3
Modelo autoregressivo de duração condicional

3.1
Formulação

Seja x_i o tempo de duração dessazonado entre transações ocorridas entre os tempos t_{i-1} e t_i, ou seja, $x_i = t_i - t_{i-1}$, o qual é usualmente medido em segundos. Engle & Russel[3] sugerem que a dependência temporal das durações pode ser modelada pela esperança condicional $E(x_i|\mathcal{F}_{i-1})$[1], tal que $\frac{x_i}{E(x_i|\mathcal{F}_{i-1})}$ seja i.i.d (independente e identicamente distribuído).

O modelo Autoregressivo de Duração Condicional (ou ACD, sigla em inglês) descreve a duração observada como

$$x_i = \psi_i \epsilon_i, \tag{3-1}$$

onde ϵ_i é uma variável aleatória i.i.d. com $E(\epsilon_i) = 1$, sendo assim, $E(x_i|\mathcal{F}_{i-1}) = \psi_i$.

A segunda equação especifica um modelo para as durações condicionais ψ_i,

$$\psi_i = \omega + \sum_{j=1}^{p} \alpha x_{i-j} + \sum_{j=1}^{q} \beta \psi_{i-j} \tag{3-2}$$

com as seguintes restrições para os coeficientes: $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$ e $\alpha + \beta < 1$. A última restrição garante a existência da média incondicional da duração e as demais a positividade das durações condicionais.

O modelo mais simples é o ACD(1,1) cujas inovações seguem uma distribuição Exponencial de parâmetro $\lambda = 1$, por este motivo é chamado de EACD(1,1). Escreve-se o modelo EACD(1,1) da seguinte forma:

$$\begin{aligned} x_i &= \psi_i \epsilon_i, \\ \psi_i &= \omega + \alpha x_{i-1} + \beta \psi_{i-1} \end{aligned} \tag{3-3}$$

[1] $\mathcal{F}_{i-1}$ denota o conjunto de informação passada até o tempo t_{i-1}.

onde $\epsilon_i \sim \exp(1)$. Usando-se os momentos da exponencial temos que $E(\epsilon_i) = 1$, $Var(\epsilon_i) = 1$, e $E(\epsilon_i^2) = Var(x_i) + E^2(x_i) = 2$. Assumindo que x_i seja fracamente estacionário (i.e., os dois primeiros momentos de x_i são invariantes no tempo), podemos derivar a variância de x_i. Primeiramente, aplicamos o operador média na equação (3-3) e obtemos:

$$
\begin{aligned}
E(x_i) &= E(E(\psi_i \epsilon_i | \mathcal{F}_{i-1})), \\
E(\psi_i) &= \omega + \alpha E(x_{i-1}) + \beta E(\psi_{i-1})
\end{aligned}
\tag{3-4}
$$

Sob a hipótese de estacionaridade fraca temos $E(\psi) = E(\psi_{i-1}) = \mu_x$. Portanto,

$$
\mu_x \equiv E(x_i) = E(\psi_i) = \frac{\omega}{1 - \alpha - \beta} \tag{3-5}
$$

Então, $E(x_i^2) = E(E(\psi_i^2 \epsilon_i^2 | \mathcal{F}_{i-1})) = 2E(\psi_i^2)$.

Seguindo, aplicamos o operador esperança ao quadrado de ψ_i na equação (3-3) e usando a hipótese de estacionaridade fraca temos:

$$
E(\psi_i^2) = \mu_x^2 \times \frac{1 - (\alpha + \beta)^2}{1 - 2\alpha^2 - \beta^2 - 2\alpha\beta} \tag{3-6}
$$

Finalizando, temos que $Var(x_i) = E(x_i^2) - E^2(x_i)$ e $E(x_i^2) = 2E(\psi_i^2)$. Sendo assim,

$$
Var(x_i) = 2E(\psi_i^2) - \mu_x^2 = \mu_x^2 \times \frac{1 - \beta^2 - 2\alpha\beta}{1 - \beta^2 - 2\alpha\beta - 2\alpha^2} \tag{3-7}
$$

onde μ_x é definido na equação (3-5). Este resultado mostra que a variância incondicional é invariante no tempo, onde $1 > 2\alpha^2 + \beta^2 + 2\alpha\beta$. A variância do modelo cujas inovações são Weibull ou Gama pode ser obtida usando o mesmo procedimento do modelo EACD(1,1).

