A studentized permutation test in group sequential designs

Long-Hao Xu[1], Tobias Mütze[2], Frank Konietschke[3], Tim Friede[1]

[1]Department of Medical Statistics,
University Medical Centre Göttingen,
Göttingen, Germany
[2]Statistical Methodology,
Novartis Pharma AG,
Basel, Switzerland
[3]Institute of Biometry and Clinical Epidemiology,
Charité – Universitätsmedizin Berlin,
Berlin, Germany

Symposium: Statistical Planning of Translational Studies

March 20, 2024

Outline

1. Motivation
2. The overview of classical group sequential designs
3. The proposed studentized permutation method
4. Simulations
5. Conclusions

Outline

Motivation

- Group sequential designs are commonly used in clinical trials. The main characteristic of this design is that the sample size is not fixed in advance and the data are sequentially obtained.
- However, there are many situations in medical research where sample sizes are rather restricted. **For example, preclinical trials, early clinical trials, and rare diseases.**
- From a statistical perspective, the above situations may lead to the **small sample size problem**. The existing methods do not control the Type I error rate well for small samples.

Outline

Group sequential designs: overview

Without loss of generality, we consider **one-sided hypothesis testing** and denote $Z^{(k)}$ as the studentized test statistic at the end of stage $k = 1, ..., K$.

After stage $k = 1, ..., K - 1$

- if $Z^{(k)} \geq c_k$ stop, reject H_0
- otherwise continue to group $k + 1$,

after stage K

- if $Z^{(K)} \geq c_K$ stop, reject H_0
- otherwise stop, accept H_0.

There is another option to consider stopping for futility. In this presentation, we do not discuss this situation.

Consider the **two-arm** comparison in the general K-stage unbalanced design.
For $i = 1, ..., K$,
m_i: the number of the new collected data of the **treatment arm** at the end of stage i
n_i: the number of the new collected data of the **control arm** at the end of stage i
Let $\widetilde{m}_i = \sum_{j=1}^{i} m_j$ and $\widetilde{n}_i = \sum_{j=1}^{i} n_j$.
Assume that $X_i \overset{iid}{\sim} F_1,\ i = 1, 2, ...,$ and $Y_i \overset{iid}{\sim} F_2,\ i = 1, 2, ...$ are from the treatment arm and the control arm, respectively. Let $\mu_1 = \mathbb{E}[X_1]$ and $\mu_2 = \mathbb{E}[Y_1]$.
From a practical view, we assume the allocation ratio between two arms is fixed, namely for all $i = 1, ..., K$

$$\frac{m_i}{n_i} = \gamma > 0. \tag{1}$$

Table: K-stage unbalanced design

T :	m_1	m_2	...	m_K
C :	n_1	n_2	...	n_K

Thus, under the null hypothesis $H_0 : \mu_1 = \mu_2$, the corresponding studentized test statistics are

$$S_{\widetilde{m}_k,\widetilde{n}_k} = \frac{\sqrt{\frac{\widetilde{m}_k\widetilde{n}_k}{\widetilde{m}_k+\widetilde{n}_k}}(\widehat{\mu}_{1,k} - \widehat{\mu}_{2,k})}{\sqrt{\frac{\widetilde{n}_k}{\widetilde{m}_k+\widetilde{n}_k}\widehat{\sigma}^2_{1,k} + \frac{\widetilde{m}_k}{\widetilde{m}_k+\widetilde{n}_k}\widehat{\sigma}^2_{2,k}}} \tag{2}$$

where

$$\widehat{\mu}_{1,k} = \frac{1}{\widetilde{m}_k}\sum_{i=1}^{\widetilde{m}_k} X_i, \quad \widehat{\mu}_{2,k} = \frac{1}{\widetilde{n}_k}\sum_{i=1}^{\widetilde{n}_k} Y_i,$$

$$\widehat{\sigma}^2_{1,k} = \frac{1}{\widetilde{m}_k - 1}\sum_{i=1}^{\widetilde{m}_k}(X_i - \bar{X}_{\widetilde{m}_k})^2, \quad \widehat{\sigma}^2_{2,k} = \frac{1}{\widetilde{n}_k - 1}\sum_{i=1}^{\widetilde{n}_k}(Y_i - \bar{Y}_{\widetilde{n}_k})^2.$$

