Optimal sequential designs in dose finding studies

David Azriel; advisors: Yosi Rinott and Micha Mandel

The Hebrew University of Jerusalem

November 25, 2010

Outline

- Example.
- Formulation of the problem.
- Existing methods: standard design; the continual reassessment method (CRM).

Infinite horizon:

- Adaptive design - statistical inference.
- The treatment vs. experimentation dilemma - consistency considerations.
- A related problem: Pitman and Bahadur approaches.

Finite horizon:

- Optimal designs computed via Dynamic Programming equations.
- Example.

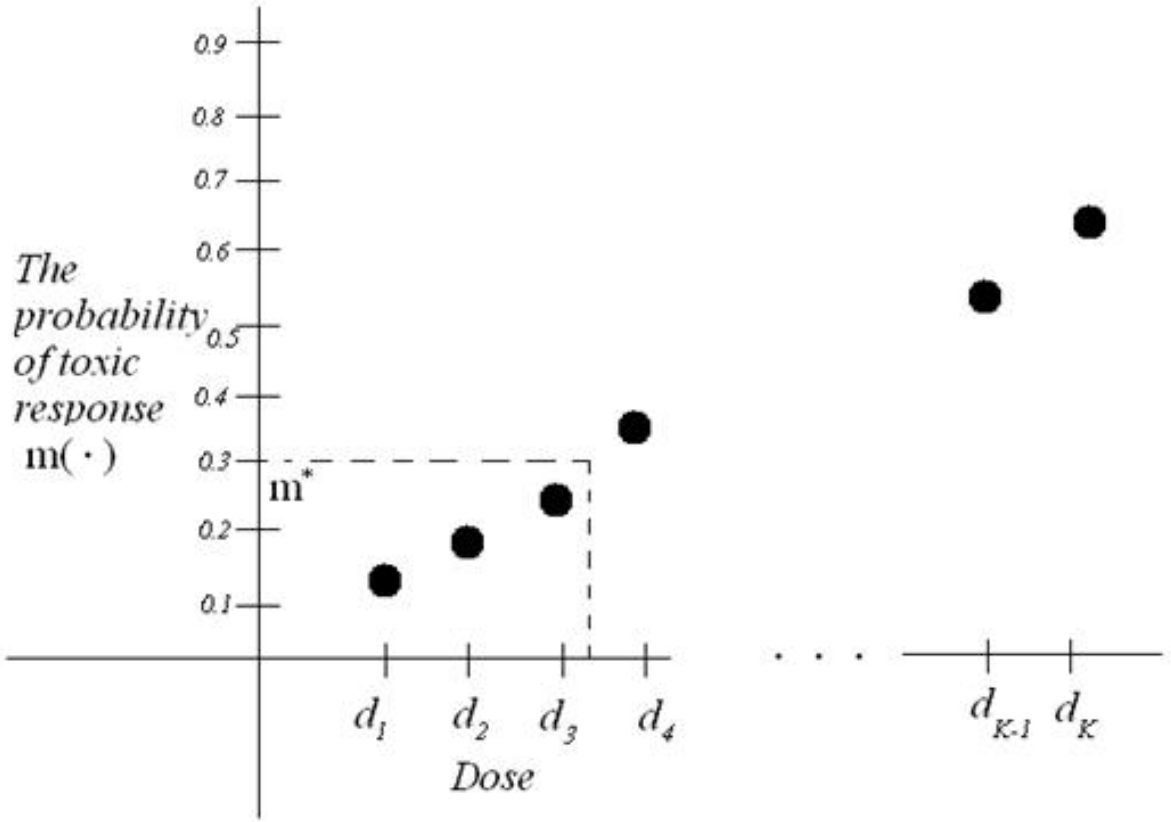
Example

- Mathew et al. (2004). Platelet-Derived Growth Factor Receptor Inhibitor Imatinib Mesylate and Docetaxel: A Modular Phase I Trial in Androgen-Independent Prostate Cancer.
- **Doses:** 20, 25, 30, 35, 40, 45 mg/m^2 (docetaxel).
- **Subjects:** Eight cohorts of six subject each.
- **Target:** To find the the maximum tolerated dose (MTD); i.e., the dose level of docetaxel in combination with oral imatinib at 600 mg daily that achieved a dose-limiting toxicity (DLT) rate closest to 30%.
- **Statistical design:** The continual reassessment method (CRM).

