

# Testes de hipótese em Modelos Multivariados de Covariância Linear Generalizada (McGLM)

Uma visão geral

Lineu Alberto Cavazani de Freitas  
lineuacf@gmail.com

PPG Informática  
Data Science & Big Data  
Universidade Federal do Paraná

<https://lineu96.github.io/st/>

TH MCGLM

Lineu Alberto

Quem sou eu

Um pouco sobre o PPGINF

Minha pesquisa

Etapas do processo de análise

Conjunto de dados

Modelos de Regressão

McGLM

Testes de hipótese em modelos de regressão

Testes de hipótese em Modelos Multivariados de Covariância Linear Generalizada (McGLM)

1

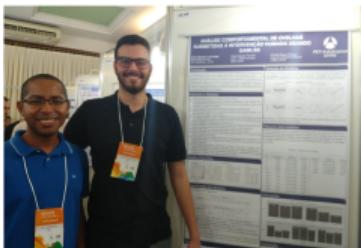
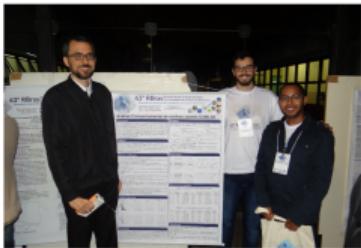
# Quem sou eu

[TH MCGLM](#)[Lineu Alberto](#)[Quem sou eu](#)[Um pouco sobre o  
PPGINF](#)[Minha pesquisa](#)[Etapas do processo de  
análise](#)[Conjunto de dados](#)[Modelos de Regressão](#)[McGLM](#)[Testes de hipótese em  
modelos de regressão](#)[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

## Quem sou eu

- ▶ Estatístico formado pela Universidade Federal do Paraná (UFPR) em 2019.
- ▶ Atividades participadas na graduação:
  - ▶ Programa de Educação Tutorial (PET-Estatística UFPR).
  - ▶ Monitoria da disciplina de Modelos Lineares Generalizados (CE225) na UFPR.
  - ▶ Estágio no Núcleo de Estatística e Monitoramento da CGJ (NEMOC-TJPR).
  - ▶ Estágio no Departamento de Inteligência de Crédito do Banco Bradesco S.A.



[TH MCGLM](#)[Lineu Alberto](#)[Quem sou eu](#)[Um pouco sobre o PPGINF](#)[Minha pesquisa](#)[Etapas do processo de análise](#)[Conjunto de dados](#)[Modelos de Regressão](#)[McGLM](#)[Testes de hipótese em modelos de regressão](#)[Testes de hipótese em Modelos Multivariados de Covariância Linear Generalizada \(McGLM\)](#)



TH MCGLM

Lineu Alberto

Quem sou eu

Um pouco sobre o  
PPGINF

Minha pesquisa

Etapas do processo de  
análise

Conjunto de dados

Modelos de Regressão

McGLM

Testes de hipótese em  
modelos de regressãoTestes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada (McGLM)

## Onde estou hoje

- ▶ Atualmente mestrando no Programa de Pós Graduação em Informática da UFPR sob a orientação dos professores:
  - ▶ Wagner Hugo Bonat.
  - ▶ Marco Antonio Zanata Alves.
- ▶ Inserido na área de concentração Ciência da Computação, linha de pesquisa Tecnologia da Informação e grupo de pesquisa Data Science & Big Data.

TH MCGLM

Lineu Alberto

Quem sou eu

Um pouco sobre o  
PPGINF

Minha pesquisa

Etapas do processo de  
análise

Conjunto de dados

Modelos de Regressão

McGLM

Testes de hipótese em  
modelos de regressão

Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada (McGLM)

## 2

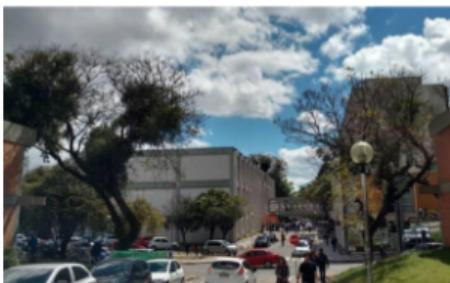
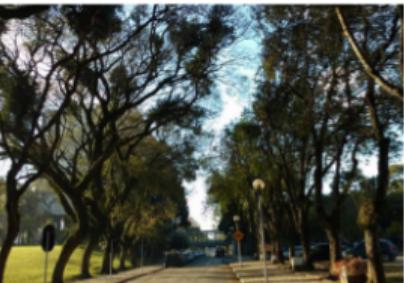
# Um pouco sobre o PPGINF

[TH MCGLM](#)[Lineu Alberto](#)[Quem sou eu](#)[Um pouco sobre o  
PPGINF](#)[Minha pesquisa](#)[Etapas do processo de  
análise](#)[Conjunto de dados](#)[Modelos de Regressão](#)[McGLM](#)[Testes de hipótese em  
modelos de regressão](#)[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

# PPGInf UFPR Centro Politécnico

Departamento de  
Informática

CURITIBA | PR  
[ppginf@inf.ufpr.br](mailto:ppginf@inf.ufpr.br)



## Visão geral

- ▶ O Programa é oferecido pelo Departamento de Informática da UFPR.
- ▶ Tem conceito 5 de 7 possíveis atribuídos na última avaliação pela CAPES (5 equivale a "muito bom" na escala).
- ▶ Oferece mestrado desde 1996.
- ▶ Oferece doutorado desde 2009.

TH MCGLM

Lineu Alberto

Quem sou eu

Um pouco sobre o  
PPGINF

Minha pesquisa

Etapas do processo de  
análise

Conjunto de dados

Modelos de Regressão

McGLM

Testes de hipótese em  
modelos de regressão

Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada (McGLM)

## Áreas, linhas e grupos de pesquisa

- ▶ No programa a área de concentração é Ciência da Computação.
- ▶ Existem 3 linhas de pesquisa:
  - ▶ Inteligência computacional.
  - ▶ Redes e sistemas distribuídos.
  - ▶ Tecnologia da Informação.
- ▶ Cada uma das linhas possuem diversos grupos de pesquisa.

