

# ECONOMETRIA

## INTRODUÇÃO

### **Professores:**

Hedibert Freitas Lopes (4ECO) - [hedibertFL@insper.edu.br](mailto:hedibertFL@insper.edu.br)

Priscila Fernandes Ribeiro (4ADM-A)

Sérgio Ricardo Martins (4ADM-B)

2<sup>o</sup> semestre de 2016

- 1 O que é Econometria?
  - Exemplo 1
  - Exemplo 2
  - Exemplo 3
- 2 Tipos de dados
- 3 Objetivos da Disciplina
- 4 Programa do curso
- 5 Bibliografia
- 6 Pacote estatístico R
  - Exemplo 4
- 7 Critério de avaliação
  - Provas
  - Atividades
  - Listas de exercícios
- 8 Regras para sala de aula

# Horários

Horários das aulas:

4<sup>as</sup> e 6<sup>as</sup> das 7h30 às 9h30

Horários de atendimento:

4<sup>as</sup> das 10h às 11h30

Sala 604

Monitora: Paloma Uribe

E-mail: PalomaVU@al.insper.edu.br

Horários: Sexta-feira das 12h às 13h30

Local: Sala BM&F Bovespa 2 - 3<sup>o</sup> andar.

# O que é Econometria?

É o ramo da Economia que trata da mensuração de relações econômicas, isto é, relações entre variáveis de natureza econômica.

A Econometria é, na verdade, uma combinação de teoria ou outro raciocínio a priori com matemática e estatística, com o objetivo de dar conteúdo empírico às formulações teóricas da Economia.

# Exemplo 1

O departamento de RH de uma empresa deseja avaliar a eficácia dos testes aplicados para a seleção de funcionários.

Para tanto, foi sorteada uma amostra aleatória de 50 funcionários que fazem parte da empresa e que passaram pelo processo de seleção que utilizou os tais testes..

Para cada um dos funcionários foi registrada a média das notas obtidas nos testes de criatividade, raciocínio mecânico, raciocínio abstrato e habilidade matemática (notas de 0 a 26).

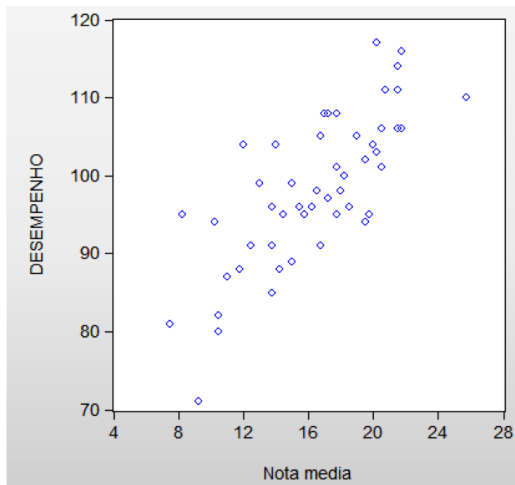
## Exemplo 1 (continuação)

Ainda, após 6 meses da contratação, foi calculado um escore que indica o seu desempenho profissional (0 a 120).

### **Pergunta:**

Existe alguma relação entre o escore de desempenho dos funcionários e a nota média nos testes?

# Exemplo 1 (continuação)



## Exemplo 1 (continuação)

### Possíveis perguntas de interesse:

- 1 Qual modelo estatístico você proporia para estudar a relação entre o escore de desempenho dos funcionários e a nota média nos testes?
- 2 Qual método de estimação você utilizaria para encontrar os estimadores dos parâmetros do modelo proposto?
- 3 Os estimadores anteriormente propostos apresentam boas propriedades?
- 4 Qual a estimativa para o escore médio de desempenho de funcionários que obtiveram nota média igual a 13 nos testes?
- 5 A variável nota média nos testes é relevante para explicar o escore médio de desempenho dos funcionários?



## Example 2

Tsay (2002) objetivou estimar os parâmetros do seguinte modelo de regressão:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

em que

$y_t$ : excessos de log-retornos mensais da Ford Motors;

$x_t$ : excessos de log-retornos mensais da S&P 500.

Ainda, para gerar as séries de interesse, o autor utilizou como ativo livre de risco a série mensal das letras do Tesouro Americano com vencimento a três meses.

## Example 3: Treinamento versus produtividade

Wooldridge (2011, Seções 1.1 e 1.2)

Você foi contratado para avaliar a eficácia de um programa de treinamento.