The trial - first cycle

- **The first cohort** was treated 30 mg/m^2. 0/6 DLT's.
- **The second cohort** was treated 45 mg/m^2. 3/4 DLT's (was not completed).
- **The third cohort** was treated 35 mg/m^2. 5/6 DLT's.
- **The fourth cohort** was treated 30 mg/m^2. 3/6 DLT's.
- 30 mg/m^2 is declared the MTD.

The formulation of the problem



Maximum Tolerated Dose (MTD)= d_3

The formulation of the problem

- Dose range: $D := \{d_1, \ldots, d_K\}$.
- $m(d_j)$ denotes the probability of toxic response at dose d_j.
- $m(\cdot)$ is some unknown increasing function.
- m^* denotes a desired response level.
- The Goal: to find the MTD.
  $MTD := d_{j^*}$, $j^* := \arg\min_j |m(d_j) - m^*|$.

Adaptive design

- $x_n \in D$ denotes the dose assigned to the nth subject ; the sequence $\{x_n\}_{n=1}^{\infty}$ is called a design.
- y_n denotes the response of the nth subject.
- x_1 is arbitrary (typically = $x_1 = d_1$).
- $\mathcal{F}_n := \sigma(x_1, y_1, \ldots, x_n, y_n)$ denotes the available data at stage n of the experiment.
- $x_n = f(x_1, y_1, \ldots, x_{n-1}, y_{n-1})$ i.e., $x_n \in \mathcal{F}_{n-1}$.
- We assume that $y_n | \mathcal{F}_{n-1} \sim \text{Bernoulli}(m(x_n))$.
- $n_j = n_j(n)$ is the number of subjects treated d_j among the first n subjects.
- **Unlike regular statistics the main question is how to collect the data such that the inference would be efficient.**

The standard '3+3' design

- The subjects are divided to cohorts of 3.
- The first cohort is treated with d_1.
- If no severe toxicity is observed then the dose is escalated to the next highest level.
- Otherwise, an additional three are treated at the same dose.
- If fewer than 2 toxicities are observed amongst the 6, then the dose is escalated to the next highest level.
- Otherwise the trial is terminated. The dose in use at trial termination is recommended as the MTD.

The continual reassessment method

The CRM assumes the one parameter working model

$$P(y = 1|x = d_j) = \xi_j^{\alpha},$$

where ξ_j $(j = 1, \ldots, K)$ are known constants and α is the unknown parameter.
e.g., $(\xi_1, \ldots, \xi_6) = (0.07, 0.16, 0.30, 0.40, 0.46, 0.53)$.
The CRM approach assigns the next dose according to the maximum likelihood estimate of the MTD.

$$x_{n+1} = d_{\tilde{j}}\ ,\ \tilde{j} = \arg\min_j |m^* - \xi_j^{\hat{\alpha}_n}|$$

where $\hat{\alpha}_n$ is the MLE.

Adaptive design - statistical inference

- Note, $\{y_n\}_{n=1}^{\infty}$ are not independent.
- However, for each $j = 1, \ldots, K$,
  $M_n := \sum_{i=1}^{n} I(x_i = d_j)\{y_i - m(d_j)\}$ is a martingale with respect to $\mathcal{F}_n$.

Lemma

$\bar{y}_n(d_j) \to m(d_j)$ almost surely on $\{n_j \to \infty\}$, $j = 1, \ldots, K$, where $\bar{y}_n(d_j) := \frac{1}{n_j} \sum_{i=1}^{n} I(x_i = d_j) y_i$.

Lemma - proof

Proof. The quadratic variation of M_n is

$$\sum_{i=1}^{n}[I(x_i = d_j)]^2 \cdot m(d_j) \cdot [1 - m(d_j)] = m(d_j) \cdot [1 - m(d_j)]n_j$$

Therefore, by the strong law of large numbers for square integrable martingales,

$$\frac{1}{n_j}\sum_{i=1}^{n} I(x_i = d_j)\{y_i - m(d_j)\} \to 0 \quad a.s.\ on\ \ \{n_j \to \infty\}.$$

Since $\bar{y}_n(d_j) = m(d_j) + \frac{1}{n_j}\sum_{i=1}^{n} I(x_i = d_j)\{y_i - m(d_j)\}$, the lemma follows.

The treatment vs. experimentation dilemma

Typically, dose-finding studies have two different purposes:

1. Treatment: ideally, treat each subject with the MTD; since it is unknown, use the best available estimate of the MTD at the time of treatment.
2. Experimentation: obtain a good estimate for the MTD at the end of the study.

