

4

Resultados

Neste capítulo, vamos analisar o comportamento dos retornos de preços intradiários do IBOVESPA e obter modelagem para as distribuições baseadas em distribuições q-Gaussianas.

Lembramos que utilizamos apenas o tempo de pregão, ou seja, o período em que o mercado está aberto e funcionando; que é de aproximadamente 6:30h da sua abertura ao fechamento. O período sem pregão tal como finais de semana, feriados e períodos de negociação suspensa por motivo extraordinário, não são levados em conta.

Para analisar as séries de retorno, construir histogramas e fazer a análise estatística dos dados empíricos, utilizamos os softwares Excell for Windows® e Origin® 7.0.

As curvas de ajuste das distribuições empíricas do IBOVESPA apresentadas neste capítulo são as curvas ótimas, ou seja, as que minimizam a soma dos erros quadráticos. Para cálculo destes erros, utilizamos a ferramenta Chi-quadrado do Origin® 7.0.

4.1

Dados Empíricos de Alta Freqüência do IBOVESPA

Os dados obtidos diretamente da Bovespa para o período estudado (novembro de 2002 a Junho de 2004) consistem de 352.489 valores do índice de preços gravados em intervalos de 30 segundos, apresentados na figura 1.1. A partir destes dados construímos a série temporal de retornos de preços, segundo a definição (1.5), para várias janelas temporais Δt em escala intradiária. Assim, nossa análise de retornos intradiários utiliza da ordem de 10^5 dados, o que permite uma boa análise estatística.

Na figura 4.1 apresentamos a série histórica de retornos do IBOVESPA para intervalos consecutivos $\Delta t = 1$ minuto.

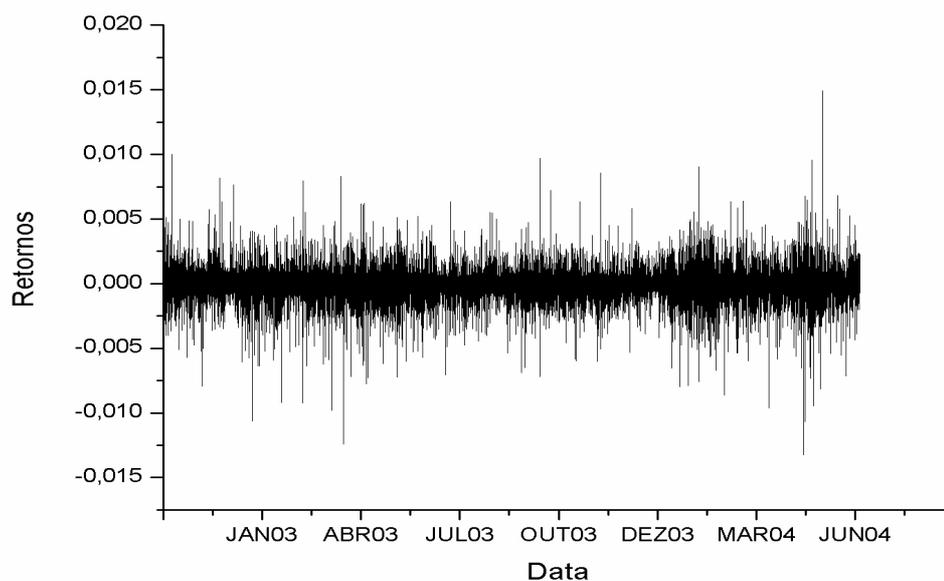


Figura 4.1 – Retornos para $\Delta t=1$ minuto do IBOVESPA

Esta série possui valor médio $\mu = 0,04 \times 10^{-4}$ e desvio padrão $\sigma = 5,7 \times 10^{-4}$, o que dá medida do tamanho típico das flutuações de retorno mostradas na figura 4.1. Como era de se esperar, os retornos são extremamente pequenos devido ao curto intervalo de tempo entre as medidas. Verificamos no entanto, que o valor médio é muito menor que o desvio padrão, mostrando que mesmo nessa escala de tempo tão pequena, devemos levar em conta as flutuações de valores em torno do valor médio. Como comparação, para dados de retorno diários, obtém-se $\mu = 7,0 \times 10^{-4}$, que é duas ordens de grandeza maior do que para escala de $\Delta t=1$ minuto. Como 1 dia de pregão possui aproximadamente 400 minutos, esperamos que a média diária seja aproximadamente 400 vezes maior do que a média de $\Delta t=1$ minuto e que esta relação valha para a escala anual também. Na escala diária, temos como média esperada $\mu \cong 16,0 \times 10^{-4}$, resultado que apresenta a mesma ordem de grandeza do valor observado empiricamente. Na escala anual esperamos uma taxa de retorno médio 252 vezes maior¹ do que a taxa de retorno diário, encontrando $\mu = 0,175$ ou uma taxa de 17,5% ao ano. A inflação anual real é da ordem de 10% [36]. Podemos atribuir a diferença entre essa taxa e a inflação devido ao fato de que no cálculo de índice de preços da bolsa de valores são contabilizadas as ações com maior volume de negócios, que valorizam-se em média, mais do que a taxa de crescimento do mercado.

¹ Um ano na bolsa de valores possui aproximadamente 252 dias de pregão.

Obviamente a operação de normalização não modifica o aspecto da série de retornos como podemos verificar pela figura 4.2. A nova série de retornos normalizados possui pela definição (1.8) desvio padrão unitário e média zero. A operação de normalização é necessária para compararmos séries de retornos em diferentes escalas de tempo pois elas terão a mesma escala de valores.

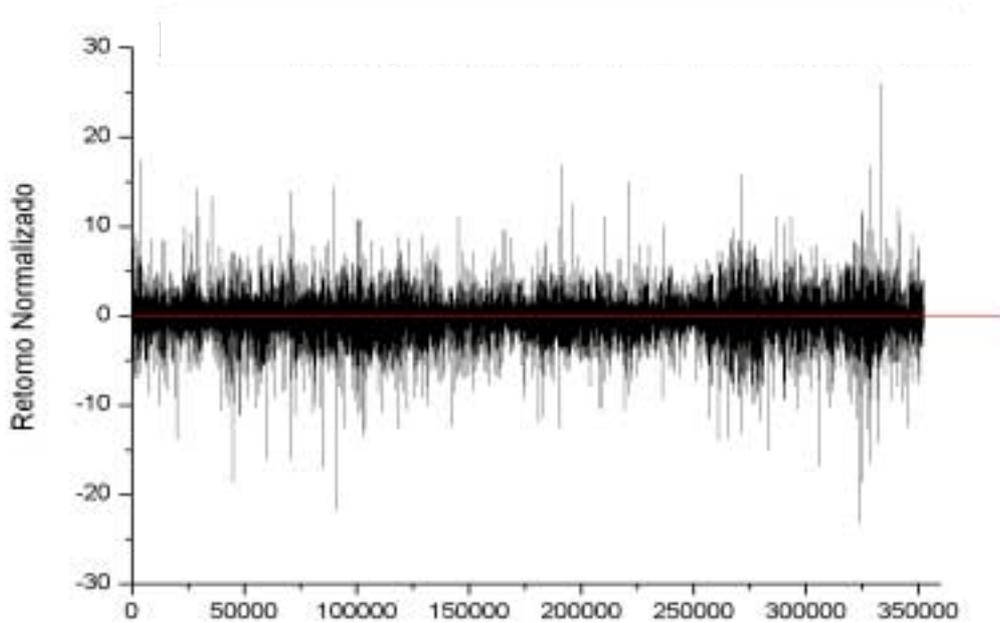
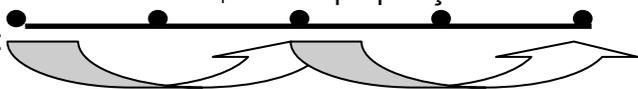
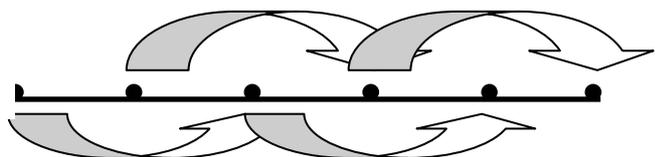


Figura 4.2 – Retorno Normalizado para $\Delta t=1$ minuto durante o período analisado

A partir dos dados disponíveis para o índice de preços a cada 30 segundos, podemos calcular retornos de preços em escala de tempo $\Delta t = 1$ minuto de duas formas: considerando períodos consecutivos Δt_i sem superposição dos intervalos temporais (esquemáticamente: ) ,

tal como utilizado na figura 4.2, ou com superposição dos intervalos Δt_i ,

(esquemáticamente: ) .

Este último caso equivale a incluir retornos considerando-se outra origem temporal dos dados. A figura 4.3 apresenta o histograma da freqüência de valores de retornos normalizados obtidos a partir das séries geradas com ou sem

superposição das janelas temporais. Verifica-se que, devido à enorme quantidade de dados não há mudança significativa de resultados.

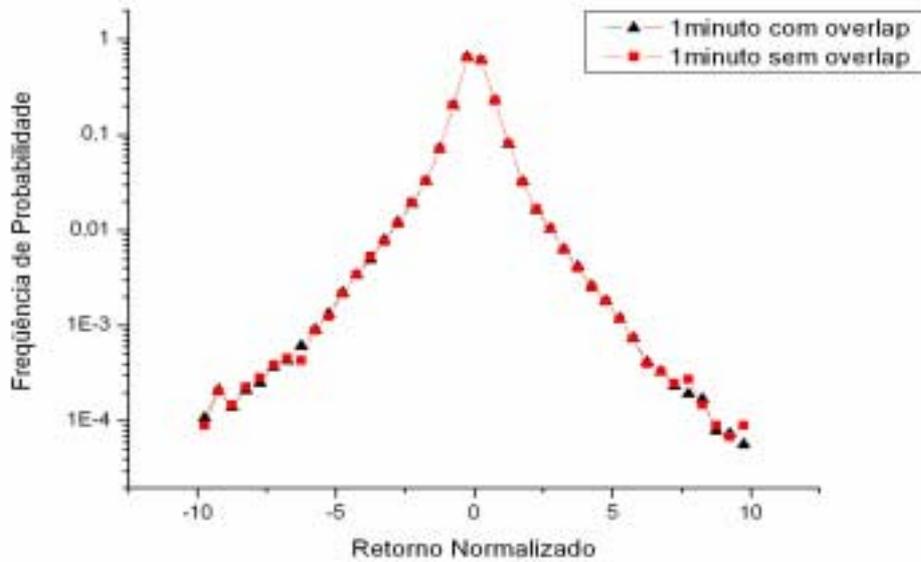


Figura 4.3 – Comparação dos histogramas de retorno para $\Delta t=1$ minuto do IBOVESPA em escala semi-logarítmica, considerando-se séries geradas com ou sem superposição dos intervalos temporais de medida.

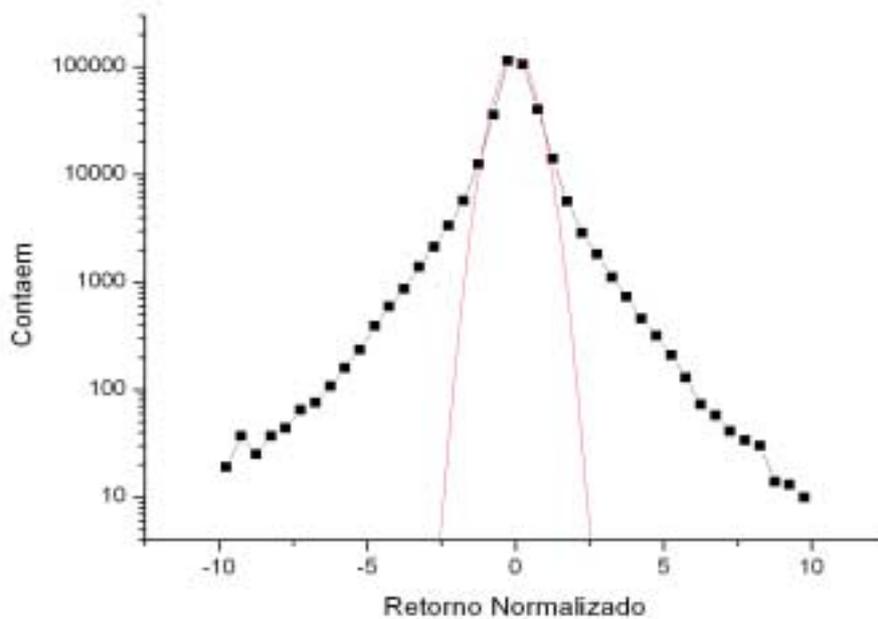


Figura 4.4 – Comparação da distribuição de retornos para $\Delta t=1$ minuto do IBOVESPA (pontos) com a distribuição Gaussiana (linha cheia)

Analisando a distribuição de retornos de $\Delta t=1$ minuto em escala semi-log, mostrado na figura 4.4, verificamos algumas características. Esta distribuição é aproximadamente simétrica, não Gaussiana e acentuadamente leptocúrtica.

Em nossa análise, consideramos as flutuações de retorno para várias escalas de tempo Δt intradiárias. Como um método de verificação da acuracidade dos valores de retorno obtidos, construímos as séries intradiárias nas várias escalas temporais de duas maneiras. Na primeira, para todas as janelas Δt , os retornos foram calculados diretamente utilizando-se a equação (1.5) com intervalos de tempo consecutivos, isto é, sem superposição. Na segunda, a partir da série de retornos de $\Delta t=1$ minuto construída com intervalos sem superposição, as séries para as demais janelas temporais Δt foram construídas somando-se retornos sucessivos de $\Delta t=1$ minuto, usando a propriedade (2.14). Como era de se esperar, não houve mudanças significativas na distribuição gerada, comprovado pela figura 4.5. Os histogramas construídos têm o mesmo número e tamanho de bins² e o número total de retornos (contagem total) é o mesmo nos dois casos.