É sabido que o programa teve a duração de 20 semanas e que diversos cursos foram oferecidos fora do horário de expediente.

Além disso, você também ficou sabendo que qualquer trabalhador horista da produção poderia participar; e a matrícula em todo o programa (ou em parte dele) era voluntária.

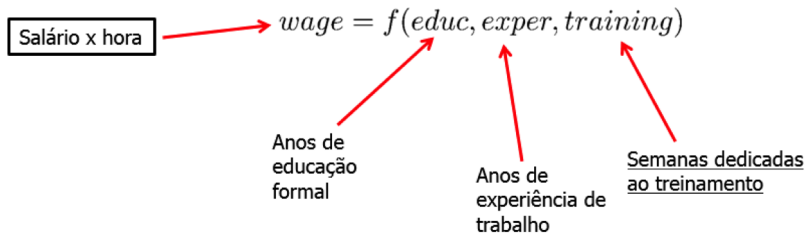
### **Pergunta:**

Qual o efeito, se houver, do programa de treinamento sobre o salário-hora de cada trabalhador?

## Example 3 (continuação)

Da sua experiência, você sabe que fatores como educação e experiência devem influenciar a produtividade dos trabalhadores e, por consequência, seus respectivos ganhos, uma vez que os trabalhadores costumam ser pagos de acordo com a sua produtividade.

Assim, tal raciocínio o levou ao seguinte modelo:



## Example 3 (continuação)

Após especificar um modelo econômico, seu próximo passo é pensar no modelo econométrico.

Porém, antes de podermos empreender uma análise econométrica, a forma da função  $f(\textit{educ}, \textit{exper}, \textit{training})$  precisa ser especificada.

**Como fazer isso?**

## Example 3 (continuação)

Suponha que, neste exemplo, o modelo econométrico pensado seja dado por:

$$wage = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 training + u$$

Diagram illustrating the regression model components:

- $wage$ : Salário x hora
- $\beta_1 educ$ : Anos de educação formal
- $\beta_2 exper$ : Anos de experiência de trabalho
- $\beta_3 training$ : Semanas dedicadas ao treinamento
- $u$ : Determinantes não-observados do salário

- 1 Qual razão nos leva a ter interesse, neste caso, no parâmetro  $\beta_3$ ?
- 2 Qual o significado prático de testar  $H_0 : \beta_3 = 0$  contra  $H_a : H_0$  é falsa?

# Tipos de dados

- 1) Cross-sectional (corte transversal)
- 2) Séries Temporais
- 3) Pooled cross sections (cortes transversais agrupados)
- 4) Painel/Longitudinais

Ao longo dessa disciplina trabalharemos com dados do tipo *cross-sectional* e de séries temporais.

# Dado cross-sectional

Amostra formada por indivíduos, firmas, cidades, estados, países ou qualquer outra unidade de interesse, observadas num determinado instante de tempo ou período de tempo.

**Exemplo:** Pesquisa Nacional por Amostrade Domicílios (PNAD)

Obtém informações anuais sobre características demográficas e socio-econômicas da população, como sexo, idade, educação, trabalho rendimento, e características dos domicílios, e, com periodicidade variável, informações sobre migração, fecundidade, nupcialidade, entre outras, tendo como unidade de coleta os domicílios. Temas específicos abrangendo aspectos demográficos, sociais e econômicos também são investigados.

# Série temporal

Uma série temporal é qualquer conjunto de observações ordenadas no tempo, espaço, volume ou algum outro parâmetro físico.

## Exemplos:

- IPCA mensal
- PIB trimestral
- Venda mensal de automóveis
- Câmbio diário
- Número de transações de determinado ativo financeiro a cada minuto

## Observações:

- Dados de séries temporais costumam ser serialmente correlacionados.
- Ordenação dos dados é importante.
- Frequência dos dados: diário, semanal, mensal, trimestral, anual, etc.
- Séries temporais costumam apresentar tendência e/ou sazonalidade.



# Objetivos da Disciplina

Ao final desta disciplina os alunos deverão ser capazes de aplicar corretamente o conteúdo da disciplina para **modelar e resolver problemas nas áreas de Administração e Economia**:

- 1 Propor variáveis para mensurar quantidades de interesse nas áreas de Administração e Economia;
- 2 Coletar bases de dados, resumir e analisar descritivamente os dados;
- 3 Propor, estimar e validar modelos de regressão linear de forma adequada, tanto para dados do tipo cross-section quanto para dados provenientes de séries temporais;
- 4 Interpretar os resultados do modelo proposto em termos do problema de interesse e usar tais resultados para tomada de decisão.

