

Análise de Séries Temporais

Aula 3

Transformações

Problema 1: em muitas situações de interesse, a distribuição da amostra é assimétrica e pode conter valores atípicos → a suposição de normalidade não está satisfeita.

Problema 2: a existência de tendências em séries econômicas e financeiras.

Procedimento utilizado: fazer uma transformação das observações de modo a obter uma distribuição mais simétrica e próxima da normal ou estabilizar a variância e tornar o efeito sazonal aditivo.

Uma transformação bastante usada é:

$$X^{(p)} = \begin{cases} X^p, & p > 0 \\ \ln(x), & p = 0 \\ -X^p, & p < 0 \end{cases}$$

$$p = 0,5 \rightarrow (x)^{1/2}$$

$$P = -1 \rightarrow -1/x$$

A razão para a mudança de sinal quando $p < 0$ é a de assegurar que os dados transformados tenham a mesma ordem relativa que os originais.

Assimetria à direita $\rightarrow P < 1$

Assimetria à esquerda $\rightarrow P > 1$

Na prática, consideramos valores de p na sequência

-3, -2, -1, -1/2, -1/3, -1/4, 0, 1/4, 1/3, 1/2, 1, 2, 3

- Para cada valor de p , obtemos gráficos apropriados: **histograma**, **box-plot** e **gráficos de simetria** para os dados originais e transformados, de modo que podemos escolher o **valor mais adequado** para p .
- para cada valor de p na sequência se calcule a média, mediana e um estimador de escala(desvio padrão ou algum estimador robusto) e então se escolher o valor que minimiza

$$d_p = (\text{média-mediana})/(\text{medida de escala})$$

que pode ser vista como uma medida de assimetria. Numa distribuição simétrica, $d_p=0$.

Para se ter uma ideia do tipo de transformação que é adequada, pode-se utilizar um gráfico que traz, no eixo das abscissas, médias de subconjuntos de observações da série original e, no eixo das ordenadas, a amplitude de cada um desses subconjuntos; se $Z_1, \dots, Z_k$ for um tal subconjunto com k observações, calculamos

$$\bar{Z} = \frac{1}{k} \sum_{i=1}^k Z_{t_i},$$

$$w = \max(Z_{t_i}) - \min(Z_{t_i}),$$

que são medidas de posição e variabilidade, respectivamente; o par $(\bar{Z}, w)$ será um ponto do gráfico. O número de elementos em cada subsérie pode ser igual ao período, no caso de séries sazonais.

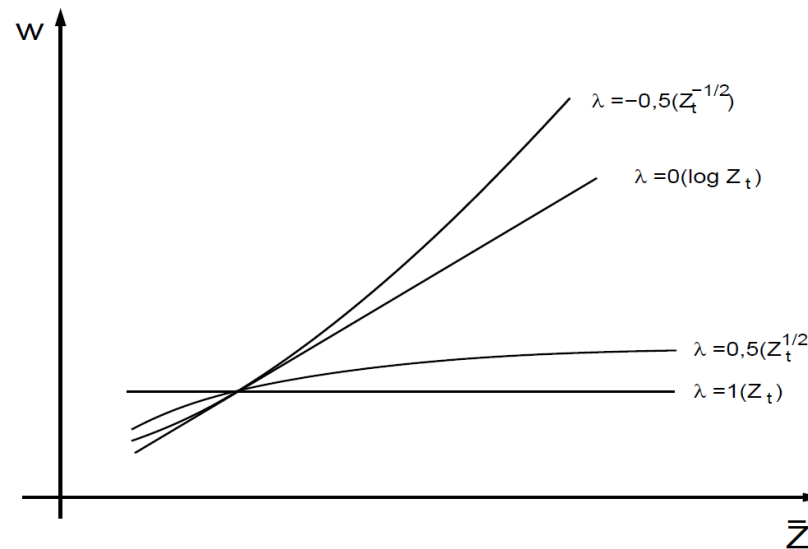


Figura 1.6: Gráficos amplitude $\times$ média, ilustrando alguns valores possíveis de λ .

Se w independer de $\bar{Z}$, obteremos pontos espalhados ao redor de uma reta paralela ao eixo das abscissas, e neste caso não haverá necessidade de transformação. Se w for diretamente proporcional a $\bar{Z}$, a transformação logarítmica é apropriada.

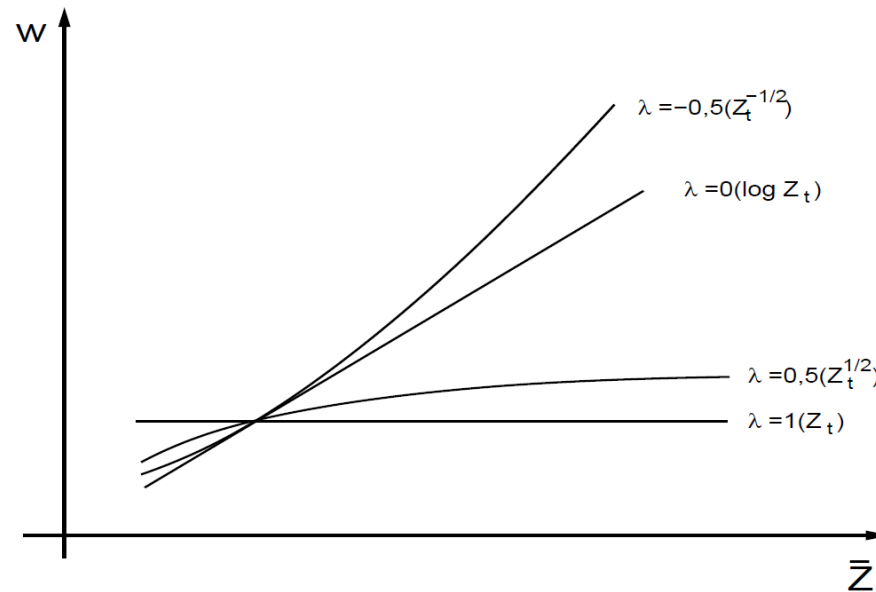


Figura 1.6: Gráficos amplitude $\times$ média, ilustrando alguns valores possíveis de λ .

Problema 3: estabilizar a **variância** às vezes é mais importante do que tornar a distribuição aproximadamente normal.

Suponha X uma v.a. com

$$E(x) = \mu, \quad \text{Var}(x) = h^2(u) \sigma^2 \text{ (não é constante)}$$

Idéia: $X \rightarrow g(X)$ tal que $\text{Var}[g(X)] = \text{cte.}$

Considere uma expansão de Taylor de $g(X)$ até primeira ordem, ao redor de $g(\mu)$:

$$g(X) \approx g(\mu) + (X-\mu)g'(\mu)$$

Então:

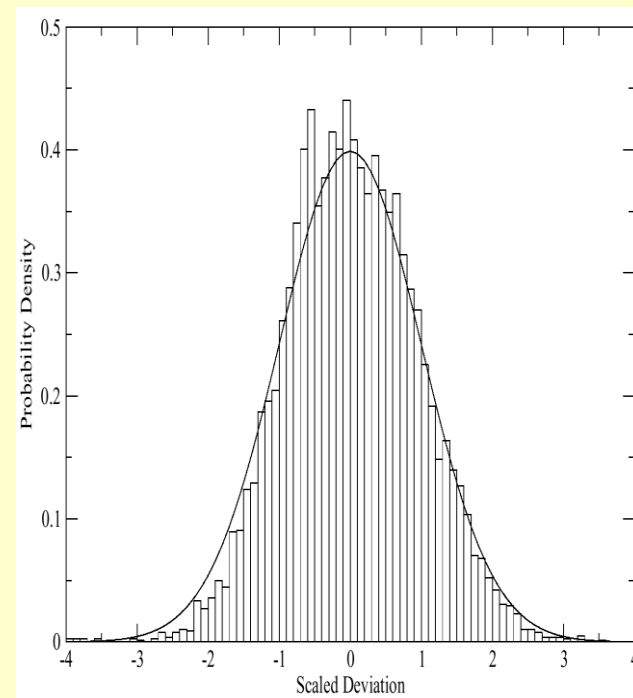
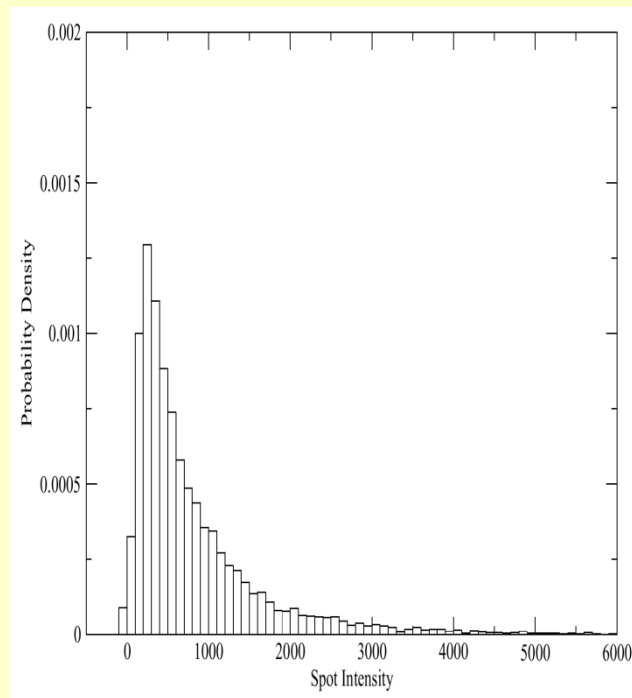
$$\text{Var}(g(X)) \approx [g'(\mu)]^2 \text{Var}(X) = g'(\mu)^2 h^2(\mu) \sigma^2$$

Assim,

$$\text{Var}(g(X)) \approx \text{cte} \rightarrow g'(\mu) = 1/h(\mu)$$

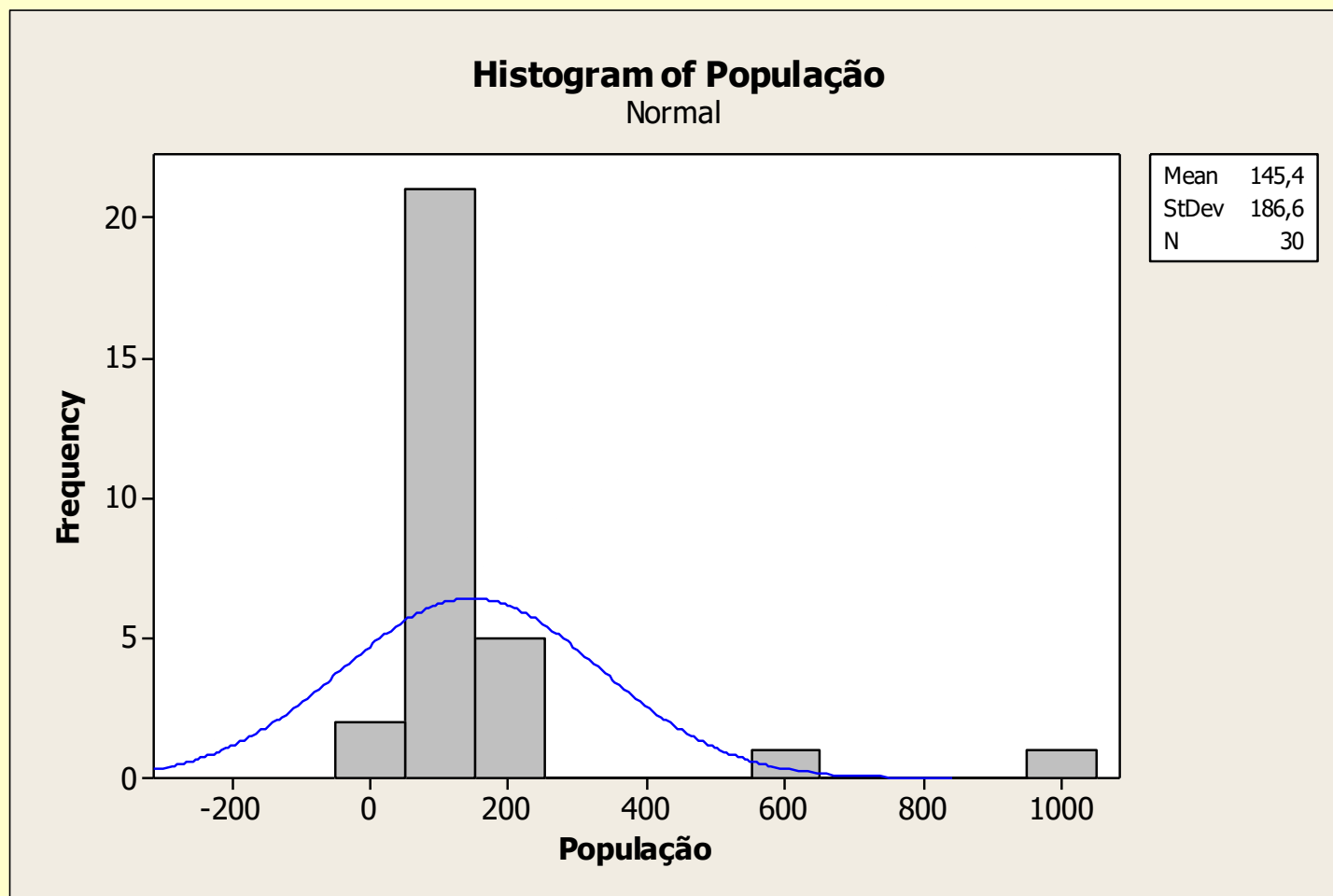
Transformação Log

- Modelo de Efeitos Multiplicativos $(y_j = \lambda e^{\beta C_j}) \Rightarrow$ Linearizar
- Distribuição Assimétrica Positiva $\Rightarrow$ Normalizar
- Estabilizar Variância