Under some assumptions, the vector $\mathbf{S} = (S_{\widetilde{m}_1,\widetilde{n}_1}, \ldots, S_{\widetilde{m}_K,\widetilde{n}_K})'$ asymptotically follows the multivariate normal distribution $N(\mathbf{0}, \mathbf{\Sigma})$ with

$$\mathbf{\Sigma}_{ij} = \mathrm{Cov}(S_{\widetilde{m}_i,\widetilde{n}_i}, S_{\widetilde{m}_j,\widetilde{n}_j}) = \sqrt{\mathcal{I}^*_{\min\{i,j\}}/\mathcal{I}^*_{\max\{i,j\}}}$$

under the null hypothesis $H_0 : \mu_1 = \mu_2$.

This corvariance structure is also known as independent increments property.

After choosing a suitable α-spending function f, we prespecify the maximum information $\mathcal{I}_{\max}$ and we find the boundary c_i that solves the equations

$$\begin{cases} \mathbb{P}\left(S_{\widetilde{m}_1,\widetilde{n}_1} \geq c_1\right) = f\left(\widehat{\mathcal{I}}_1/\mathcal{I}_{\max}\right) \\ \mathbb{P}\left(S_{\widetilde{m}_1,\widetilde{n}_1} < c_1, S_{\widetilde{m}_2,\widetilde{n}_2} \geq c_2\right) = f\left(\widehat{\mathcal{I}}_2/\mathcal{I}_{\max}\right) - f\left(\widehat{\mathcal{I}}_1/\mathcal{I}_{\max}\right) \\ \cdots \\ \mathbb{P}\left(S_{\widetilde{m}_1,\widetilde{n}_1} < c_1, \ldots, S_{\widetilde{m}_{K-1},\widetilde{n}_{K-1}} < c_{K-1}, S_{\widetilde{m}_K,\widetilde{n}_K} \geq c_K\right) = \alpha - f\left(\widehat{\mathcal{I}}_{K-1}/\mathcal{I}_{\max}\right), \end{cases} \tag{3}$$

where $\widehat{\mathcal{I}}_i = (\widehat{\mathrm{Var}}(\widehat{\mu}_{1,i} - \widehat{\mu}_{2,i}))^{-1}$, and the vector $(S_{\widetilde{m}_1,\widetilde{n}_1}, \ldots, S_{\widetilde{m}_K,\widetilde{n}_K})$ converges in distribution to the multivariate normal distribution $N(\mathbf{0}, \boldsymbol{\Sigma})$.

There are two commonly-used α-spending functions in practice.

- Pocock type:

$$f(t) = \min\{\alpha \log(1 + (e - 1)t), \alpha\}$$

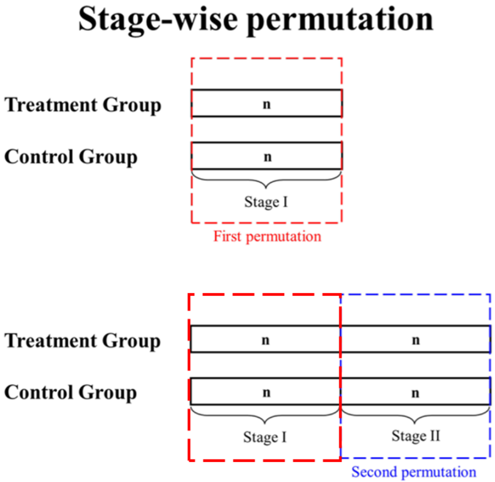
- O'Brien-Fleming type (for one-sided hypothesis testing):

$$f(t) = \min\{2 - 2\Phi(\Phi^{-1}(1 - \alpha/2)/\sqrt{t}), \alpha\}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

Outline

The main idea of this proposed stage-wise permutation is that **the observations stay within their original stage after the permutation**.



