

Aplicação de métodos estatísticos para comparação de modelos de otimização

Gustavo Marques Zeferino

SAC SIS 2013

22 de novembro de 2013

Otimização

Conceito

Em matemática, o termo otimização, ou programação matemática, refere-se ao estudo de problemas em que se busca minimizar ou maximizar uma função através da escolha sistemática dos valores de variáveis reais ou inteiras dentro de um conjunto viável.

Tipos de problemas de otimização

■ Otimização contínua

$$\begin{array}{ll}\min_{\mathbf{x}} & f(\mathbf{x}) \\ \text{sujeito a} & g(\mathbf{x})_i \leq 0, \quad i = 1, \dots, m \\ & h(\mathbf{x})_i = 0, \quad i = 1, \dots, p\end{array}$$

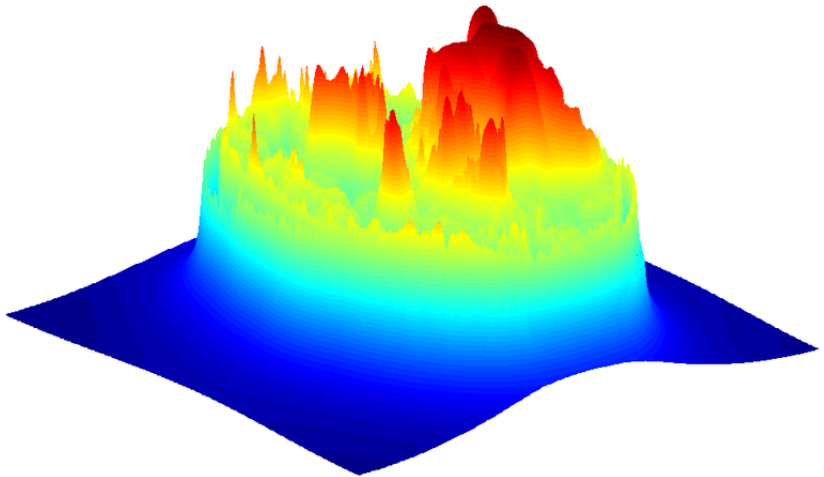
onde

- $f(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$ é a função objetivo
- $g(\mathbf{x})_i \leq 0$ são restrições de desigualdade
- $h(\mathbf{x})_i = 0$ são restrições de igualdade

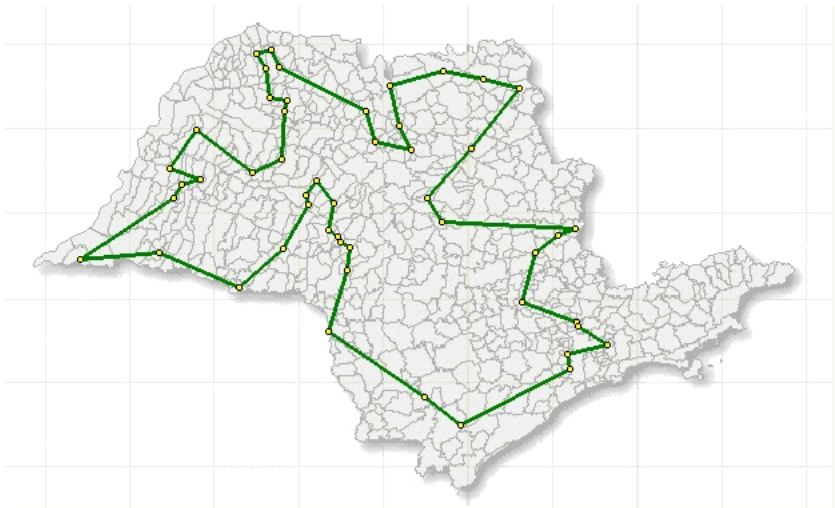
Tipos de problemas de otimização

- Otimização discreta ou combinatória é quádruplo (I, f, m, G) , onde
 - I é um conjunto de instâncias
 - Dada uma instância, $x \in I$, $f(x)$ é um conjunto de soluções viáveis
 - Dada uma instância x e uma solução viável y de x , $m(x, y)$ denota a medida de y , que é usualmente um número real positivo
 - G é a função objetivo, e é tanto mínimo ou máximo.

