

6 Reconstrução de Séries Financeiras

6.1 Introdução

Muitos fenômenos físicos que envolvem grandezas flutuantes podem ser descritos por uma equação diferencial estocástica do tipo de Itô-Langevin com a forma apresentada na Eq. (2.36). Sabe-se que a FDP associada é governada pela EFP (2.33) correspondente [24].

Para uma série temporal ideal, regida pela Eq. (2.36) e colecionada por um período suficientemente longo e com alta resolução, a dinâmica do processo pode ser reconstruída diretamente obtendo $D^{(1)}$ e $D^{(2)}$, como foi feito nos Capítulos 3 e 4, usando sua definição estatística, dada pela Eq. (2.34). Nesta seção da tese reconstruiremos mediante esse método e utilizando um algoritmo numérico, uma série de retornos intradiários do IBOVESPA.

6.2 Cálculo dos coeficientes de KM

Utilizamos como exemplo de reconstrução de uma série temporal, a série de retornos normalizados do IBOVESPA para a escala temporal linear de quatro minutos. Eles foram calculados a partir de uma série de preços catalogados a cada 30 segundos, sem sobreposição dos dados. O período adotado foi de 01 de novembro de 2002 a 19 de julho de 2006, totalizando 794.310 dados. Desta forma, a série de retornos de quatro minutos possui 99.288 dados.

Definimos como unidade de tempo a escala de 4 minutos, para a qual, portanto, $\tau \equiv \Delta t = 1$.

Da obtenção dos momentos condicionais, calculamos as funções $\tilde{D}^{(1)}(x, \tau)$ e $\tilde{D}^{(2)}(x, \tau)$ na Eq. (2.35), obtendo formas lineares e quadráticas em x respectivamente, como apresentado na Fig. 6.1. Para $\tau = 1$, obtivemos $(\tilde{a}_1, \tilde{b}_2, \tilde{b}_0) = (0, 85 \pm 0, 05; 0, 45 \pm 0, 03; 0, 40 \pm 0, 02)$.

Para calcular os coeficientes de KM necessários para a integração da Eq. (2.36), consideramos como primeira aproximação, as funções $\tilde{D}^{(1)}$ e $\tilde{D}^{(2)}$,

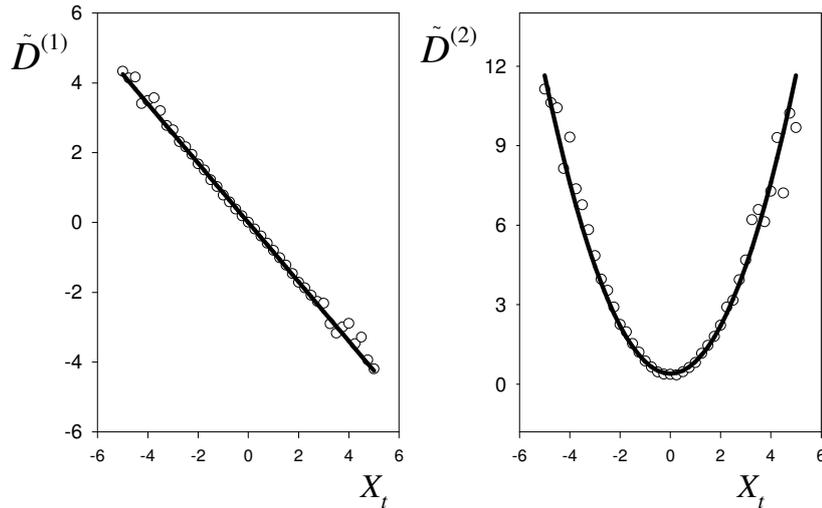


Figura 6.1: Coeficientes $\tilde{D}^{(1)}$ e $\tilde{D}^{(2)}$ dos dados numéricos para $\tau = 1$. Linhas correspondem aos melhores ajustes, linear e quadrático respectivamente, por mínimos quadrados aos dados.

