

Séries Temporais Financeiras

aula 02

Agregação de Retornos

O log-retorno de período k (*agregação temporal* dos retornos)

$$r_t[k] = \log \frac{P_t}{P_{t-k}} = \log(1 + R_t[k]) = \sum_{j=0}^{k-1} \log(1 + R_{t-j}) = \sum_{j=0}^{k-1} r_{t-j}.$$

Agregação *cross-section* (transversal), para diversos ativos de uma carteira de investimentos, c :

Suponha que esta contenha N instrumentos A_1, \dots, A_N , com pesos w_1, \dots, w_N , com $\sum w_i = 1$.

Denotemos por R_i os retornos simples e por r_i os log-retornos desses ativos, $i = 1, \dots, N$.

Se P_0 indicar o preço inicial da carteira, após um período teremos, para **retornos continuamente compostos**,

$$\frac{P_1}{P_0} = \sum_{i=1}^N w_i e^{r_i}.$$

O log-retorno da carteira é $r_c = \log \frac{P_1}{P_0}$, logo obtemos

$$r_c = \log \sum_{i=1}^N w_i e^{r_i},$$

O retorno simples da carteira é

$$\begin{aligned}R_c &= \frac{P_1}{P_0} - 1 \\&= \sum_{i=1}^N w_i e^{r_i} - 1 \\&= \sum_{i=1}^N w_i (1 + R_i) - 1,\end{aligned}$$

$$R_c = \sum_{i=1}^N w_i R_i.$$

No caso de **composição discreta** teremos

$$\frac{P_1}{P_0} = \sum_{i=1}^N w_i (1 + r_i),$$

de modo que o retorno simples da carteira é $R_c = (P_1 - P_0)/P_0$, ou seja,

$$R_c = \sum_{i=1}^N w_i (1 + r_i) - 1 = \sum_{i=1}^N w_i r_i.$$

O retorno simples é uma soma ponderada de retornos simples, no caso de composição contínua, e uma soma ponderada de log-retornos, no caso de composição discreta.

(i) Agregação temporal: para $i = 1, \dots, N$,

$$R_{i,t}[k] = \prod_{j=0}^{k-1} (1 + R_{i,t-j}) - 1,$$

$$r_{i,t}[k] = \sum_{j=0}^{k-1} r_{i,t-j},$$

para retornos simples e log-retornos, respectivamente.

(ii) Agregação cross-section: para a carteira c e período t ,

$$R_{c,t} = \sum_{i=1}^N w_i R_{i,t},$$

$$r_{c,t} = \log \left(\sum_{i=1}^N w_i e^{r_{i,t}} \right).$$

$$r_{c,t} \approx \sum_{i=1}^N w_i r_{i,t}.$$

Para **agregação temporal** é mais conveniente trabalhar com log-retornos, enquanto que para **agregação cross-section** os retornos simples são mais convenientes.

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 (arquivo m-ibv94.01.dat). Esses retornos mensais são obtidos usando-se a fórmula,

$$r_t[k] = \log \frac{P_t}{P_{t-k}} = \log(1 + R_t[k]) = \sum_{j=0}^{k-1} \log(1 + R_{t-j}) = \sum_{j=0}^{k-1} r_{t-j}.$$

ou seja, somando-se os retornos diários. Observe que obtemos uma série mais suave, ou seja, com menor variabilidade do que a série de retornos diários.