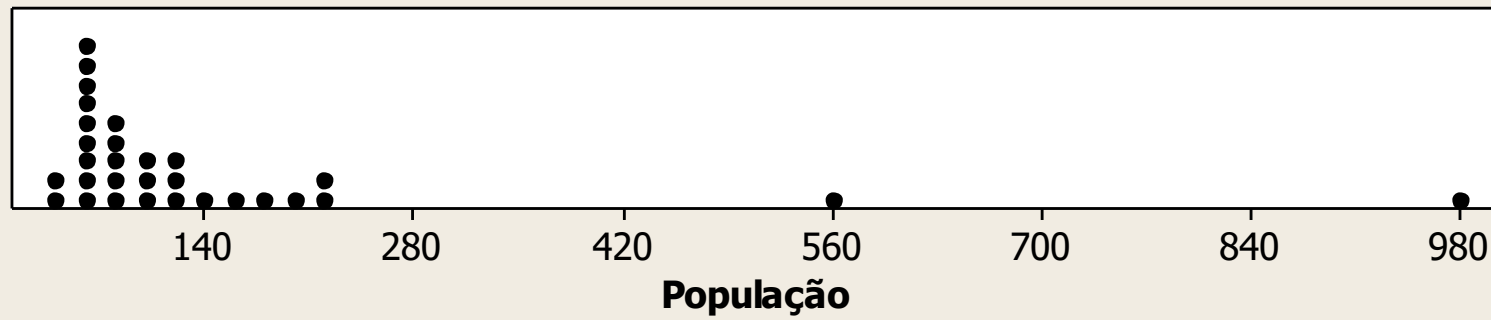


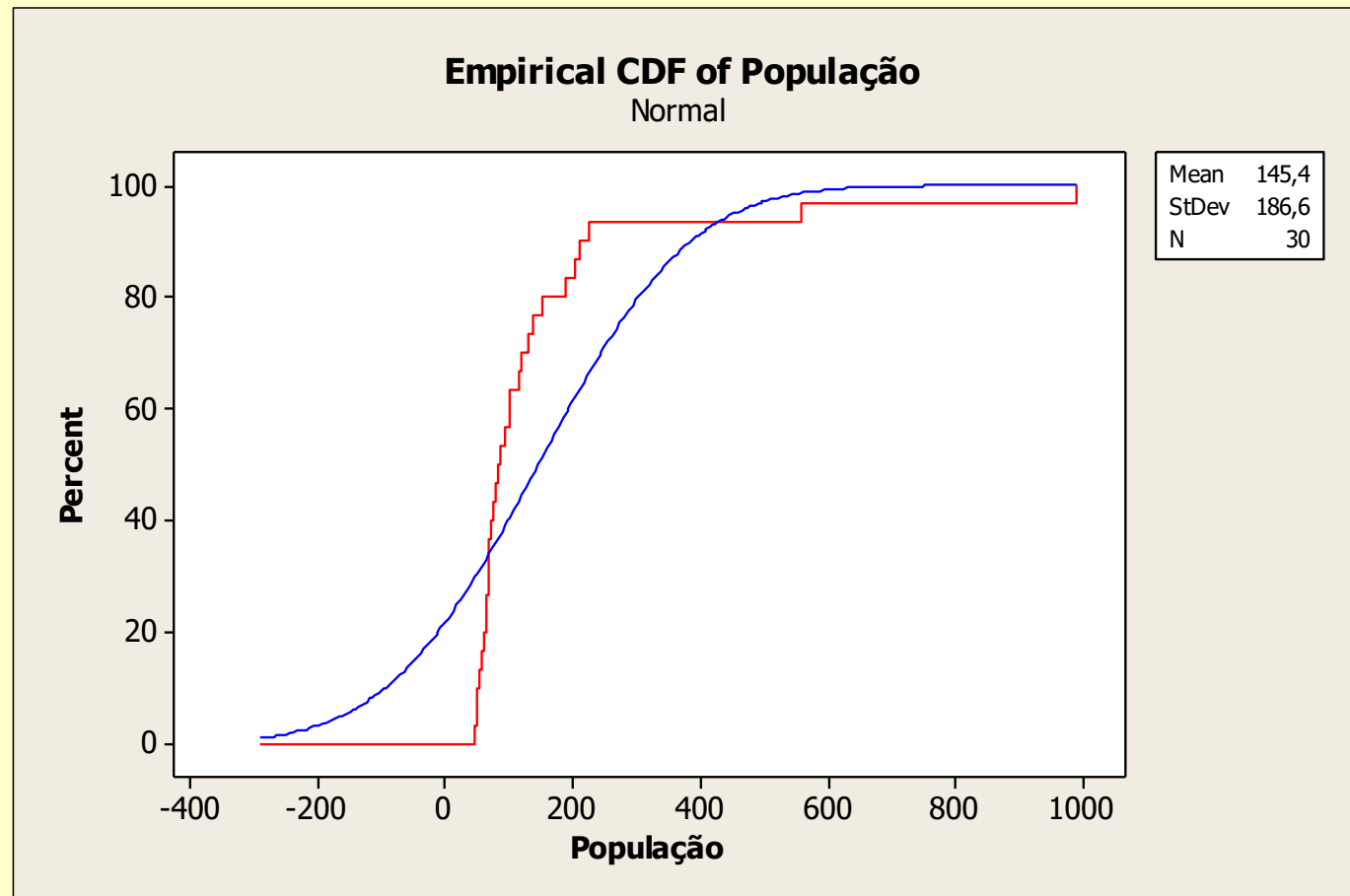
Populações(em 10.000 habitantes) dos 30 municípios mais populosos do Brasil

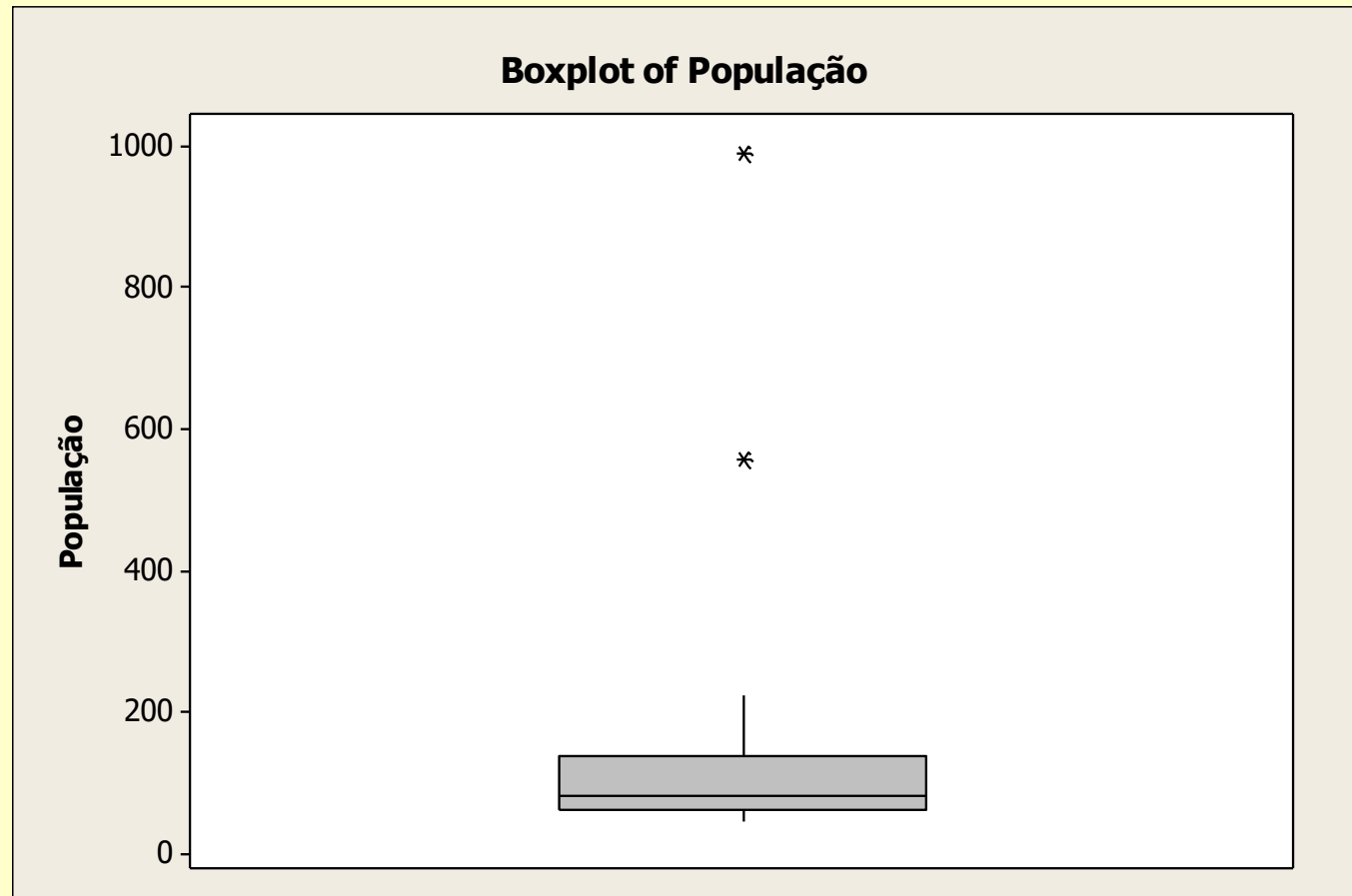
Município	População	Município	População	
São Paulo	988,8	Nova Iguaçu	83,9	
Rio de Janeiro	556,9	São Luis	80,2	
Salvador	224,6	Maceió	74,7	
Belo Horizonte	210,9	Duque de Caxias	72,7	
Fortaleza	201,5	São Bernardo do Campo	68,4	
Brasília	187,7	Natal	66,8	
Curitiba	151,6	Teresina	66,8	
Recife	135,8	Osasco	63,7	
Porto Alegre	129,8	Santo André	62,8	
Manaus	119,4	Campo Grande	61,9	
Belem	116,0	João Pessoa	56,2	
Goiania	102,3	Jaboatão	54,1	
Guarulhos	101,8	Contagem	50,3	
Campinhas	92,4	São Jose dos Campos	49,7	
São Gonçalo	84,7	Ribeirão Preto	46,3	