The test statistics at the end of stage $k = 1, ..., K$ is computed as

$$S^{\tau}_{\widetilde{m}_k,\widetilde{n}_k} = \frac{\sqrt{\frac{\widetilde{m}_k \widetilde{n}_k}{\widetilde{m}_k + \widetilde{n}_k}}(\widehat{\mu}^{\tau}_{1,k} - \widehat{\mu}^{\tau}_{2,k})}{\sqrt{\frac{\widetilde{n}_k}{\widetilde{m}_k + \widetilde{n}_k}\widehat{\sigma}^2_{1,k,\tau} + \frac{\widetilde{m}_k}{\widetilde{m}_k + \widetilde{n}_k}\widehat{\sigma}^2_{2,k,\tau}}} \tag{4}$$

where these values $\widehat{\mu}^{\tau}_{1,k}$, $\widehat{\mu}^{\tau}_{2,k}$, $\widehat{\sigma}^2_{1,k,\tau}$, and $\widehat{\sigma}^2_{2,k,\tau}$ denote the quantities $\widehat{\mu}_{1,k}$, $\widehat{\mu}_{2,k}$, $\widehat{\sigma}^2_{1,k}$, and $\widehat{\sigma}^2_{2,k}$ **being computed with the permuted variables**.

Next, we consider how to compute the corresponding boundary c_j^{τ} based on the test statistics $S^{\tau}_{\widetilde{m}_j,\widetilde{n}_j}$. We adopt the α-spending function approach and find the boundary c_j^{τ} that solves the equations

$$\begin{cases} \widetilde{\mathbb{P}}\left(S^{\tau}_{\widetilde{m}_1,\widetilde{n}_1} \geq c_1^{\tau}\right) = f\left(\widehat{\mathcal{I}}_1/\mathcal{I}_{\max}\right) \\ \widetilde{\mathbb{P}}\left(S^{\tau}_{\widetilde{m}_1,\widetilde{n}_1} < c_1^{\tau}, S^{\tau}_{\widetilde{m}_2,\widetilde{n}_2} \geq c_2^{\tau}\right) = f\left(\widehat{\mathcal{I}}_2/\mathcal{I}_{\max}\right) - f\left(\widehat{\mathcal{I}}_1/\mathcal{I}_{\max}\right) \\ \cdots \\ \widetilde{\mathbb{P}}\left(S^{\tau}_{\widetilde{m}_1,\widetilde{n}_1} < c_1^{\tau}, \ldots, S^{\tau}_{\widetilde{m}_{K-1},\widetilde{n}_{K-1}} < c_{K-1}^{\tau}, S^{\tau}_{\widetilde{m}_K,\widetilde{n}_K} \geq c_K^{\tau}\right) = \alpha - f\left(\widehat{\mathcal{I}}_{K-1}/\mathcal{I}_{\max}\right). \end{cases} \tag{5}$$

In the above equations, **we use the permutation distribution to approximate the distribution of the test statistics under the null hypothesis.** Thus, the proposed method provides a new way to generating the boundaries, and the critical values depend on the observed data.

In the next steps, we will show the asymptotic properties of the proposed method. We will show that the conditional distribution of $\mathbf{S}^\tau = (S^\tau_{\widetilde{m}_1,\widetilde{n}_1}, \ldots, S^\tau_{\widetilde{m}_K,\widetilde{n}_K})'$ given the observed data approximates the unconditional distribution of $\mathbf{S} = (S_{\widetilde{m}_1,\widetilde{n}_1}, \ldots, S_{\widetilde{m}_K,\widetilde{n}_K})'$.

Theorem

Under some assumptions, the conditional permutation distribution of

$$\mathbf{S}^\tau = (S^\tau_{\widetilde{m}_1,\widetilde{n}_1}, S^\tau_{\widetilde{m}_2,\widetilde{n}_2}, \ldots, S^\tau_{\widetilde{m}_K,\widetilde{n}_K})'$$

given the observed data $\biguplus_{i=1}^K \mathbf{Z}_i$ *weakly converges to a multivariate normal distribution* $N(\mathbf{0}, \mathbf{\Sigma})$ *in probability, where the* (i,j)*-th element of* $\mathbf{\Sigma}$ *is*

$$\mathbf{\Sigma}_{ij} = Cov(S^\tau_{\widetilde{m}_i,\widetilde{n}_i}, S^\tau_{\widetilde{m}_j,\widetilde{n}_j}) = \sqrt{\mathcal{I}^*_{\min\{i,j\}}/\mathcal{I}^*_{\max\{i,j\}}}.$$