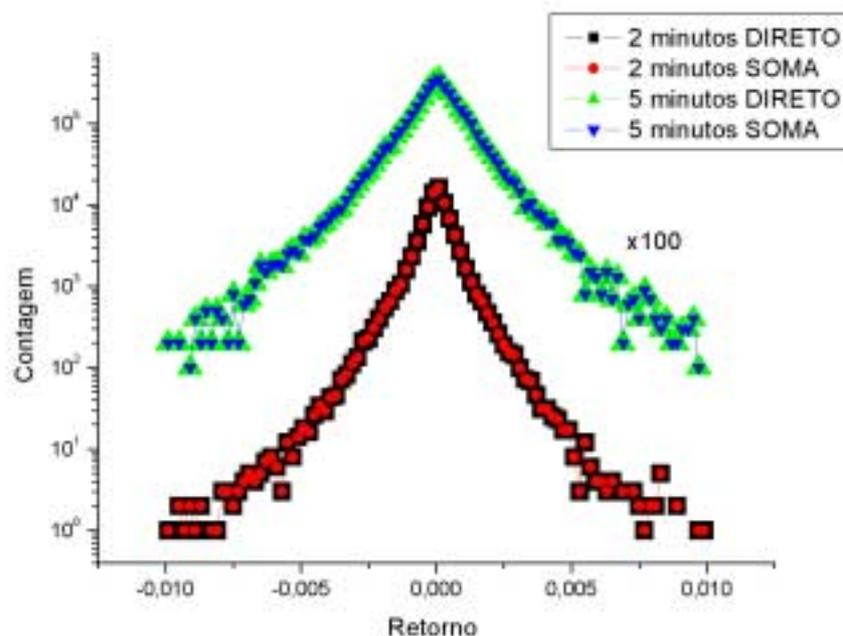


Figura 4.5 – Histogramas de retorno para $\Delta t= 2$ e 5 minutos construídos com dois métodos: cálculo direto (DIRETO), utilizando a equação (1.5) ; e através da soma de retornos de $\Delta t=1$ minuto (SOMA). Os gráficos para $\Delta t=5$ minutos estão multiplicados por um fator de 100 para melhor visualização.

² Intervalos de classe de mesmo tamanho originados da divisão do intervalo total de dados. Em cada um desses intervalos (bins), soma-se a frequência de dados ocorridos.

Uma vez encontradas as séries de forma robusta, passamos a analisar suas propriedades. Considere a figura 4.6 na qual apresentamos a série de retornos do IBOVESPA para três janelas temporais diferentes. Verificamos que a magnitude média desses retornos aumenta conforme aumentamos o intervalo de tempo de medida, como era de se esperar, pois as variações de preços são em geral maiores.

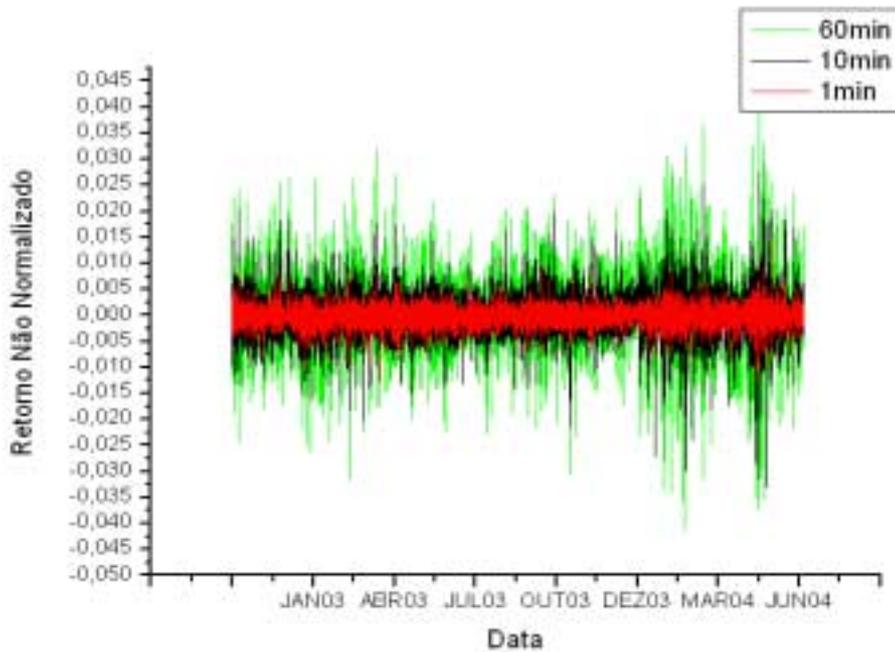


Figura 4.6 – Série de retornos intradiários do IBOVESPA no período analisado

Apresentamos na figura 4.7 o histograma sobreposto de dados de retorno para escalas de tempo de $\Delta t=1,10$ e 60 minutos. De acordo com o observado na série temporal da figura 4.6, verificamos que a distribuição de retornos para $\Delta t = 60$ minutos é mais larga do que nas escalas de tempo menores.

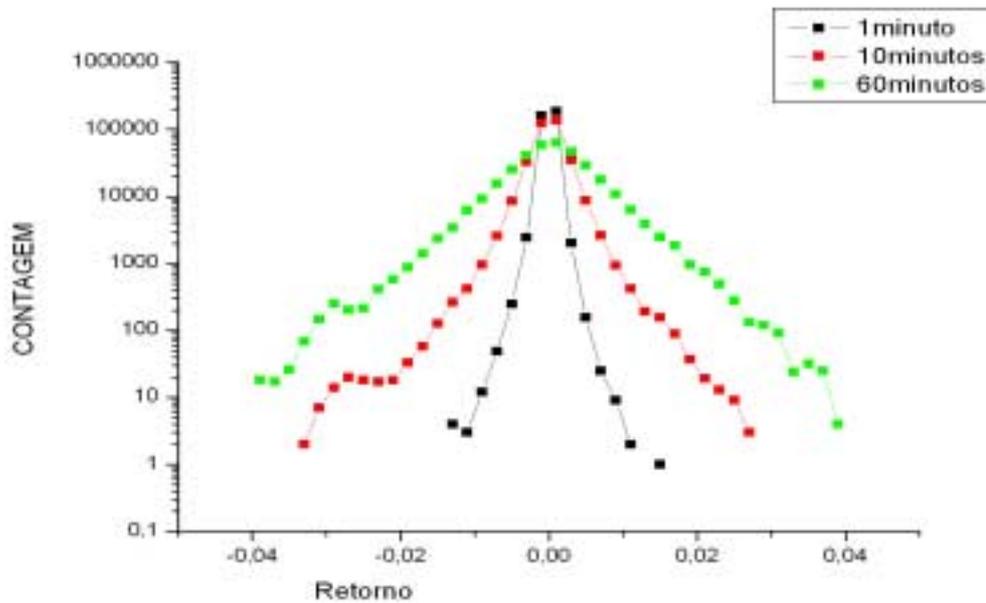


Figura 4.7 – Histograma de retornos do IBOVESPA para escalas de tempo $\Delta t = 1, 10$ e 60 minutos.

Na figura 4.8 apresentamos a série temporal de retornos normalizados de acordo com a definição (1.8) para as mesmas escalas de tempo da figura 4.6. Lembrando que estas séries possuem média nula e desvio padrão unitário, verificamos que a série de $\Delta t=1$ minuto é a que possui maior frequência de grandes flutuações relativas entre as escalas de tempo apresentadas.

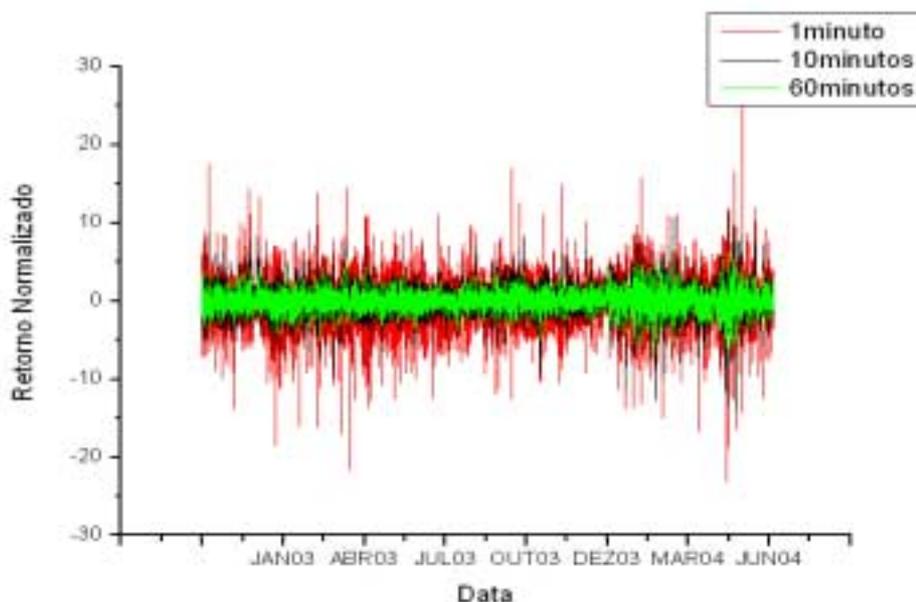


Figura 4.8 – Série histórica de retornos normalizados do IBOVESPA

A figura 4.9 apresenta para comparação os histogramas para retornos normalizados nas escalas de tempo $\Delta t = 1$ minuto e 60 minutos. Vê-se que a distribuição normalizada para a escala de tempo de $\Delta t = 1$ minuto, é a mais leptocúrtica, significando que é a série que apresenta maior ocorrência de grandes flutuações relativas, de acordo com o observado na série temporal da figura 4.8.

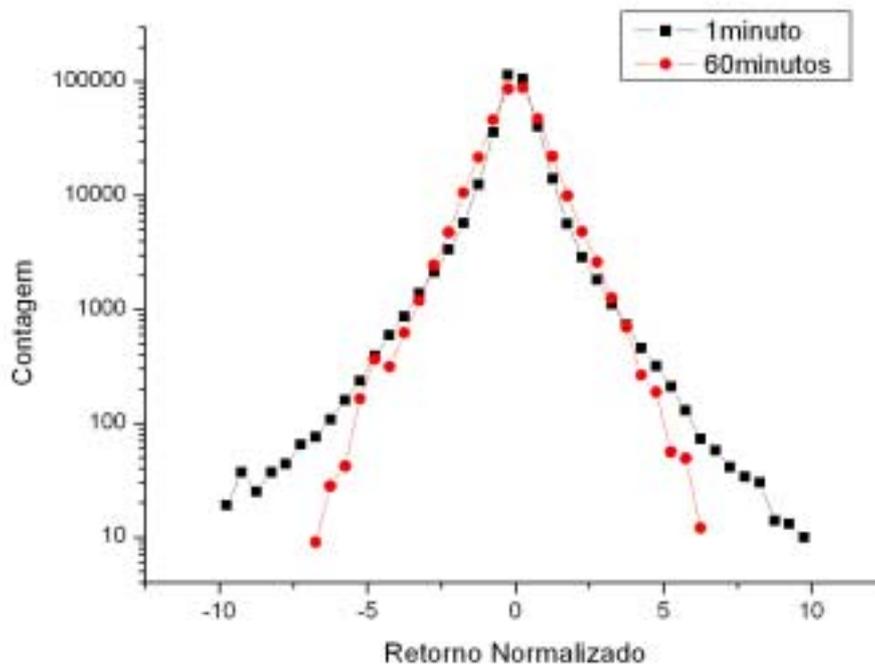


Figura 4.9 – Histograma de retornos normalizados do IBOVESPA para escalas de tempo $\Delta t = 1$ e 60 minutos.

Considere agora o histograma de retornos normalizados em várias escalas microscópicas de tempo, ou seja, intervalos de tempo da ordem de minutos (figura 4.10). Verificamos um colapso de dados das distribuições sobre uma mesma forma, uma curva característica. A invariância das distribuições dos preços de alta-freqüência por mudança de escala temporal sugere que a variação de preços neste horizonte de eventos é governada pelos mesmos mecanismos de mercado. Este resultado indica que devemos modelar as flutuações de preços do mercado intradiário por uma equação de Itô-Langevin na qual a solução da equação de Fokker-Planck associada possua forma invariante por mudança de escala temporal.

