

## 5

### Análise do Volume de Ações Negociadas



#### 5.1

##### Introdução

Neste capítulo investigamos a estatística dos volumes negociados em mercados financeiros. Mostramos que processos estocásticos de volumes negociados podem ser entendidos tendo como base processos de Poisson misturados em uma escala mesoscópica.

#### 5.2

##### Estatística descritiva dos volumes negociados

Utilizamos como exemplo representativo o volume negociado intradiário da BOVESPA, de 3 de janeiro de 2005 a 13 de setembro de 2007. Os dados foram obtidos em intervalos de 30 minutos, consistindo de 9970 pontos. O tempo é contado apenas em dias de negócio, descartando finais de semana e feriados. A série original de volume negociado,  $V(t)$ , foi reescalada pelo valor

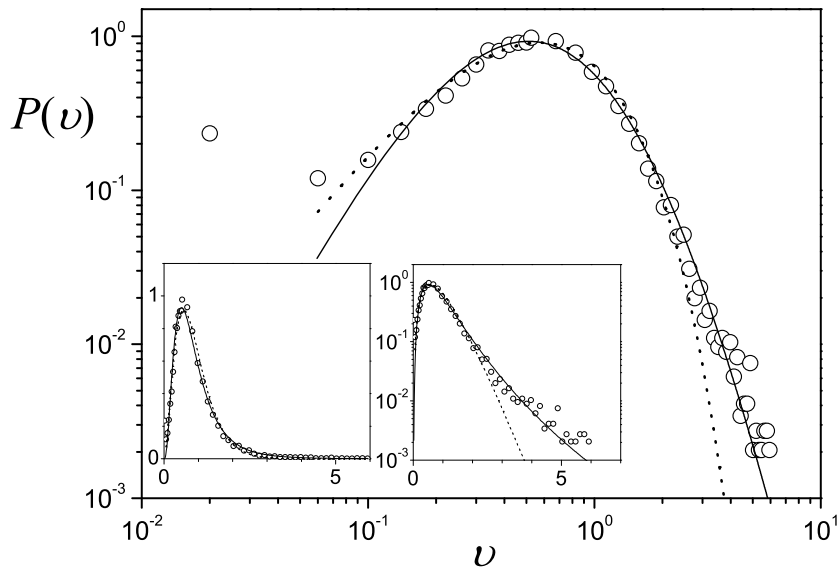
empírico médio  $\langle V \rangle$ . A série de  $v = V/\langle V \rangle$  está representada na Fig. 1.3, e mostra o padrão intermitente da atividade de negociação.

Verificamos que nossos histogramas de volumes negociados empíricos em diferentes escalas temporais de agregação  $\Delta t$ , podem ser bem aproximados pela FDP  $q$ -Gama dada na Eq. (2.15), como ilustrado na Fig. 5.1 (linhas cheias).

Os ajustes dos histogramas dos volumes de negociação usando a FDP Gama dada por

$$P(v) = \frac{\eta}{\Gamma(\beta)} (\eta v)^{\beta-1} \exp(-\eta v), \quad (5.1)$$