TH MCGLM

Lineu Alberto

Quem sou eu

Um pouco sobre o  
PPGINF

Minha pesquisa

Etapas do processo de  
análise

Conjunto de dados

Modelos de Regressão

McGLM

Testes de hipótese em  
modelos de regressão

Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada (McGLM)

## Grupos de pesquisa

### Inteligência Computacional

C-Bio      Computação Bioinspirada, Meta-heurísticas e Aprendizagem de Máquina

IMAGO      Visão Computacional, Computação Gráfica e Processamento de Imagens

LIAMF      Inteligência Artificial e Métodos Formais

LIC      Sistemas Tutores Inteligentes

TEORIA      Teoria da Computação, Otimização e Combinatória

VRI      Visão, Robótica e Imagem

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o  
PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de  
análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em  
modelos de regressão](#)

[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

## Grupos de pesquisa

### Redes e Sistemas Distribuídos

CCSC Centro de Ciência de Segurança Computacional

LARSIS Laboratório de Redes, Sistemas Distribuídos e Segurança

NR2 Gerenciamento de Redes de Computadores, Sistemas Distribuídos, Redes Sem Fio, Segurança, Modelagem e Análise de Desempenho

LSE Laboratório de Sistemas Embarcados

HiPES High Performance and Efficient Systems. (Computer Architecture, Security and Networks)

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o  
PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de  
análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em  
modelos de regressão](#)

[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

## Grupos de pesquisa

### Tecnologia da Informação

BD Banco de Dados

BDGE Banco de Dados de Grande Escala

FAES Fundamentos e Aplicações em Engenharia de Software

C3SL Centro de Computação Científica e Software Livre

GrES Engenharia de Software

IHC Interação Humano-Computador

DSBD Data Science and Big Data

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o  
PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de  
análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em  
modelos de regressão](#)

[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

## Por que é um programa interessante para nós estatísticos?

- ▶ Porque é interessante se arriscar em outra área.
- ▶ Pela variedade de grupos de pesquisa.
- ▶ Pela flexibilidade do currículo:
  - ▶ 2 disciplinas obrigatórias.
  - ▶ 3 disciplinas optativas.
  - ▶ Prática em docência (para bolsistas).
- ▶ Pelo DSBD.

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o  
PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de  
análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

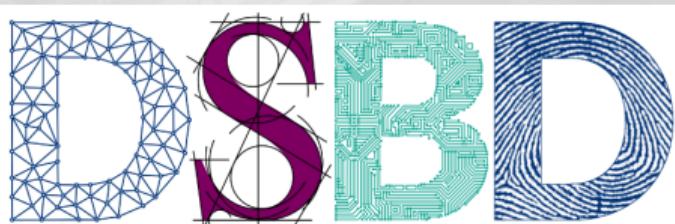
[McGLM](#)

[Testes de hipótese em  
modelos de regressão](#)

[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

## O DSBD

- ▶ [Laboratório de Data Science & Big Data.](#)
- ▶ Criado em 2018 em uma parceria entre os departamentos de Informática ([DINF](#)) e Estatística ([DEST](#)).
- ▶ É uma possibilidade para quem quer seguir na academia, em Curitiba, seguindo trabalhando de alguma forma com métodos estatísticos (ou não), sem perder o contato com o ([DEST](#)).
- ▶ É uma oportunidade para levar nossos conhecimentos adquiridos no bacharelado em Estatística e combiná-los com conhecimentos a serem adquiridos na Ciência da Computação.

[TH MCGLM](#)[Lineu Alberto](#)[Quem sou eu](#)[Um pouco sobre o PPGINF](#)[Minha pesquisa](#)[Etapas do processo de análise](#)[Conjunto de dados](#)[Modelos de Regressão](#)[McGLM](#)[Testes de hipótese em modelos de regressão](#)[Testes de hipótese em Modelos Multivariados de Covariância Linear Generalizada \(McGLM\)](#)

## 3

# Minha pesquisa

[TH MCGLM](#)[Lineu Alberto](#)[Quem sou eu](#)[Um pouco sobre o  
PPGINF](#)[Minha pesquisa](#)[Etapas do processo de  
análise](#)[Conjunto de dados](#)[Modelos de Regressão](#)[McGLM](#)[Testes de hipótese em  
modelos de regressão](#)[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

## Onde tudo começou

- ▶ O projeto teve início em 2018 quando eu e uma colega de curso ([Jhenifer Caetano Veloso](#)), desenvolvemos nosso TCC sob orientação do professor Wagner.
- ▶ O título do trabalho foi "[Análise de Variância Multivariada para Dados Não Gaussianos via Teste Wald](#)".

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em modelos de regressão](#)

[Testes de hipótese em Modelos Multivariados de Covariância Linear Generalizada \(McGLM\)](#)



TH MCGLM

Lineu Alberto

Quem sou eu

Um pouco sobre o  
PPGINF

Minha pesquisa

Etapas do processo de  
análise

Conjunto de dados

Modelos de Regressão

McGLM

Testes de hipótese em  
modelos de regressão

Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada (McGLM)

## Plano para o mestrado

- ▶ A ideia da minha pesquisa de mestrado é dar continuidade e ampliar o foco do trabalho que teve início na graduação.
- ▶ O título atual do trabalho é "Testes de hipótese em Modelos Multivariados de Covariância Linear Generalizada (McGLM)".

