

Estatística Básica

Introdução à Análise Exploratória de Dados

Renato Dourado Maia

Instituto de Ciências Agrárias

Universidade Federal de Minas Gerais



Pergunta Inicial

O que é Estatística?

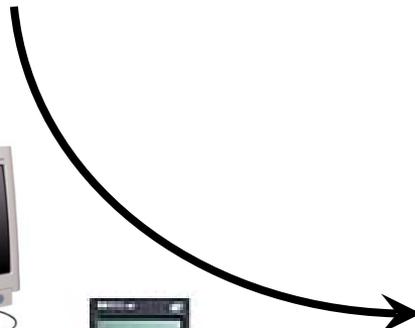


Contexto Moderno

Grande
Quantidade de
Informações



Quantificação
em nossas
vidas



**FERRAMENTAS
COMPUTACIONAIS**

Vantagens?

Perigos?

Ferramentas Computacionais

- ▣ As ferramentas computacionais permitem que grandes quantidades de informações sejam rápida e confortavelmente analisadas, contribuindo positivamente para a difusão e utilização de métodos estatísticos.
- ▣ Todavia, a automação pode levar um indivíduo sem preparo específico a utilizar técnicas inadequadas para analisar e/ou resolver um dado problema.

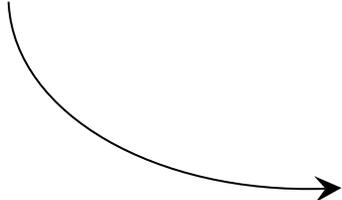
Conclusão Importante

Ensino, Aprendizagem e a Vida Universitária: um Bate-Papo com os Acadêmicos



Conclusão Importante

- É necessária a **compreensão** dos **conceitos** básicos de estatística, bem como das **suposições** básicas **necessárias** para a sua utilização de forma **criteriosa!**



Uma base teórica sólida permite a criação e o entendimento de coisas práticas: só se consegue aplicar na prática a teoria que realmente se entende!

Lamento de um Estudante:

- Se eu tivesse somente um dia de vida, gostaria de ficar numa aula de estatística. Assim o tempo passaria mais vagarosamente.

Conclusão Importante



AS ESTATÍSTICAS SÃO COMO UM BIQUINI: O QUE É REVELADO É INTERESSANTE; O QUE ESTÁ OCULTO É ESSENCIAL.



Usa-se por vezes a Estatística como um bêbedo usa um poste de luz: Mais para suporte do que para iluminação...

Conclusão Importante

- Lei de Griffin na Estatística:
 - Estatísticas são a maneira mais **lógica** e **precisa** de falar uma **meia-verdade** sem **exatidão**.

- Princípio de James:
 - Não há **mentira maior** do que uma **verdade mal interpretada**.

- **Torture** os **dados** por um tempo suficiente, e eles **contam tudo!**

Mas e a Primeira Pergunta?

- O que é Estatística?
 - Conjunto de técnicas que permite, de forma sistemática, organizar, descrever, analisar e interpretar dados oriundos de estudos ou experimentos, realizados em qualquer área do conhecimento.
 - **Dados**: um ou mais conjuntos de valores, numéricos ou não.

Para que Estudar Estatística?

- A Estatística permite descrever e compreender **relações** entre **variáveis** de forma imediata: a informação é apresentada de modo a possibilitar uma **rápida interpretação** e identificação das **relações** mais importantes.
- A Estatística permite também tomar **melhores decisões**, e de forma mais **rápida**: é possível controlar **mais informações** num espaço de **tempo** mais curto.

Para que Estudar Estatística?

- A Estatística **facilita** a tomada de decisões: o conhecimento de situações passadas e presentes, acompanhado por uma **previsão fundamentada** da **evolução futura**, é a **base** para as tomadas de decisões.

ALGUÉM ACREDITA QUE O ESTUDO DA ESTATÍSTICA
NÃO É IMPORTANTE PARA A SUA FORMAÇÃO?



Contexto e Abstração

- Tem-se afirmado que **mais de 70%** dos estudantes de uma grande universidade particular são contrários a um plano de aumentar as taxas de alunos para permitir a construção de novos estacionamentos. **Se 15 de 18** alunos **selecionados ao acaso** naquela universidade se opõem ao plano, teste a afirmação ao nível **0,05** de significância.

Contexto e Abstração

- Tem-se afirmado que mais de 70% das laranjeiras num município paulista foram severamente danificadas por uma geada recente. Se 15 de 18 laranjeiras selecionadas ao acaso naquele município foram severamente danificadas por aquela geada, teste a afirmação ao nível 0,05 de significância.