obtidas para a menor escala disponível, ou seja, $\tau = 1$. Dado que $D^{(1)}(x) = -a_1x$ e $D^{(2)}(x) = b_0 + b_2x^2$, a solução estacionária da Eq. (2.33) é

$$P(x) = \frac{P_0}{[1 + \frac{b_2}{b_0}x^2]^{1+\frac{a_1}{b_2}}}, \quad (6.1)$$

que possui forma q -gaussiana e constante de normalização P_0 dada por $\sqrt{\frac{b_2}{\pi b_0} \frac{\Gamma[\frac{a_1}{b_2}+1]}{\Gamma[\frac{a_1}{b_2}+\frac{1}{2}]}}$. A condição

$$b_0 = a_1 - b_2 \quad (6.2)$$

surge do fato de que, uma vez que os dados estão normalizados, a variância $\sigma^2 = b_0/(a_1 - b_2)$, deve ser unitária.

Os valores dos parâmetros dos coeficientes de KM calculados com $\tau = 1$ produzem uma FDP estacionária, Eq. (6.1), em boa concordância com o histograma empírico, conforme mostrado na Fig. 6.2.

A Eq. (2.36) foi integrada numericamente através do método de Euler [63]. Para comparação, utilizamos também os métodos de Milstein e Runge-Kutta [63], e os resultados não foram significativamente diferentes. O passo de integração foi $dt = 10^{-3}$, e os dados foram gravados a cada $1/dt$ passos, consistentemente com nossa escolha de $\tau = 1$. Apresentamos a série empírica na Fig. 6.3(a) e a gerada numericamente na Fig. 6.3(b). O histograma da série artificial coincide, naturalmente, com a distribuição estacionária dada pela Eq. (6.1).

Além de obter os histogramas, em concordância com os valores teóricos, também monitoramos as correlações. Na Fig. 6.4, observamos que a função de

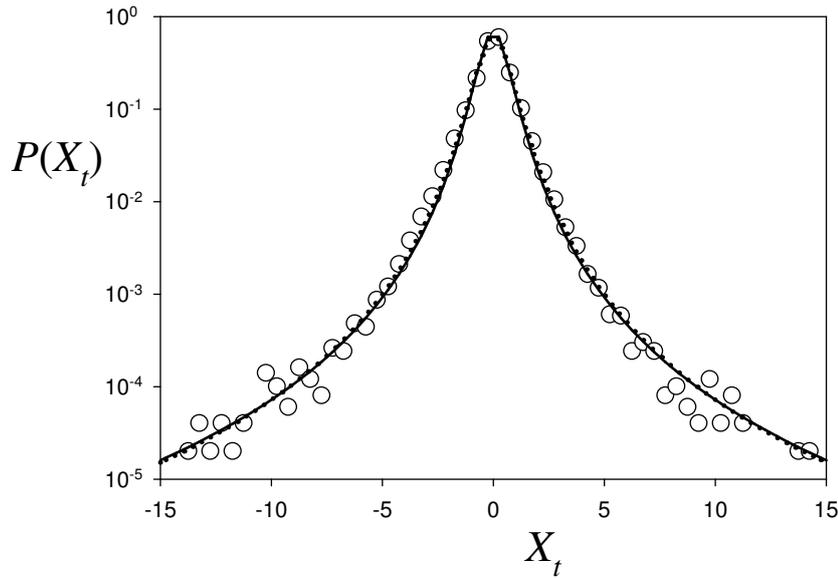


Figura 6.2: FDPs dos retornos normalizados. Símbolos correspondem aos histogramas empíricos. FDPs teóricas dadas pela Eq. (3.12) com $(\tilde{a}_1, \tilde{b}_2, \tilde{b}_0) = (0, 85; 0, 45; 0, 40)$ obtidas do cálculo dos coeficientes de tendência e difusão de acordo com as Eqs. (6.3) e (6.4) (linhas pontilhadas), e $(a_1, b_2, b_0) = (1, 9; 1, 10; 0, 81)$ obtidos da Eqs. (6.15) e (6.16) (linhas cheias).

autocorrelação da série obtida a partir dos coeficientes para τ -finito possuía um tempo de correlação maior que o da série empírica, apesar das FDPs serem muito semelhantes.