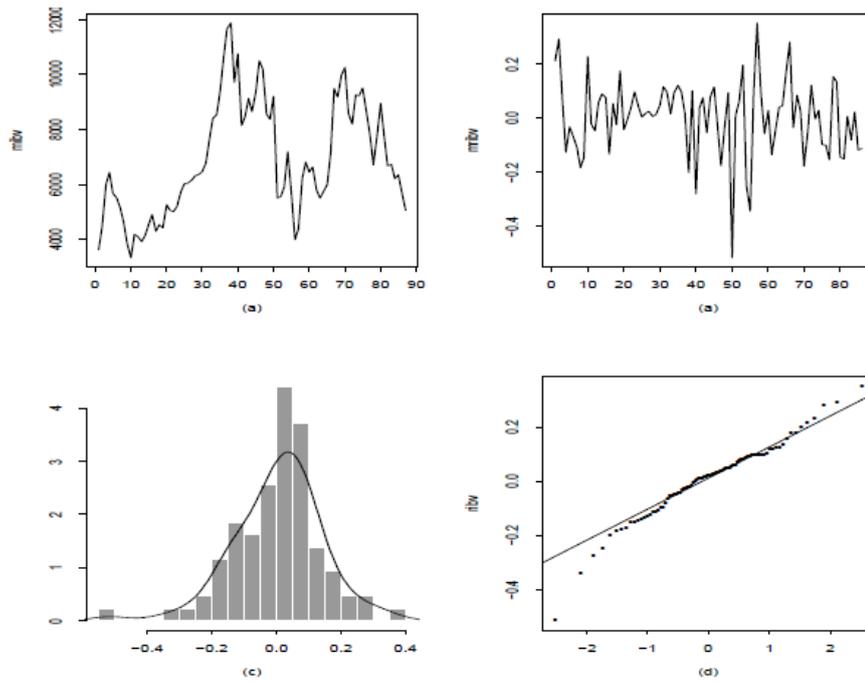


Figura 1.5: (a) Gráfico da série dos retornos mensais do Ibovespa (b) histograma dos retornos com densidade ajustada (c) gráfico $Q \times Q$

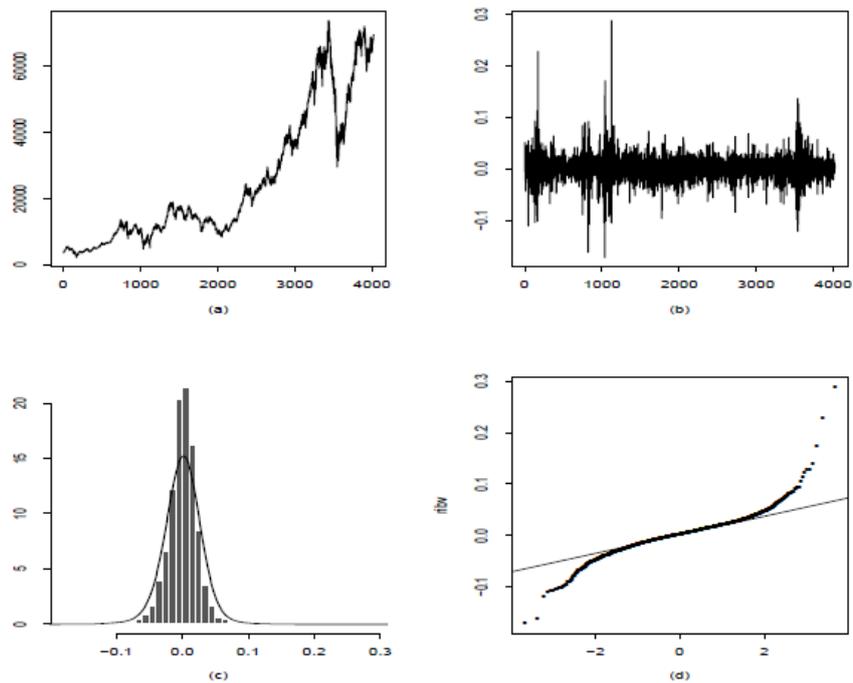


Figura 1.1: (a) Gráfico da série Ibovespa (b) série dos retornos (c) histograma dos retornos com densidade ajustada (d) gráfico $Q \times Q$

Distribuição de Retornos

Considere uma série de retornos $\{r_t, t = 1, \dots, T\}$, observados em instantes de tempo igualmente espaçados. Essa série pode ser considerada parte de uma realização de um processo estocástico $\{r_t, t \in Z\}$, onde $Z = \{0, \pm 1, \pm 2, \dots\}$. O processo estará especificado completamente se conhecermos as distribuições finito-dimensionais

$$F(x_1, \dots, x_n; t_1, \dots, t_n) = P(r(t_1) \leq x_1, \dots, r(t_n) \leq x_n),$$

para quaisquer instantes de tempo t_1, \dots, t_n e qualquer $n \geq 1$.

Caracterizar o processo por momentos até determinada ordem, como a média e a função de autocovariância

$$E(r_t) = \int_{-\infty}^{\infty} r dF(r; t)$$

$$\gamma(t_1, t_2) = E(r_{t_1} r_{t_2}) - E(r_{t_1})E(r_{t_2}), \quad t_1, t_2 \in Z.$$

Outras suposições simplificadoras podem ser introduzidas, como condições de estacionariedade, ergodicidade ou normalidade do processo.

