INTRODUÇÃO À ECONOMETRIA

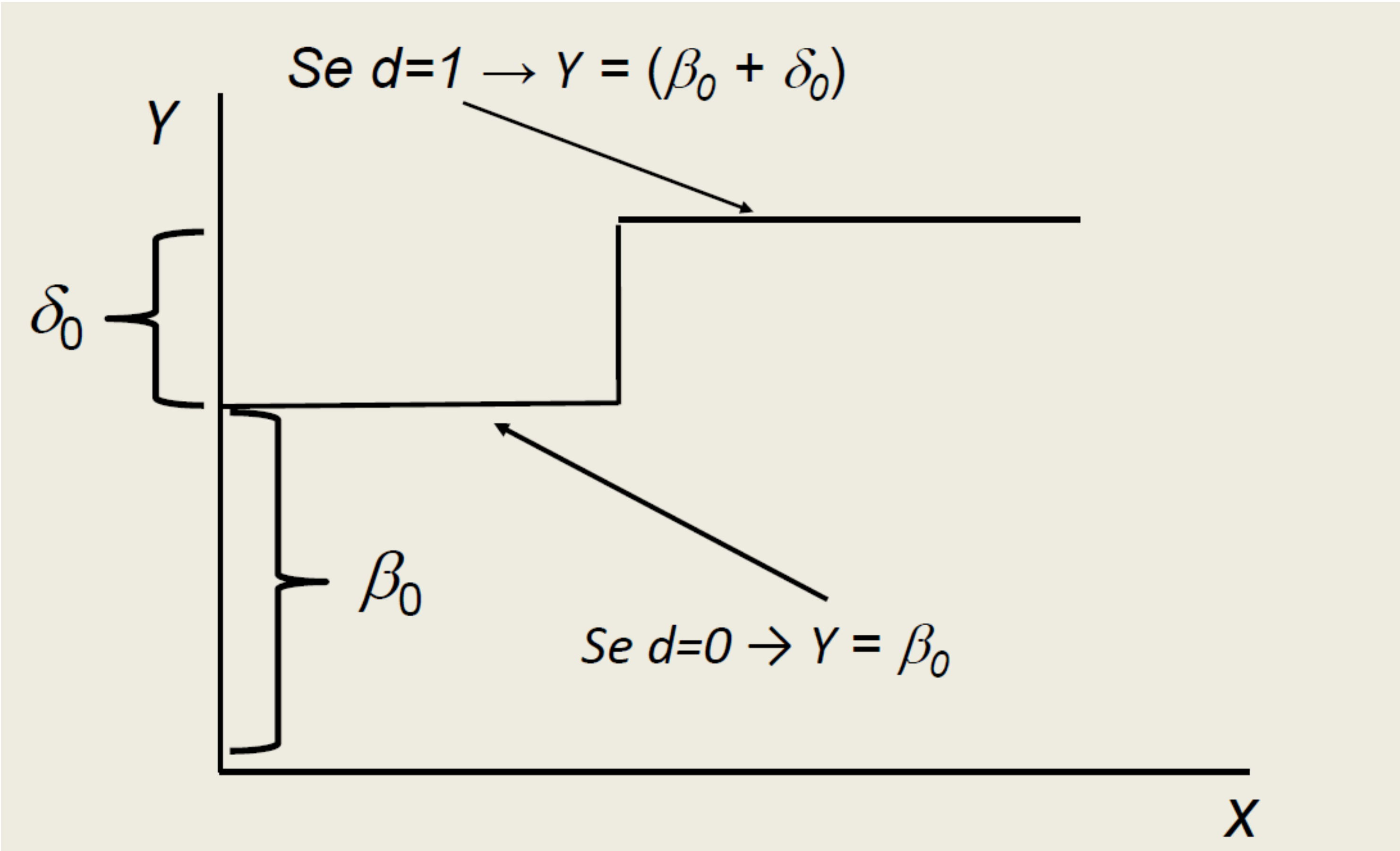
Análise de Regressão Múltipla

Aula 9

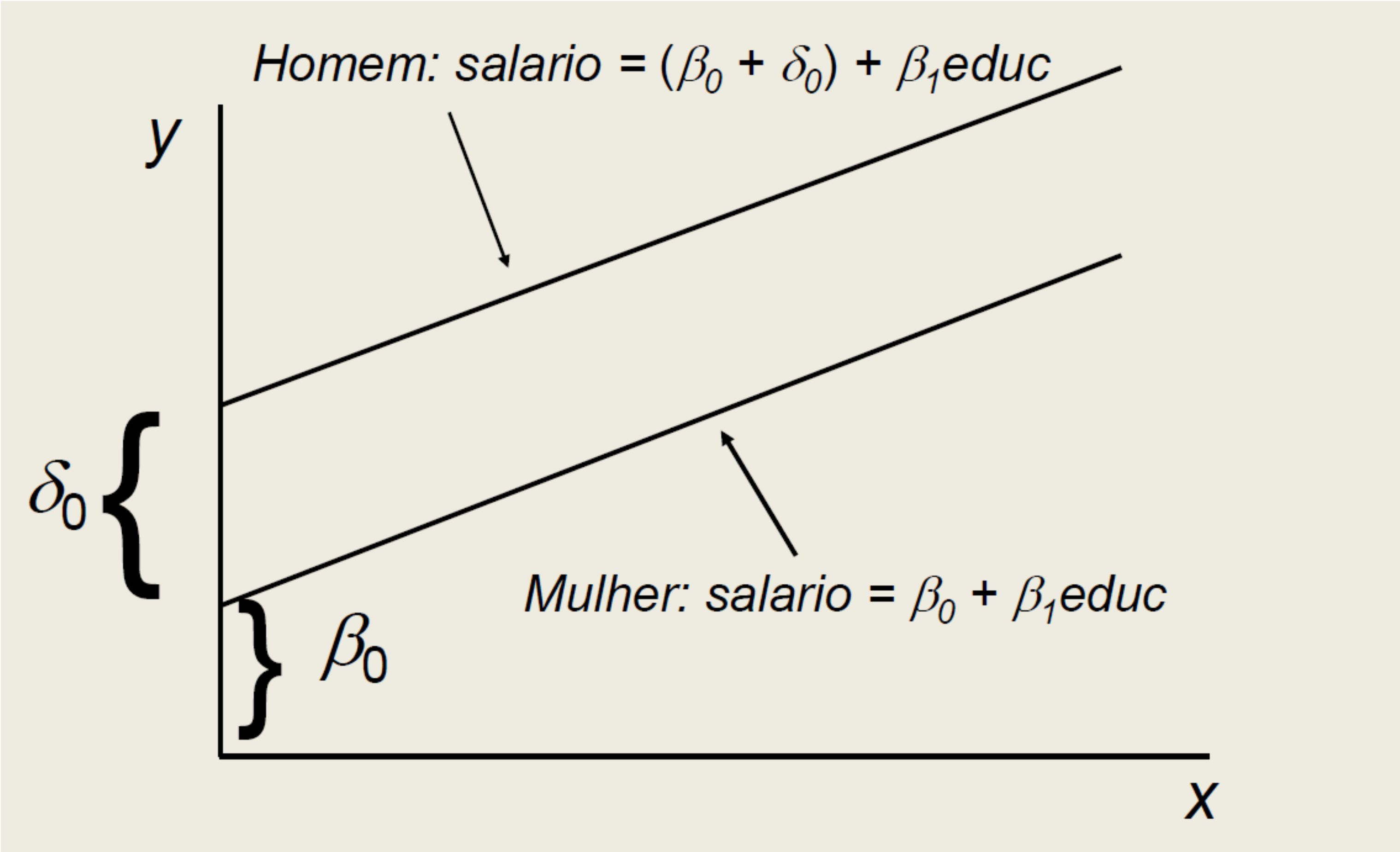
Escola Nacional de Administração Pública