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o  
PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de  
análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em  
modelos de regressão](#)

[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

## Nosso objetivo é:

- ▶ Adaptar o teste Wald para realização de testes de hipótese gerais sobre parâmetros de Modelos Multivariados de Covariância Linear Generalizada (McGLM).
- ▶ Implementar funções para efetuar tais testes, bem como funções para efetuar Análises de Variância e Análises de Variância Multivariadas para os McGLM.
- ▶ Demonstrar as propriedades e comportamento dos testes propostos com base em estudos de simulação.
- ▶ Demonstrar o potencial de aplicação das metodologias discutidas com base na aplicação a conjuntos de dados reais.

## Plano de hoje

- ▶ Devido ao público alvo da palestra vamos tentar falar pouco de teoria (especificações, fórmulas, etc) e mais sobre ideias (pra que as coisas servem e porque utilizá-las é interessante).
- ▶ Vamos tentar construir uma linha de raciocínio falando desde o básico até chegar na classe de modelos que eu trabalho pra poder falar sobre o que estamos desenvolvendo.

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em modelos de regressão](#)

[Testes de hipótese em Modelos Multivariados de Covariância Linear Generalizada \(McGLM\)](#)

## Vamos conversar sobre:

- ▶ Quais as etapas do processo de análise de um conjunto de dados.
- ▶ Modelos de regressão de uma forma bastante didática (para que servem, quais os tipos, pressupostos gerais).
- ▶ McGLM que é classe de modelos de regressão sobre a qual trabalhamos.
- ▶ Como podemos aplicar teste de hipótese sobre os parâmetros do modelo desta classe.

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o  
PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de  
análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em  
modelos de regressão](#)

[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

## 4

# Etapas do processo de análise

TH MCGLM

Lineu Alberto

Quem sou eu

Um pouco sobre o  
PPGINF

Minha pesquisa

Etapas do processo de  
análise

Conjunto de dados

Modelos de Regressão  
McGLM

Testes de hipótese em  
modelos de regressão

Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada (McGLM)

## O processo de análise consiste em:

- ▶ Definição do problema.
- ▶ Planejamento do estudo.
- ▶ Coleta de dados.
- ▶ Análise dos dados:
  - ▶ Análise exploratória.
  - ▶ Aplicação de métodos mais sofisticados que permitam generalizar os resultados para a população.
- ▶ Interpretação dos resultados.

TH MCGLM

Lineu Alberto

Quem sou eu

Um pouco sobre o  
PPGINF

Minha pesquisa

Etapas do processo de  
análise

Conjunto de dados

Modelos de Regressão

McGLM

Testes de hipótese em  
modelos de regressão

Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada (McGLM)

## 5

# Conjunto de dados

[TH MCGLM](#)[Lineu Alberto](#)[Quem sou eu](#)[Um pouco sobre o  
PPGINF](#)[Minha pesquisa](#)[Etapas do processo de  
análise](#)[Conjunto de dados](#)[Modelos de Regressão](#)[McGLM](#)[Testes de hipótese em  
modelos de regressão](#)[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

- ▶ Um conjunto de dados ou "dataset" é uma coleção de dados normalmente tabulados em que organiza-se elementos ou indivíduos e suas características (variáveis).
- ▶ Uma forma conveniente de se organizar os dados é da seguinte forma:
  - ▶ Cada coluna representa uma variável.
  - ▶ Cada linha representa uma observação.
  - ▶ Cada célula representa o valor observado no elemento i na variável j.

country	year	cases	population
Afghanistan	2000	775	193071
Afghanistan	2000	2666	2059360
Brazil	1999	31737	17206362
Brazil	2000	80488	17404898
China	1999	21258	127215272
China	2000	21666	128018583

Variáveis

country	year	cases	population
Afghanistan	1999	775	193071
Afghanistan	2000	2666	2059360
Brazil	1999	31737	17206362
Brazil	2000	80488	17404898
China	1999	21258	127215272
China	2000	21666	128018583

Observações

country	year	cases	population
Afghanistan	1999	775	193071
Afghanistan	2000	2666	2059360
Brazil	1999	31737	17206362
Brazil	2000	80488	17404898
China	1999	21258	127215272
China	2000	21666	128018583