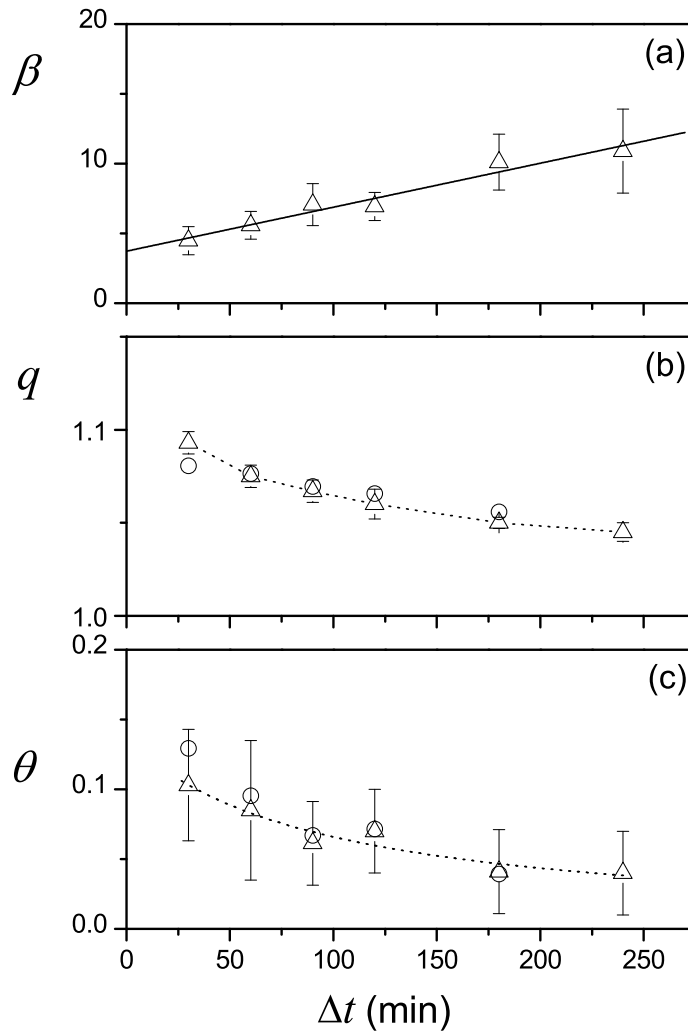
são bastante bons para valores intermediários, como ilustrado na Fig. 5.1 (linhas pontilhadas). Contudo, eles manifestam desvios para volumes pequenos e grandes.



**Figura 5.1:** Gráfico log-log da FDP dos volumes negociados a cada 30 minutos na BOVESPA no período entre 3/01/05 a 13/09/07. Os símbolos representam o histograma empírico e a linha sólida o ajuste por mínimos quadrados da Eq. (2.15) com parâmetros mostrados na Fig. 5.2. A linha pontilhada corresponde ao ajuste da Eq. (2.15) com  $q = 1$ , representado para comparação. Os gráficos inseridos correspondem às representações do gráfico principal em escalas linear-linear e log-linear.

O ajuste da  $q$ -Gama apresenta um desvio notável somente para volumes muito pequenos (ver também Refs. [12, 15, 16]). O excesso na frequência empírica pode ser devido a um patamar mínimo para transações nos mercados, cujas estatísticas não fazem parte do escopo da tese. Todavia, há uma concordância muito boa para toda a região de ajuste englobando a maioria dos volumes significantes. Um ajuste de mínimos quadrados na Eq. (2.15) foi realizado ao longo da região dos dados da amostra, descartando a região de menor volume (tipicamente abaixo de  $v \simeq 10^{-1}$ ). Neste ajuste, foi considerado

o peso estatístico  $1/P(v)$  com o objetivo de ponderar volumes extremos. Os valores resultantes do ajuste dos parâmetros são apresentados na Fig. 5.2 e na tabela 5.1. De forma consistente com um volume reescalonado médio unitário, enquanto os três parâmetros variam, a média da distribuição  $q$ -Gama,  $\beta\theta/((1-q)\beta+2-q)$ , permanece próxima a 1 (com a discrepância de no máximo 6%).



