

Tratamento estatístico para pesquisa acadêmica

Juliana Cobre

Instituto de Ciências Matemáticas e de Computação - ICMC

Universidade de São Paulo - USP



Informação não é conhecimento, mas não há conhecimento sem informação.

Pedro Luis do Nascimento Silva - IBGE

Informação não é conhecimento, mas não há conhecimento sem informação.

Pedro Luis do Nascimento Silva - IBGE

Dados

Informação não é conhecimento, mas não há conhecimento sem informação.

Pedro Luis do Nascimento Silva - IBGE

Dados

- ▶ Censos

Informação não é conhecimento, mas não há conhecimento sem informação.

Pedro Luis do Nascimento Silva - IBGE

Dados

- ▶ Censos
- ▶ Pesquisas amostrais

Informação não é conhecimento, mas não há conhecimento sem informação.

Pedro Luis do Nascimento Silva - IBGE

Dados

- Censos
- Pesquisas amostrais
- Registros administrativos

Informação não é conhecimento, mas não há conhecimento sem informação.

Pedro Luis do Nascimento Silva - IBGE

Dados

- Censos
- Pesquisas amostrais
- Registros administrativos
- Dados "orgânicos"

Informação não é conhecimento, mas não há conhecimento sem informação.

Pedro Luis do Nascimento Silva - IBGE

Dados

- Censos
- Pesquisas amostrais
- Registros administrativos
- Dados "orgânicos"
- Big data

Estatística

- ▶ Planeja a amostra de dados.
- ▶ Extrai conhecimento útil dos dados.
- ▶ Apoia tomada de decisões.

Estatística

- ▶ Planeja a amostra de dados.
- ▶ Extrai conhecimento útil dos dados.
- ▶ Apoia tomada de decisões.

Problema real →

- ▶ Planeja a amostra de dados.
- ▶ Extrai conhecimento útil dos dados.
- ▶ Apoia tomada de decisões.

Problema real → Coletar dados →

Estatística

- ▶ Planeja a amostra de dados.
- ▶ Extrai conhecimento útil dos dados.
- ▶ Apoia tomada de decisões.

Problema real → Coletar dados → Formular perguntas →
→

- ▶ Planeja a amostra de dados.
- ▶ Extrai conhecimento útil dos dados.
- ▶ Apoia tomada de decisões.

Problema real → Coletar dados → Formular perguntas →
→ Explorar, resumir e analisar os dados →

- ▶ Planeja a amostra de dados.
- ▶ Extrai conhecimento útil dos dados.
- ▶ Apoia tomada de decisões.

Problema real → Coletar dados → Formular perguntas →
→ Explorar, resumir e analisar os dados → Responder perguntas

Problemas

- ▶ Quanto espero lucrar com a venda do produto em um ano?

Problemas

- ▶ Quanto espero lucrar com a venda do produto em um ano?
- ▶ Qual o preço ideal de uma passagem aérea?

Problemas

- Quanto espero lucrar com a venda do produto em um ano?
- Qual o preço ideal de uma passagem aérea?
- Qual o risco de o cliente ser inadimplente?

Problemas

- Quanto espero lucrar com a venda do produto em um ano?
- Qual o preço ideal de uma passagem aérea?
- Qual o risco de o cliente ser inadimplente?
- O medicamento é eficaz?

Problemas

- ▶ Quanto espero lucrar com a venda do produto em um ano?
- ▶ Qual o preço ideal de uma passagem aérea?
- ▶ Qual o risco de o cliente ser inadimplente?
- ▶ O medicamento é eficaz?
- ▶ A metodologia desenvolvida atinge o objetivo?

Problemas

- Quanto espero lucrar com a venda do produto em um ano?
- Qual o preço ideal de uma passagem aérea?
- Qual o risco de o cliente ser inadimplente?
- O medicamento é eficaz?
- A metodologia desenvolvida atinge o objetivo?
- Os usuários estão satisfeitos com o novo sistema?

Problemas

- Quanto espero lucrar com a venda do produto em um ano?
- Qual o preço ideal de uma passagem aérea?
- Qual o risco de o cliente ser inadimplente?
- O medicamento é eficaz?
- A metodologia desenvolvida atinge o objetivo?
- Os usuários estão satisfeitos com o novo sistema?
- Experimentos investigativos nas aulas de química motivam os alunos?

Problemas

- Quanto espero lucrar com a venda do produto em um ano?
- Qual o preço ideal de uma passagem aérea?
- Qual o risco de o cliente ser inadimplente?
- O medicamento é eficaz?
- A metodologia desenvolvida atinge o objetivo?
- Os usuários estão satisfeitos com o novo sistema?
- Experimentos investigativos nas aulas de química motivam os alunos?
- Qual a percepção da população a respeito do *stalking*?

Estatística para não estatísticos

O que se aprende?

