

Macroeconometria - Séries de tempo

FAUSTO JOSÉ ARAÚJO VIEIRA

Aula 1

17 de abril a 22 de maio de 2018

OBJETIVO GERAL

- ▶ Esperamos que no final do curso, cada aluno seja capaz de estimar modelos de séries de tempo e possa projetar variáveis econômicas;
 - ▶ O aluno será capaz de escolher dentre os diferentes modelos econométricos, o melhor que se adequa a sua demanda;
- ▶ No final do curso, o aluno saberá usar o software “R” (R-Studio);
- ▶ Propomos uma abordagem de ensino mais prática;
 - ▶ O foco não é na demonstração teórica dos modelos, mas no entendimento dos conceitos e sua aplicação;
- ▶ Entender e analisar os resultados dos modelos e projeções do ponto de vista estatístico e econômico;
 - ▶ Replicaremos modelos propostos pelo BCB, FMI e artigos acadêmicos.

EMENTA

- ▶ Introdução e revisão do R-Studio e do curso de introdução a Econometria;
- ▶ Processos estacionários;
 - ▶ Modelos ARMA;
 - ▶ Sazonalidade;
- ▶ Processos não estacionários;
 - ▶ Raiz unitária - tendência estacionária e estocástica;
- ▶ Vetor Autorregressivo;
 - ▶ Estimação do VAR
 - ▶ Causalidade de Granger
- ▶ Estimação do FAVAR;
 - ▶ Análise de componentes principais
- ▶ Modelos Econométricos dinâmicos;
- ▶ Vetor de Correção de Erros;
 - ▶ Cointegração;
 - ▶ Estimação do VECM;

BIBLIOGRAFIA

BÁSICA

- ▶ Bueno, Rodrigo L. S. **Econometria de séries temporais**. São Paulo: Cengage Learning, 2015.
- ▶ Shumway, Robert e Stoffer, David. **Time Series Analysis and Its Applications with R Examples**. New York: Springer, 2011.

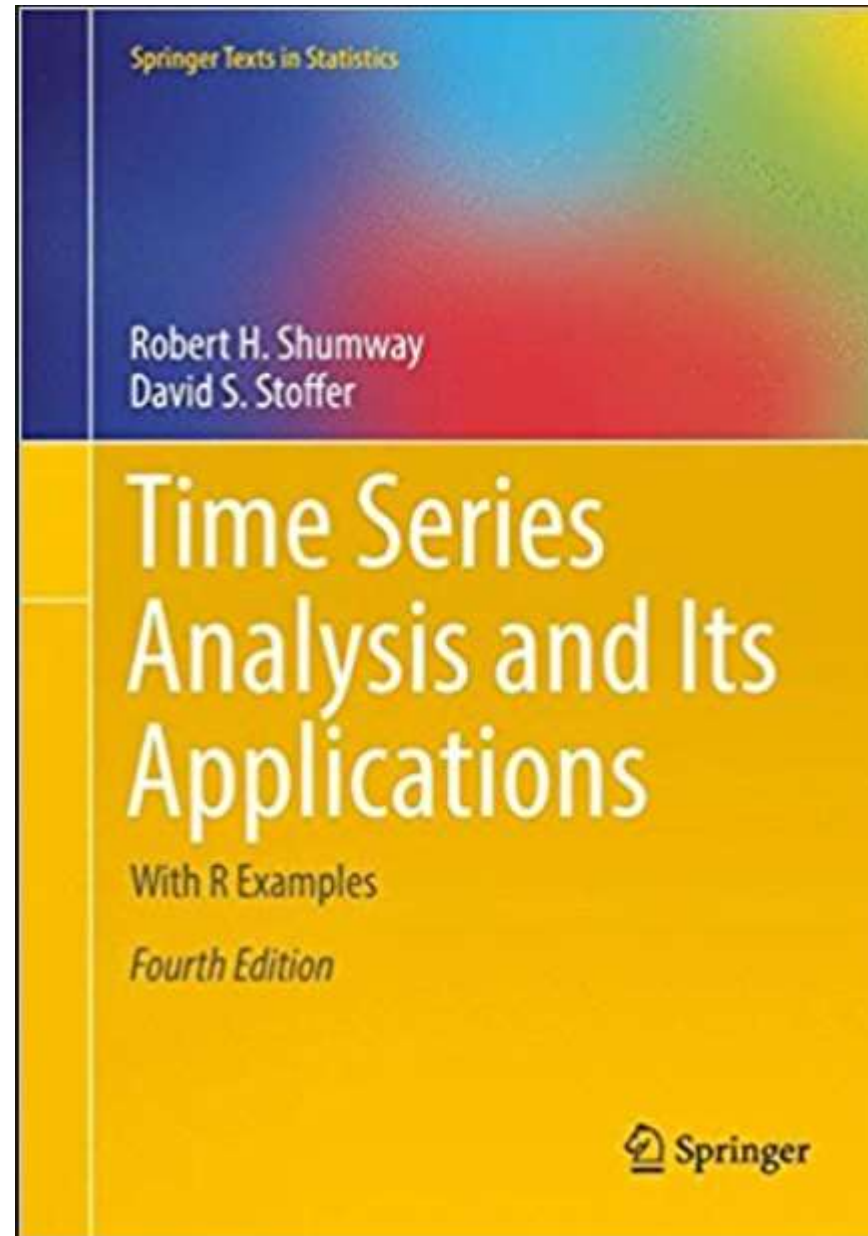
COMPLEMENTAR

- ▶ Hamilton, James. **Time Series Analysis**. Princenton University Press, 1994.
- ▶ Gujarati, Damodar N. **Econometria Básica**. Makron Books Ltda, 2000.
- ▶ Artigos - aplicação:
 - ▶ Trabalhos para discussão - BCB;
 - ▶ ANPEC;
 - ▶ NBER working papers;
 - ▶ Outros.

BIBLIOGRAFIA



BIBLIOGRAFIA



BIBLIOGRAFIA

<http://www.anpec.org.br/novosite/br/encontros>



Associação Nacional dos Centros de Pós-Graduação em Economia

Secretaria Centros Exame Encontros Revista Prêmio Contato

f t

Encontros

ENCONTRO NACIONAL DE ECONOMIA

O Encontro Nacional de Economia é promovido anualmente pela ANPEC com o objetivo de estimular o intercâmbio entre economistas e profissionais de áreas afins. Realizado no mês de dezembro, consiste no mais importante evento científico nacional na sua área. Confira as informações sobre o 45º Encontro Nacional de Economia, que será realizado em dezembro de 2017.

Durante o Encontro, são apresentados trabalhos inéditos selecionados por uma equipe designada para tal fim. Os textos exploram as fronteiras do conhecimento científico na teoria econômica, na economia política e na econometria. Há também a preocupação com a discussão da realidade nacional, que é objeto de painéis e sessões temáticas, além de temas de interesse regional.

O evento conta também com a participação de renomados pesquisadores estrangeiros.

Para assegurar a ampla circulação dos trabalhos apresentados no Encontro, a ANPEC promove a sua publicação em Anais do Encontro, que contém todos os textos discutidos em plenário, e os divulga no seu site.

▫ [Leia mais...](#)

Trabalhos

Estão disponíveis para consulta os trabalhos apresentados durante os últimos encontros nacionais:

[2016](#) [2015](#) [2014](#)

[2013](#) [2012](#) [2011](#)

[2010](#) [2009](#) [2008](#)

[2007](#) [2006](#) [2005](#)

[2004](#) [2003](#) [2001](#)

Confira também a relação dos encontros nacionais promovidos pela ANPEC.

BIBLIOGRAFIA

<http://www.bcb.gov.br/pec/wps/port/default.asp>

BANCO CENTRAL DO BRASIL

Assegurar a estabilidade do poder de compra da moeda e um sistema financeiro sólido e eficiente.

Busca: Busca avançada

[Acesso à Informação do BCB](#) | [Sistema de Metas para a Inflação](#) | [Economia e finanças](#) | [Câmbio e Capitais Internacionais](#) | [Sistema de Pagamentos Brasileiro](#) | [Sistema Financeiro Nacional](#) | [Supervisão do SFN](#) | [Regimes de Resolução e Privatizações](#)

[Início](#) » [Economia e finanças](#) » [Trabalhos para discussão](#) » [Trabalhos para discussão](#)

Trabalhos para discussão

Os Trabalhos para Discussão não devem ser citados como representando as opiniões do Banco Central do Brasil. As opiniões expressas nos trabalhos são exclusivamente do(s) autor(es) e não refletem, necessariamente, a visão do Banco Central.

[Regulamento da Série Trabalhos para Discussão](#) | [Prêmio de Melhores Trabalhos para Discussão](#) | [Cadastramento em Lista de Divulgação](#)

2018 | 2017 | 2016 | 2015 | 2014 | 2013 | 2012 | 2011 | 2010 | 2009 | 2008 | 2007 | 2006 | 2005 | 2004 | 2003 | 2002 | 2001 | 2000

BIBLIOGRAFIA

<http://www.nber.org/papers>



NBER Working Papers

Bibliographic Search for Working Papers only:

Full Text Search for Working Papers:

You can search on expressions of two or more words by putting the expression in quotation marks: e.g., "corporate tax" or "Chinese monetary policy."

AValiação

- ▶ A avaliação será composta dos exercícios práticos realizados em sala de aula;
- ▶ No final de cada tópico, haverá estimação e projeção de variáveis econômicas, exemplos:
 - ▶ Usaremos um modelo ARIMA para projetar o IPCA;
 - ▶ Estimar se há raiz unitária para PPC;
 - ▶ Sistema dinâmico de um modelo semi-estrutural do Banco Central;
 - ▶ Modelo VAR para o IPCA livres;
 - ▶ Modelo VECM para as exportações brasileiras agregadas;
 - ▶ Modelo FAVAR para o PIB
 - ▶ Outros;
- ▶ Cada aluno poderá deixar de fazer até 2 exercícios;
- ▶ Enviar o resultado dos trabalhos para: aperfcarreiras@enap.gov.br

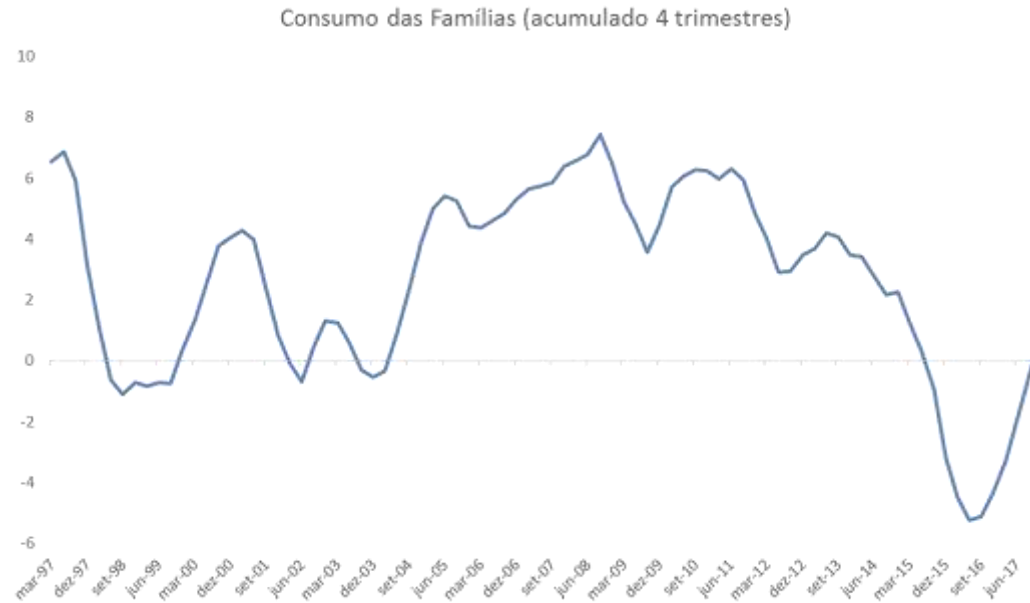
REVISÃO - CONCEITOS BÁSICOS DE ECONOMETRIA E COMANDOS NO “R”

SUMÁRIO DESTA AULA

- ▶ Revisão dos conceitos básicos de Econometria:
 - ▶ Análise de Regressão;
 - ▶ Mínimos Quadrados Ordinários (MQO);
 - ▶ Hipóteses básicas para estimação do modelo;
 - ▶ Análise dos Resíduos;
 - ▶ Relaxando as hipóteses básicas.
- ▶ Revisão dos principais comandos do R-Studio.
 - ▶ Modelo consumo das famílias;
 - ▶ Modelo da balança comercial;

EXEMPLO DA ECONOMIA BRASILEIRA

- ▶ Estimar relação entre consumo das famílias e outras variáveis econômicas;



- ▶ Quais variáveis posso utilizar?