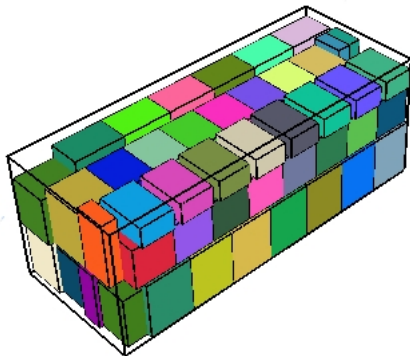
Otimização de uma função contínua



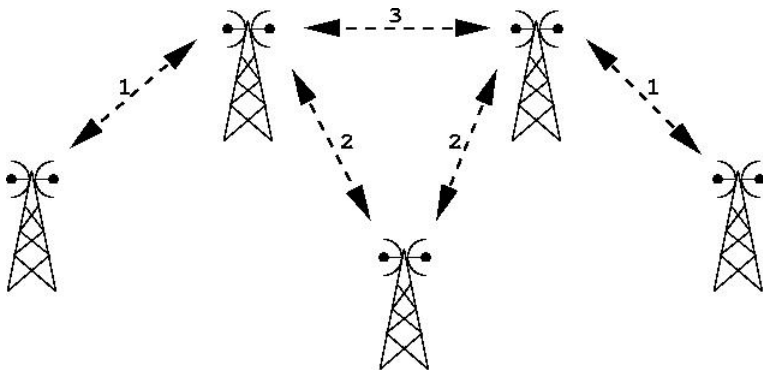
Problema do caixeiro viajante



Problema de corte e empacotamento



Atribuições de Frequências em Telefonia de Celular



Métodos de otimização

■ Métodos determinísticos

- força-bruta
- simplex
- branch & bound
- etc.

■ Métodos estocásticos ou probabilísticos

- algoritmo genético
- colônia de formigas
- enxame de partículas
- simulating annealing
- GRASP
- evolução diferencial
- etc.

Distribuição de probabilidade

Uma solução de um método estocástico representa uma variável aleatória pertencente a uma distribuição de probabilidade

Distribuição de probabilidade

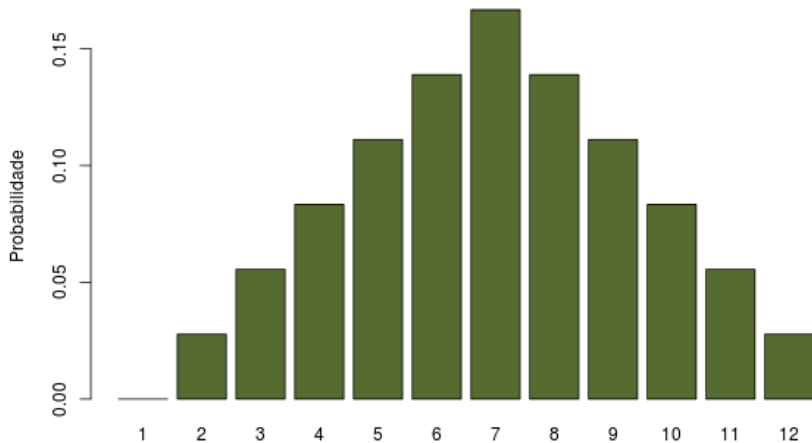
Uma solução de um método estocástico representa uma variável aleatória pertencente a uma distribuição de probabilidade

Distribuição de probabilidade

Uma distribuição de probabilidade descreve a chance de uma variável assumir um valor dentro um espaço de valores possíveis

Exemplo

Soma no lançamento de dois dados



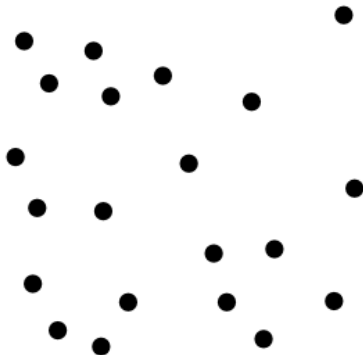
Problema das P-Medias

Descrição do problema

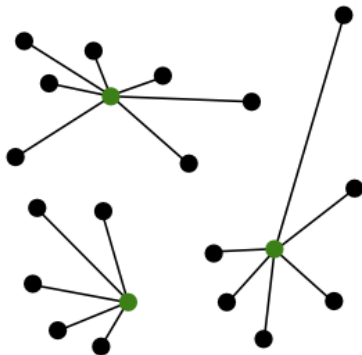
Dado um grafo ponderado:

- escolha p vértices para serem as medianas (fornecedores)
- para vértice restante (cliente), escolha somente um mediana para atendê-lo
- Some os pesos (distâncias) de cada vértice cliente ao seu vértice fornecedor
- Encontre uma solução que minimize a soma dos pesos

Problema das P-Medianas



Problema das P-Medianas



Modelos de otimização

Duas propostas para resolver esse problema:

- (Método 1) Meta-heurística Multi-Start
- (Método 2) Meta-heurística Multi-Start com um operador local aplicado sobre as soluções encontradas