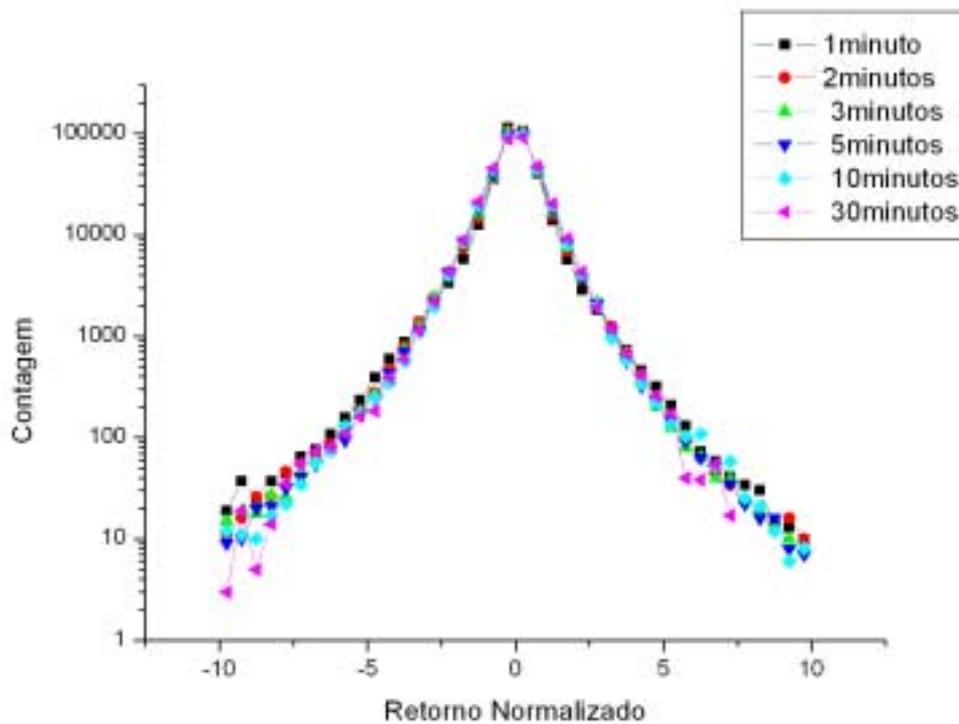


Figura 4.10 – Histograma de retornos normalizados do IBOVESPA para escalas de tempo apresentadas na legenda da figura (colapso de dados).

No entanto, como mostrado na figura 4.9, o colapso de dados não é verificado entre a escala de tempo de 1 minuto e a escala de tempo de 60 minutos. Isto indica que a estacionariedade do comportamento do mercado de alta frequência é quebrada para escalas de tempo da ordem de 1 hora.

O objetivo desta dissertação é modelar as propriedades estatísticas e dinâmicas dos retornos de alta frequência em escala de tempo $\Delta t < 60$ minutos, caracterizando o regime do mercado de ações brasileiro em escalas temporais microscópicas.

4.2

Modelagem da Distribuição de Retornos na Escala de 1 minuto

Nesta seção, apresentaremos as distribuições de retorno de alta frequência para várias escalas temporais Δt e a modelagem destas distribuições baseadas em distribuições de Tsallis, definidas pela equação (2.15).

A seguir mostramos o gráfico para a distribuição de retornos normalizados na janela temporal de $\Delta t=1$ minuto em escala semi-logarítmica. Este gráfico foi construído utilizando 40 bins e considerando retornos reais de amplitude máxima de 10σ (em relação à média).

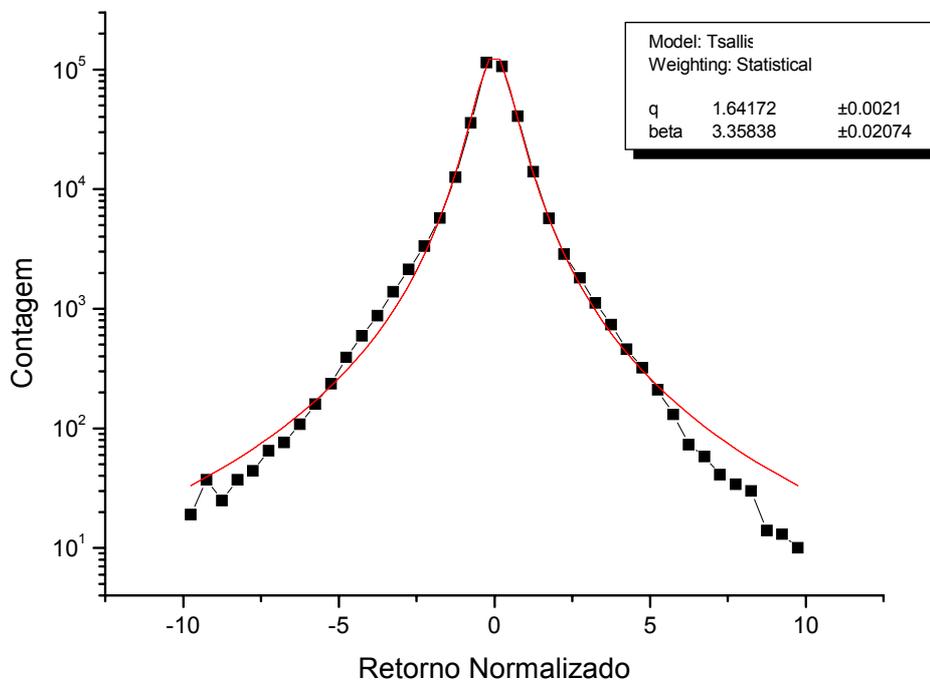


Figura 4.11 – Histograma de retornos normalizados para intervalo de tempo $\Delta t=1$ minuto (linha e pontos pretos) e distribuição de Tsallis que apresenta melhor descrição do comportamento dos dados. Parâmetros ótimos apresentados na legenda.

Através da figura 4.11, observa-se que a q-Gaussiana de ajuste para toda a faixa de valores empíricos de retorno descreve bem o centro da distribuição porém não fornece uma boa aproximação para os valores da cauda, em especial para valores de retorno normalizados com módulos maiores do que 5σ .

Como a q-Gaussiana ótima superestima a probabilidade de ocorrência de grandes valores de retornos, é necessário realizar um truncamento na cauda da distribuição para suavizá-la. Para descobrir o tipo de truncamento gradual (2.25), se em lei de potência ou tipo exponencial, realizamos a análise da distribuição acumulada dos retornos normalizados. Desta forma, podemos inferir com maior precisão o caráter do comportamento da cauda, pois as flutuações estatísticas da distribuição original são atenuadas. A seguir, apresentamos respectivamente nas figuras 4.12 e 4.13, os gráficos semi-logarítmico e log-log da distribuição acumulada de retornos de 1 minuto com magnitude até 10σ .

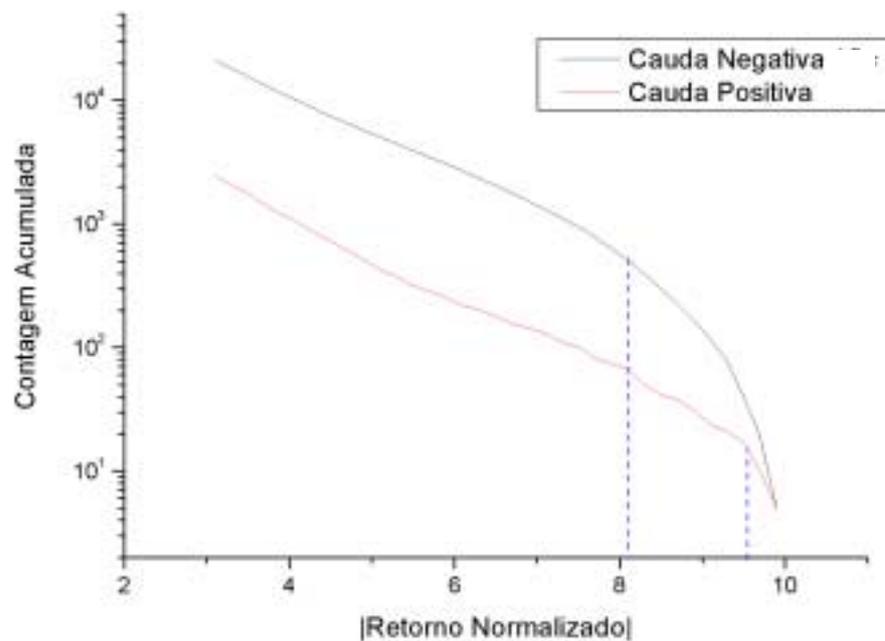


Figura 4.12 – Gráfico semi-logarítmico da distribuição acumulada de retornos normalizados em intervalo de tempo $\Delta t= 1$ minuto. A linha preta representa a cauda de retornos negativos (transladada para melhor visualização) e a linha vermelha, os retornos positivos. A linha pontilhada azul delimita o limite considerado para existência de um comportamento exponencial.

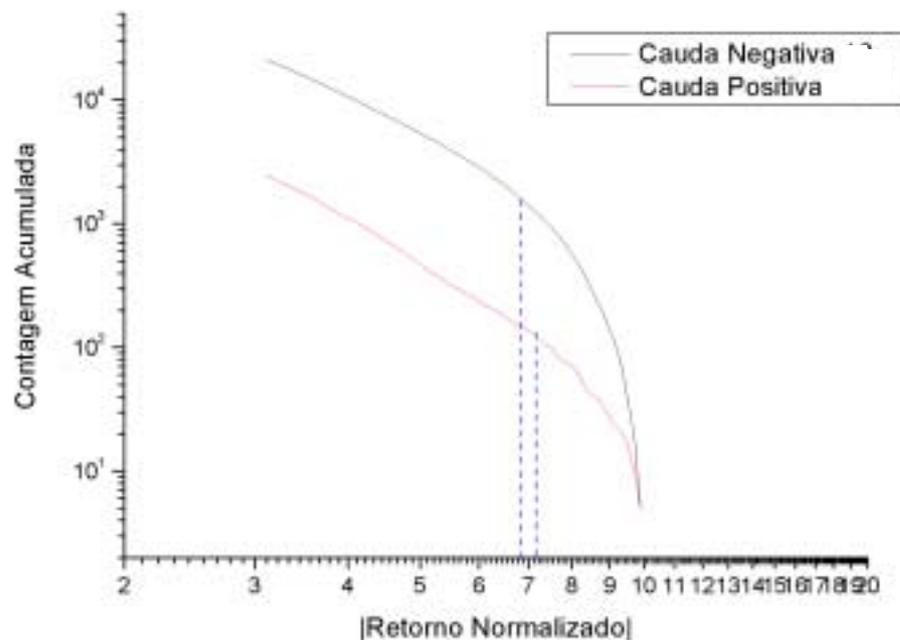


Figura 4.13 – Gráfico log-log da distribuição acumulada de retornos normalizados em intervalo de tempo $\Delta t=1$ minuto. A linha preta representa a cauda de retornos negativos (transladada para melhor visualização) e a linha vermelha, os retornos positivos. A linha pontilhada azul delimita o limite considerado para existência de um comportamento em lei de potência.

Na figura 4.12, observa-se que o limite de persistência de uma reta, que caracteriza o caráter exponencial da cauda da distribuição de $\Delta t=1$ minuto, se estende até aproximadamente 8σ para a cauda negativa e para praticamente toda a faixa de valores de retorno na cauda positiva. Na figura 4.13, observa-se um comportamento aproximadamente linear apenas até 7σ , característico de cauda em lei de potência. O decaimento abrupto da distribuição acumulada para valores extremos de retornos nas figuras (4.12) e (4.13) deve-se à insuficiência de dados e não devem ser levados em conta na análise.

Através da comparação dos resultados apresentados nas figuras 4.12 e 4.13, observa-se que o truncamento exponencial descreve uma faixa mais larga de valores de retorno. No entanto, como esta faixa ainda cobre menos do que uma década de valores, ainda não se pode inferir o tipo de truncamento que melhor caracteriza a cauda da distribuição de retornos de $\Delta t=1$ minuto. Analisando porém as distribuições de retornos para escalas de tempo maiores, como será mostrado adiante, observa-se que de fato o truncamento exponencial é o que melhor caracteriza a cauda das distribuições em escala de tempo intradiária.

Para analisarmos quantitativamente o comportamento exponencial das caudas das distribuições, é necessário identificar inicialmente a região central de valores de retornos que é bem descrita por q-Gaussianas. Para isso, vamos fazer a seguir algumas análises comparativas dos parâmetros ótimos, buscando obter robustez de resultados.

Nas figuras 4.14 e 4.15 são mostradas as distribuições de retornos normalizados para as escalas de tempo de $\Delta t=1,2$ e 3 minutos, considerando-se retornos de magnitude máxima igual a 10σ e 20σ respectivamente assim como as q-Gaussianas que melhor representam estas distribuições empíricas na faixa de retornos considerada.

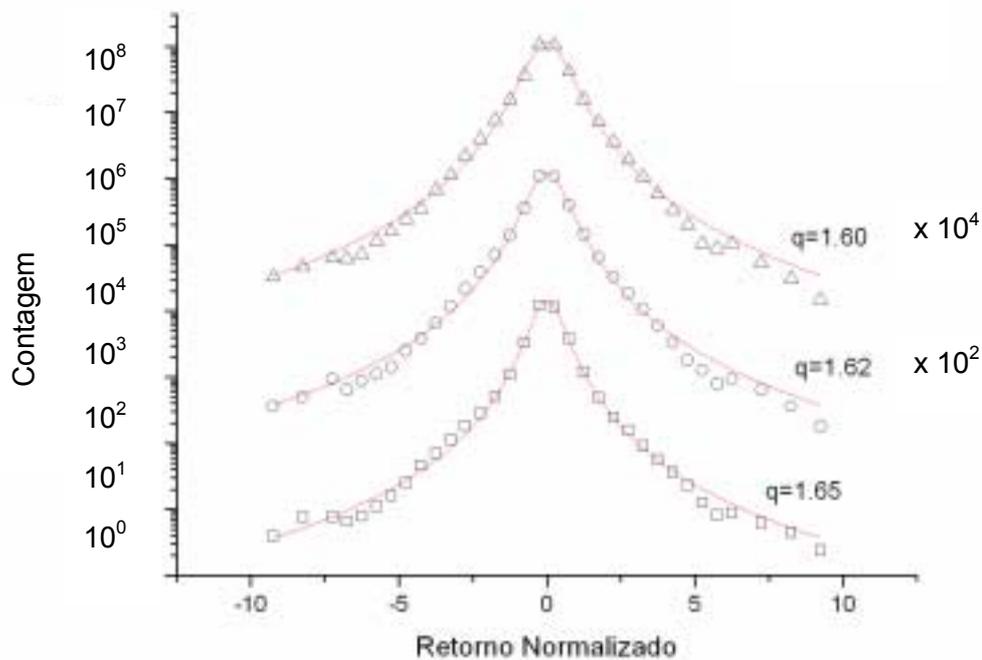


Figura 4.14 – Distribuições empíricas e q-Gaussianas ótimas (linhas vermelhas) para retornos normalizados do IBOVESPA de magnitude até 10σ , em escalas de tempo $\Delta t=1$ minuto (\square), 2 minutos (\circ) e 3 minutos (\triangle). As curvas para $\Delta t = 2$ e 3 minutos foram movidas verticalmente para melhor visualização.

