PCC170 - Projeto e Análise de Experimentos Computacionais

Marco Antonio M. Carvalho

Departamento de Computação
Instituto de Ciências Exatas e Biológicas
Universidade Federal de Ouro Preto

Conteúdo

Fonte

Este material é parcialmente baseado no conteúdo de:

- Felipe Campelo (2018), Lecture Notes on Design and Analysis of Experiments. Online: `http://git.io/v3Kh8` Version 2.12; Creative Commons BY-NC-SA 4.0.
- Marcelo Menezes Reis. Conceitos elementares de estatística. Departamento de Informática e Estatística, Universidade Federal de Santa Catarina, 2003. Disponível em: `https://bityli.com/vrgmc`.
- Fernanda Peres. Estatística aplicada à vida real. 2022. Disponível em: `https://bityli.com/fMepl`.

Licença

População

*Uma **população** é um grande conjunto de objetos de natureza semelhante que é de interesse como um todo.*[a]

Pode ser um conjunto real (por exemplo, de pessoas) ou hipotético (por exemplo, todos os resultados possíveis para um experimento).

[a] Glossary of statistical terms.

Amostra

Uma **amostra** é um subconjunto de uma população.

> *Uma amostra é escolhida para fazer inferências sobre a população examinando ou medindo os elementos da amostra.*[a]

[a] Glossary of statistical terms.

Revisão de conceitos estatísticos

fonte: `https://bityli.com/fMepl`

Amostra

Por que tomamos uma amostra? Por que não usamos a população toda?

- Custo alto para obter informação da população toda.
- Tempo muito longo para obter informação da população toda, ou algumas vezes impossível, e.g., estudo de poluição atmosférica.
- Algumas vezes é logicamente impossível, e.g., em ensaios destrutivos.

Observação

Uma **observação** é um elemento único de uma determinada amostra, um ponto de dados coletados individualmente.

Uma observação também pode ser considerada como uma amostra de tamanho um.

Revisão de conceitos estatísticos

Inferência estatística

A **inferência estatística** utiliza amostras para chegar a conclusões sobre populações.

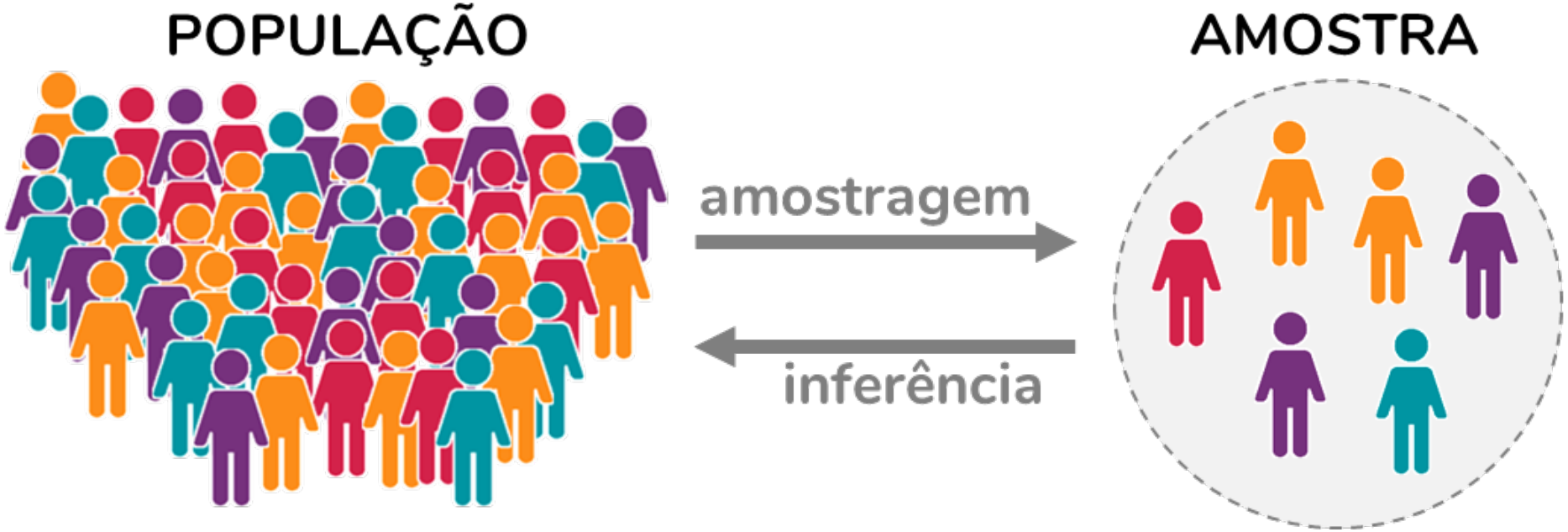
Probabilidade

Dada a população de pessoas, quais são as chances de obter uma certa combinação de alturas?