# Programa do curso (até a prova intermediária)

## Parte I: Análise de Regressão Linear Simples

- 1 Introdução (O Modelo de Regressão Linear Simples)
- 2 Método de Estimação (Mínimos Quadrados Ordinários)
- 3 Medida de Qualidade de Ajuste
- 4 Formas Funcionais
- 5 Suposições
- 6 Propriedade dos Estimadores
- 7 Inferência

## Parte II: Análise de Regressão Linear Múltipla

- 1 Introdução (O Modelo de Regressão Linear Múltipla)
- 2 Método de Estimação (Mínimos Quadrados Ordinários)
- 3 Medida de Qualidade de Ajuste
- 4 Suposições e Propriedade dos Estimadores
- 5 Interação
- 6 Variável Dummy

# Programa do curso (após a prova intermediária)

## Parte II: Mais Análise de Regressão Linear Múltipla

- 1 Teste F-parcial
- 2 Análise Gráfica de Resíduos
- 3 Heterocedasticidade
  - 1 Teste de White
  - 2 Estimador Robusto

## Parte III: Endogeneidade

- 1 Definição de endogeneidade
- 2 Variáveis Instrumentais
- 3 Mínimos Quadrados em 2 Estágios (2SLS)
- 4 Teste de Hausman e Teste de Sargan

# Programa do curso (após a prova intermediária)

## Parte IV: Modelos de Regressão com uso de dados de séries temporais

- ① Introdução aos dados de Séries Temporais
- ② Correlação Serial
- ③ Modelos autorregressivos
- ④ Modelos de Regressão com dados de séries temporais
  - ① Estimação
  - ② Análise de resíduos
- ⑤ Tendência
- ⑥ Sazonalidade

# Bibliografia

Wooldridge (2011) *Introdução à Econometria: uma abordagem moderna*, 4ª edição. São Paulo: Cengage Learning.

[Clique aqui para os dados da 4ª edição americana do Wooldridge.](#)  
[Clique aqui para os dados da 5ª edição americana do Wooldridge.](#)

## Bibliografia complementar

- Gujarati and Porter (2011) *Econometria Básica*, 5ª ed. AMGH Editora Ltda.
- Heij, Boer, Franses, Kloek and Van Dijk (2004) *Econometric methods with applications in business and economics*. Oxford University Press.
- Johnston e Dinardo (1997) *Econometric Methods*. McGraw Hill.
- Maddala and Lahiri (2009) *Introduction to Econometrics*, 4th ed. Wiley.
- Pindyck and Rubenfeld (2000) *Econometric Models and Economic Forecasts*, 4th ed. McGraw-Hill Publishing Co.
- Stock and Watson (2004) *Econometria*. Addison Wesley.

# Pacote estatístico R

Todos os cálculos apresentados ao longo do curso foram obtidos através da linguagem/pacote de estatística computacional R. [Clique aqui para saber mais.](#)

O R tem várias características atraentes e modernas:

- 1 É gratuito e *open-source*.
- 2 É extremamente fácil de agregar/alterar rotinas pre-existentes.
- 3 É muito fácil de usar;
- 4 Fácil comunicação com Cpp, Python, etc.
- 5 Usuário é motivado (leia-se: obrigado) a pensar!

A Paloma Uribe preparou um [material didático](#) bem legal.  
Outros manuais/tutoriais você poderá encontrar [aqui](#) e [aqui](#).  
Para aqueles que gostam de *reference cards*, veja [este aqui](#).

## Exemplo 4

O arquivo `salario.txt` contém 220 linhas (indivíduos) e 4 colunas (características):

- 1 `salario` (valores entre 110 e 172)
- 2 `posicao` (categorias de 1 a 9)
- 3 `anosexperiencia` (idade em anos de 1.7 a 26.1)
- 4 `sexo` (variável *dummy*: 0 =mulher,1 =homen)