- Introdução à Teoria das Probabilidades.
- Análise descritiva.
- Introdução à Inferência.

Necessário: Base matemática.

Estatística para não estatísticos

O que se aprende?

- Introdução à Teoria das Probabilidades.
- Análise descritiva.
- Introdução à Inferência.

Qual o objetivo?

- Conhecer a base da Estatística.
- Aprender a importância/interpretação da visualização de dados.
- Aprender conceitos, linguagem, interpretação da estatística.
- Realizar inferência.

Necessário: Base matemática.

Inferência

- ▶ Estimação de parâmetros
 - média e variância de distribuição normal.
 - proporção (modelos binomial, geométrico, binomial negativo).
 - regressão linear simples.

- ▶ Teste de hipóteses
 - média e variância de população com distribuição normal (uma ou duas populações).
 - proporção (uma ou duas populações).
 - média e variância de uma população qualquer.
 - aderência.
 - parâmetros da regressão linear simples.

Inferência

- ▶ Estimação de parâmetros
 - média e variância de distribuição normal.
 - proporção (modelos binomial, geométrico, binomial negativo).
 - regressão linear simples.

- ▶ Teste de hipóteses
 - média e variância de população com distribuição normal (uma ou duas populações).
 - proporção (uma ou duas populações).
 - média e variância de uma população qualquer.
 - aderência.
 - parâmetros da regressão linear simples.

Estatística para não estatísticos

Qual o problema?

- ▶ Alguns desses testes são resultados assintóticos, ou seja, valem para uma amostra suficientemente grande.
- ▶ Falha nas suposições \Rightarrow f alta de garantia do resultado.
- ▶ Falha nas conclusões e falhas nas tomadas de decisão.

A Estatística falhou?

Estatística para não estatísticos

Qual o problema?

- ▶ Alguns desses testes são resultados assintóticos, ou seja, valem para uma amostra suficientemente grande.
- ▶ Falha nas suposições \Rightarrow f alta de garantia do resultado.
- ▶ Falha nas conclusões e falhas nas tomadas de decisão.

A Estatística falhou?

Solução

- ▶ Testes não-paramétricos.
- ▶ Inferência bayesiana.
- ▶ Entre outras.

Estatística na pesquisa acadêmica

Questionários

- ▶ Comitê de ética.
- ▶ Itens abertos (respostas em forma de texto).
 - Diferentes grafias.
 - Dificuldade em extrair informação.
 - Necessidade de grandes conjuntos de dados.
- ▶ Itens fechados (respostas em alternativas).
 - Quantas alternativas?
 - Permitir mais de uma alternativa?
- ▶ Quem responderá o questionário?
 - Amostra?

Experimentos

- ▶ Comitê de ética.
- ▶ Planejamento (do experimento).
- ▶ Amostra.

Estatística na pesquisa acadêmica

Tamanho amostral

- Um dos passos iniciais da (pré-)análise estatística.
- É algo que pode comprometer os resultados.
- Há dificuldade de colocar em prática o que é ideal.

Amostra

- É algo que pode comprometer os resultados.
- Há dificuldade de colocar em prática o que é ideal.
- Devemos trabalhar dentro da realidade com cautela.

Núcleo de Estatística Aplicada (NEA)

O que é? É um Projeto de Extensão vinculado ao Bach. em Estatística e Ciência de Dados (BECD) do ICMC/USP.

Objetivos: aprendizado e lapidação das competências dos Estatísticos em formação no BECD do ICMC/USP através da análise de projetos da sociedade.

Como funciona? Projeto é inscrito. Se aprovado é encaminhado para análise, e uma equipe de aprendizes, supervisionada por um professor, realiza a análise.

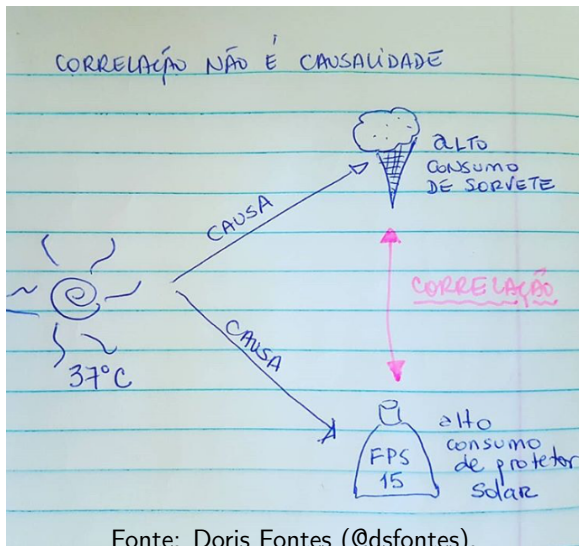
Condição: Os dados ficam disponíveis para fins didáticos e de pesquisa, seguindo os termos de confidencialidade quando necessário.

