

**Testes para Avaliação das Previsões
do Valor em Risco**

Jaime Enrique Lincovil Curivil

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
MESTRE EM CIÊNCIAS

Programa: Estatística
Orientador: Prof^ª. Dr^ª. Chang Chiann

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CNPq

São Paulo, Mayo de 2015

Testes para Avaliação das Previsões do Valor em Risco

Esta versão da dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho, realizada em 27/02/2015. Uma cópia da versão original está disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão Julgadora:

- Prof^a Dr^a Chang Chiann (orientadora) -IME-USP
- Prof^a Dr^a Airlane Pereira Alencar -IME-USP
- Prof. Dr. Guilherme de Oliveira Lima Cagliari Marques - UFABC

Agradecimentos

Em primeiro lugar agradeço à Universidade de São Paulo pela oportunidade de fazer o mestrado no Brasil. Em segundo lugar, agradeço infinitamente a minha professora orientadora, pela paciência e dedicação nas correções feitas no trabalho. Em terceiro lugar, agradeço a todos meus amigos chilenos, brasileiros e de outras nacionalidades pelo apoio moral.

Resumo

Curivil, J.E.L. **Testes para avaliação das previsões do Valor em Risco**. 2015. 56 f. Dissertação (Mestrado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2015.

Neste trabalho, apresentamos alguns métodos para avaliação das previsões do Valor em Risco (VaR). Estes métodos testam um tipo de eficiência, denominada cobertura condicional correta. O poder empírico e a probabilidade do erro de tipo I são comparados através de simulações de Monte Carlo. Além disso, avaliamos um novo método de previsão do VaR, o qual é aplicado nos retornos diários do Ibovespa. Os resultados obtidos mostram que a nova classe de testes, baseados em uma regressão Weibull discreta, em muitos casos, tem poder empírico maior comparando com outros métodos apresentados neste trabalho.

Palavras-chave: Valor em Risco, Teste de avaliação, Cobertura Condicional, Poder empírico.

Abstract

Curivil, J.E.L. **Backtesting for Value at Risk Models**. 2015. 56 f. Tesis (Master) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2015.

In this paper, we present some procedures for assessing forecasts for the Value at Risk (VaR). These procedures test a type of efficiency, referred as correct conditional coverage. The empirical power and type I error probability are compared through a Monte Carlo simulation. The results show that a new class of tests based on a discrete Weibull regression in most cases has greater power empirical to other methods available in this paper.

Keywords: Value at Risk, Backtesting , Condition Coverage, Empirical power.

Sumário

Lista de Abreviaturas	ix
Lista de Símbolos	xi
Lista de Figuras	xiii
Lista de Tabelas	xv
1 Introdução	1
1.1 Objetivos	2
1.2 Contribuições	2
1.3 Organização do trabalho	3
2 Conceitos e definições básicas.	5
2.1 Retornos	5
2.1.1 Fatos Estilizados dos Retornos	5
2.2 Valor em Risco (VaR)	6
2.3 Teoria dos Valores Extremos	6
2.4 Teste da Razão de Verossimilhanças Generalizada (TRVG)	7
2.5 Sequência de Previsões	8
2.6 Análise de Tempo de Falha Discreto	9
3 Métodos de previsão do Valor em Risco.	11
3.1 Introdução	11
3.2 Estimação do VaR utilizando a Simulação Histórica (SH)	11
3.3 Previsão do VaR utilizando os modelos ARMA e GARCH	11
3.4 Previsão do VaR utilizando a Teoria dos Valores Extremos	12
3.5 Previsão do VaR utilizando o Método POT	13
3.6 O método POT baseado em Durações entre Excessos (DPOT)	14
3.7 Agrupamento das Violações do VaR	15
4 Testes para avaliação de previsão do VaR	17
4.1 Introdução	17
4.2 Eficiência dos Intervalos de Previsão	17
4.3 Testes Baseados na Sequência I_T	19
4.3.1 Teste do Tipo Bernoulli	19

4.3.2	Ampliação da Informação Passada	21
4.3.3	Teste de Quantis Dinâmicos	22
4.3.4	Testes do Tipo CAViaR	23
4.4	Testes Baseados em Durações	26
4.4.1	Testes Baseados na Distribuição Weibull Discreta	26
4.4.2	Testes usando Modelos de Regressão Weibull Discretos	29
4.4.3	Testes baseados no Método dos Momentos Generalizado (MMG)	32
4.5	Considerações finais	34
5	Simulações e Aplicações	35
5.1	Introdução	35
5.2	Estudo de Simulação Comparativa	35
5.2.1	Probabilidade do Erro de Tipo I e Poder Empírico	36
5.3	Aplicações	40
6	Conclusões e Trabalhos Futuros	45
6.1	Conclusões	45
6.2	Trabalhos futuros	45
A	Apêndice	47
A.1	Domínio Máximo de Atração	47
A.2	P-valor de Monte Carlo	47
A.3	Programação em R: Teste Geométrico VaR	47
	Referências Bibliográficas	51

Lista de Abreviaturas

TRVG	Teste da Razão de Verossimilhanças Generalizada.
VaR	Valor em Risco.
POT	Peaks - over - threshold.
SH	Simulação Histórica.
TVE	Teoria dos Valores Extremos.
DGP	Distribuição Generalizada de Pareto.
MMG	Método dos Momentos Generalizado.
EMV	Estimador de Máxima Verossimilhança.
GARCH	Generalized autoregressive conditional heteroskedasticity.
APARCH	Asymmetric power autoregressive conditional heteroskedasticity.

Lista de Símbolos

$\mathcal{L}(\theta, \mathbf{X})$	A função de verossimilhança de θ correspondente à amostra observada \mathbf{X} .
$\log\mathcal{L}(\theta, \mathbf{X})$	A função de log-verossimilhança de θ correspondente à amostra observada \mathbf{X} .
χ_k^2	Variável aleatória qui-quadrado com k graus de liberdade.
\xrightarrow{D}	Convergência em Distribuição.
\mathcal{R}	Conjunto dos números reais.
\mathcal{N}	Conjunto dos números naturais.
\mathcal{N}^*	Conjunto dos números naturais não-nulos.

