

# Optimal Allocation in Multi-Armed Clinical Trials

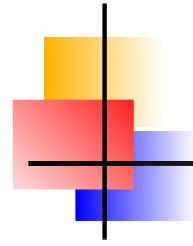
Yevgen Tymofyeyev

Department of Mathematics and Statistics  
University of Maryland, Baltimore County  
Baltimore, MD

## Joint work with

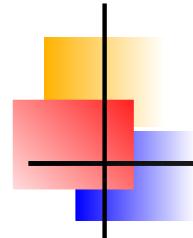
- William F. Rosenberger
- Feifang Hu

Workshop on Adaptive Designs  
September 26, 2003



# *Outline*

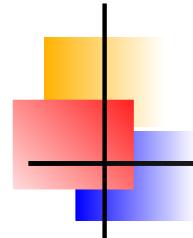
---



## *Outline*

---

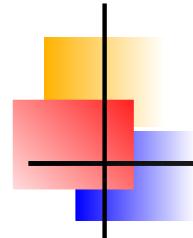
- ▶ Optimal allocation for 2 treatments



## Outline

---

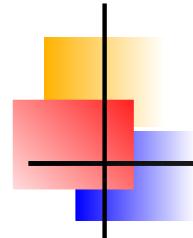
- ▶ Optimal allocation for 2 treatments
- ▶ “Skewed” allocation



## Outline

---

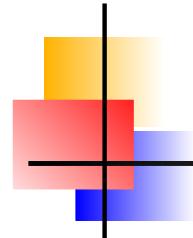
- ▶ Optimal allocation for 2 treatments
- ▶ “Skewed” allocation
- ▶ Formulation and interpretation of the optimization problems



## Outline

---

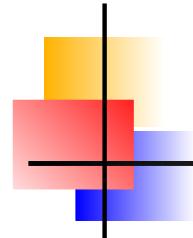
- ▶ Optimal allocation for 2 treatments
- ▶ “Skewed” allocation
- ▶ Formulation and interpretation of the optimization problems
- ▶ Optimal solutions



## Outline

---

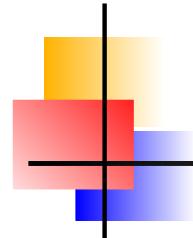
- ▶ Optimal allocation for 2 treatments
- ▶ “Skewed” allocation
- ▶ Formulation and interpretation of the optimization problems
- ▶ Optimal solutions
- ▶ Comparison of allocation ratios



# Outline

---

- ▶ Optimal allocation for 2 treatments
- ▶ “Skewed” allocation
- ▶ Formulation and interpretation of the optimization problems
- ▶ Optimal solutions
- ▶ Comparison of allocation ratios
- ▶ Practical implementation and conclusion



# Optimal allocation for K=2 treatments

Hayre (1979), Jennison and Turnbull (2000)

Problem:

$$\begin{aligned} & \min_{R=n_A/n_B} w_A n_A + w_B n_B, \\ & \text{while } \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B} = \text{const.} \end{aligned}$$

Then optimal allocation is

$$R^* = \frac{\sigma_A}{\sigma_B} \sqrt{\frac{w_A}{w_B}}$$

Rosenberger, Stallard, and Ivanova *et al.* (2001)

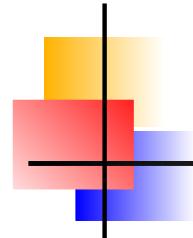
For  $p_A - p_B$  as the treatment effect

if  $w_A = w_B = 1$

$$R^* = \sqrt{\frac{p_A q_A}{p_B q_B}} \quad (\text{Neyman allocation}),$$

if  $w_A = q_A$ ,  $w_B = q_B$

$$R^* = \sqrt{\frac{p_A}{p_B}} \quad (\text{RSIHR allocation}).$$



# The contrast test of homogeneity

Let  $p_i$  be the prob. of success, and  $n_i$  be number of patients for  $i$ -th treatment ( $i = 1, \dots, K$ ).

Define  $\mathbf{p}_c = (p_1 - p_K, p_2 - p_K, \dots, p_{K-1} - p_K)$ .

Hypothesis

$$H_0 : \mathbf{p}_c = 0 \quad \text{versus} \quad H_A : \mathbf{p}_c \neq 0.$$

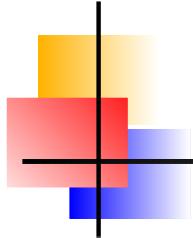
Test statistics

$$\hat{\mathbf{p}}'_c \hat{\Sigma}_n \hat{\mathbf{p}}_c$$

Noncentrality parameter

$$\phi(n_1, \dots, n_K) = \mathbf{p}'_c \Sigma_n^{-1} \mathbf{p}_c,$$

$$\Sigma_n = \begin{bmatrix} p_1 q_1 / n_1 & 0 & \dots & 0 \\ 0 & p_2 q_2 / n_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p_K q_K / n_K \end{bmatrix} + \frac{p_K q_K}{n_K} \mathbf{e} \mathbf{e}'$$



## "Fixed trace allocation"

Denote  $R_i = n_i/n_K, \forall i$

Fix

$$tr(\Sigma_n) = \sum_{i=1}^K p_i q_i / n_i + (K-1)p_K q_K = const$$

Minimize  $\sum_{i=1}^K w_i n_i$ ; here  $w_i$ 's are some objective weights.