O que a teoria nos diz? (Artigo do BCB)

- ▶ Medium-Size Macroeconomic Model for the Brazilian Economy (2003)

$$\ln\left(\frac{C_t}{\bar{Y}_{t-1}^d}\right) = \alpha_0 + \alpha_1 \cdot \ln\left(\frac{C_{t-1}}{\bar{Y}_{t-2}^d}\right) + \alpha_2 \cdot r_{t-1} + \sum_{i=1}^3 \beta_i \cdot Seas_i + \alpha_3 \cdot D_{96}$$

Where:

C_t household consumption;

\bar{Y}_t^d disposable potential income: $\bar{Y}_t^d = \bar{Y}_t - T_t^d$;

T_t^d direct taxes;

\bar{Y}_t potential output;

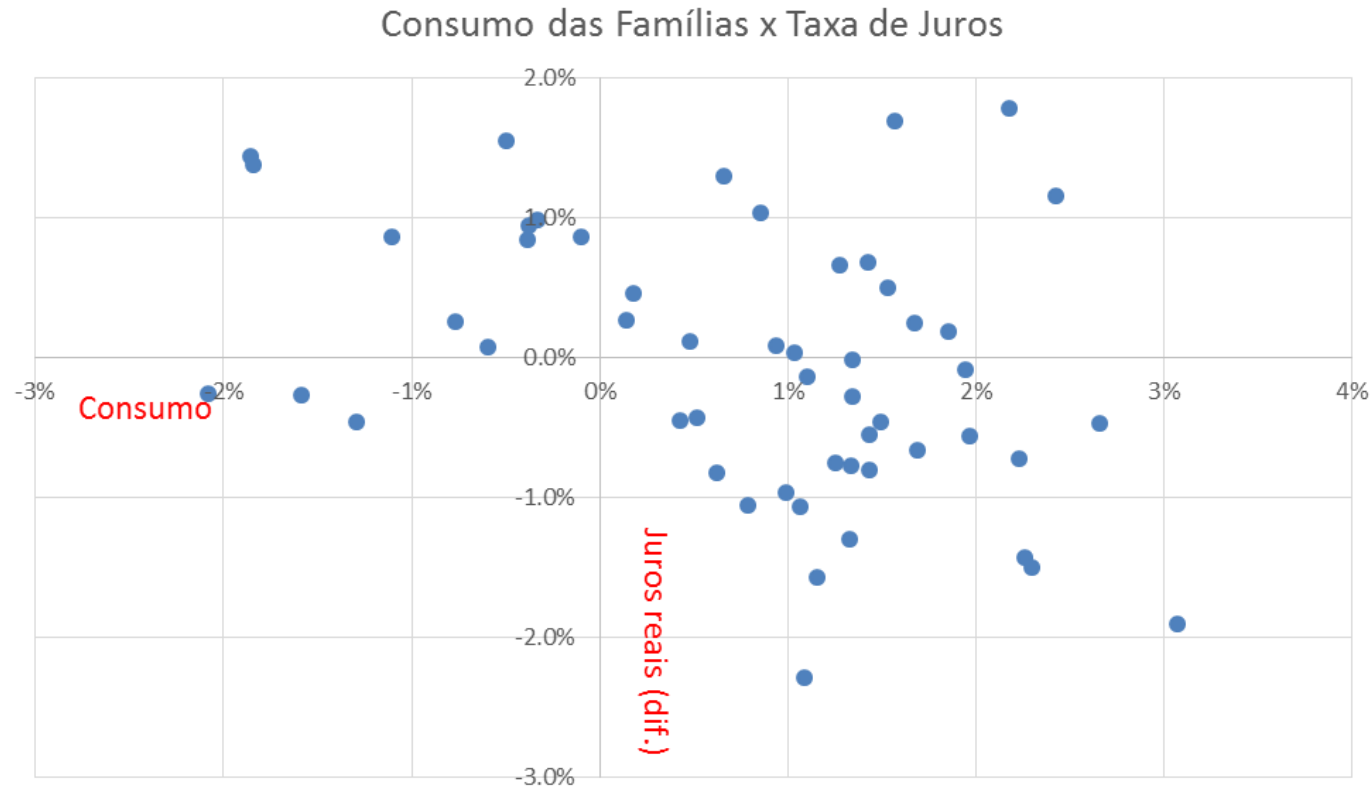
r_t short term real interest

$$rate^{\circ} : r_t = \frac{Selic_t}{4} - \begin{cases} \pi_t & , \text{ from 1991Q1 to 1995Q1} \\ \sum_{i=0}^3 \frac{\pi_{t-i}}{4} & , \text{ from 1995Q2 on} \end{cases} ;$$

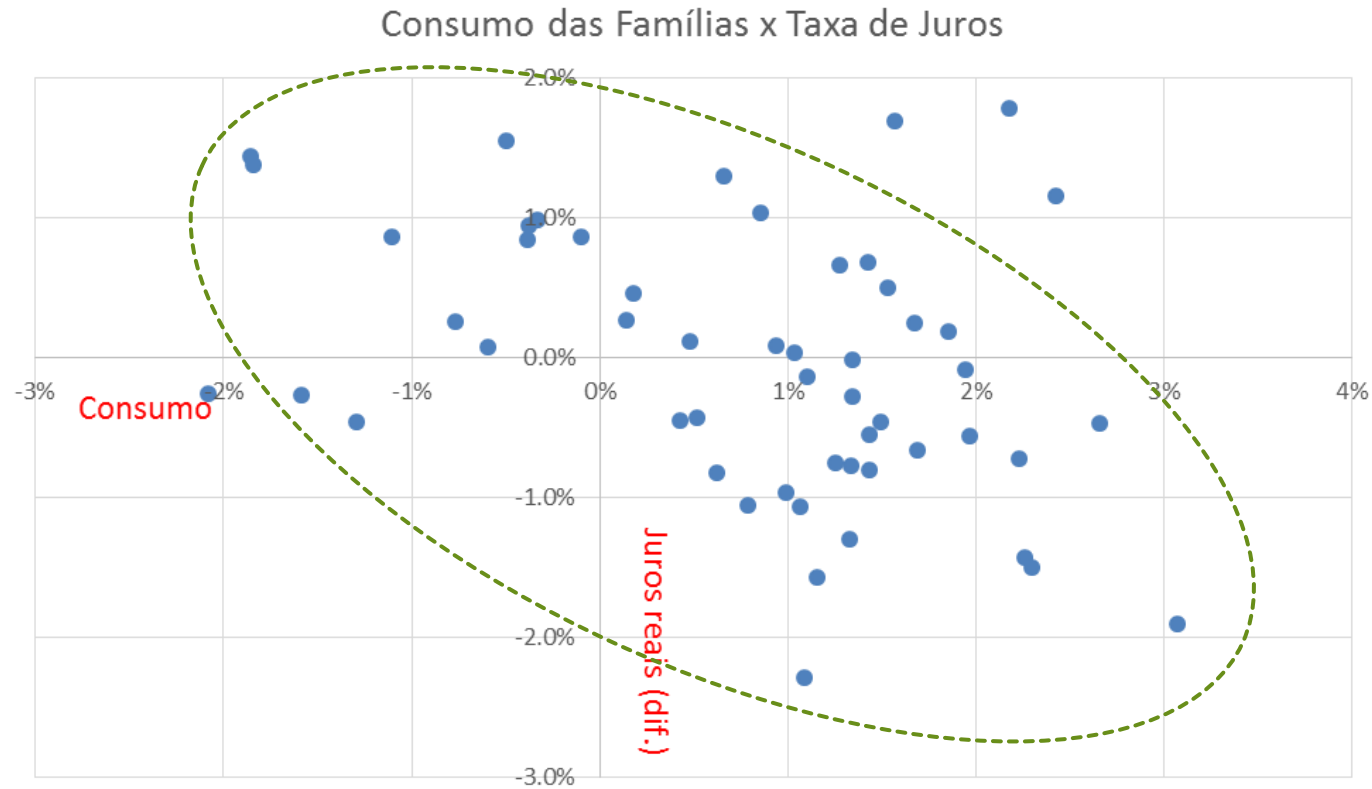
$Selic_t$ annualized Brazilian over night nominal interest rate;

D_{96} step dummy from year 1996 on.

Juros reais x Consumo das famílias



Juros reais x Consumo das famílias

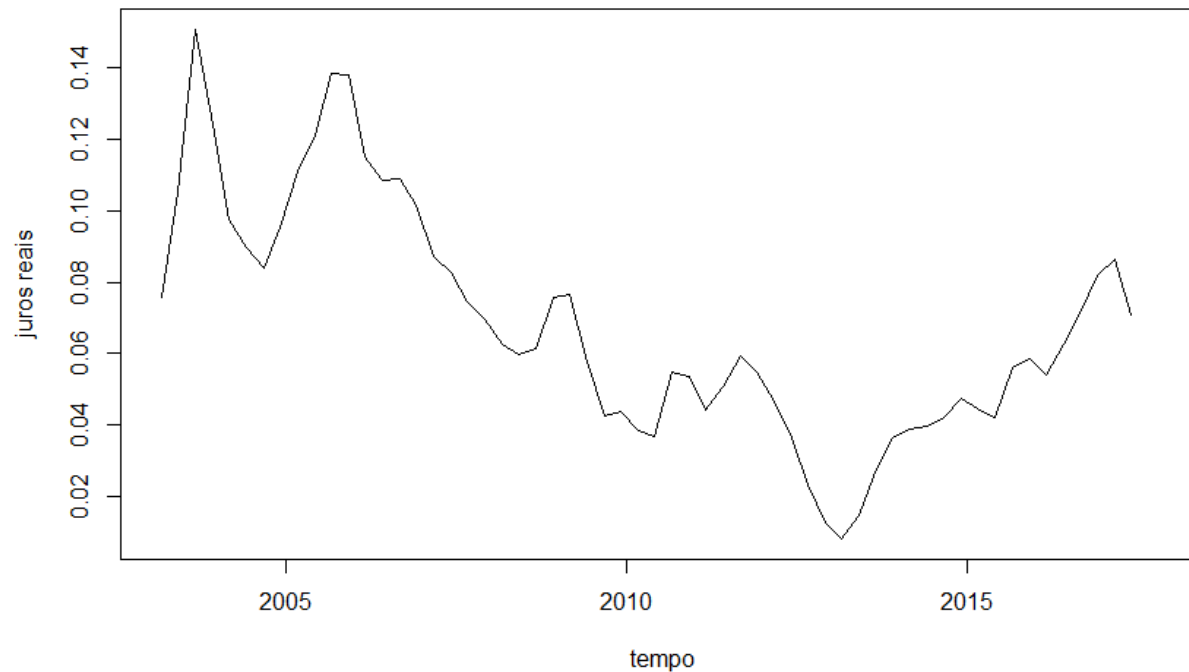


Vamos para o R Studio

- ▶ Baixando os dados...
 - ▶ `setwd("C:/diretorio1/diretorio2")` - determinar o diretório: Não se esqueça, **inverter** as barras!!!
 - ▶ `list.files()` - quais arquivos que estão no seu diretório
 - ▶ `X<-read.csv("aula1_dados.csv",sep=";", dec=".", head=TRUE)` - baixando os dados.
 - ▶ Em Data, o X pode estar somente com uma coluna. Isso tem a ver com o decimal.
 - ▶ `X <Enter>` - mostra todos os dados do data.frame
 - ▶ `X["cons_familias"]` ou `X$cons_familias`
 - ▶ `summary(X)` ou `summary(X$juros_reais)`
 - ▶ `cons_familias=ts(X$cons_familias,start=c(2003,2),frequency=4) # série temporal`
- ▶ O que é um dataframe
 - ▶ Dataframe pode ser visto como um objeto de forma retangular com linhas representado as observações e colunas representando as variáveis;
 - ▶ Diferente de uma matriz, um objeto do tipo dataframe pode conter variáveis de diferentes tipos (números e textos).