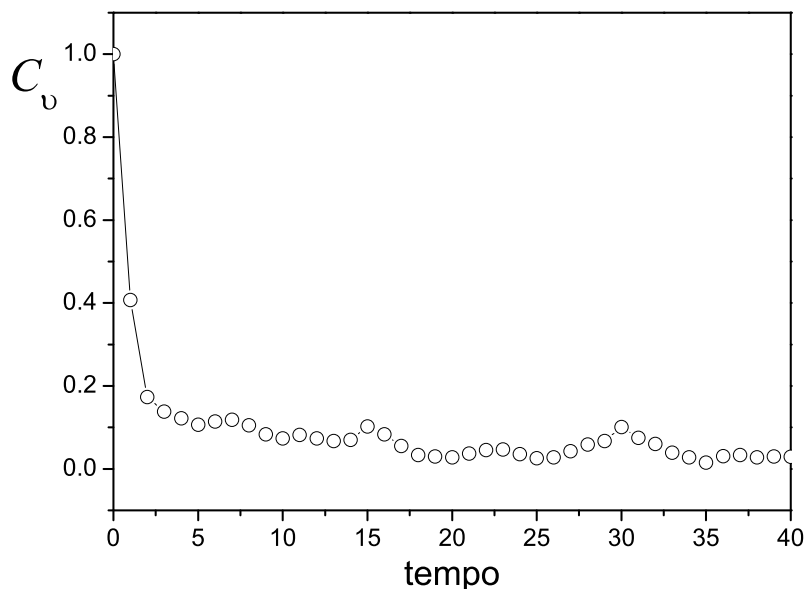
**Figura 5.2:** Parâmetros de ajuste da distribuição  $q$ -Gama,  $\beta$  (a),  $q$  (b) e  $\theta$  (c), para a FDP dos volumes normalizados entre 3/01/05 e 13/09/07, em função de  $\Delta t$  (triângulos). Os círculos correspondem aos valores teóricos obtidos das Eqs. (5.3). As linhas pontilhadas são um guia visual. As linhas cheias em (a) correspondem a um ajuste linear.

Para esclarecer o surgimento das FDPs  $q$ -Gama neste cenário, é importante verificar a independência de sucessivos volumes. Dessa forma, calculamos as funções de autocorrelação de primeira e segunda ordem, e não encontramos níveis importantes de correlação para nenhuma janela de tamanho maior do que 30 minutos. Na Fig. 5.3 apresentamos o função de autocorrelação de primeira ordem do volume  $C_v$  em função do intervalo de tempo . Ela

Escala	$\beta$	$\theta$	$q$	$10^{-9}\langle V \rangle$
30min	4,48	0,103	1,093	1,61
60min	5,58	0,085	1,075	3,21
90min	7,06	0,061	1,067	4,81
120min	6,93	0,070	1,060	6,42
180min	10,1	0,041	1,050	9,63
240min	10,9	0,040	1,045	12,8

**Tabela 5.1:** Parâmetros de ajuste da função  $q$ -Gama obtidos para as FDPs dos volumes de negociação em diferentes escalas de agregação para a BOVESPA no período entre 3.jan.05 e 13.set.2007. Incluímos também os valores médios  $\langle V \rangle$  das séries originais. A  $P(v)$  ajustada foi calculada com peso estatístico e descartando volumes ( $v \leq 10^{-1}$ ).

mostra que os efeitos de memória em sucessivos ticks de volumes neste nível de agregação de dados é fraco.



**Figura 5.3:** Função de autocorrelação para os dados de volume de negociação analisados nesta tese.

### 5.3 Modelagem

Nesta seção, modelamos o processo que gera as distribuições empíricas dos volumes negociados. Nossa argumentação se dará em termos de processos de ponto ou de chegada.

Em um determinado nível de agregação de dados, cada intervalo de largura  $\Delta t$  contém o volume total negociado naquele intervalo. Em  $\Delta t$ , ocorre um número de negociações, cada uma envolvendo um dado volume de ações.

Uma vez que nossa base de dados não fornece nenhuma informação sobre os volumes individuais em cada negociação, pensamos, alternativamente, na chegada de intervalos através do fluxo de volumes cumulativos. O intervalo de volume acumulado entre sucessivos ticks é  $\Delta v_c = v$ . Então, deduzimos a estatística de  $v$  em analogia àquela dos intervalos de tempo (ou tempos de espera) em processos de chegada usuais.