Variável Binária – Dependente

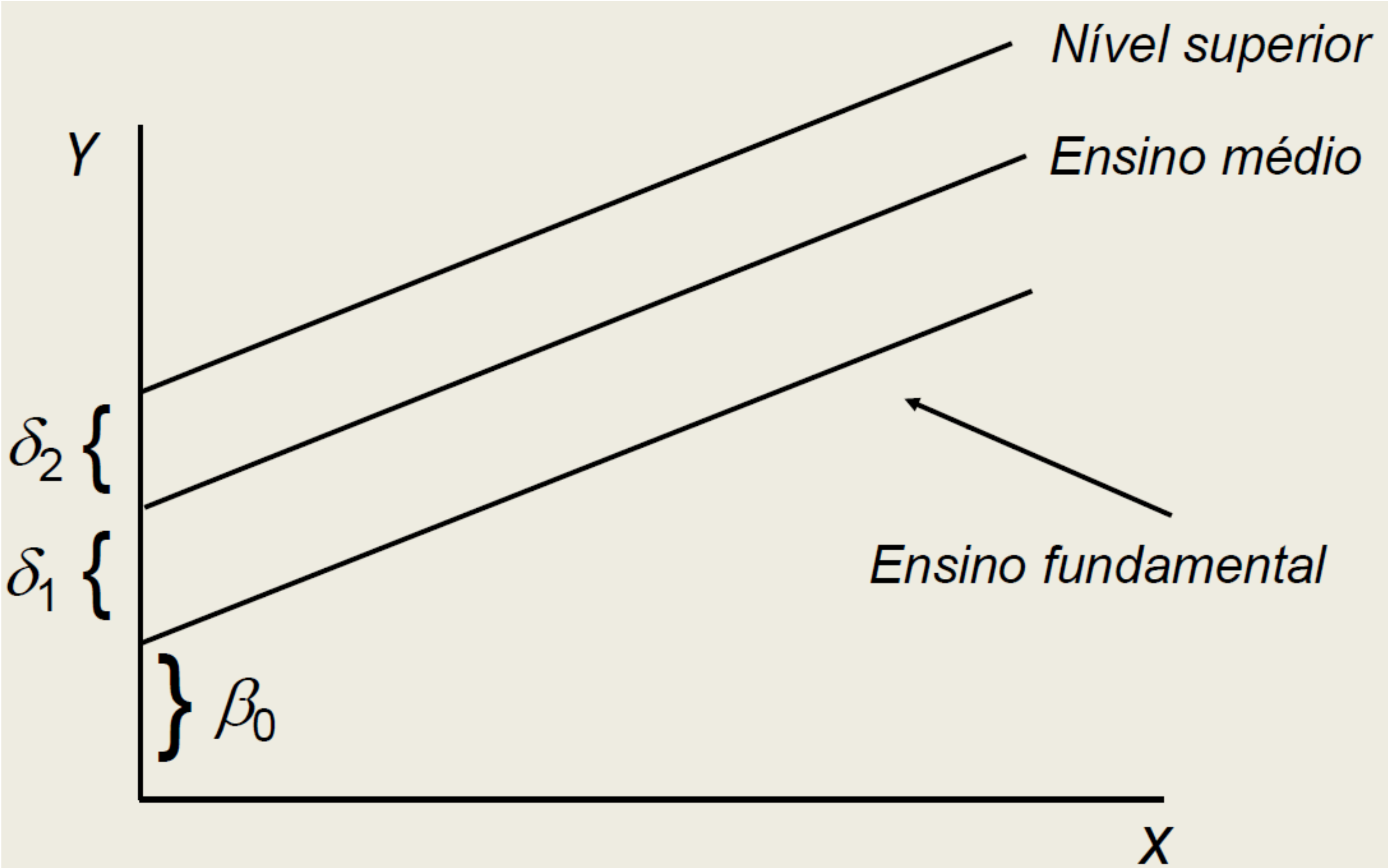
Variável Binária – Revisão
Regressão Simples



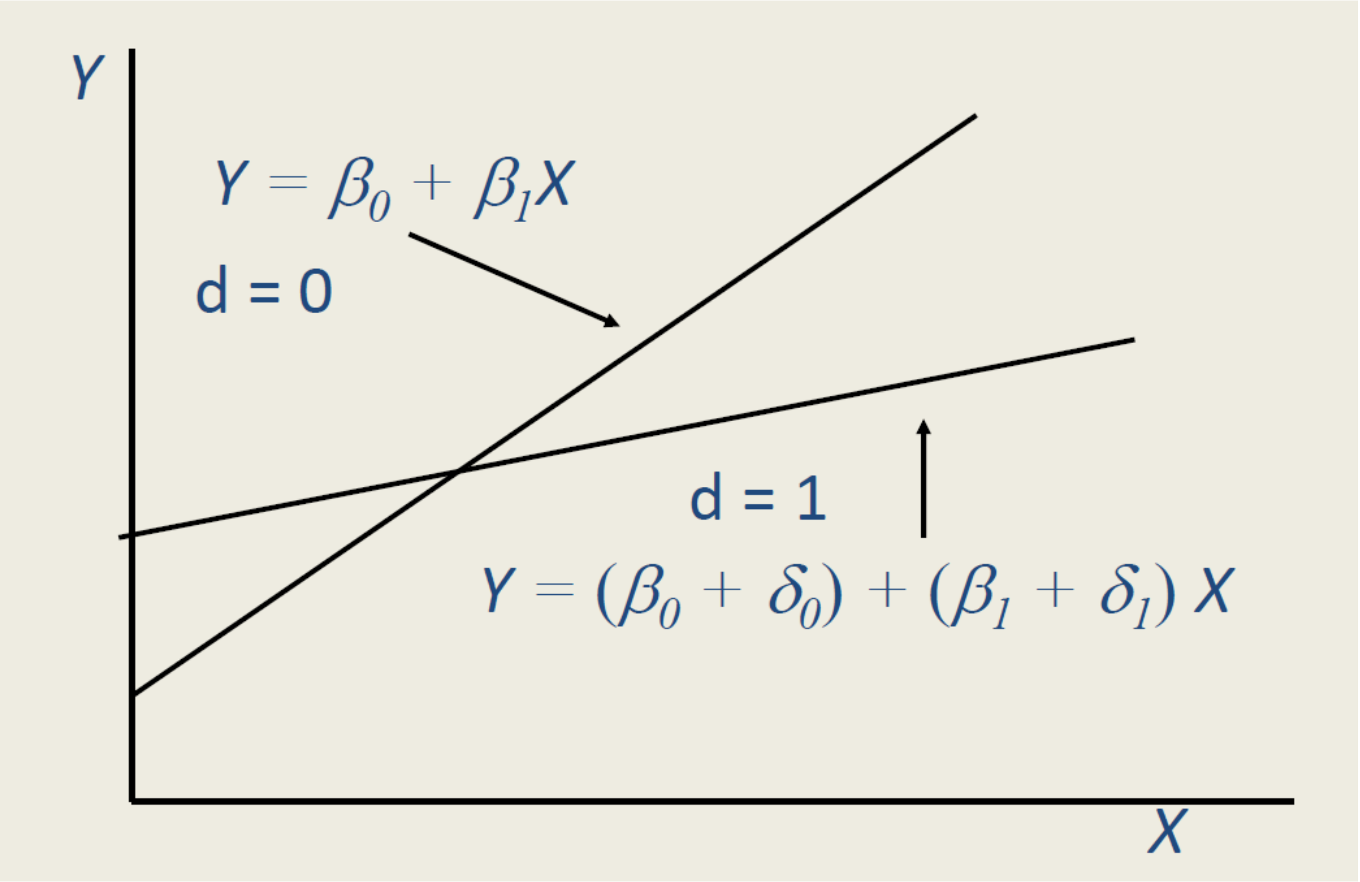
Variável Binária – Revisão
Regressão Múltipla



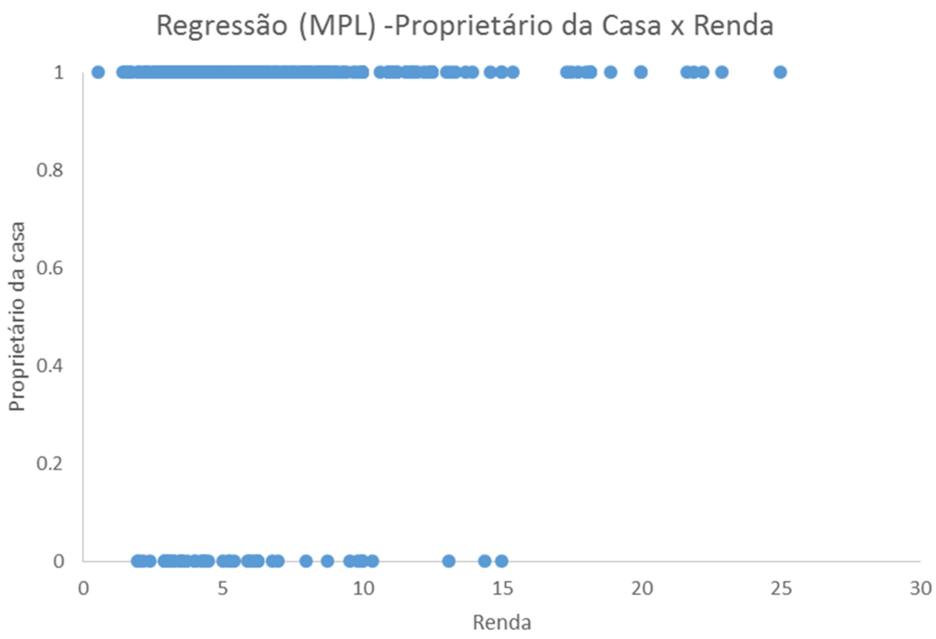
Variável Binária – Revisão
Várias dummies



