

Econometria e Séries Temporais - Aula 7 -

Prof. Mestre. Omar Barroso

Instituto Brasileiro de Educação, Pesquisa e Desenvolvimento

ARCH

- O modelo ARCH significa heterocedasticidade condicional auto-regressiva (**AutoRegressive Conditional Heteroskedasticity**).
- O ARCH é um modelo estatístico usado em econometria para analisar dados de séries temporais, particularmente aqueles que exibem períodos de agrupamento de volatilidade — onde grandes mudanças tendem a ser seguidas por grandes mudanças (de qualquer sinal), e pequenas mudanças tendem a ser seguidas por pequenas mudanças.

Principais características (ARCH)

- **Heterocedasticidade:** O modelo ARCH é projetado para abordar a **volatilidade variável no tempo** (heterocedasticidade), onde a **variância dos termos de erro não é constante**, mas varia ao longo do tempo.
- **Variância condicional:** Em um modelo ARCH, a variância do termo de erro atual é modelada como uma função dos termos de erro ao quadrado de períodos anteriores. **Isso significa que os erros passados influenciam a volatilidade de hoje.**

Heterocedasticidade no ARCH

- Na econometria, Heteroscedasticidade se refere ao fenômeno em que a variância dos termos de erro (resíduos) em um modelo de regressão não é constante entre observações.
- No contexto de dados de séries temporais, isso geralmente se manifesta como períodos de alta volatilidade (**grandes flutuações**) alternando com períodos de baixa volatilidade (**pequenas flutuações**), um padrão conhecido como agrupamento de volatilidade.

Heterocedasticidade no ARCH

- Em modelos tradicionais (sem o ARCH), a heterocedasticidade é frequentemente vista como um problema, implicando ineficiência ou viés nas estimativas do modelo.
- No entanto, o modelo ARCH incorpora diretamente a heterocedasticidade em sua estrutura, permitindo que a variância dos erros dependa de observações passadas.

Heterocedasticidade no ARCH

- Supondo, que temos um modelo de série temporal da determinada forma:
- $y_t = \mu + \epsilon_t$
- No qual,
- y_t = A variável dependente em 't'
- μ = a média constante ou **condicional** sob outras variáveis
- A média condicional opera no qual existem informações de outras variáveis média.
- $E[\mu_t | y_{t-1}, \dots, y_{t-n}]$
- ϵ_t = O termo de erro.

O termo de erro no ARCH

- Podemos denominar a variância da determinada maneira:
- $\epsilon = \sigma_t^2 z_t$
- σ_t^2 : é a variância condicional variável no tempo de ϵ .
- z_t : é o ruído branco, tipicamente assumido em seguir uma distribuição normal e uma variância unitária $z_t \sim N(0,1)$.

O termo de erro no ARCH

- O modelo ARCH supõe que a variância σ_t^2 é uma função de erros passados.

$$\mathbb{E}[e_t | I_{t-1}] = \mathbb{E}[\rho e_{t-1} | I_{t-1}] + \mathbb{E}[v_t] = \rho e_{t-1}$$

- Desta maneira,
- $\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2 + \dots + \alpha_q \epsilon_{t-q}^2$
- Se houver grandes resíduos (ou seja, grandes erros) em períodos anteriores σ_t^2 , eles serão altos, indicando alta volatilidade no período atual.
- Se resíduos previos forem baixos, σ_t^2 a variância também será (implica uma baixa volatilidade).

Heterocedasticidade manifestada

- **Agrupamento da volatilidade:** Períodos de grandes erros tendem a se agrupar.
- Isso ocorre quando ϵ_t^2 é grande, ele aumenta a volatilidade σ_t^2 , o que significa que o erro atual ϵ_t provavelmente também terá alta volatilidade.
- **Incerteza Variável no Tempo:** O nível de incerteza na série muda ao longo do tempo, dependendo de erros passados. Isso cria uma série temporal com níveis variáveis de risco, onde períodos de alto risco seguem outros períodos de alto risco, e períodos de baixo risco seguem períodos de baixo risco.