Lista de Figuras

3.1	Agrupamento das violações do VaR	16
5.1	Gráfico do Ibovespa e os log-retornos no período de 04/07/1994 a 07/08/02	41
5.2	Pós-amostra e previsões do VaR utilizando os métodos POT e DPOT	41

Lista de Tabelas

5.1	Probabilidade empírica do erro de tipo I, para os testes de independência. PGD 2, $\alpha^* = 0,1, p = 0,05$	37
5.2	Probabilidade empírica do erro de tipo I, para os testes de cobertura condicional. PGD 2, $\alpha^* = 0,1, p = 0,05$	37
5.3	Poder empírico para os testes de independência. PGD 1, $\alpha^* = 0,1, ws = 250$	38
5.4	Poder empírico para os testes de independência. PGD 2, $\alpha^* = 0,1, p = 0,05$	38
5.5	Poder empírico para os testes de cobertura condicional. PGD 1, $\alpha^* = 0,1, ws = 250$	39
5.6	Poder empírico para os testes de cobertura condicional. PGD 2, $\alpha^* = 0,1, p = 0,05$	39
5.7	Poder empírico para os testes de cobertura condicional. PGD 3, $\alpha^* = 0,1, p = 0,01$	40
5.8	Poder empírico para os testes de cobertura condicional. PGD 3, $\alpha^* = 0,1, p = 0,05$	40
5.9	Estatísticas dos log-retornos do índice da Ibovespa	41
5.10	Avaliação das previsões do VaR gerada pelos métodos POT e DPOT.	42
5.11	Tabela de rejeições para $\alpha^* = 0,05$	42

Capítulo 1

Introdução

No contexto do mercado financeiro, uma atividade muito importante é a medição do risco. O Comitê de Supervisão Bancária de Basileia (BCBS, sigla de *Basel Committee on Banking Supervision*) é uma organização que congrega autoridades de supervisão bancária, que tem o objetivo de fortalecer o funcionamento dos sistemas financeiros. Em 1996, uma modificação ao acordo de Basileia, estabeleceu como medida básica para o risco de mercado o Valor em Risco (VaR).

A estimação do modelo de previsão do VaR apresenta grandes dificuldades, pois a distribuição condicional dos retornos é geralmente desconhecida e difícil de modelar. Na literatura, várias abordagens foram desenvolvidas para estimar o VaR. Alguns desses métodos se baseiam em pressuposições irreais; eles assumem que os retornos são independentes e identicamente distribuídos, ignorando os fatos estilizados como o agrupamento de volatilidade e variância dinâmica dependente do tempo. Os principais métodos para a previsão do VaR existentes na literatura são :

- Métodos não-paramétricos: uma das vantagens destes métodos é que não necessita de suposições sobre a distribuição do retorno. Os dois métodos mais utilizados são a Simulação Histórica e o Método de Estimação de Densidade. A Simulação Histórica é um método bastante popular, que utiliza a distribuição empírica de retornos como uma aproximação para a função de distribuição, assim o VaR é o quantil desta distribuição empírica, suas vantagens e desvantagens foram documentadas por Dowd [2007]. Uma outra opção é o Método de Estimação de Densidade, que utiliza um kernel para calcular os percentis ou estimar o VaR.
- Métodos Paramétricos: O primeiro modelo paramétrico para estimar o VaR foi *Riskmetrics* de Morgan [1996]. Posteriormente, surgiram os métodos baseados em modelos para volatilidade. Destacamos aqueles baseados (i) na família GARCH, (ii) os modelos de volatilidade estocástica e (iii) modelos de volatilidade realizada, ver Galdi e Pereira [2007].
- Métodos semi paramétricos: combinam as abordagens não-paramétrica e paramétrica. Alguns dos métodos mais importantes são (i) a Simulação Histórica Filtrada, apresentada em Barone-Adesi *et al.* [2002], (ii) o método de CAViaR de (Engle e Manganelli [2004]), (iii) o método do Máximo dos Blocos, utilizando a Teoria dos Valores Extremos de McNeil *et al.* [1998] e (iv) a abordagem do *peak-over-threshold* (POT) semi paramétrico de Danielsson *et al.* [1998] .

Dada a popularidade do VaR e a falta de uma abordagem eficaz, avaliar seu desempenho é de grande importância. Na prática, esta avaliação, geralmente, é realizada através de *testes*, denominados comumente na literatura como *backtesting*.

É importante destacar que tais testes avaliam o conjunto de previsões do VaR e não o modelo em si. Considerando uma certa definição de eficiência, os testes podem rejeitar a eficácia das previsões, mas serem inconclusivos com respeito à adequação do modelo.

Os métodos de avaliação do VaR podem ser considerados como casos particulares dos testes de avaliação de intervalos de previsão. No começo, alguns autores como McNees e Fine [1995] e Baillie e Bollerslev [1992] avaliaram esses intervalos comparando a frequência de violações (número

de observações fora do respectivo intervalo) e a probabilidade de cobertura p . Kupiec [1995] propôs uma formalização desse procedimento, construindo um procedimento de teste. Posteriormente, o trabalho proposto por Christoffersen [1998] foi de grande relevância, uma vez que formalizou uma definição de eficiência para intervalos de previsão, denominada *correta cobertura condicional*, além de propor procedimentos de testes dessa eficiência. Uma das primeiras aplicações desse procedimento foi a avaliação das previsões do VaR. No começo, esta avaliação foi baseada nas violações das previsões do VaR e as abordagens recentes utilizam as durações entre estas violações.