Primeiras linhas do arquivo `salario.txt`

|   | salario | posicao | anosexperiencia | sexo |
|---|---------|---------|-----------------|------|
| 1 | 148     | 7       | 16.7            | 1    |
| 2 | 165     | 7       | 6.7             | 1    |
| 3 | 145     | 5       | 14.8            | 1    |
| 4 | 139     | 7       | 13.9            | 0    |
| 5 | 142     | 6       | 6.4             | 0    |
| 6 | 144     | 5       | 9.1             | 1    |
| 7 | 128     | 3       | 8.5             | 0    |

# Código R

```

dados = read.table("http://hedibert.org/wp-content/uploads/2014/02/salario.txt",header=TRUE)

attach(dados)

boxplot(salario[sexo==0],salario[sexo==1],names=c("Mulheres","Homens"),ylab="Salario",col=c(2,3))

boxplot(anosexperiencia[sexo==0],anosexperiencia[sexo==1],names=c("Mulheres","Homens"),ylab="Anos de experiencia",col=c(2,3))

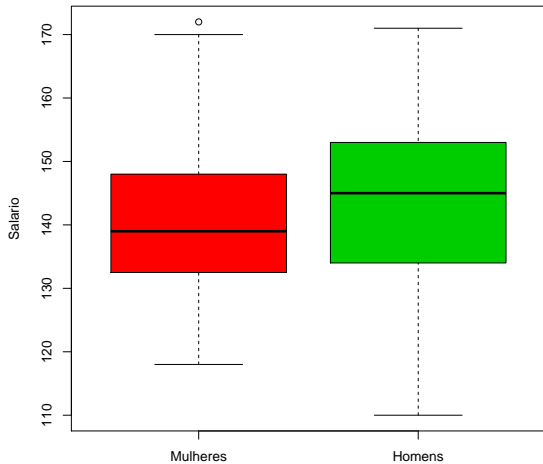
plot(anosexperiencia,salario,xlab="Anos de experiencia",ylab="Salario",col=sexo+1)
abline(lm(salario[sexo==0]~anosexperiencia[sexo==0]),col=1)
abline(lm(salario[sexo==1]~anosexperiencia[sexo==1]),col=2)
legend("bottomright",legend=c("mulheres","homens"),col=1:2,bty="n",cex=1.5,pch=16)

e = rnorm(220,0,0.1)
plot(posicao+e,salario,xlab="Posicao",ylab="Salario",col=sexo+1,axes=FALSE)
axis(2);box();axis(1,at=1:9)
abline(lm(salario[sexo==0]~posicao[sexo==0]),col=1)
abline(lm(salario[sexo==1]~posicao[sexo==1]),col=2)
legend("topleft",legend=c("mulheres","homens"),col=1:2,bty="n",cex=1.5,pch=16)
medias.mulher = rep(0,9)
medias.homen = rep(0,9)
for (i in 1:9){
  medias.mulher[i] = mean(salario[(sexo==0)&(posicao==i)])
  medias.homen[i] = mean(salario[(sexo==1)&(posicao==i)])
}
points(1:9,medias.mulher,col=1,pch=17,cex=1.5)
points(1:9,medias.homen,col=2,pch=17,cex=1.5)

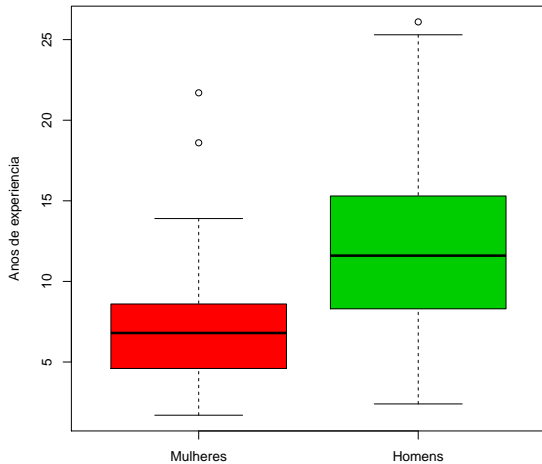
```