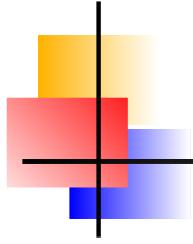
$$\min_{R_1, \dots, R_{K-1}} Q(R_1, \dots, R_{K-1}),$$

where

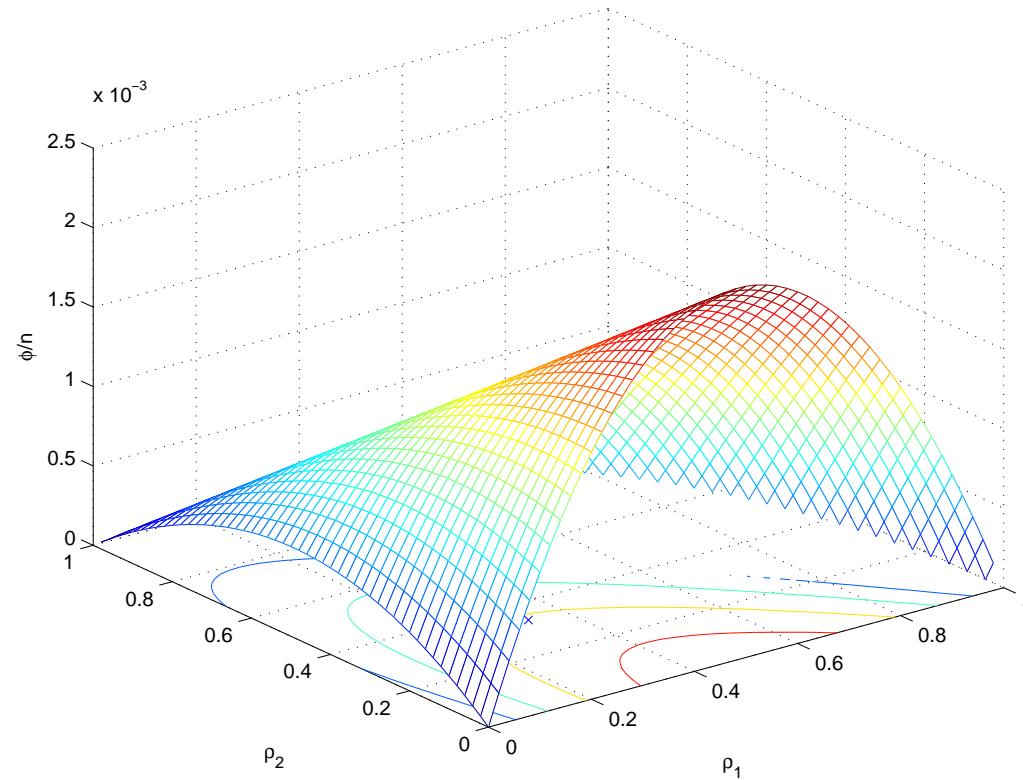
$$Q(R_1, \dots, R_{K-1}) = \sum_{i=1}^K w_i R_i \left( \sum_{j=1}^{K-1} \frac{p_j q_j}{R_j} + (K-1)p_K q_K \right).$$

Solution:

$$R_i^* = \sqrt{\frac{p_i q_i w_K}{(K-1)p_K q_K w_i}}, \quad i = 1, \dots, K-1$$

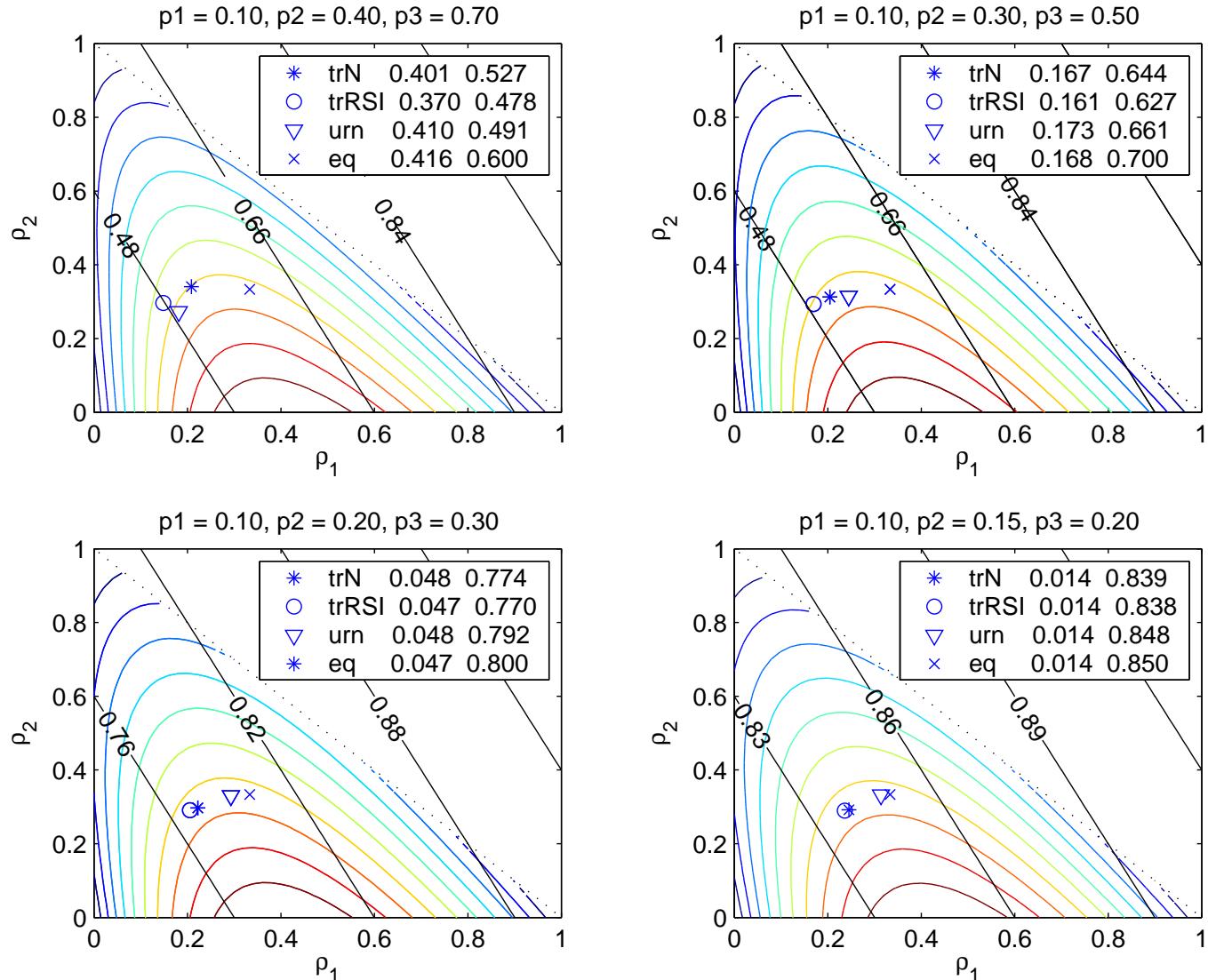


# Noncentrality parameter surface

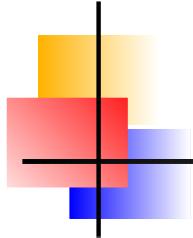


$$\frac{\phi(n_1, n_2, n_3)}{n_1 + n_2 + n_3} = \sum_{j=1}^2 \frac{\rho_j}{p_j q_j} (p_j - p_K)^2 - \frac{\left[ \sum_{j=1}^2 \frac{\rho_j}{p_j q_j} (p_j - p_K) \right]^2}{\sum_{j=1}^3 \frac{\rho_j}{p_j q_j}}, \quad \rho_i = n_i / \sum_{j=1}^K n_j.$$

# Comparison of the allocation ratios



(urn allocation:  $\rho_i = (1/q_i)/\prod_{j=1}^3 1/q_j$ )



# Optimization problems

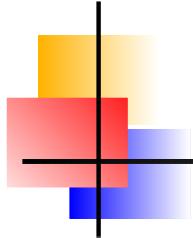
---

## Problem 1:

$$\begin{aligned} \min_{n_1, \dots, n_K} \quad & \sum_{i=1}^K w_i n_i, \\ \text{s.t.} \quad & n_i \geq 0, \quad i = 1, \dots, K, \\ & \phi(n_1, \dots, n_K) \geq C, \end{aligned}$$

## Problem 2:

$$\begin{aligned} \max_{m_1, \dots, m_K} \quad & \phi(m_1, \dots, m_K), \\ \text{s.t.} \quad & m_i \geq 0, \quad i = 1, \dots, K, \\ & \sum_{i=1}^K w_i m_i \leq M, \end{aligned}$$



# Concavity of the noncentrality parameter

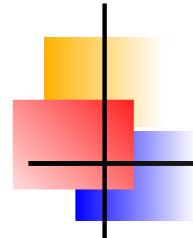
---

We have

$$\phi(n_1, \dots, n_K) = \sum_{j=1}^{K-1} \frac{n_j}{p_j q_j} (p_j - p_K)^2 - \frac{\left[ \sum_{j=1}^{K-1} \frac{n_j}{p_j q_j} (p_j - p_K) \right]^2}{\sum_{j=1}^K \frac{n_j}{p_j q_j}}.$$

## Lemma

The function  $\phi(n_1, \dots, n_K)$ ,  $n_i \geq 0$ ,  $i = 1, \dots, K$  is concave.



