Choosing a suitable randomization procedure with randomizeR

Diane Uschner  David Schindler  Ralf-Dieter Hilgers

RWTH Aachen University

November 2, 2015
Acknowledgment

This research is part of the IDeAl project which has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under Grant Agreement no 602552.

Please visit

http://www.ideal.rwth-aachen.de/

to get the latest news.
Introduction

- Randomized Controlled Clinical Trial with $K \geq 2$ treatment arms
- Restricted Randomization is used for the allocation of treatments to patients.
- Aim: Choose suitable randomization procedure according to problems that might occur during the trial
- Propose a tool for the design of a clinical trial to
  - Assess and compare randomization procedures sequence wise wrt issues (e.g. selection bias)
  - Calculate the exact distribution of the issue (e.g. distribution of the type-I-error for the sequences).
Aim: Exact distribution of the issue

Example: Exact distribution of the type-I-error in case the responses are influenced by selection bias (Convergence Strategy).

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Probability</th>
<th>P(rej)(CS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBAA</td>
<td>0.1666667</td>
<td>0.04229902</td>
</tr>
<tr>
<td>BABA</td>
<td>0.1666667</td>
<td>0.18880215</td>
</tr>
<tr>
<td>ABBA</td>
<td>0.1666667</td>
<td>0.04972876</td>
</tr>
<tr>
<td>BAAB</td>
<td>0.1666667</td>
<td>0.04972876</td>
</tr>
<tr>
<td>ABAB</td>
<td>0.1666667</td>
<td>0.18880215</td>
</tr>
<tr>
<td>AABB</td>
<td>0.1666667</td>
<td>0.04229902</td>
</tr>
</tbody>
</table>
Suitable Randomization Procedure

Process

1. Identify constraints that impact the validity of the trial.
2. Define **issues** that measure the constraint.
3. Assess randomization procedures according to the issues.
4. Select appropriate randomization procedure on the basis of the assessment.

**Definition of issue**

An *issue* is a criterion for the assessment of randomization procedures that can be measured for each randomization sequence.
Structure of the package

issues

randomization procedures

randomization sequences

gensSeq
getAllSeq

assessment/comparison
A randomization procedure $\mathcal{M}$ is a probability distribution on $\Gamma = \{0, 1\}^N$. $t \in \Gamma$ is called randomization sequence. $\mathcal{M}$ produces the sequences

$$\Gamma_\mathcal{M} = \{ t \in \Gamma \mid \mathbb{P}_\mathcal{M}(t) \neq 0 \}$$

```
install.packages('randomizeR')
library(randomizeR)
N<-8
```
Choosing a suitable randomization procedure with randomizeR

Random Allocation Rule

Equally probable final balance sequences:

\[
P_{RAR}(t) = \begin{cases} 
\left( \frac{N}{N/2} \right)^{-1} & \sum_{i=1}^{N} (2 \cdot t_i - 1) = 0 \\
0 & \text{else.}
\end{cases}
\]

\text{rarPar}(N)
Permutted Block Randomization

Equally probable balance sequences that attain balance after each block:

\[ p_{PBR}(t) = \begin{cases} 
    \left( \frac{k}{2} \right)^{-N/k} & \sum_{i=1}^{j \cdot k} (2 \cdot t_i - 1) = 0 \\
    0 & \text{else.} 
\end{cases} \]

for \( j = 1, \ldots, N/k \).

\begin{verbatim}
 k <- 4  #block length
 bc <- rep(k, N/k)  #block constellation
 pbrPar(bc)
\end{verbatim}
Maximal Procedure (Berger (2005))

Equally probable final balance sequences that do not exceed an imbalance boundary $b$.

$$\mathbb{P}_{MP}(t) = \begin{cases} \frac{1}{|\Gamma_{MP}|} & D_N = 0, \forall i : |D_i| \leq b \\ 0 & \text{else.} \end{cases}$$

```r
b <- 2
mpPar(N, b)
```
Big Stick Design

Toss a fair coin in until you hit the imbalance boundary. Then make a deterministic allocation.

\[ P_{BSD}(t) = \begin{cases} 0.5^{N-da} & \sum_{i=1}^{N} |2 \cdot t_i - 1| \leq b \\ 0 & \text{else.} \end{cases} \]

with imbalance boundary \( b \) and number of deterministic allocations

\( da := |\{j : \sum_{i=1}^{j} t_i = b\}|. \)

\[ b <- 2 \]

\textbf{bsdPar(N,b)}
Let’s get random!

createParam() Creates a <.>Par object according to user input.
createSeq() Generates a random sequence according to user input.
genSeq() Generate a random sequence from a <.>Par object.
getAllSeq(myPar) Compute $\Gamma_M$ for $N < 20$.
getProb(seqs) Compute the theoretical probabilities for an object seqs of type randSeq.
saveRand(seqs) Save the randomization protocol including a the randomization sequence(s) to .csv.
Performance of generating $10^x$ RAR sequences, $x \in \{3, 4, 5, 6\}$.

<table>
<thead>
<tr>
<th></th>
<th>user</th>
<th>system</th>
<th>elapsed</th>
</tr>
</thead>
<tbody>
<tr>
<td>system.time(genSeq(rarPar(100),10^3))</td>
<td>0.06</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td>system.time(genSeq(rarPar(100),10^4))</td>
<td>0.70</td>
<td>0.00</td>
<td>0.71</td>
</tr>
<tr>
<td>system.time(genSeq(rarPar(100),10^5))</td>
<td>6.16</td>
<td>0.05</td>
<td>6.23</td>
</tr>
<tr>
<td>system.time(genSeq(rarPar(100),10^6))</td>
<td>62.95</td>
<td>0.44</td>
<td>63.48</td>
</tr>
</tbody>
</table>
Definition

An issue is a criterion for the assessment of randomization procedures that can be measured for each randomization sequence.