Dotplot of População



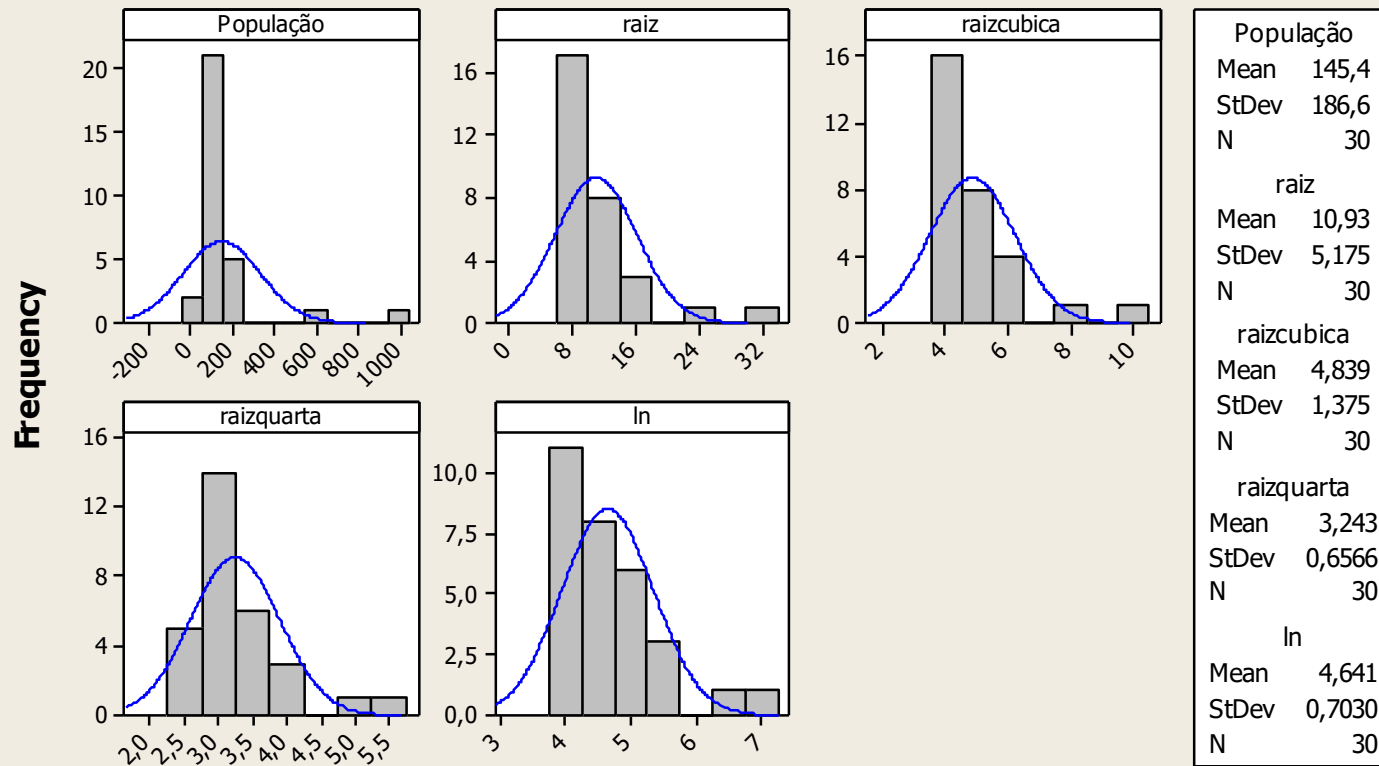




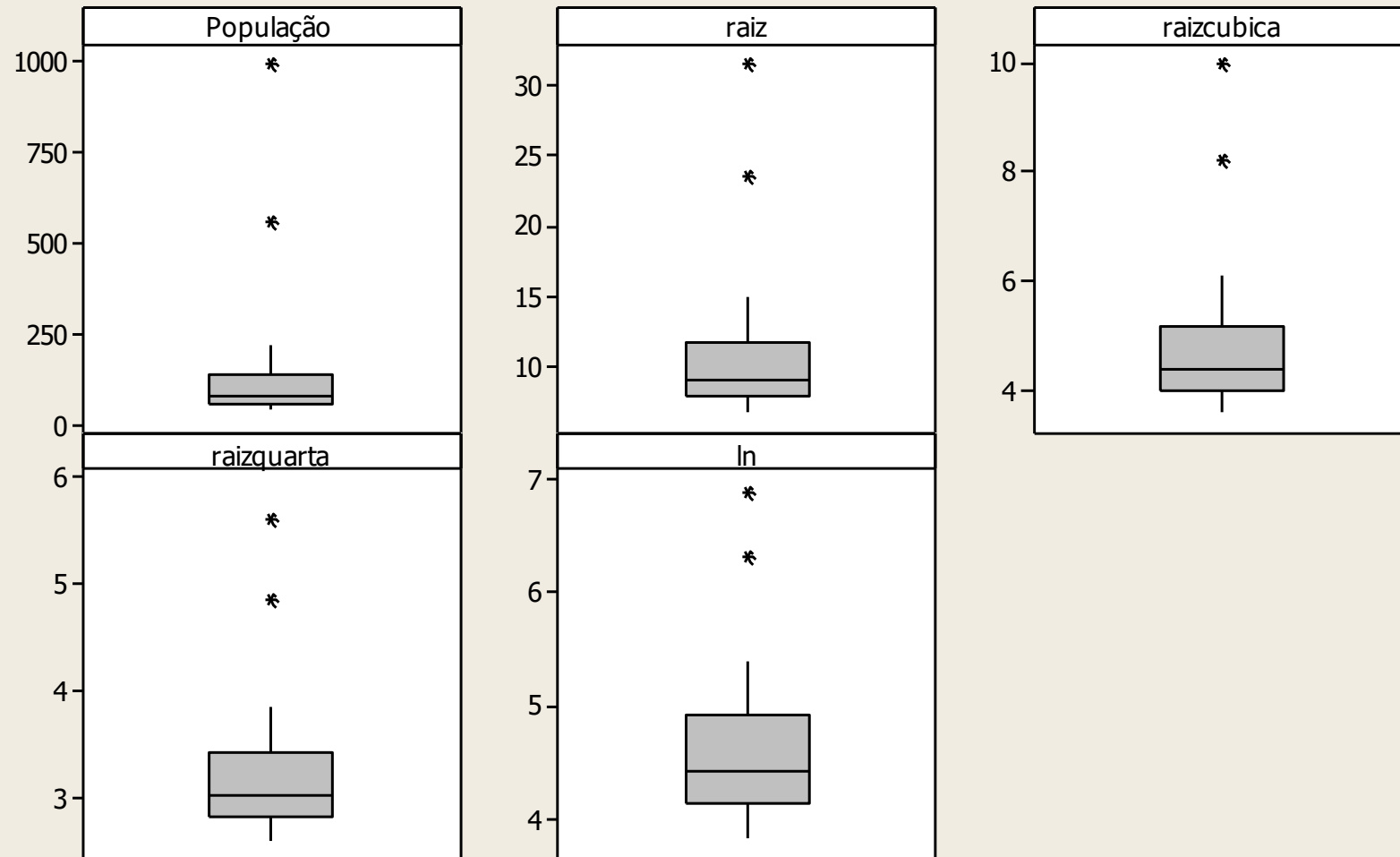
Descriptive Statistics: População

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
População	30	0	145,4	34,1	186,6	46,3	63,5	84,3	139,8	988,8

Histogram of População; raiz; raizcubica; raizquarta; ln Normal



Boxplot of População; raiz; raizcubica; raizquarta; ln



Descriptive Statistics: População; raiz; raizcubica; raizquarta; ln

Variable	N	N*	Mean	SE Mean	StDev	CoefVar	Minimum	Q1	Median
População	30	0	145,4	34,1	186,6	128,31	46,3	63,5	84,3
raiz	30	0	10,933	0,945	5,175	47,33	6,804	7,967	9,181
raizcubica	30	0	4,839	0,251	1,375	28,41	3,591	3,989	4,385
raizquarta	30	0	3,243	0,120	0,657	20,25	2,609	2,823	3,030
ln	30	0	4,641	0,128	0,703	15,15	3,835	4,151	4,434

Variable	Q3	Maximum	Skewness	Kurtosis
População	139,8	988,8	3,76	15,40
raiz	11,818	31,445	2,76	8,78
raizcubica	5,188	9,963	2,38	6,62
raizquarta	3,438	5,608	2,18	5,60
ln	4,939	6,896	1,61	3,00

Retorno

Um dos objetivos em finanças é a avaliação de **riscos** de uma carteira de ativos financeiros. O risco é medido em termos de **variações** de preços dos ativos.