# Equivalence of formulation

---

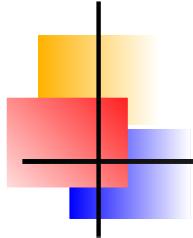
## Theorem 1

There exist solutions to problems 1 and 2.

Denote the optimum of problem 1 by  $\mathbf{n}^* = (n_1^*, \dots, n_K^*)$  and the optimum of problem 2 by  $\mathbf{m}^* = (m_1^*, \dots, m_K^*)$ .

Then,

$$\frac{n_i^*}{\sum_{j=1}^K n_j^*} = \frac{m_i^*}{\sum_{j=1}^K m_j^*}, \forall i = 1, \dots, K.$$



# Neyman allocation

---

## Theorem 2

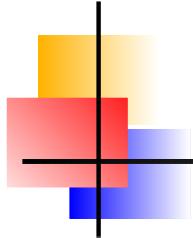
Consider the special case of problems 1 and 2, when

$$\mathbf{w} = (1, \dots, 1)' \in \mathbb{R}^K.$$

Let  $p_1 > p_2 > \dots > p_K$ , then the vector  $\mathbf{n}^* = (n_1^*, \dots, n_K^*)$  such that

$$n_2^* = \dots = n_{K-1}^* = 0 \text{ and } \frac{n_1^*}{n_K^*} = R^* = \sqrt{\frac{p_1 q_1}{p_K q_K}}$$

solves both optimization problems.



## RSIRH allocation

---

### Theorem 3

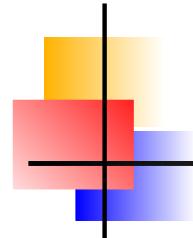
Consider the special case of problems 1 and 2, when

$$\mathbf{w} = (q_1, \dots, q_K)' \in \mathbb{R}^K.$$

Let  $p_1 > p_2 > \dots > p_K$ , then the vector  $\mathbf{n}^* = (n_1^*, \dots, n_K^*)$  such that

$$n_2^* = \dots = n_{K-1}^* = 0 \text{ and } \frac{n_1^*}{n_K^*} = R^* = \sqrt{\frac{p_1}{p_K}}$$

solves both optimization problems.



## Doubly adaptive biased coin design

Eisele (1994), Hu and Zhang (2003)

The goal of the allocation scheme is to have

$$\mathbf{n}/N = \mathbf{v} = \rho(\boldsymbol{\theta}),$$

where  $\rho(\boldsymbol{\theta}) \in \mathbb{R}^K$  is the vector of proportion which depend on unknown parameter  $\boldsymbol{\theta}$ .

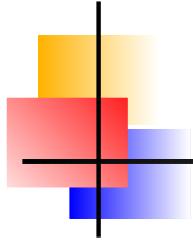
The allocation rule is

$$\mathbf{g}(\mathbf{x}, \boldsymbol{\rho}) = (g_1(\mathbf{x}, \boldsymbol{\rho}), \dots, g_K(\mathbf{x}, \boldsymbol{\rho})) : [0, 1]^K \times [0, 1]^K \rightarrow [0, 1]^K.$$

The  $m + 1$ -th patient is assigned to the treatment  $k$  with probability

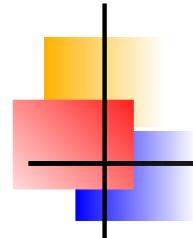
$$p_{m,k} = g_k\left(\frac{n_m}{m}, \hat{\boldsymbol{\rho}}_m\right), \quad k = 1, \dots, K,$$

where  $\hat{\boldsymbol{\rho}}_m = \boldsymbol{\rho}(\hat{\boldsymbol{\theta}}_m)$  is the sample estimate of  $\mathbf{v} = \rho(\boldsymbol{\theta})$ .