Os preços P_t em geral não são estacionários, ao passo que os log-retornos o são, donde o interesse nesses últimos. Todavia, a suposição de normalidade dos log-retornos, em geral, não é válida.

se tivermos N ativos com retornos r_{it} em T instantes de tempo, teríamos que considerar as distribuições

$$F(r_{1,1}, \dots, r_{N,1}; \dots; r_{1,T}, \dots, r_{N,T}),$$

que usualmente podem depender de outras variáveis e parâmetros desconhecidos.

O estudo dessas distribuições é muito geral e há necessidade de introduzir restrições. Por exemplo, podemos supor que a distribuição é a mesma para todo instante de tempo (invariância temporal).

As distribuições finito-dimensionais

$$F(x_1, \dots, x_n; t_1, \dots, t_n) = P(r(t_1) \leq x_1, \dots, r(t_n) \leq x_n),$$

Podem ser escritas como (tomando-se $t_i = i$, $i = 1, \dots, n$ e omitindo a dependência de F sobre esses tempos):

$$F(r_1, \dots, r_n) = F_1(r_1)F_2(r_2|r_1) \dots F_n(r_n|r_1, \dots, r_{n-1}).$$

Uma hipótese muitas vezes formulada é que os retornos são **temporalmente independentes**, ou seja, não são previsíveis usando retornos passados. Nessa situação, teremos que

$$F_t(r_t|r_1, \dots, r_{t-1}) = F_t(r_t).$$

Ergodicidade é uma propriedade mais difícil de estabelecer. Basicamente, um processo é **ergódico** se pudermos estimar

características de interesse (média, autocovariância etc) a partir de uma única trajetória do processo. Assim, um processo é ergódico na média se a média amostral convergir, em probabilidade, para a média verdadeira do processo.

Uma outra suposição que às vezes é feita sobre a distribuição dos retornos é que esta seguiria uma **distribuição estável**.

A função de distribuição de

$$F(r_1, \dots, r_n) = F_1(r_1)F_2(r_2|r_1) \dots F_n(r_n|r_1, \dots, r_{n-1}).$$

depende, em geral, de co-variáveis Y e de um vetor de parâmetros, Θ , que a caracterizam. Supondo retornos com distribuição contínua, podemos obter a função de verossimilhança e, a partir dela, estimar Θ .

Por exemplo, supondo-se que as distribuições condicionais $f_t(r_t|r_1, \dots, r_{t-1})$ sejam normais, com média μ_t e variâncias σ_t^2 , então $\theta = (\mu_t, \sigma_t^2, t = 1, \dots, n)$ e a função de verossimilhança ficará

$$f(r_1, \dots, r_n; \theta) = f_1(r_1; \theta) \prod_{t=2}^n \frac{1}{\sigma_t \sqrt{2\pi}} \exp\left(-\frac{(r_t - \mu_t)^2}{2\sigma_t^2}\right).$$

O estimador de máxima verossimilhança de θ é obtido maximizando-se essa função ou o logaritmo dela.

Podemos considerar N ativos ao longo do tempo, r_{1t}, \dots, r_{Nt} ,

$$\mathbf{r}_t = (r_{1t}, r_{2t}, \dots, r_{Nt})'$$

O interesse reside no estudo das distribuições condicionais

$$F_t(\mathbf{r}_t | \mathbf{r}_1, \dots, \mathbf{r}_{t-1}, \mathbf{Y}, \boldsymbol{\theta}).$$

Há vários dispositivos gráficos que podemos utilizar para avaliar a forma da distribuição dos retornos: o **histograma**, **estimativas da função densidade** e **gráficos quantis-quantis** (os chamados “Q×Q plots”).

Histogramas

A partir dos dados x_1, \dots, x_n o histograma pode ser definido por

$$H(x) = \sum_{i=1}^n I\{x - \tilde{x}_i; h\},$$

onde \tilde{x}_i é o centro do intervalo onde a observação x_i cai e $I\{z; h\}$ é o indicador do intervalo $[-h, h]$. Algum tipo de escalamento é feito para que a área do histograma seja um.