Estatística

Dadas as alturas de algumas pessoas sorteadas, o que posso saber sobre a população?

Revisão de conceitos estatísticos



fonte: `https://bityli.com/fMepl`

Revisão de conceitos estatísticos

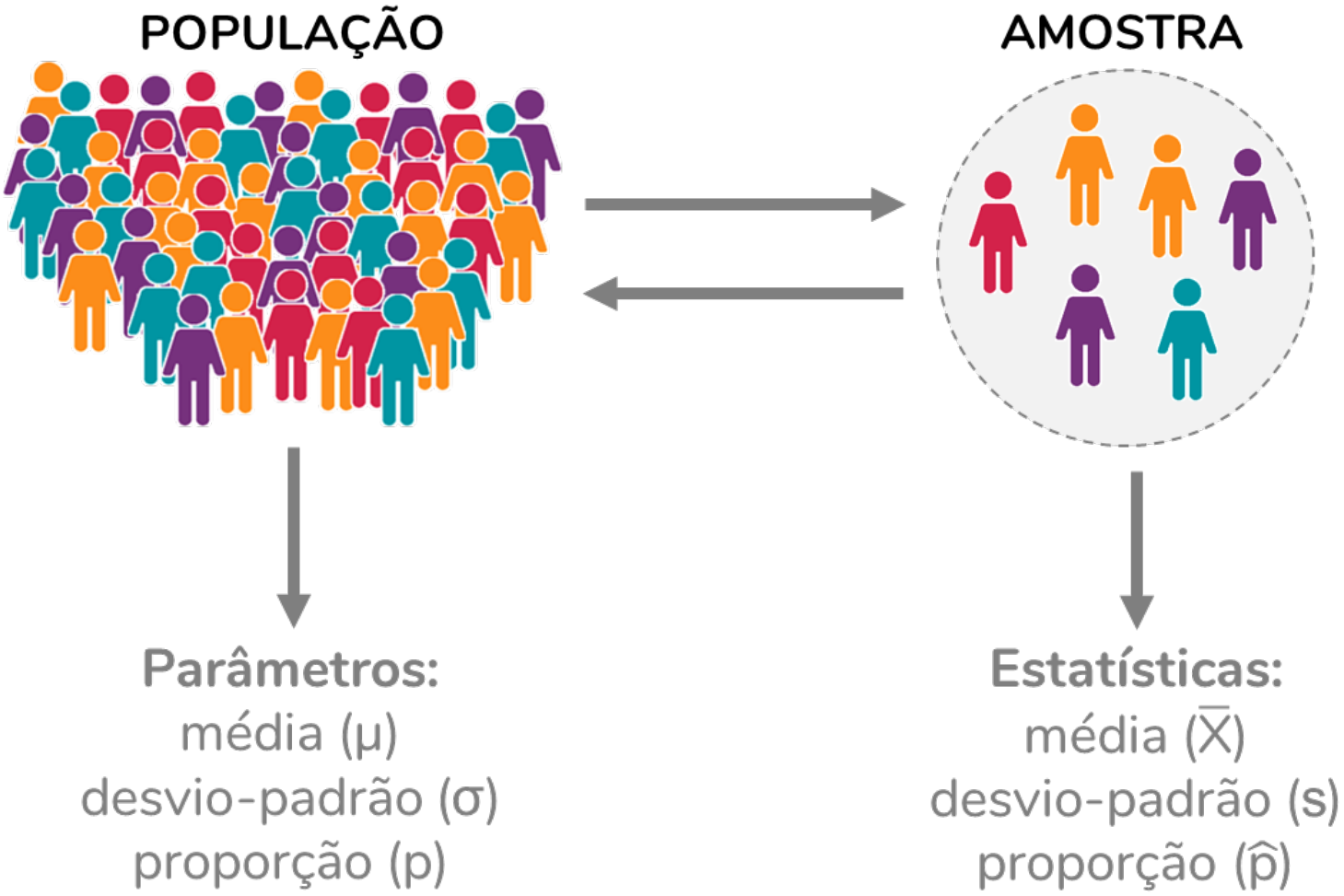
Parâmetros

As informações extraídas da população, como média, desvio padrão, proporções, são chamadas de **parâmetros** e são, em sua maioria, representadas por letras gregas.