# *Practical implementation*

---

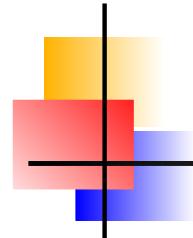


## Practical implementation

---

Obtained allocation function:

$$\rho(\mathbf{p}) = \begin{bmatrix} \rho_1(p_1, \dots, p_K) \\ \rho_2(p_1, \dots, p_K) \\ \vdots \\ \rho_{K-1}(p_1, \dots, p_K) \\ \rho_K(p_1, \dots, p_K) \end{bmatrix} = \begin{bmatrix} R^*/(1 + R^*) \\ 0 \\ \vdots \\ 0 \\ 1/(1 + R^*) \end{bmatrix}, \text{ where } p_1 > p_2 > \dots > p_K$$

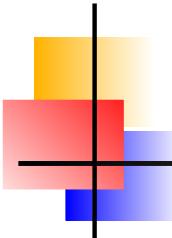


## Practical implementation

Obtained allocation function:

$$\rho(\mathbf{p}) = \begin{bmatrix} \rho_1(p_1, \dots, p_K) \\ \rho_2(p_1, \dots, p_K) \\ \vdots \\ \rho_{K-1}(p_1, \dots, p_K) \\ \rho_K(p_1, \dots, p_K) \end{bmatrix} = \begin{bmatrix} R^*/(1 + R^*) \\ 0 \\ \vdots \\ 0 \\ 1/(1 + R^*) \end{bmatrix}, \text{ where } p_1 > p_2 > \dots > p_K$$

Use Doubly Adaptive Biased Coin Design to target  $\rho(\mathbf{p})$ .



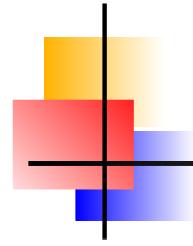
## Practical implementation

Obtained allocation function:

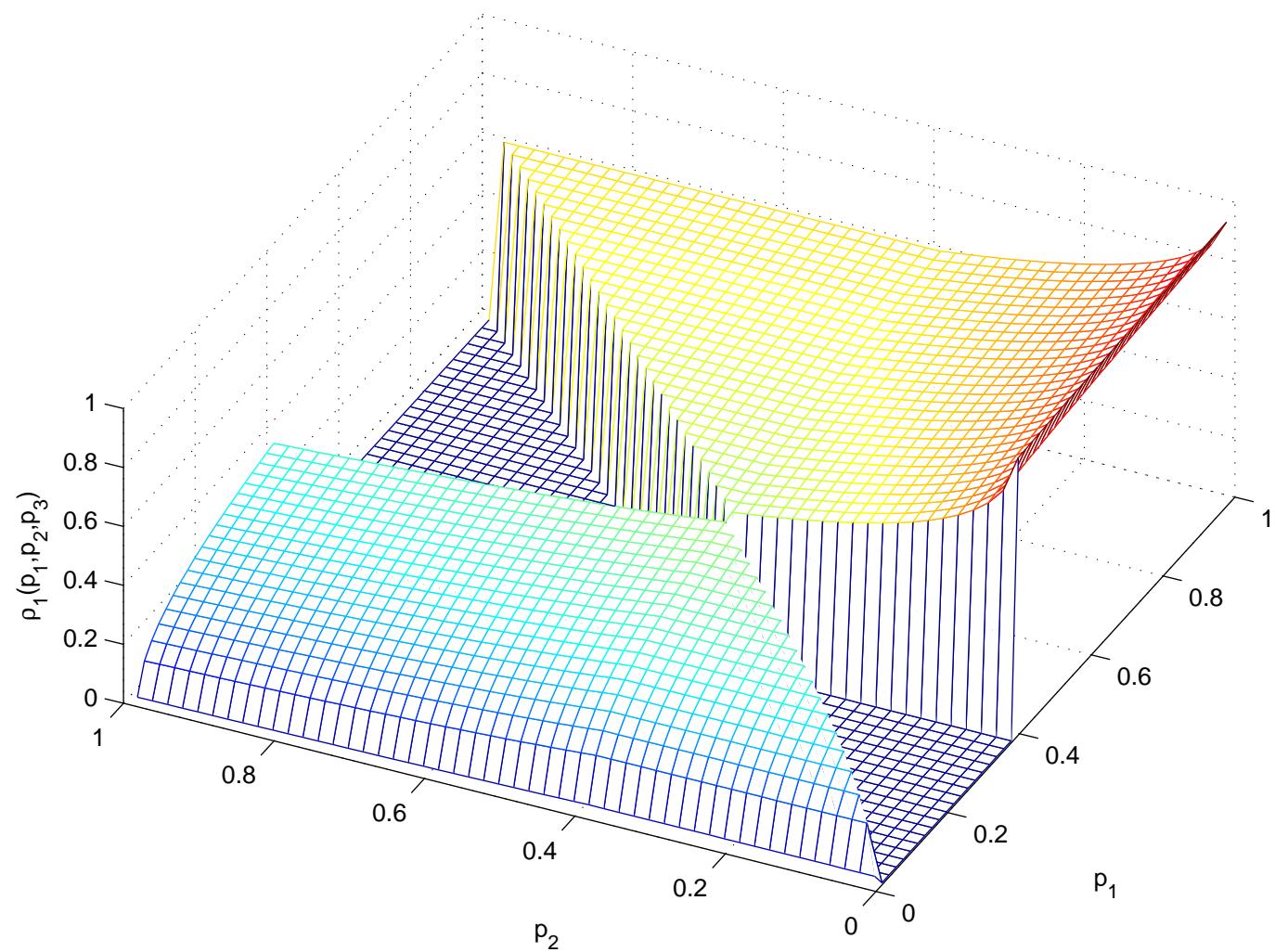
$$\rho(\mathbf{p}) = \begin{bmatrix} \rho_1(p_1, \dots, p_K) \\ \rho_2(p_1, \dots, p_K) \\ \vdots \\ \rho_{K-1}(p_1, \dots, p_K) \\ \rho_K(p_1, \dots, p_K) \end{bmatrix} = \begin{bmatrix} R^*/(1 + R^*) \\ 0 \\ \vdots \\ 0 \\ 1/(1 + R^*) \end{bmatrix}, \text{ where } p_1 > p_2 > \dots > p_K$$