Multiplicador Lagrangiano \leftrightarrow LM

- O teste do Multiplicador de Lagrange (LM) para o modelo ARCH, é um procedimento estatístico usado para detectar a **presença de heterocedasticidade** condicional autorregressiva (ARCH) **nos resíduos** de um modelo de série temporal.
- Ele ajuda a identificar se existe volatilidade variável no tempo, o que justifica o uso de um modelo ARCH para um determinado conjunto de dados.

Multiplicador Lagrangiano \leftrightarrow LM

- Em séries temporais econômicas/financeiras (como retornos de ações), a volatilidade (variância) frequentemente muda ao longo do tempo, agrupando-se em certos períodos.
- De acordo com Engle (1982), modelo ARCH, modela essa volatilidade variável no tempo permitindo que a **variância dos termos de erro** (ou resíduos) **dependa de termos de erro quadrático passados**. Antes de ajustar um modelo ARCH, é importante testar se tal heterocedasticidade condicional existe nos dados. É aqui que entra o teste LM.

Multiplicador Lagrangiano \leftrightarrow LM

- H_0 : Não há efeito ARCH, ou seja, os resíduos têm uma variância constante (homocedástico).
- H_a : Existe um efeito ARCH (heterocedasticidade), significando que a variância dos resíduos depende de erros passados (heterocedástico).

Multiplicador Lagrangiano (Teoria)

- Primeiramente, vamos estimar a série temporal dos dados para obter os resíduos $\hat{\epsilon}_t$.
- Segundo, vamos elevar os resíduos ao quadrado para observar se existe dependência dos dados passados em relação ao presente ϵ_t^2 .
- $\epsilon_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2 + \dots + \alpha_p \epsilon_{t-p}^2 + u_t$

Multiplicador Lagrangiano (Teoria)

- A estatística de teste é baseada no R-quadrado desta regressão auxiliar. Especificamente, a estatística de teste LM é:
- $LM = T \cdot R^2$
- Onde T é o número de observações e R² é o coeficiente de determinação (variabilidade de Y explicada por X) da regressão do último slide.

Multiplicador Lagrangiano (Teoria)

- O ML segue uma distribuição Chi-quadrado (χ^2).
- Se a estatística LM for maior que o valor crítico da distribuição qui-quadrado para um determinado nível de significância (por exemplo, 5%), você rejeita a hipótese nula e conclui que os efeitos ARCH estão presentes nos resíduos.

Decisão LM

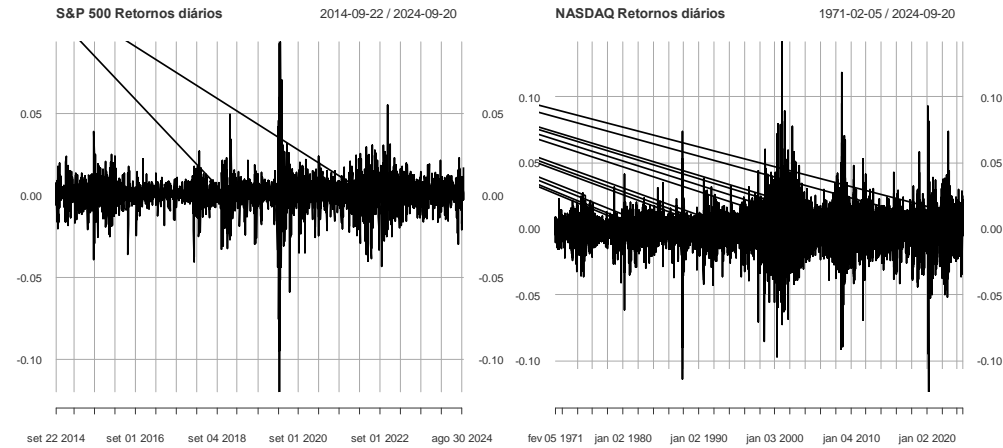
- H_0 : Se o valor crítico $LM \geq \chi^2$ não existem evidências de variância ao longo do tempo.
- H_0 : Se o valor crítico $LM \leq \chi^2$ existem evidências de variância ao longo do tempo.
- Por convêniencia o valor crítico utilizado sera do p-valor (0.05).