Nesse cenário, a estatística do processo de ponto se constitui, em essência, em um processo de chegada poissoniano. De fato, é uma aproximação válida considerar:

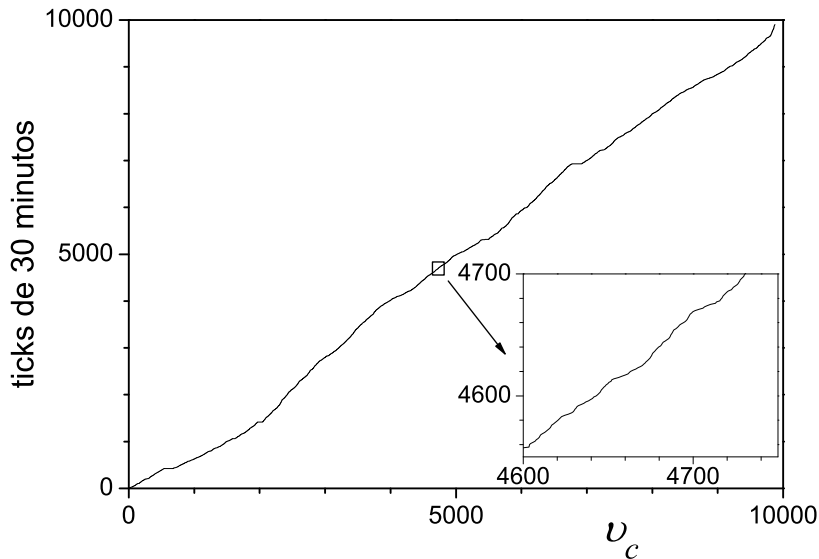
- (i) o número de eventos (ticks) em incrementos de volumes sem sobreposição como independentes e estacionários (nas escalas de tempo intradiárias analisadas);
- (ii) que para  $\delta v_c$  suficientemente pequeno, a probabilidade de acontecimento de um tick é proporcional a  $\delta v_c$ ;
- (iii) que a probabilidade de mais de um tick chegar em  $\delta v_c$  é desprezível.

Nesse caso, se a taxa de ocorrência (número médio de ticks por volume acumulado) é  $\eta$ , então a FDP dos intervalos  $v$  seguirá uma lei exponencial simples  $P(v) = \eta \exp(-\eta v)$ , se assemelhando aos tempos de espera entre ocorrências sucessivas em processos de chegada poissonianos [21]. Contudo, as transações no mercado são atividades complexas envolvendo muitos passos para tomar uma decisão no nível microscópico de eventos individuais, e os volumes em qualquer escala maior são construídos de volumes na escala desses eventos. Então, é razoável assumir que, ao invés de ser um processo poissoniano simples, o processo poissoniano dos volumes é descrito por  $\beta$  estágios ocultos, cada um associado a uma estatística exponencial simples, com taxas similares. Como vimos no Capítulo 2 (2.38), para estágios independentes, ocorrendo a uma taxa  $\eta$ , a distribuição dos volumes será dada pela Eq. (5.1), também conhecida como a distribuição Gama,  $\Gamma_{\beta,1/\eta}$ .

Nesse caso, o parâmetro de forma  $\beta$  pode ser interpretado como um número efetivo de estágios do processo combinado. Lembramos também que a Eq. (5.1) é equivalente à Eq. (2.15) quando  $q \rightarrow 1$ .

Entretanto, o comportamento ligeiramente não exponencial na cauda pode ter uma origem mais fundamental. A análise das correlações indica que os desvios não são devidos à presença de fortes correlações. Então, passamos a investigar outras fontes das caudas longas observadas. Por um lado, a negociação não é homogênea. Por outro, como vimos no Capítulo 2, foi mostrado por Wilk e Wlodarczyk [30] que as distribuições com cauda em lei de potência emergem de forma natural de flutuações lentas de um parâmetro

relacionado ao inverso de uma medida de temperatura aparecendo em uma distribuição originalmente exponencial. Assim, observamos cuidadosamente o possível caráter estocástico da taxa  $\eta$  na Eq. (5.1), que quantifica o número médio de eventos (ticks) por volume acumulado  $v_c$  de cada estágio. Para tal, construímos um gráfico do número de ticks em função do  $v_c$ , que é apresentado na Fig. 5.4.