- Usually the emphasis is on the first purpose. Shu and O'Quigley (2008): "being optimal for anything other than the best estimated treatment for the next patient, or group of patients, to be included in the study is not acceptable".
- Purpose 2 is the core of MTD studies, but may require to treat subjects with high doses in order to find the MTD as fast as possible.

The treatment vs. experimentation dilemma (cont.)

In the case of a continuous response that follows a simple linear regression model

$$y = \alpha + \beta x + \varepsilon$$

and a continuous dose space, Lai and Robbins (1982) show that this dilemma can be resolved asymptotically by treating each subject with the estimated MTD based on a truncated version of the least squares estimators.

The treatment vs. experimentation dilemma (cont.)

Theorem

There exists no design that satisfies for all increasing functions m:

$$P(\exists N \; s.t. \; \forall n \geq N : \; x_n = d_{j*}) = 1,$$

or equivalently that $P(x_n \neq d_{j^*} \; i.o. \;) = 0$.

The idea of the proof is that a design that concentrates eventually on one dose, say d_j, can yield a consistent estimator for $m(d_j)$, but cannot estimate well $m(d_i)$ for $i \neq j$; therefore, such a design may miss the MTD.

The treatment vs. experimentation dilemma (cont.)

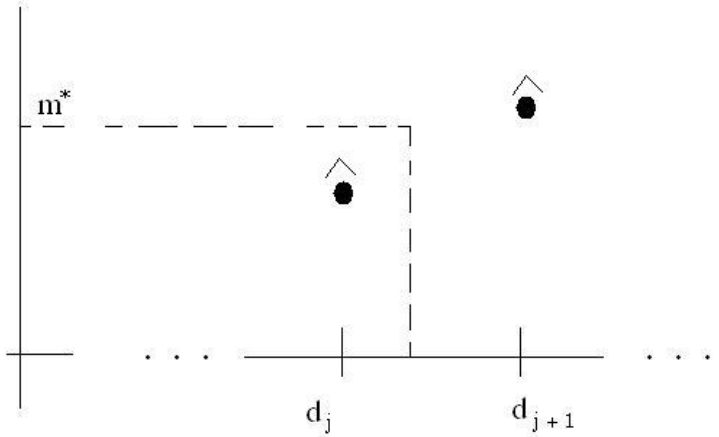
Corollary

Let $\{\widehat{MTD}_n\}_{n=1}^{\infty}$ *be any sequence of estimators of the MTD. If for all n,* $x_{n+1} = \widehat{MTD}_n$ *then* $\{\widehat{MTD}_n\}_{n=1}^{\infty}$ *is not strongly consistent.*

The corollary has important implications for phase I studies because many designs assign the estimated MTD to the next subjects. Such designs cannot yield consistent estimation of the MTD unless severe parametric assumptions on m are imposed. Hence, in our framework, it is not obvious that the aforementioned ethical requirement of Shu and O'Quigley should be accepted.

How to build a consistent design?

Let $\hat{m}_n$ be the estimator for m at stage n. Typically, $\hat{m}_n(d_j) \leq m^* \leq \hat{m}_n(d_{j+1})$ for some j.



Choose $x_n = d_j$ or $x_n = d_{j+1}$ such that if $\hat{m}_n(d_j) \leq m^* \leq \hat{m}_n(d_{j+1})$ is infinitely often then both are infinitely often.

Consistent design (cont.)

What should be the balance?
Mukerjee (1981) suggested equal balance but this is not necessarily optimal.
Here is an option:

$$\text{if } m^* \leq (>) \frac{\hat{m}_n(d_j) + \hat{m}_n(d_{j+1})}{2} \quad \text{then}$$

$$x_{n+1} = \begin{cases} d_j & \text{with probability } 1 - \frac{1}{k} \quad (\frac{1}{k}) \\ d_{j+1} & \text{with probability } \frac{1}{k} \quad (1 - \frac{1}{k}) \end{cases},$$

where $k := k(n, j)$ is the number of times that $\{\hat{m}_n(d_j) \leq m^* \leq \hat{m}_n(d_{j+1})\}$ occurred among the first n subjects.

Consistent design (cont.)

For practical purposes (small n), we found that the algorithm performs better when the rate of choosing the estimated MTD is reduced by replacing k with $a \cdot k + 2$, where a is a (small) constant.