Testando o ARCH

- Vamos importar da Biblioteca FRED as séries temporais da S&P 500 e Nasdaq.

Testando o ARCH

- Testando o retorno dos índices financeiros, podemos sugerir que existe uma clara evidência de variância nos dados.



Teste LM

- Conforme os resultados do teste LM, ambas as séries apresentam heterocedasticidade.
- Podemos rejeitar H_0 .
- ARCH LM-test; Null hypothesis: no ARCH effects
- data: sp500_returns
- Chi-squared = 1001.2, df = 12, p-value < $2.2e-16$
- ARCH LM-test; Null hypothesis: no ARCH effects
- data: nasdaq_returns
- Chi-squared = 2975.6, df = 12, p-value < $2.2e-16$

Testando o ARCH (S&P 500) – Parâmetros

- (Intercept – a_0): $6.716e-05$. Isso representa o termo constante no modelo, indicando o nível médio de volatilidade de longo prazo.
- O pequeno erro padrão ($1,718e-06$) e o valor t extremamente alto (39,10) sugerem que a_0 é altamente significativo estatisticamente (valor $p < 2e-16$).
- Coefficient(s):
 - Estimate Std. Error t value Pr(>|t|)
 - a_0 $6.716e-05$ $1.718e-06$ 39.10 $<2e-16$ ***
 - a_1 $4.755e-01$ $2.518e-02$ 18.88 $<2e-16$ ***

Testando o ARCH (S&P 500) – Parâmetros

- a_1 (coeficiente ARCH(1)): $4,755e-01$. Isso representa a sensibilidade da volatilidade de hoje aos retornos quadrados do período anterior.
 - O valor t (18,88) também é grande, e o valor p ($< 2e-16$) mostra que o coeficiente é estatisticamente significativo. Isso indica que os retornos quadrados passados impactam significativamente a volatilidade atual.
 - $a_1 = 0.475$ representa que aproximadamente 47.5% da volatilidade atual depende da volatilidade do dia anterior.
- Coefficient(s):

	Estimate	Std. Error	t value	Pr(> t)
a0	$6.716e-05$	$1.718e-06$	39.10	$<2e-16$ ***
a1	$4.755e-01$	$2.518e-02$	18.88	$<2e-16$ ***

Testando o ARCH (S&P 500) – Resíduos

- Os resíduos mostram uma distribuição ampla, com a maioria dos resíduos centralizados em torno de zero. Os resíduos máximos e mínimos são bem grandes, o que pode indicar extrema volatilidade nos dados.

- Residuals:

•	Min	1Q	Median	3Q	Max
•	-6.56612	-0.39910	0.06779	0.59729	5.46510

Testando o ARCH (S&P 500) – Resíduos

- O teste Jarque-Bera verifica a normalidade nos resíduos.
- Valor $p < 2,2e-16$ significa que os resíduos se desviam significativamente de uma distribuição normal, sugerindo a presença de **assimetria ou curtose** nos dados. Isso é típico em séries temporais financeiras, onde os retornos geralmente têm caudas grossas e não são distribuídos normalmente.
- Jarque Bera Test
- data: Residuals
- X-squared = 1374.3, df = 2, p-value < 2.2e-16