Valores

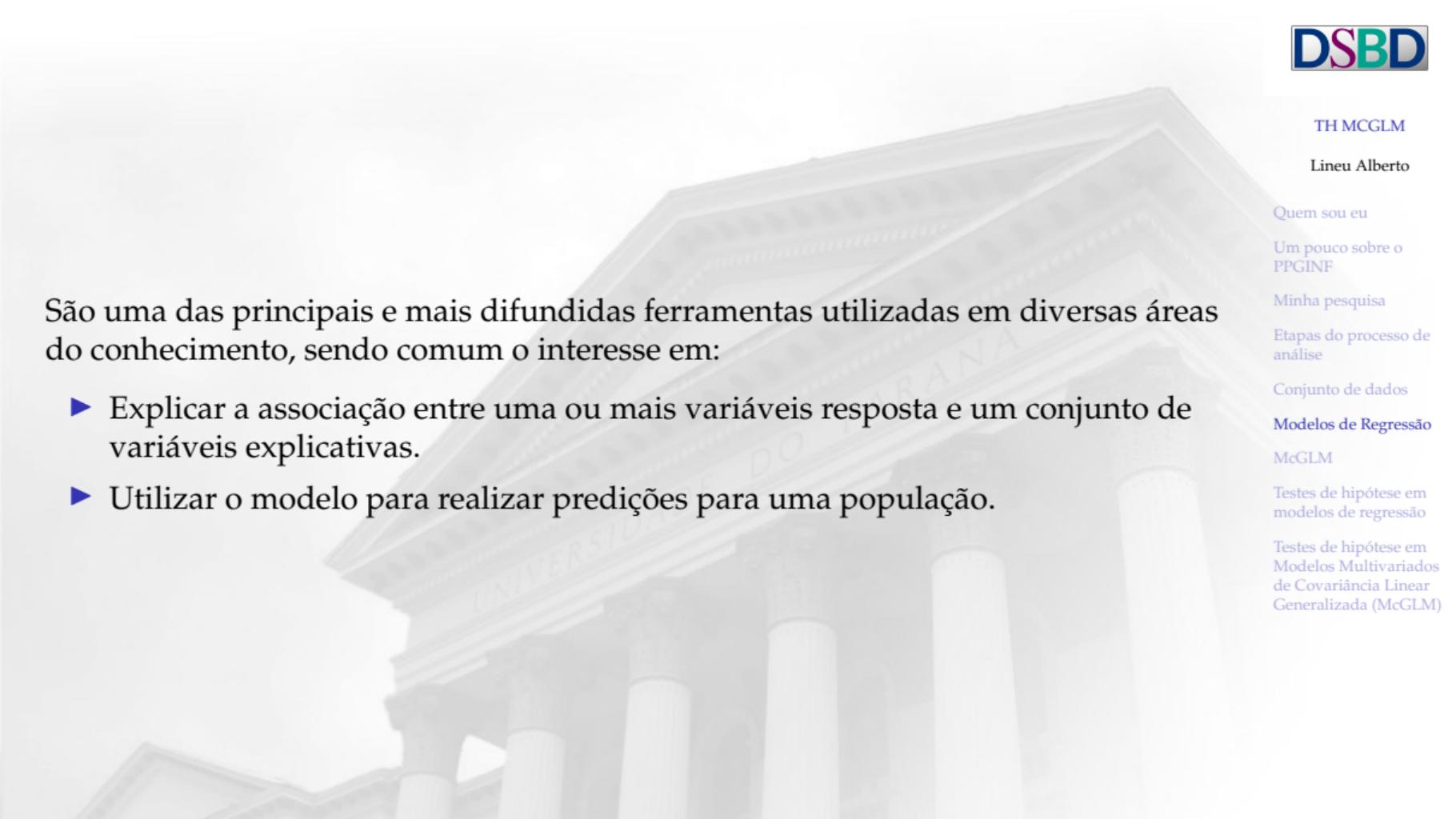
## 6

# Modelos de Regressão

[TH MCGLM](#)[Lineu Alberto](#)[Quem sou eu](#)[Um pouco sobre o  
PPGINF](#)[Minha pesquisa](#)[Etapas do processo de  
análise](#)[Conjunto de dados](#)[Modelos de Regressão](#)[McGLM](#)[Testes de hipótese em  
modelos de regressão](#)[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

- ▶ Nos casos univariados, estes modelos associam uma única variável resposta, a uma ou mais variáveis explicativas.
- ▶ De forma geral, um modelo de regressão é uma expressão matemática que relaciona a média da variável resposta às variáveis preditoras (covariáveis).
- ▶ A grosso modo, estimamos parâmetros que medem a força de associação da variável explicativa sobre a variável resposta.
- ▶ A variável resposta segue uma distribuição de probabilidade condicional às covariáveis e a média é descrita por um preditor linear.

- ▶ Há casos em que são coletadas mais de uma resposta por unidade experimental.
- ▶ E há o interesse de modelá-las em função de um conjunto de variáveis explicativas.
- ▶ Saímos então do cenário univariado (apenas uma resposta) e passamos para o multivariado (mais de uma resposta).
- ▶ Existem classes de modelos univariados e multivariados.



São uma das principais e mais difundidas ferramentas utilizadas em diversas áreas do conhecimento, sendo comum o interesse em:

- ▶ Explicar a associação entre uma ou mais variáveis resposta e um conjunto de variáveis explicativas.
- ▶ Utilizar o modelo para realizar previsões para uma população.

[TH MCGLM](#)[Lineu Alberto](#)[Quem sou eu](#)[Um pouco sobre o PPGINF](#)[Minha pesquisa](#)[Etapas do processo de análise](#)[Conjunto de dados](#)[Modelos de Regressão](#)[McGLM](#)[Testes de hipótese em modelos de regressão](#)[Testes de hipótese em Modelos Multivariados de Covariância Linear Generalizada \(McGLM\)](#)

## Por exemplo...

- ▶ Considere o exemplo disponível na seção 4.3.6 do livro [Modelos de Regressão com Apoio Computacional](#) do professor [Gilberto A. Paula](#):
- ▶ Os dados referem-se a um estudo sobre demanda de TV's a cabo em 40 regiões dos Estados Unidos.

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em modelos de regressão](#)

[Testes de hipótese em Modelos Multivariados de Covariância Linear Generalizada \(McGLM\)](#)

**Foram coletadas as variáveis como:**

- ▶ Número de assinantes de TV a cabo.
- ▶ Número de domicílios.
- ▶ Percentual de domicílios com TV a cabo.
- ▶ Renda per capita por domicílio com TV a cabo (em USD).
- ▶ Dentre outras.

**O objetivo do estudo é verificar se existe associação entre o número de assinantes e as demais variáveis.**

TH MCGLM

Lineu Alberto

Quem sou eu

Um pouco sobre o  
PPGINF

Minha pesquisa

Etapas do processo de  
análise

Conjunto de dados

Modelos de Regressão

McGLM

Testes de hipótese em  
modelos de regressão

Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada (McGLM)

## E agora?

- ▶ É um problema que faz sentido ser analisado com um modelo de regressão?
- ▶ Qual a variável resposta?
- ▶ Quais as variáveis explicativas?
- ▶ Devo usar um modelo univariado ou multivariado?
- ▶ Qual a natureza da minha variável resposta?
- ▶ Que distribuição de probabilidade devo usar para modelar o problema?
- ▶ Seria um modelo focado em explicação ou predição?