Estatísticas

As informações provenientes de uma amostra recebem o nome de **estatísticas**, e são representadas por letras do nosso alfabeto, o romano.

Revisão de conceitos estatísticos



fonte: `https://bityli.com/fMepl`

Média

É uma medida de tendência central dos dados, sensível a dados *outliers*.

Seja n o número total de valores x_i, $1 \leq i \leq n$. A média aritmética é dada por

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Mediana

É uma medida de tendência central dos dados. Indica o valor central dos dados, uma alternativa robusta à média.

A mediana de uma lista finita de números pode ser encontrada ordenando os números do menor para o maior.

Se houver um número ímpar de elementos, o número do meio é o valor do meio $\frac{n+1}{2}$.

Se houver um número par de elementos, a mediana é definida como a média dos dois valores do meio $\frac{n+1}{2}$ e $\frac{n}{2}$.

Variância

É uma medida de variabilidade. A variância representa o desvio quadrado médio em relação a média dos valores dos dados.

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

Desvio padrão

É uma medida de variabilidade. Indica o desvio médio dos valores, nos dados, em relação ao valor médio geral.

$$\sigma = \sqrt[2]{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2}$$

Inferência estatística

Dois dos conceitos centrais da inferência estatística são as estimativas e os intervalos estatísticos.

Ambos os termos referem-se ao uso de informações obtidas de uma amostra para inferir valores prováveis sobre parâmetros populacionais.

Estimativa: valor estimado para um determinado parâmetro populacional.

Intervalo estatístico: intervalo estimado de valores possíveis/prováveis para um determinado parâmetro populacional.

Estimativa

Uma estimativa é uma estatística que fornece o valor de máxima plausibilidade para um dado parâmetro populacional, desconhecido.

Utilizamos estimativas de uma amostra como nosso "melhor chute" para os verdadeiros valores populacionais.

Exemplos são a média amostral, o desvio padrão amostral e a mediana amostral, os quais estimam a verdadeira média, desvio padrão e mediana da população, que são desconhecidos.

Estimativa

Problemas de estimativa surgem com frequência em todas as áreas da ciência e engenharia, sempre que há necessidade de estimar, por exemplo:

- Uma média populacional;
- Uma variância populacional;
- Uma proporção da população;
- **A diferença nas médias de duas populações**;
- Etc.

Variáveis aleatórias

Suponha que queremos obter uma estimativa para um parâmetro arbitrário, e.g., a média de uma dada população.

A amostragem aleatória de uma população resulta em uma **variável aleatória**, e qualquer função dessas observações, i.e., qualquer estatística, é uma variável aleatória também.

Sendo variáveis aleatórias, as estatísticas também têm suas próprias distribuições de probabilidade, chamadas de **distribuições de amostragem**.

Intervalo estatístico

Os intervalos estatísticos definem regiões que provavelmente conterão o valor real de uma quantidade estimada.

Tais intervalos são usados para quantificar a incerteza associada a uma determinada estimativa, permitindo a derivação de declarações em níveis de confiança quantificáveis.

Reportar intervalos é sempre melhor do que reportar estimativas pontuais, pois fornecem as informações necessárias para quantificar a localização e a incerteza de seus valores estimados.