6.3 Correção dos coeficientes de KM

O resultado apresentado acima nos leva a algumas considerações. Em alguns casos, estimativas numéricas não podem ser sempre extrapoladas em direção ao limite $\tau \rightarrow 0$, devido à taxa limitada de aquisição dos dados reais. Dessa forma, como vimos, temos acesso apenas às estimações finitas em τ dos parâmetros $\tilde{D}^{(1)}$ e $\tilde{D}^{(2)}$, onde τ nem sempre é pequeno em comparação com os tempos característicos do processo. Assim, para algumas séries temporais reais, como as séries financeiras, os métodos de reconstrução precisam ser revisados.

O erro devido ao uso de τ -finito pode ser estimado da expansão da Eq. (2.36). Recentemente, Sura e Barsugli [64] obtiveram as correções de primeira ordem para os coeficientes dados pela Eq. (2.35). Para um dado incremento τ , as aproximações em diferenças finitas apresentadas na Eq. (2.35) podem ser reescritas como:

$$\tilde{D}^{(1)}(x, \tau) = \frac{1}{\tau} \langle X(t + \tau) - X(t) \rangle |_{X(t)=x}, \quad (6.3)$$

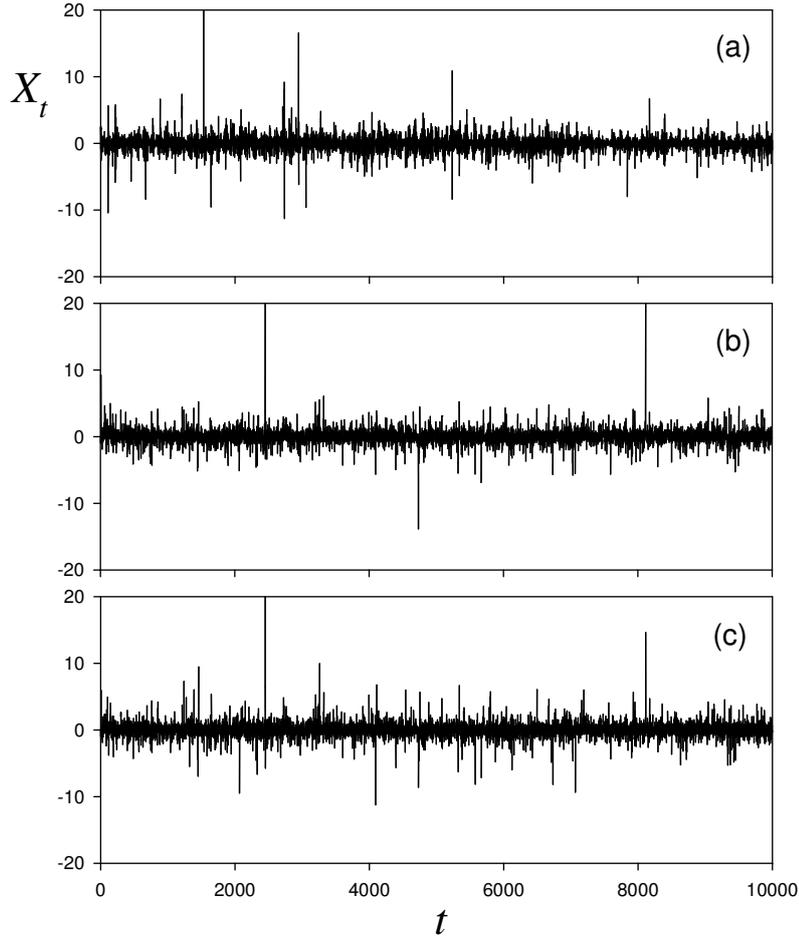


Figura 6.3: Séries temporais de retornos: empírica (a), artificiais geradas com parâmetros $(\tilde{a}_1, \tilde{b}_2, \tilde{b}_0) = (0, 85; 0, 45; 0, 40)$ (b) e $(a_1, b_2, b_0) = (1, 9; 1, 10; 0, 81)$ (c).

$$2\tilde{D}^{(2)}(x, \tau) = \frac{1}{\tau} \langle (X(t + \tau) - X(t))^2 \rangle |_{X(t)=x}. \quad (6.4)$$