Thus, the choice between d_j and d_{j+1} is done with probability $\frac{1}{a \cdot k + 2}$ which is $\approx 1/2$ for small a.

This modification does not change the asymptotic behavior of the estimator, while improving the learning rate of the response curve in early stages of the experiment.

Consistent design (cont.)

Theorem

Assume that $m(d_{j'}) < m^* < m(d_{j'+1})$. *The design satisfies*

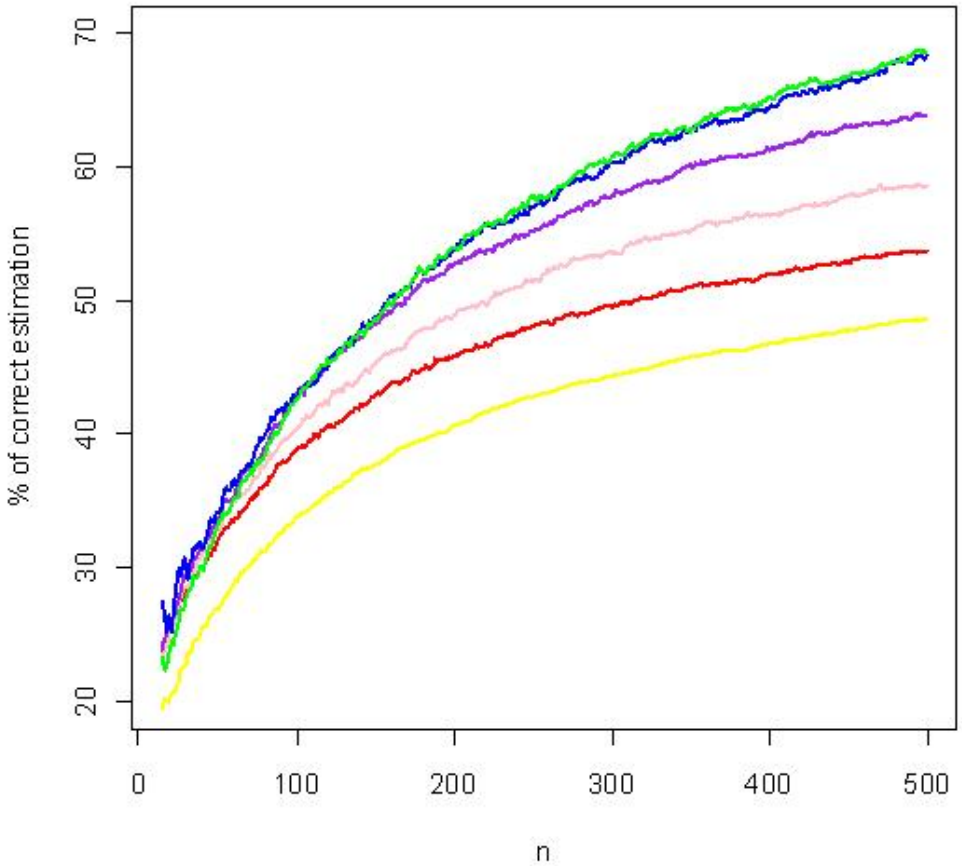
I. $P(\exists N \text{ s.t. } \forall n \geq N : \ \{\hat{m}_n(d_{j'}) \leq m^* \leq \hat{m}_n(d_{j'+1})\} \ \text{occurs}) = 1$.