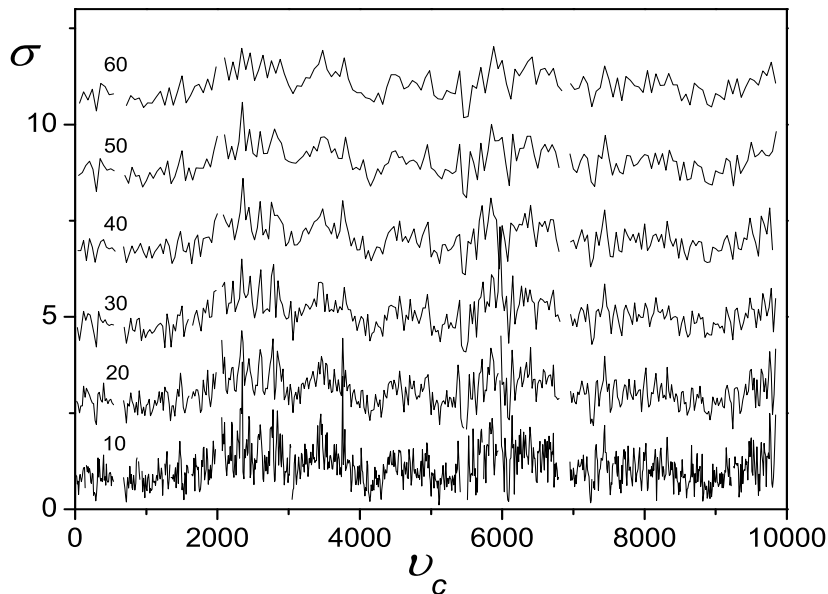


**Figura 5.4:** Número de ticks de 30 minutos em função do volume acumulado  $v_c$  na BOVESPA durante o período entre 3/01/05 e 13/09/07. Gráfico inserido: ampliação da curva.

Pode-se observar que o declive local efetivo, o qual chamaremos de  $\sigma$ , flutua em diferentes escalas, como mostrado na Fig. 5.4. Podemos notar que  $\sigma^{-1}$  é aproximadamente uma medida local de  $\langle v \rangle$ , que, de acordo com a Eq. (5.1), satisfaz  $\langle v \rangle = \beta \eta^{-1}$ . A taxa local é então  $\eta = \beta \sigma$  correspondendo à taxa de cada estágio (escondido) do processo. De fato, a natureza composta do processo de volume implica que a taxa efetiva observada é diminuída quando comparada com a taxa real de estágio simples.

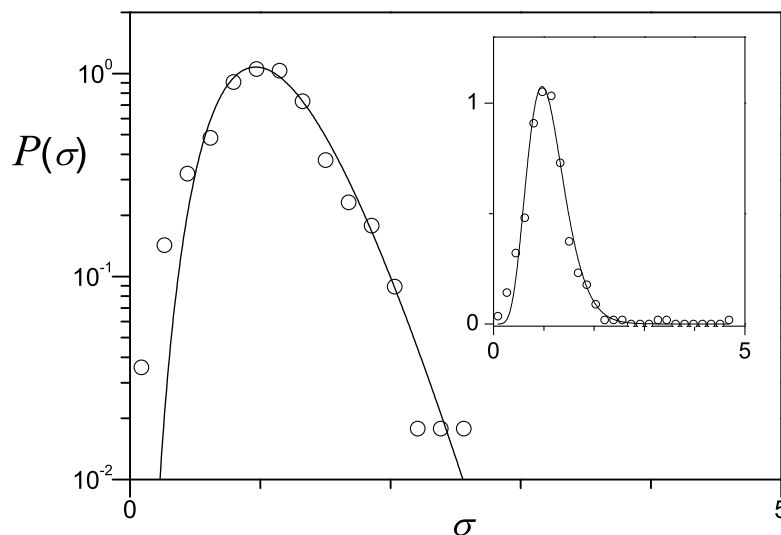
Com o objetivo de quantificar as flutuações da taxa, os declives locais da Fig. 5.4 foram calculados para janelas de igual comprimento  $\Delta v_c$ . Pequenas janelas são muito ruidosas enquanto janelas mais largas escondem a estrutura das flutuações. Portanto, consideramos janelas de tamanho intermediário correspondentes tipicamente a intervalos de 30 ticks (cerca de 2 dias de negociação). Na Fig. 5.5 apresentamos a série temporal de  $\sigma$  em função do número de pontos por janela.