Gráfico dos juros reais

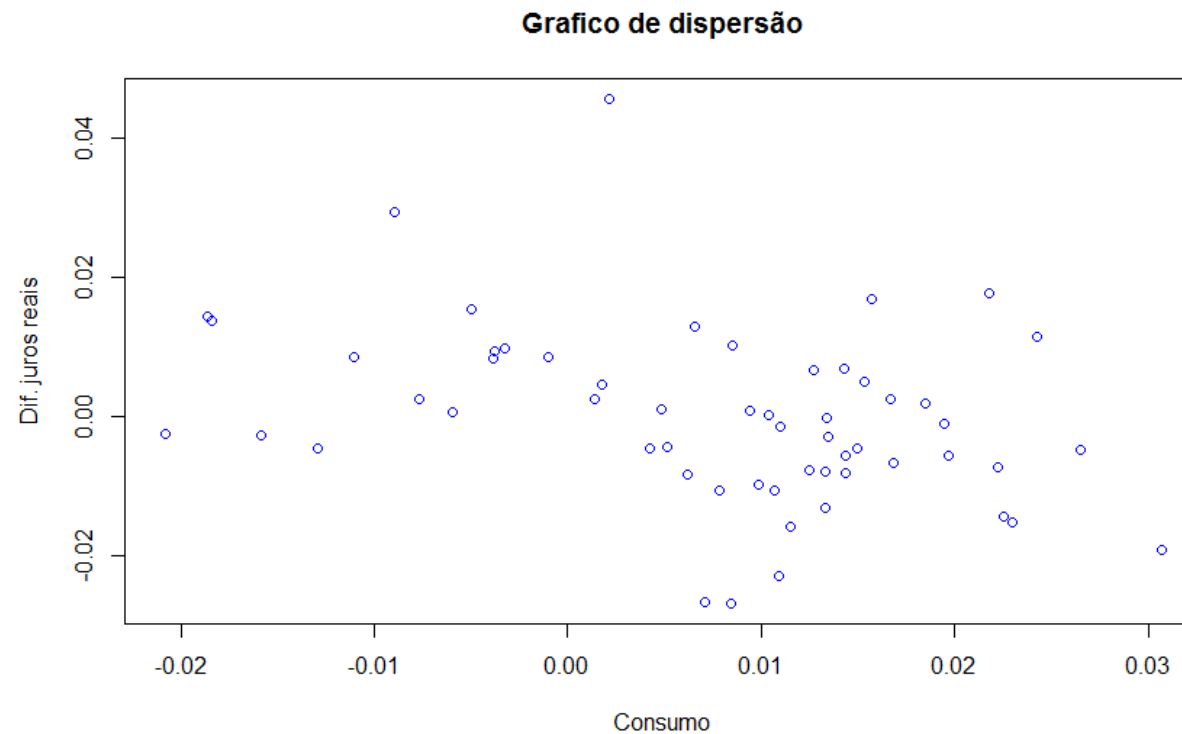
- ▶ `X$data<-as.Date(X$data,'%d/%b/%y')` - 'b' é mês (mmm - abr), 'm' (mm-04)
- ▶ `plot(X$data,X$juros_reais,type="l",xlab="tempo",ylab = "juros reais")`



- ▶ Opções (type = "l", "p", "b")

Gráfico de dispersão

- ▶ `plot(X$cons_familias,X$dif_juros_reais,type="p",main="Gráfico de dispersão",xlab = "Consumo",ylab = "Dif. juros reais",col="blue")`



Regressão Linear (Teoria)

- ▶ Função de Regressão Populacional (FRP):

- ▶ $Y_t = \beta_0 + \beta_1 X_t + u_t$

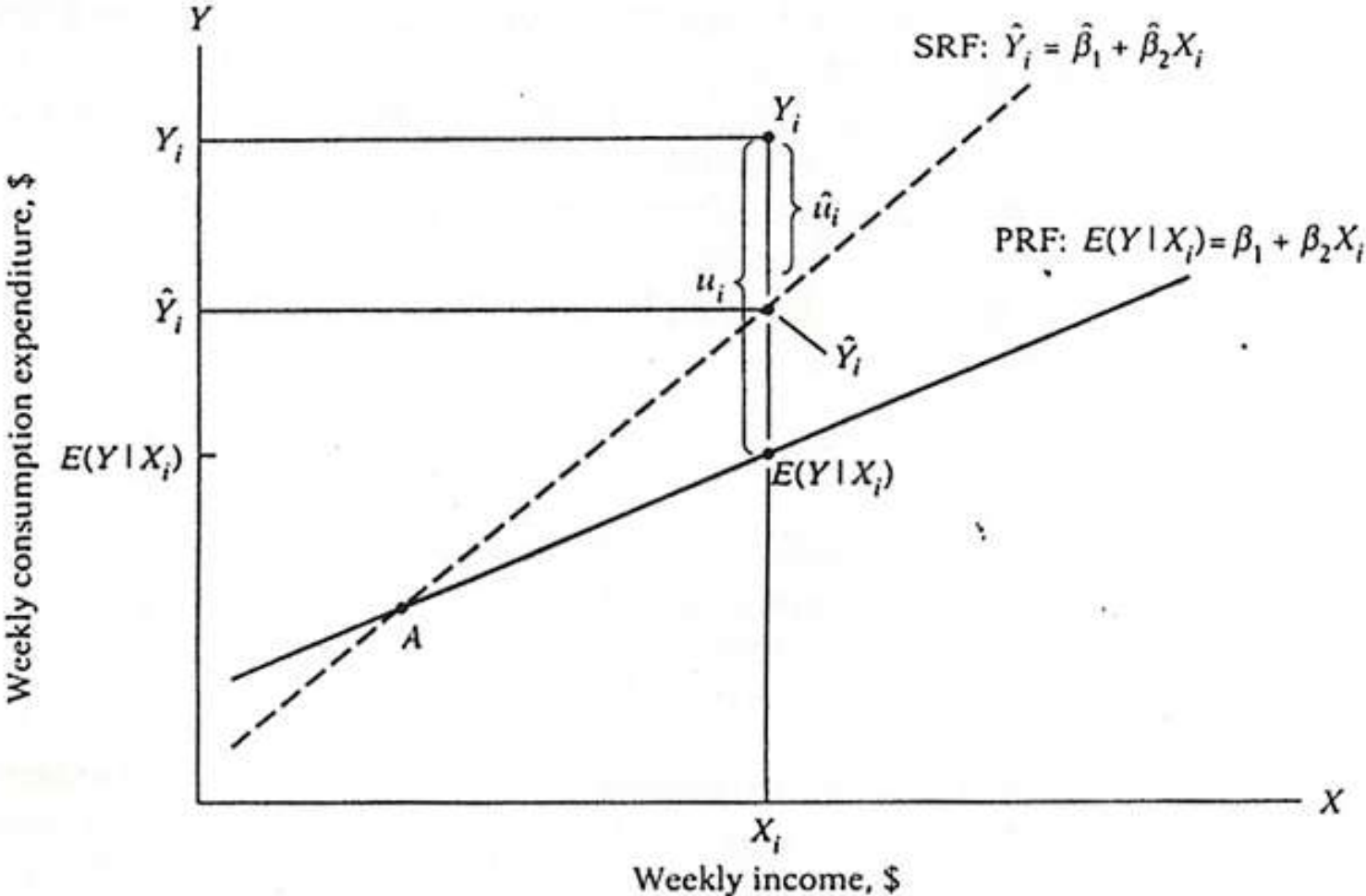
- ▶ Essa função não é diretamente observável.

- ▶ Função de Regressão Amostral (FRA):

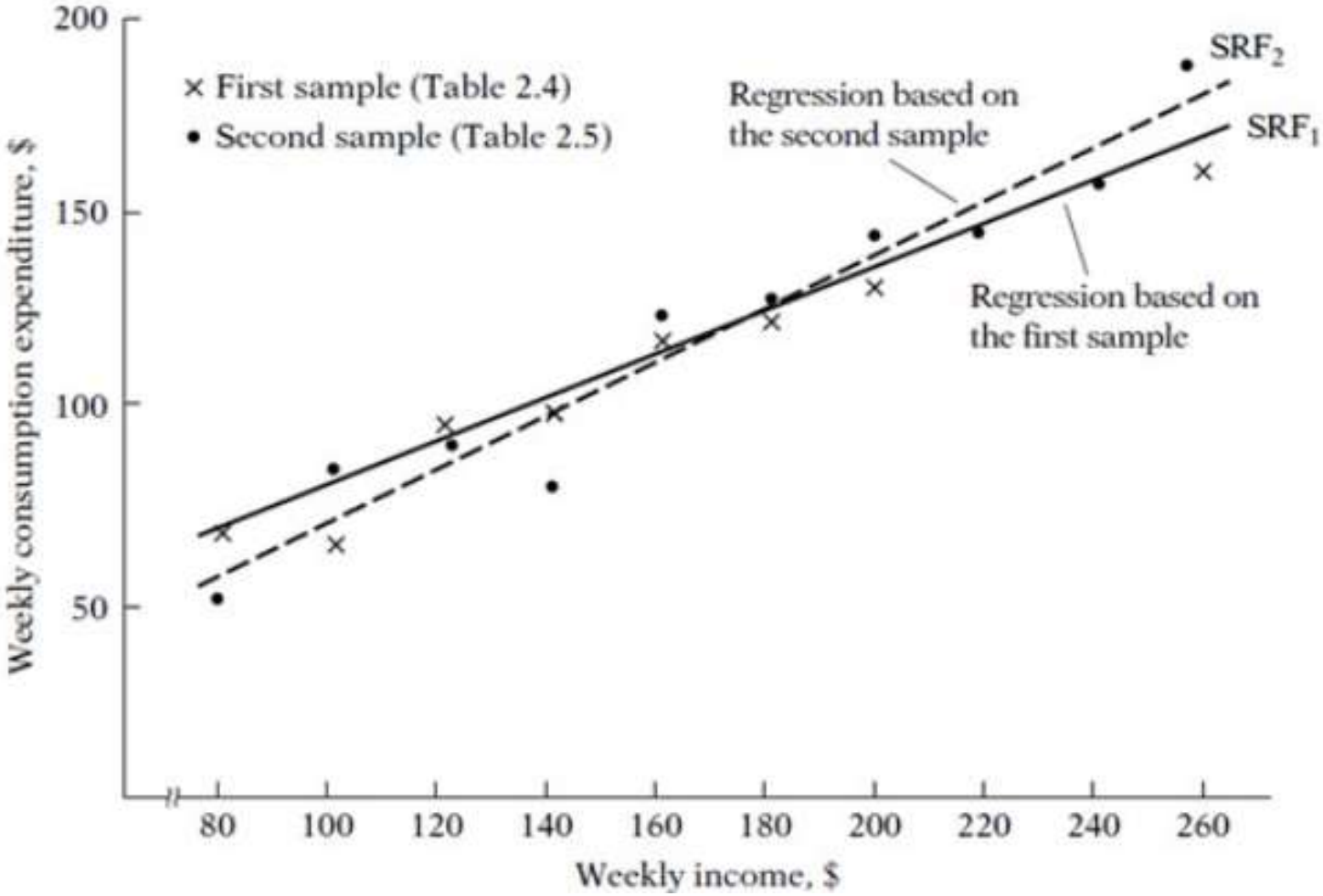
- ▶ $Y_t = \hat{\beta}_0 + \hat{\beta}_1 X_t + \hat{u}_t$

- ▶ Devido as alterações amostrais, os coeficientes da função de regressão linear se alteram.
- ▶ Por isso, consideramos intervalos de confiança para que a estimação fique mais “próxima” da relação real.

Regressão Linear (Teoria) - FRP x FRA



Regressão Linear (Teoria) - FRA1 x FRA2



Regressão Linear Simples

▶ Como estimar a Função de Regressão Amostral:

▶ Minimizar a seguinte equação:

$$\text{▶ } \sum \hat{u}_t^2 = \sum (Y_t - \hat{Y}_t)^2 = \sum (Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_t)^2$$

▶ Resíduos maiores, exercerão maior influência na estimação.