P_t : preço de um ativo no instante t ;

$\Delta P_t = P_t - P_{t-1}$: a variação do preço.

Retorno liquido simples (a variação relativa de preços):

$$R_t = (P_t - P_{t-1}) / P_{t-1} = \Delta P_t / P_{t-1}$$

$1 + R_t = P_t / P_{t-1}$: **retorno bruto simples**

$$p_t = \log P_t$$

log retorno (ou simplesmente retorno):

$$r_t = \log(P_t / P_{t-1}) = \log(1 + R_t)$$

Para um u pequeno, temos $\log(1+u) \approx u$

Portanto, $r_t = \log(1+R_t) \approx R_t$

O retorno simples de período k ($t-k, t$):

$$1 + R_t(k) = (P_t - P_{t-k}) / P_{t-k} = (1+R_t)(1+R_{t-1})\dots(1+R_{t-k+1})$$

O log-retorno de período k :

$$\begin{aligned} r_t(k) &= \log(P_t/P_{t-k}) = \log(1 + R_t(k)) = \\ &= \log(1+R_t) + \log(1+R_{t-1}) + \dots + \log(1+R_{t-k+1}) \\ &= r_t + r_{t-1} + \dots + r_{t-k+1} \end{aligned}$$

Exemplo:

Um mês compreende normalmente cerca de 21 dias de transações, de modo que o log-retorno continuamente composto em um mês é dado por:

$$r_t(21) = r_t + r_{t-1} + \dots + r_{t-20}$$

para todo t .

Exemplo 1.1. Na Figura 1.1 (a), temos o gráfico dos Índices diários da Bolsa de Valores de São Paulo (Ibovespa) no período de 4 de julho de 1994 a 29 de setembro de 2010, num total de $T = 4019$ observações. Estes são os valores de fechamento do Índice. Podemos ter outras informações, como os valores de abertura, mínimo e máximo, por exemplo.

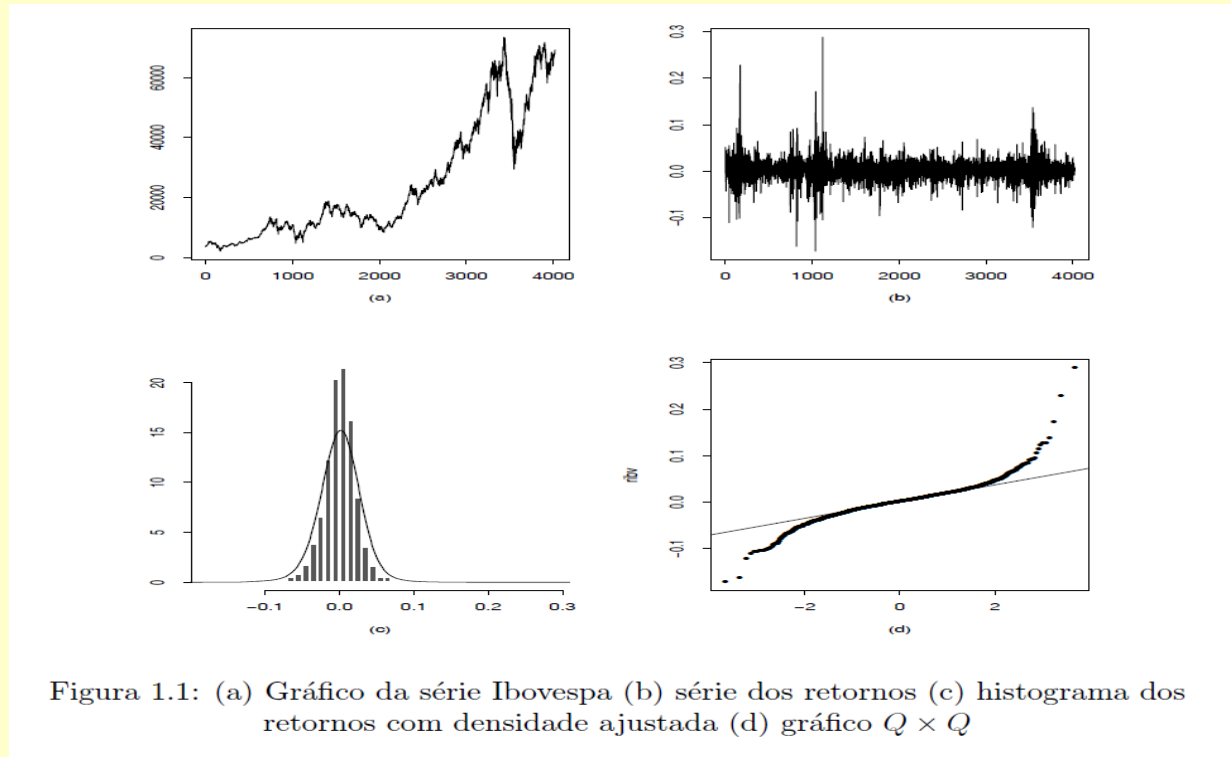
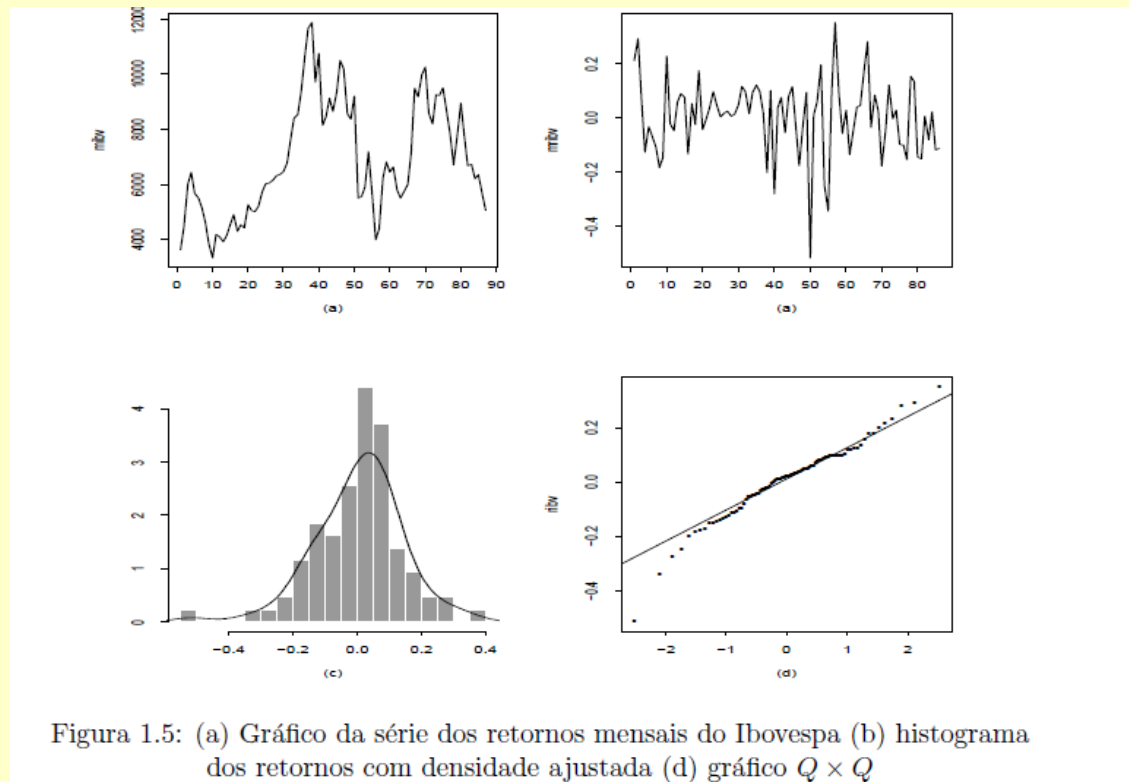


Figura 1.1: (a) Gráfico da série Ibovespa (b) série dos retornos (c) histograma dos retornos com densidade ajustada (d) gráfico $Q \times Q$

Exemplo 1.4. Considere os índices diários do Ibovespa do exemplo 1.1 e sejam $P_1 = 3580,80, \dots, P_5 = 4051,90$. Então,

$$\begin{aligned} R_2 &= \frac{P_2 - P_1}{P_1} = -0,004608, \\ r_2 &= \log \frac{P_2}{P_1} = -0,004619, \\ R_5[3] &= \frac{P_5 - P_2}{P_2} = 0,136801, \\ r_5[3] &= r_5 + r_4 + r_3 = \log \frac{P_5}{P_2} = 0,128218. \end{aligned}$$

Exemplo. Na Figura 1.5 (a), temos os índices **mensais** do Ibovespa, enquanto que na Figura 1.5 (b) temos os respectivos retornos, no período de junho de 1994 a agosto de 2001, com $T = 86$ dados. Observe que obtemos uma **série mais suave**, ou seja, com menor variabilidade do que a série de retornos diários.