Com o passar do tempo tem surgido novas propostas de testes, mencionamos apenas aqueles mais relevantes em relação ao presente trabalho. Possivelmente, os primeiros autores que utilizaram a amostra de durações para avaliar intervalos de previsão foram Danielsson *et al.* [2000], eles propuseram um teste do tipo qui-quadrado para verificar a qualidade de ajuste. Com respeito à avaliação do VaR, Christoffersen e Pelletier [2004] propuseram a abordagem baseada em durações, considerando uma distribuição contínua. Haas [2005] prosseguiu com esta ideia, mas considerou uma distribuição discreta do tipo Weibull para as durações. Berkowitz *et al.* [2011] apresentaram uma ampla classe de testes, principalmente os testes de tipo CAViaR e o teste de tipo Geométrico. Candelson *et al.* [2010] apresentaram a abordagem baseada no Método dos Momentos Generalizado (MMG). Recentemente Pelletier e Wei [2014] apresentaram um novo método considerando o teste Geométrico, revendo a consideração da influência da previsão do VaR na probabilidade de acontecer uma violação.

Uma forma de validar os testes ao longo do desenvolvimento desta dissertação é a apresentação de estudos de simulação que evidenciem as propriedades dos testes sobre amostras finitas, por exemplo, o poder empírico.

O objetivo principal deste trabalho é comparar o poder da classe de testes apresentados em Pelletier e Wei [2014] com outros procedimentos, que tiveram um bom desempenho em estudos de simulações passados. Além disso, apresentaremos uma aplicação utilizando dados reais.

1.1 Objetivos

Os principais objetivos deste trabalho são:

1. Mostrar como uma sequência de previsões do VaR pode ser avaliada segundo a definição de eficiência de intervalos de previsão apresentada em Christoffersen [1998];
2. Apresentar métodos clássicos de avaliação das previsões do VaR e também uma nova classe de testes, derivada de Pelletier e Wei [2014];
3. Apresentar técnicas frequentemente utilizadas para calcular o poder empírico e a probabilidade empírica do erro de tipo I, dos testes apresentados;
4. Avaliar as previsões do VaR geradas pelos métodos *Peaks-over-threshold* e sua extensão apresentada em Araújo Santos e Fraga Alves [2013], utilizando os retornos diários do Ibovespa.

1.2 Contribuições

As principais contribuições do trabalho são:

1. Apresentação e implementação (no *software* R) de três novas opções para o teste proposto em Pelletier e Wei [2014];
2. A comparação do poder empírico dos testes selecionados fornece conclusões práticas, que podem ser útil na escolha dos testes de avaliação.

1.3 Organização do trabalho

No Capítulo 2, descrevemos os conceitos básicos utilizados no decorrer deste trabalho. No Capítulo 3, apresentamos e discutimos alguns modelos de previsão do VaR comumente utilizados. Os testes de avaliação das previsões são apresentados no Capítulo 4. No Capítulo 5 discutimos os resultados obtidos pelas simulações e a aplicação aos dados reais. Finalmente, no Capítulo 6, concluímos sobre os resultados obtidos e propomos possíveis extensões deste trabalho.

Capítulo 2

Conceitos e definições básicas.

Neste capítulo, apresentaremos alguns conceitos e definições que serão utilizados ao longo deste trabalho.

2.1 Retornos

O Risco do mercado financeiro está relacionado à mudança de preços dos ativos. Seja P_t o preço de um ativo no instante t , supondo que não haja dividendos pagos no período, definimos a variação de preços deste ativo do instante $t - 1$ até o instante t como :

$$\Delta P_t = P_t - P_{t-1},$$

e a variação relativa de preços deste ativo é dada por:

$$R_t = \frac{\Delta P_t}{P_{t-1}}.$$

Denominamos a $1 + R_t = \frac{P_t}{P_{t-1}}$ como *retorno bruto simples*, geralmente expressado em porcentagem. Seja $p_t = \log(P_t)$, o log-retorno é definido como (ao menos que se especifique o contrario \log é considerado o logaritmo com base e):

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(1 + R_t) = p_t - p_{t-1}.$$

Dado que para um x pequeno $\log(1 + x) \approx x$ segue que os retornos simples R_t e os log-retornos r_t são valores muito próximos.

2.1.1 Fatos Estilizados dos Retornos

As séries econômicas e financeiras têm características comuns a outras séries temporais como: tendência, sazonalidade, pontos influentes, heterocedasticidade condicional e não linearidade.

Os retornos financeiros apresentam algumas características particulares, entre elas os fatos estilizados. Os quais resumimos a seguir: (a) em geral, não são autocorrelacionados. (b) Os quadrados dos retornos são autocorrelacionados. (c) Apresentam agrupamentos de volatilidade ao longo do tempo; ou seja, em sua dinâmica durante um período prolongado a variância da série é elevada para depois ter uma queda prolongada. (d) Sua distribuição não condicional apresenta caudas mais pesadas do que uma distribuição normal, além disso, em geral é leptocúrtica. (e) Alguns retornos são não lineares, no sentido de responder de maneira diferente, por exemplo, a choques positivos ou negativos. Para maiores detalhes ver [Morettin \[2008\]](#).

2.2 Valor em Risco (VaR)

Neste trabalho estaremos interessados em calcular uma medida de um tipo particular de risco, o chamado risco de mercado financeiro ou Valor em Risco (VaR). Neste caso, o VaR é uma medida do grau de incerteza sobre retornos líquidos futuros segundo a posição financeira considerada, a qual será ou uma posição comprada ou vendida. Uma posição comprada significa possuir determinado ativo (ou carteira de ativos) e uma posição vendida envolve vender um ativo que não se possui, mas que é alugado.

Informalmente, o VaR é definido como uma medida da variação potencial máxima do valor de um ativo que quantifica quanto se pode perder sobre um período pré-fixado h e com dada probabilidade p . Para uma definição mais formal suponhamos que no instante t precisamos calcular o risco de uma posição financeira para o horizonte $h > 0$. Seja

$$\Delta P(h) = P(t+h) - P(t),$$

a variação dos ativos entre o instante t e $t+h$. A quantidade $\Delta P(h)$ representa o lucro ou a perda ($L\&P$) da posição sobre o horizonte h . Seja $F_h(x)$ a função de distribuição de $\Delta P(h)$, a definição do VaR para uma posição comprada é dada a seguir.