Modelos de otimização

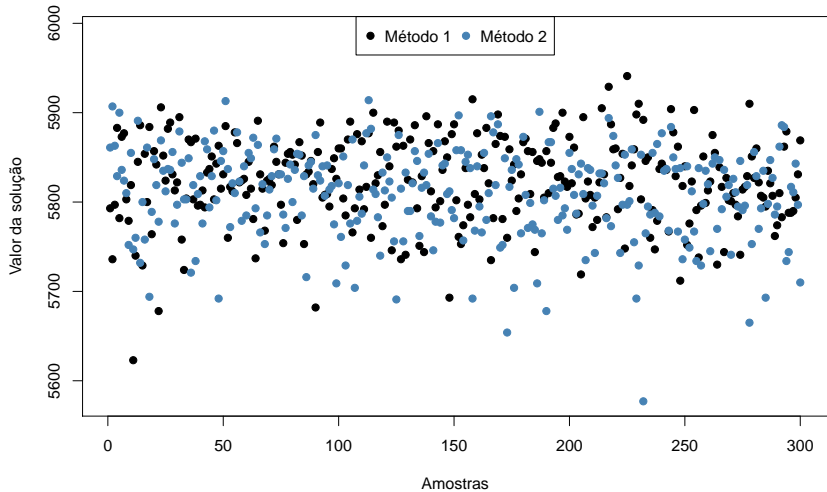
Duas propostas para resolver esse problema:

- (Método 1) Meta-heurística Multi-Start
- (Método 2) Meta-heurística Multi-Start com um operador local aplicado sobre as soluções encontradas

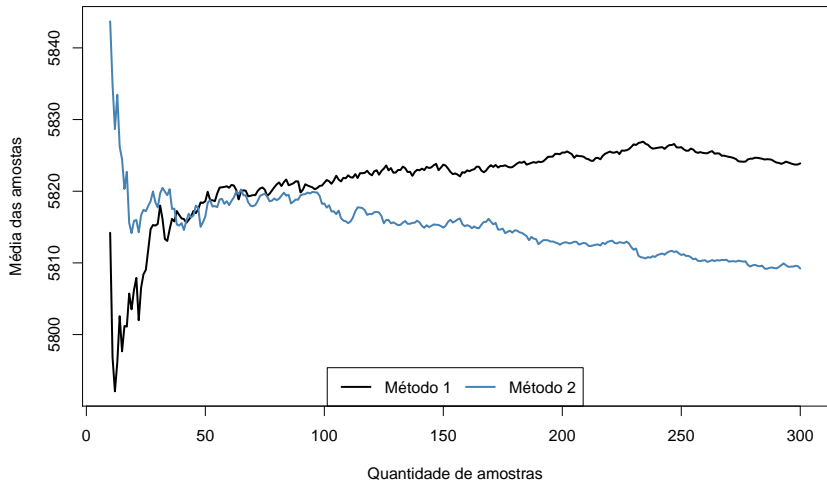
Comparação

- Melhor solução encontrada em até 3 minutos de execução

Modelos de otimização



Modelos de otimização



Modelos de otimização

- Qual o número de amostras a serem utilizadas?
- Como saber se o resultado já é conclusivo?

Estatística

Teorema Central do Limite

Dado um número suficientemente grande de amostras geradas de forma independente, a distribuição da média dessas amostras tende a distribuição normal.

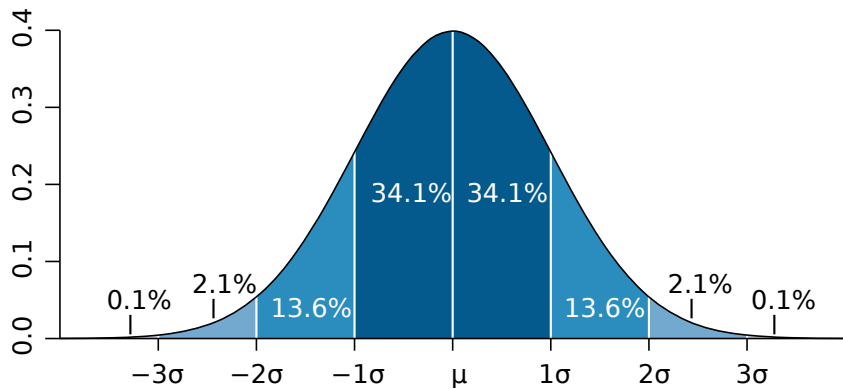
Estatística

Distribuição normal

Distribuição de probabilidade paramétrica definida pelos parâmetros média (μ) e desvio-padrão (σ) com a seguinte função de densidade de probabilidade:

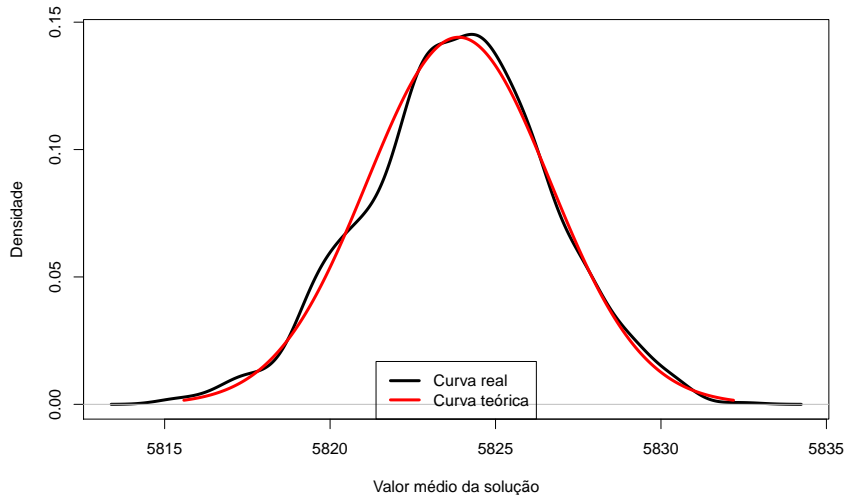
$$f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x - \mu)^2}{2\sigma^2}}$$

Distribuição normal



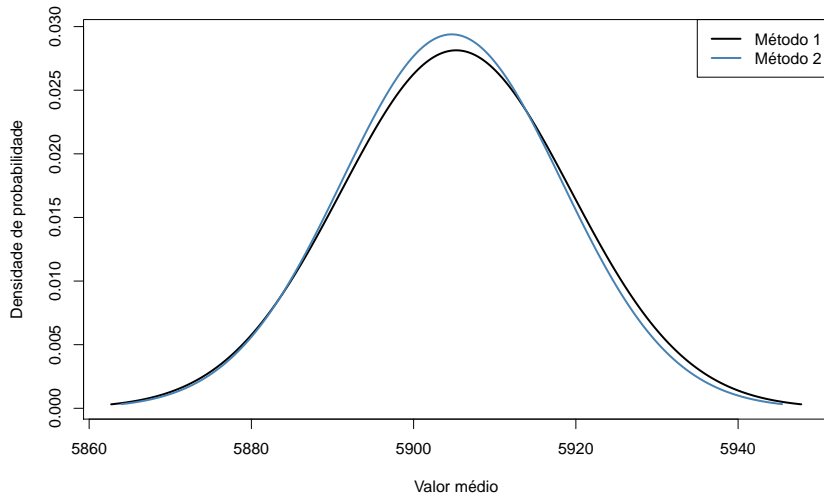
Distribuição normal

Método 1



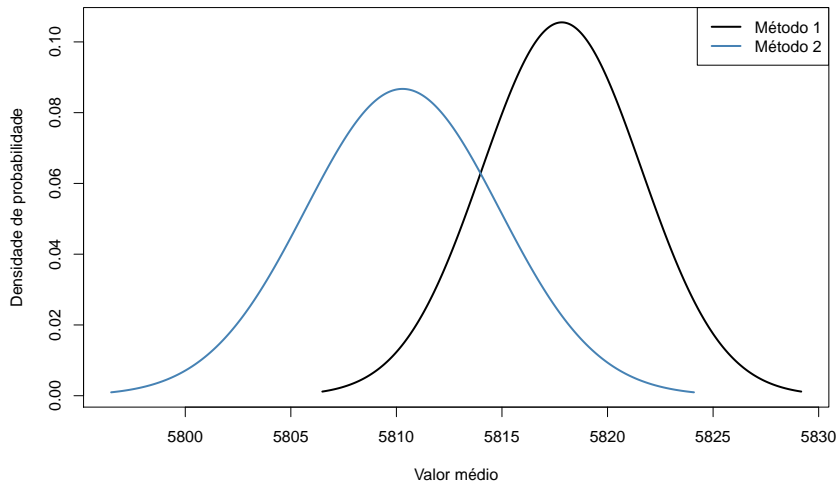
Comparação dos métodos

Distribuição do valor médio das soluções (20 amostras)