Contexto e Abstração

- Tem-se afirmado que **mais de 70%** de certos aviões apresentam fissuras em seus lemes de direção devidas à fadiga do metal. **Se 15 de 18** desses aviões **selecionados ao acaso** apresentam fissuras em seus lemes de direção devidas à fadiga do metal, **teste a afirmação ao nível 0,05 de significância**.

Contexto e Abstração

- Tem-se afirmado que **mais de 70%** de todos médicos associados a planos de saúde estão insatisfeitos com seus honorários. **Se 15 de 18** médicos **escolhidos ao acaso** dentre os associados a planos de saúde estão insatisfeitos com seus honorários, **teste a afirmação ao nível 0,05 de significância**.

Contexto e Abstração – Reflexão

- ▣ Os livros **básicos** que pretendem ensinar a jogar xadrez são **normalmente** escritos na seguinte ordem. **Inicialmente**, expõem-se o **objetivo do jogo**, os **movimentos das peças** e os **xequeres elementares**. Assim, aprende-se logo a dar xeque-mate com dama e rei contra rei ou torre e rei contra rei. Em seguida, apresentam-se **finais um pouco mais complexos**, em que cada jogador possui algumas peças; por exemplo, torre, peão e rei contra torre e rei.

Contexto e Abstração – Reflexão

- Agora, o jogador deve aprender a, partindo dessa situação, cair num daqueles finais simples que já aprendeu a ganhar. No exemplo dado, o jogador com peão a mais pode vencer se ele conseguir promover o tal peão, transformando-o numa dama. Posteriormente, discutem-se posições mais complicadas, com mais peças. A idéia é sempre a mesma: mostrar ao jogador como simplificar a posição, de modo a cair num daqueles finais com poucas peças que ele já aprendeu a vencer.

Contexto e Abstração – Reflexão

- Somente nas últimas páginas de um livro didático sobre xadrez é que se apresentam as aberturas, ou seja, sugestões para os primeiros lances do jogo. A abertura é mais difícil de se entender. Seus conceitos são estratégicos; seus frutos só aparecem depois de várias jogadas. Assim, aprende-se a jogar xadrez pelo final, numa ordem contrária àquela na qual o jogo se desenvolve. Entretanto, **DIDATICAMENTE**, essa é a ordem mais lógica para se ensinar o jogo.

MONTEIRO, L. H. A. **Sistemas Dinâmicos**. 2. ed. São Paulo: Editora Livraria da Física, 2006. (Trecho do prefácio da primeira edição)

Contexto e Abstração – Reflexão



Divisão Simplificada da Estatística

- A Estatística pode ser dividida em três áreas principais:
 - Estatística Descritiva.
 - Probabilidade.
 - Inferência Estatística.