Define the decision function of group sequential designs as

$$\varphi_n = \sum_{i=1}^{K} \varphi_{n,i}, \tag{6}$$

where

$$\varphi_{n,1} = \begin{cases} 1 & \text{if } S_{\widetilde{m}_1,\widetilde{n}_1} \geq c_1 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\varphi_{n,i} = \begin{cases} 1 & \text{if } S_{\widetilde{m}_1,\widetilde{n}_1} < c_1 \text{ and } \dots \text{ and } S_{\widetilde{m}_{i-1},\widetilde{n}_{i-1}} < c_{i-1} \text{ and } S_{\widetilde{m}_i,\widetilde{n}_i} \geq c_i \\ 0 & \text{otherwise} \end{cases}$$

for $i = 2, \dots, K$.
With the definition of φ_n, it is easy to show $\mathbb{E}[\varphi_n] \to \alpha$ under the null hypothesis and $\mathbb{E}[\varphi_n] \to 1$ under the alternative hypothesis as $n \to \infty$.

With the definition of φ_n in (6), we give the following definition of its permutated version. Define the permutated version of φ_n as

$$\varphi_n^{\mathcal{T}} = \sum_{i=1}^{K} \varphi_{n,i}^{\mathcal{T}}, \tag{7}$$

where

$$\varphi_{n,1}^{\mathcal{T}} = \begin{cases} 1 & \text{if } S_{\widetilde{m}_1,\widetilde{n}_1} \geq c_1^{\mathcal{T}} \\ 0 & \text{otherwise} \end{cases}$$

and

$$\varphi_{n,i}^{\mathcal{T}} = \begin{cases} 1 & \text{if } S_{\widetilde{m}_1,\widetilde{n}_1} < c_1^{\mathcal{T}} \text{ and } \ldots \text{ and } S_{\widetilde{m}_{i-1},\widetilde{n}_{i-1}} < c_{i-1}^{\mathcal{T}} \text{ and } S_{\widetilde{m}_i,\widetilde{n}_i} \geq c_i^{\mathcal{T}} \\ 0 & \text{otherwise} \end{cases}$$

for $i = 2, \ldots, K$. The main difference between φ_n in (6) and $\varphi_n^{\mathcal{T}}$ in (7) is that the critical value c_i depends on the multivariate normal distribution while the critical value $c_i^{\mathcal{T}}$ depends on the permutational distribution of the observed data.

In the next theorem, we will show that the unconditional test φ_n and the conditional permutation test φ_n^τ are asymptotically equivalent, which means that both tests have asymptotically the same power.

Theorem

Suppose that the assumptions of Theorem 1 are fulfilled.

1. *Under the null hypothesis $H_0 : \mu_1 = \mu_2$, the permutation test φ_n^τ is* **asymptotically exact** *at α level of significance, i.e. $\mathbb{E}[\varphi_n^\tau] \to \alpha$, and* **asymptotically equivalent** *to φ_n, i.e.*

$$\mathbb{E}[|\varphi_n - \varphi_n^\tau|] \to 0, \ \text{as } n \to \infty.$$

2. *The permutation test φ_n^τ is* **consistent**, *i.e.*

$$\mathbb{E}[\varphi_n^\tau] \to \alpha \mathbf{1}\{\mu_1 - \mu_2 = 0\} + \mathbf{1}\{\mu_1 - \mu_2 \neq 0\}, \ \text{as } n \to \infty.$$

In the framework of permutation, we do not need the assumption of normality for the observations.

Thus, the null hypothesis $H_0 : \mu_1 = \mu_2$ does not imply that F_1 and F_2 have the same distribution.

The permutation distribution is one way to approximate or mimic the distribution of the test statistics under the null hypothesis.