II. *The sequence* $\widehat{MTD}_n$ *is strongly consistent.*

III. $P(x_n = d_{j^*}) \to 1$.

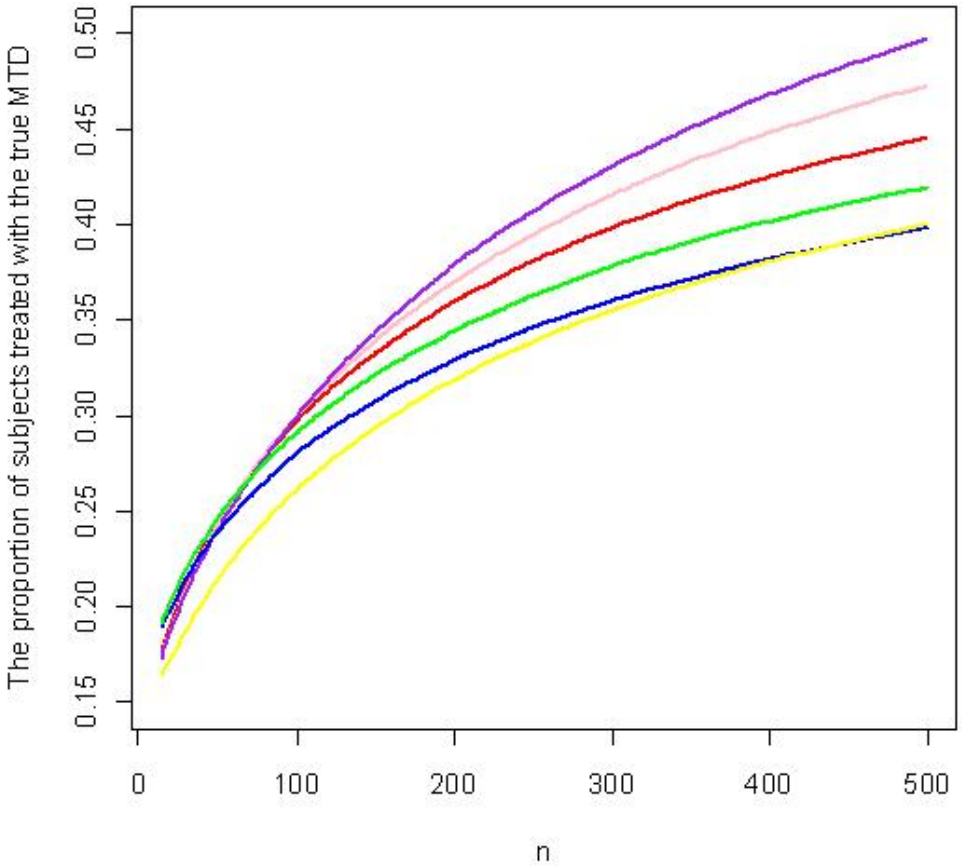
In other words, the probability of the event $x_n = d_{j^*}$ approaches 1. Such a design 'almost' resolves the treatment versus experimentation dilemma.

Simulations - experimentation



Green: up and down method. Blue: Mukerjee. Yellow:CRM. Purple: $a = (10 - 2)/100$. Pink: $a = (10 - 2)/50$. Red: $a = (10 - 2)/30$.

Simulations - treatment.



Green: up and down method. Blue: Mukerjee. Yellow:CRM. Purple: $a = (10 - 2)/100$. Pink: $a = (10 - 2)/50$. Red: $a = (10 - 2)/30$.

Different approach - experimentation

- Consider only two doses, d_1 and d_2, with expected responses $m(d_1)$ and $m(d_2)$, where $m(d_1) < m(d_2)$ and one aim at finding the treatment having probability of response closest to m^*.
- A natural criterion is to select d_2 if $(\hat{m}_n(d_1) + \hat{m}_n(d_2))/2 < m^*$ and to select d_1 otherwise.
- An optimal design may be defined as an allocation rule of $n = n_1 + n_2$ subjects to doses d_1 and d_2 that maximizes $P((\hat{m}_n(d_1) + \hat{m}_n(d_2))/2 < m^*)$ if $m(d_2)$ is closer to m^*.

Pitman approach

- We look at a sequence of parameters $m_n(d_1), m_n(d_2)$ and $m^{n,*}$ such that $(m_n(d_1) + m_n(d_2))/2 - m^{n,*} = K/\sqrt{n}$ for a constant K, $F^n(d_1) \to F(d_1)$, $F^n(d_2) \to F(d_2)$.
- Let $M_n = M_n(F^n(d_1), F^n(d_2), p^{n,*}, \alpha)$ be the minimal number of observations required such that the probability of error is smaller than α when the parameters are $F^n(d_1), F^n(d_2), p^{n,*}$.

Pitman approach (cont.)

Theorem

$$\lim_{n\to\infty}\frac{M_n}{n}=\left\{\frac{z_{1-\alpha}}{2K}\right\}^2\left[\frac{m(d_1)(1-m(d_1))}{\gamma}+\frac{m(d_2)(1-m(d_2))}{1-\gamma}\right],$$

where γ is the limiting allocation in dose d_1, i.e., $\frac{n_1}{n}\to\gamma$.

Therefore, the asymptotically optimal allocation rule is Neyman allocation, that is, choosing $\gamma=\frac{\sqrt{m(d_1)(1-m(d_1))}}{\sqrt{m(d_1)(1-m(d_1))}+\sqrt{m(d_2)(1-m(d_2))}}$ minimizes the limit of $\frac{m_n}{n}$.