Variável Binária – Revisão
Efeitos da Interação



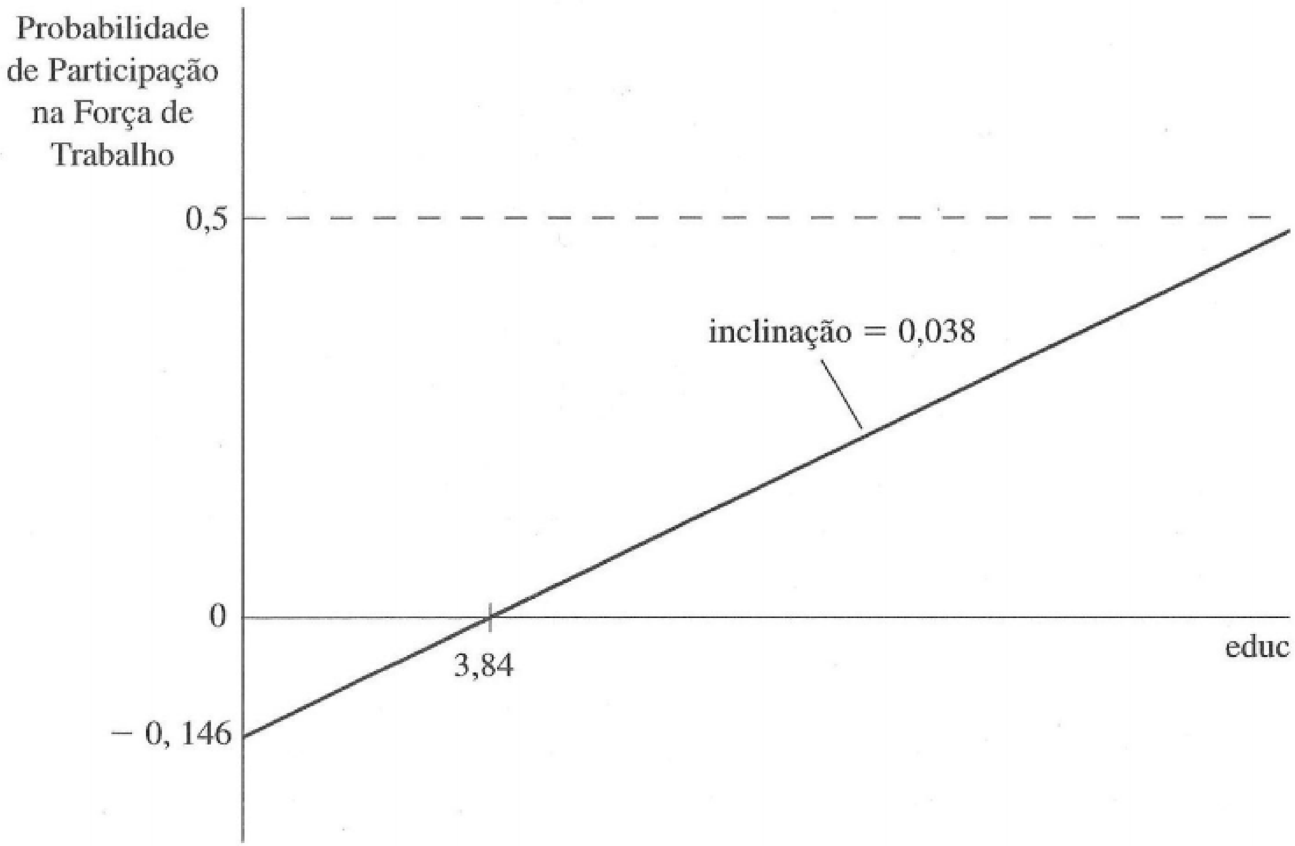
Variável Binária
Modelo de Probabilidade Linear



$$Proprietário\ da\ casa_i = \beta_0 + \beta_1 Renda\ familiar_i$$

$$pr(Proprietário\ da\ casa_i = 1\ |Renda\ familiar_i)$$

Variável Binária
Modelo de Probabilidade Linear

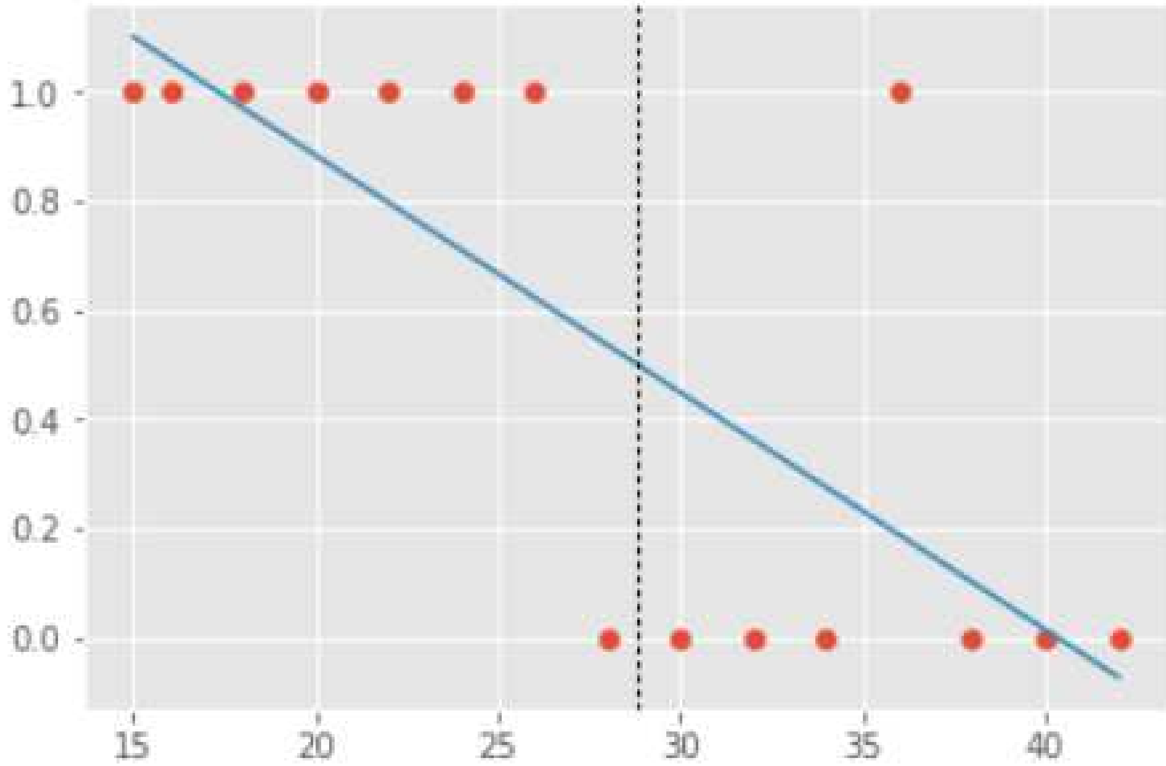


Variável Binária
Modelo de Probabilidade Linear

Problemas com o MPL

- Resíduos não tem distribuição Normal (binomial) – há problema da inferência;
- Pode assumir valores negativos. Isso faz sentido?
    - Não satisfação de $0 \leq pr(Y_i \,|X_i) \leq 1$
- Heterocedasticidade;