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o  
PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de  
análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em  
modelos de regressão](#)

[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

**Tendo essas informações somos capazes de selecionar a classe de modelos de regressão que mais se adequa ao problema que temos em mãos. E passar por etapas como:**

- ▶ Ajustar o modelo.
- ▶ Verificar se está bem ajustado.
- ▶ Em caso de mal ajuste, verificar que mudanças são possíveis na especificação.
- ▶ Em caso de ajuste satisfatório, podemos estudar os parâmetros:
  - ▶ Lembre-se, os parâmetros são quantidades que associam as variáveis explicativas com as variáveis resposta.
  - ▶ Podemos formular hipóteses para verificar se estes parâmetros são iguais a determinados valores de interesse, como 0, por exemplo.

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em modelos de regressão](#)

[Testes de hipótese em Modelos Multivariados de Covariância Linear Generalizada \(McGLM\)](#)

## Existem inúmeras classes de modelos de regressão, algumas delas são:

- ▶ Modelo Linear Normal.
- ▶ Modelos Lineares Generalizados (GLM).
- ▶ Modelos de regressão local.
- ▶ Modelos de regressão de splines.
- ▶ Modelos aditivos generalizados (GAM).
- ▶ Modelos de efeitos aleatórios.
- ▶ Modelos Aditivos Generalizados para Locação, Escala e Forma (GAMLSS).
- ▶ Modelos de regressão multivariados.
- ▶ Modelos Multivariados de Covariância Linear Generalizada (McGLM).

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o  
PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de  
análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em  
modelos de regressão](#)

[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

## Existem inúmeras classes de modelos de regressão, algumas delas são:

- ▶ Modelo Linear Normal.
- ▶ Modelos Lineares Generalizados (GLM).
- ▶ Modelos de regressão local.
- ▶ Modelos de regressão de splines.
- ▶ Modelos aditivos generalizados (GAM).
- ▶ Modelos de efeitos aleatórios.
- ▶ Modelos Aditivos Generalizados para Locação, Escala e Forma (GAMLSS).
- ▶ Modelos de regressão multivariados.
- ▶ **Modelos Multivariados de Covariância Linear Generalizada (McGLM).**

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o  
PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de  
análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em  
modelos de regressão](#)

[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

7

# McGLM

[TH MCGLM](#)[Lineu Alberto](#)[Quem sou eu](#)[Um pouco sobre o  
PPGINF](#)[Minha pesquisa](#)[Etapas do processo de  
análise](#)[Conjunto de dados](#)[Modelos de Regressão](#)[McGLM](#)[Testes de hipótese em  
modelos de regressão](#)[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

## Os Modelos Lineares Generalizados (GLM)

- ▶ São uma forma de modelagem univariada para dados de diferentes naturezas, tais como dados contínuos simétricos e assimétricos, contagens, dentre outras.
- ▶ Tais características tornam essa classe de modelos uma flexível ferramenta de modelagem aplicável a diversos tipos de problemas.
- ▶ Contudo, por mais flexível e discutida na literatura, essa classe apresenta duas principais restrições:
  1. A incapacidade de lidar com observações dependentes.
  2. A incapacidade de lidar com múltiplas respostas simultaneamente.

## Os Modelos Multivariados de Covariância Linear Generalizada (McGLM)

- ▶ Trata-se de uma classe que contorna as duas principais restrições presentes nos GLM.
- ▶ Esta classe de modelagem comporta:
  - ▶ Múltiplas respostas.
  - ▶ Respostas de diferentes naturezas.
  - ▶ Respostas correlacionadas.
  - ▶ Observações não independentes.
  - ▶ Extensões multivariadas para modelos de:
    - ▶ Séries temporais.
    - ▶ Dados longitudinais.
    - ▶ Dados espaciais.

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o  
PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de  
análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em  
modelos de regressão](#)

[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

# Testes de hipótese em modelos de regressão

[TH MCGLM](#)[Lineu Alberto](#)[Quem sou eu](#)[Um pouco sobre o PPGINF](#)[Minha pesquisa](#)[Etapas do processo de análise](#)[Conjunto de dados](#)[Modelos de Regressão](#)[McGLM](#)[Testes de hipótese em modelos de regressão](#)[Testes de hipótese em Modelos Multivariados de Covariância Linear Generalizada \(McGLM\)](#)

## Qual é a importância disso?

- ▶ Como já discutimos, nos modelos de regressão estimamos parâmetros que associam variáveis explicativas às variáveis resposta.
- ▶ Caso tenhamos um modelo para um problema e este modelo esteja bem ajustado, podemos estudar estas estimativas a fim de melhor compreender o fenômeno estudado.
- ▶ E extrair informações como:
  - ▶ Quais variáveis são importantes.
  - ▶ Quais não são.
  - ▶ Quais delas estão positivamente ou negativamente associadas à resposta.
  - ▶ etc.

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o  
PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de  
análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em  
modelos de regressão](#)

[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

## Fazemos isto por meio de testes de hipótese.

Com estes testes podemos responder perguntas como:

- ▶ Existe evidência suficiente que nos permita afirmar que o parâmetro que associa determinada variável explicativa à variável resposta é igual a 0?
- ▶ Em outras palavras: existe evidência que nos permita afirmar que existe efeito desta variável sobre a variável resposta?
- ▶ Esta é a ideia dos testes de hipótese aplicados no contexto dos modelos de regressão.