▶ Derivando em relação aos coeficientes, temos que:

$$\text{▶ } \frac{\partial(\sum \hat{u}_t^2)}{\partial \hat{\beta}_0} = -2 \sum (Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_t) = -2 \sum \hat{u}_t = 0$$

$$\text{▶ } \sum Y_t = T \hat{\beta}_0 + \hat{\beta}_1 \sum X_t$$

$$\text{▶ } \frac{\partial(\sum \hat{u}_t^2)}{\partial \hat{\beta}_1} = -2 \sum (Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_t) X_t = -2 \sum \hat{u}_t X_t = 0$$

$$\text{▶ } \sum Y_t X_t = \hat{\beta}_0 \sum X_t + \hat{\beta}_1 \sum X_t^2$$

Regressão Linear Simples

▶ Resolvendo as equações:

$$\widehat{\beta}_1 = \frac{\sum(Y_t - \bar{Y})(X_t - \bar{X})}{\sum(X_t - \bar{X})^2} = \frac{Cov(Y, X)}{Var(X)}$$

$$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1 \bar{X}$$

▶ O coeficiente de correlação tem relação próxima com $\widehat{\beta}_1$

$$\rho = \frac{Cov(Y, X)}{\sqrt{Var(X)var(Y)}}$$

▶ O coeficiente de correlação está diretamente ligado ao R^2 .

$$\rho = \pm\sqrt{R^2}$$

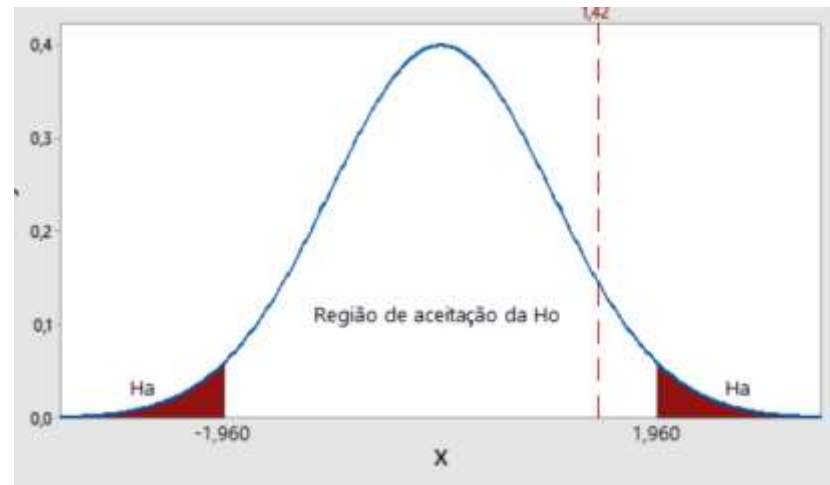
Teste de significância

► Variância do erro:

$$\hat{\sigma}^2 = \frac{\sum(\hat{u}_t)^2}{T-k}$$

$$\hat{\sigma}_{\hat{\beta}_1}^2 = \frac{\hat{\sigma}^2}{\sum(X_{1t} - \bar{X})^2}$$

$$Z = \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}}$$



Estimando a regressão linear simples

- ▶ `reg1=lm(X$cons_familias~X$dif_juros_reais)`
- ▶ `summary(reg1)`

```
call:
lm(formula = X$cons_familias ~ X$dif_juros_reais)

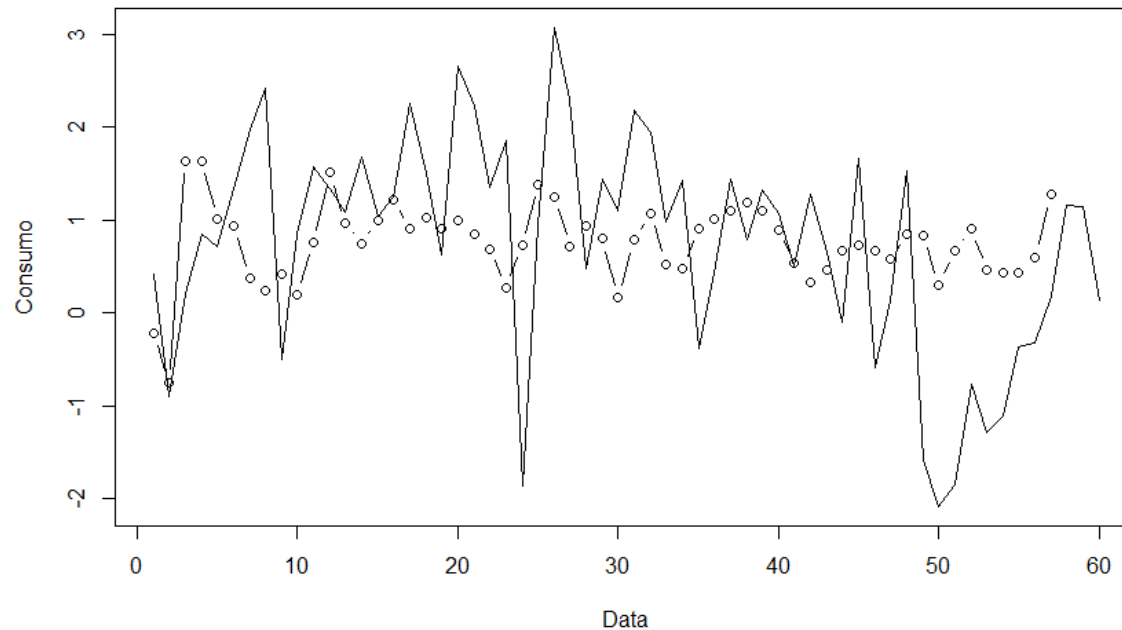
Residuals:
    Min       1Q   Median       3Q      Max
-2.9181 -0.6817  0.2157  0.8997  2.0540

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.7525     0.1505   4.999 6.22e-06 ***
X$dif_juros_reais -0.3300     0.1185  -2.784 0.00735 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.137 on 55 degrees of freedom
Multiple R-squared:  0.1235,    Adjusted R-squared:  0.1076
F-statistic:  7.75 on 1 and 55 DF,  p-value: 0.00735
```

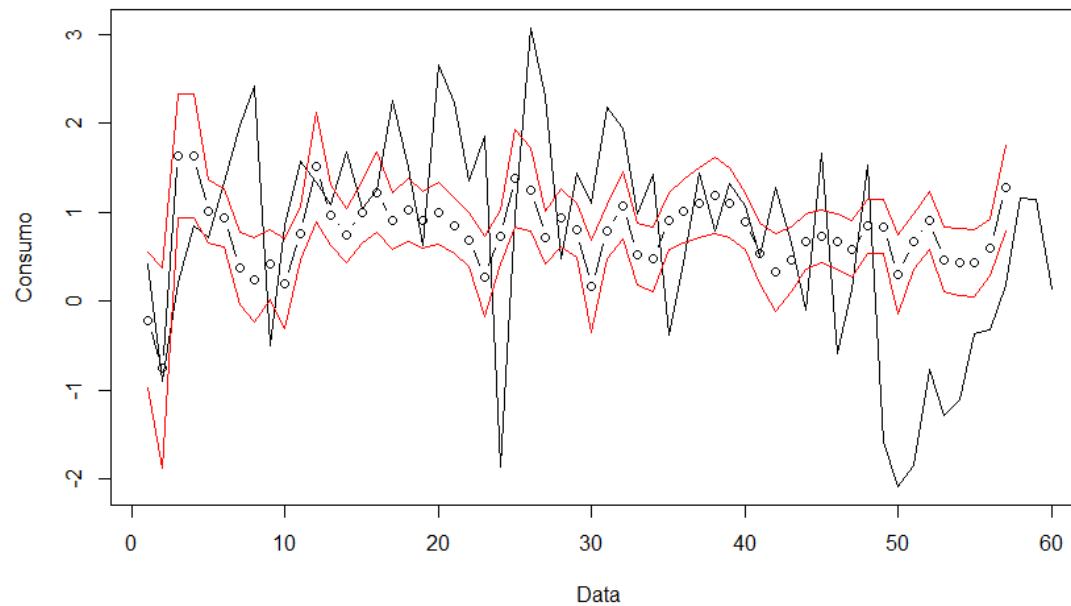
ESTIMADO X OBSERVADO

- ▶ Resultado da estimação - séries projetadas e o intervalo de confiança:
 - ▶ `X1fit<-predict.lm(reg1,interval="confidence")` #resultado da estimação
- ▶ Colocando os dados em um dataframe:
 - ▶ `X1fit<-data.frame(X1fit)`
- ▶ Gráfico do valor real e estimado:
 - ▶ `plot(X$cons_familias,type='l',xlab="Data",ylab="Consumo")`
 - ▶ `points(X1fit$fit,type='b')`



ESTIMADO X OBSERVADO

- ▶ Colocando os intervalos no gráfico
 - ▶ `points(X1fit$lwr,type='l',col="red")`
 - ▶ `points(X1fit$upr,type='l',col="red")`



- ▶ Foi um bom ajuste?

O QUE A TEORIA NOS DIZ? (Sach & Larrain)

- ▶ Restrição Orçamentária no modelo de dois períodos
 - ▶ Período 1: $Y_1 - C_1 = W_1 - C_1 = D_1$
 - ▶ Período 2: $Y_2 - C_2 = W_2 - C_2 + rD_1 = D_2 = 0$
 - ▶ Onde C é o consumo, W o salário, r a taxa de juros reais, D é a dívida (crédito)
- ▶ Solução do sistema das equações acima:
 - ▶ $C_1 + \frac{C_2}{(1+r)} = W_1 + \frac{W_2}{(1+r)}$
- ▶ Podemos especificar no nosso modelo da seguinte forma:
- ▶ Consumo é função do salário, da taxa de juros reais e do crédito.

ESTIMANDO A REGRESSÃO MÚLTIPLA

▶ `reg2<-lm(X$cons_familias~X$dif_juros_reais+X$massa_sal+X$credito)`

▶ `summary(reg2)`

```
call:
lm(formula = x$cons_familias ~ x$dif_juros_reais + x$massa_sal +
    x$credito)

Residuals:
    Min       1Q   Median       3Q      Max
-2.92142 -0.54472  0.01643  0.57462  2.10780

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      0.35104    0.16694   2.103 0.040252 *
x$dif_juros_reais -0.20957    0.10746  -1.950 0.056460 .
x$massa_sal      -0.02827    0.08156  -0.347 0.730223
x$credito         0.22284    0.05668   3.931 0.000247 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

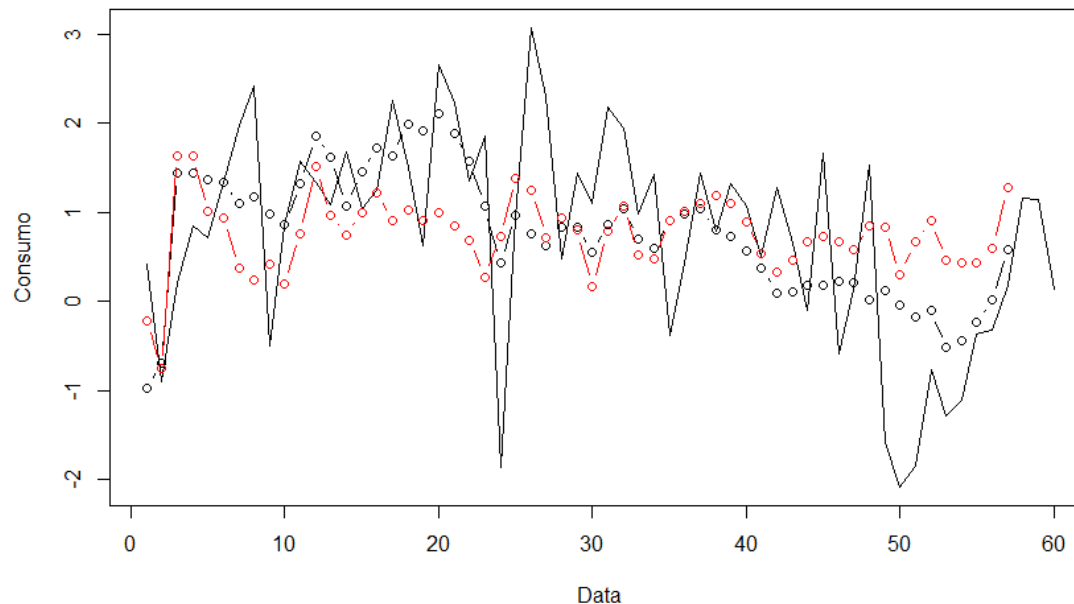
Residual standard error: 0.9911 on 53 degrees of freedom
Multiple R-squared:  0.3578,    Adjusted R-squared:  0.3214
F-statistic: 9.842 on 3 and 53 DF,  p-value: 2.913e-05
```

▶ `resid1<-residuals(reg1)`

▶ `resid2<-residuals(reg2)`

COMPARANDO OS MODELOS

- ▶ Resultado da estimação - series projetadas e o intervalo de confiança:
 - ▶ `X2fit<-predict.lm(reg2,interval="confidence")` #resultado da estimação
- ▶ Colocando os dados em um dataframe:
 - ▶ `X2fit<-data.frame(X2fit)`
- ▶ Comparando os modelos estimados:
 - ▶ `plot(X$cons_familias,type='l',xlab="Data",ylab="Consumo")`
 - ▶ `points(X2fit$fit,type='b')`
 - ▶ `points(X1fit$fit,type='b',col="red")`



O QUE FALTA NO MODELO?