Intervalo de confiança

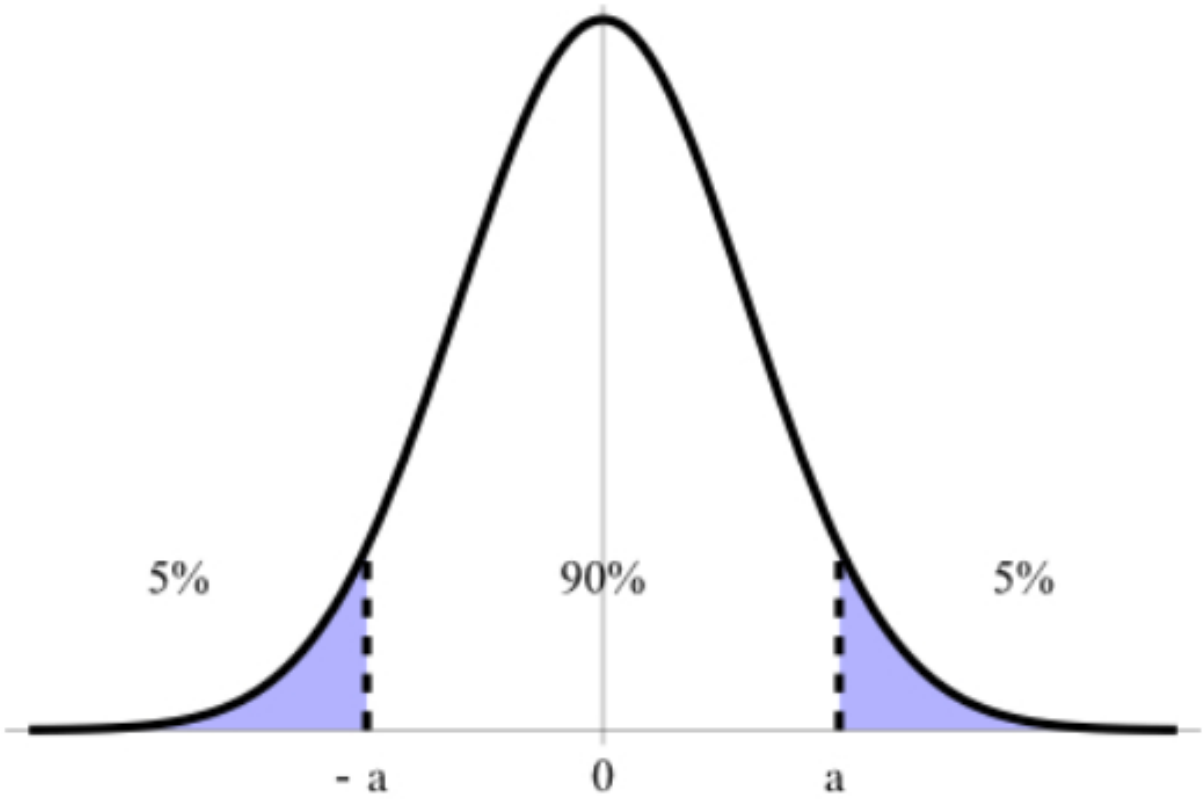
Os intervalos de confiança quantificam o grau de incerteza associado à estimativa de parâmetros populacionais, como a média ou a variância.

Uma definição útil é pensar nos intervalos de confiança em termos de confiança no método:

"O método usado para derivar o intervalo tem uma taxa de acerto de 95%"

Em outras palavras, o intervalo gerado tem 95% de chance de 'capturar' o parâmetro verdadeiro da população.

Revisão de conceitos estatísticos



Intervalo de Confiança de 90% entre seus limites inferior (-a) e superior (a).
fonte: `https://bityli.com/ngAuf`

Intervalo de confiança

Suponha que estejamos interessados num parâmetro populacional verdadeiro, e desconhecido, θ.

Podemos estimar o parâmetro θ usando informação de nossa amostra.

Contudo, sabemos que o valor estimado, em parte das vezes, não será exatamente igual ao valor verdadeiro.

Seria interessante encontrar um intervalo de confiança que forneça um intervalo de valores plausíveis para o parâmetro baseado nos dados amostrais.

Intervalo de confiança

Um intervalo de confiança de 95% para um parâmetro populacional fornece um intervalo no qual estaríamos 95% confiantes de cobertura do verdadeiro valor do parâmetro.

Tecnicamente, 95% de todos os intervalos de confiança que construirmos conterão o verdadeiro valor do parâmetro, dado que todas as suposições envolvidas estejam corretas.

Se obtivermos um intervalo de confiança para o parâmetro θ para cada uma dentre 100 amostras aleatórias da população, somente 5, em média, destes intervalos de confiança não conterão θ.

Intervalo de confiança

Podemos obter intervalos de confiança de 95% para:

- Médias;
- **Diferenças de médias;**
- Diferenças em proporções;
- Etc.

Podemos também criar intervalos de confiança de 90%, 99%, 99.9%, etc, mas os intervalos de confiança de 95% são os mais utilizados.

Comparação de duas populações

Frequentemente, temos interesse em coletar informações sobre duas populações para compará-las.

Assim como na inferência estatística para um parâmetro populacional, intervalos de confiança e testes de significância são ferramentas estatísticas úteis para a diferença entre dois parâmetros populacionais.