Testando o ARCH (S&P 500) – Resíduos

- O teste Box-Ljung verifica a correlação serial nos resíduos quadrados.
- $p\text{-valor} = 0,01205$ sugere que ainda há alguma autocorrelação nos resíduos quadrados, o que indica que o modelo ARCH(1) pode não capturar totalmente a volatilidade variável no tempo nos dados. Isso pode sugerir que um modelo mais complexo (por exemplo, um modelo ARCH ou GARCH de ordem superior) pode fornecer um ajuste melhor.
- Box-Ljung test
- data: Squared.Residuals
- X-squared = 6.3033, df = 1, p-value = 0.01205

Testando o ARCH (S&P 500)

- O modelo ARCH(1) é estatisticamente significativo para os retornos do S&P 500, conforme indicado pelo coeficiente a_1 significativo.
- A volatilidade condicional (efeito ARCH) é influenciada por 47,5% dos retornos quadrados do dia anterior, confirmando a presença de volatilidade variável no tempo.

Testando o ARCH (S&P 500)

- No entanto, os testes diagnósticos residuais sugerem que os resíduos não são distribuídos normalmente (teste Jarque-Bera) e que alguma autocorrelação ainda permanece nos resíduos quadrados (teste Box-Ljung).
- Isso pode indicar que um modelo mais sofisticado como um GARCH(1,1) ou um modelo ARCH de ordem superior poderia fornecer um ajuste melhor para os dados (*próxima aula!!*)

Testando o ARCH (NASDAQ)

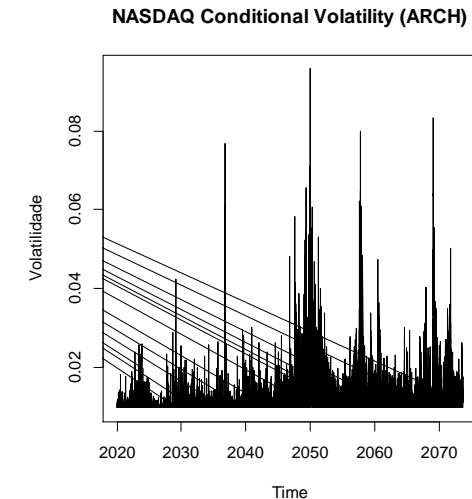
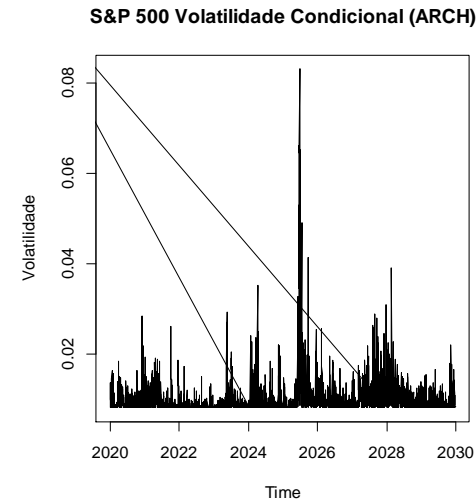
- Agrupamento de volatilidade: semelhante ao S&P 500, os retornos do NASDAQ exibem um agrupamento de volatilidade significativo, onde uma grande parte da volatilidade atual é explicada por retornos quadrados passados. No entanto, o nível de base de volatilidade (intercepto a_0) é maior para o NASDAQ.
- Movimentos extremos: os resíduos do NASDAQ mostram extremos maiores (valores mínimos/máximos mais altos), refletindo que o NASDAQ experimentou oscilações mais drásticas nos retornos do que o S&P 500.

Testando o ARCH (NASDAQ)

- Não normalidade: ambos os mercados mostram forte não normalidade nos resíduos (teste Jarque-Bera), mas a não normalidade do NASDAQ é mais pronunciada, indicando que os retornos têm caudas grossas ainda mais pronunciadas.
- Ajuste do modelo: ambos os modelos exibem alguma autocorrelação nos resíduos quadrados (teste Box-Ljung), mas o modelo NASDAQ sugere uma presença mais forte de autocorrelação, o que implica que o modelo ARCH(1) pode ser menos adequado para o NASDAQ do que para o S&P 500.