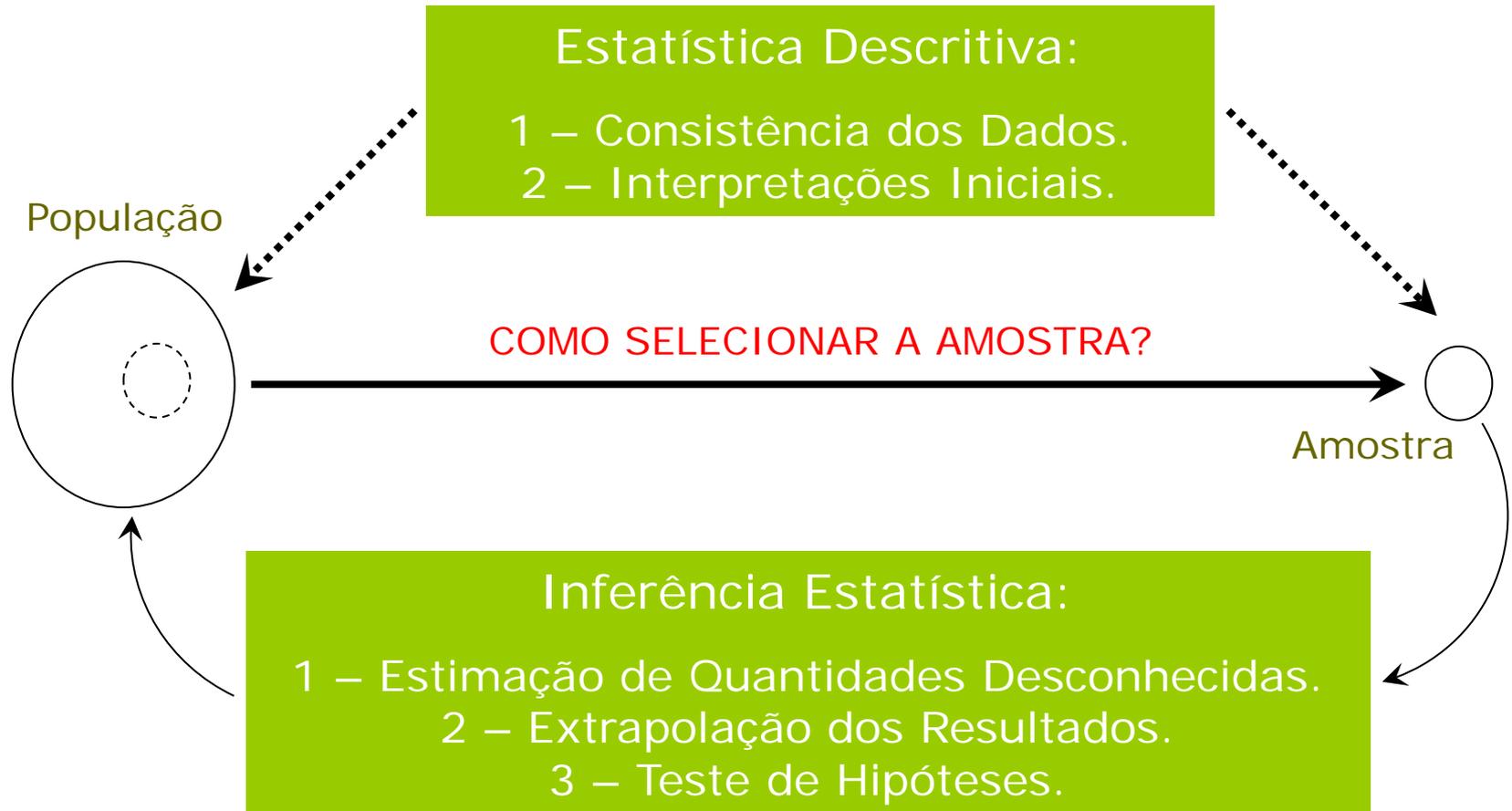
Divisão Simplificada da Estatística

- A **Estatística Descritiva**, que é em geral utilizada na etapa inicial da análise, tem como objetivo de aplicar alguma técnica para **resumir** a informação daquele particular conjunto de valores, para que possam ser tiradas **conclusões** a respeito de **características de interesse**.
- A **Probabilidade** pode ser pensada como a teoria matemática utilizada para se estudar a **incerteza** de fenômenos de caráter **aleatório**.

Divisão Simplificada da Estatística

- A **Inferência Estatística** é o estudo de técnicas que possibilitam a **extrapolação**, a um **grande** conjunto de dados, das informações e conclusões obtidas a partir de **subconjuntos** de valores, usualmente de **dimensão muito menor**.

Divisão Simplificada da Estatística



População e Amostra

- Ao **grande** conjunto de dados que contém a característica de **interesse** dá-se o nome de **População**.
- Ao **subconjunto** da **População**, em geral com dimensão sensivelmente **menor**, que será utilizado para as análises, dá-se o nome de **amostra**.

NA PRÓXIMA AULA, DISCUTIREMOS ALGUNS PROCEDIMENTOS QUE PODEM SER UTILIZADOS PARA ORGANIZAR E DESCRIVER UM CONJUNTO DE DADOS, SEJA EM UMA POPULAÇÃO, SEJA EM UMA AMOSTRA...

Importante

- ❑ As explicações realizadas em sala **nunca substituem uma boa leitura!** Recomendo a todos desenvolver o hábito de estudar por **um material formal de referência**, que apresente o conteúdo utilizando a **linguagem apropriada** para o domínio em questão.
- ❑ Sempre que possível, além dos materiais de apoio, disponibilizarei na página da disciplina **links interessantes e materiais complementares!**

Exercício

1. Classifique em verdadeiro ou falso:
 - a. Estatística é um conjunto de técnicas destinadas a organizar um conjunto de valores numéricos.
 - b. Sempre que estivermos trabalhando com números, deveremos utilizar a Inferência Estatística
 - c. A Estatística Descritiva fornece uma maneira adequada de tratar um conjunto de valores, numéricos ou não, com a finalidade de conhecermos o fenômeno de interesse.
 - d. Qualquer amostra representa, de forma adequada, uma população.
 - e. As técnicas estatísticas não são adequadas para casos que envolvam experimentos destrutivos, como, por exemplo, queima de equipamentos, destruição de corpos de provas, etc.