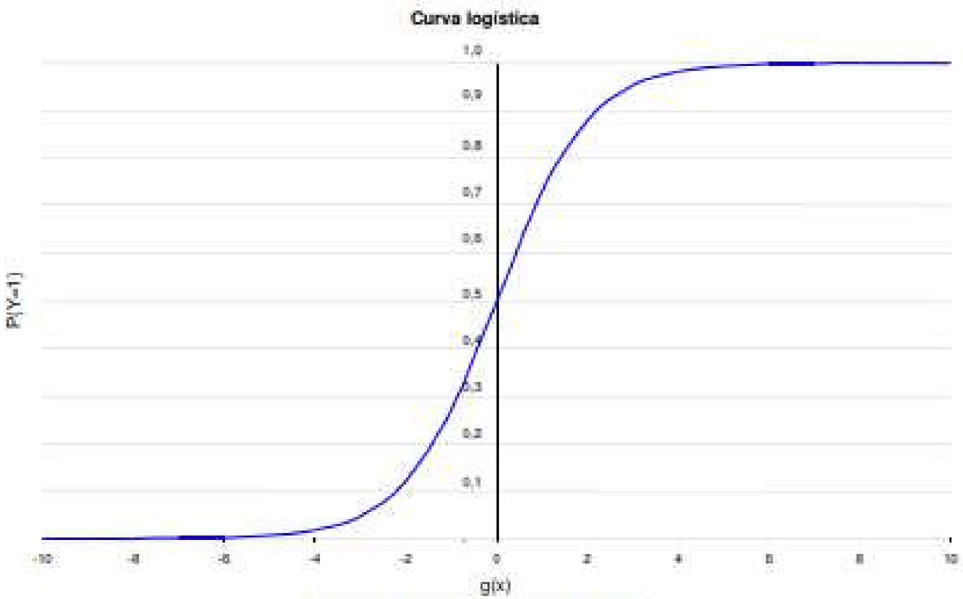
Variável Binária
MPL (problemas)



Variável Binária
Probit

- Emerge de uma função normal acumulada;
- Embora linear nos parâmetros, a variável independente não é linear;
- A variável dependente está limitada [0,1];

Variável Binária
Probit



$g(x) = B_0 + B_1X_1 + \cdots + B_pX_p$

$$P(Y=1) = \frac{1}{1+e^{-g(x)}}$$

$g(x) \rightarrow +\infty$, então $P(Y=1) \rightarrow 1$

$g(x) \rightarrow -\infty$, então $P(Y=1) \rightarrow 0$

Variável Binária
Exemplo de MPL e Probit

Equação MPL

- eua<-read.csv("EUA.csv",sep=";", dec=".", head=TRUE)
- summary(eua)
- reg_recss=lm(eua$recessao~eua$juro_real+eua$Prod_industrial)
- summary(reg_recss)

O que significam os coeficientes?

Variável Binária
Exemplo de MPL e Probit

Equação PROBIT

- install.packages("aod")
- require(aod)
- probit <- glm(eua$recessao ~ eua$Prod_industrial + eua$juro_real, family = binomial(link = "probit"),data=eua)
- summary(probit)
- confint(probit)

INTRODUÇÃO – SÉRIES DE TEMPO

Algumas séries

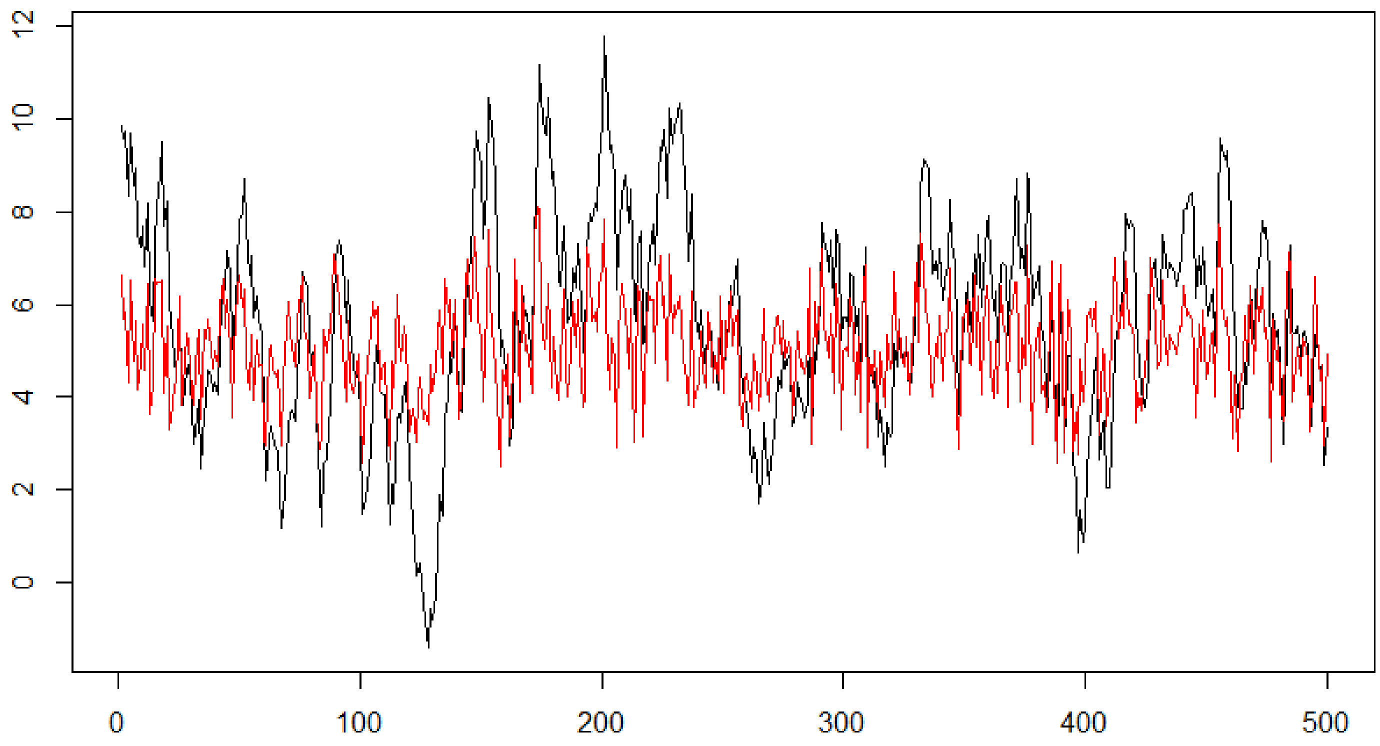
- A série é estacionária ou há tendência?

- 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



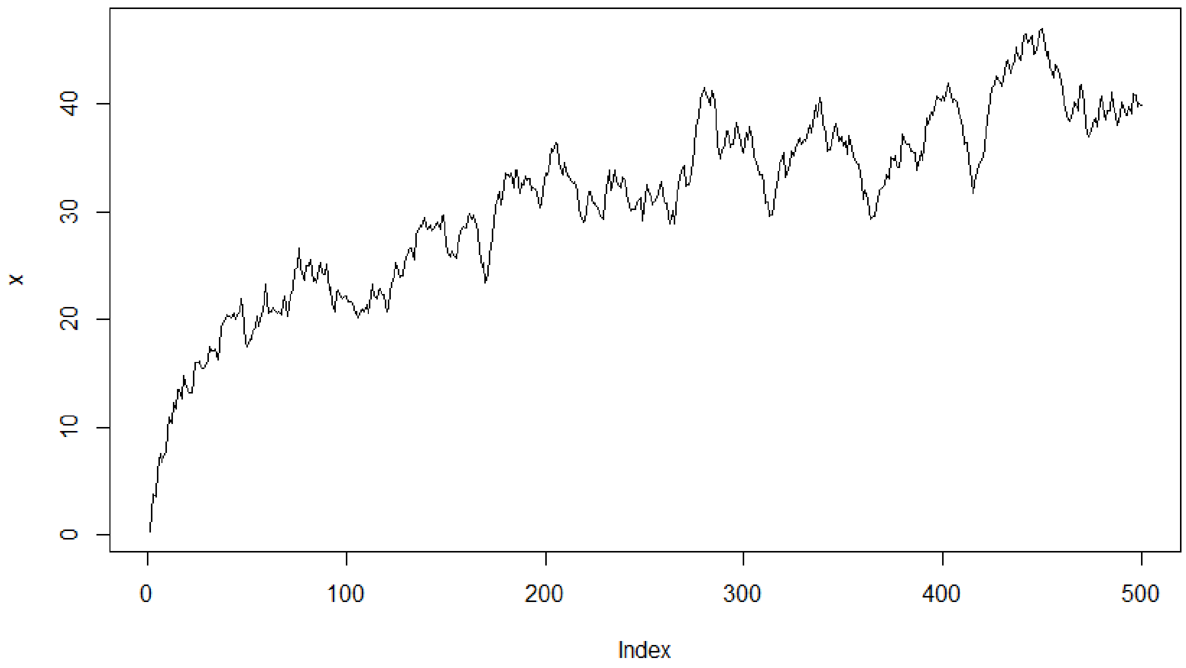
- Maior persistência da série preta - volatilidade