Bahadur approach

On the other hand, if the parameters are considered fixed, then the CLT is not relevant, and a probability such as $P((\hat{F}_n(d_1) + \hat{F}_n(d_2))/2 < p^*)$ should be approximated by Large Deviations theory, which is related to the notion of Bahadur efficiency.

Assuming that $m(d_2)$ is closer to m^*, the experimenter wishes to minimize the probability of choosing d_1:

$P[\{\hat{m}_n(d_1) + \hat{m}_n(d_2)\}/2 \geq m^*]$.

Bahadur approach (cont.)

Theorem

Let $\gamma_n = n_1/n$, $0 < \gamma < 1$, *and assume that* $\gamma_n \to \gamma$. *Then,*

$$\lim_{n\to\infty} \frac{1}{n} \log P[\{\hat{m}_n(d_1) + \hat{m}_n(d_2)\}/2 \geq m^*] = \psi(\gamma),$$

where, $\psi(\gamma) = \inf_t\{\gamma \log(1 - m(d_1) + m(d_1)e^{t/\gamma}) + (1 - \gamma)\log(1 - m(d_2) + m(d_2)e^{t/(1-\gamma)}) - t2m^*$.
Moreover, let $\gamma^* = \arg\min \psi(\gamma)$, *and* γ_n^* *be the value of the allocation minimizing* $P[\{\hat{m}_n(d_1) + \hat{m}_n(d_2)\}/2 \geq m^*]$. *Then* $\gamma_n^* \to \gamma^*$.

Comparison of Bahadur and Pitman

Table: The optimal allocation γ^* for different parameters compared to Neyman allocation.

$m(d_1)$	$m(d_2)$	m^*	γ^* (Bahadur)	Neyman allocation (Pitman)
0.1	0.3	0.28	0.4201951	0.3956439
0.2	0.35	0.3	0.4600556	0.4561162
0.22	0.33	0.3	0.4711939	0.4683612
0.25	0.35	0.33	0.4786154	0.4758474
0.2	0.4	0.33	0.4553075	0.4494897
0.1	0.4	0.3	0.3998936	0.3797959

