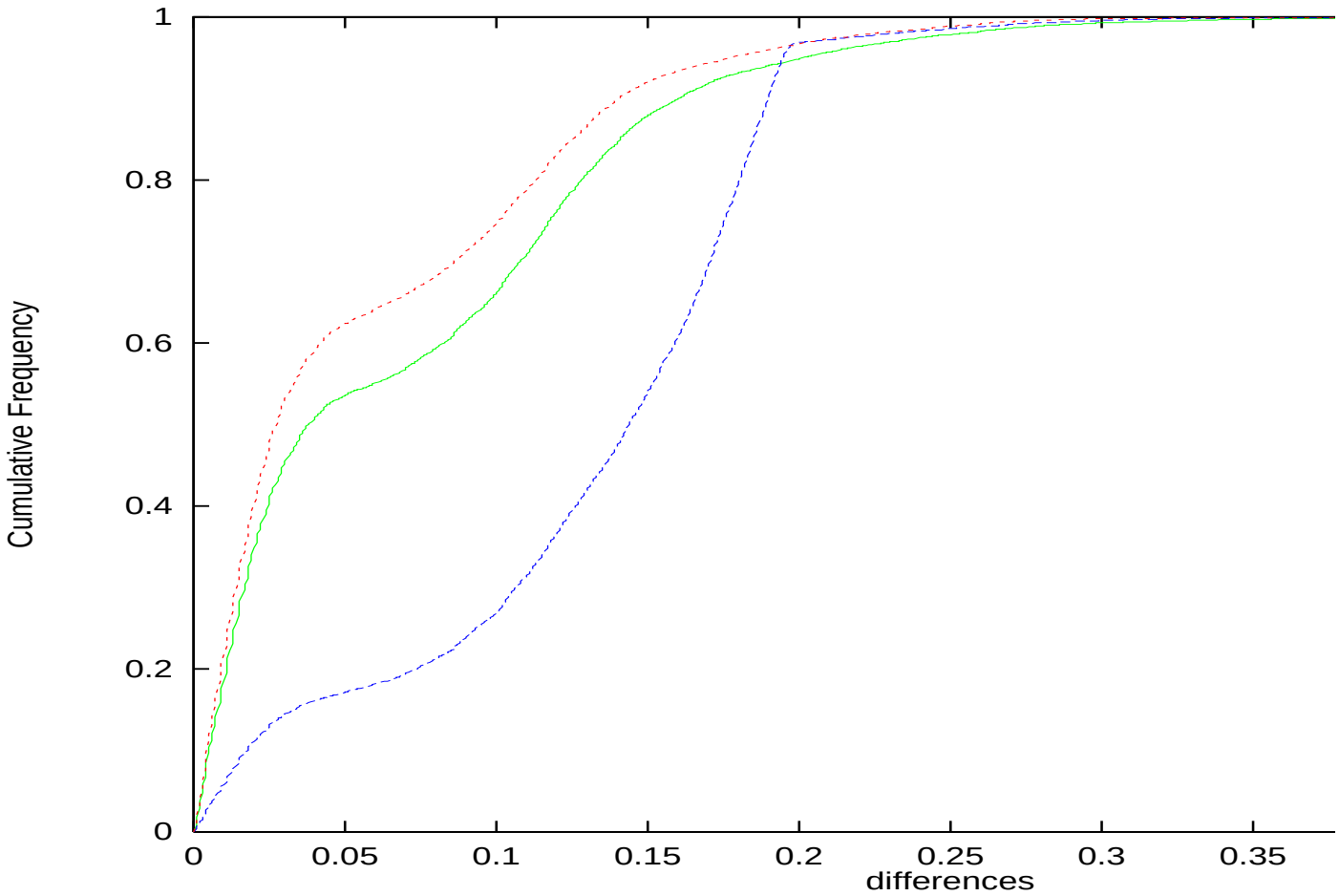
Graph of cumulative errors



Optimal .... , CRM1 .... , CRM2 ....



Optimal .... , CRM .... , 3+3 ....



Optimality of CRM

- Large sample optimality of CRM (Shen, L. and O'Quigley, J. *Biometrika* 1996)
- Finite sample optimality based on simulations (Paoletti et al *Comp. Stat. Data Analysis.* 2002)
- CRM from viewpoint of design optimality. (Tian, T. *Statist. Probability. Letters.* 2016)

Matched comparisons and optimal design

Subject j	v_j	s_j	Different methods	
			mTPI	mTPI2
1	.53	6	1	1
2	.08	2	1	1
3	.29	4	1	1
4	.41	5	2	2
5	.79	-	2	2
6	.04	1	2	2
7	.87	-	2	2
8	.15	3	2	2
9	.63	-	2	3
10	.56	6	3	3
11	.32	4	3	3
12	.09	2	2	3