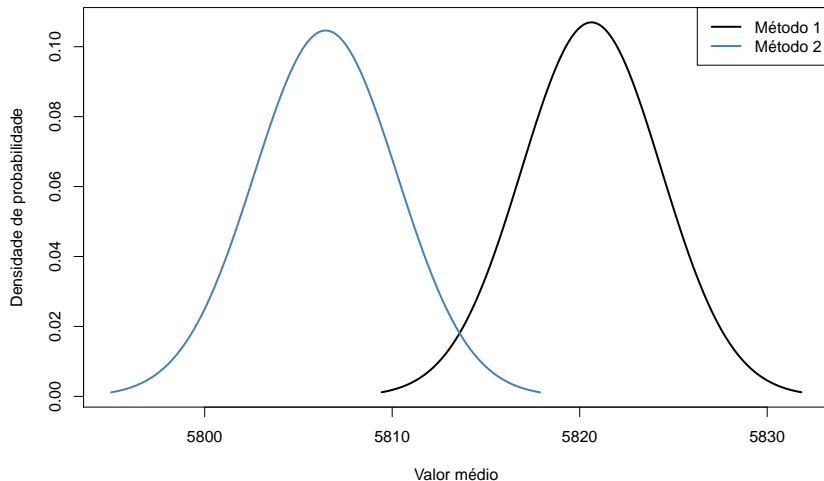
Comparação dos métodos

Distribuição do valor médio das soluções (120 amostras)



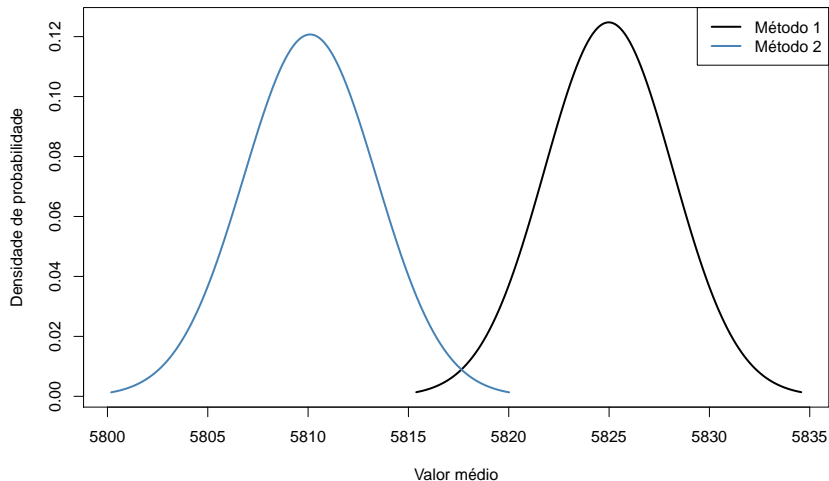
Comparação dos métodos

Distribuição do valor médio das soluções (180 amostras)



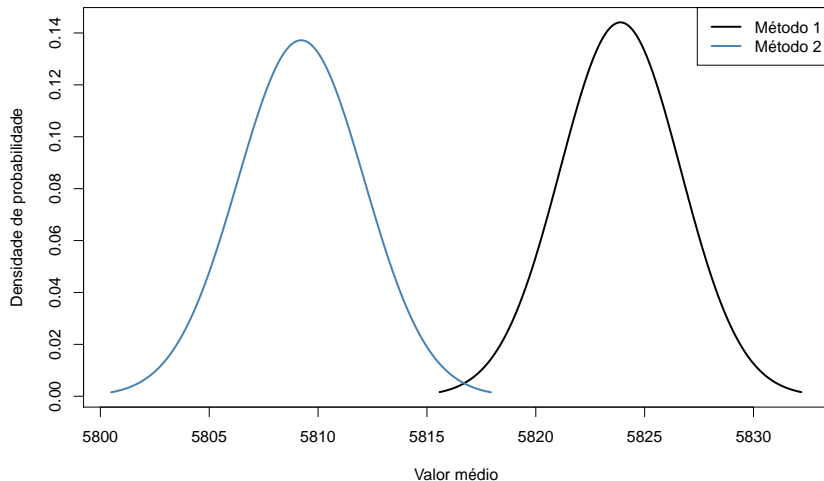
Comparação dos métodos

Distribuição do valor médio das soluções (240 amostras)

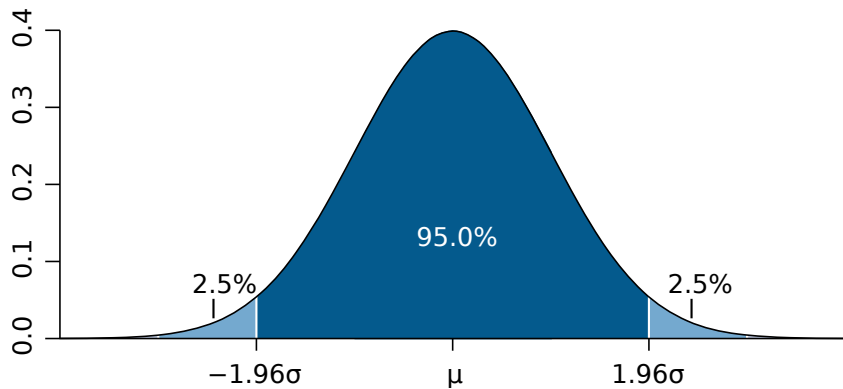


Comparação dos métodos

Distribuição do valor médio das soluções (300 amostras)