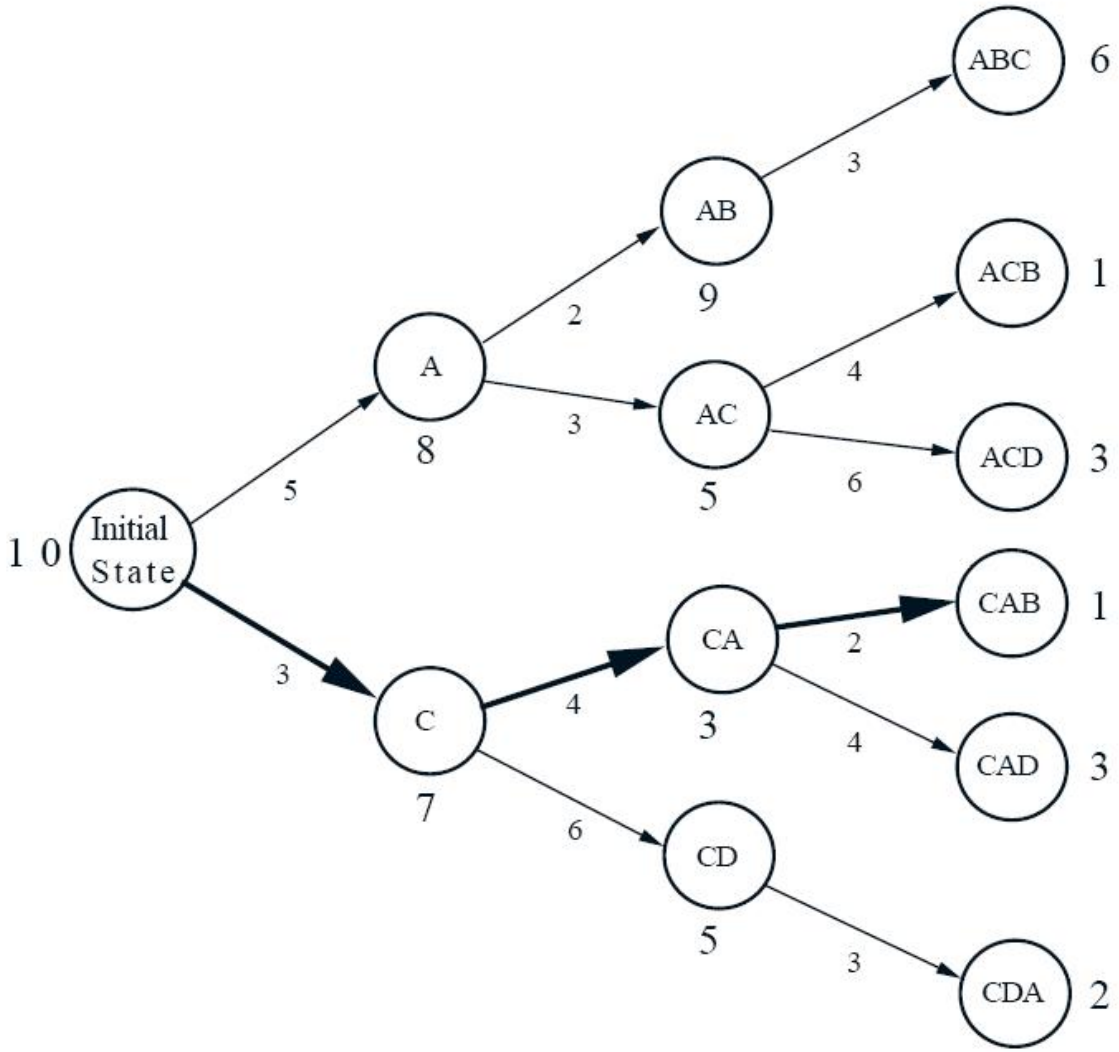
Back to adaptive design

- In practice, the optimal allocation depends on unknown parameters and an adaptive allocation design should be used.
- The effect of this should be explored.
- This will probably not change the allocation according to Pitman type criterion, but may affect the allocation according to Bahadur.
- However, both approaches yield similar allocations.

Finite horizon

- Consider a sequential design of N stages, that is, N subjects (N is small, typically ≤ 50).
- The challenge is to find the optimal sequential design that minimizes a certain loss function.
- Leung and Wang (2002) suggest a dynamic programming (DP) algorithm for this problem, but computational complexities enforce them to use myopic policies.
- An approximate DP algorithm is introduced by Bartroff and Lai (2010).
- But a full DP that assigns subjects to treatments in an optimal way has yet to be developed.
- **In order to be able to find optimal design a specific Bayesian model should be assumed.**

Dynamic Programming



Back to our problem

- At each stage of the experiment the optimal decision (i.e., the dose for the next subject) should be computed.
- Each stage is associated with a state $(r_1, n_1, \ldots, r_K, n_K)$, where n_j and r_j are the number of patients treated with dose d_j and the number of toxic reactions at that dose, respectively.
- DP algorithm: start with the last stage ($\sum n_j = N$) and then go backwards.
- The problem: there are too many possible states, typically $\approx 10^{11}$.

The model

The model:

$$P(y = 1|x, \boldsymbol{\theta}) = m(x, \boldsymbol{\theta}) = 1/\{1 + \exp(-\alpha - \beta x)\},$$

where $\boldsymbol{\theta} = (\alpha, \beta)$, with the conjugate prior

$$\pi(\boldsymbol{\theta}) = C \prod_{j=1}^{K} m(d_j, \boldsymbol{\theta})^{\rho_j} \{1 - m(d_j, \boldsymbol{\theta})\}^{\nu_j - \rho_j} I\{\theta \in \Theta\}$$

where C is a normalizing constant, ρ_j, ν_j are parameters that can be specified from a prior belief on the expected number of toxicities, ρ_j out of ν_j subjects treated with dose d_j ($j = 1, \ldots, K$), and Θ is the support of θ.