Críticas:

- o comportamento do histograma depende da escolha de h e da posição inicial da grade,
- informação é perdida, pois substituímos x_i pelo ponto médio do intervalo ao qual ele pertence.

Para evitarmos estas dificuldades, estimadores mais suaves da densidade $f(x)$ podem ser usados.

Um dos mais utilizados substitui retângulos por uma função

núcleo (“kernel”) mais suave, obtendo-se

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n k(x - x_i; h),$$

onde k é o núcleo, em geral também uma função densidade, cuja variância é controlada por h , chamada *largura de faixa* (*bandwidth*). O comportamento de \hat{f} vai depender de h , de modo que uma crítica ao histograma permanece. Mas é mais fácil comparar estimadores deste tipo do que histogramas. Uma escolha usual para k é $k(z; h) = \phi(z; h)$, a densidade normal com média zero e desvio padrão h .

O programa SPlus usa a função `hist` para obter o histograma de um conjunto de dados e as funções `ksmooth` e `density` para obter \hat{f} e pode-se escolher entre os núcleos retangular, triangular, cosseno e gaussiano, usando a opção `window`.

Q×Q Plots

Suponha que a v.a. X tenha distribuição contínua, com f.d.a. F . Então, para $0 \leq p \leq 1$, o p -quantil de F é o valor Q_p satisfazendo $F(Q_p) = p$, ou seja,

$$F(Q_p) = P(X \leq Q_p) = p.$$

Se existir a inversa de F , então $Q_p = F^{-1}(p)$. No caso de X ser discreta, a definição tem que ser modificada: o p -quantil é o valor Q_p satisfazendo

$$\begin{aligned} P(X \leq Q_p) &\geq p \\ P(X \geq Q_p) &\geq 1 - p. \end{aligned}$$

Dado um conjunto de observações podemos calcular os *quantis empíricos*. Uma maneira é considerar a *função de distribuição empírica* \hat{F}_n como estimador de F , ou seja, dadas as observações X_1, \dots, X_n de X ,

$$\hat{F}_n(x) = \frac{1}{n} \#\{i : 1 \leq i \leq n, X_i \leq x\}.$$

Então, o quantil Q_p é estimado pelo p -quantil de \hat{F}_n . Ou seja, o p -quantil estimado, q_p , seria definido por $\hat{F}_n(q_p) = p$. Contudo, usaremos um enfoque um pouco diferente.

Chamemos de r_1, \dots, r_T os retornos observados e considere as estatísticas de ordem $r_{(1)} \leq r_{(2)} \leq \dots \leq r_{(T)}$. Um estimador consistente de Q_p é dado pelo p -quantil empírico, definido por

$$q_p = \begin{cases} r_{(i)}, & \text{se } p = p_i = (i - 0,5)/T, \quad i = 1, \dots, T \\ (1 - f_i)r_{(i)} + f_i r_{(i+1)}, & \text{se } p_i < p < p_{i+1} \\ r_{(1)}, & \text{se } 0 < p < p_1 \\ r_{(T)}, & \text{se } p_T < p < 1, \end{cases} \quad (1.21)$$

onde $f_i = (p - p_i)/(p_{i+1} - p_i)$.

Ou seja, ordenados os dados, q_p é uma das estatísticas de ordem, se p for da forma $p_i = (i - 0,5)/T$ e está na reta ligando os pontos $(p_i, r_{(i)})$ e $(p_{i+1}, r_{(i+1)})$, se p estiver entre p_i e p_{i+1} . Tomamos p_i da forma escolhida e não como i/T para que, por exemplo, a mediana calculada segundo esta definição coincida com a definição usual.

Há dois tipos de gráficos $Q \times Q$: teóricos e empíricos. O primeiro tipo é usado para verificar se um conjunto de dados vem de determinada distribuição. O segundo tipo é usado para verificar se dois conjuntos de dados têm uma mesma distribuição. Para verificar se um conjunto de dados provêm de uma distribuição especificada, consideramos o gráfico em que, no eixo horizontal, colocamos os quantis teóricos da distribuição hipotetizada para os dados, e no eixo vertical, os quantis empíricos dos dados, ambos calculados nos pontos p_i acima. Se as observações realmente são provenientes da distribuição em questão, os pontos deverão estar distribuídos ao longo de uma reta.