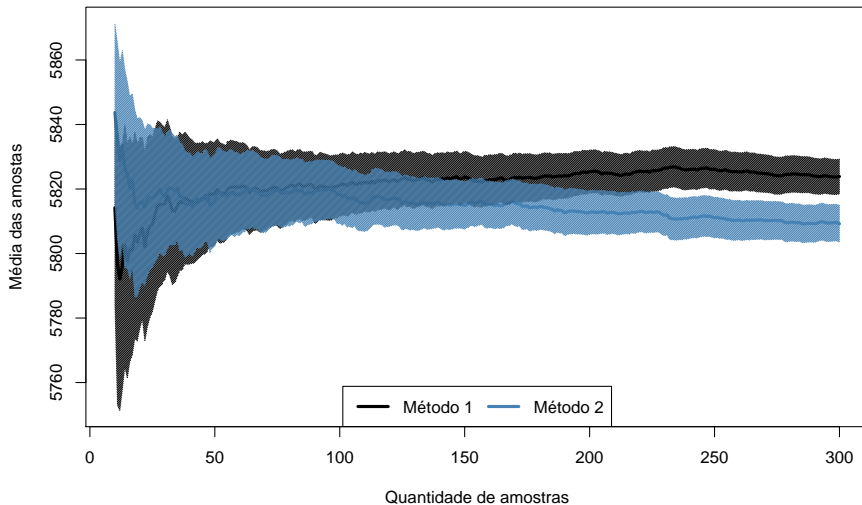


Intervalo de Confiança



Intervalo de Confiança

Intervalo de confiança de 95%



Aplicações Web

- entrada/instâncias: usuários, pageviews
- amostras: interações geradas (cliques, tempo de sessão, download, compras, etc.)
- objetivo: otimizar métricas relacionadas a interação do usuário

KPI - Key Performance Indicator

Exemplo de métricas:

- **CTR (Click-through rate)**: taxa de ocorrência de cliques
- **RPU (revenue per user)**: receita por usuário
- **Installations / New Users**: número de vezes que um aplicativo é instalado em um dispositivo ou browser. A métrica é utilizada como indicativo do número de novos usuários um aplicativo pode atrair dado um intervalo de tempo (dia, semana, mês, etc.)
- **Sessions Per User**: Compara o número média de sessões um usuário gera por unidade de tempo. Exmplos: número médio de vezes que um usuário interage com um jogo, média de pesquisas feitas pelo usuário num site de buscas.

Como realizar otimização em aplicações web

- Cada vez que um usuário interage com a aplicação gera uma amostra única

Como realizar otimização em aplicações web

- Cada vez que um usuário interage com a aplicação gera uma amostra única
- Não é possível aplicar dois modelos ao mesmo usuário ao mesmo tempo.

Como realizar otimização em aplicações web

- Cada vez que um usuário interage com a aplicação gera uma amostra única
- Não é possível aplicar dois modelos ao mesmo usuário ao mesmo tempo.
- Os modelos precisam ser testados ao mesmo tempo

Teste A/B



Teste A/B

- Separe de forma aleatória o tráfego do site entre duas versões
 - **Controle:** Versão atual do site
 - **Experimento:** Versão nova a ser testada
- Colete métricas de interesse (KPI's)
- Análise estatística dos dados coletados

Casos de uso Teste A/B

- Amazon
 - Recomendações personalizadas
 - a cada 100ms de lentidão do site, o impacto negativo nas vendas chega a 100k
- Google: cerca de 20k experimentos controlados por ano
- Microsoft: realiza testes controlados até para encontrar erros no backend

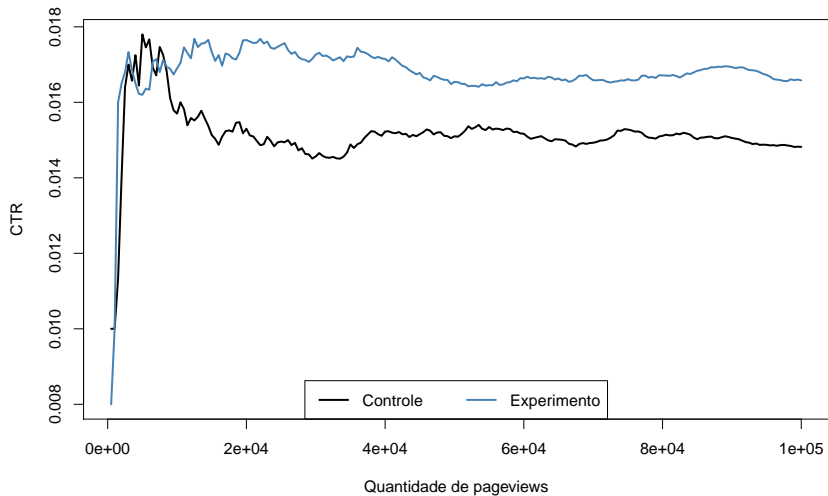
História dos experimentos aleatórios controlados