The posterior

The posterior density of θ is

$$= C \prod_{j=1}^{K} m(d_j, \theta)^{\rho_j + r_j} \{1 - m(d_j, \theta)\}^{n_j + \nu_j - (r_j + \rho_j)} I\{\theta \in \Theta\}$$

$$= Ce^{\alpha \sum_{j=1}^{K} (\rho_j + r_j) + \beta \sum_{i=1}^{K} (\rho_j + r_j) d_j} \prod_{j=1}^{K} \{1 - m(d_j, \theta)\}^{\nu_j + n_j} I\{\theta \in \Theta\},$$

The posterior density is updated at each stage; it depends on the parameters of the prior and on

$$(\sum_{j=1}^{K} r_j, \sum_{j=1}^{K} r_j d_j, n_1, \ldots, n_K) =: (a_1, a_2, n_1, \ldots, n_K).$$

This model reduces dramatically the number of states

- Our model enables the implementation of DP to states of the form $(a_1, a_2, n_1, \ldots, n_K)$, which has only $K + 1$ dimensions.
- A further simplification is obtained when the dose range D is a lattice $\{d_0, d_0 + \Delta, \ldots, d_0 + (K-1)\Delta\}$, as the number of values that should be considered for a_2 is relatively small: $KN + 1$.
- Furthermore, most practical designs require assignment of doses to a cohort of patients rather than to a single individual in each stage.
- Leung and Wang (2002) consider 8 cohorts of 3 subjects each and $K = 6$ dose levels; the total number of states in the final stage is $\approx 10^6$ which is computationally feasible and is much less than $\approx 10^{11}$, the corresponding number in Leung and Wang's model.

DP equations

The DP equations for cohorts of size H are:

$$J(a_1, a_2, n_1, \ldots, n_K) =$$

$$= \begin{cases} E_\pi[L_1\{\hat{g}(\theta), g(\theta)\} \mid a_1, a_2, n_1, \ldots, n_K] & \sum n_j = N \\ \min\{J_1, \ldots, J_K\} & \sum n_j < N, \end{cases}$$

$$J_i = E_\pi[L_2\{d_i, g(\theta)\} | a_1, a_2, n_1, \ldots, n_K]$$

$$+ \sum_{h=0}^{H} \binom{H}{h} p_i^h (1 - p_i)^{H-h} J(a_1 + h, a_2 + h d_i, \ldots, n_i + H, \ldots)$$

J_i is the expected loss when choosing dose d_i and following the optimal strategy, $g(\theta) = \frac{\text{logit}(m^*) - \alpha}{\beta}$, $\hat{g}$ is an estimate of g based on the data in the final stage, $p_i = E_\pi[m(d_i, \theta) | a_1, a_2, n_1, \ldots, n_k]$, and L_1 and L_2 are loss functions.

DP equations (cont.)

- The loss function L_1 corresponds to the estimation of the MTD at the end of the trial.
- L_2 represents, for each cohort, the penalty for adverse reaction to a dose that is too high or too low.
- The DP equations can be solved by backwards induction.
- The dose d_i where the minimum in is obtained is the decision in state $(a_1, a_2, n_1, \ldots, n_k)$.
- $J(0, \ldots, 0)$ is the expected loss under the optimal policy, that is, the smallest possible expected loss.

An example

Consier the following simple scenario: $D = \{1, 2, 3, 4\}$, and $N = 16$ patients that were divided to 4 cohorts with $H = 4$ patients in each. We used the prior with parameters $(\nu_1, \nu_2, \nu_3, \nu_4) = (1, 0, 0, 1)$, and $(\rho_1, \rho_2, \rho_3, \rho_4) = (0.1, 0, 0, 0.5)$. We considered two loss structures, representing the two extreme aspects of the estimation vs. treatment dilemma:

$$(i) \qquad L_1\{\hat{g}(\theta), g(\theta)\} = \{\hat{g}(\theta) - g(\theta)\}^2 \text{ and } L_2 = 0$$

$$(ii) \qquad L_1 = 0 \text{ and } L_2\{d_i, g(\theta)\} = \{d_i - g(\theta)\}^2,$$

where here the estimator $\hat{g}(\theta)$ is the posterior mean of $g(\theta)$. the first loss targets estimation of the MTD at the end of the trial without penalizing for inappropriate treatment, while the second targets only optimal treatment of subjects during the trial.

Results

- We obtained $J(0, \ldots, 0) = 2.90$ for the first loss and 35.76 for the second.
- A “one step look ahead" approach for the first loss structure yields expected loss of 3.03.
- A myopic policy for the second loss structure yields expected loss of 36.84.
- The optimal policy under the first loss and calculated the second loss obtained loss of 40.78.
- Similarly, applying the optimal design under the second loss and calculating the first yielded a loss of 3.97.

Results - discussion

- A design that treats in an optimal way does not estimate well the MTD
- and a design that focuses on estimating the MTD may perform poorly in terms of treatment.
- This finding is consistent with the asymptotic analysis.
- The improvement of the optimal policy over the myopic policy is small.