- ▶ Hábitos?
 - ▶ O consumo passado importa.
 - ▶ Defasagem da variável dependente.
- ▶ Efeito da taxa de juros e da massa salarial
 - ▶ O impacto da taxa de juros afeta as decisões de consumo com certa defasagem.

Regressão Linear Múltipla

▶ Como estimar a Função de Regressão Amostral:

▶ Minimizar a seguinte equação:

$$\sum \hat{u}_t^2 = \sum (Y_t - \hat{Y}_t)^2 = \sum (Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_{1t} - \hat{\beta}_2 X_{2t})^2$$

▶ Resíduos maiores, exercerão maior influência na estimação.

▶ Derivando em relação aos coeficientes, temos que:

$$\frac{\partial(\sum \hat{u}_t^2)}{\partial \hat{\beta}_0} = -2 \sum (Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_{1t} - \hat{\beta}_2 X_{2t}) = -2 \sum \hat{u}_t = 0$$

$$\sum Y_t = T \hat{\beta}_0 + \hat{\beta}_1 \sum X_t + \hat{\beta}_2 \sum X_{2t}$$

$$\frac{\partial(\sum \hat{u}_t^2)}{\partial \hat{\beta}_{jt}} = -2 \sum (Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_{1t} - \hat{\beta}_2 X_{2t}) X_{jt} = -2 \sum \hat{u}_t X_{jt} = 0$$

$$\sum Y_t X_{jt} = \hat{\beta}_0 \sum X_{jt} + \hat{\beta}_1 \sum X_{1t} X_{jt} + \hat{\beta}_2 \sum X_{2t} X_{jt} \quad \text{Para } j=1,2$$

Variância x Covariância

Teoria - Regressão Múltipla

- Solução das equações:

Covariância de X_{1t} e Y_{2t} e variância X_{2t}

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_{i1} \sum_{i=1}^n x_{i2}^2 - \sum_{i=1}^n y_i x_{i2} \sum_{i=1}^n x_{i1} x_{i2}}{\sum_{i=1}^n x_{i1}^2 \sum_{i=1}^n x_{i2}^2 - \left(\sum_{i=1}^n x_{i1} x_{i2} \right)^2}$$

Variância de X_{1t} e X_{2t} Covariância de X_{1t} e X_{2t}

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n y_i x_{i2} \sum_{i=1}^n x_{i1}^2 - \sum_{i=1}^n y_i x_{i1} \sum_{i=1}^n x_{i2} x_{i1}}{\sum_{i=1}^n x_{i1}^2 \sum_{i=1}^n x_{i2}^2 - \left(\sum_{i=1}^n x_{i1} x_{i2} \right)^2}$$

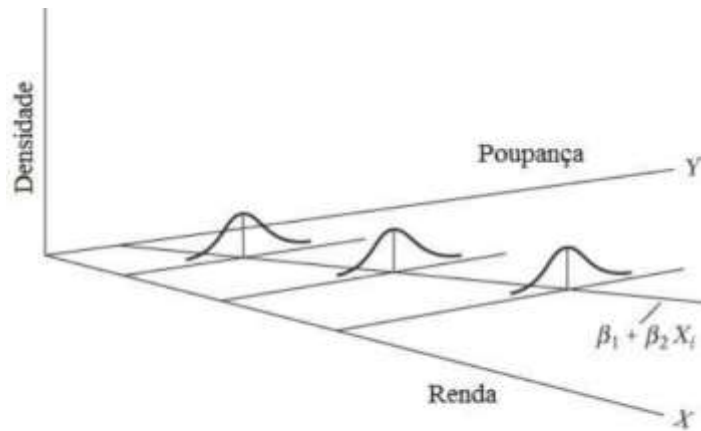
Teoria - Regressão Múltipla

▶ Propriedades:

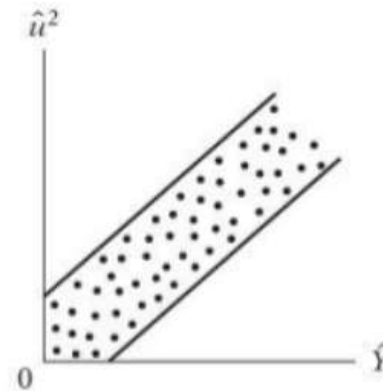
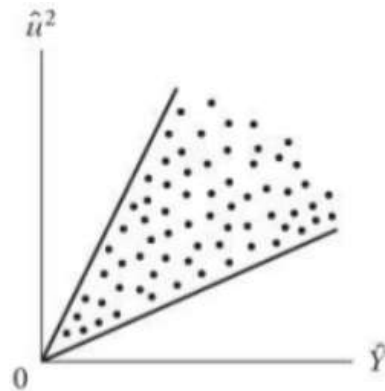
- ▶ A reta da regressão passa pelas médias de Y e Xs;
 - ▶ Quando se retira a média das variáveis, a reta de regressão passa pela origem;
- ▶ $\sum \hat{u}_t = 0$ - média dos resíduos é igual a zero;
- ▶ O resíduo não tem correlação com as variáveis explicativa, $\text{cov}(\hat{u}_t, X_{jt}) = 0$
- ▶ Os estimadores de MQO são não lineares, como também têm variância mínima.
 - ▶ Melhor estimador MELNV
- ▶ Podemos dividir a variação de Y em componentes explicados (SQE) e parte não explicada (SQR);
 - ▶ $R^2 = \frac{SQE}{SQE+SQR}$

Hipóteses básicas para estimação

1. Modelo de regressão é linear nos parâmetros;
2. Os valores de X são fixados em amostragem repetida.
 - ▶ Os valores da regressão são condicionados aos valores de X .
3. Valor médio zero do resíduo:
 - ▶ $E(u_t|X_t) = 0$
4. Homocedasticidade ou variância igual de u_t :



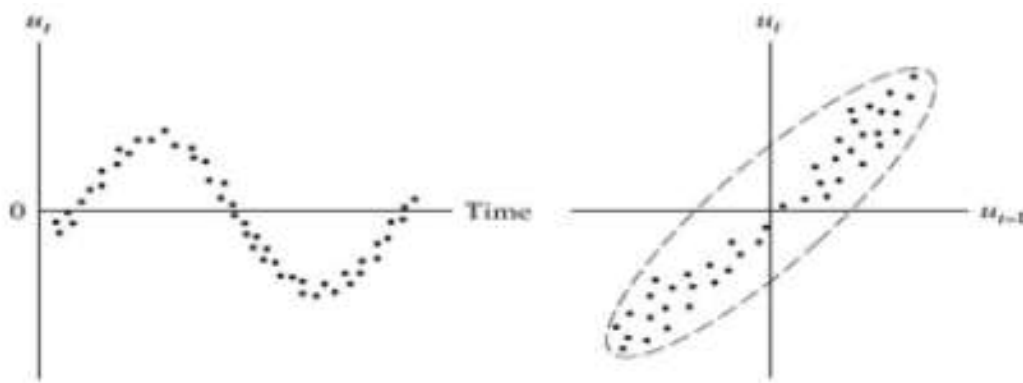
HOMOCEDASTICIDADE



Hipóteses básicas para estimação

5. Nenhuma autocorrelação entre os resíduos.

► $E(u_t, u_{t+k} | X_t, X_{t+k}) = E[u_t - E(u_t) | X_t] E[u_{t+k} - E(u_{t+k}) | X_{t+k}] = 0$



6. O número de observações T tem que ser maior do que o total de parâmetros

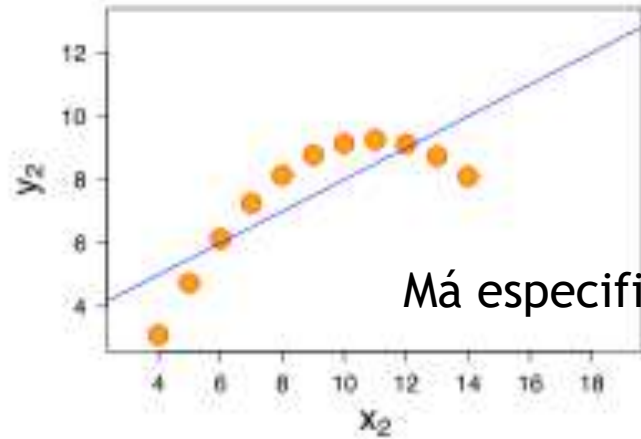
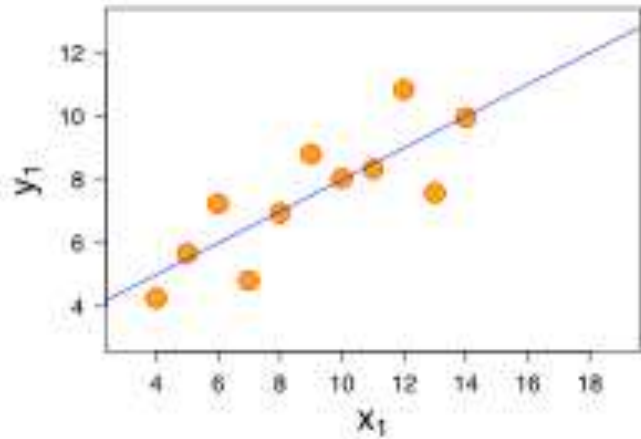
7. Variabilidade dos valores X, ou seja $\text{var}(X) > 0$;

8. Não há multicolinearidade perfeita: $X_{1t} = kX_{2t}$

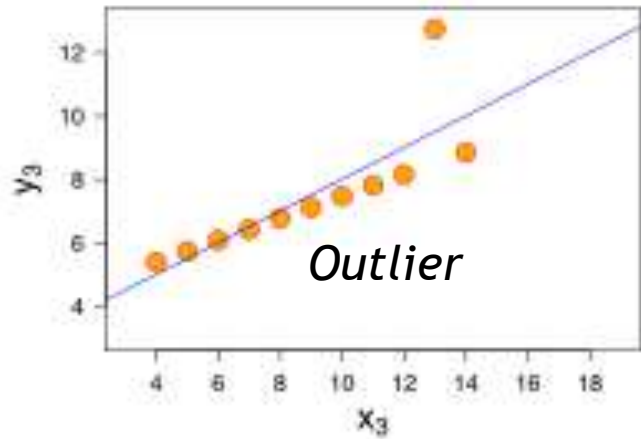
► R^2 elevado, mas estatística t baixa.

9. Modelo corretamente especificado.

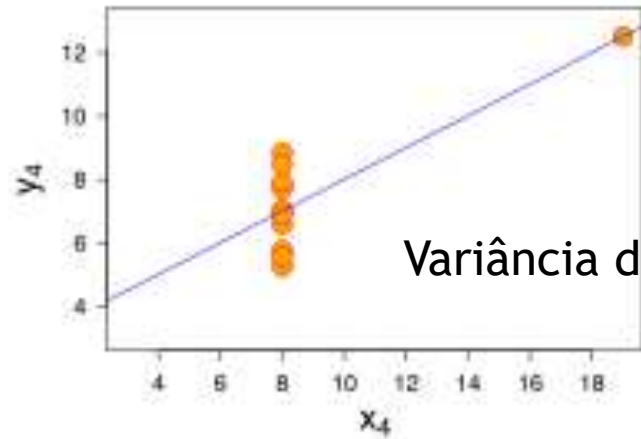
Diagnosticos gráficos



Má especificação



Outlier



Variância de X

Erro de especificação

▶ Problema:

▶ Equação verdadeira (FRP):

$$\text{▶ } Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t}$$

▶ Equação estimada (FRA):

$$\text{▶ } Y_t = \hat{\beta}_0 + \hat{\beta}_1 X_{1t}$$

▶ Há problema?

$$\text{▶ Sim: } \text{corr}(X_{1t}, X_{2t}) \neq 0 \rightarrow \hat{\beta}_1 = \beta_1 + k\beta_2$$

$$\text{▶ Não: } \text{corr}(X_{2t}, X_{2t}) = 0 \rightarrow \hat{\beta}_1 = \beta_1$$

▶ Coeficientes viesados!

▶ Problemas de má especificação;

▶ Variável não observada;

▶ Erro de medida.

