

# Clinical Trial Optimization

## Introduction

Assigning subjects to treatment groups is pivotal in experimental clinical trials for novel drugs. A critical component involves identifying distinct subject cohorts suitable for specific treatments. These cohorts should exhibit maximum similarity based on measurable *covariates* (or attributes), which influence individual responses to treatment, thus enabling the differentiation of treatment effects from confounding factors.

On average, clinical trials for newly approved drugs incur costs of \$1 billion, with a failure rate of 70%. In 2010, global pharmaceutical companies allocated \$32.5 billion to conduct clinical trials out of a total research and development expenditure of \$46.4 billion. The financial and temporal constraints of this complex process requires accurate patient stratification to enhance the likelihood of success.

## 1 Setting

Each patient is associated with  $r = 3$  covariates  $\vec{w} * i = (w_{i1}, w_{i2}, w_{i3})$  that are relevant for predicting the patient's outcome. There are  $n$  patients in total. The decision maker knows the total number of subjects of the experiment  $n$  and the respective covariates  $\vec{w} = (\vec{w}_1, \dots, \vec{w}_n)$ . The decision maker will assign  $k := N/m$  subjects to each of  $m \geq 2$  treatment groups. In this case, we will consider the scenario where only two groups are used ( $m = 2$ ).

The binary decision variables are  $\vec{x} = \{x_{ip} | i = 1, \dots, N, p = 1, 2\}$ , where  $x_{ip} = 1$  if patient  $i$  is assigned to group  $p$ , and  $x_{ip} = 0$  otherwise.

## 2 Objective

The objective of the assignment is to *minimize the discrepancy* between any two groups in the weighted sum of the first and second moments of the covariates. That is, the discrepancy is calculated as follows:

$$d = \sum_{s=1}^3 |\Delta\mu_s| + \rho \sum_{s=1}^3 |\Delta\sigma_{ss}| + 2\rho \sum_{s=1}^3 \sum_{s'=s+1}^3 |\Delta\sigma_{ss'}|$$

where

$$\Delta\mu_s = \frac{1}{n} \sum_{i=1}^n w_{is}(x_{i1} - x_{i2}), \quad \forall s \in \{1, 2, 3\}$$

and

$$\Delta\sigma_{ss'} = \frac{1}{n} \sum_{i=1}^n w_{is} w_{is'} (x_{i1} - x_{i2}), \quad \forall s \in \{1, 2, 3\}, s' \in \{s, \dots, 3\}$$

The parameter  $\rho$  regulates the relative weight of the first and the second moments. In this case, we will set  $\rho = 0.5$ .

### 3 Constraints

There are several constraints that needed to ensure an appropriate patient stratification:

- **Number of patients in each group:** Groups 1 and 2 need to have the same number of patients,  $n/2$ . This constraint can be written as:

$$\sum_i x_{ip} = N/2, \quad \forall p \in \{1, 2\}.$$

- **Each patient should be assigned to exactly one group.** This constraint can be written as:

$$x_{i1} + x_{i2} = 1 \quad \forall i \in \{1, \dots, n\}$$

- **Symmetry redundancy:** this constraint reduces the redundancy as a result of permutation symmetry in group numbering (changing group 1 by group 2). This constraint can be written as:

$$x_{12} = 0$$

### 4 Problem Formulation

Putting it all together, this problem can be formulated as follows:

$$\min_{\vec{x}} \sum_{s=1}^3 |\Delta\mu_s| + \rho \sum_{s=1}^3 |\Delta\sigma_{ss}| + 2\rho \sum_{s=1}^3 \sum_{s'=s+1}^3 |\Delta\sigma_{ss'}| = d$$

subject to

$$\begin{aligned} \sum_i x_{ip} &= n/2, \quad \forall p \in \{1, 2\} \\ x_{i1} + x_{i2} &= 1, \quad \forall i \in \{1, \dots, n\} \\ x_{12} &= 0 \end{aligned}$$

This corresponds to a mixed-integer optimization problem with 10 continuous variables ( $\{d, \Delta\mu_1, \Delta\mu_2, \Delta\mu_3, \Delta\sigma_{11}, \Delta\sigma_{22}, \Delta\sigma_{33}, \Delta\sigma_{12}, \Delta\sigma_{13}, \Delta\sigma_{23}\}$ ) and  $2n - 1$  binary variables ( $\{x_{11}, \dots, x_{n1}, x_{22}, \dots, x_{n2}\}$ ).

Solving this optimization problem classically is challenging (or even impossible!) when the number of patients is large. For this reason, quantum computing can provide significant advantage to solve this kind of problems.