- Marinha inglesa (1700)
- Tratamento médico via sangramento (1836)
- Higiene hospitalar (1800)

Otimização estocástica

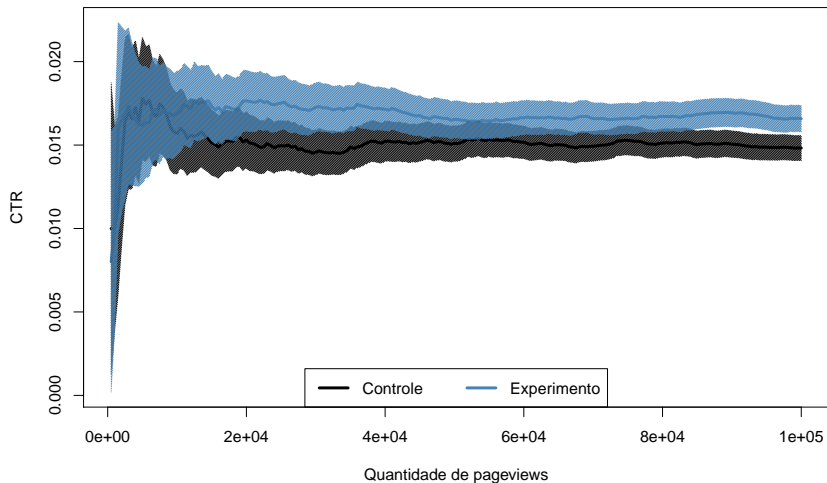
- usuários diferentes geram interações diferentes
- um mesmo usuário gera interações diferentes ao longo do tempo
- métricas possuem uma distribuição de probabilidade

Exemplo CTR 50/50



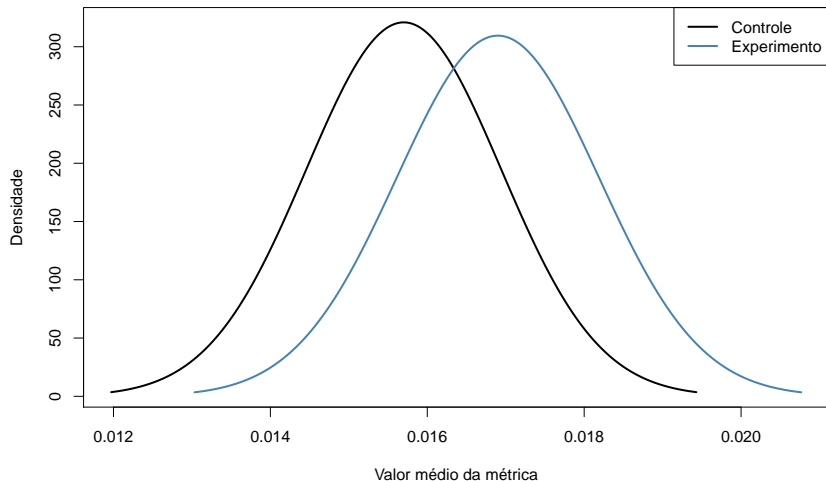
Exemplo CTR 50/50

Intervalo de confiança de 95%



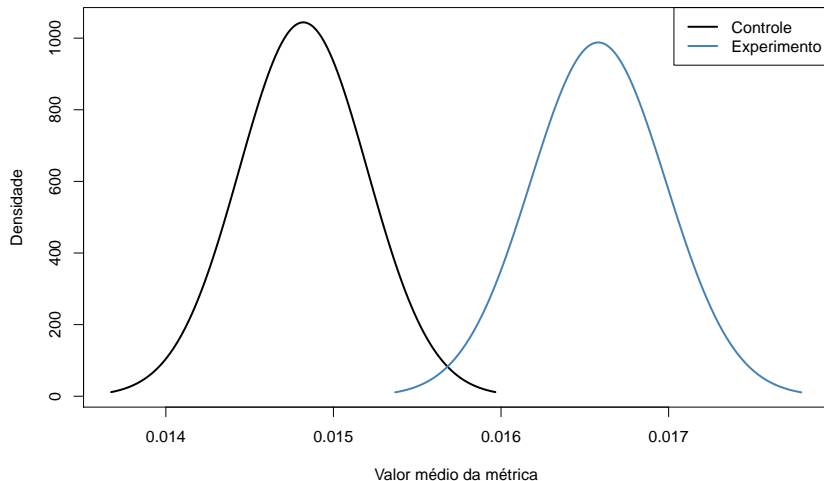
Exemplo CTR 50/50

Distribuição da métrica (10.000 amostras)