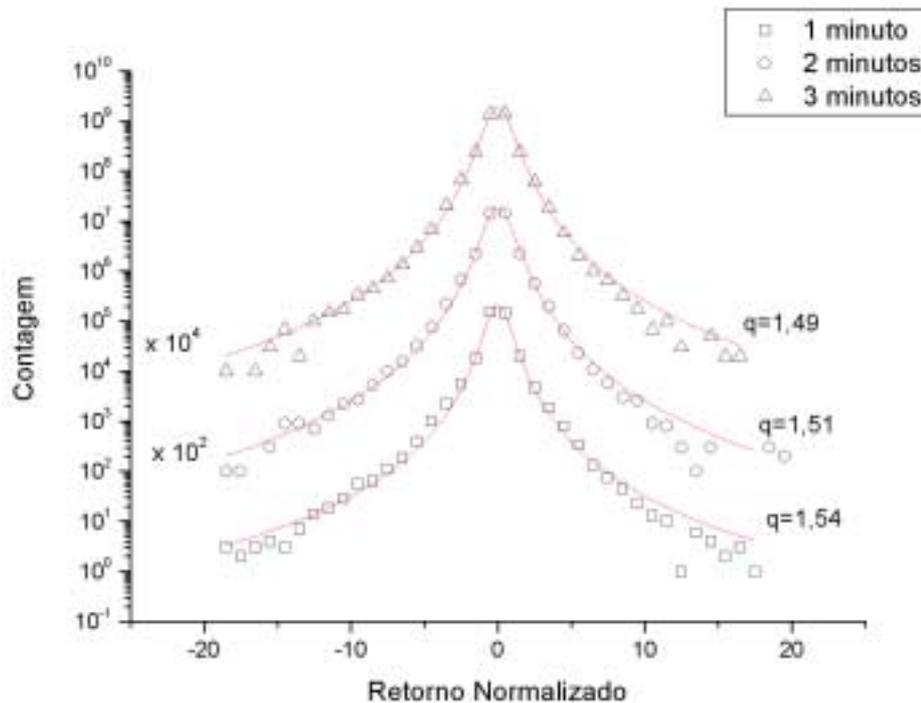


Figura 4.15 – Distribuições empíricas e q-Gaussianas (linhas vermelhas) para retornos normalizados do IBOVESPA com magnitude até 20σ . As curvas para $\Delta t = 2$ e 3 minutos foram movidas verticalmente para melhor visualização.

Comparando-se os parâmetros não-extensivos q ótimos mostrados nas figuras, observa-se um decréscimo sistemático do valor de q quando são consideradas faixas mais largas de valores de retorno. Isto se deve à tentativa do ajuste de incluir regiões cada vez maiores da cauda, cujo decaimento é exponencial, através de q-Gaussianas, cuja cauda possui comportamento em lei de potência. Para cumprir tal objetivo, o valor do q ótimo é diminuído sistematicamente. Caso a cauda da distribuição fosse em forma de lei de potência, esperaria-se a convergência do valor de q ótimo à medida que mais dados da cauda fossem incluídos, de acordo com (2.21).

Passamos a investigar a largura da região central de valores de retorno (considerada como o “bulk” de valores) cujas distribuições são bem descritas por q-Gaussianas. Para obter o limite da região do *bulk*, apresentamos a seguir nas tabelas 4.1 e 4.2 os ajustes dos parâmetros q e β em (2.15) para as distribuições de retornos normalizados de $\Delta t=1$ minuto ao longo dos meses, de acordo com a faixa de retornos considerada.

MÊS	$ Z < 10\sigma$	$ Z < 5\sigma$	$ Z < 3\sigma$
Novembro 2002	1,58	1,64	1,65
Dezembro 2002	1,54	1,60	1,60
Janeiro 2003	1,59	1,64	1,66
Fevereiro 2003	1,58	1,66	1,66
Março 2003	1,68	1,78	1,78
Abril 2003	1,64	1,74	1,75
Mai 2003	1,68	1,80	1,83
Junho 2003	1,66	1,76	1,79
Julho 2003	1,62	1,69	1,70
Agosto 2003	1,63	1,75	1,77
Setembro 2003	1,59	1,73	1,82
Outubro 2003	1,57	1,66	1,73
Novembro 2003	1,58	1,61	1,68
Dezembro 2003	1,56	1,74	1,75
Janeiro 2004	1,64	1,77	1,77
Fevereiro 2004	1,67	1,74	1,89
Março 2004	1,66	1,71	1,76
Abril 2004	1,61	1,77	1,80
Mai 2004	1,65	1,76	1,86
Junho 2004	1,66	1,76	1,81

Tabela 4.1 – Parâmetros q ótimos para a distribuição de retornos normalizados no intervalo de $\Delta t = 1$ minuto para cada mês entre novembro de 2002 e junho de 2004.

MÊS	$ Z < 10\sigma$	$ Z < 5\sigma$	$ Z < 3\sigma$
Novembro 2002	2,46	3,03	3,17
Dezembro 2002	2,36	2,81	2,80
Janeiro 2003	2,77	3,29	3,46
Fevereiro 2003	2,63	3,32	3,31
Março 2003	4,42	5,25	5,40
Abril 2003	3,57	4,74	4,93
Mai 2003	4,45	6,12	6,74
Junho 2003	3,44	4,63	5,14
Julho 2003	3,45	4,30	4,46
Agosto 2003	3,13	4,26	4,51
Setembro 2003	2,79	3,27	4,55
Outubro 2003	2,42	3,29	3,51
Novembro 2003	2,61	2,62	3,39
Dezembro 2003	2,26	4,26	4,35
Janeiro 2004	3,23	4,88	4,51
Fevereiro 2004	3,58	4,52	6,06
Março 2004	3,56	3,94	4,71
Abril 2004	2,97	4,34	4,68
Mai 2004	3,94	5,34	6,45
Junho 2004	3,63	4,75	5,34

Tabela 4.2 – Parâmetros de escala β ótimos para a distribuição de retornos normalizados no intervalo de $\Delta t = 1$ minuto para cada mês entre novembro de 2002 e junho de 2004.

Observa-se que ao longo dos meses há em geral uma diferença menor dos parâmetros ótimos entre a segunda e terceira coluna, havendo maior discrepância em relação aos resultados da primeira coluna. Isto nos leva a considerar o limite de magnitude de 5σ para a região central da distribuição de retornos. Retornos de magnitude maior do que 5σ serão tratados nesta dissertação como valores de cauda e não serão incluídos nos ajustes das q-Gaussianas.

Considerando-se magnitudes até 5σ , verifica-se que de fato, a região central do histograma de retornos de $\Delta t = 1$ minuto é bem descrita por uma q-Gaussiana, como mostrado na figura 4.16.

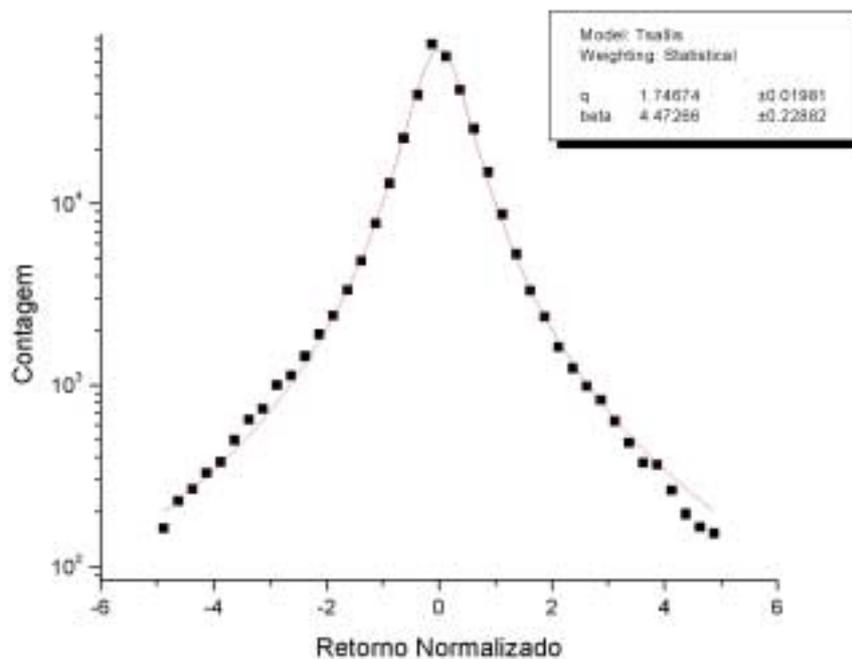


Figura 4.16 – Distribuição empírica e q-Gaussiana ótima (linha vermelha) para retornos normalizados do IBOVESPA com magnitude até 5σ . Na legenda estão os parâmetros relativos a q-Gaussiana ótima

Comparando-se as figuras 4.11 e 4.16, verifica-se, que de fato, o parâmetro ótimo q aumenta conforme diminuimos a extensão dos valores de retornos considerados no ajuste.

Por outro lado, comparando-se os parâmetros ótimos para a distribuições de $\Delta t=1$ minuto mostradas nas figuras 4.11 e 4.14, observa-se uma pequena diferença de resultados, apesar de ter sido considerada mesma faixa de valores de retornos normalizados. Enquanto o histograma da figura 4.11 foi construído com 40 bins, o histograma da figura 4.14 possui apenas 32 bins. Isto nos leva a analisar um segundo fator que influencia os resultados de ajuste, que é o número de bins utilizados nos histogramas analisados.

Ao modificar a largura dos bins, agrupamos dados em diferentes classes e consideramos cada classe através de um único valor de frequência total. Ao tomarmos grupos diferentes de dados, teremos uma sutil variação na aparência do histograma. Essa diferença se reflete nas variações respectivas dos parâmetros q e β ótimos. Na tabela 4.3 apresentamos os valores obtidos de acordo com o número de bins utilizados para distribuições de retornos em várias

escalas temporais. Em todos os ajustes, foram considerados apenas retornos de módulo menor do que 5σ .

$\Delta t(\text{min})$	q	β	q	β	q	β	q	β
	20bins	20bins	40bins	40bins	100bins	100bins	1000bins	1000bins
1	1,70	3,70	1,75	4,47	1,78	4,88	1,72	4,16
2	1,68	3,03	1,72	3,55	1,73	3,68	1,70	3,43
3	1,64	2,60	1,68	3,01	1,69	3,11	1,66	2,96
4	1,61	2,33	1,65	2,67	1,66	2,75	1,63	2,64
5	1,58	2,12	1,62	2,40	1,62	2,47	1,61	2,40
6	1,57	2,02	1,60	2,25	1,60	2,31	1,58	2,25
7	1,55	1,91	1,58	2,13	1,58	2,18	1,56	2,12
8	1,53	1,86	1,56	2,06	1,57	2,10	1,55	2,05
9	1,53	1,82	1,56	2,00	1,56	2,05	1,54	2,00
10	1,51	1,72	1,55	1,95	1,56	2,00	1,54	1,95
15	1,50	1,65	1,52	1,79	1,53	1,82	1,51	1,79
20	1,49	1,61	1,51	1,73	1,52	1,76	1,50	1,72
25	1,49	1,57	1,51	1,69	1,52	1,71	1,50	1,68
30	1,49	1,52	1,51	1,63	1,51	1,66	1,49	1,62
45	1,48	1,40	1,49	1,48	1,49	1,50	1,47	1,47
60	1,47	1,34	1,48	1,41	1,48	1,43	1,47	1,40

Tabela 4.3 – Variação dos parâmetros q e β de acordo com o número de bins utilizados, considerando-se retornos tais que $|x| \leq 5\sigma$

Devido às diferenças entre os parâmetros ótimos das distribuições ao modificarmos a quantidade de bins, é necessária uma padronização de sua utilização, com a escolha adequada de um número fixo de bins. Através dos resultados da tabela 4.3 conclui-se que com 20 bins, tem-se um número muito pequeno de pontos para ajuste, gerando limite inferior para o parâmetro q . Por outro lado, nos histogramas com 1000 bins, verificamos muitas flutuações (ver figura 4.17).