Na comparação de algoritmos, isto nos fornece uma idéia de quanta diferença existe entre os resultados de dois ou mais métodos diferentes.

Amostras independentes

Em **amostras estatisticamente independentes**, temos duas amostras e cada observação de uma amostra não possui relacionamento direto com as observações da outra amostra.

Amostras independentes são selecionadas aleatoriamente para que as suas observações não dependam dos valores de outras observações.

Cada amostra contém observações diferentes e não há uma maneira significativa de emparelhá-las.

Por exemplo, um estudo de medicação que possui um grupo de controle e um grupo de tratamento que contém diferentes indivíduos.

Amostras dependentes

Em **amostras estatisticamente dependentes**, as observações de uma amostra fornecem informações sobre as observações de outra amostra

As amostras contêm o mesmo conjunto de observações ou contêm observações diferentes, que emparelharam de forma significativa.

As amostras são frequentemente dependentes porque contêm as mesmas observações – esse é o exemplo mais comum.

No entanto, outros estudos usam observações **pareadas**.

Amostras pareadas

Em **amostras pareadas**, temos duas amostras, mas cada observação da primeira amostra é pareada com uma observação da segunda amostra.

Nesses estudos, os pesquisadores deliberadamente emparelham observações com características muito semelhantes.

Embora os pares sejam compostos por observações diferentes, a análise estatística as trata como a mesma porque são intencionalmente muito semelhantes.

Como resultados de métodos diferentes para uma mesma instância.

Diferença entre amostras dependentes e independentes

As amostras dependentes são medições para **um mesmo conjunto de itens**.

As amostras independentes são medições feitas em **dois conjuntos de itens diferentes**.

Quando realizamos um teste de hipóteses usando duas amostras, é necessário escolher o tipo de teste dependendo de se as amostras são dependentes ou independentes.

Comparação de duas populações

Quando temos **amostras independentes** de cada uma de duas populações, podemos sumarizá-las pelas suas médias e desvios padrão.

Sejam estas medidas denotadas por $\bar{x}_1$, s_1 para a amostra um e $\bar{x}_2$, s_2 para a amostra dois.

Sejam as correspondentes médias populacionais e desvios padrão denotados por μ_1, μ_2, σ_1 e σ_2 respectivamente.

Comparação de duas populações

Uma estimativa natural da diferença entre médias na população, $\mu_1 - \mu_2$, é dada pela diferença nas médias amostrais:

$$\bar{x}_1 - \bar{x}_2,$$

É necessária uma medida do erro padrão para esta estimativa para que seja possível construir um intervalo de confiança ou realizar um teste de hipóteses.

Comparação de duas populações

Com **dados pareados**, consideramos a seguinte notação:

$$\begin{aligned} x_{1i} &= \text{medida 1 no par } i, \\ x_{2i} &= \text{medida 2 no par } i \end{aligned}$$

E então escrevemos as diferenças nas medidas de cada par como

$$d_i = x_{2i} - x_{1i}.$$

Comparação de duas populações

Temos então uma amostra de diferenças d_i, e podemos usar os métodos estatísticos que conhecemos.

Podemos calcular um intervalo de confiança para a diferença média e testar se a diferença média é igual a um valor específico, e.g. zero, ou não.

Note que, tendo duas amostras pareadas, estamos interessados na **diferença média** enquanto que, tendo duas amostras independentes, estamos interessados na **diferença nas médias**.

Ainda que numericamente estas quantidades sejam as mesmas, conceitualmente elas são diferentes.

Comparação de duas populações

Por si só, a diferença média não diz muito, além de fornecer um número para a diferença.

O número pode ser **estatisticamente significativo**, ou pode ser apenas devido a variações aleatórias ou ao acaso.

Para testar a hipótese de que os resultados podem ser significativos, realizamos um **teste de hipóteses** para diferenças entre médias.

Comparação de duas populações

É importante discernimos entre significância estatística e significância prática.

Um efeito pode ser estatisticamente significante mas não ter qualquer importância prática e vice-versa.

Por exemplo, um estudo muito grande pode estimar a diferença entre a média de peso de plantas como sendo 0,0001 gramas e concluir que a diferença é estatisticamente significativa.

Contudo, na prática, esta diferença é desprezível e provavelmente de pouca importância real.

O mesmo pode ocorrer ao compararmos algoritmos em experimentos computacionais.

Dúvidas?