Definição 2.1 *Definimos o VaR de uma posição comprada sobre um horizonte h , com probabilidade p , $0 < p < 1$, por meio de:*

$$p = Pr(\Delta P(h) < VaR) = F_h(VaR).$$

Notemos que o VaR depende de p e de h , assim poderíamos escrever $VaR_h(p)$. Além disso, consideramos as seguintes observações:

- (i) Geralmente o VaR é dado em unidades monetárias, por exemplo em reais, e os retornos simples R_t são dados em porcentagem, lembramos que os log-retornos r_t são aproximadamente iguais a R_t , por isso podemos supor que r_t mede, aproximadamente, as variações percentuais. Neste trabalho consideraremos a variação dos ativos entre o instante t e $t+1$ dada por $\Delta p(1) = \log(P_{t+1}) - \log(P_t) = r_{t+1}$.
- (ii) A definição (2.1) mostra que o VaR é o p -quantil da distribuição $F_h(\cdot)$ que na prática deve ser estimado.
- (iii) O VaR para uma posição comprada tem valor negativo, pois acontece uma perda se $\Delta p(h) < 0$.

Acontece uma perda no caso de uma posição vendida se $\Delta p(h) \geq 0$, ou seja, o preço do ativo aumenta. Neste caso o VaR é definido por

$$p = Pr(\Delta p(h) \geq VaR) = 1 - F_h(VaR),$$

que comumente é positivo para p pequeno.

As definições anteriores implicam que o VaR para uma posição comprada é calculado utilizando a cauda esquerda da função de distribuição $F_h(\cdot)$ e a cauda direita para uma posição vendida. Também podemos aplicar a definição (2.1) para calcular o VaR de uma posição vendida se usarmos a distribuição de $-\Delta p(h)$. Portanto, basta considerar o cálculo do VaR para uma dessas posições.

Neste trabalho estaremos interessados na previsão do VaR feita no instante t com um horizonte $h = 1$ com probabilidade p , denotado por $VaR_{t+1|t}(p)$.

2.3 Teoria dos Valores Extremos

A teoria de valores extremos (TVE) clássica estuda o comportamento de máximos, mínimos e outras estatísticas de ordem de sequências de variáveis aleatórias independentes e identicamente

distribuídas. Também foram feitas extensões para o caso de séries estacionárias com dependência fraca e séries não estacionárias foram consideradas na literatura.

Dada uma série observada de retornos r_1, \dots, r_L , definimos $r_{(L)} = \max\{r_1, \dots, r_L\}$. Em resumo, a Teoria dos Valores Extremos (TVE) clássica centrada em $r_{(L)}$, estuda a distribuição aproximada para o máximo padronizado $r_{(L)}^*$ dado por:

$$r_{(L)}^* = \frac{r_{(L)} - \beta_L}{\alpha_L},$$

para as sequências de constantes $\{\alpha_L > 0\}$ e $\{\beta_L\}$, que são escolhidas de modo a estabilizar a posição e escala do máximo $r_{(L)}$, quando $L \rightarrow +\infty$. Supondo que os retornos são independentes com distribuição F , se existirem sequências de constantes $\{\alpha_L > 0\}$ e $\{\beta_L\}$, tal que a distribuição de $r_{(L)}^*$ converge para uma distribuição não degenerada $G(z)$, quando $L \rightarrow +\infty$, então G pertence a uma das três famílias da distribuição generalizada de valores extremos (GVE), com função de distribuição dada por:

$$G(z) = \exp \left[- \left(1 + \xi \left(\frac{z - \tau}{\sigma} \right) \right)^{-1/\xi} \right],$$

definida sobre $\{z : 1 + \xi \left(\frac{z - \tau}{\sigma} \right) > 0\}$, para $-\infty < \tau < +\infty$, $-\infty < \xi < +\infty$ e $\sigma > 0$, em que τ é o parâmetro de posição, σ o parâmetro de escala e ξ o parâmetro de forma. A família de distribuição é determinada pelo parâmetro ξ , se $\xi = 0$ é do tipo Gumbel, se $\xi > 0$ é do tipo Frechét e se $\xi < 0$ é do tipo Weibull.

2.4 Teste da Razão de Verossimilhanças Generalizada (TRVG)

Seja Θ o espaço paramétrico, $\mathbf{X} = (x_1, \dots, x_n)^\top$ os valores observados da variável aleatória X e $\mathcal{L}(\theta; \mathbf{X})$ a função de verossimilhança de θ dada \mathbf{X} . Consideremos uma situação geral em que as hipóteses de interesse são:

$$H_0 : \theta \in \Theta_0,$$

$$H_1 : \theta \in \Theta_1,$$

em que $\Theta = \Theta_0 \cup \Theta_1$, $\Theta_0 \neq \emptyset$ e $\Theta_1 \neq \emptyset$.

O TRVG pode ser definido como um teste com a região crítica dada por:

$$\mathbf{A}_1 = \{ \mathbf{X} : \lambda(\mathbf{X}) = \frac{\sup_{\theta \in \Theta_1} \mathcal{L}(\theta; \mathbf{X})}{\sup_{\theta \in \Theta_0} \mathcal{L}(\theta; \mathbf{X})} \geq c \}.$$

Por facilidades computacionais, a região crítica do TRVG também pode ser definida como:

$$\mathbf{A}_1^* = \{ \mathbf{X} : \lambda(\mathbf{X}) = \frac{\sup_{\theta \in \Theta_0} \mathcal{L}(\theta; \mathbf{X})}{\sup_{\theta \in \Theta} \mathcal{L}(\theta; \mathbf{X})} \leq c \}.$$

Em relação à interpretação do TRVG, observemos que $0 \leq \lambda(\mathbf{X}) \leq 1$, pois o numerador é o supremo com relação a θ que pertence a um subconjunto de Θ ($\Theta_0 \in \Theta$), enquanto que o denominador é o supremo sobre todo o conjunto Θ . Se H_0 for verdadeira, $\lambda(\mathbf{X})$ estaria próximo de 1, se for falsa, esperamos que o denominador seja grande em relação ao numerador, e, portanto $\lambda(\mathbf{X})$ deve ser próximo de 0.