Teste RESET

- ▶ Ramsey Regression Equation Specification Error Test

- ▶ Equação principal

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + u$$

- ▶ Equação auxiliar:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \delta_1 \hat{y}^2 + \delta_2 \hat{y}^3 + u$$

- ▶ Hipótese:

$$H_0 : \delta_1 = 0, \delta_2 = 0$$

Distribuição F(2,n-k-3)

Teste RESET

- ▶ Baixar o pacote “lmtest”
 - ▶ `install.packages("lmtest")`
- ▶ Ativar o pacote
 - ▶ `library("lmtest")`
- ▶ Calcular para regressão 1 e 2:
 - ▶ `resettest(reg1,power=2:3,type="regressor",data="d")`
 - ▶ `resettest(reg2,power=2:3,type="regressor",data="d")`

```
> resettest(reg1,power=2:3,type="regressor",data="d")
```

```
RESET test
```

```
data: reg1  
RESET = 0.92633, df1 = 2, df2 = 53, p-value = 0.4023
```

```
> resettest(reg2,power=2:3,type="regressor",data="d")
```

```
RESET test
```

```
data: reg2  
RESET = 2.512, df1 = 6, df2 = 47, p-value = 0.03433
```

Teste BPG - heterocedasticidade

- ▶ Breusch-Pagan-Godfrey Test
- ▶ Utiliza o modelo estimado

$$y = X\beta + \varepsilon.$$

- ▶ Equação auxiliar:

$$e_i^2 = \gamma_1 + \gamma_2 z_{2i} + \cdots + \gamma_p z_{pi} + \eta_i$$

- ▶ Hipótese:

$$\gamma_2 = \cdots = \gamma_p = 0$$

- ▶ Resultado do teste tem a distribuição χ_{p-1}^2

$$LM = nR^2$$

Onde p é o número de parâmetros

Teste BPG - heterocedasticidade

- ▶ Pacote “lmtest” já está ativo
- ▶ Calcular para regressão 1 e 2:
 - ▶ `bptest(reg1)`
 - ▶ `bptest(reg2)`

```
> bptest(reg1)
          studentized Breusch-Pagan test
data:  reg1
BP = 1.3172, df = 1, p-value = 0.2511
> bptest(reg2)
          studentized Breusch-Pagan test
data:  reg2
BP = 1.0426, df = 3, p-value = 0.7909
```

- ▶ Descrição do comando

Usage

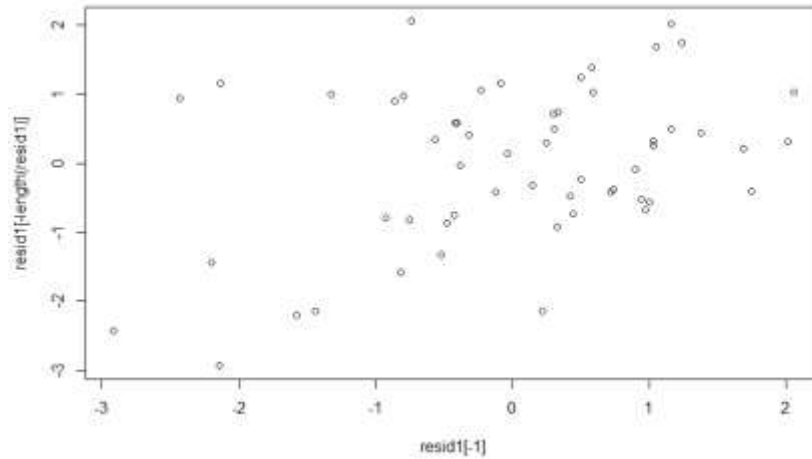
```
bptest(formula, varformula = NULL, studentize = TRUE, data = list())
```

Arguments

<code>formula</code>	a symbolic description for the model to be tested (or a fitted "lm" object).
<code>varformula</code>	a formula describing only the potential explanatory variables for the variance (no dependent variable needed). By default the same explanatory variables are taken as in the main regression model.

Gráfico de autocorrelação dos resíduos

- ▶ `plot(resid1[-length(resid1)]~resid1[-1])`



- ▶ `plot(resid2[-length(resid2)]~resid2[-1])`

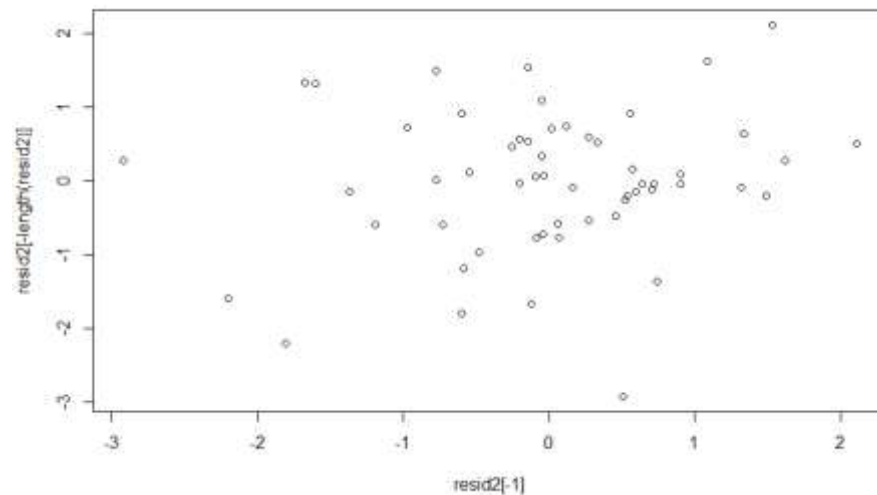


Gráfico de autocorrelação dos resíduos

► `rd1<-lm(resid1[-length(resid1)]~resid1[-1])`

► `summary(rd1)`

```
Call:
lm(formula = resid2[-length(resid2)] ~ resid2[-1])

Residuals:
    Min       1Q   Median       3Q      Max
-2.98649 -0.49465  0.02278  0.57533  1.89032

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.01005    0.12932  -0.078   0.938
resid2[-1]   0.14826    0.13414   1.105   0.274

Residual standard error: 0.9678 on 54 degrees of freedom
Multiple R-squared:  0.02212,    Adjusted R-squared:  0.004014
F-statistic: 1.222 on 1 and 54 DF,  p-value: 0.2739
```

► `rd2<-lm(resid2[-length(resid2)]~resid2[-1])`

► `summary(rd2)`

```
Call:
lm(formula = resid2[-length(resid2)] ~ resid2[-1])

Residuals:
    Min       1Q   Median       3Q      Max
-2.98649 -0.49465  0.02278  0.57533  1.89032

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.01005    0.12932  -0.078   0.938
resid2[-1]   0.14826    0.13414   1.105   0.274

Residual standard error: 0.9678 on 54 degrees of freedom
Multiple R-squared:  0.02212,    Adjusted R-squared:  0.004014
F-statistic: 1.222 on 1 and 54 DF,  p-value: 0.2739
```

Teste BG - correlação serial

- ▶ Breusch-Godfrey Test

- ▶ Utiliza o modelo estimado

$$y = X\beta + \varepsilon.$$

- ▶ Equação auxiliar:

$$\hat{u}_t = \alpha_0 + \alpha_1 X_{t,1} + \alpha_2 X_{t,2} + \rho_1 \hat{u}_{t-1} + \rho_2 \hat{u}_{t-2} + \dots + \rho_p \hat{u}_{t-p} + \varepsilon_t$$

- ▶ Hipótese:

$$H_0 : \{\rho_i = 0 \text{ for all } i\}$$

- ▶ Resultado do teste tem a distribuição

$$nR^2 \sim \chi_p^2.$$

Onde p é o número de parâmetros

Teste BG - autocorrelação

- ▶ Pacote “lmtest” já está ativo
- ▶ Calcular para regressão 1 e 2:
 - ▶ `bgtest(reg1)`
 - ▶ `bgtest(reg2)`

```
> bgtest(reg1, order=4)
```

```
      Breusch-Godfrey test for serial correlation of order up to 4
```

```
data:  reg1
```

```
LM test = 13.791, df = 4, p-value = 0.007992
```

```
> bgtest(reg2, order=4)
```

```
      Breusch-Godfrey test for serial correlation of order up to 4
```

```
data:  reg2
```

```
LM test = 3.501, df = 4, p-value = 0.4777
```

Teste Durbin Watson

- ▶ Verifica se há autocorrelação na primeira defasagem.

```
> dwtest(reg2)
```

```
      Durbin-watson test
```

```
data:  reg2
```

```
DW = 1.697, p-value = 0.08588
```

```
alternative hypothesis: true autocorrelation is greater than 0
```


Inserindo defasagens na estimação

- ▶ `reg2_1<-lm(X$cons_familias[-length(X$cons_familias)]~X$cons_familias[-1]+X$juros_reais[-1]+X$massa_sal[-1]+X$credito[-1])`
- ▶ `summary(reg2_1)`

```
lm(formula = X$cons_familias[-length(X$cons_familias)] ~ X$cons_familias[-1] +
    X$juros_reais[-1] + X$massa_sal[-1] + X$credito[-1])

Residuals:
    Min       1Q   Median       3Q      Max
-2.39124 -0.32029  0.07342  0.35607  2.35116

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.62713    0.30331   2.068 0.043768 *
X$cons_familias[-1] 0.14174    0.12141   1.167 0.248458
X$juros_reais[-1] -0.09093    0.04155  -2.188 0.033245 *
X$massa_sal[-1]   0.09168    0.07741   1.184 0.241722
X$credito[-1]    0.24986    0.06680   3.740 0.000467 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8938 on 51 degrees of freedom
Multiple R-squared:  0.4964,    Adjusted R-squared:  0.4569
F-statistic: 12.57 on 4 and 51 DF,  p-value: 3.463e-07
```

Séries temporais

- ▶ Quantas defasagens das variáveis dependentes usar?
- ▶ Devo usar a defasagem da variável explicativa?
- ▶ Usar o nível ou a variação?
- ▶ Resíduo passado ajuda a explicar as variações futuras?

- ▶ **Modelos teóricos e modelos estatísticos!**

Exercícios

- ▶ Estimar uma forma funcional da balança comercial de 1 dos 4 países:
 - ▶ Brasil;
 - ▶ Colômbia;
 - ▶ Chile;
- ▶ Variáveis:
 - ▶ Variação da balança comercial em US\$ milhões: $(t)-(t-4)$;
 - ▶ Câmbio efetivo real: variação percentual de $t/t-4$
 - ▶ Preço das commodities: variação percentual de $t/t-4$
 - ▶ PIB agregado dos países emergentes e desenvolvidos;
 - ▶ PIB: variação percentual de $t/t-4$

Exercícios

- ▶ Usar uma função com 1 variável independente e outra com pelo menos 2 variáveis (pode ser a variável dependente defasada);
- ▶ Fazer o gráfico do valor realizado e dos preditos pelos dois modelos;
- ▶ Fazer o teste de autocorrelação.
- ▶ Enviar para:

aperfcarreiras@enap.gov.br

Exercícios - EUA

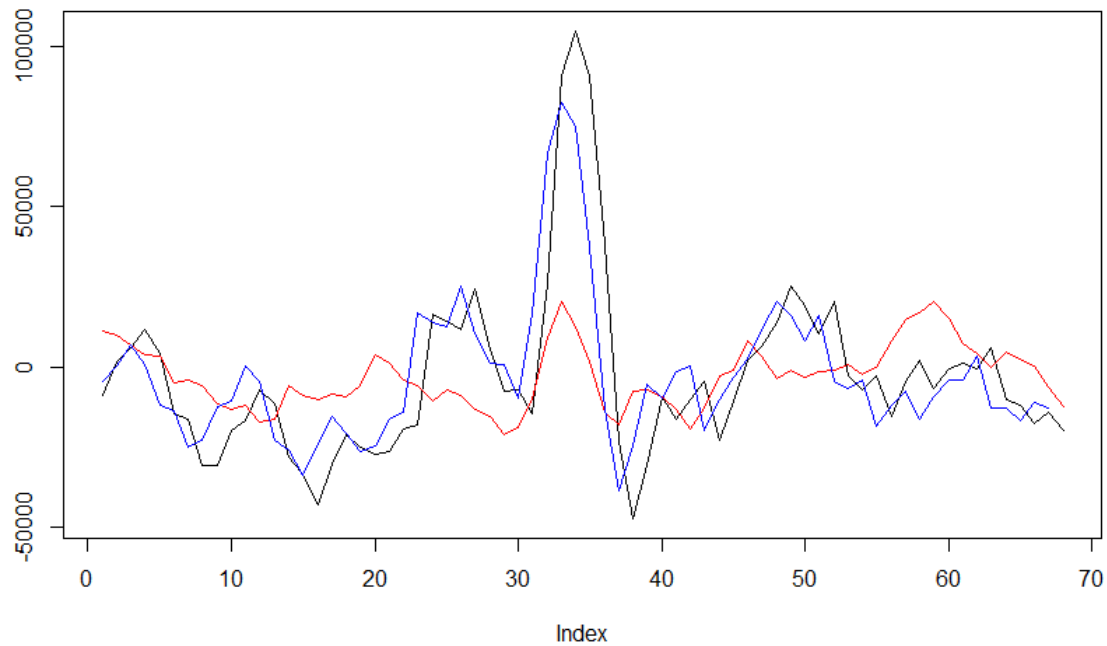
- ▶ `Z<-read.csv("aula1_eua.csv",sep=";", dec=".", head=TRUE) # baixando os dados.`
 - ▶ `pib_eua<-ts(Z$pib_eua,start=c(2001,1),frequency=4)`
 - ▶ `saldo<-ts(Z$eua_balanca,start=c(2001,1),frequency=4)`
 - ▶ `cambio<-ts(Z$eua_cambefet,start=c(2001,1),frequency=4)`
 - ▶ `commdty<-ts(Z$comdty,start=c(2001,1),frequency=4)`
 - ▶ `pib_dm<-ts(Z$pib_dm,start=c(2001,1),frequency=4)`
- ▶ `eq1<-lm(saldo~cambio)`
 - ▶ `summary(eq1)`
- ▶ `eq2<-lm(saldo[-length(saldo)]~saldo[-1]+cambio[-1])`
 - ▶ `summary(eq2)`