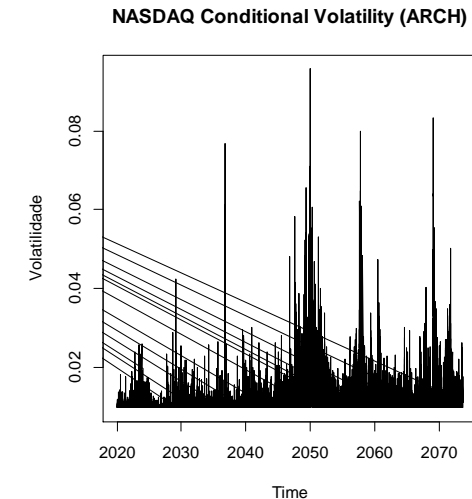
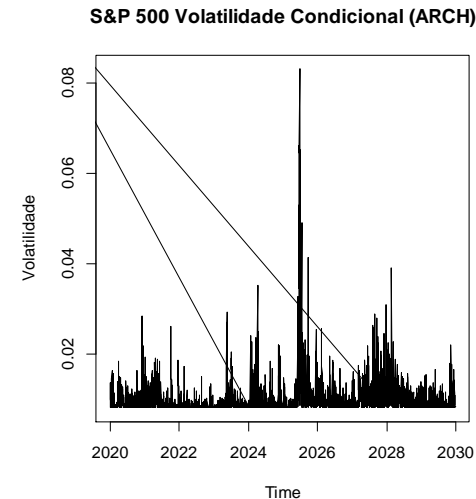
Volatilidade Condicional – S&P 500

- Mín (0,0082): O menor nível de volatilidade condicional durante o período foi de cerca de 0,0082, indicando que o mercado teve períodos de relativa calma com baixa volatilidade.
- Mediana (0,0089): A volatilidade mediana foi ligeiramente maior, em torno de 0,0089, o que sugere que em um dia de negociação médio, o S&P 500 teve esse nível de risco.
- Média (0,0104): A volatilidade média é ligeiramente maior que a mediana, indicando que houve períodos de maior volatilidade que empurraram a média acima da tendência central (mediana).
- Máx (0,0830): O valor máximo (0,083) indica um pico significativo na volatilidade. Isso provavelmente corresponde a um período de turbulência do mercado, possivelmente durante uma quebra do mercado ou eventos econômicos inesperados (como a crise da COVID-19 em 2020).



Volatilidade Condicional – NASDAQ

- Min (0,0097): A volatilidade condicional mínima para o NASDAQ é maior do que para o S&P 500, sugerindo que o NASDAQ tem experimentado uma volatilidade relativamente mais persistente, mesmo em seus períodos mais calmos.
- Mediana (0,0104): A mediana também é maior do que a do S&P 500, indicando que o NASDAQ geralmente tem maior volatilidade ao longo do período, o que pode ser devido ao fato de que as ações do NASDAQ são mais pesadas em tecnologia e, portanto, inerentemente mais arriscadas.
- Média (0,0121): A volatilidade média do NASDAQ também é maior do que a do S&P 500. Isso reforça a ideia de que o NASDAQ é um índice mais arriscado, com mais flutuações nos retornos.
- Máx. (0,0957): A volatilidade máxima do NASDAQ é significativamente maior do que a do S&P 500. Isso aponta para um evento extremo em que o NASDAQ enfrentou um grande pico de volatilidade, provavelmente durante crises ou choques de mercado, talvez até maior do que a reação do S&P 500.



Perguntas p/ Revisão

- Como podemos brevemente definir a Heterocedasticidade?
- O que é uma variância condicional e como ela afeta a volatilidade de hoje?
- O que é o agrupamento de volatilidade?
- Como podemos implementar o ARCH e como podemos interpretar os resultados?

Referência(s)

- Engle, R.F., 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4), pp.987-1007.