O valor de c é obtido resolvendo a equação:

$$\alpha = \sup_{\theta \in \Theta_0} P(\lambda(\mathbf{X}) \leq c),$$

para isso, precisamos da distribuição de $\lambda(\mathbf{X})$ que, em geral, não é simples de ser obtida ou é desconhecida.

Seja $\boldsymbol{\theta} = (\theta_1, \dots, \theta_r)^\top$ o vetor de parâmetro r -dimensional e $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \dots, \hat{\theta}_r)^\top$ seu estimador de máxima verossimilhança (EMV). Bickel e Doksum [1977] apresentam o TRVG usando o EMV $\hat{\boldsymbol{\theta}}$ de $\boldsymbol{\theta} \in \Theta$ e $\hat{\boldsymbol{\theta}}^0$ de $\boldsymbol{\theta} \in \Theta_0$ para calcular

$$\lambda(\mathbf{X}) = \mathcal{L}(\hat{\boldsymbol{\theta}}^0; \mathbf{X}) / \mathcal{L}(\hat{\boldsymbol{\theta}}; \mathbf{X}).$$

Se podemos encontrar uma função h tal que $h(\lambda(\mathbf{X}))$ tenha uma forma simples e distribuição conhecida sob H_0 , é possível obter c , resolvendo a equação $\alpha = P_{H_0}(h(\lambda(\mathbf{X}))) \leq c$.

Se a distribuição de $\lambda(\mathbf{X})$ é desconhecida ou não é possível encontrar a função h , podemos usar a distribuição assintótica da estatística do TRVG, dada no seguinte teorema.

Teorema 2.1 (Sen e Singer [1994]) *Seja X_1, \dots, X_n uma amostra aleatória da variável X com função de densidade de probabilidade $f(x|\theta)$. Sob as condições de regularidade, se $\theta \in \Theta_0$, então a distribuição da estatística $-2\log(\lambda(\mathbf{X})) \xrightarrow{D} \chi_v^2$, quando $n \rightarrow +\infty$, em que v é a diferença entre o número de parâmetros não especificados em Θ e Θ_0 .*

Neste trabalho, consideramos a seguinte expressão para $-2\log(\lambda(\mathbf{X}))$:

$$-2 \left(\log \mathcal{L}(\hat{\boldsymbol{\theta}}^0; \mathbf{X}) - \log \mathcal{L}(\hat{\boldsymbol{\theta}}; \mathbf{X}) \right).$$

2.5 Sequência de Previsões

Neste trabalho consideramos a avaliação de sequência de previsões, utilizando valores pós-amostrais. A seguir apresentamos um procedimento que consideramos para obter tal sequência:

- (a) Dividimos uma amostra $\mathbf{Z}_t = \{z_1, z_2, \dots, z_{t_0}, z_{t_0+1}, \dots, z_{T-1}, z_T, z_{T+1}\}$ do processo Z_t , em duas sub-amostras dadas por $\mathbf{Z}_{1,t} = \{z_t\}_{t=1}^{t_0}$ e $\mathbf{Z}_{2,t} = \{z_{t+1}\}_{t=t_0}^T$.
- (b) Ajustamos um modelo adequado, utilizando inicialmente os dados da amostra $\mathbf{Z}_{1,t}$. Seja $\hat{z}_t(h) = \hat{z}_{t_0+h|t_0}$ a previsão de z_{t_0+h} de origem t_0 e horizonte h . Consideramos um horizonte $h = 1$, uma origem de previsão dinâmica t_{orig} e um tamanho amostral fixo ws . Para obter a primeira previsão, usamos $t_{orig} = t_0$ e $ws = t_0$, deste modo usamos $\{z_1, z_2, \dots, z_{t_0-1}, z_{t_0}\}$ para ajustar o modelo de previsão e obter $\hat{z}_{t_0+1|t_0}$.

Para a segunda previsão, consideramos $t_{orig} = t_0 + 1$ e para manter o mesmo tamanho amostral (ws) usamos os valores $\{z_2, z_3, \dots, z_{t_0}, z_{t_0+1}\}$ para obter $\hat{z}_{t_0+2|t_0+1}$, análogo ao caso anterior.

O processo segue até a última previsão, que considera $t_{orig} = T$ e uma amostra dada por: $\{z_{T-ws}, z_{T-ws-1}, \dots, z_{T-1}, z_T\}$. No final do processo é obtida a seguinte sequência de previsões:

$$\{\hat{z}_{t_0+1|t_0}, \hat{z}_{t_0+2|t_0+1}, \dots, \hat{z}_{T|T-1}, \hat{z}_{T+1|T}\}.$$

O número de previsões obtidas considerando um horizonte de $h = 1$ é $T + 1 - t_0$.

- (c) A sub-amostra $\mathbf{Z}_{2,t} = \{z_t\}_{t=t_0+1}^{T+1}$ será usada para avaliar a qualidade da previsão. Na literatura $\mathbf{Z}_{1,t}$ e $\mathbf{Z}_{2,t}$ são denominadas como valores amostrais e pós-amostrais, respectivamente. Para mais detalhes sobre estes experimentos, ver Tashman [2000].

Dada uma sequência de previsões do VaR obtidas pelo procedimento anterior, este trabalho considera a avaliação da seguinte sequência de intervalos de previsão:

$$\{(-\infty, VaR_{t+1|t}(p))\}_{t=t_0}^T. \quad (2.1)$$

Dada uma pós-amostra de retornos $r_{t_0+1}, \dots, r_{T+1}$, dizemos que ocorreu uma violação do VaR, para uma posição comprada, no tempo $t+1$ se $r_{t+1} < VaR_{t+1|t}(p)$, deste modo definimos a variável indicadora da violação do VaR, por:

$$I_{t+1} = \begin{cases} 1 & , \text{ se } r_{t+1} < VaR_{t+1|t}(p) \\ 0 & , \text{ caso contrario.} \end{cases}$$

Também será importante considerar a duração entre duas violações do VaR consecutivas, dada por:

$$D_i = t_i - t_{i-1}, \quad (2.2)$$

em que t_i denota o instante da i -ésima violação e $t_0 = 0$ implica que D_1 é o tempo até a primeira violação.