Uma das formas para estimarmos as correções introduzidas quando se adota um incremento de tempo τ -finito nas Eqs. (6.3) e (6.4), é a expansão em série de Itô-Taylor [63] de $X(t + \tau)$. Para obtê-la, vamos considerar a Eq. (2.36) em sua forma integral:

$$X_{t+\tau} = X_t + \int_t^{t+\tau} D^{(1)}(X_s) ds + \int_t^{t+\tau} \sqrt{2D^{(2)}(X_s)} dW_s \quad (6.5)$$

Na Eq. (6.5), os valores de $D^{(1)}$ e $D^{(2)}$ em instantes intermediários são obtidos utilizando-se a fórmula de Itô [63] para funções arbitrárias da variável estocástica X_t (ver Apêndice A.1.3):

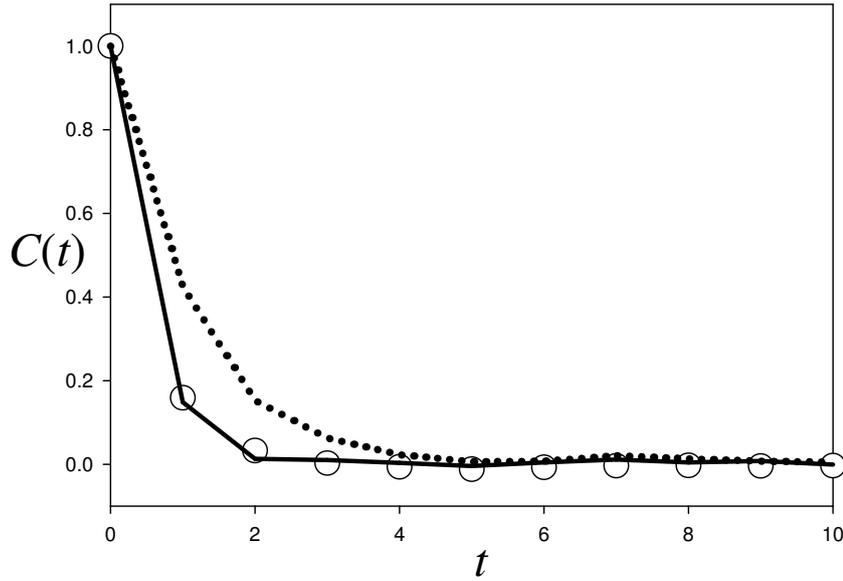


Figura 6.4: Função de autocorrelação linear $C(t)$ em função do tempo. Os símbolos correspondem aos dados empíricos, as linhas correspondem aos processos artificiais gerados com parâmetros $(\tilde{a}_1, \tilde{b}_2, \tilde{b}_0) = (0, 85; 0, 45; 0, 40)$ (linhas pontilhadas) e $(a_1, b_2, b_0) = (1, 9; 1, 10; 0, 81)$ (linhas cheias).

$$dF(X_t, t) = (\partial_t F + D^{(1)} \partial_{X_t} F + D^{(2)} \partial_{X_t X_t} F) dt + (\sqrt{2D^{(2)}} \partial_{X_t} F) dW \equiv L^0 F dt + L^1 F dW . \quad (6.6)$$

A primeira aplicação da fórmula de Itô à Eq. (6.5) fornece (ver Apêndice A.1.3):

$$X_{t+\tau} = X_t + \int_t^{t+\tau} \left(D^{(1)}(X_t) + \int_t^s L_0(D^{(1)}(X_{s'})) ds' + \int_t^s L_1(D^{(1)}(X_{s'})) dW_{s'} \right) ds + \int_t^{t+\tau} \left(\sqrt{2D^{(2)}}(X_t) + \int_t^s L_0 \sqrt{2D^{(2)}}(X_{s'}) ds' + \int_t^s L_1 \sqrt{2D^{(2)}}(X_{s'}) dW_{s'} \right) dW_s . \quad (6.7)$$

Após aplicações sucessivas, podem ser obtidas correções de ordem superior.