3.2 Estimação

Nesta seção descreveremos como estimar o modelo ACD(1,1) cujas inovações seguem uma função de densidade Gama Generalizada (chamado GACD(1,1)) pois, segundo Zhang, Russel & Tsay[9], tal distribuição consegue capturar o excesso de dispersão[2] observada nas séries de duração. A

[2]Dizemos que existe excesso de dispersão quando os dados apresentam maior variância do que o valor predito pela distribuição assumida. Isto pode ocorrer porque as observações são positivamente correlacionadas e apresentam pouco grau de independencia ou porque a verdadeira distribuição dos dados assume uma forma mais complexa. Quando o contrário ocorre, ou seja, os dados apresentam menor variância do que o valor predito pela

seguir apresentamos a função de densidade Gama Generalizada.

A variável aleatória X tem distribuição Gama Generalizada com parâmetros $\gamma > 0, \lambda > 0, \kappa > 0$ se a função de densidade é dada por:

$$f(y|\gamma,\kappa,\beta) = \begin{cases} \frac{\gamma y^{\kappa\gamma-1}}{\beta^{\kappa\gamma}\Gamma(\kappa)} \exp\left[-\left(\frac{y}{\beta}\right)^{\gamma}\right] & se\ y \geq 0 \\ 0 & caso\ c.c. \end{cases} \tag{3-8}$$

onde λ é o parâmetro de escala, α e κ são parâmetros de forma. A distribuição pode ser escrita como

$$G = \left(\frac{Y}{\beta}\right)^{\gamma} \tag{3-9}$$

onde G é uma variável aleatória Gama padrão com parâmetro κ ($G \sim Gama(\kappa)$). A função de densidade de Y pode ser obtida por uma parametrização de G, assim como seus momentos são dados por:

$$E(G^m) = \frac{\Gamma(\kappa+m)}{\Gamma(\kappa)} \tag{3-10}$$

Então, os momentos de Y são obtidos por:

$$E(Y^m) = E[(\beta G^{\frac{1}{\gamma}})^m] = \beta^m E(G^{\frac{m}{\gamma}}) = \frac{\beta^m \Gamma\left(\kappa + \frac{m}{\gamma}\right)}{\Gamma(\kappa)} \tag{3-11}$$

Quando $\kappa = 1$, a Gama Generalizada se reduz a uma Weibull e quando $\gamma = 1$ e $\kappa = 1$ se reduz a uma exponencial. Desta forma, a Exponencial e Weibull são casos especiais da função Gama Generalizada.

A esperança da distribuição Gama Generalizada é $E(Y) = \frac{\beta\Gamma(\kappa+\frac{1}{\gamma})}{\Gamma(\kappa)}$. Em modelos de duração a componente aleatória ϵ precisa ter igual a 1. Portanto, definimos uma nova variável aleatória $\epsilon = X = \frac{\lambda Y}{\beta}$, onde $\lambda = \frac{\Gamma(\kappa)}{\Gamma(\kappa+\frac{1}{\gamma})}$, desta forma, temos $E(X) = 1$ e a função de densidade:

$$f(x|\gamma,\kappa) = \begin{cases} \frac{\gamma x^{\kappa\gamma-1}}{\lambda^{\kappa\gamma}\Gamma(\kappa)} \exp\left[-\left(\frac{x}{\lambda}\right)^{\gamma}\right] & se\ x \geq 0 \\ 0 & caso\ c.c. \end{cases} \tag{3-12}$$

Assim, a função de log-verossimilhança resulta em:

$$\ell = \sum_{i=1}^{N} \left[log(\gamma) + (\kappa\gamma - 1) log\left(\frac{x_i}{\lambda\psi_i}\right) - log(\lambda\psi_i\Gamma(\kappa)) - \left(\frac{x_i}{\lambda\psi_i}\right)^{\gamma} \right] \tag{3-13}$$

onde

distribuição assumida, dizemos que há sub-dispersão.