A sequência binária $\{I_{t+1}\}_{t=t_0}^T$ e as respectivas durações $\{D_i\}_{i=1}^{N(T)}$ serão parte da informação utilizada pelos testes de avaliação. Neste caso denotamos o tamanho da amostra de durações por $N(T)$, pois depende do tamanho das violações do VaR.

2.6 Análise de Tempo de Falha Discreto

A duração entre as violações no tempo t_i e t_{i-1} , dada por D_i , pode ser considerada como tempo de falha discreto e pode ser estudado através de sua função de risco. Consideramos a D_i como uma variável aleatória discreta, que assume valores em $\{1, 2, \dots\}$, com função de probabilidade dada por $f(d) = P(D_i = d)$ e função de sobrevivência dada por:

$$S(d) = P(D_i > d) = \sum_{j=d+1}^{+\infty} f(j), \quad (2.3)$$

para $i = 1, \dots, N(T)$.

A função de risco de D_i no tempo d é definida como a probabilidade condicional de falhar ao tempo d , dado que a i -ésima duração sobreviveu até o tempo $d-1$, dada por:

$$\lambda_d = P(D_i = d | D_i \geq d) = \frac{P(D_i = d)}{P(D_i \geq d)} = \frac{f(d)}{S(d-1)}, \quad (2.4)$$

para $d=1, 2, \dots$.

As seguintes propriedades mostram a relação de λ_d com a função de probabilidade e sobrevivência de D_i :

$$S(d) = \prod_{j|j \leq d} (1 - \lambda_j) \quad (2.5)$$

e

$$f(d) = \lambda_d \prod_{j=1}^{d-1} (1 - \lambda_j). \quad (2.6)$$

Como no caso contínuo, a função de risco discreta λ_d , determina uma única distribuição para o tempo de falha D_i . Para maiores detalhes ver [Kalbfleisch e Prentice \[2011\]](#).

Capítulo 3

Métodos de previsão do Valor em Risco.

3.1 Introdução

Na literatura há vários métodos de previsão do VaR . Neste capítulo apresentamos os seguintes: o método de Simulação Histórica (SH), a abordagem econométrica que utiliza os modelos da família $ARMA-GARCH$ e uma nova abordagem do método Peaks - over - threshold (POT), apresentada em Araújo Santos e Fraga Alves [2013], denominada método POT baseado em durações.

3.2 Estimação do VaR utilizando a Simulação Histórica (SH)

Seja $VaR_{t+1}(p)$ o p -quantil da distribuição condicional do retorno r_{t+1} , definido como $p = P(r_{t+1} < VaR_{t+1})$, o qual também pode ser expressado por:

$$VaR_{t+1}(p) = F_{t+1}^-(p),$$

em que $F_{t+1}(\cdot)$ denota a função de distribuição condicional de r_{t+1} . O método de Simulação Histórica utiliza os valores passados (um ou dois anos, por exemplo) dos retornos para construir um estimador empírico de $F_{t+1}^-(p)$ dado por:

$$\hat{VaR}_{t+1|t}^{sh}(p) = \text{percentil}(\{r_s\}_{s=t-ws+1}^t, p), \quad (3.1)$$

em que ws é o tamanho da amostra móvel (*rolling window*) que é usado para aproximar a distribuição condicional.

Na prática, usamos $ws = 250$ ou 500 que corresponde ao período de 1 ou 2 anos de negociações (ou dias trabalhados) aproximadamente.

O método SH é amplamente utilizado pelas instituições financeiras, uma vez que este método é não paramétrico e de fácil implementação. No entanto, o método tem algumas deficiências como:

- (i) Os retornos passados usados para a previsão são assumidos como independentes e identicamente distribuídos.
- (ii) O método tem uma baixa sensibilidade à dinâmica de uma série com uma volatilidade evoluindo no tempo.

3.3 Previsão do VaR utilizando os modelos ARMA e GARCH

Em geral, uma série de retornos é não correlacionada, mas dependente, por isso, a volatilidade é modelada por um dos modelos heterocedásticos. Além disso, algumas delas ainda exibem a presença de auto-correlação, havendo a necessidade de eliminá-la por meio do ajuste inicial de um modelo linear. A estratégia é, portanto, modelar a média da série dos retornos utilizando modelos lineares do

tipo *ARMA* e depois modelar os resíduos deste modelo por um membro da família *GARCH*. Nesta abordagem supomos que o retorno r_t segue um processo $ARMA(m, q)$ - $GARCH(u, s)$, de forma que:

$$\begin{aligned} r_t &= \phi_0 + \sum_{i=1}^m \phi_i r_{t-i} + a_t + \sum_{j=1}^q \theta_j a_{t-j}, \\ a_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^u \alpha_i a_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2. \end{aligned}$$

Seja $\Psi_t = \{r_t, r_{t-1}, \dots\}$ e assumindo que os parâmetros são conhecidos, as previsões, a um passo da média condicional e da variância condicional de r_{t+1} , $E(r_{t+1}|\Psi_t) = \hat{r}_t(1)$ e $Var(r_{t+1}|\Psi_t) = \hat{\sigma}_t^2(1)$, respectivamente, são :

$$\hat{r}_t(1) = \phi_0 + \sum_{i=1}^m \phi_i r_{t-i+1} + \sum_{j=1}^q \theta_j a_{t-j+1},$$

e

$$\hat{\sigma}_t^2(1) = \alpha_0 + \sum_{i=1}^u \alpha_i a_{t-i+1}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j+1}^2.$$

Se ϵ_t segue uma distribuição $\mathcal{N}(0, 1)$, então

$$r_{t+1}|\Psi_t \sim \mathcal{N}(\hat{r}_t(1), \hat{\sigma}_t^2(1)),$$

logo, de

$$P(r_{t+1} \leq VaR_{t+1|t}(p)) = P\left(\frac{r_{t+1} - \hat{r}_t(1)}{\hat{\sigma}_t(1)} \leq \frac{VaR_{t+1|t}(p) - \hat{r}_t(1)}{\hat{\sigma}_t(1)}\right) = p,$$

temos que

$$\frac{VaR_{t+1|t}(p) - \hat{r}_t(1)}{\hat{\sigma}_t(1)} = z_p.$$

Portanto, o VaR é dado por $\hat{r}_t(1) + z_p \hat{\sigma}_t(1)$, em que z_p corresponde ao p -ésimo quantil da distribuição normal padrão.