# Salário por sexo



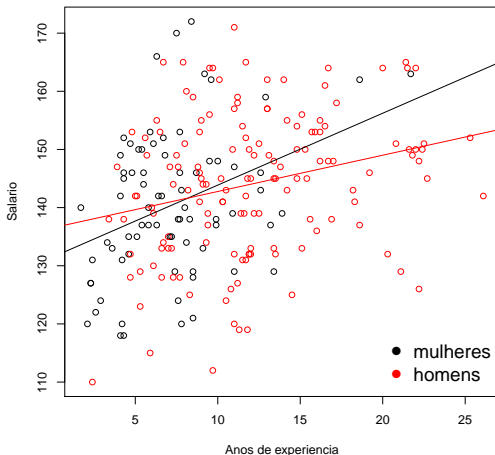
# Anos de experiência por sexo



# Salário vs anos de experiência

Mulheres:  $\text{salario} = 131.5 + 1.24 \text{ anosexperienca}$

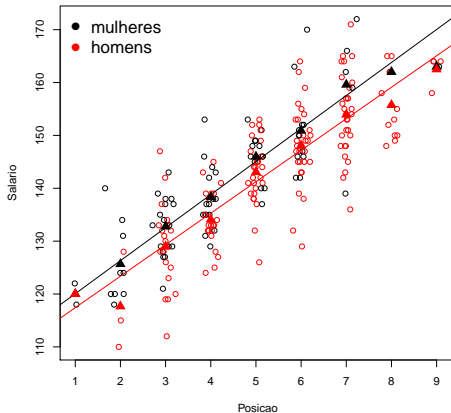
Homens:  $\text{salario} = 136.5 + 0.62 \text{ anosexperienca}$



# Salário vs posição

Mulheres:  $\text{salario} = 113.8 + 6.25 \text{ posicao}$

Homens:  $\text{salario} = 111.5 + 5.96 \text{ posicao}$



**Cuidado:** Retas não válidas pois posicao é uma variável categórica!

# Critério de avaliação

A nota final será uma média ponderada de três componentes:

- Prova Intermediária (PI): 30%
- Prova Final (PF): 60%
- Média simples das 2 maiores notas das 3 atividades (AT): 10%

# Provas

- **Matéria:** compreende a matéria dada em sala de aula e a apresentada nas indicações de leitura.
- Traga o material necessário, que inclui lápis, caneta, borracha e calculadora.
- **As provas serão com consulta exclusivamente ao Wooldridge (2011).**
- O professor não tirará dúvidas durante a prova: faz parte da avaliação compreender a questão. Se ela contiver erros que tornem a resolução impossível, indique-os na sua prova e você receberá a nota completa pela questão, podendo esta ser anulada, se for o caso.

# Atividades

Ao longo do semestre, 3 (três) atividades serão desenvolvidas (ao longo de 24 horas) em equipes de até 8 (oito) integrantes. Nas aulas seguintes às liberações das atividades, 3 grupos serão aleatoriamente sorteados e 2 membros desses grupos (também aleatoriamente sorteados) apresentarão, em nome de seus grupos, os resultados obtidos.

As atividades objetivam avaliar o grau de entendimento que os membros de cada equipe apresentam no que diz respeito à **Análise e Resolução de Problemas**.

Ainda, para desenvolver habilidades na compreensão e uso de teorias inferenciais, essas atividades têm por interesse apresentar um problema da maneira como ele é encontrado na prática. Além disso, trazem o propósito de avaliar a autonomia dos membros da equipe em buscar ferramentas estatísticas que foram ou não diretamente apresentadas em sala de aula e utilizá-las de forma análoga às aquelas já vistas.

# Lista de exercícios

Ao longo do semestre serão distribuídas algumas listas de exercícios e nelas existirão grupos de exercícios que:

- 1 Serão resolvidos pelo monitor, nas sessões de monitoria.
- 2 Poderão ser resolvidos pelos alunos, durante as sessões de monitoria.

Todos os exercícios das listas serão provenientes do livro Wooldrige (2011).

Gabaritos de exercícios **não resolvidos** durante as sessões de monitorias **não** serão liberados.

Também **não será liberado gabarito** dos exercícios resolvidos nas sessões de monitoria.



# Regras para sala de aula

- **Horário de entrada em aula:** até fechamento da porta.
- **Presença:** As chamadas serão feitas em horários aleatórios.
- **Laptop:** siga o **programa de ensino** para saber os dias marcados com o uso do R. Laptops deverá permanecer fechado até ser solicitado por mim.
- **Acesso à rede:** proibido durante o período de aula, a menos que seja necessário para alguma atividade didática.
- **Celulares:** desligados durante a aula.
- **Placa de identificação:** de uso obrigatório.
- Acompanhem o **programa de ensino** da disciplina que traz os assuntos que serão discutidos em cada aula! **Façam a leitura indicada previamente!**
- Todos os slides discutidos em sala estarão em [www.hedibert.org](http://www.hedibert.org).