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de análise](#)

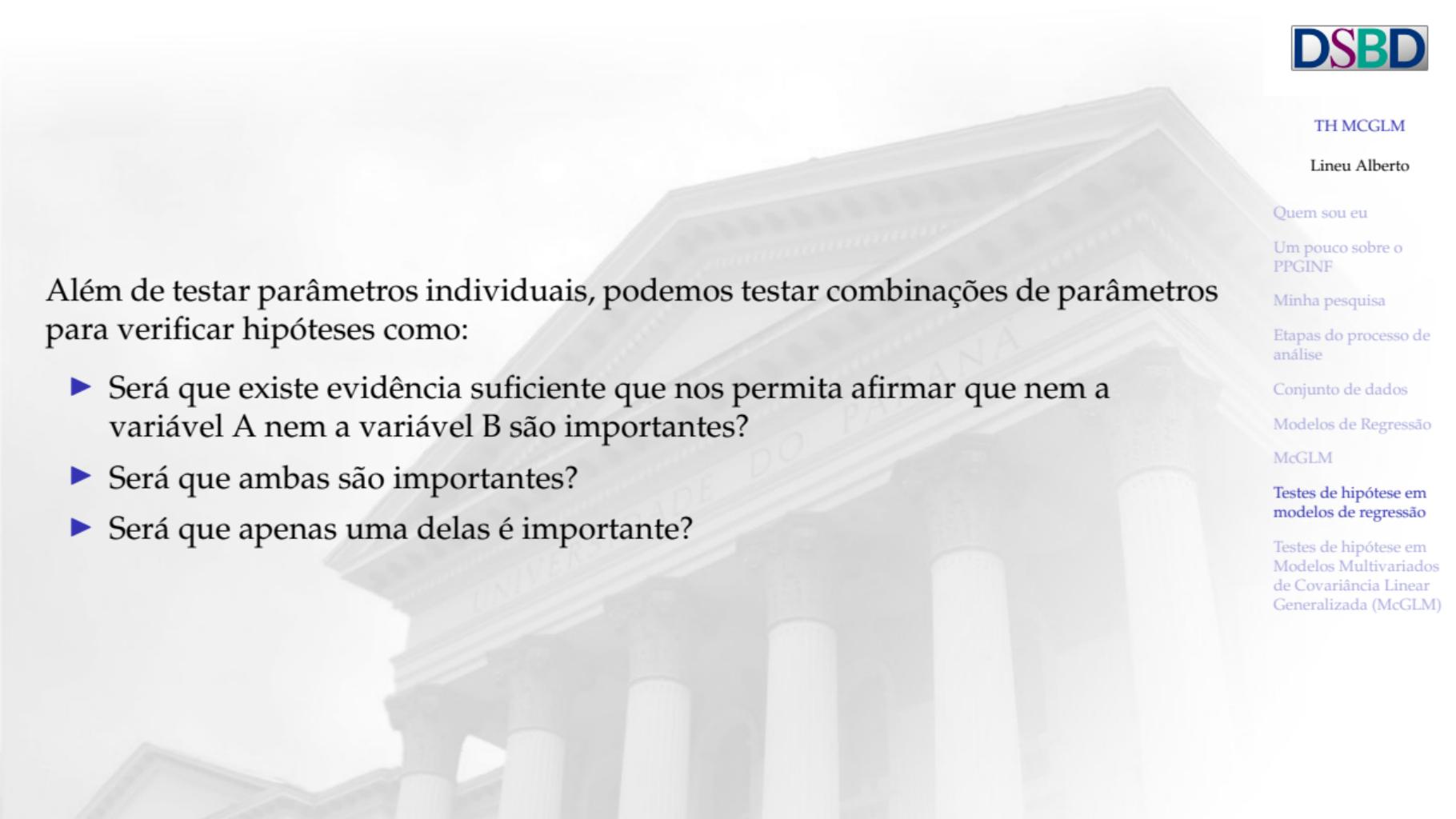
[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em modelos de regressão](#)

[Testes de hipótese em Modelos Multivariados de Covariância Linear Generalizada \(McGLM\)](#)