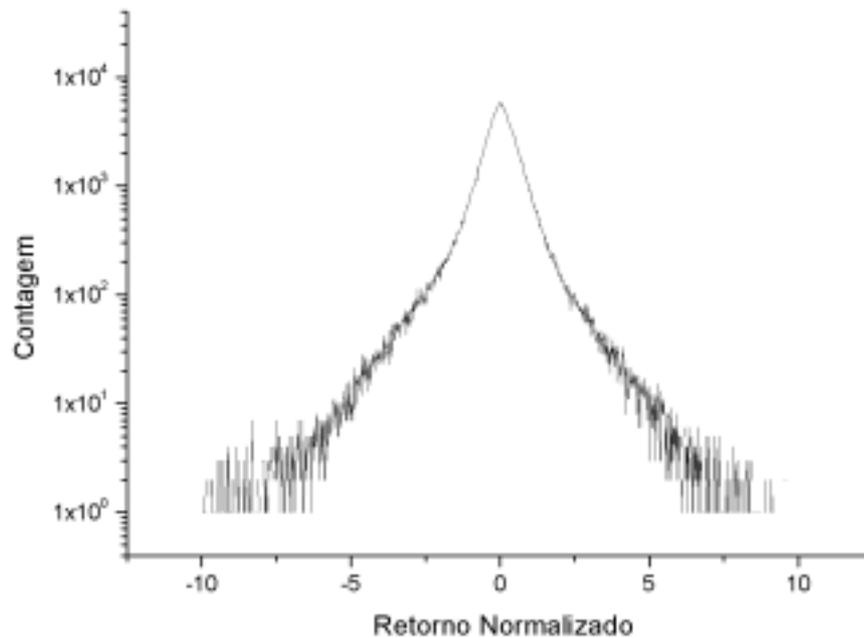


Figura 4.17 – Histograma de retorno normalizado para $\Delta t= 1$ minuto construído com 1000 bins.

A partir da tabela 4.3, vê-se ainda que a utilização de 100 bins majora os parâmetros ótimos. Dessa forma, adotamos para nossa análise distribuições com 40 bins.

Tendo definido o limite do *bulk* da distribuição de $\Delta t=1$ minuto como 5σ e o número de bins dos histogramas como sendo igual a 40, obtemos o valor ótimo do parâmetro não extensivo $q=1,75$ para a região central desta distribuição.

Mostramos a seguir, gráficos com o comportamento de q e β ótimos ao longo dos meses, considerando para o ajuste a faixa de valores de *bulk* para retornos de $\Delta t=1$ minuto normalizados. As barras de erro representadas nas figuras 4.18 e 4.19 correspondem ao desvio padrão dos valores mensais encontrados. É feita também comparação com os parâmetros ótimos obtidos para a distribuição global de todo o período analisado.

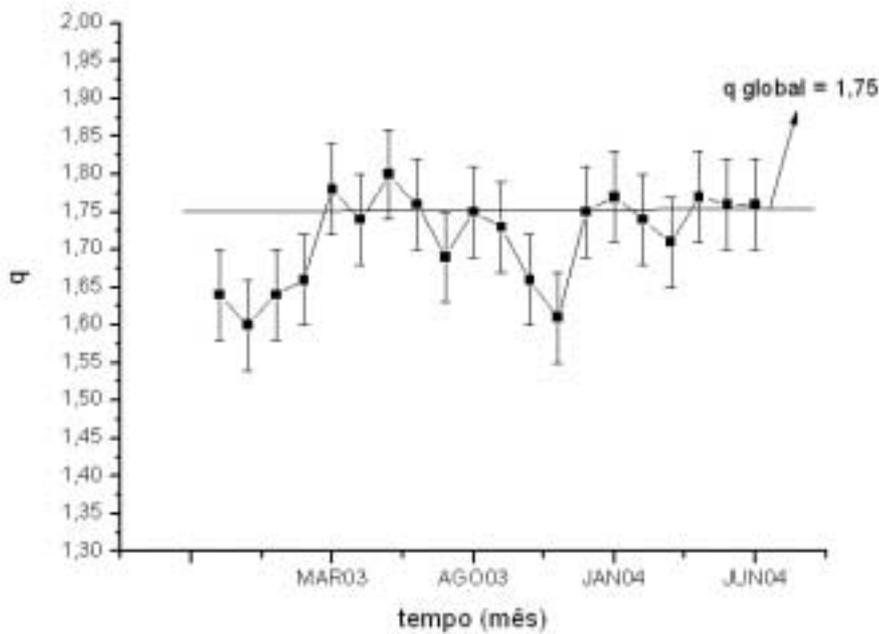


Figura 4.18 – Comportamento do parâmetro ótimo q em cada mês estudado em nossa análise (de Novembro de 2002 a Junho de 2004) e barra de erro do conjunto dos valores mensais encontrados. A linha representa o parâmetro ótimo $q = 1,75$ da distribuição global.

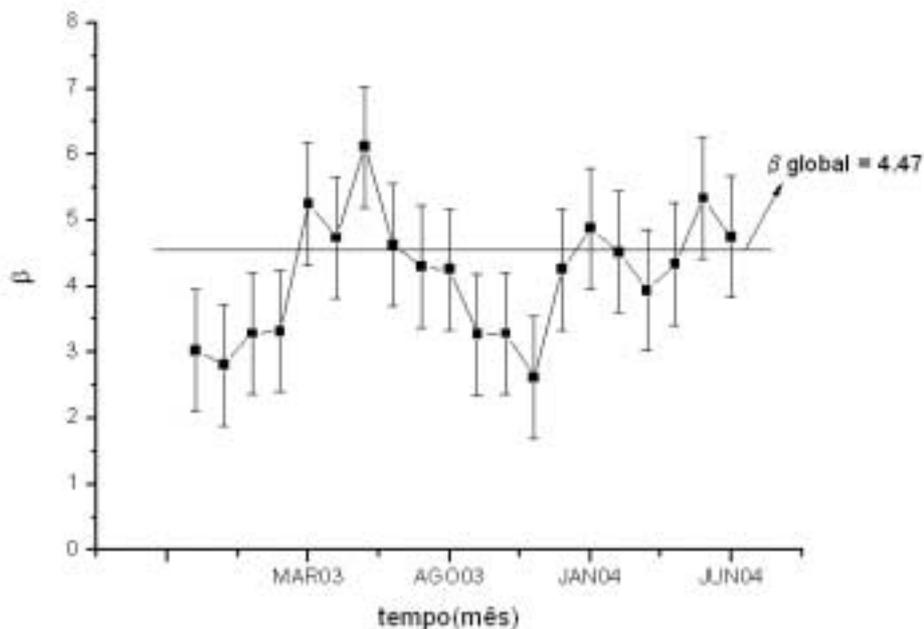


Figura 4.19 – Comportamento do parâmetro ótimo de escala β em cada mês estudado em nossa análise (de Novembro de 2002 a Junho de 2004) e barra de erro do conjunto dos meses. A linha representa o parâmetro ótimo $\beta = 4,47$ da distribuição global.

Podemos constatar que os parâmetros encontrados para a maioria dos meses do período analisado são consistentes com o comportamento da distribuição global. Isto significa que esta modelagem é razoavelmente estável no período estudado.

Uma vez definida a região central dos valores de retornos de $\Delta t = 1$ minuto que são regidos por q-Gaussianas, passaremos à obtenção dos dados relativos à cauda exponencial para retornos de módulo maior do que 5σ , representada pela função de truncamento exponencial do tipo (2.25a). Segue-se que:

$$P_{ACUM}(x) \sim \exp\left(-\frac{(|x| - x_C)}{\xi}\right) \quad \text{para } |x| > x_C \quad (4.1)$$

A partir da inclinação do gráfico semi-logarítmico da distribuição acumulada de retornos mostrada na figura 4.20, obtemos $\xi \cong 4$:

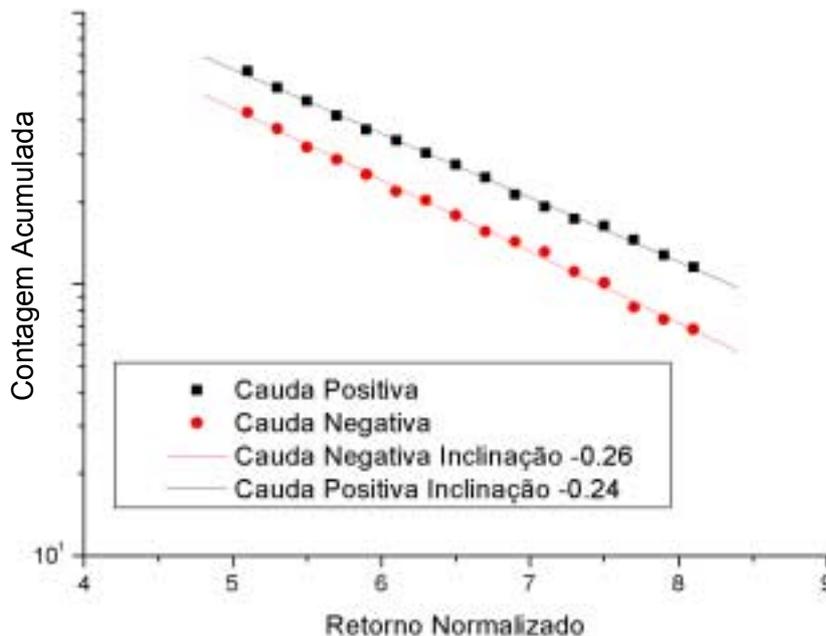


Figura 4.20 – Gráfico semi-logarítmico da distribuição acumulada para os valores de cauda da distribuição de retorno normalizado para $\Delta t = 1$ minuto.

Desta forma, nossa análise indica modelagem para a distribuição de retornos normalizados de $\Delta t = 1$ minuto dada pela distribuição de Tsallis Truncada (P_q^T), descrita em analogia com (2.23) e (2.25a) por:

$$\begin{cases} P_q^T(x) = CP_q(x) & \text{para } |x| \leq x_C \\ P_q^T(x) = CP_q(x) \exp\left(-\frac{|x| - x_C}{\xi}\right) & \text{para } |x| > x_C \end{cases} \quad (4.2)$$

Apresentamos na figura abaixo o gráfico da distribuição de retornos normalizados de $\Delta t = 1$ minuto e a respectiva distribuição de Tsallis truncada com parâmetros $q=1,75$ e $\beta= 4,47$. Foi utilizado o valor $x_C = 4\sigma$, consistente com a constante de decaimento exponencial $\xi = 4\sigma$ e a região $|x| > 5\sigma$ considerada para ajuste puramente exponencial da cauda. Como se pode concluir pela figura 4.21, a distribuição empírica fica descrita de forma acurada por esta modelagem.

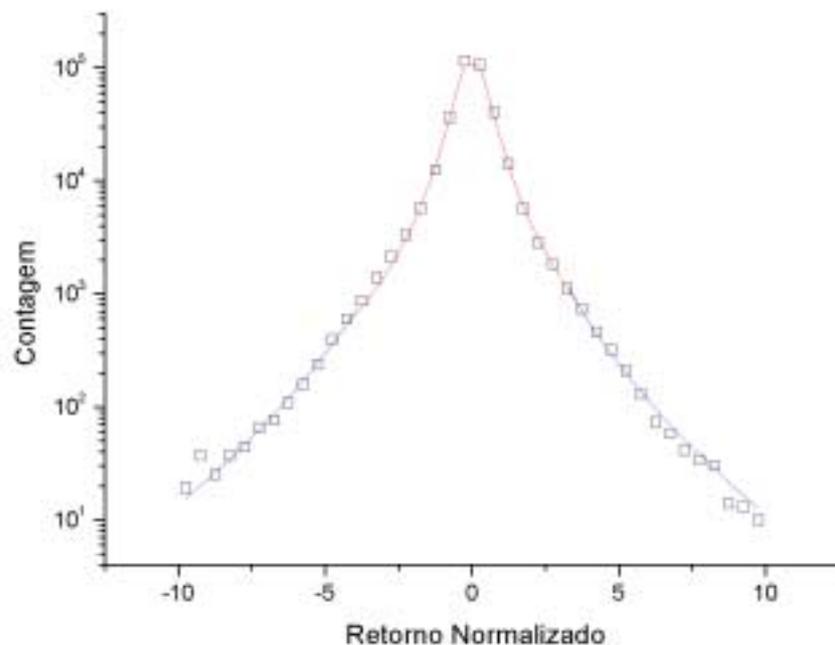


Figura 4.21 – Gráfico semi-logarítmico do histograma de retorno normalizado para $\Delta t = 1$ minuto modelado por uma distribuição de Tsallis truncada: q-Gaussiana com parâmetro $q = 1,75$ para centro da distribuição (linha vermelha) e q-Gaussiana exponencialmente truncada com parâmetro $\xi = 4$ para a cauda (linha azul).

4.3

Análise das Correlações Temporais Intradiárias

Na escala intradiária utilizada, as distribuições empíricas ainda estão bem distantes da distribuição Gaussiana, que é caracterizada pelo parâmetro $q=1$. A aparente violação do T.L.C. nos mercados financeiros se deve predominantemente à presença de correlação entre retornos de preços sucessivos, fenômeno que retarda a convergência para a distribuição Gaussiana atratora. Para provar a presença de correlação intrínseca à série de retornos intradiários, analisaremos comparativamente, para várias escalas de tempo, as distribuições de retornos empíricos (que denominaremos com a sigla EMP) e as distribuições obtidas a partir de uma série artificial gerada pela mistura da ordem temporal dos retornos reais (que denominaremos MIST). Desta forma, eliminamos efeitos de memória entre retornos sucessivos que porventura existam na série real. A série artificial obtida é então não-correlacionada. A série de retornos de $\Delta t=1$ minuto foi misturada aleatoriamente utilizando a subrotina do Fortran77 (ver apêndice B). Como era de se esperar, esta série possui os mesmos valores de média e desvio padrão da série original, confirmados pela tabela 4.4.

	Dados	Média	Desvio Padrão
1minuto MIST	163.830	3,39E-6	1,51E-4
1minuto EMP	163.830	3,39E-6	1,51E-4

Tabela 4.4 – Média e Desvio Padrão das séries de $\Delta t=1$ minuto real (EMP) e misturada (MIST).