Outline

Simulation setting

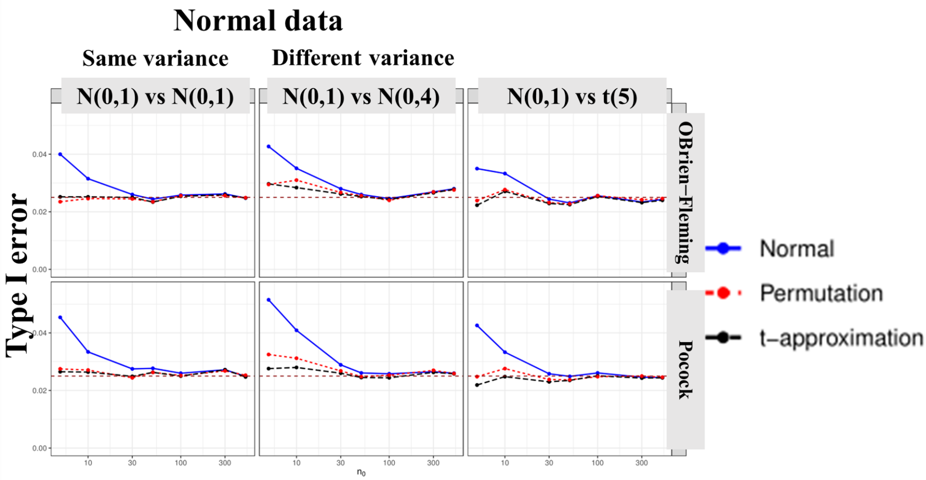
(nperm $= 10000$, nsim $= 10000$, $\alpha = 0.025$)

Here, n_0 is the sample size per treatment at any given stage.

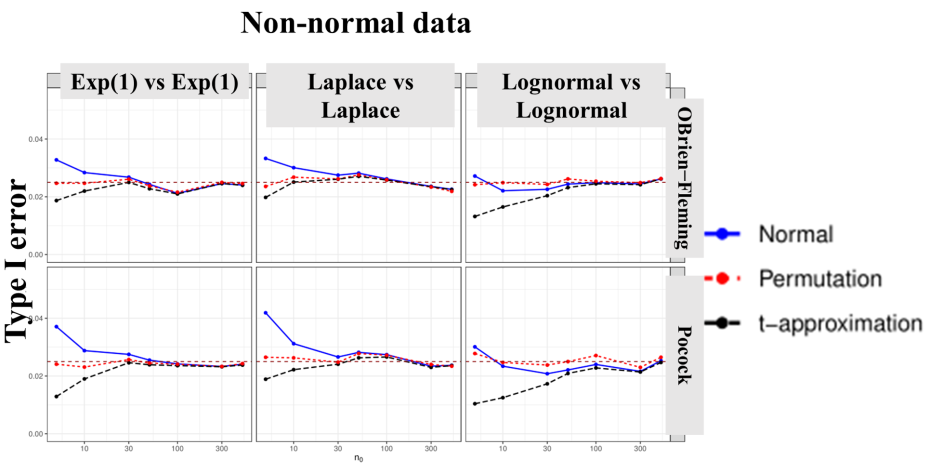
In this presentation, we only focus on **two-stage** group sequential designs.

As a competitor, we include the t-approximation method in our simulation results.

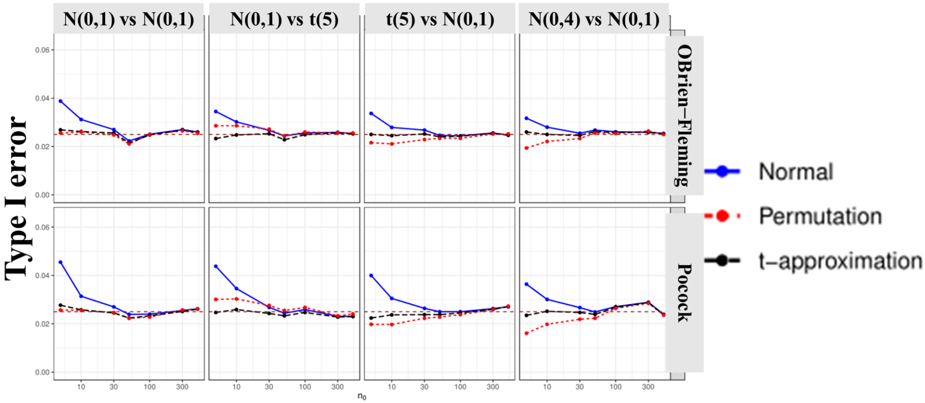
Type I error (1:1 allocation ratio)