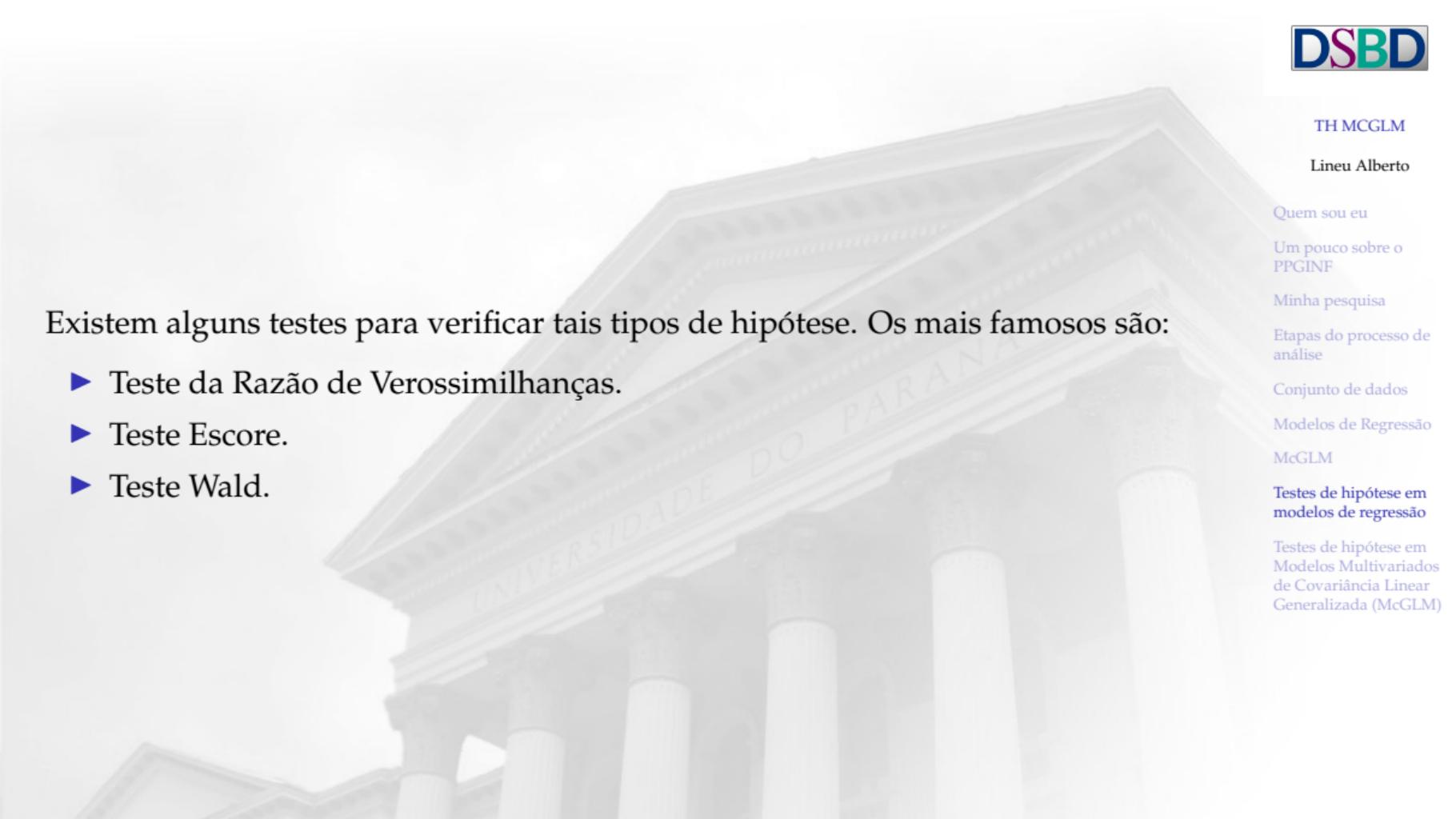


Além de testar parâmetros individuais, podemos testar combinações de parâmetros para verificar hipóteses como:

- ▶ Será que existe evidência suficiente que nos permita afirmar que nem a variável A nem a variável B são importantes?
- ▶ Será que ambas são importantes?
- ▶ Será que apenas uma delas é importante?

[TH MCGLM](#)[Lineu Alberto](#)[Quem sou eu](#)[Um pouco sobre o PPGINF](#)[Minha pesquisa](#)[Etapas do processo de análise](#)[Conjunto de dados](#)[Modelos de Regressão](#)[McGLM](#)[Testes de hipótese em modelos de regressão](#)[Testes de hipótese em Modelos Multivariados de Covariância Linear Generalizada \(McGLM\)](#)

- ▶ No caso dos modelos multivariados isso se torna especialmente útil pois podemos verificar a validade de hipóteses como:
  - ▶ Será que determinada variável é importante em todas as respostas simultaneamente?
  - ▶ Será que determinada variável é importante em ao menos uma resposta?
  - ▶ Será que determinada variável não tem efeito em nenhuma das respostas?
- ▶ Note que se trata de uma hipótese sobre múltiplos parâmetros, pois temos um valor que associa a variável para cada uma das respostas.



Existem alguns testes para verificar tais tipos de hipótese. Os mais famosos são:

- ▶ Teste da Razão de Verossimilhanças.
- ▶ Teste Escore.
- ▶ Teste Wald.

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o  
PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de  
análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em  
modelos de regressão](#)

[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

## Vamos retomar o exemplo...

Considerando o estudo sobre demanda de TV's a cabo em 40 regiões dos Estados Unidos. Em que temos as seguintes variáveis:

- ▶ Número de assinantes de TV a cabo.
- ▶ Número de domicílios.
- ▶ Percentual de domicílios com TV a cabo.
- ▶ Renda per capita por domicílio com TV a cabo (em USD).
- ▶ Dentre outras.

TH MCGLM

Lineu Alberto

Quem sou eu

Um pouco sobre o  
PPGINF

Minha pesquisa

Etapas do processo de  
análise

Conjunto de dados

Modelos de Regressão

McGLM

Testes de hipótese em  
modelos de regressão

Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada (McGLM)

- ▶ Digamos que agora nosso interesse seja modelar o número de assinantes de TV a cabo E o percentual de domicílios com TV a cabo.
- ▶ Note que temos duas variáveis resposta, ou seja, caberia utilizar uma classe de modelos multivariados (como o McGLM).
- ▶ Digamos que tenhamos ajustado um modelo, verificado os pressupostos, realizado uma análise de diagnóstico e que, por fim, concluímos que nosso modelo está bem ajustado.
- ▶ Passamos então para a etapa do estudo das estimativas dos parâmetros que associam variáveis explicativas às variáveis resposta.

## Vamos considerar a variável explicativa número de domicílios:

- ▶ Teríamos em mãos as estimativas de 2 parâmetros referentes ao número de domicílios:
  - ▶ Um deles que associa a variável à primeira resposta.
  - ▶ E outro que associa a variável à segunda resposta.
- ▶ Ou seja, teremos um parâmetro que associa o número de domicílios ao número de assinantes de TV a cabo e outro que associa o número de domicílios com o percentual de domicílios com TV a cabo.

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o  
PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de  
análise](#)

