

# Inferência por abordagens de Monte Carlo

Fundamentos e aplicações

Prof. Walmes Zeviani  
walmes@ufpr.br

Laboratório de Estatística e Geoinformação  
Departamento de Estatística  
Universidade Federal do Paraná

Atualizado em 2018-10-10

## Justificativas

- ▶ Métodos computacionalmente intensivos para inferência estatística são usados quando as abordagens tradicionais não são adequadas.
  - ▶ Resultados assintóticos em pequenas amostras.
  - ▶ Violação de pressupostos.
  - ▶ Não existência de mecanismos de inferência específicos.
- ▶ Tais métodos se baseiam em reamostragem e/ou simulação.
- ▶ Podem ser aplicados em muitos contextos.

## Objetivos

- ▶ Definir o que são os métodos de Monte Carlo.
- ▶ Apresentar aplicações.

# Questões históricas

- ▶ Foi um método introduzido durante a II Guerra Mundial.
- ▶ No projeto de construção da bomba atômica, Stanisław Ulam, von Neumann e Fermi consideraram a possibilidade de utilizar o método, que envolvia a simulação direta de problemas de natureza probabilística relacionados com o coeficiente de difusão do neutron em certos materiais.
- ▶ O nome Monte Carlo é uma referência ao resort town (cassino) de Mônaco por causa da natureza aleatória da abordagem.

# O que são métodos de Monte de Carlo

- ▶ Corresponde a métodos baseados em simulação estocástica massiva para:
  - ▶ Aproximação de funções e de integrais.
  - ▶ Estimação de valores médios ou obtenção de distribuições amostrais.
  - ▶ Avaliação de propriedades de um estimador pontual/intervalar.
  - ▶ Avaliação de propriedades de testes de hipótese.
  - ▶ Determinação de tamanhos amostrais e curvas de poder.
  - ▶ Dentre outras várias aplicações.
- ▶ Em outras palavras, envolvem a simulação de experimentos ou sistemas em que pelo menos um componente aleatório esteja presente (Ferreira 2013).
- ▶ É imprescindível, portanto, recursos para geração de números aleatórios das distribuições envolvidas no problema.
- ▶ Métodos de bootstrap e de testes de aleatorização são casos particulares de Monte Carlo.

# Exemplo: avaliar a taxa de erro tipo I de um teste de hipótese

1. Gerar uma realização do modelo assumido para os dados sob a hipótese nula.
2. Sob a hipótese nula, delimitar a região de aceitação e rejeição considerando nível de significância  $\alpha$ .
3. Aplicar o teste sob os dados simulados.
4. A partir do teste, tomar a decisão correspondente e guardar o resultado.
5. Repetir de 1 a 4  $M$  vezes.
6. Calcular a proporção de vezes em foi feita a rejeição da hipótese nula.
7. Se a proporção for superior a  $\alpha$ , o teste é liberal, caso contrário é conservador, para o nível de significância adotado.

## Exemplo: avaliar a taxa de cobertura de um estimador intervalar

1. Gerar uma realização do modelo assumido para os dados.
2. Definir o nível de confiança  $1 - \alpha$  para obtenção dos intervalos de confiança.
3. Determinar o intervalo de confiança com os dados simulados.
4. Verificar se o intervalo construído contém o verdadeiro valor do parâmetro usado para simular os dados e guardar o resultado.
5. Repetir de 1 a 4  $M$  vezes.
6. Calcular a proporção de vezes em que o intervalo conteve o valor do parâmetro.
7. Se a proporção for superior a  $1 - \alpha$ , o intervalo é conservador, caso contrário é liberal, para o nível de confiança adotado.

# Exemplo: construção da curva de poder do teste

Curva de poder sob um tamanho de amostra fixo.

1. Defina um conjunto de valores  $\Delta = \{0, \delta, 2\delta, \dots, k\delta\}$  em que 0 corresponde à hipótese nula  $H_0 : \theta = \theta_0$ . Os demais correspondem a incrementos  $\delta$  em  $\theta_0$  e fazem o afastamento da  $H_0$  com relação ao parâmetro sob hipótese  $\theta$ .
2. Sob  $H_0$ , delimitar a região de aceitação e rejeição considerando nível de significância  $\alpha$ .
3. Para cada valor em  $\Delta$ , definir  $\theta_a = \theta_0 + \Delta_i$  ( $i = 0, \dots, k$ ), então
  - 3.1 Gerar uma realização do modelo assumido para os dados usando  $\theta_a$ .
  - 3.2 Aplicar o teste para  $H_0$  com os dados simulados.
  - 3.3 A partir do teste, tomar a decisão correspondente e guardar o resultado.
  - 3.4 Repetir de 1 a 3  $M$  vezes.
  - 3.5 Calcular a proporção de vezes em foi feita a rejeição de  $H_0$ .
4. Fazer o gráfico da taxa de rejeição de  $H_0$  para cada valor em  $\Delta$ .

# Considerações até aqui

- ▶ Nos métodos de Monte Carlo deve-se assumir distribuição de probabilidades para as variáveis aleatórias do modelo/sistema.
- ▶ Essa é uma das principais desvantagens, conforme Ferreira 2013.





## Próxima aula

- ▶ Testes de hipótese Monte Carlo.
- ▶ Comparação de testes concorrentes.
- ▶ Avaliação do desempenho de testes com fuga dos pressupostos.
- ▶ Outras aplicações.

## Avisos

- ▶ Sabatina estará disponível a partir de Amanhã às 19h.

# Referências bibliográficas

Ferreira, D. F. (2013). *Estatística Computacional em Java*. Editora UFLA.