$\psi_i = \omega + \sum_{j=1}^{p} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-1}$ para $i > 1$ e $\psi_0 = \frac{\omega}{(1-\beta)}$ e $\lambda = \frac{\Gamma(\kappa)}{\Gamma(\kappa+\frac{1}{\gamma})}$.

3.3
Modelo de Volatilidade Instantânea

A fim de estimar a volatilidade utilizando os dados de negociação, primeiramente consideramos que os preços dos ativos são medidos por um 'preço médio', o qual é a média do preço de venda e compra no instante da negociação, conforme dito anteriormente. Esta medida reduz o número de casos de preços discrepantes.

O modelo é formulado com as durações seguindo um processo do tipo ACD, e os preços seguindo um processo GARCH (*Generalized Autoregressive Conditional Heteroscedastic*). Assim, a volatilidade é dependente dos movimentos dos preços e das durações.

Seja r_i o retorno geométrico entre os negócios i e $i-1$. Definida h_i como variância condicional por negociação:

$$h_i = V(r_i | x_i, \mathcal{F}_{t-1}) \tag{3-14}$$

Observa-se que esta variância é condicional às durações contemporâneas e também à variação de preços passados.

Adicionalmente defini-se a variância condicional por segundo para o i-ésimo negócio da seguinte forma:

$$V(\frac{r_i}{\sqrt{x_i}} \mid x_i, \mathcal{F}_{t-1}) = \sigma_i^2 \tag{3-15}$$

Assim, a relação entre (3-14) e (3-15) é

$$h_i = x_i \sigma_i^2 \tag{3-16}$$

A volatilidade por unidade de tempo é modelada como um processo GARCH. Engel & Russel [3] propõem um modelo ARMA(p,q) para a série $\frac{r_i}{\sqrt{x_i}}$ na tentativa de filtrar a autocorrelação dos retornos induzida pela Microestrutura do Mercado, vide Engel[**?**].

Se as durações não possuem informação sobre a variância por unidade de tempo, então, o modelo GARCH(1,1) para dados irregularmente espaçados no tempo é simplesmente:

$$\sigma_i^2 = \omega + \alpha e_{i-1}^2 + \beta \sigma_{i-1}^2 \tag{3-17}$$

onde e_{i-1}^2 é a inovação da série $\frac{r_i}{\sqrt{x_i}}$.

Modelos mais gerais são propostos baseados nos modelos teóricos de Easley e O'hara *apud* Engle & Russel [10], Admati e Pleiderer *apud* Engle & Russel [10]. Engle[8] sugere a seguinte especificação:

$$\sigma_i^2 = \omega + \alpha e_{i-1}^2 + \beta \sigma_{i-1}^2 + \gamma_1 x_i + \gamma_2 \frac{x_i}{\psi_i} + \gamma_3 \psi_i^{-1} + \gamma_4 \xi_{i-1} \tag{3-18}$$

onde ψ_i é a duração esperada obtida do modelo ACD e ξ_{i-1} caracteriza a persistência calculada via amortecimento exponencial do quadrado do retorno por unidade de tempo ($\xi_{i-1} = (1-\varrho)r_{i-1}^2 + \varrho\xi_{i-1}$, onde $\varrho = 0.995$, sugerido por Engel[8]). O impacto das durações na volatilidade é incorporado pelas variáveis $\frac{x_i}{\psi_i}$, x_i e ψ_i^{-1}, as quais são o efeito de causalidade entre a volatilidade e as durações, o impacto contemporâneo e o efeito inversamente proporcional médio, respectivamente.

No capítulo subseqüente, apresentamos a estimativa da volatilidade instantânea proposta por Engle[10], ver equação (3-18), e como esta é impactada pelas durações.