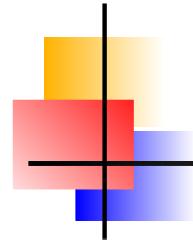
Use Doubly Adaptive Biased Coin Design to target  $\rho(\mathbf{p})$ .  
Employ smoothing techniques to derive  $\rho_{smoothed}(\mathbf{p}) : \mathbb{R}^K \rightarrow \mathbb{R}^K$ .  
For example,

$$\rho_{smoothed}(\mathbf{p}_0) = \int_{\Omega(\mathbf{p}_0)} \frac{\rho(\mathbf{p})}{|\Omega(\mathbf{p}_0)|} d\mathbf{p}$$



# *Plot of the theoretical allocation function*





# *Plot of the smoothed allocation function*

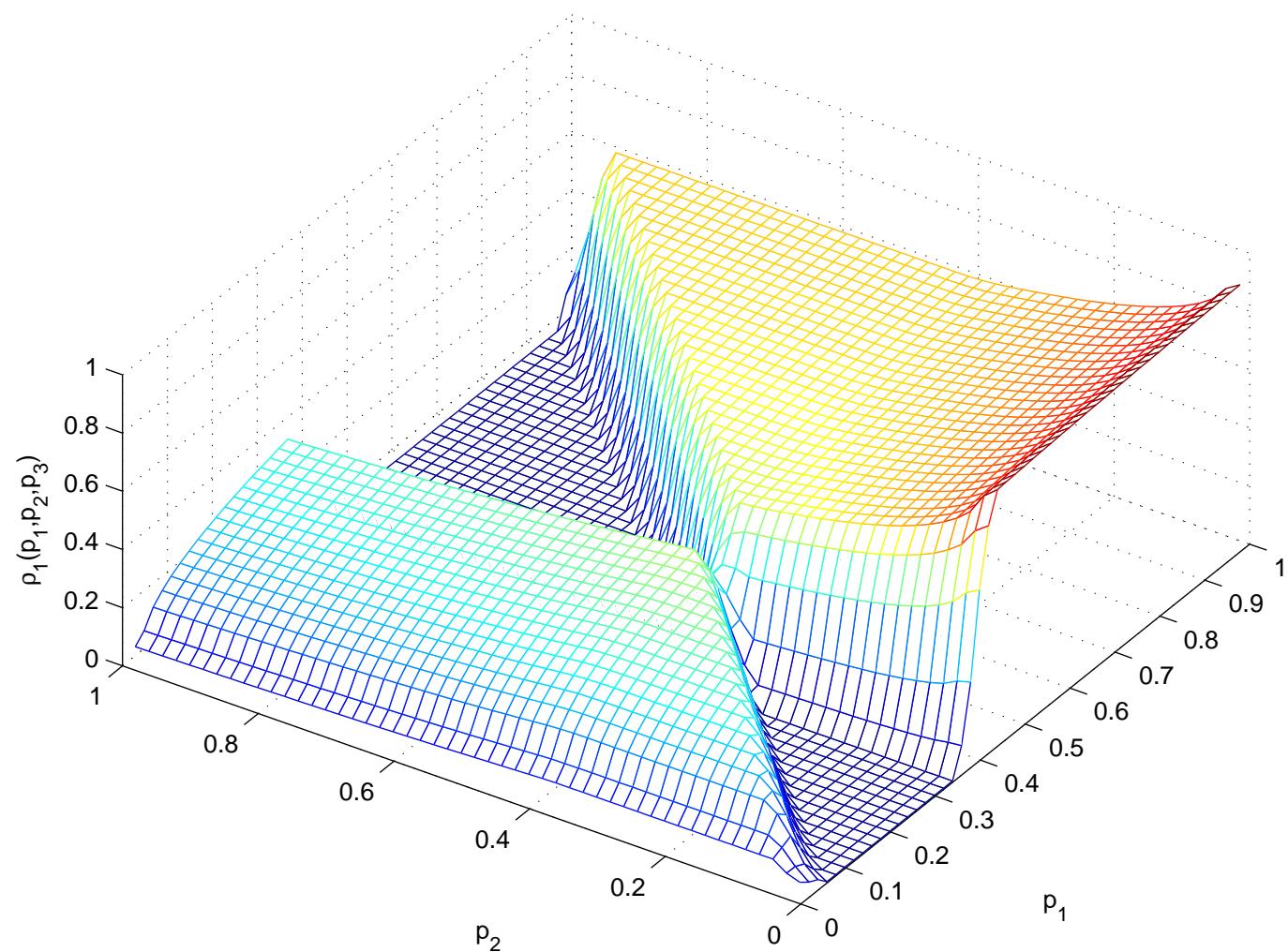


Table: “theoretical” vs “smoothed” allocation ratios

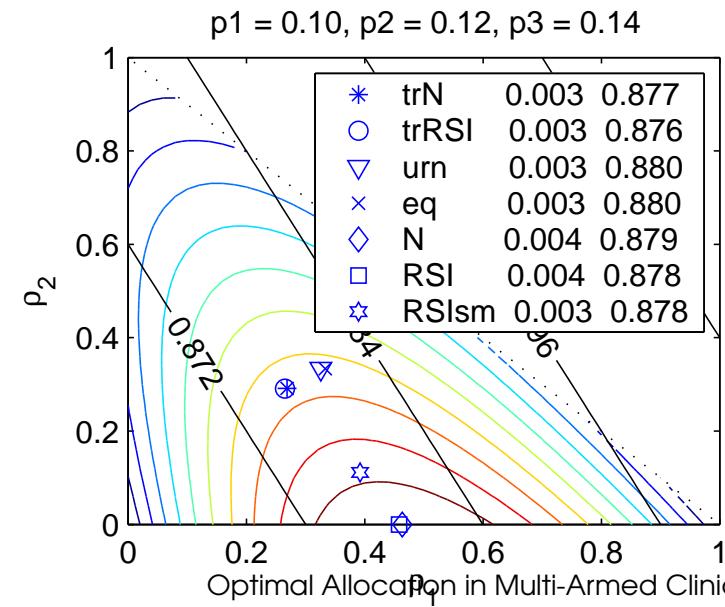