Exercícios - EUA

- ▶ `eq1<-lm(saldo~cambio)`
 - ▶ `summary(eq1)`
 - ▶ `eqfit<-predict.lm(eq1,interval="confidence")`
 - ▶ `eqfit1<-data.frame(eqfit)`
 - ▶ `erro1<-resid(eq1)`
- ▶ `eq2<-lm(saldo[-length(saldo)]~saldo[-1]+saldo[-2]+cambio[-1])`
 - ▶ `summary(eq2)`
 - ▶ `eqfit<-predict.lm(eq2,interval="confidence")`
 - ▶ `eqfit2<-data.frame(eqfit)`
 - ▶ `erro2<-resid(eq2)`

Exercícios - EUA

- ▶ `plot(Z$eua_balanca,type = "l")`
- ▶ `points(eqfit1$fit,type = "l",col= "red")`
- ▶ `points(eqfit2$fit,type = "l",col= "blue")`



Exercícios - EUA

```
> bgtest(eq1)
```

```
      Breusch-Godfrey test for serial correlation of order up to 1
```

```
data: eq1
```

```
LM test = 41.713, df = 1, p-value = 1.057e-10
```

```
> bgtest(eq2)
```

```
      Breusch-Godfrey test for serial correlation of order up to 1
```

```
data: eq2
```

```
LM test = 12.338, df = 1, p-value = 0.0004439
```

- ▶ A primeira defasagem não resolveu!

INTRODUÇÃO - SÉRIES DE TEMPO

The background features abstract, overlapping geometric shapes in various shades of green, ranging from light lime to dark forest green. The shapes are primarily triangles and polygons, creating a dynamic, layered effect. The overall composition is clean and modern, with the text centered on a white background.

Algumas séries

- ▶ A série é estacionária ou há tendência?
- ▶ A variável é estocástica ou determinística?
 - ▶ Raramente as séries são puramente determinística, mas normalmente há parte estocástica e determinística

Série estacionária

▶ Estocástica

▶ $x_t = \mu + \beta x_{t-1} + u_t$, onde $\beta < 1$

▶ `w = rnorm(550,0,1) #criando 550 observações normais N(0,1);`

▶ `plot(w,type="l")`

▶ $x_t = 5 + 0.3 x_{t-1} + u_t$

▶ `x1 = filter(w, filter=c(0.3), method="recursive")[-(1:50)]+5`

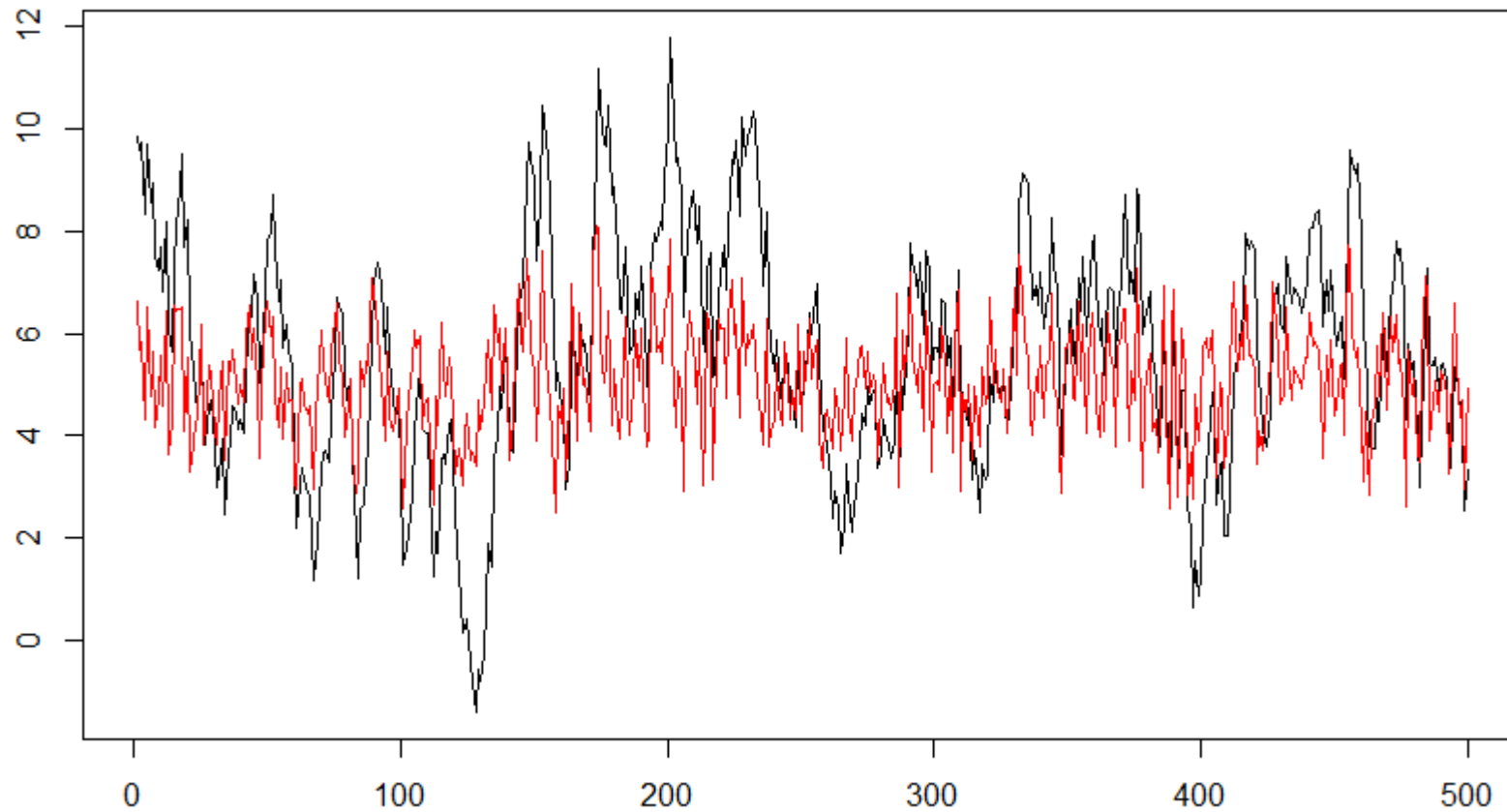
▶ $x_t = 5 + 0.9 x_{t-1} + u_t$

▶ `x2 = filter(w, filter=c(0.9), method="recursive")[-(1:50)]+5`

▶ `plot(x2,type="l")`

▶ `points(x1,type="l",col="red")`

Série estacionária



- ▶ Maior persistência da série preta - volatilidade

Passeio Aleatório

- ▶ $x_t = x_{t-1} + u_t$

- ▶ Recursivamente:

- ▶ $x_t = x_{t-2} + u_t + u_{t-1}$

- ▶ $x_t = x_0 + \sum_{i=1}^t u_j$

- ▶ Longa memória - impacto dos resíduos passados tem forte relação com o valor de hoje

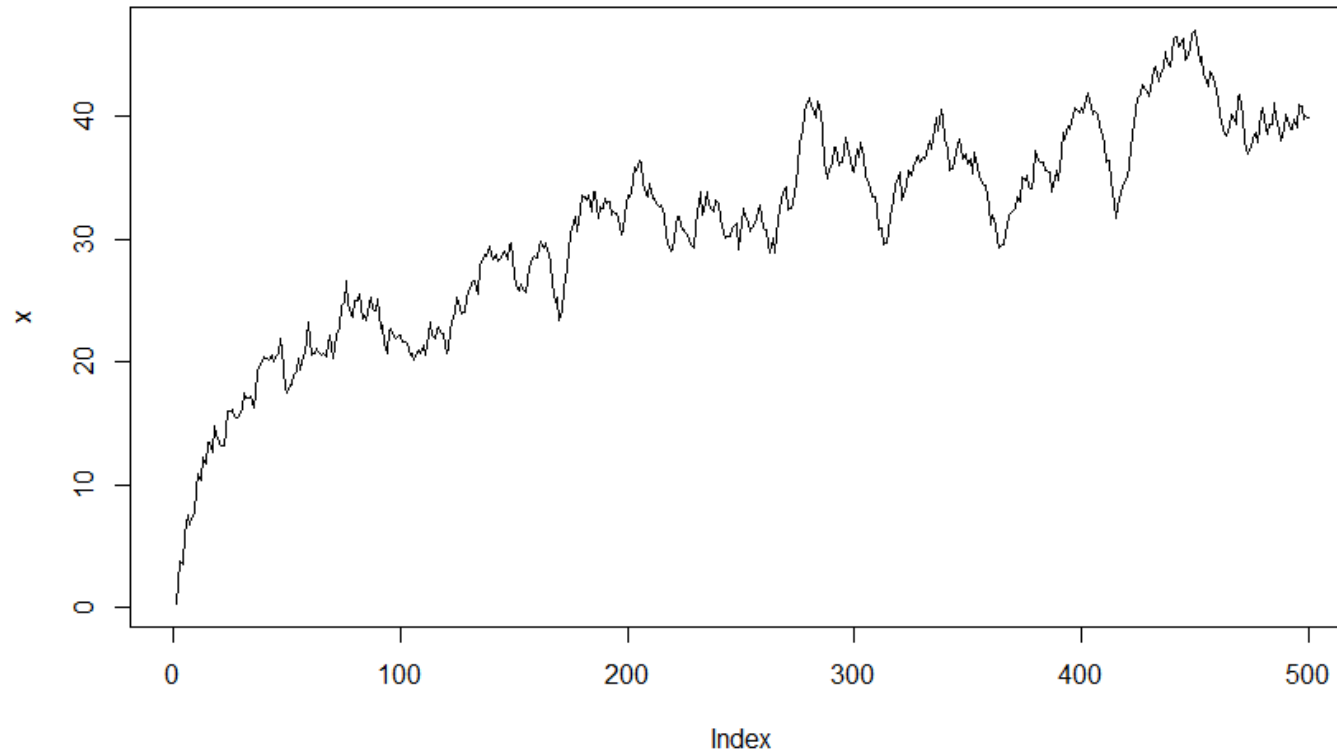
- ▶ Criando um passeio aleatório

- ▶ `w=rnorm(550,0,1)`

- ▶ `x = filter(w, filter=c(1), method="recursive")[-(1:50)]`

Passeio Aleatório

► `plot(x,type="l")`

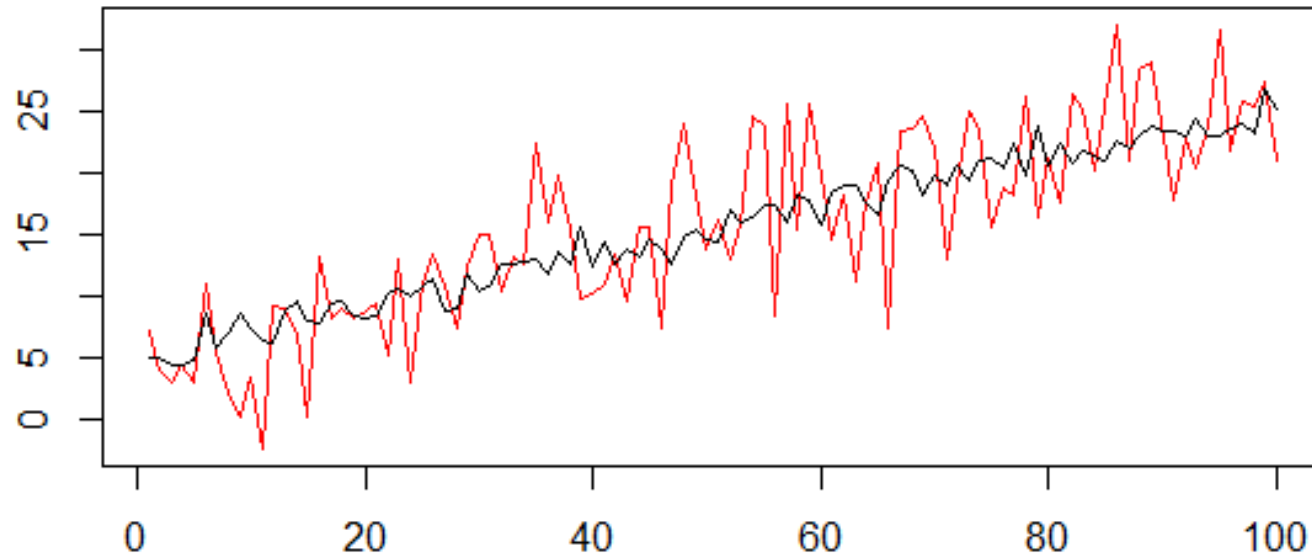


Série com tendência determinística

- ▶ $x_t = \mu + \delta t + u_t$
- ▶ $x_t = 5 + 0.2t + u_t$
- ▶ Variância menor de $u \sim N(0,1)$ #fazer um script
 - ▶ `u1= rnorm(100,0,1)`
 - ▶ `t<-1:100`
 - ▶ `y1<-5+0.2*t+u1`
- ▶ Variância maior de $u \sim N(0,1)$
 - ▶ `u2= rnorm(100,0,4)`
 - ▶ `y2<-5+0.2*t+u2`