Exemplo 1.1. (continuação) Na Figura 1.1 (c), temos o histograma dos retornos diários do Ibovespa, com uma densidade estimada a partir dos dados. Vemos que o histograma tem a parte central mais alta do que uma normal e há a presença de valores bastante afastados da posição central dos dados. Esses fatos são característicos de retornos financeiros e são descritos pela chamada medida de curtose, a ser estudada na seção seguinte. Dizemos que os retornos são leptocúrticos, com caudas mais pesadas que a normal. Na Figura 1.1 (d), temos o gráfico $Q \times Q$ com respeito aos quantis da distribuição normal padrão. Se os dados fossem aproximadamente normalmente distribuídos, os pontos estariam sobre uma reta, o que não acontece no caso em questão.

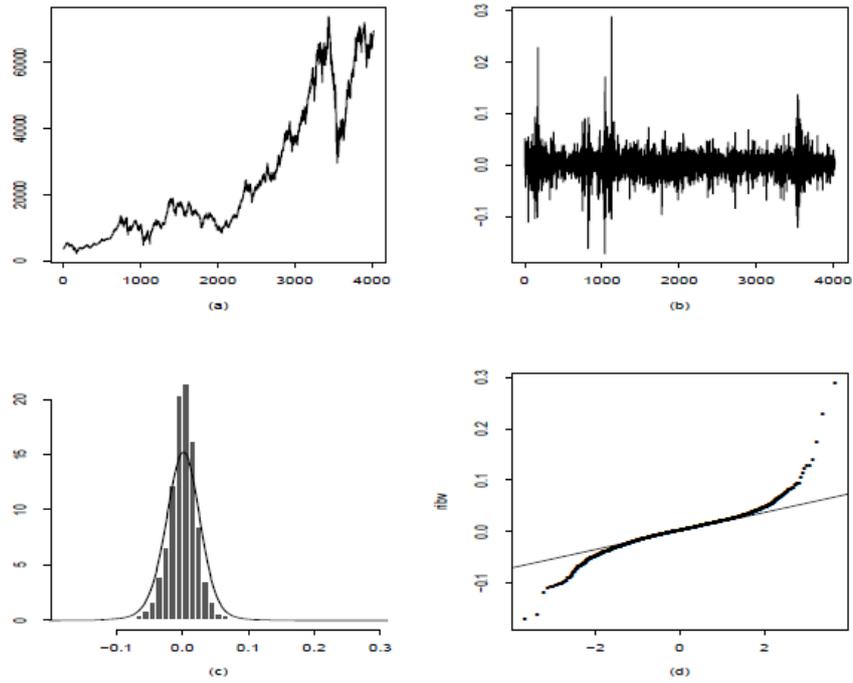


Figura 1.1: (a) Gráfico da série Ibovespa (b) série dos retornos (c) histograma dos retornos com densidade ajustada (d) gráfico $Q \times Q$