Se ϵ_t tem distribuição t-Student padronizada com v graus de liberdade, a previsão do VaR é dada por $\hat{r}_t(1) + t_v^*(p) \hat{\sigma}_t(1)$ (em que $t_v^*(p)$ é o p -quantil da distribuição t_v padronizada).

Seja $t_v(p)$ o p -quantil de uma t-Student(v), a relação de $t_v(p)$ com $t_v^*(p)$, é dada a seguir:

$$p = P(t_v \leq t_v(p)) = P\left(\frac{t_v}{\sqrt{v/(v-2)}} \leq \frac{t_v(p)}{\sqrt{v/(v-2)}}\right) = P\left(t_v^* \leq \frac{t_v(p)}{\sqrt{v/(v-2)}}\right),$$

se $v > 2$, logo $t_v^*(p) = t_v(p)/\sqrt{v/(v-2)}$. Por tanto, se ϵ_t segue uma distribuição t-Student com v graus de liberdade, a previsão do VaR no instante t com um horizonte $h = 1$ denotado por $\hat{VaR}_{t+1|t}^{A.G}(p)$ é dada por:

$$\hat{VaR}_{t+1|t}^{A.G}(p) = \hat{r}_t(1) + \frac{t_v(p) \hat{\sigma}_t(1)}{\sqrt{v/(v-2)}}. \quad (3.2)$$

3.4 Previsão do VaR utilizando a Teoria dos Valores Extremos

Dada uma sequência $\mathbf{r}_L = \{r_1, \dots, r_L\}$, definimos $r_{(1)} = \min\{r_1, \dots, r_L\}$ e $r_{(L)} = \max\{r_1, \dots, r_L\}$. O mínimo e o máximo da sequência são relevantes para o cálculo do VaR para as posições financeiras comprada e vendida, respectivamente.

Na prática, basta considerar um dos casos, devido ao fato de que

$$r_{(1)} = -\max\{r_1^c, \dots, r_L^c\},$$

em que r_t^c é o retorno simétrico de r_t dado por: $r_t^c = -\log(P_{t+1}/P_t) = -r_t$. Neste caso a previsão do VaR é positiva, pois feita para uma posição comprada e notamos que a violação do VaR dada por $r_{t+1}^c > VaR_{t+1|t}(p)$ é equivalente a $r_{t+1} < -VaR_{t+1|t}(p)$.

O procedimento para obter a previsão do VaR, aplicando a TVE e utilizando a série de retornos r_L , é dado a seguir (ver Morettin [2008]):

- (a) Dividimos r_L em m blocos de tamanho n , tal que $L = mn$.
- (b) Obtemos os máximos $r_{(n),i} = \max_{1 \leq j \leq n} \{r_{(i-1)n+j}\}$ respectivos de cada bloco, em que $r_{(i-1)n+j}$ denota o j -ésimo retorno do i -ésimo bloco para $i = \{1, \dots, m\}$. Logo, obtemos os EMV de τ , σ e ξ ajustando uma distribuição generalizada dos valores extremos para esses máximos e um n fixo.
- (c) É possível obter uma expressão para os quantis da distribuição, a partir da qual podemos obter a previsão do VaR, dada por:

$$VaR_{t+1|t}^{TVE}(p) = \begin{cases} \hat{\tau}_n - \frac{\hat{\sigma}_n}{\hat{\xi}_n} \{1 - (-n \log(1-p))^{-\hat{\xi}_n}\} & , \text{ se } \hat{\xi}_n \neq 0 \\ \hat{\tau}_n - \hat{\sigma}_n \log(-n \log(1-p)) & , \text{ se } \hat{\xi}_n = 0, \end{cases}$$

em que $\hat{\tau}_n$, $\hat{\sigma}_n$ e $\hat{\xi}_n$ são os EMV de τ , σ e ξ respectivamente, considerando um tamanho n dos blocos.

Duas críticas feitas a essa abordagem são: a escolha dos blocos (o tamanho n) não é claramente definida e dado que é um método não condicional os efeitos de variáveis explicativas não podem ser considerados.

3.5 Previsão do VaR utilizando o Método POT

O método POT baseia-se na medida de excesso dos retornos com relação a um limiar ótimo denotado por μ dado por $y_t = r_t - \mu$. Para uma posição vendida, por exemplo, se fixamos $\mu = 0.025$, o i -ésimo excesso y_i acontece no tempo k_i se $r_{k_i} \geq 0.025$. Esta abordagem utiliza a informação contida em $(k_i, r_{k_i} - \mu)$.

Para modelar o excesso utilizamos a distribuição Generalizada de Pareto (DGP), a qual tem a seguinte função de distribuição:

$$G_{\gamma,\sigma}(y) = \begin{cases} 1 - (1 + \gamma y/\sigma)^{-1/\gamma} & , \text{ se } \gamma \neq 0, \\ 1 - \exp(-y/\sigma) & , \text{ se } \gamma = 0, \end{cases}$$

em que $\sigma > 0$, o suporte da distribuição é $y \geq 0$ quando $\gamma \geq 0$ e $0 \leq y \leq -\sigma/\gamma$ quando $\gamma < 0$.

A função de distribuição do excesso $Y = r - \mu$ dado que $r > \mu$ é dada por:

$$F_\mu(y) = P(r - \mu \leq y | r > \mu) = \frac{P(r \leq y + \mu) - P(r \leq \mu)}{1 - P(r \leq \mu)} = \frac{F(y + \mu) - F(\mu)}{1 - F(\mu)}, \quad (3.3)$$

para $0 \leq y < r^F - \mu$, em que $r^F := \sup\{r : F(r) < 1\}$. A distribuição do excesso pode ser aproximada à DGP para um limiar μ suficientemente grande, utilizando o seguinte teorema:

Teorema 3.1 *É possível encontrar uma função $\sigma(\mu) > 0$ tal que:*

$$\lim_{\mu \rightarrow r^F} \sup_{0 \leq y < r^F - \mu} |F_\mu(y) - G_{\gamma,\sigma}(y)| = 0,$$

se e somente se F está no domínio máximo de atração de uma distribuição de valores extremos.