Considerando apenas os termos da expansão até a primeira ordem nas Eqs. (6.3) e (6.4), obtemos (ver Apêndice A.1.3):

$$\tilde{D}^{(1)}(x, \tau) = D^{(1)} + \left(\frac{D^{(1)}(D^{(1)})'}{2} + \frac{D^{(2)}(D^{(1)})''}{4} \right) \tau + \mathcal{O}(\tau^2) \quad (6.8)$$

e

$$\begin{aligned} \tilde{D}^{(2)}(x, \tau) = & D^{(2)} + \left[\frac{1}{2}(D^{(1)})^2 + D^{(2)}(D^{(1)})' + D^{(1)}\sqrt{D^{(2)}} \left(\sqrt{D^{(2)}}\right)' \right. \\ & \left. + \{D^{(2)} \left(\sqrt{D^{(2)}}\right)'^2 + \left(\sqrt{D^{(2)}}\right)^3 \left(\sqrt{D^{(2)}}\right)'' \} \right] \tau + \mathcal{O}(\tau^2). \end{aligned} \quad (6.9)$$

As Eqs. (6.8) e (6.9) nos fornecem uma alternativa direta para corrigir os erros devido ao τ -finito. Apesar disso, precisamos especificar as formas de $D^{(1)}$, $(D^{(1)})'$, $D^{(2)}$ e $(D^{(2)})'$. Outras maneiras de se determinar o erro em coeficientes de KM são referenciadas em [64].

Embora as correções de primeira ordem geralmente sejam suficientes, há situações onde esta correção não é apropriada. Neste caso, consideraremos uma correção exata desenvolvida recentemente para os dois primeiros coeficientes de KM calculados com τ -finitos [65]:

$$\tilde{a}_1 = a_1 \sum_{j \geq 0} \frac{[-a_1]^j}{(j+1)!} \tau^j, \quad (6.10)$$

$$\tilde{b}_0 = b_0 \sum_{j \geq 0} \frac{[-2(a_1 - b_2)]^j}{(j+1)!} \tau^j, \quad (6.11)$$

$$\tilde{b}_2 = b_2 \sum_{j \geq 0} \frac{\frac{1}{2}[2(a_1 - b_2)]^{j+1} - [a_1]^{j+1}}{(j+1)!} \tau^j. \quad (6.12)$$

Se a Eq. (6.2) é válida, obtemos a relação $\tilde{b}_0 = \tilde{a}_1 - \tilde{b}_2$, válida também para qualquer ordem de truncamento das séries acima.

Somando as séries infinitas e levando em conta a condição dada pela Eq. (6.2), obtém-se

$$\tilde{a}_1 = \frac{1 - \exp(-a_1\tau)}{\tau}, \quad (6.13)$$

$$\tilde{b}_0 = \frac{1 - \exp(-2b_0\tau)}{2\tau}. \quad (6.14)$$

As Eqs. (6.13) e (6.14) podem ser invertidas para se extrair os parâmetros exatos em função dos estimados:

$$a_1 = \frac{\ln(1 - \tilde{a}_1\tau)}{-\tau}, \quad (6.15)$$

$$b_0 = \frac{\ln(1 - 2\tilde{b}_0\tau)}{-2\tau}. \quad (6.16)$$

Também podemos extrair valores aproximados dos parâmetros

verdadeiros a partir das séries truncadas apresentadas nas Eqs. (6.10), (6.11) e (6.12). Chamando de $a_1^{(n)}$ o valor obtido pela inversão da Eq. (6.10) truncada em uma ordem n para um dado \tilde{a}_1 , podemos analisar a convergência desse parâmetro. Na Fig. 6.5, observamos que a convergência, dentro de uma margem de erro de 5%, depende do valor de \tilde{a}_1 . É mais rápida para valores pequenos de \tilde{a}_1 , ou seja, tão logo $1/\tilde{a}_1$ torna-se grande em comparação com a escala $\tau = 1$. Contudo, este não parece ser o caso dos dados empíricos das séries financeiras analisadas nesta tese: para $\tilde{a}_1 \approx 0,85$ a convergência é atingida somente na sexta ordem.

