

# Estudo de Caso 3 - Comparação de desempenho de duas configurações de um algoritmo de otimização

Ana Júlia de Lima Martins, Antônio Carlos da Anunciação, Melchior Augusto Syrio de Melo

16 de dezembro de 2024

```
##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:stats':
##
##   filter, lag
##
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
##
## Attaching package: 'MASS'
##
## The following object is masked from 'package:dplyr':
##
##   select
##
## Attaching package: 'TH.data'
##
## The following object is masked from 'package:MASS':
##
##   geyser
```

## Resumo

Este relatório descreve uma análise estatística do desempenho de duas configurações distintas de um algoritmo de otimização baseado em evolução diferencial (DE). Neste experimento, aplicou-se um Experimento Completamente Aleatorizado com Blocos (RCBD) com o intuito de avaliar o desempenho das duas configurações em diferentes instâncias do problema. A análise teve como objetivo identificar se há diferenças estatisticamente significativas no desempenho médio das configurações, determinando qual delas apresenta o melhor desempenho médio e avaliando a magnitude das diferenças encontradas. Os resultados são apresentados no contexto das questões técnicas de interesse, proporcionando recomendações baseadas nos achados experimentais.

## 1. Design do Experimento

A análise será conduzida a partir de um Experimento Completamente Aleatorizado com Blocos (RCBD), em que as instâncias do algoritmo são consideradas os blocos, e as configurações do algoritmo representam os tratamentos (níveis do fator de interesse). O objetivo é verificar se há diferenças estatisticamente significativas no desempenho médio entre as configurações do algoritmo, levando em conta a variabilidade atribuída às instâncias (blocos).

## 1.1. Hipótese

Para esse experimento, estamos interessados em investigar se há alguma diferença no desempenho médio do algoritmo quando equipado com diferentes configurações, para a classe de problemas de interesse. Portanto, para realizar a análise estatística, estabelecemos as seguintes hipóteses:

- **Hipótese Nula ( $H_0$ ):** Não há diferença no desempenho médio entre as configurações, isto é, não existe diferença no tamanho dos efeitos  $\tau_i$ .

$$\left\{ H_0 : \tau_i = 0, \forall i \in \{1, 2, \dots, a\} \right\}$$

- **Hipótese Alternativa ( $H_a$ ):** Existe pelo menos uma configuração que apresenta um efeito significativamente diferente de zero, ou seja, com um desempenho superior.

$$\left\{ H_1 : \exists \tau_i \neq 0 \right\}$$

Se decidirmos pela rejeição da hipótese nula e as premissas do teste forem validadas, precisaremos determinar qual configuração em termos de desempenho médio. Para responder a essa pergunta, pode-se realizar uma comparação todos contra todos, utilizando o teste de Tukey devido a sua sensibilidade superior ao fazer esse tipo de comparação.

## 1.2. Tamanho amostral

### 1.2.1. Estimando o número de instâncias (blocos)

Choosing the sample size, or the number of blocks to run, is an important decision when using an RCBD. Increasing the number of blocks increases the number of replicates and the number of error degrees of freedom, making design more sensitive.

Escolher o tamanho da amostra ou o número de blocos para rodar é uma decisão importante ao usar um Desenho de Bloqueamento Completo Randomizado (RCBD). Aumentar o número de blocos aumenta o número de replicações e os graus de liberdade de erro, tornando o design mais sensível.

Para determinar o número de blocos para o experimento,

O tamanho amostral para este teste pode ser calculado como o menor número inteiro tal que seja igual ou maior do que o poder desejado.

### Hypothesis under Alternative H1

Sob a hipótese alternativa H1, a estatística  $t_0$  segue uma distribuição t não central (Mathews, 2010) com parâmetro de não centralidade:

$$ncp = \frac{(\mu_D - \mu_0)\sqrt{N}}{\hat{\sigma}_\Phi} = \frac{\delta\sqrt{N}}{\hat{\sigma}_\Phi} = d\sqrt{N}.$$

Assumindo uma medida de relevância mínima  $d^* = |\delta^*|/\sigma_\Phi$ , o poder do teste é dado pela integral da distribuição t não central com  $ncp^* = d^*\sqrt{N}$  sobre os valores de  $t_0$  para os quais a hipótese nula  $H_0$  é rejeitada.

O tamanho da amostra para este teste pode ser calculado como o menor inteiro tal que a potência  $\pi^*$  é igual ou maior que a potência desejada. No caso da hipótese alternativa unilateral, temos que:

```
d = 0.5          # Mínima diferença de importância prática (padronizada)
alpha <- 0.05     # Nível de significância
p <- 1 - alpha    # Nível de confiança
```

```

n <- 2
power <- 0
while (power < 0.8)
{
  df <- n - 1      # n - 1 graus de liberdade
  ncp <- d * sqrt(n) # Non-centrality parameter

  power <- pt(qt(p,df), df, ncp=ncp, lower.tail=F)
  n <- n + 1
}
n

```

```
## [1] 28
```

```
power
```

```
## [1] 0.8118316
```

### 1.2.2. Estimando o número de repetições por bloco