Fatos estilizados sobre retornos

Séries econômicas e financeiras apresentam algumas características que são comuns a outras séries temporais, como:

- a) tendências;
- b) sazonalidade;
- c) Pontos influentes;
- d) Heteroscedasticidade condicional;
- e) Não linearidade.

Os retornos financeiros apresentam, outras características peculiares, que muitas séries não apresentam.

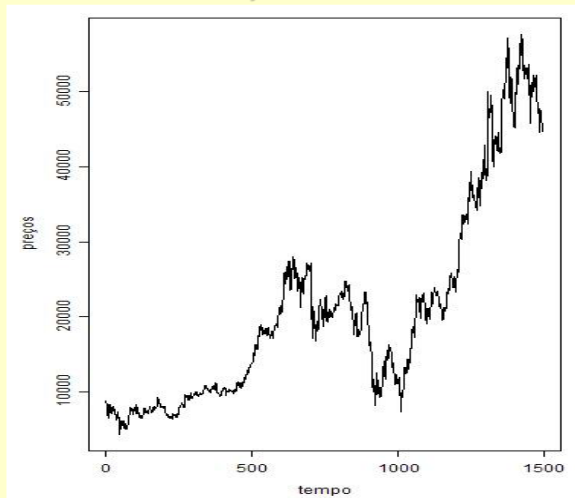
Retorno: não tendência, não sazonalidade, com exceção de retorno intra-diários.

Séries de preços, de taxa de câmbio e de taxa de juros: tendência que variam no tempo.

Os principais fatos estilizados relativos a retornos financeiros:

1. Não auto-correlacionados;
2. Os quadrados dos retornos são auto-correlacionados;
3. Agrupamentos de volatilidades ao longo do tempo;
4. Em geral, não são normais.

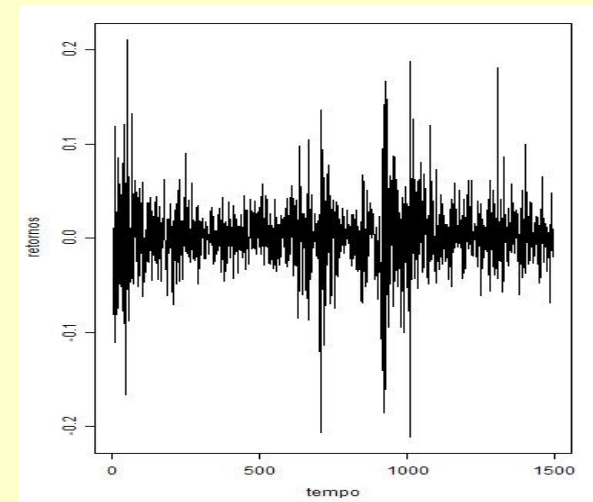
Série ações Petrobras



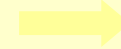
Interesse: Medir variações de preços de ativos para avaliar o risco.

- Livres de escala
- Não auto-correlacionado
- Quadrado auto-correlacionado
- Distribuição com caudas pesadas
- Não lineares
- Grupamento de volatilidades

Retornos Petrobras

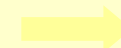


retorno



$$r_t = \ln \frac{P_t}{P_{t-1}} = \ln(1 + R_t) = \ln(P_t) - \ln(P_{t-1}) = p_t - p_{t-1}$$

retorno de período k



$$r_t(k) = \ln \frac{P_t}{P_{t-k}} = \ln(1 + R_t(k)) = p_t - p_{t-k}$$

O gráfico dos Índices diários da Bolsa de Valores de São Paulo (Ibovespa) no período de 4 de julho de 1994 a 29 de setembro de 2010, num total de $T = 4019$ observações. Estes são os valores de fechamento do Índice. Podemos ter outras informações, como os valores de abertura, mínimo e máximo, por exemplo.

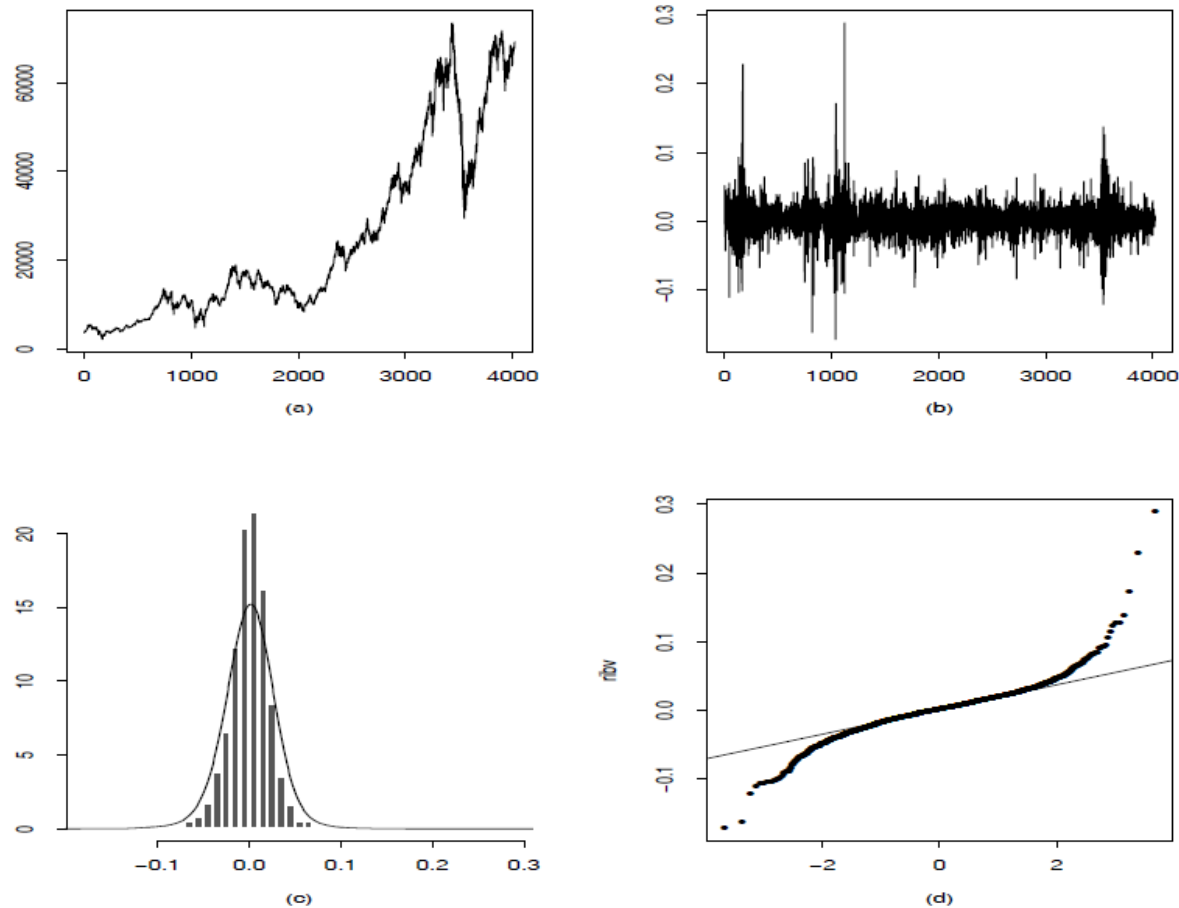


Figura 1.1: (a) Gráfico da série Ibovespa (b) série dos retornos (c) histograma dos retornos com densidade ajustada (d) gráfico $Q \times Q$

o gráfico dos Índices diários do Dow Jones Industrial Average (fechamento), DJIA, no período de 3 de janeiro de 1995 a 26 de dezembro de 2002, com $T = 1992$ observações.