O histograma de declives locais é apresentado na Fig. 5.6. Esta figura mostra que a estatística dos declives  $\sigma$  (logo, aquela das taxas locais  $\eta$ ) segue, em boa aproximação, uma distribuição Gama. O teste de ajuste de



**Figura 5.5:** Série temporal de  $\sigma$  em função de  $\nu_c$  para diferentes números de pontos por janela  $\Delta\nu_c$  (indicado na figura).

Kolmogorov-Smirnoff (teste KS) [37], mostra como resultado uma adequação forte à hipótese da distribuição Gama para  $\sigma$ . De forma contrária, outras distribuições padrões de dois parâmetros foram rejeitadas, por exemplo a Gama inversa, com valor de referência (p-valor) menor que 0,035.



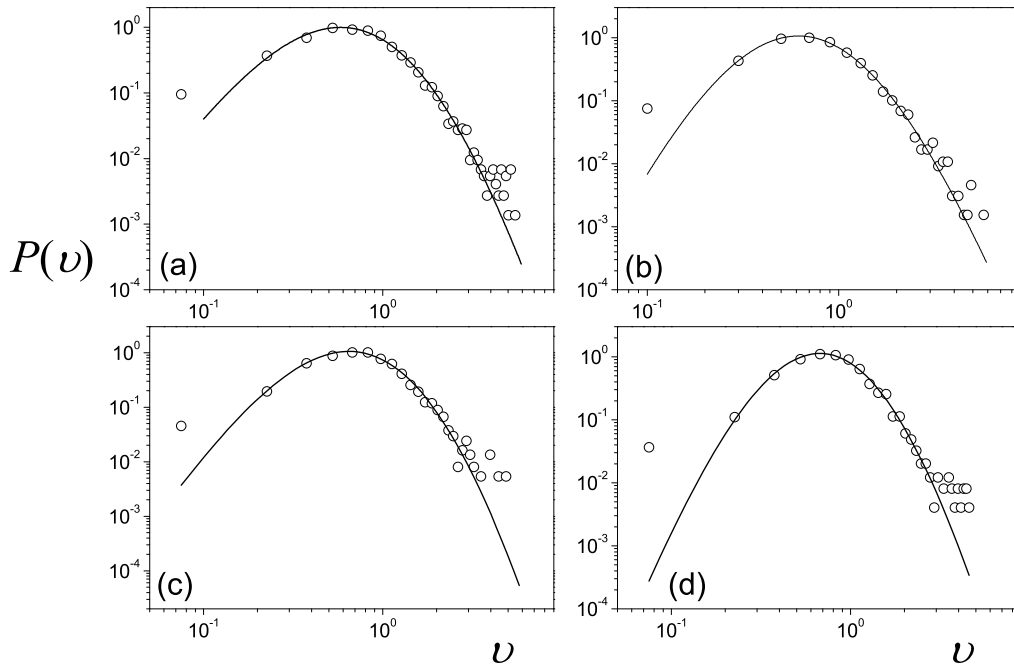
**Figura 5.6:** Histograma dos declives locais do gráfico mostrado na Fig. 5.4, calculados em janelas de 30 ticks (símbolos). Linhas sólidas correspondem ao melhor ajuste de mínimos quadrados de uma FDP Gama. Gráfico inserido: o mesmo gráfico em escala linear.