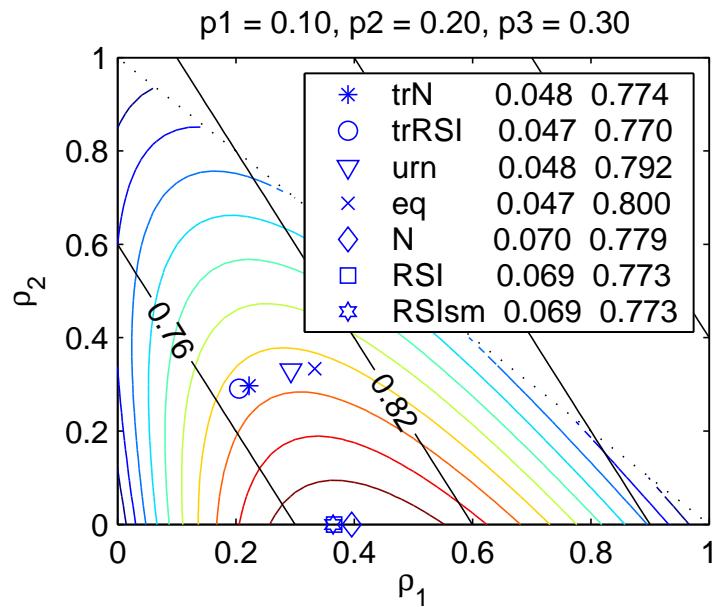
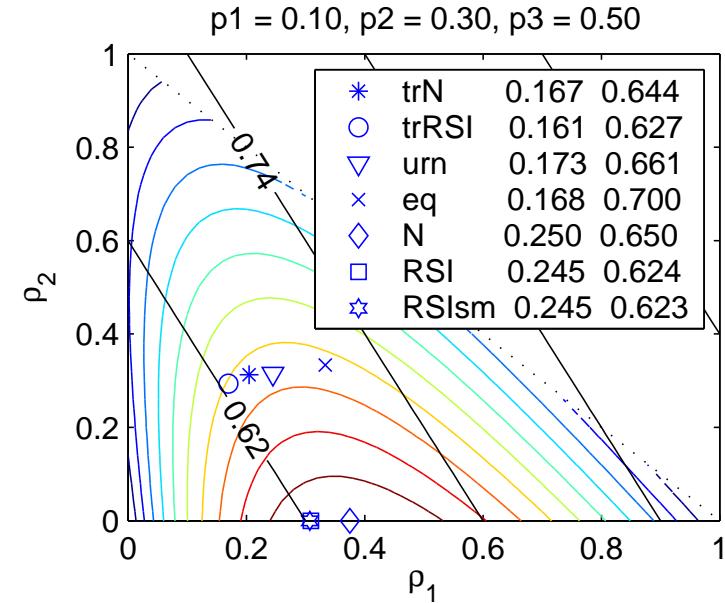
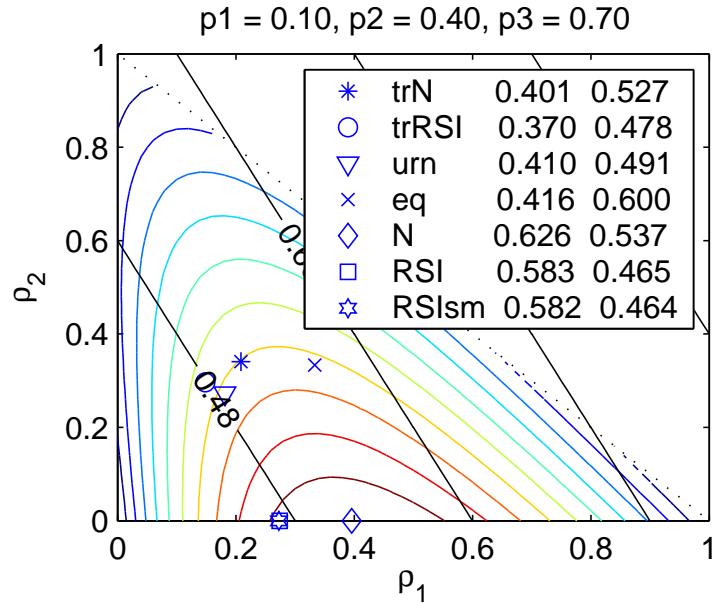
	probabilities			theoretical			smoothed		
	$p_1$	$p_2$	$p_3$	$\rho_1$	$\rho_2$	$\rho_3$	$\rho_1$	$\rho_2$	$\rho_3$
a	0.20	0.22	0.24	0.476	0	0.523	0.420	0.111	0.468
b	0.20	0.22	0.24	0.476	0	0.523	0.417	0.120	0.462
a	0.22	0.20	0.24	0	0.476	0.523	0.218	0.314	0.467
b	0.22	0.20	0.24	0	0.476	0.523	0.177	0.359	0.462
a	0.381	0.382	0.383	0.499	0	0.500	0.349	0.296	0.354
b	0.381	0.382	0.383	0.499	0	0.500	0.338	0.320	0.341
a	0.20	0.24	0.28	0.457	0	0.542	0.457	0	0.542
b	0.20	0.24	0.28	0.457	0	0.542	0.457	0	0.542