A existência de correlação nos retornos empíricos posará ser detectada através da comparação das distribuições EMP e MIST para escalas de tempo maiores. Estas distribuições são obtidas, utilizando (2.14), a partir da soma de retornos sucessivos de $\Delta t=1$ minuto proveniente da série empírica ou artificial, respectivamente. Como podemos verificar na figura 4.22, estas distribuições possuem comportamentos que se diferenciam cada vez mais à medida que a escala temporal aumenta.

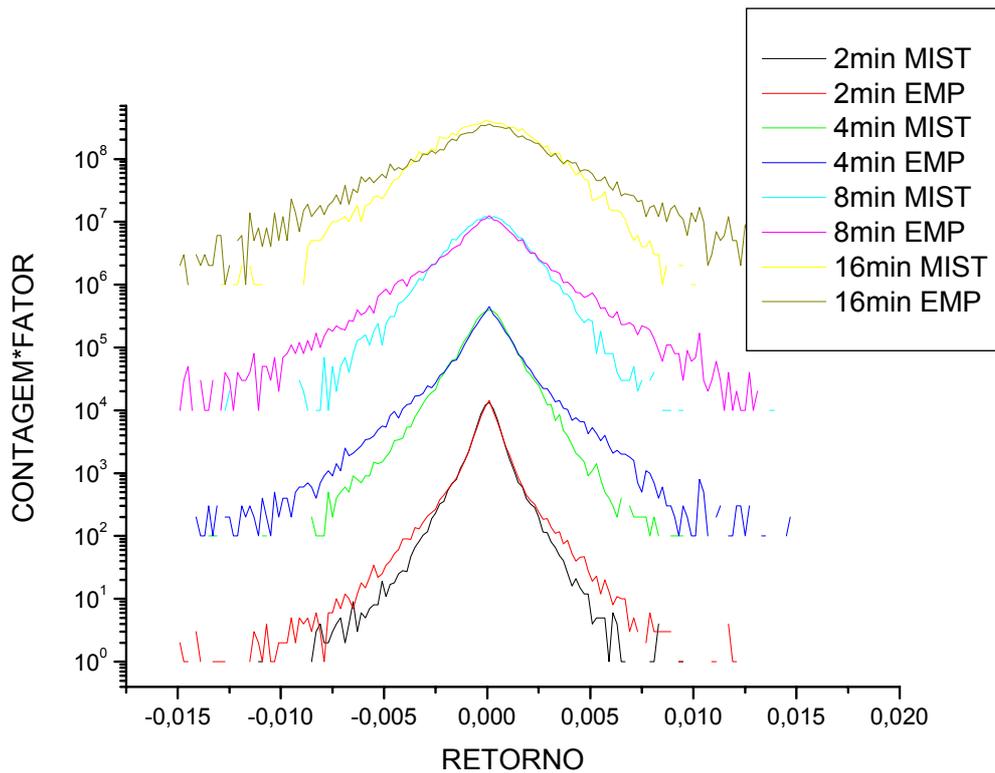


Figura 4.22 – Distribuições de retorno (EMP) e as construídas a partir da agregação de retornos de $\Delta t=1$ minuto misturados (MIST). Os gráficos estão multiplicados por um fator para melhor visualização.

A figura 4.23 mostra ainda em gráfico semi-logarítmico ampliado, os resultados para $\Delta t=16$ minutos, onde é nítida a diferença de comportamento entre as distribuições. O gráfico comparativo das distribuições EMP e MIST para $\Delta t=16$ minutos em escala linear utilizado na figura 4.24 tem como objetivo mostrar que, apesar das diferenças entre estas distribuições, ambas satisfazem à condição de normalização.

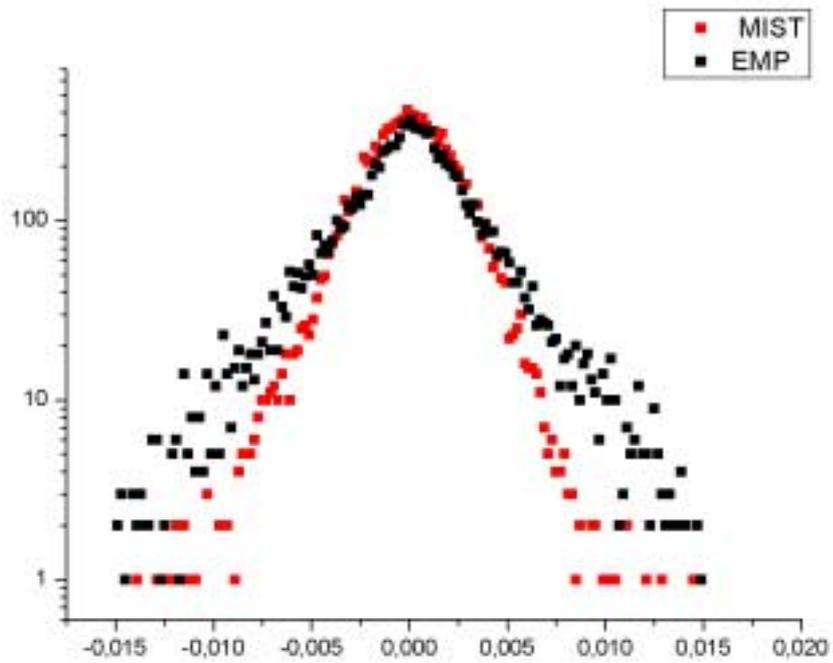


Figura 4.23 – Distribuições de retorno EMP e MIST para $\Delta t = 16\text{min}$ em escala semi-logarítmica.

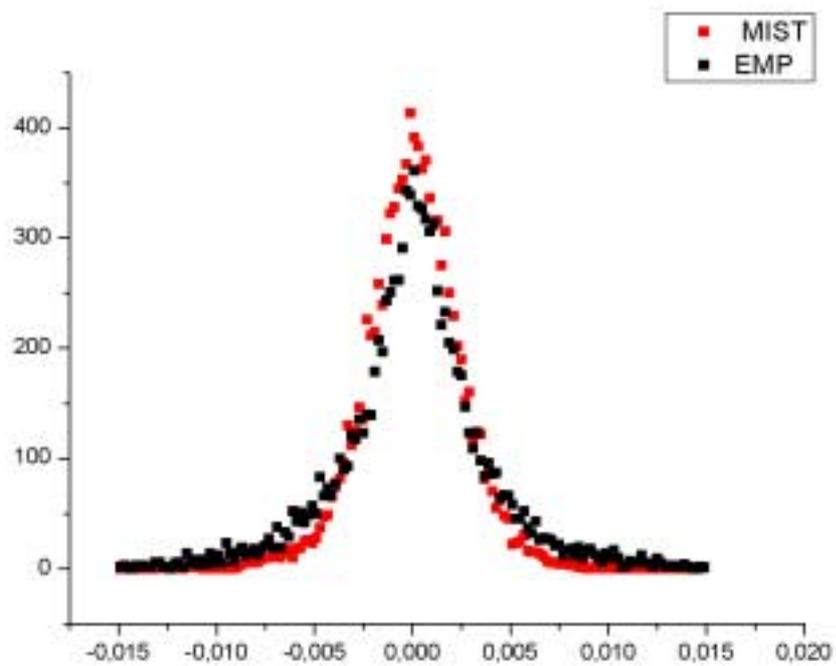


Figura 4.24 – Distribuições de retorno EMP e MIST para $\Delta t = 16\text{min}$ em escala linear.

Através da obtenção de um parâmetro q efetivo ótimo das distribuições de retornos EMP e MIST, podemos caracterizar a convergência para a distribuição Gaussiana em cada caso. Foram feitos ajustes para os histogramas nas escalas de tempo $\Delta t = 2, 4, 8, 16, 30, 45$ e 60 minutos. Estes histogramas foram construídos na forma escolhida como padrão: módulos de retorno menores que 5σ e número de bins igual a 40. Os valores encontrados para q_{efetivo} estão apresentados na tabela 4.5 juntamente com o resultado anteriormente obtido do valor ótimo $q = 1,75$ para $\Delta t = 1$ minuto. No caso MIST, a convergência para a Gaussiana é rápida devido à ausência de correlação. Por outro lado, o regime Gaussiano está longe de ser alcançado na escala temporal intradiária dos dados reais. De fato, o ajuste do histograma de retornos diários ainda fornece o valor $q_{\text{efetivo}} \cong 1,4$. Com este procedimento comprovou-se a importância das correlações na evolução temporal das distribuições financeiras.

Escala Δt (min)	MIST	EMP
1	1,75	1,75
2	1,65	1,75
4	1,50	1,73
8	1,33	1,64
16	1,24	1,52
30	1,18	1,51
45	1,07	1,49
60	1,06	1,48

Tabela 4.5 – Valores dos parâmetros q efetivos das distribuições de retorno EMP e MIST para os intervalos $\Delta t = 1, 2, 4, 8, 16, 30, 45$ e 60 minutos.

Para ilustrar a grande diferença de comportamento das distribuições EMP e MIST para escalas de tempo intradiárias, apresentamos a seguir na figura 4.25, um gráfico comparativo para $\Delta t = 45$ minutos, onde no caso MIST, a distribuição é bem ajustada por uma Gaussiana ($q = 1$).

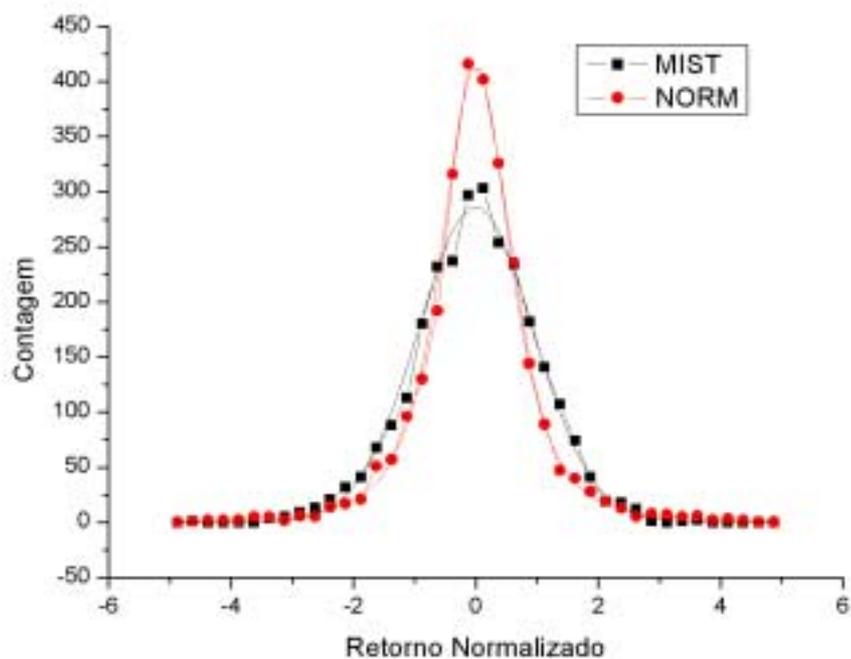


Figura 4.25– Distribuições de retorno EMP e MIST para $\Delta t=45$ minutos em escala linear. As distribuições estão delineadas pelas respectivas q-Gaussianas ótimas.

A existência ou não de correlação na série histórica pode ser inferida a partir do comportamento do desvio padrão das flutuações de preços com o intervalo tempo, descrita por $\sigma(\Delta t) \propto \Delta t^H$, onde H é chamado de expoente de Hurst. A figura 4.26 mostra em escala logarítmica o comportamento $\sigma \times \Delta t$. As distribuições MIST fornecem inclinação $H=0,5$ característica dos processos aleatórios sem memória. As distribuições reais de retornos (EMP) fornecem inclinação $H > 0,5$ caracterizando um processo correlacionado. No entanto, esta inclinação diminui para escalas de tempo mais longas, indicando um declínio das correlações. A presença de correlação, tendo o efeito de retardar o processo de convergência para a Gaussiana, permite o surgimento de um regime quase-estacionário característico das escalas de tempo microscópicas.

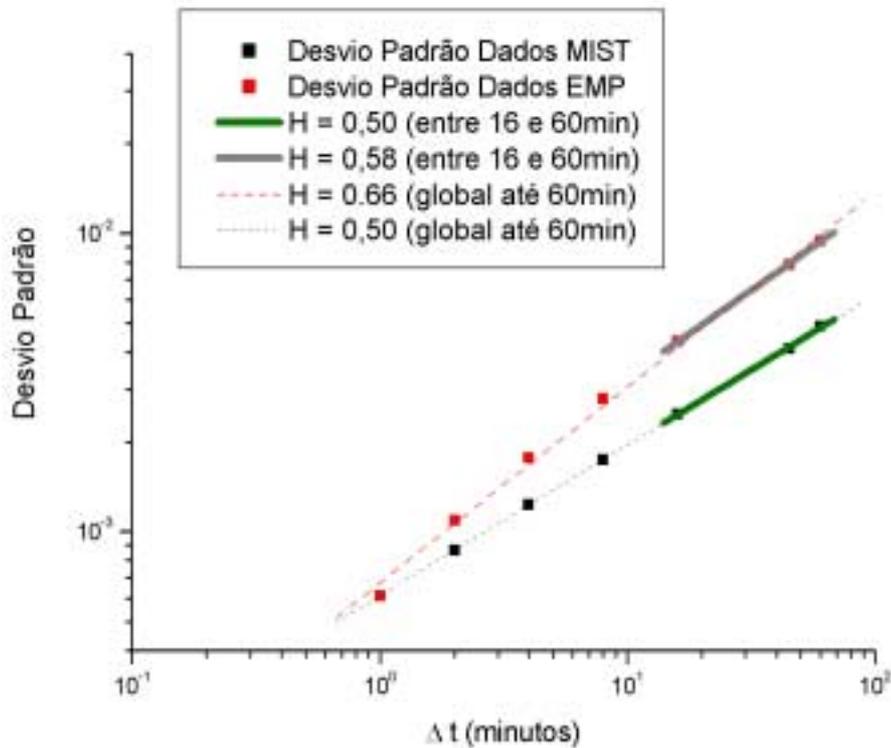


Figura 4.26 – Dependência do desvio padrão das distribuições de retorno EMP e MIST com a escala temporal Δt .