Os parâmetros da FDP  $\Gamma_{\alpha,1/\kappa}$  associada aos histogramas de  $\eta$  foram

obtidos do ajuste da FDP  $\Gamma_{\alpha,1/\beta\kappa}$  ao histograma de  $\sigma = \eta/\beta$ . Valores para diferentes escalas de agregação analisadas podem ser encontrados na tabela 5.2. Espera-se que a distribuição dos declives dependa somente fracamente do intervalo de ticks  $\Delta t$ . De fato, os valores  $\alpha$  e  $\kappa\beta$  não são afetados significativamente pela escolha de  $\Delta t$ . Além disso, os resultados empíricos estão de acordo com a média teórica dos declives,  $\langle\sigma\rangle = \alpha/(\kappa\beta)$ , que é unitária por construção ( $\langle\sigma\rangle = \langle v\rangle \equiv 1$ ).

A partir de séries de volumes de 30 minutos, podemos gerar novas séries em diferentes escalas, agregando os dados, obtendo assim, séries de volumes acumulados a cada 60 minutos, 120 minutos, etc.

Observamos que nas diferentes escalas de agregação a distribuição dos volumes segue a forma  $q$ -Gama. Apresentamos a seguir, na Fig. 5.7, as  $q$ -Gamas obtidas para as escalas utilizadas na tese e apresentadas na tabela 5.1.



**Figura 5.7:** Ajustes da função  $q$ -Gama obtidos para as FDPs dos volumes de negociação para diferentes escalas para a BOVESPA no período de 3/01/05 a 13/09/07: 60min(a), 90min(b), 120min(c), 180min(d). A  $q$ -Gama ajustada foi calculada com peso estatístico e desconsiderando volumes  $\leq 10^{-1}$ .

O fato da taxa de ocorrência  $\eta$  não ser uma constante, e sim uma quantidade flutuante, pode explicar o comportamento das caudas das distribuições dos volumes negociados. Se  $\eta$  flutua com uma escala característica maior que o processo estocástico subjacente, podemos interpretar a PF  $\Gamma_{\beta,1/\eta}$  em (5.1) como a probabilidade condicional  $f_{v|\eta}$  da variável  $v$  dado  $\eta$ , onde  $\eta$  é uma variável estocástica. Além disso, vamos considerar que  $\eta$  possui



Escala	$\alpha$	$\kappa$	$1/\beta\kappa$	$q_{th}$	$\theta_{th}$	$\theta_{th}/\theta$	$q_{th}/q$
30min	7,93	1,605	0,139	1,081	0,129	1,25	0,99
60min	7,50	1,248	0,144	1,076	0,095	1,12	1,00
90min	7,33	0,966	0,147	1,069	0,067	1,09	1,00
120min	8,30	1,090	0,132	1,066	0,072	1,02	1,00
180min	7,78	0,704	0,141	1,056	0,039	0,96	1,00

**Tabela 5.2:** Parâmetros obtidos para o ajuste da distribuição Gama para os histogramas das taxas em diferentes escalas para a BOVESPA no período de 3/01/05 a 13/09/07. A  $P(\sigma)$  foi ajustada utilizando toda a região dos dados. Para comparação, incluímos os valores teóricos dos parâmetros da FDP  $q$ -Gama, obtidos a partir da Eq. (5.3).

uma distribuição Gama, i.e.,  $f_\eta(x) = \Gamma_{\alpha,\kappa}(x)$ . Embora, essa seja apenas uma estimativa da distribuição das taxas locais, ela leva em conta suas características principais, com pequenos desvios nas caudas, como observado na Fig. 5.6. Com esses pressupostos, a PF marginal dos volumes passa a ter a forma:

$$P(v) = \int dx f_\eta(x) f_{v|\gamma}(v, x) = \mathcal{N} v^{\beta-1} e_q^{-\eta v}, \quad (5.2)$$

onde  $\mathcal{N}$  é um fator de normalização e a função  $q$ -exponencial ( $e_q$ ) foi definida no Capítulo 2.