Passeio Aleatório

- $\boldsymbol{x_t = x_{t-1} + u_t}$
- Recursivamente:
    - $\boldsymbol{x_t = x_{t-2} + u_t + u_{t-1}}$
    - $\boldsymbol{x_t = x_0 + \sum_{i=1}^{t} u_j}$
- Longa memória – impacto dos resíduos passados tem forte relação com o valor de hoje
- Criando um passeio aleatório
    - x = filter(w, filter=c(1), method="recursive")[-(1:50)]

Passeio Aleatório

- plot(x,type="l")

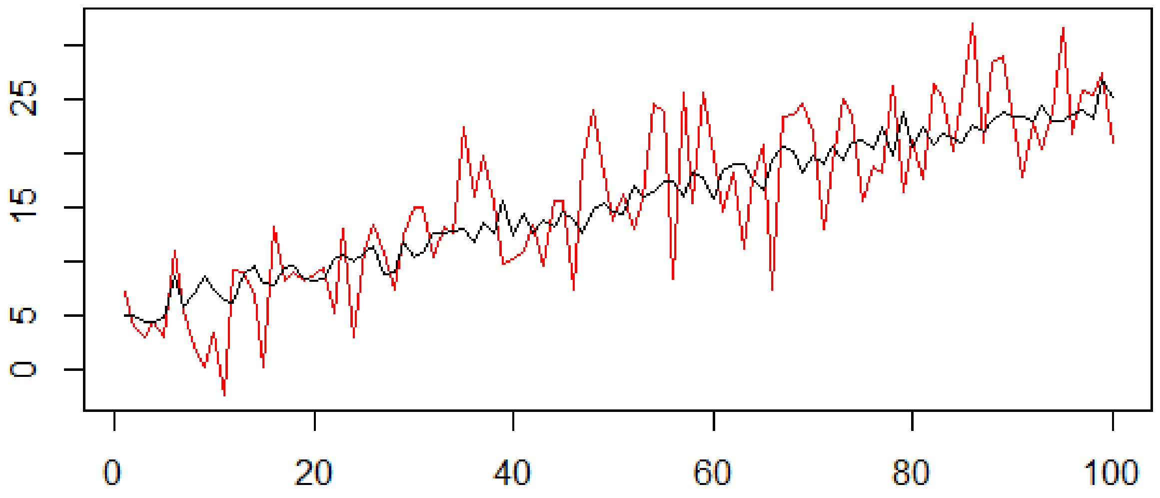


Série com tendência determinística

- $\boldsymbol{x_t = \mu + \delta t + u_t}$
- $\boldsymbol{x_t = 5 + 0.2t + u_t}$
- Variância menor de u~N(0,1) #fazer um script
  - u1= rnorm(100,0,1)
  - t<-1:100
  - y1<-5+0.2*t+u1
- Variância maior de u~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")

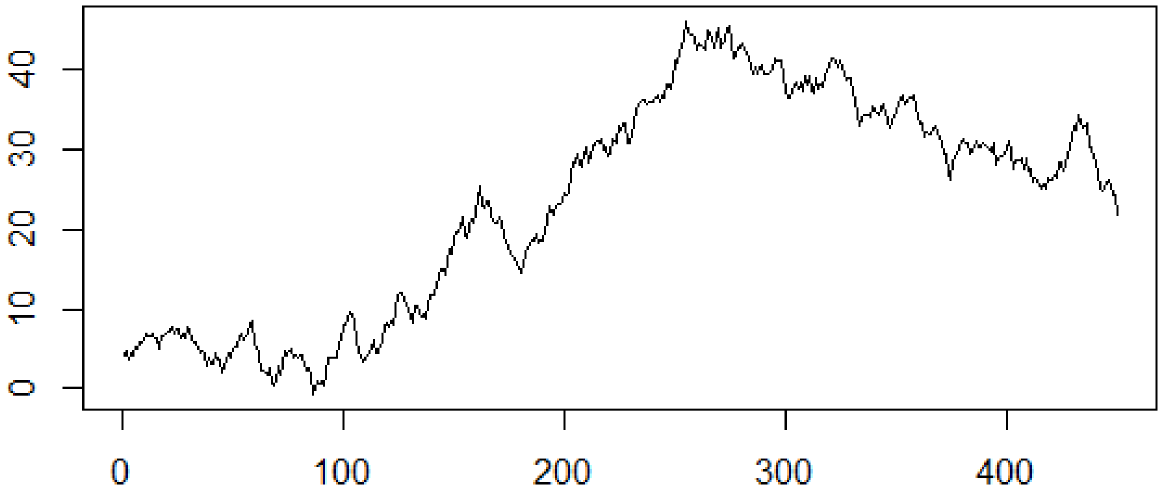


Série com tendência estocástica

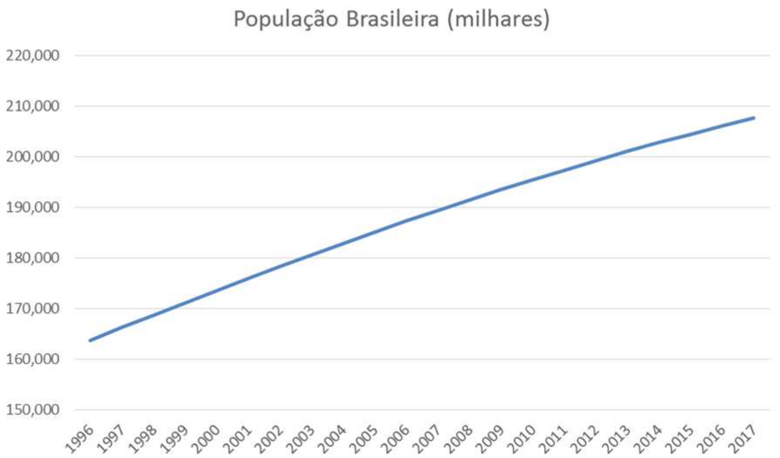
- $x_t = \mu + x_{t-1} + u_t$
- **Recursivamente:**
  - $x_t = 2\mu + x_{t-2} + u_t + u_{t-1}$
  - $x_t = t\mu + x_0 + \sum_{i=1}^{t} u_j$
- $x_t = 5 + x_{t-1} + u_t$
  - u= rnorm(500,0,1)
  - x = filter(u, filter=c(1), method="recursive")[-(1:50)]+5

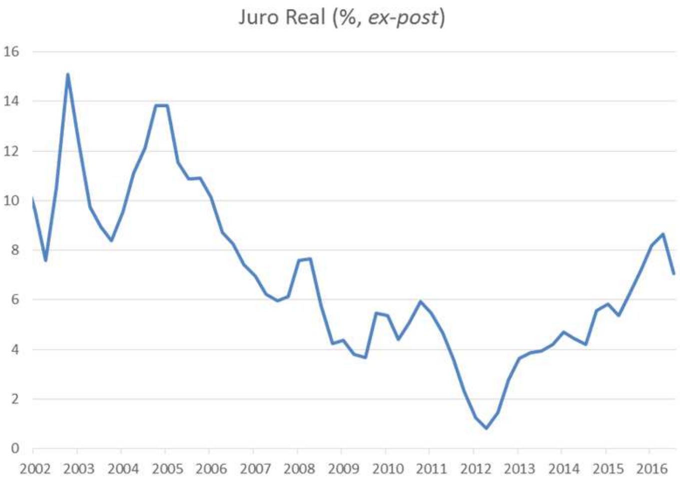
Série com tendência estocástica

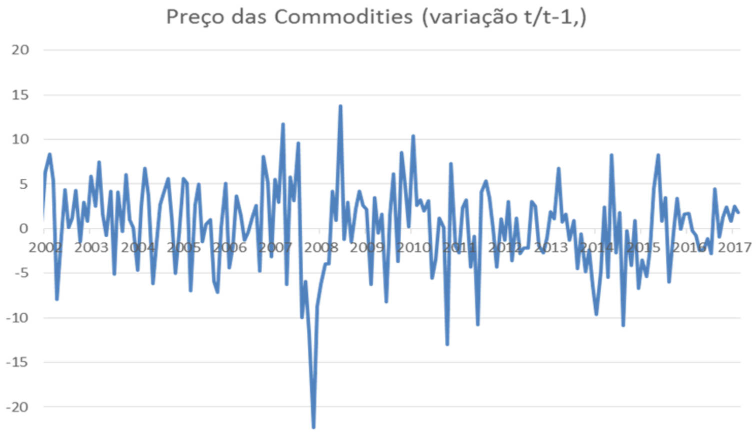
- plot(x,type="l")