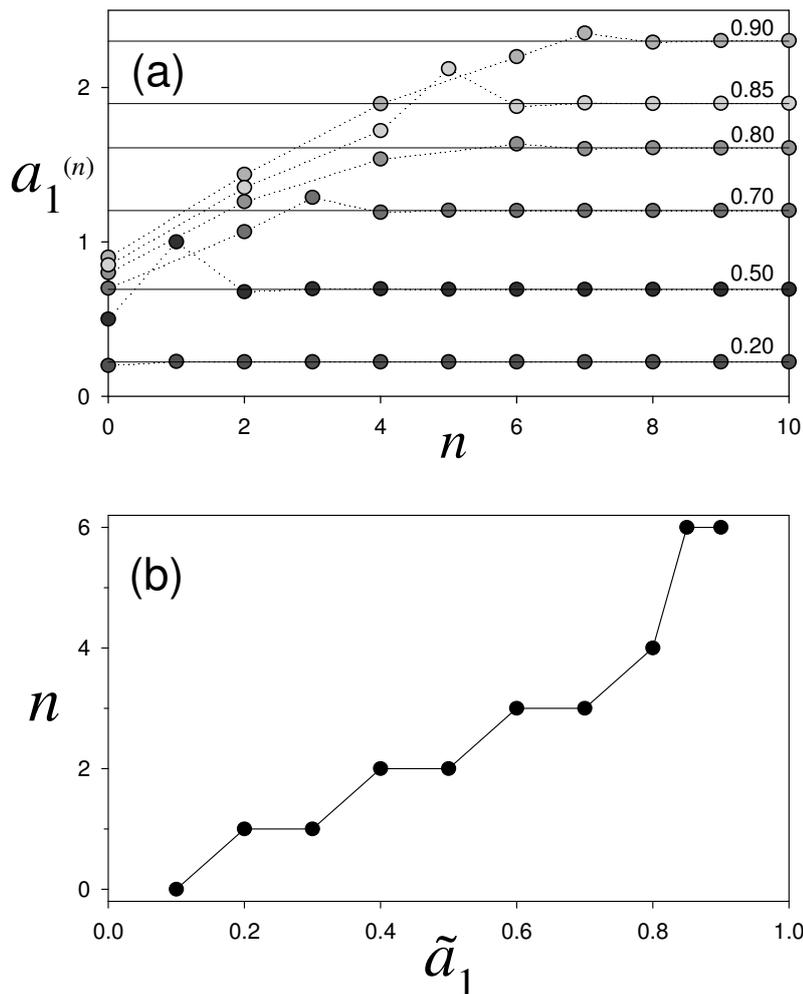


Figura 6.5: (a) Dependência de $a_1^{(n)}$ em função da ordem de aproximação n dada na Eq. (6.10), para diferentes valores de \tilde{a}_1 . (b) Ordem na qual o valor limite é atingido em função de \tilde{a}_1 , com erro de 5%.

A partir das Eqs. (6.15), (6.16) e (6.2), obtivemos os parâmetros exatos $(a_1, b_0 \text{ e } b_2) = (1, 90; 1, 10; 0, 81)$.

Também comparamos os resultados teóricos obtidos para $\tau = 1$ com a análise de séries temporais geradas numericamente usando as estimativas

teóricas de $D^{(1)}$ e $D^{(2)}$. A Eq. (2.36) também foi resolvida mediante o método de Euler, com o mesmo procedimento adotado para gerar a série artificial de retornos a partir de $\tilde{D}^{(1)}$ e $\tilde{D}^{(2)}$. Como podemos verificar na Fig. 6.3(c), a série artificial também possui boa concordância com a série empírica da Fig. 6.3(a).

Uma vez que $a_1/b_2 \approx \tilde{a}_1/\tilde{b}_2$, ambos os conjuntos de valores, os exatos e os para τ -finito produzem uma FDP teórica em muito boa aproximação com os histogramas empíricos, como ilustrado na Fig. 6.2.

A principal diferença está no cálculo das correlações. Na Fig. 6.4, observamos a boa aproximação entre a função de autocorrelação gerada a partir das séries artificiais com os parâmetros limites e com os empíricos, enquanto que os coeficientes obtidos com τ -finito geravam um tempo de correlação maior, apesar das FDPs serem muito semelhantes.