Assim, vemos que para a variável aleatória  $\eta$ , regida pela distribuição Gama de parâmetros  $\alpha$  e  $\kappa$ , a FDP incondicional (ou misturada) dos volumes se torna uma  $q$ -Gama.

Os valores teóricos dos parâmetros da FDP  $q$ -Gama resultante generalizada (2.15) são:

$$\beta_{th} = \beta, \quad q_{th} = 1 + \frac{1}{\alpha + \beta} \quad \text{e} \quad \theta_{th} = \frac{\kappa}{\alpha + \beta}. \quad (5.3)$$

Note-se que, enquanto  $\beta$  controla o aumento da FDP para volumes menores,  $\alpha$  controla a cauda: lei de potência com expoente  $\alpha + 1$  para  $\alpha$  finito, e lei exponencial para  $\alpha \rightarrow \infty$ .

Obtivemos uma boa concordância entre os valores  $q_{th}$  e  $\theta_{th}$  e os resultados dos ajustes da distribuição empírica dos volumes, nas diferentes escalas de agregação. Esta comparação é apresentada na tabela 5.2.

Em resumo, nossa análise leva à FDP  $q$ -Gama que descreve, em muito boa aproximação os histogramas empíricos (veja Fig. 5.1), em todas as escalas de agregação.

Na Fig. 5.2(a), pode-se observar que  $\beta$  aumenta linearmente com  $\Delta t$  a partir de um valor  $\beta_0 > 1$ , nas diferentes escalas de agregação. Isso implica a existência de um número invariante efetivo de estágios  $\beta_0$  associado com o

nível de tempo mais fundamental da descrição das observáveis financeiras. A linearidade caracteriza as propriedades de agregação do processo do volume. Na Fig. 5.2(b) pode-se observar que  $q$  decai enquanto  $\Delta t$  aumenta. Isso implica que inomogeneidades se reduzem para grandes intervalos de acumulação  $\Delta t$ . É notável que, para as 10 ações de maior volume na Nasdaq e NYSE para intervalos de 1, 2 e 3 minutos, são observados [12, 15] comportamentos similares.

Das Eqs. (5.3),  $q$  é determinado de uma maneira única a partir do número efetivo  $\alpha + \beta$  de estágios ou graus de liberdade independentes que especificam ambas as distribuições Gama. Então, por um lado,  $q$  reflete o grau de granularidade na descrição do processo de negociação (através de  $\beta$ ). Por outro, isso também reflete a presença de inomogeneidades (através de  $\alpha$ ), decorrentes das flutuações no ambiente do mercado, tais como o fluxo não estacionário de notícias e rumores. Ambos os efeitos se dissipam para grandes intervalos de acumulação  $\Delta t$ .

Investigamos também uma outra grandeza observável financeira, a série da quantidade de transações realizadas, ou ações negociadas, para diferentes escalas de agregação. Ao contrário da FDP dos volumes, a FDP obtida segue uma distribuição Gama, e não foram observadas flutuações no gráfico da taxa local do gráfico do número de ticks em função da grandeza acumulada. Apresentamos este resultado no Apêndice A.4.

## 5.4

### Considerações finais

Em resumo, investigamos um banco de dados real dos volumes negociados em uma escala temporal mesoscópica onde a natureza granular das transações e propriedades de agregação estão presentes. Além de fornecer novas perspectivas para a compreensão da dinâmica dos volumes, o estudo presente provê uma demonstração consistente com dados reais sobre como as FDPs com leis de potência emergem de efeitos de flutuação observáveis, de acordo com a visão superestatística.

O conteúdo deste capítulo pode também ser encontrado na Ref. [62].