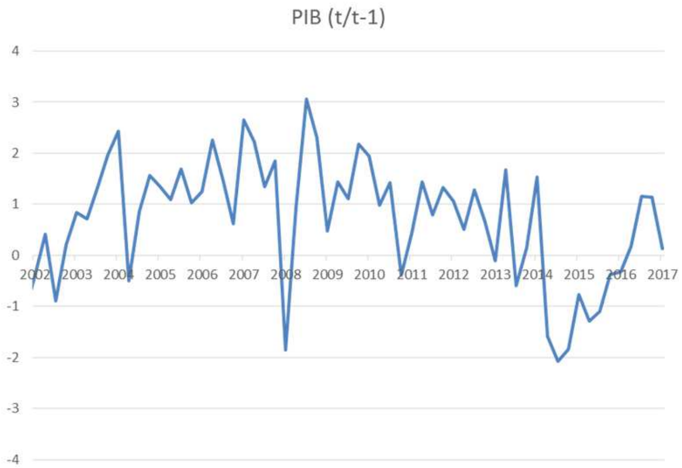


COMO CLASSIFICARIA AS SÉRIES ABAIXO?









Alguns conceitos básicos

1. Autocovariância;

2. Estacionariedade;

3. Ruído branco;

4. Processo de médias móveis (MA);

5. Processo autoregressivo (AR);

Autocovariância e autocorrelação

- Autocovariância
  - $\gamma_k = E(Y_t - \mu_t)(Y_{t-k} - \mu_{t-k}) = \mathrm{cov}(Y_t, Y_{t-k})$
- Autcorrelação
  - $$\rho(k) = \frac{\gamma(k)}{\gamma(0)}$$
- Matriz de convariância (dividido pelo desvio padrão – autocorrelação):

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

  - Pressupõe estacionariedade

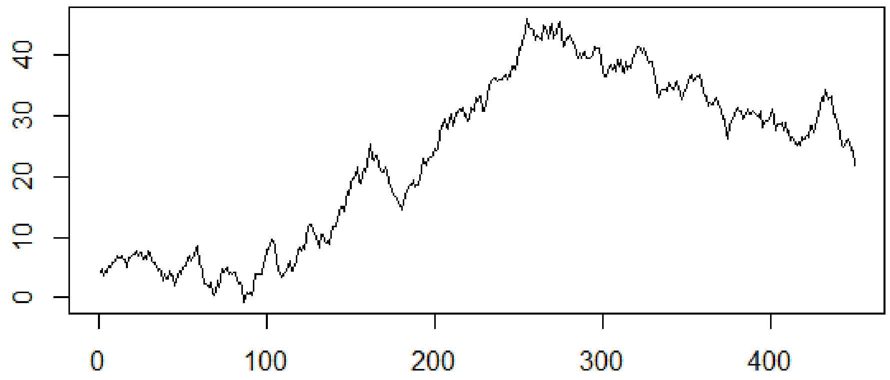
Estacionaridade

- 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
  - $E(Y_t - \mu)(Y_{t-j} - \mu)$= γ_j para todo t e j
- Ambos independem de t e são finitos – fracamente estacionário;
- Descrição de fracamente estacionário x estritamente estacionário
  - Estritamente estacionário: as funções de distribuição da série com mesmo tamanho de intervalo temporal terão propriedades estatísticas similares.
  - Para todo j, A funçã $(Y_{t_1}, Y_{t_2}, \cdots, Y_{t_n})$ é semelhante a função $(Y_{t_{1-j}}, Y_{t_{2-j}}, \cdots, Y_{t_{n-j}})$

Por que não é estacionário?

- Ruído Branco

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

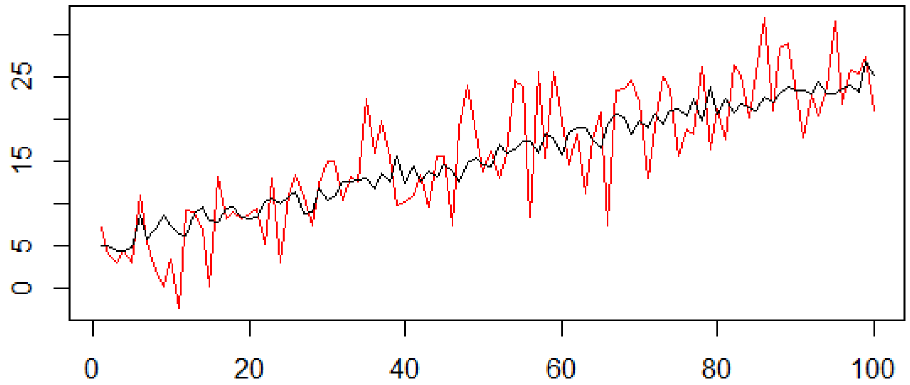


  - Média e variância
  - Solução?

- Tendência

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

  - Média depende do tem
  - Solução?



Ruído branco

- Média zero, variância constante e finta, 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

Estimar um modelo ARMA

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

Médias Móveis – MA(1)

- Considere o seguinte processo estocástico:

$$y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}$$

onde ε_t é um ruído branco ~N(0, σ^2);

- Média móvel mais geral do que normalmente escutamos:
    - Exemplo: média móvel de 3 meses da produção industrial.
    - Pesos não são necessariamente idênticos.
- Processo é estacionário?
    - Média:
    - $E(y_t) = \mu + E(\varepsilon_t) + \theta E(\varepsilon_{t-1}) = \mu$

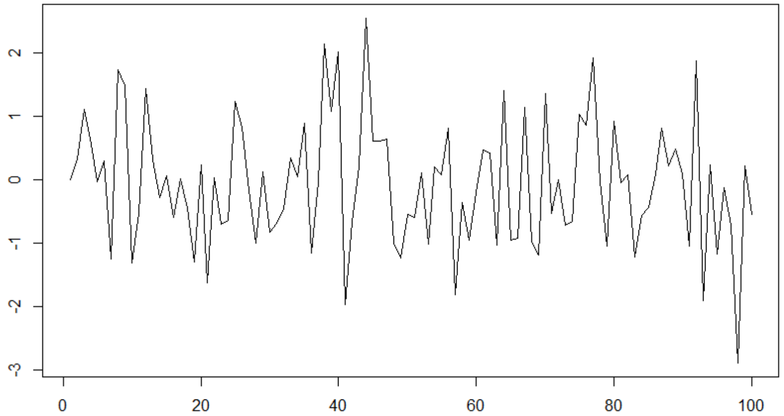
Médias Móveis – MA(1)

- Criar um script no R
- Digitar as seguintes linhas:
  - y<- e <- rnorm(100,0,1)
  - for (t in 2:100) y[t] <- e[t]+0.3*e[t-1]
  - plot(y,type= "l")
- Colocar a primeira linha como comentário “#” para usar diferentes coeficientes, mas com a mesma amostra aleatória.
- Salvar os valores: y_ma1_0_3=y

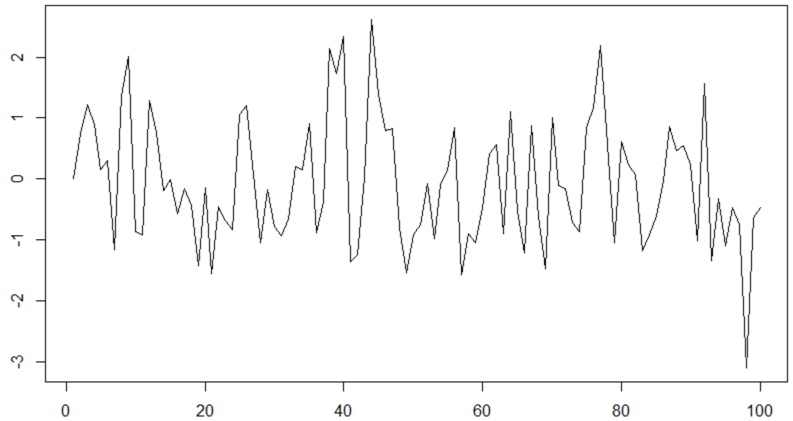
  y_ma1_0_9=y

Médias Móveis – MA(1)