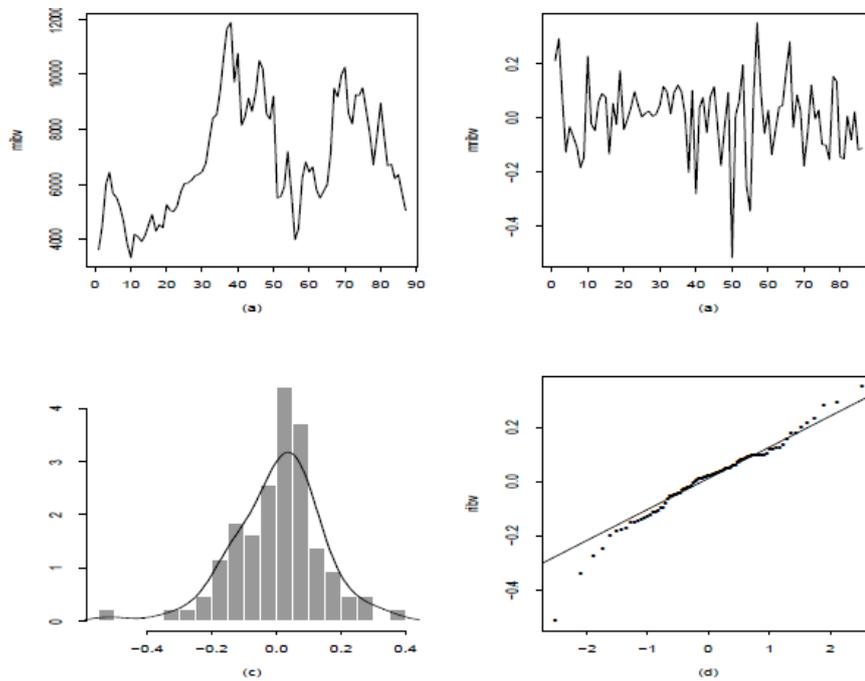


Figura 1.5: (a) Gráfico da série dos retornos mensais do Ibovespa (b) histograma dos retornos com densidade ajustada (c) gráfico $Q \times Q$

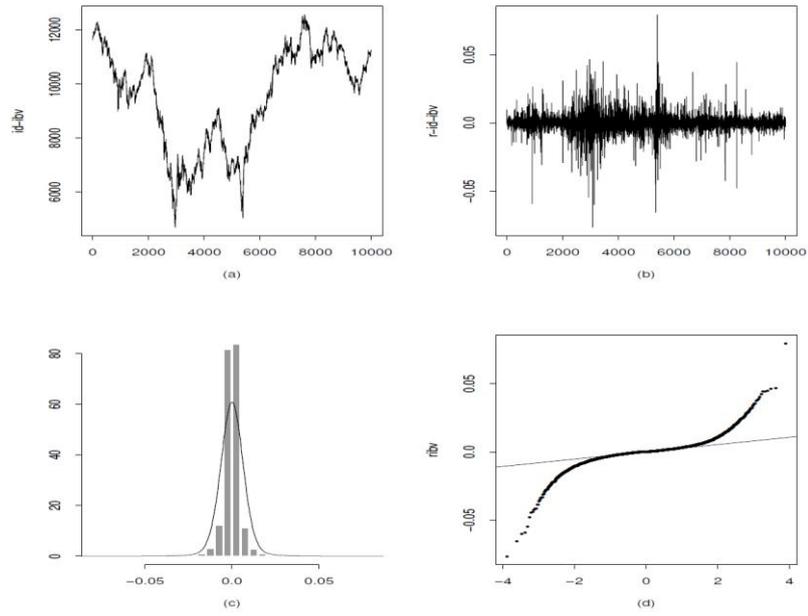


Figura 1.3: (a) Gráfico da série Ibovespa intradiária (b) série dos retornos (c) histograma dos retornos com densidade ajustada (d) gráfico $Q \times Q$.

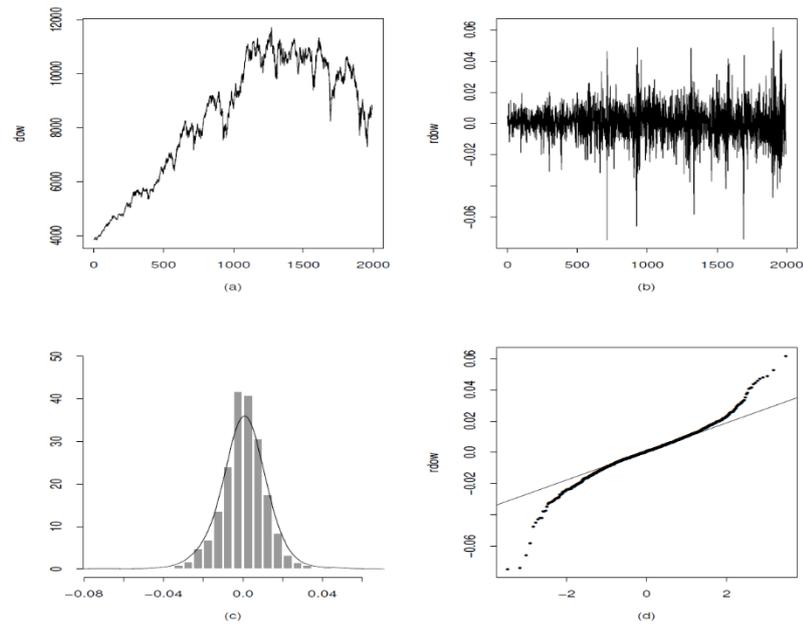


Figura 1.2: (a) Gráfico da série DJIA (b) série dos retornos (c) histograma dos retornos com densidade ajustada (d) gráfico $Q \times Q$.