Mais informações

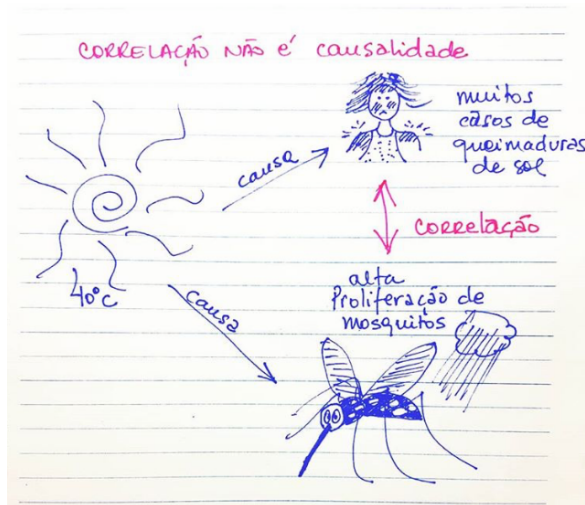
www.nea.icmc.usp.br

Submeta seu projeto!

Alguns exemplos



Alguns exemplos



Fonte: Doris Fontes (@dsfontes).

Alguns exemplos

Variável de confundimento: está associada a um evento e uma exposição

Alguns exemplos

Variável de confundimento: está associada a um evento e uma exposição

Exemplo

- ▶ Evento: Câncer de boca.
- ▶ Exposição: beber exageradamente.
- ▶ Variável de confundimento: Fumo?

Alguns exemplos

Variável de confundimento: está associada a um evento e uma exposição

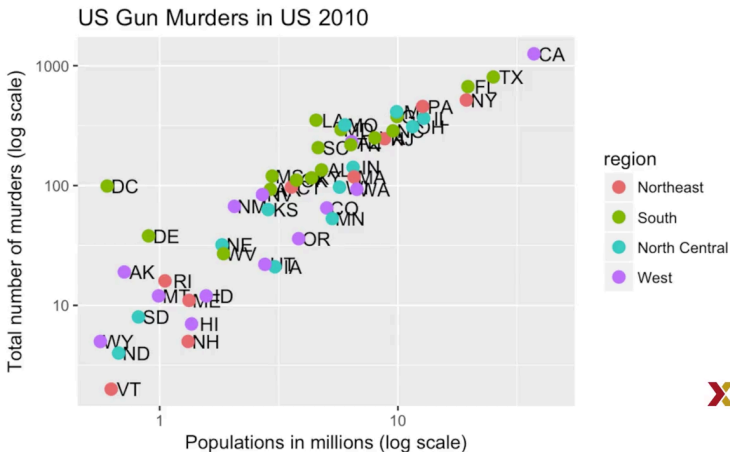
Exemplo

- ▶ Evento: Câncer de boca.
- ▶ Exposição: beber exageradamente.
- ▶ Variável de confundimento: Fumo?

Exemplo

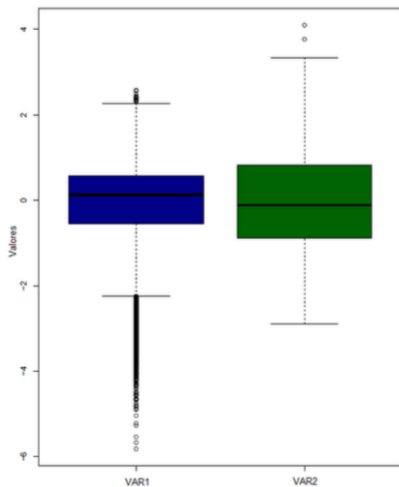
- ▶ Evento: Diminuição da fertilidade feminina.
- ▶ Exposição: consumo exagerado de café.
- ▶ Variável de confundimento: Fumo? Idade? Estresse?

Alguns exemplos

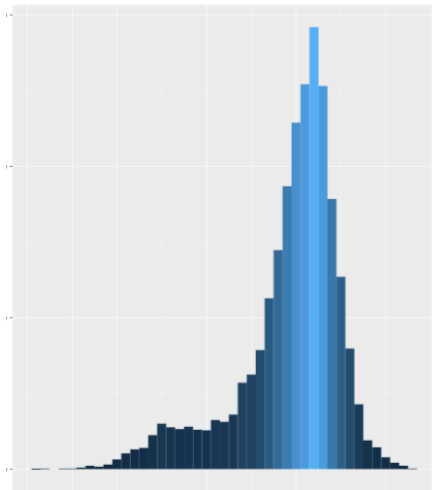


Fonte: Harvard X.

Alguns exemplos



Alguns exemplos: TCC Vitor Correa Yoshida - 2018/2



Analisar dados requer uma profunda investigação!

jucobre@icmc.usp.br ou jucobre@gmail.com

jucobre.com

Muito obrigada!