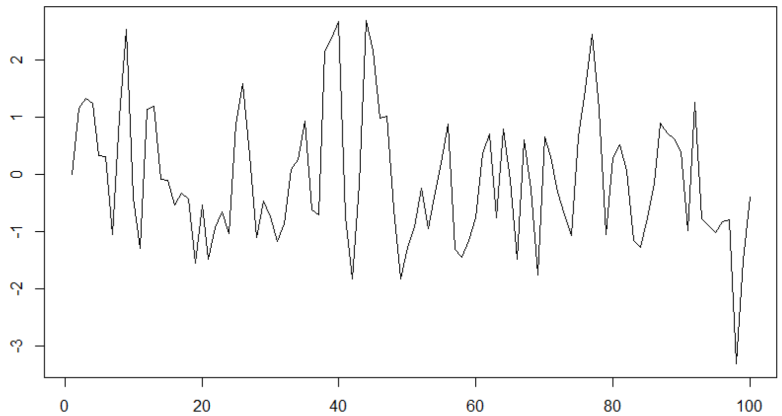
$$y_t = \mu + \varepsilon_t$$



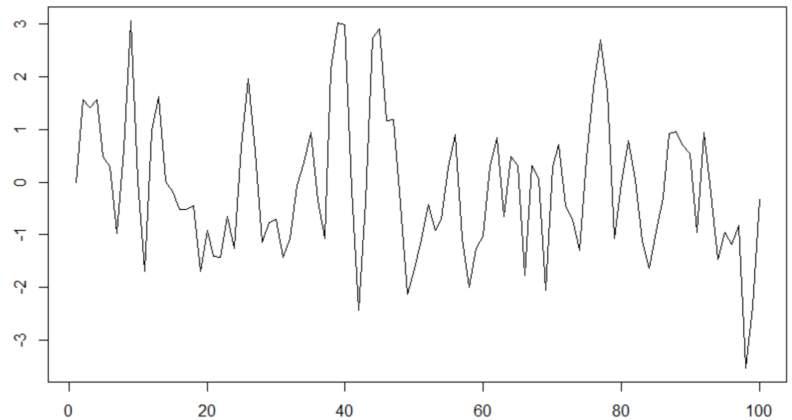
$$y_t = \mu + \varepsilon_t + 0.3\varepsilon_{t-1}$$



$$y_t = \mu + \varepsilon_t + 0.6\varepsilon_{t-1}$$



$$y_t = \mu + \varepsilon_t + 0.9\varepsilon_{t-1}$$



Médias Móveis – MA(1)

- Séries ficaram mais persistentes com a elevação do coeficiente do MA(1)

- Variância também aumenta:
  - A escala dos gráficos aumenta com o coeficiente do MA(1)
  - = $\sigma^2(1+\theta^2)$

- Autocorrelação aumenta com o valor maior de θ

Exemplos de MA(2)

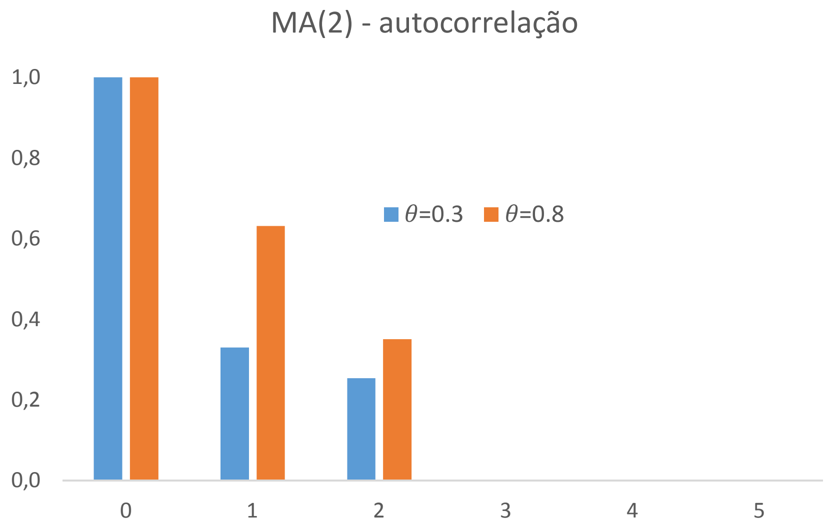
- MA(2)

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2)\sigma^2$$

$$\gamma_1 = (\theta_1 + \theta_1\theta_2)\sigma^2$$

$$\gamma_2 = (\theta_2)\sigma^2$$

$$\gamma_j = 0, para\ j > 2$$



- Autocorrelação é dividir por γ_0

- **Importante!!!**

$$\rho_j = \begin{cases} 1, & j = 0 \\ \frac{\theta_1+\theta_1\theta_2}{1+\theta_1^2+\theta_2^2}, & j = 1 \\ \frac{\theta_2}{1+\theta_1^2+\theta_2^2}, & j = 2 \\ 0, & j > 2. \end{cases}$$

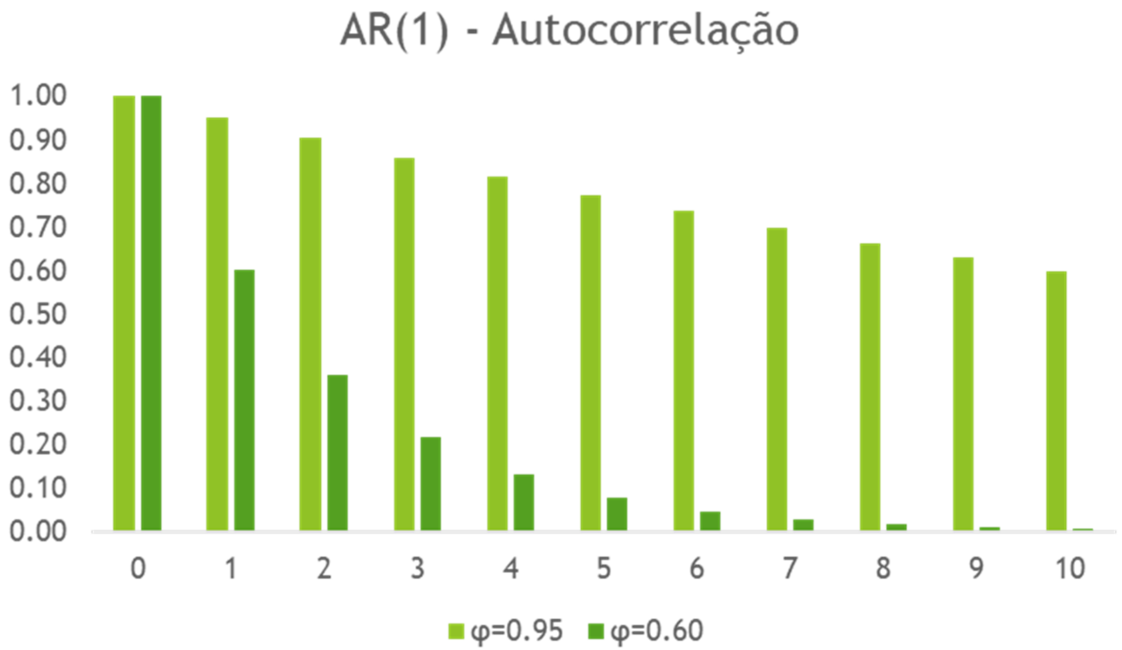
é determinada pelo

Autorregressivo – AR(1)

- Considere o seguinte processo estocástico:

$y_t = c + \varphi y_{t-1} + \varepsilon_t$ onde ε_t é um ruído branco ~N(0, σ^2) e | φ|<1;

Autocorrelação - AR(1)



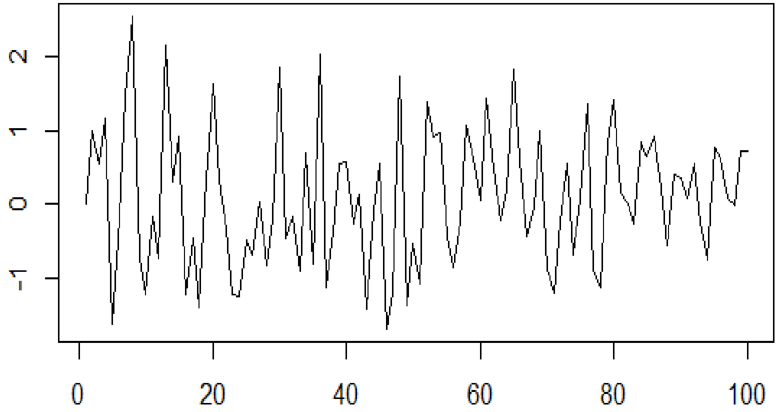
Autorregressivo AR(1)