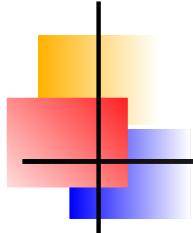
Smoothing parameters:

a - grid cell size = 0.02, step size = 1

b - grid cell size = 0.01, step size = 2

# Comparison of the allocations (cont.)

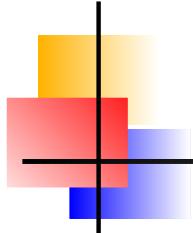




## Concluding Remarks

---

- \* Target allocation is an essential part of response-adaptive design
- \* The obtained “skewed” allocations allow some benefit with respect to power and other objectives, compared to equal allocation
- \* The generalized Neyman and RSIHR allocation can be considered as the solution of corresponding convex optimization problem in multi-armed clinical trials
- \* Theoretically optimal solution “eliminates” all but the extreme treatments and, in general, is a discontinuous function of the parameters
- \* The derived optimal allocation ratios can be applied in a response-adaptive randomization procedure



## References

---

- Hayre, L. S. (1979), "Two-Population Sequential Tests with Three Hypotheses", *Biometrika*, 66,465-474.
- Jennison, C., and Turnbull, B. W. (2000), Group Sequential Methods With Application to Clinical Trials, Boca Raton, FL: Chapman and Hall/CRC.
- Rosengerger, W. F., Stallard, N., Ivanova, A., Harper, C. N., and Ricks, M. L. (2001) "Optimal Adaptive Designs for Binary Response Trials", *Biometrics*, 57,909-913.
- Hu, F., and Rosenberger, W. F. (2003), "Optimality, Variability, Power: Evaluating Response- Adaptive Randomization Procedure for Treatment Comparisons", JASA, in press.
- Hu, F., and Zhang L. X. (2004), "Asymptotic Properties of Double Adaptive Biased Coin Design for Multi Treatment Clinical Trials", *The Annals of Statistics*, in press.