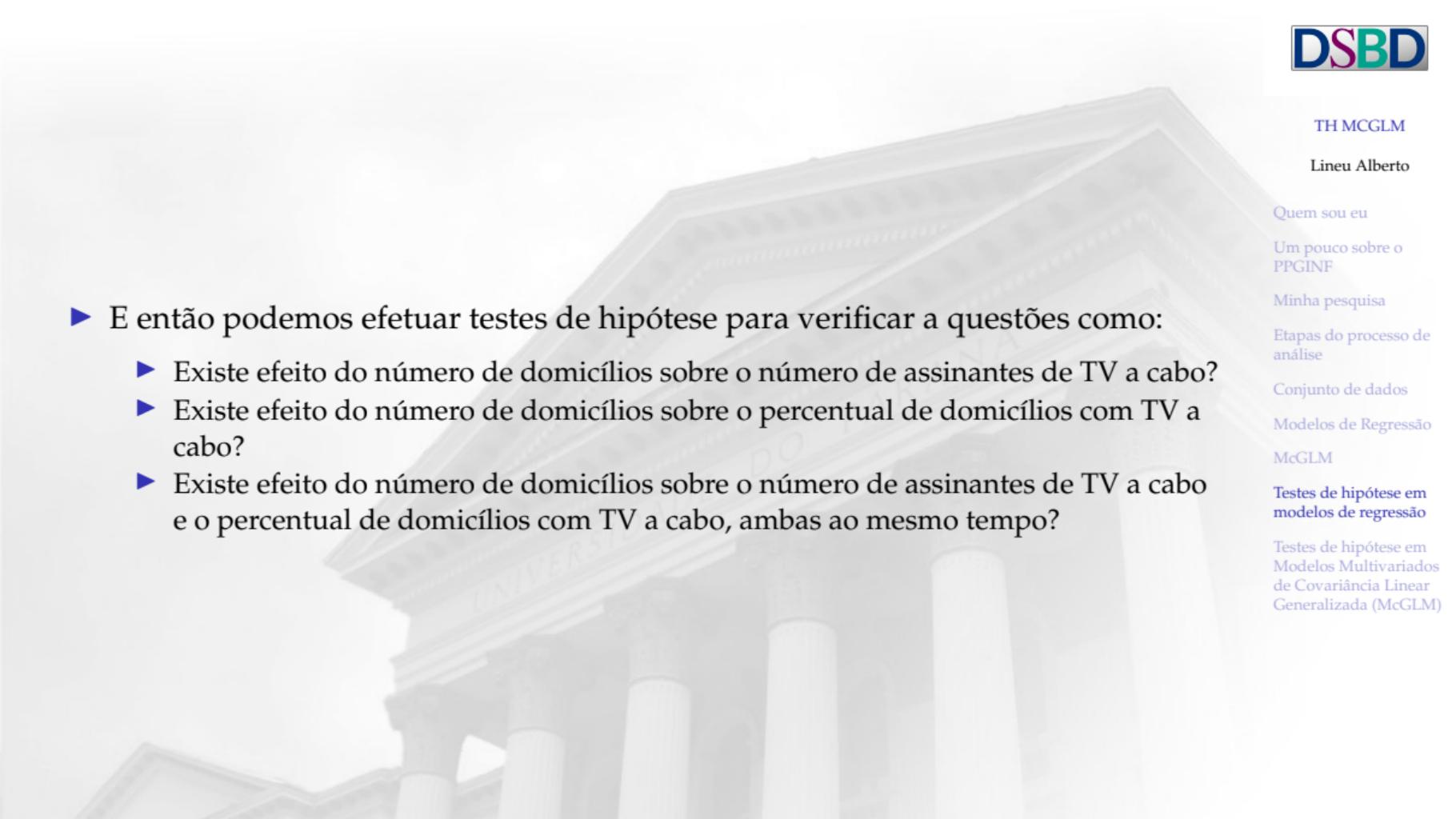
[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em  
modelos de regressão](#)

[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

- 
- ▶ E então podemos efetuar testes de hipótese para verificar a questões como:
    - ▶ Existe efeito do número de domicílios sobre o número de assinantes de TV a cabo?
    - ▶ Existe efeito do número de domicílios sobre o percentual de domicílios com TV a cabo?
    - ▶ Existe efeito do número de domicílios sobre o número de assinantes de TV a cabo e o percentual de domicílios com TV a cabo, ambas ao mesmo tempo?

[TH MCGLM](#)[Lineu Alberto](#)[Quem sou eu](#)[Um pouco sobre o  
PPGINF](#)[Minha pesquisa](#)[Etapas do processo de  
análise](#)[Conjunto de dados](#)[Modelos de Regressão](#)[McGLM](#)[Testes de hipótese em  
modelos de regressão](#)[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

# Testes de hipótese em Modelos Multivariados de Covariância Linear Generalizada (McGLM)

[TH MCGLM](#)[Lineu Alberto](#)[Quem sou eu](#)[Um pouco sobre o PPGINF](#)[Minha pesquisa](#)[Etapas do processo de análise](#)[Conjunto de dados](#)[Modelos de Regressão](#)[McGLM](#)[Testes de hipótese em modelos de regressão](#)[Testes de hipótese em Modelos Multivariados de Covariância Linear Generalizada \(McGLM\)](#)

## Até o momento estes testes não foram propostos para os McGLM.

- ▶ O que nós estamos fazendo no trabalho é trabalhar com o teste Wald.
- ▶ As etapas do trabalho são:
  - ▶ Adaptar o teste para que seja aplicável no McGLM.
  - ▶ Implementar funções que permitam com que usuário faça seus testes de hipótese de uma forma simples.
  - ▶ Verificar, através de estudos de simulação, se o teste tem um comportamento adequado.
  - ▶ Fazer aplicações a conjuntos de dados reais.

Considerando a área de pesquisa, o trabalho teria as seguintes contribuições:

1. Adaptar um teste existente para uma classe de modelos não usual mas com alto potencial de aplicação.
2. Realizar um estudo pesado de simulação para verificar o funcionamento da forma que estamos propondo.
3. Analisar de dados provenientes de estudos reais para demonstrar a aplicabilidade das funcionalidades.

[TH MCGLM](#)

[Lineu Alberto](#)

[Quem sou eu](#)

[Um pouco sobre o  
PPGINF](#)

[Minha pesquisa](#)

[Etapas do processo de  
análise](#)

[Conjunto de dados](#)

[Modelos de Regressão](#)

[McGLM](#)

[Testes de hipótese em  
modelos de regressão](#)

[Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada \(McGLM\)](#)

# Obrigado!

Lineu Alberto Cavazani de Freitas  
lineuacf@gmail.com  
<https://lineu96.github.io/st/>  
PPG Informática



TH MCGLM

Lineu Alberto

Quem sou eu

Um pouco sobre o  
PPGINF

Minha pesquisa

Etapas do processo de  
análise

Conjunto de dados

Modelos de Regressão

McGLM

Testes de hipótese em  
modelos de regressão

Testes de hipótese em  
Modelos Multivariados  
de Covariância Linear  
Generalizada (McGLM)