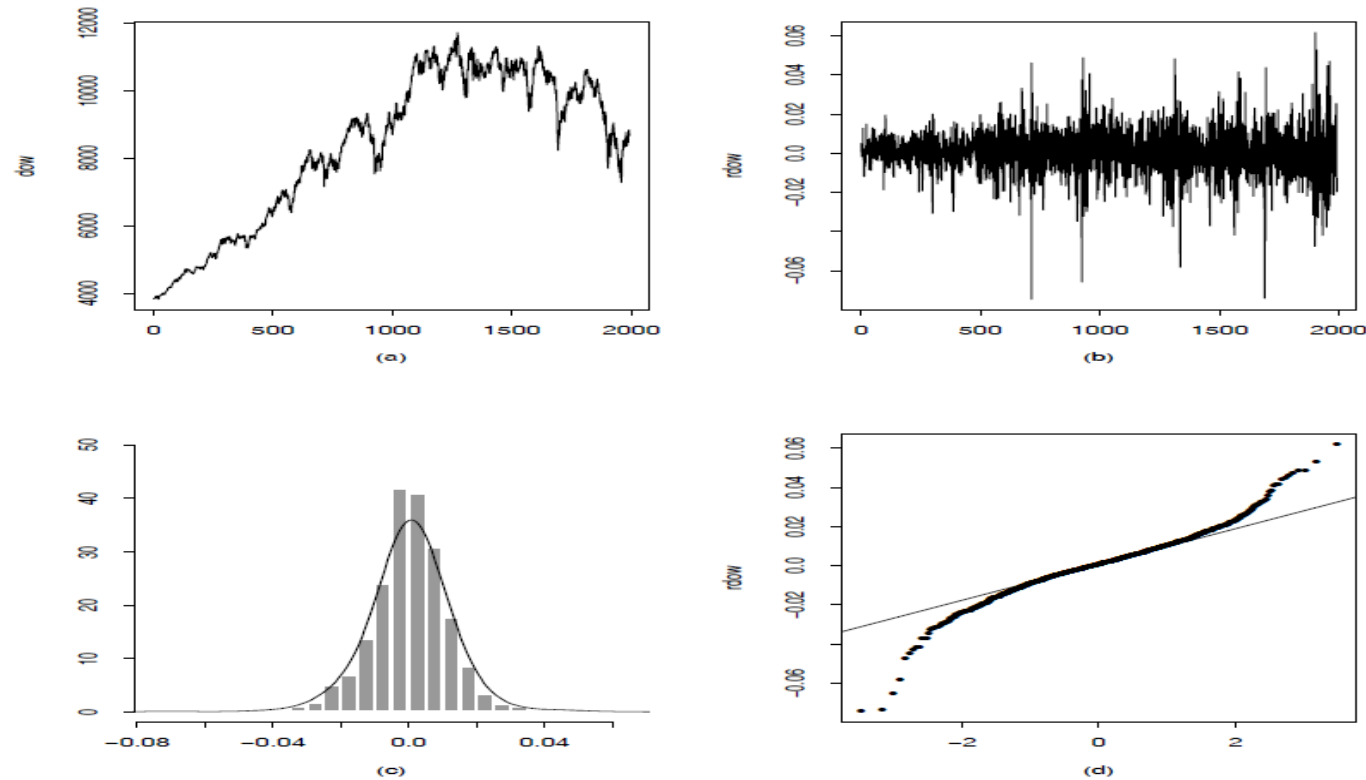


Figura 1.2: (a) Gráfico da série DJIA (b) série dos retornos (c) histograma dos retornos com densidade ajustada (d) gráfico $Q \times Q$

o gráfico de parte dos dados do Ibovespa, observados a cada quinze minutos, de 6 de abril de 1998 a 13 de agosto de 2003, num total de 1309 dias e $T = 37.961$ observações

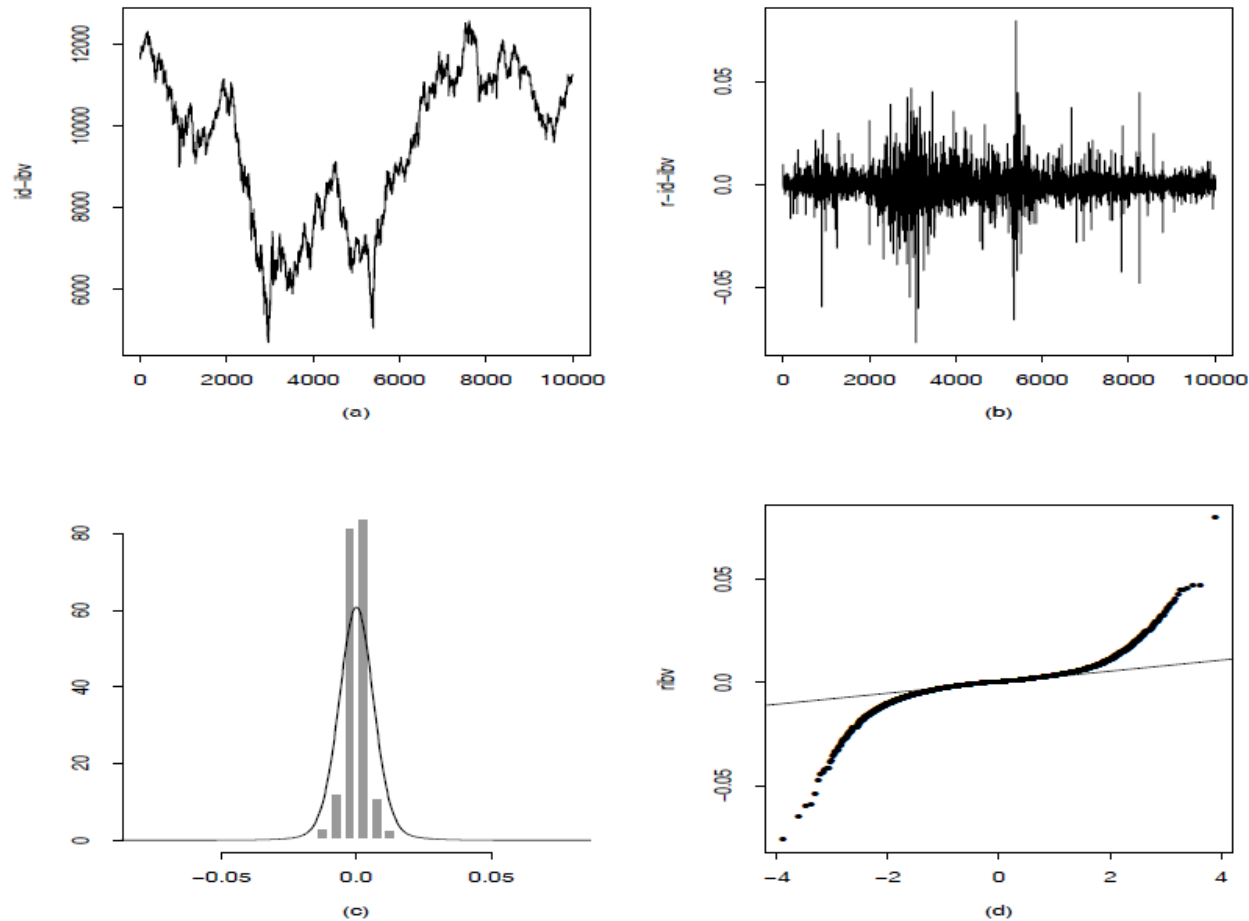
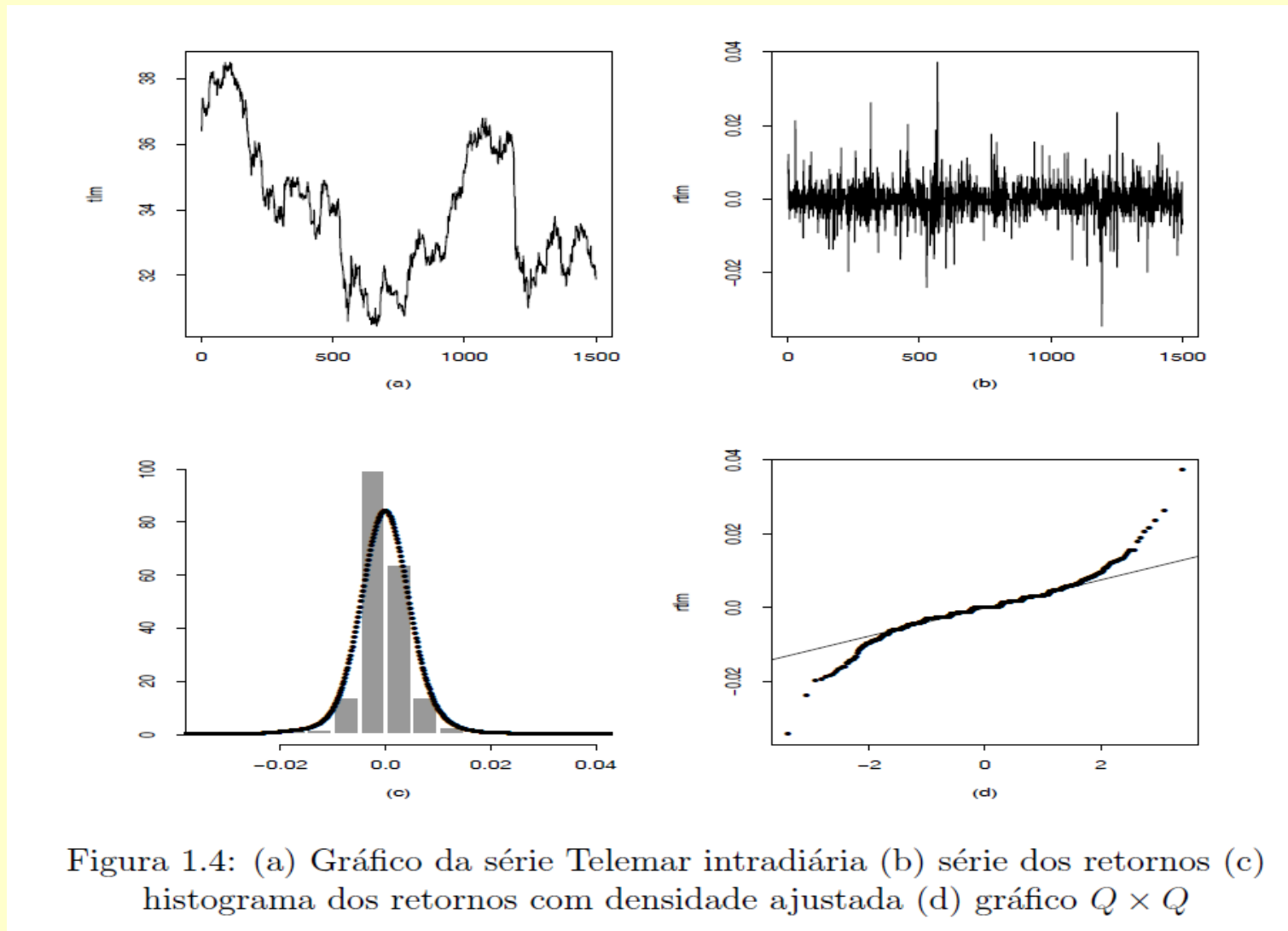