Finalmente, se o processo é estacionário e markoviano, verificado através da equação de CK e dos testes de estacionariedade PP e ADF, como mostrado nos capítulos anteriores, então a probabilidade conjunta $P(x_t, x_{t+\tau})$ contém toda a informação do processo estocástico. Na Fig. 6.6, apresentamos as curvas de nível dos FDPs conjuntas obtidas das séries temporais empíricas e artificiais para $\tau = 1$.

Vemos que o processo artificial, gerado levando-se em conta as correções às estimativas de τ -finito dos coeficientes de KM, fornece comportamento mais próximo aos dados empíricos, quando comparado ao processo artificial sem correção. Em particular, a probabilidade conjunta na Fig. 6.6(c) mostra uma dependência mais fraca entre retornos sucessivos de acordo com o obtido na Fig. 6.4.

6.4

Conclusões preliminares

Neste capítulo da tese, reconstruímos artificialmente uma série empírica de retornos intradiários. Além de reproduzir os histogramas empíricos, a série artificial possui níveis de correlação linear realistas. Para fazer esta reconstrução, vimos a necessidade de realizar correções de ordens superiores aos cálculos dos coeficientes de KM para τ -finito. Para tal, utilizamos uma expansão teórica de ordem infinita apresentada em [65]. O conteúdo deste capítulo pode também ser encontrado na Ref. [66].

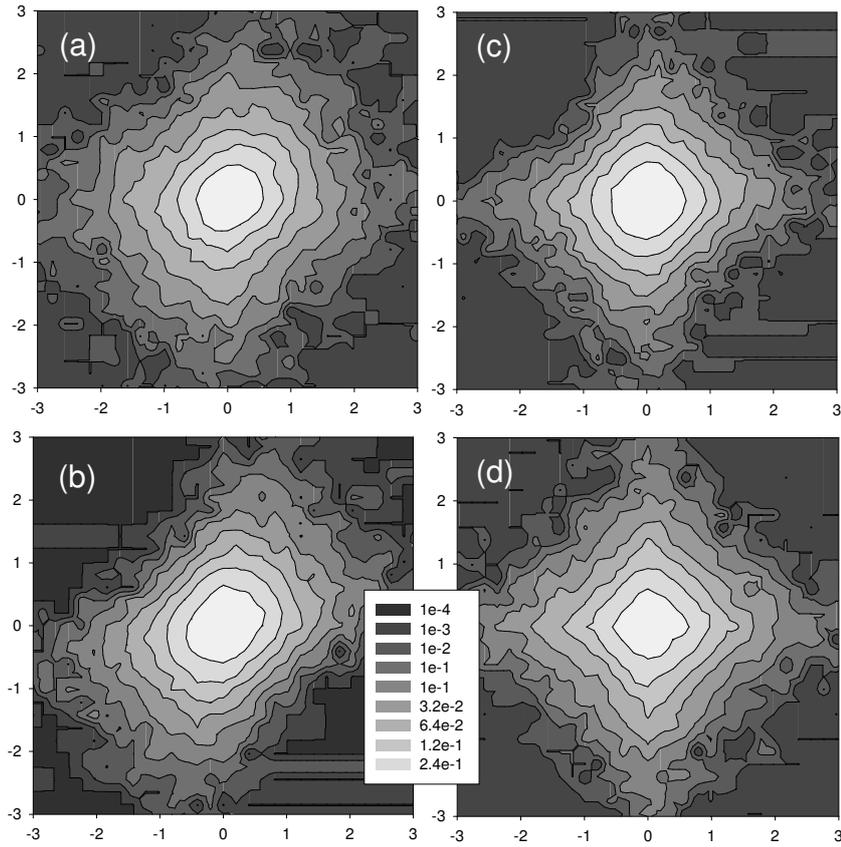


Figura 6.6: FDPs conjuntas para retornos consecutivos, construídos a partir de: (a) dados empíricos, (b) processo artificial gerado com parâmetros $(\tilde{a}_1, \tilde{b}_2, \tilde{b}_0) = (0, 85; 0, 45; 0, 40)$, (c) $(a_1, b_2, b_0) = (1, 9; 1, 10; 0, 81)$ e (d) dados empíricos misturados.