Considere a função de autocorrelação linear $R(\tau)$, dada pela equação (1.15b). A figura 4.27 apresenta o comportamento de $R(\tau)$ para os retornos normalizados na escala de $\Delta t=1$ minuto em um gráfico semi-logarítmico. O gráfico mostra que a correlação linear dos retornos é de curto alcance, caracterizado por decaimento exponencial $R(\tau) \propto \exp(-\tau/\tau_c)$. A partir da inclinação da reta mostrada no gráfico, determina-se o valor de $\tau_c = 4,3$ minutos, que caracteriza o tempo de decaimento típico da correlação no mercado brasileiro.

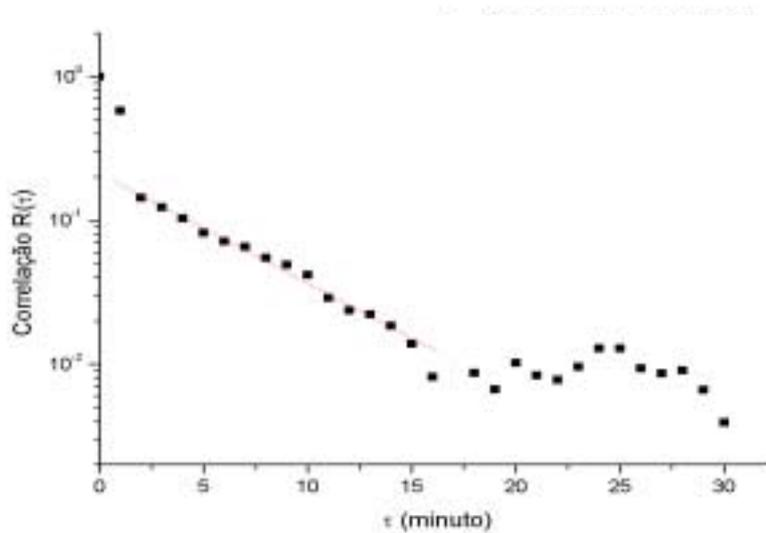


Figura 4.27– Correlação Linear entre retornos de $\Delta t=1$ minuto à distância temporal τ . A linha vermelha representa o ajuste linear entre $\tau=2$ e 15 minutos. A partir de $\tau=20$ minutos, a correlação linear está em nível de ruído.

Como mostrado no capítulo 1, funções não lineares de retornos de preços, tais como o valor absoluto ou o valor quadrático, apresentam memória de longo alcance. A seguir apresentamos na figura 4.28 o gráfico em escala logarítmica do comportamento da função de autocovariância $C(\tau)$ entre os módulos dos retornos normalizados, dada pela equação (1.16).

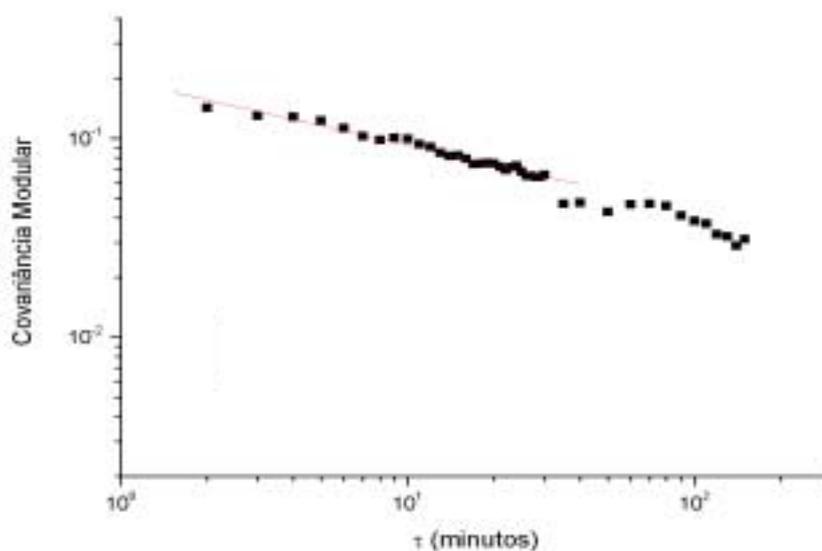


Figura 4.28 – Covariância Modular entre retornos de $\Delta t=1$ minuto à distância temporal τ . A linha vermelha representa o ajuste linear entre $\tau = 2$ e 60 minutos.

A figura 4.28 mostra que a covariância modular tem decaimento em lei de potência $C(\tau) \sim \tau^{-\eta}$. A partir da inclinação da reta mostrada no gráfico encontra-se o valor $\eta \cong 0,32$. Como $\eta < 1$, o processo tem memória de longo alcance, caracterizando retornos dependentes em todas as escalas de tempo intradiárias analisadas.

A seguir apresentamos na figura 4.29 uma comparação entre a persistência da covariância modular, que resiste a uma escala de tempo maior, e a correlação linear, que em poucos minutos, se encontra em nível de ruído.

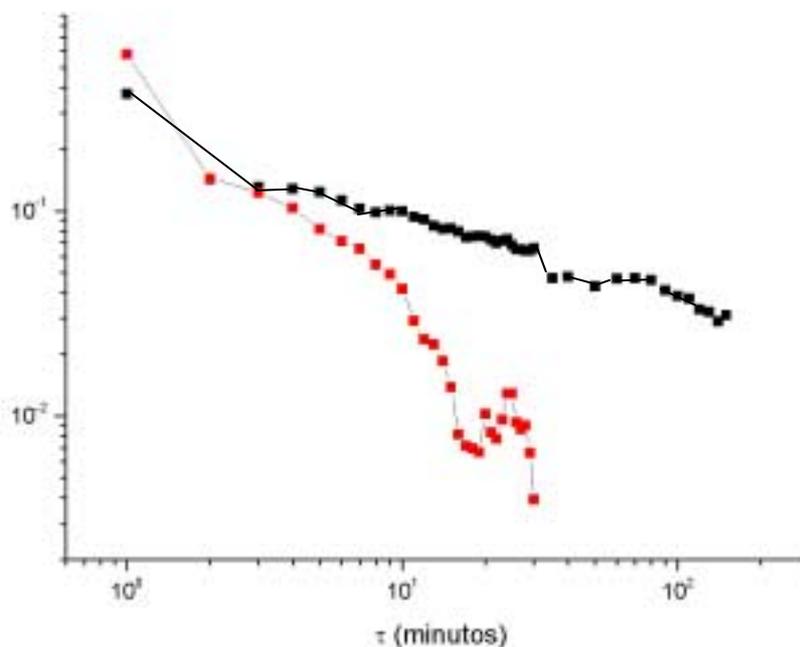


Figura 4.29– Correlação Linear (vermelho) e Covariância Modular (preto) entre retornos de $\Delta t=1$ minuto ocorrendo em intervalo temporal τ .

Podemos comparar os resultados anteriores com os do mercado americano S&P mostrado nas figuras 1.8 e 1.9. Verifica-se um comportamento semelhante entre esses dois mercados tanto na correlação linear quanto na dos módulos:

Mercado	τ_c	η
S&P 500	4 minutos	0,3
IBOVESPA	4,3 minutos	0,32

Tabela 4.6– Comparação dos parâmetros de correlação linear e modular entre os mercados americano[13] e brasileiro[trabalho presente].

4.4

Evolução Temporal das Distribuições de Retornos

A tabela 4.3 parece indicar uma diminuição inicial rápida dos valores efetivos de q à medida que a escala temporal de medida de retornos aumenta. Todos os ajustes apresentados nesta tabela consideram mesma faixa de valores de retornos normalizados isto é, até magnitude de 5σ , que mostramos ser a faixa de valores de retornos de $\Delta t = 1$ minuto cuja distribuição de ocorrência é bem descrita por uma q -Gaussiana.

No entanto, observando a forma das distribuições de retorno em gráfico semi-logarítmico mostradas na figuras 4.30, vê-se que, à medida que a escala temporal aumenta, maior é o domínio da forma triangular, indicando que o regime q -Gaussiano central é cada vez mais consumido pela cauda exponencial. Notar também que à medida que este regime central se confina, a forma da distribuição torna-se parabólica, representando comportamento Gaussiano característico de q -Gaussianas para retornos próximos da origem, de acordo com (2.18).

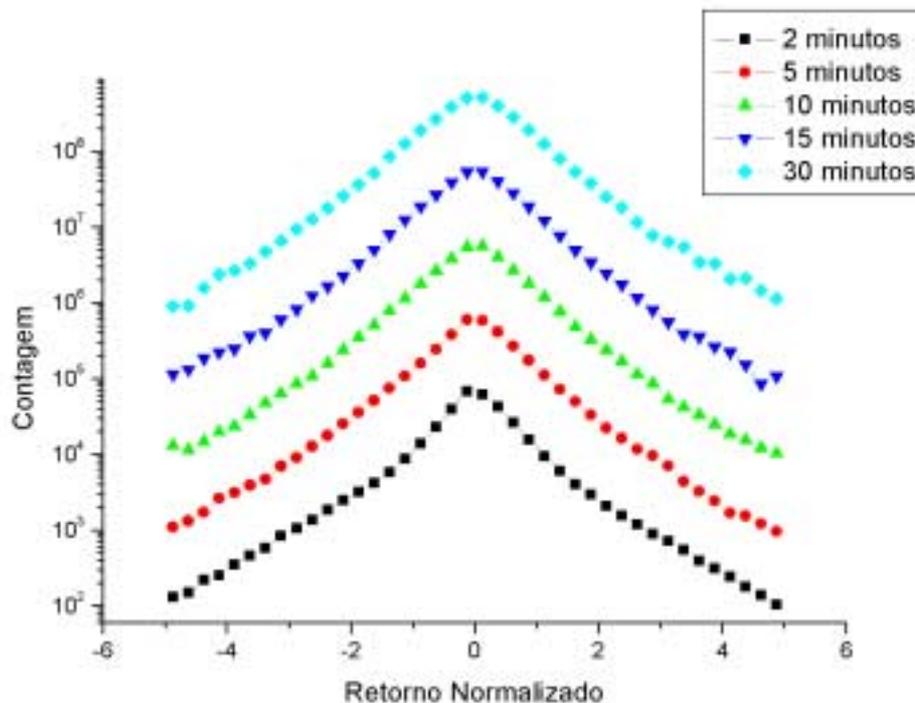


Figura 4.30 – Gráfico semi-logarítmico das distribuições empíricas para retorno normalizado do IBOVESPA em faixa de meia-largura 5σ . Os gráficos estão multiplicados por um fator acumulativo de 10 para melhor visualização. As escalas temporais estão mostradas na legenda.

Isto significa que os parâmetros não-extensivos ótimos para escalas temporais maiores que $\Delta t = 1$ minuto devem ser obtidos para regiões de retornos normalizados cada vez menores, o que implica na obtenção de parâmetros ótimos maiores do que os apresentados na tabela 4.3, de acordo com a análise apresentada na seção 4.2.

O colapso de dados mostrados na figura 4.9 sugere ainda que os dados de *bulk* possam ser modelados por uma mesma q-Gaussiana.

Variando a largura da região central de retornos a ser ajustada por uma q-Gaussiana, verificamos a possibilidade de obtenção do parâmetro ótimo $q=1,75$ para retornos nas demais escalas de tempo, estabelecendo-se assim, a estacionariedade desse valor por mudança de escala temporal. Na tabela 4.7 apresentamos, para vários intervalos de tempo Δt , o limite de valores de retornos, em unidade dos respectivos desvios-padrão, para os quais o parâmetro ótimo é $q=1,75$. Resultados para *bulk* menores do que 1σ não foram considerados pois não seriam significantes.

Δt (min)	<i>bulk</i>
1	5σ
2	4σ
3	$3,5\sigma$
4	3σ
5	2σ
6	$1,5\sigma$
7	1σ

Tabela 4.7 – Limite da região central da distribuição de retornos normalizados para várias escalas temporais nas quais o valor ótimo é $q=1,75$.

No entanto, à medida que Δt aumenta, o desvio padrão em escala real das distribuições aumenta. Assim, apesar de na escala normalizada, escalas de tempo a partir de $\Delta t = 10$ minutos apresentarem região central pequena, isto não significa esta região não possa ser empiricamente observada.

Como ilustração, na tabela 4.8 apresentamos as faixas de valores de retornos descritas pela q-Gaussiana estacionária, para as mesmas escalas temporais da tabela 4.7, porém medidas em uma mesma unidade de retornos, a do desvio padrão de $\Delta t = 1$ minuto (denotado a partir daqui por σ_1). O limite destas

faixas é obtido a partir da multiplicação do fator f mostrado nesta tabela pela largura em escala normalizada apresentada na tabela 4.7.