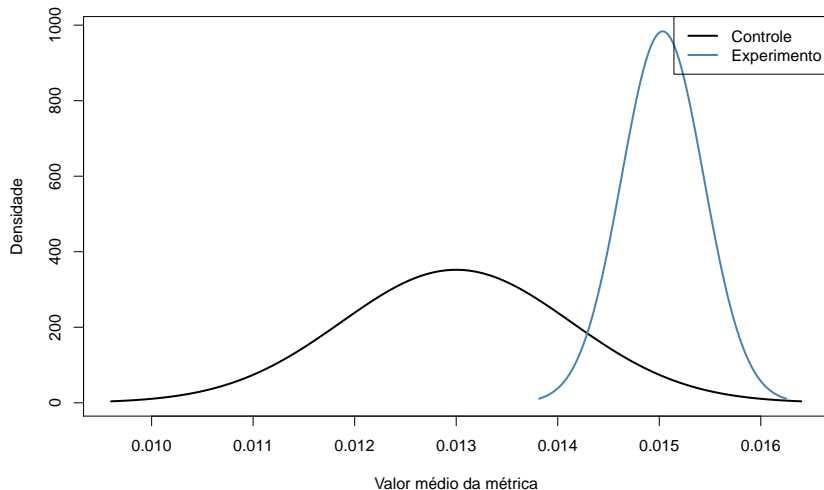
Exemplo CTR 50/50

Distribuição da métrica (100.000 amostras)

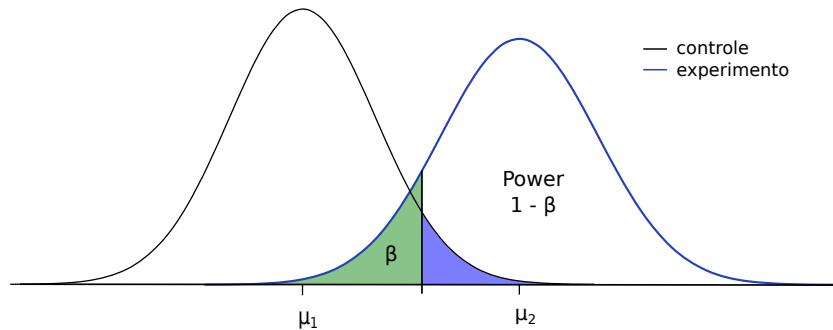


Exemplo CTR 10/90

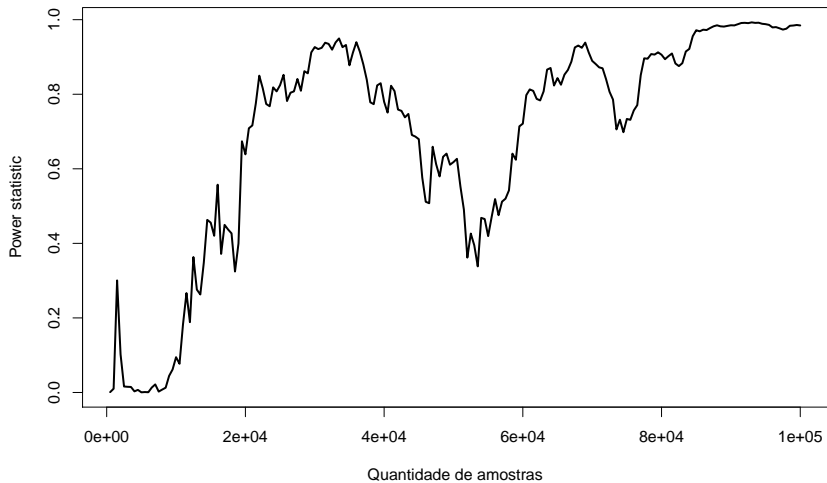
Distribuição da métrica (10% controle)



Power statistic



Power statistic



Outros métodos

- Teste de hipótese
- Bandit arms

Melhores práticas

- Defina de forma clara qual o objetivo do experimento (aumentar vendas, melhorar usabilidade, etc) e defina as métricas baseado nisso.
- Desconsidere informação gerada por bots (> 50% do tráfego do Bing)
- Comece novos experimentos com uma parte pequena do tráfego (até 5%)
- Quanto mais amostras, maior a confiança dos testes
- Não economize em teste e comece o mais rápido possível: mate as idéias ruins rapidamente e teste várias idéias para encontrar logo a que melhora a métrica em questão (cerca de 70% dos experimentos não resultam em melhorias)

Perguntas?

Obrigado!

Gustavo M. Zeferino
gumaze@gmail.com