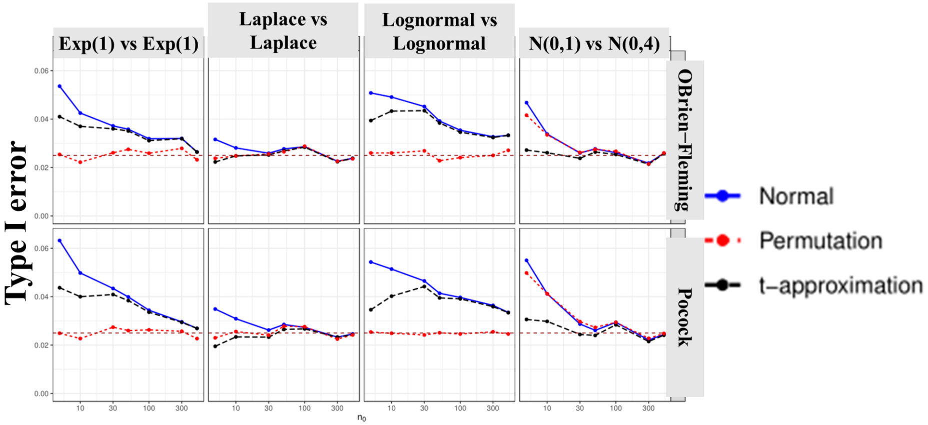
Type I error (1:1 allocation ratio)



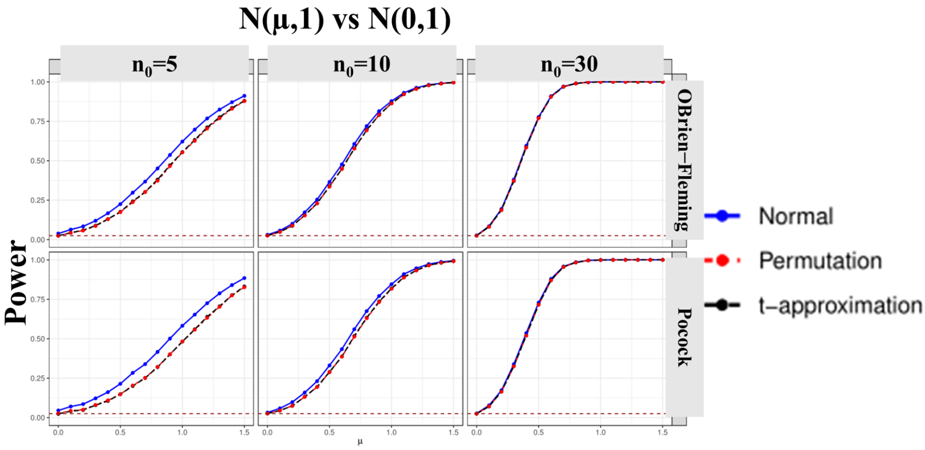
Type I error (2:1 allocation ratio)



Type I error (2:1 allocation ratio)



Power (1:1 allocation ratio)



Outline

- We develop a general robust method for both small sample sizes and large sample sizes, which controls the Type I error rate well for testing the equality of means in group sequential designs.
- For the asymptotic results, we prove Theorem 1 and Theorem 2. We show that the conditional distribution of $\mathbf{S}^{\tau} = (S^{\tau}_{\widetilde{m}_1,\widetilde{n}_1}, \ldots, S^{\tau}_{\widetilde{m}_K,\widetilde{n}_K})'$ given the observed data approximates the unconditional distribution of $\mathbf{S} = (S_{\widetilde{m}_1,\widetilde{n}_1}, \ldots, S_{\widetilde{m}_K,\widetilde{n}_K})'$. In addition, the unconditional test φ_n and the conditional permutation test φ^{τ}_n are asymptotically equivalent, and both tests have asymptotically the same power
- We carry out simulations to investigate finite sample size properties including Type I error rate and power, in particular for small samples. The results for the three-stage and five-stage designs are similar to those for the two-stage design.

Thank you for your attention!