- **selBias**: Represent exact rejection probability (size/power) in case the responses are influenced by selection bias.
- **corGuess**: Represent the proportion of correct guesses.
- **chronBias**: Represent exact rejection probability (size/power) in case the responses are influenced by chronological bias.
- **setPower**: Represent the power for a given detectable effect and size.
- **imbal**: Represent the imbalance in allocation numbers.

Table: Issues implemented in randomizeR
Model for the responses (unbiased)

Response

Let $E$ and $C$ be treatments that influence a continuous outcome $Y$. For $i = 1, \ldots, N$, we write

$$Y_i \sim \mathcal{N}(\mu + d \cdot T_i, \sigma^2) \quad (1)$$

where $d \in \mathbb{R}$ denotes the treatment effect, $\mu > 0$ the overall mean and $\sigma^2 > 0$ the equal but unknown variance. $Y_i$ is called response of patient $i$. Higher values of $Y$ are regarded as better.

Represent normal endpoints in randomizeR

```r
normEndp(mu=c(0,0), sigma=c(1,1))
```
Hypothesis of no treatment effect

Test Model:

\[ Y_i \sim \mathcal{N}(\mu + d \cdot T_i, \sigma^2) \]

Test the hypothesis **under model miss-specification**!

True Model:

\[ Y_i \sim \mathcal{N}(\mu + d \cdot T_i + g(\theta, i), \sigma^2) \]

**Null hypothesis**

We test the null hypothesis that the expected effect of the experimental treatment does not differ from the expected effect of the control treatment

\[ H_0 : d = 0 \]

against the two-sided alternative that the expected treatment effects differ

\[ H_1 : d \neq 0 \]
Selection bias under convergence strategy

Third order selection bias

- Trial is randomized.
- Allocation list is concealed.
- But: the investigator can guess the next treatment assignment due to
  - unmasking of past assignments (e.g. due to side effects).
  - restrictions of the randomization procedure.
- Investigator can deny enrollment due to soft inclusion criteria.

Berger (2005)
Selection bias under convergence strategy

Third order selection bias

- Trial is randomized.
- Allocation list is concealed.
- But: the investigator can guess the next treatment assignment due to unmasking of past assignments (e.g. due to side effects).
- restrictions of the randomization procedure.
- Investigator can deny enrollment due to soft inclusion criteria.

Berger (2005)
Biasing Policy

Choose patient $i + 1$ with expected response

$$E(Y_{i+1}) = \begin{cases} 
\mu - \eta & N_E(i) > N_C(i) \\
\mu & N_E(i) = N_C(i) \\
\mu + \eta & N_E(i) < N_C(i)
\end{cases}$$

with selection effect $\eta > 0$.

Proschan (1994)
Exact rejection probability in case of selection bias

**Given the randomization sequence** $t \in \Gamma$ and using Student’s $t$-test in order to test the hypothesis $H_0 : d = 0$ of no treatment effect, the test statistic

$$S = \frac{\sqrt{\frac{N_E N_C}{N_E + N_C}} (\bar{y}_E - \bar{y}_C)}{\frac{1}{N_E + N_C - 2} \left( \sum_{i=1}^{N} t_i (y_i - \bar{y}_E)^2 + \sum_{i=1}^{N} (1 - t_i) (y_i - \bar{y}_C)^2 \right)}$$

with $\bar{y}_E = \frac{1}{N_E} \sum_{i=1}^{N} y_i t_i$, $\bar{y}_C = \frac{1}{N_C} \sum_{i=1}^{N} y_i (1 - t_i)$ and $N = N_E + N_C$ is doubly noncentrally $t$-distributed with parameters $\delta$ and $\lambda$. 

Uschner et al

Choosing a suitable randomization procedure with randomizeR
Exact rejection probability in case of selection bias (2)

The noncentrality parameters can be determined as follows:

\[
\delta = \eta \sqrt{\frac{1}{\sigma^2 N} \sum_{i=1}^{N} 2 \cdot (t_i - \frac{1}{2}) \cdot \text{sign}(D_{i-1})}
\]

\[
\lambda = \frac{\eta^2}{\sigma^2} \left( \sum_{i=1}^{N} \text{sign}(D_{i-1})^2 - \frac{2}{N} \left( \sum_{i=1}^{N} t_i \cdot \text{sign}(D_{i-1}) \right)^2 \right)
\]

\[
- \frac{2}{N} \left( \sum_{i=1}^{N} (1 - t_i) \cdot \text{sign}(D_{i-1}) \right)^2
\]

Langer (2014)

Figure: Doubly noncentral t-distribution, \( N = 12 \)
Assess randomization procedure with randomizeR

```r
pbr <- getAllSeq(pbrPar(bc))

sb <- selBias("CS", eta = 0.6, method = "exact")

endp <- normEndp(mu=c(0,0), sigma = c(1,1))

assess(pbr, sb, endp = endp)
```
Comparison of randomization procedures

pbr <- getAllSeq(pbrPar(bc))
mp <- getAllSeq(mpPar(N,2))
bsd <- getAllSeq(bsdPar(N,2))
C <- compare(sb, pbr, mp, bsd, endp = endp)
plot(C)
Conclusions

- randomizeR makes it easy to generate randomization sequences for a large number of randomization procedures.
- Easy to assess and compare randomization procedures for a large number of issues.
- Assessment should be done before conducting a clinical trial.
Want some more?

Try it yourself! Just type

```r
install.packages("randomizeR")
library("randomizeR")
vignette("comparison-example")
```

in your R command line.


Lehmann (1975) *Nonparametrics: Statistical Methods Based on Ranks*
References