Série com tendência determinística

- ▶ Apresentar o gráfico
 - ▶ `plot(y2,type="l",col="red")`
 - ▶ `points(y1,type="l")`



Série com tendência estocástica

▶ $x_t = \mu + x_{t-1} + u_t$

▶ Recursivamente:

▶ $x_t = 2\mu + x_{t-2} + u_t + u_{t-1}$

▶ $x_t = t\mu + x_0 + \sum_{i=1}^t u_j$

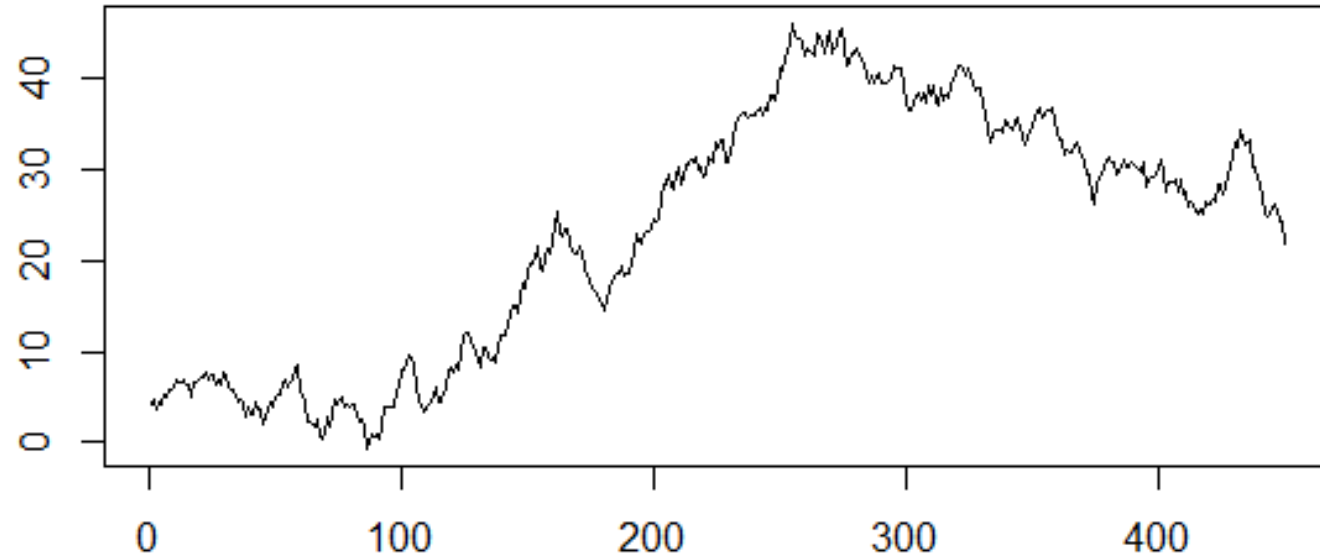
▶ $x_t = 5 + x_{t-1} + u_t$

▶ `u = rnorm(500,0,1)`

▶ `x = filter(u, filter=c(1), method="recursive")[-(1:50)]+5`

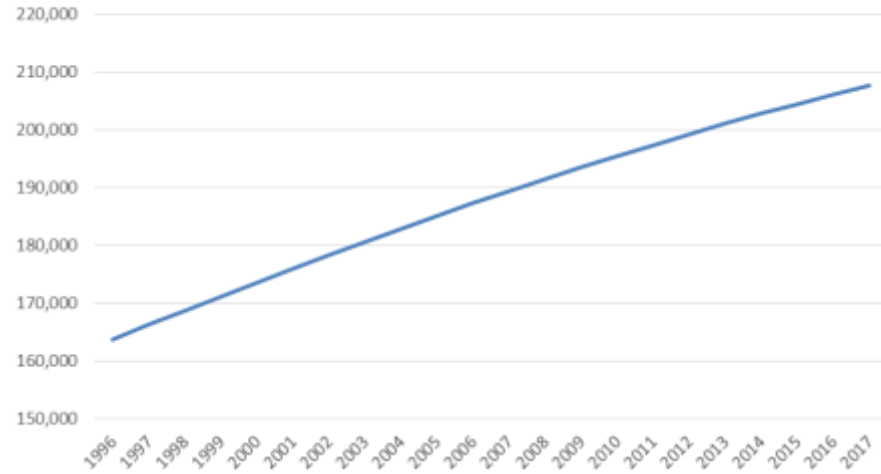
Série com tendência estocástica

► `plot(x,type="l")`



COMO CLASSIFICARIA AS SÉRIES ABAIXO?

População Brasileira (milhares)



Juro Real (% ex-post)



Preço das Commodities (variação t/t-1)



PIB (t/t-1)



Alguns conceitos básicos

1. Autocovariância;
2. Estacionariedade;
3. Ruído branco;
4. Ergodicidade;
5. Processo de médias móveis;
6. Processo autoregressivo;
7. Invertibilidade.

Autocovariância e autocorrelação

- ▶ Autocovariância

- ▶ $\gamma_k = E(Y_t - \mu_t)(Y_{t-k} - \mu_{t-k}) = \text{cov}(Y_t, Y_{t-k})$

- ▶ Autocorrelação

- ▶ $\rho(k) = \frac{\gamma(k)}{\gamma(0)}$

- ▶ Matriz de covariância (dividido pelo desvio padrão - autocorrelação):

$$\Sigma_{XX} = \sigma^2 \begin{bmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(p-1) \\ \rho(1) & 1 & \rho(1) & \dots & \rho(p-2) \\ \rho(2) & \rho(1) & 1 & \dots & \\ \vdots & & & \dots & \rho(1) \\ \rho(p-1) & & & \rho(1) & 1 \end{bmatrix}$$

- ▶ Pressupõe estacionariedade

Estacionaridade

- ▶ Propriedades estatísticas são invariantes no tempo
- ▶ Média finita e constante
 - ▶ $E(Y_t) = \mu$ para todo t
- ▶ Variância finita
 - ▶ $E(Y_t)^2 < \infty$
- ▶ Autocovariância
 - ▶ $E(Y_t - \mu)(Y_{t-j} - \mu) = \gamma_j$ para todo t e j
- ▶ Ambos independem de t e são finitos - fracamente estacionário;
- ▶ Descrição de fracamente estacionário x estritamente estacionário
 - ▶ Estrictamente estacionário: as funções de distribuição da série com mesmo tamanho de intervalo temporal terão propriedades estatísticas similares.
 - ▶ Para todo j , A função $(Y_{t_1}, Y_{t_2}, \dots, Y_{t_n})$ é semelhante a função de $(Y_{t_1-j}, Y_{t_2-j}, \dots, Y_{t_n-j})$

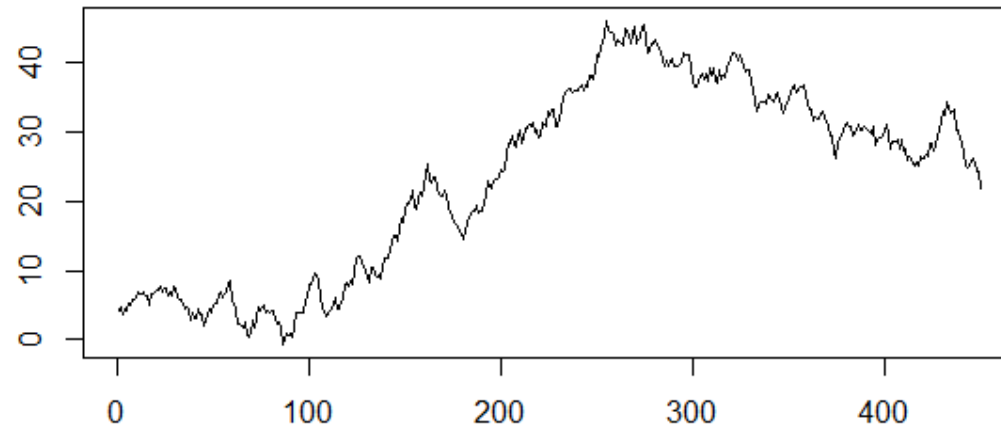
Por que não é estacionário?

▶ Ruído Branco

$$x_t = t\mu + x_0 + \sum_{i=1}^t u_j$$

▶ Média e variância

▶ Solução?

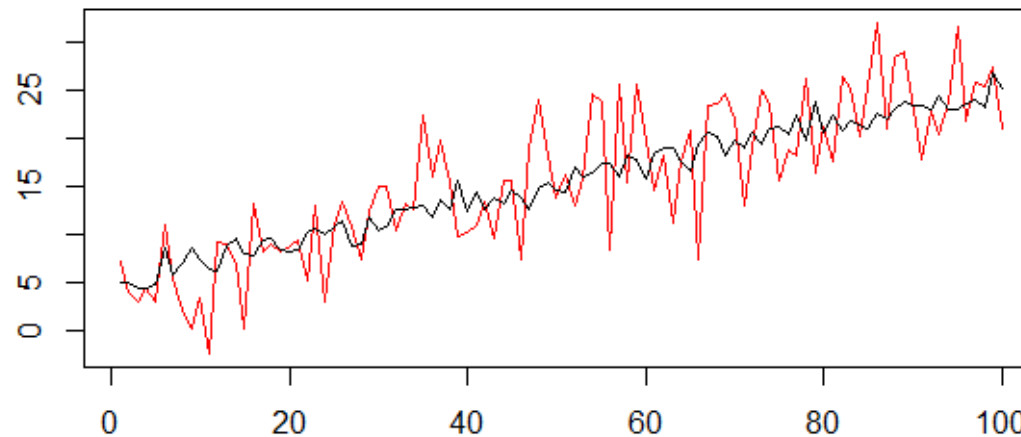


▶ Tendência

$$x_t = \mu + \delta t + u_t$$

▶ Média depende do tempo

▶ Solução?



Ruído branco

- ▶ Média zero, variância constante e finita, não há autocorrelação
- ▶ $E(\varepsilon_t) = 0$
- ▶ $E(\varepsilon_t^2) = \sigma_t^2$
- ▶ $E(\varepsilon_t \varepsilon_k) = 0$ para todo t diferente de k

Ergodicidade

- ▶ O processo aleatório manterá as propriedades estatísticas da população;
- ▶ Ou seja, como as séries temporais econômicas são únicas, espera-se que a média temporal das observações seja uma boa aproximação da média populacional.
 - ▶ $E(\bar{y}^a) = \text{p} \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{i=0}^T y_t^a = E(y_t)$, onde “a” representa a amostra
- ▶ Em termos práticos, um processo é ergódico para a média desde que a autocovariância convirja suficientemente rápido para zero.
- ▶ Ruído branco é ergódico para todos os momentos.
- ▶ Ergodicidade não implica necessariamente em estacionariedade, nem o contrário.

Estimar um modelo ARMA

- ▶ Para estimar uma série temporal, é necessário que o processo seja:
 - ▶ Estocástico;
 - ▶ Estacionário;
 - ▶ Ergótico.