Figura 1.3: (a) Gráfico da série Ibovespa intradiária (b) série dos retornos (c) histograma dos retornos com densidade ajustada (d) gráfico $Q \times Q$

o gráfico de parte dos dados da Telemar PN, observados a cada 15 minutos, de 2 de janeiro de 2002 a 31 de março de 2005, com $T = 21.429$ observações.



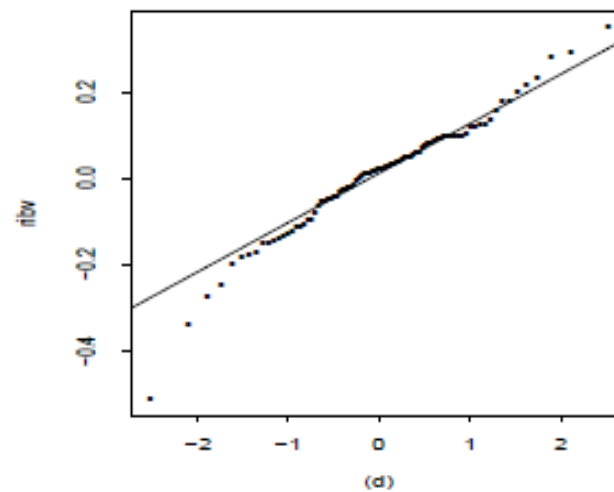
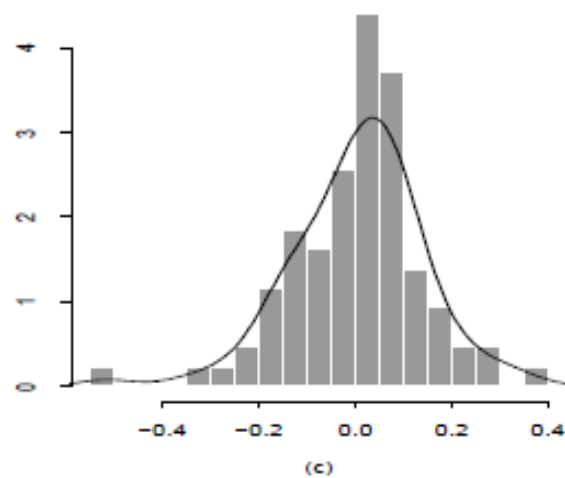
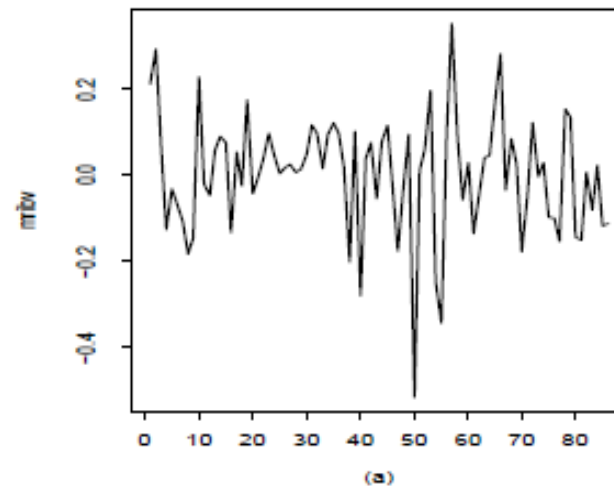
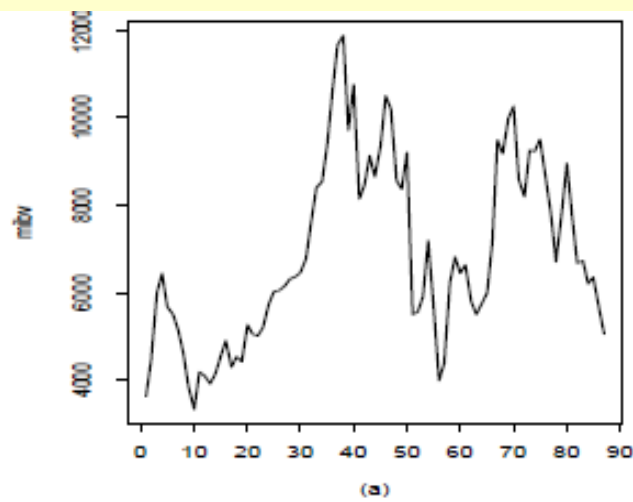


Figura 1.5: (a) Gráfico da série dos retornos mensais do Ibovespa (b) histograma dos retornos com densidade ajustada (d) gráfico $Q \times Q$