- Criar um script no R
- Digitar as seguintes linhas:
  - y<- e <- rnorm(100,0,1)
  - y[0]=0
  - y[1]=0
  - for (t in 2:100) y[t] <- e[t]+0.1*y[t-1]
  - plot(y,type= "l")
- Colocar a primeira linha como comentário “#” para usar diferentes coeficientes, mas com a mesma amostra aleatória.
- Salvar os valores: y_ar1_0_4=y

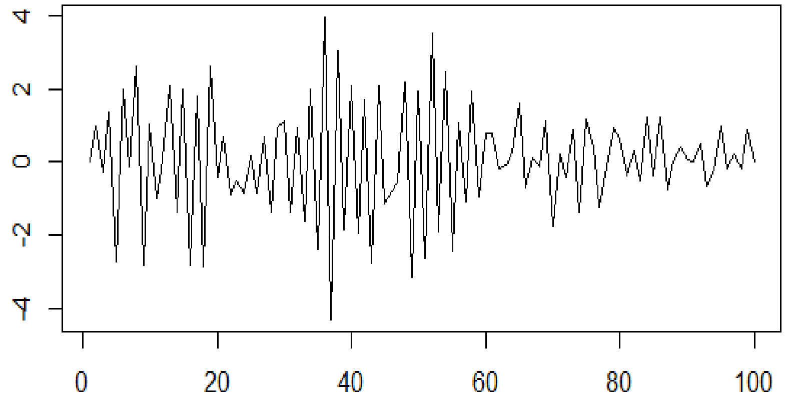
  y_ar1_0_8=y

Autorregressivo AR(1)

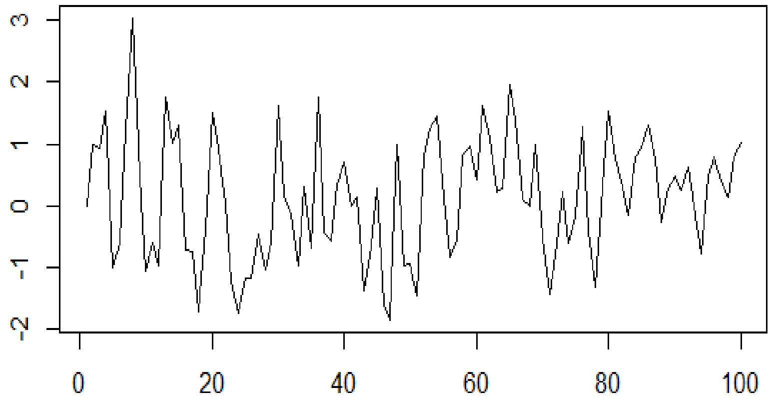
$$y_t = \mu + \varepsilon_t$$



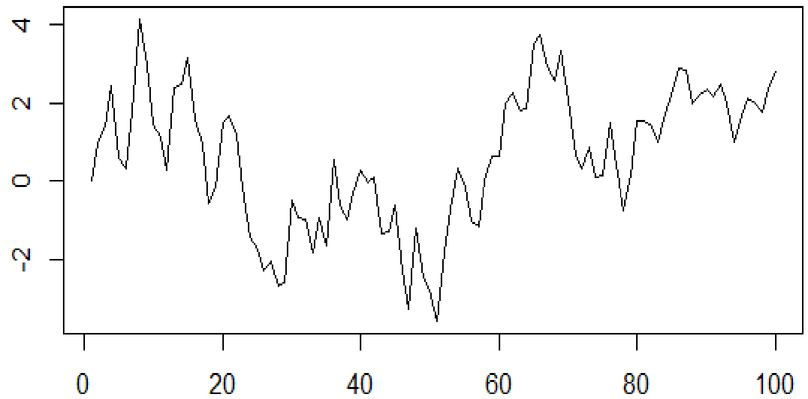
$$y_t = \mu + \varepsilon_t - 0.8y_{t-1}$$



$$y_t = \mu + \varepsilon_t + 0.4y_{t-1}$$



$$y_t = \mu + \varepsilon_t + 0.8y_{t-1}$$



Autorregressivo AR(1)

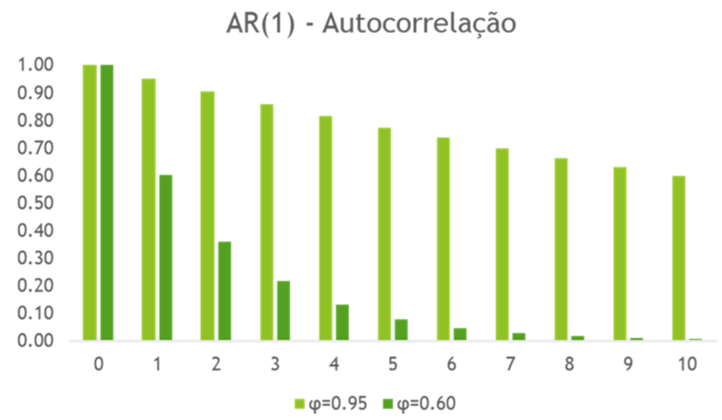
- Séries ficaram mais persistentes com a elevação do coeficiente do AR(1)
  - Coeficiente negativo têm maior influência das mudanças.
  - Dos dois gráfico à direita, tem-se a impressão de maior volatilidade do gráfico superior, no entanto, a volatilidade das séries é a mesma.

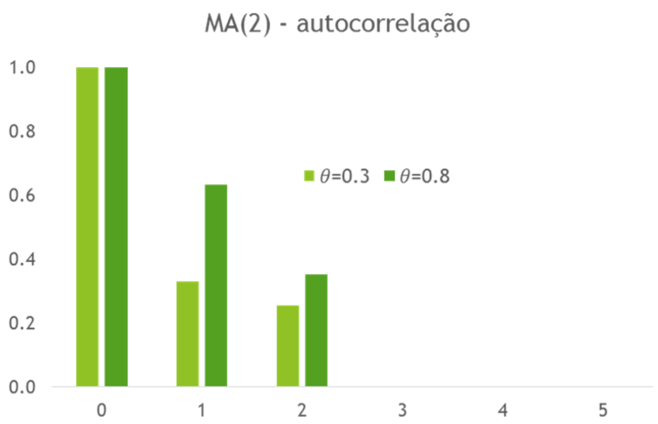
- Variância também aumenta:
  - A escala dos gráficos aumenta com o coeficiente do AR(1) – valor absoluto

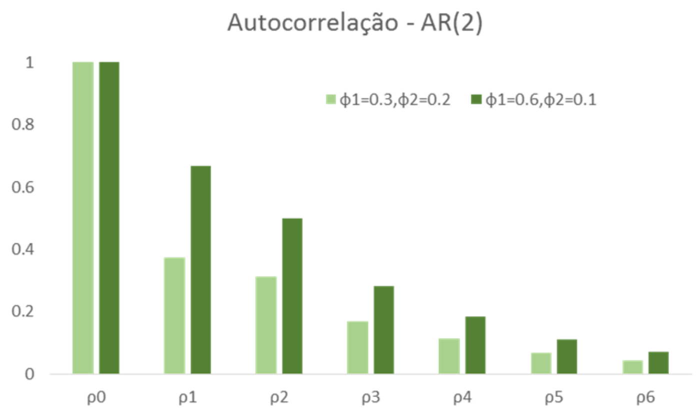
- Autocorrelação aumenta com o valor maior de φ
  - O decaimento é lento quando o valor $|\varphi|$ é próximo de 1
  - Todo AR tem um decaimento infinito, pois é pode ser representado por um MA(∞)

Autocorrelação – Não se esqueça!

- MA – ela interrompe na última defasagem do MA;
- AR – declinante com as defasagens;
- Maior o coeficiente, mais lento o declínio da autocorrelação;
- Variância aumento com o coeficiente.







EXERCÍCIO

Séries de tempo
Exercício

1. Estime a regressão de recessão nos EUA usando a taxa de inflação
2. Crie uma série estacionária e outra com passeio aleatório
3. Crie uma série AR(1) com coeficiente de defasagem de 0,5