Δt (min)	$\sigma (\Delta t)$	$f = \sigma (\Delta t) / \sigma_1$	<i>bulk</i>
1	5,74 E-4	1	$5\sigma_1$
2	8,69 E-4	1,52	$6,3\sigma_1$
3	0,00112	1,95	$6,6\sigma_1$
4	0,00134	2,34	$6,8\sigma_1$
5	0,00154	2,69	$5,2\sigma_1$
6	0,00173	3,02	$4,8\sigma_1$
7	0,00190	3,32	$3,7\sigma_1$

Tabela 4.8 – Desvio padrão das distribuições de retorno $\sigma (\Delta t)$ em várias escalas temporais (coluna 2), respectivo fator f de alargamento em relação ao desvio padrão σ_1 de 1 minuto (coluna 3) e limite da região central da distribuição de retornos para o qual o valor ótimo é $q=1.75$ (coluna 4).

Pelos resultados mostrados na tabela 4.8, vemos que ainda é possível encontrar regiões de comportamento q -Gaussiano estacionário para escalas temporais ainda maiores.

Apresentamos nas figuras a seguir algumas distribuições empíricas $P_{\Delta t}(x)$ de retornos reescalados, isto é, medidos em escala de σ_1 e diminuídos da média, e as respectivas q -Gaussianas estacionárias. As figuras 4.31 e 4.32 ilustram que $\ln P_{\Delta t}(x)$ é linear para valores de x grandes, de acordo com a equação (4.2), e quadrático para x pequenos, de acordo com a equação (2.18).

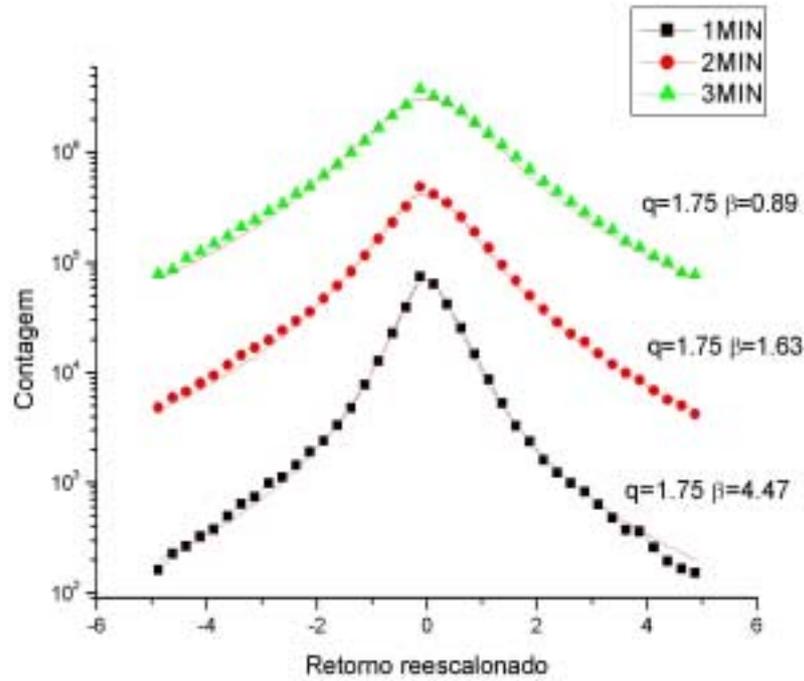


Figura 4.31 – Distribuição de retorno reescalonado para escalas de tempo $\Delta t=1,2,3$ minutos e q-Gaussiana ótima com parâmetro $q=1,75$. Os gráficos foram deslocados para melhor visualização.

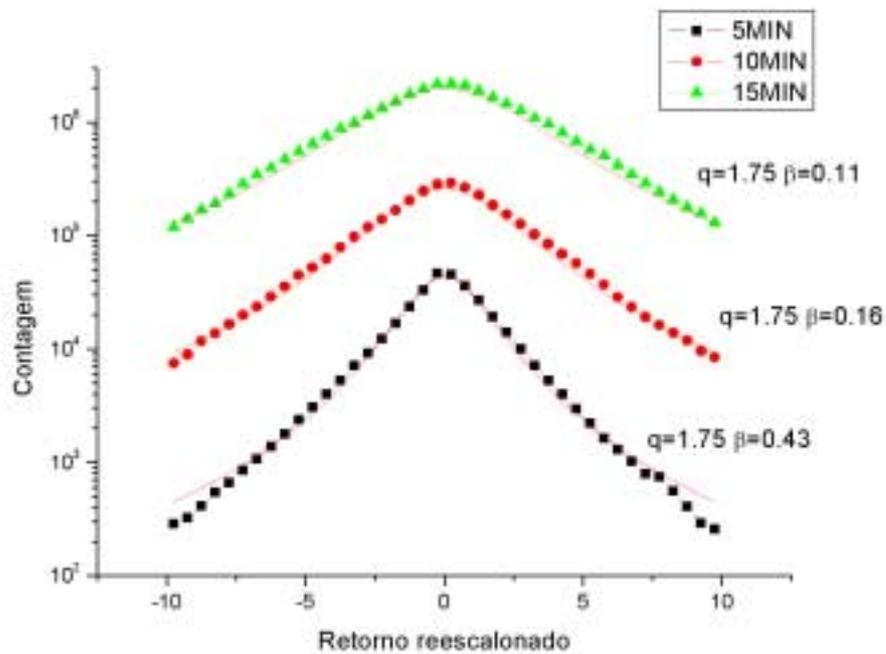


Figura 4.32– Distribuição de retorno reescalonado para escalas de tempo $\Delta t=5,10,15$ minutos e q-Gaussiana ótima com parâmetro $q=1,75$. Os gráficos foram deslocados para melhor visualização.

Para escalas de tempo mesoscópicas ($\Delta t > 20$ minutos), esta quase-estacionariedade é quebrada, não sendo possível o ajuste acurado com $q=1,75$ no centro das distribuições devido ao predomínio do comportamento exponencial, como pode ser observado na figura 4.33:

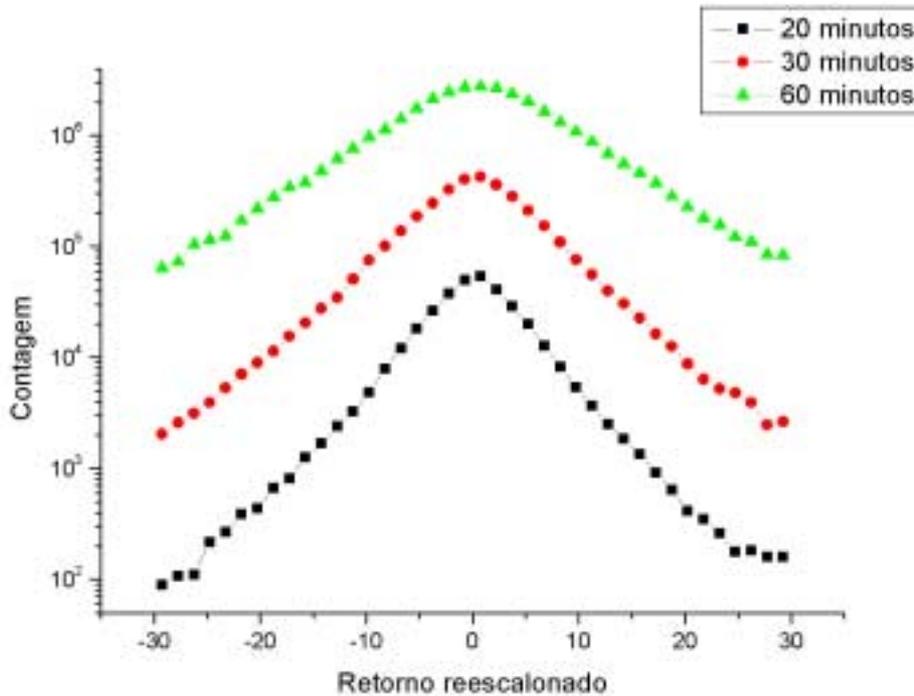


Figura 4.33– Distribuição de retorno reescalonado para escalas de tempo $\Delta t=20,30,60$ minutos. Os gráficos foram deslocados para melhor visualização.

Os resultados apresentados até aqui permitem a utilização da hipótese de estacionariedade de q -Gaussianas apenas em escalas de tempo microscópicas ($\Delta t < 20$ minutos). Como as correlações são mais fortes quanto menor for o intervalo de tempo, e em particular, de acordo com o resultado mostrado na figura 4.27, a correlação linear dura cerca de 20 minutos, vamos concentrar a análise da dinâmica intradiária nesta região de escala temporal.

Como verificado no capítulo 3, as distribuições de Tsallis com $q > 1$ são soluções da E.F.P. não-linear (3.29), que descreve uma dinâmica superdifusiva. Para identificação dos parâmetros desta modelagem com os dados empíricos obtidos, considera-se como variável estocástica x a variação logarítmica dos preços (retornos) e como variável t o intervalo Δt das medidas de retorno, em minutos. A partir dos parâmetros $\beta(\Delta t)$ ótimos obtidos para as distribuições empíricas reescaloadas (ver tabela 4.9), verificamos a consistência de nossos

dados com o comportamento previsto para $\beta(t)$ segundo esta modelagem, dada pela equação (3.44c).

Δt (min)	$\beta_{\text{REESCALONADO}}(\Delta t)$
1	4.47
2	1.63
3	0.89
4	0.58
5	0.43
6	0.32
7	0.26
8	0.21
9	0.18
10	0.16
15	0.11
20	0.07

Tabela 4.9 – Parâmetro β (Δt) das distribuições de retornos reescalados em várias escalas temporais microscópicas.

A evolução temporal das distribuições empíricas de retornos reescalados foi ajustada por (3.44c) usando ao todo dois parâmetros livres. O parâmetro não-extensivo q foi fixado no valor estacionário $q^*=1,75$, obtido pelo ajuste direto das distribuições empíricas nas escalas de tempo ultra-curtas ($t < 20$ minutos). O instante inicial considerado é $t_0=1$. Portanto, $\beta(t_0)$ é o parâmetro de escala da distribuição de 1 minuto. $Z(t_0)$ é obtido a partir de $\beta(t_0)$ usando o vínculo de normalização da distribuição que leva à (2.16). Logo, os dois parâmetros de controle para a evolução temporal são: a constante de difusão D e a taxa de reversão à média b . O resultado está mostrado na figura (4.34).

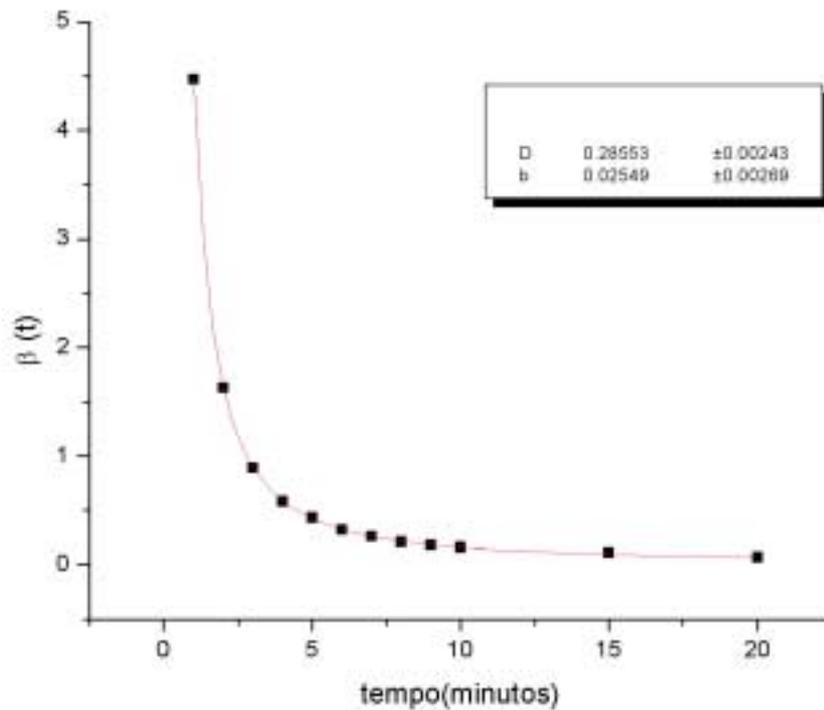


Figura 4.34 – Evolução temporal do parâmetro $\beta(t)$ para os dados do IBOVESPA de novembro 2002 a junho 2004 e previsão de comportamento segundo a E.F.P. não-linear (3.29) ótima.

O ajuste ótimo segundo (3.44c) fornece $D = 0,286 \pm 0,002$ e $b = 0,025 \pm 0,003$ para o mercado intradiário brasileiro, em escalas de tempo microscópicas. O valor do parâmetro de reversão à média a é nulo, uma vez que a média dos retornos reescalados é nula por construção.

Finalmente, através da solução (3.41) e (3.44a), vê-se que o parâmetro τ definido em (3.42a) representa o tempo característico do processo difusivo para atingir a solução de equilíbrio. Substituindo-se os valores de ajuste q^* e b , encontra-se $\tau = 32$ minutos. Este resultado é consistente com a forte dependência temporal de $\beta(t)$ para $t < 20$ minutos. Além disso, o valor β^* da solução de equilíbrio, dada por (3.45), para o nosso mercado corresponde ao valor $\beta^* \approx 0.02$, valor que ainda não foi atingido nas escalas temporais analisadas, de acordo com a tabela 4.9, consistente com o valor de τ encontrado.

Dos resultados apresentados podemos concluir que a dinâmica do mercado de ações brasileiro em escalas de tempo ultra-curtas é bem descrita por uma dinâmica superdifusiva e correlacionada governada pela E.F.P. não-linear (3.19), com tempo força de reversão à média.