A definição de domínio máximo de atração é apresentada no Apêndice A.1 e a demonstração do teorema é dada em Embrechts *et al.* [1997]. Em geral o limiar μ é selecionado utilizando os retornos observados, antes de ajustar a DGP.

Seja $\bar{F}(x) = 1 - F(x)$ e $r^* = y + \mu$ um valor particular de r , de (3.3) temos que:

$$\bar{F}_\mu(y) = \frac{\bar{F}(r^*)}{\bar{F}(\mu)},$$

logo

$$\bar{F}(r^*) = \bar{F}(\mu)\bar{F}_\mu(y).$$

Consideremos a proposta de Smith [1987] de um estimador de cauda para $\bar{F}(r^*) = P(r > r^*)$ baseado em uma aproximação da distribuição do excesso pela DGP. Seja M o número de excessos acima de μ em uma amostra r_1, \dots, r_L , utilizamos M/L como um estimador para $\bar{F}(\mu)$ e $\bar{F}_\mu(y)$ é estimada usando a DGP aproximada. Para estimar os parâmetros γ e σ da $G_{\gamma, \sigma}(y)$ ajustamos uma DGP para o excesso sobre um limiar fixo μ . O estimador de cauda de $\bar{F}(r^*)$ é obtido por:

$$\begin{aligned} \hat{\bar{F}}(r^*) &= \hat{\bar{F}}(\mu)\hat{\bar{F}}_\mu(y) \\ &= \hat{\bar{F}}(\mu)\hat{\bar{F}}_\mu(r^* - \mu), \text{ fazendo } y = r^* - \mu \\ &= (M/L)(1 - G_{\hat{\gamma}, \hat{\sigma}}(r^* - \mu)) = \frac{M}{L} \left(1 + \hat{\gamma} \frac{r^* - \mu}{\hat{\sigma}} \right)^{-1/\hat{\gamma}}, r^* > \mu. \end{aligned}$$

Para $p < \bar{F}(\mu)$, invertendo $\hat{\bar{F}}(r^*)$ obtemos o estimador POT do VaR :

$$\hat{VaR}_{t+1|t}^{POT}(p) = \mu + \frac{\hat{\sigma}}{\hat{\gamma}} \left(\left(\frac{M}{Lp} \right)^{\hat{\gamma}} - 1 \right).$$

Para maiores detalhes ver Tsay [2005].

3.6 O método POT baseado em Durações entre Excessos (DPOT)

O seguinte método considera uma distribuição condicional do excesso que permite incluir covariáveis na estimação do VaR. Sejam y_1, \dots, y_M os excessos sobre o limiar μ obtidos até o instante t , d_1 a duração até o primeiro excesso (instante até acontecer pela primeira vez $r_t \geq \mu$) e d_2, \dots, d_M definidas por:

$$d_i = k_i - k_{i-1},$$

em que k_i denota o tempo do i -ésimo excesso. Considerando as durações $d_i, \dots, d_{i-\lambda+1}$, a duração ao tempo k_i dos λ excessos anteriores, é definida por:

$$d_{i,\lambda} = d_i + \dots + d_{i-\lambda+1} = k_i - k_{i-\lambda}.$$

Seja k_M o instante do último excesso y_M , ao instante t , d^t denota a duração desde k_M até t dada por $d^t = t - k_M + 1$ e $d_M = k_M - k_{M-1}$ a duração do último excesso. Finalmente, considerando $d_{t,1} = d^t$, $d_{t,2} = d^t + d_M$, para $\lambda = 3, 4, \dots$, a duração até o tempo t a partir dos λ excessos passados é dada por:

$$d_{t,\lambda} = d^t + d_{M,\lambda-1} = d^t + d_M + \dots + d_{M-\lambda+2}.$$

Considerando $\Psi_t^* = \{d^t, d_M, d_{M-1}, \dots, d_{M-\lambda+2}\}$, assumimos a seguinte distribuição condicional para o excesso Y_t sobre μ :

$$Y_t | \Psi_t^* \sim \mathcal{DGP}(\gamma, \sigma_t = g(\alpha_1, \dots, \alpha_k, \dots, d^t, d_M, d_{M-1}, \dots, d_{M-\lambda+2})),$$

em que $\gamma, \alpha_1, \dots, \alpha_k$ são parâmetros a serem estimados. É proposta a seguinte classe de estimadores para o Var :

$$\hat{Va}R_{t+1|t}^{DPOT}(p) = \mu + \frac{\hat{\sigma}_t}{\hat{\gamma}} \left(\left(\frac{M}{Lp} \right)^{\hat{\gamma}} - 1 \right),$$

com $\hat{\sigma}_t = g(\hat{\alpha}_1, \dots, \hat{\alpha}_k, \dots, d^k, d_M, d_{M-1}, \dots, d_{M-\lambda+2})$. O método $DPOT$ proposto implica um valor esperado e variância condicional para o excesso :

$$E(Y_t | \Psi_t^*) = \frac{\sigma_t}{1 - \gamma}, \text{ para } \gamma < 1,$$

e

$$Var(Y_t | \Psi_t^*) = \frac{\sigma_t^2}{(1 - 2\gamma)}, \text{ para } \gamma < \frac{1}{2}.$$

Resultados empíricos obtidos sugerem uma relação inversa entre o excesso y_i e $d_{i,\lambda}$, e uma correlação positiva com $1/(d_{i,\lambda})^c$, para $c > 0$. É proposta a especificação $\sigma_t = \alpha/(d_{t,\lambda})^c$ e o VaR estimado por:

$$\hat{Va}R_{t+1|t}^{DPOT}(\lambda, c)(p) = \mu + \frac{\hat{\alpha}}{\hat{\gamma}(d_{t,\lambda})^c} \left(\left(\frac{M}{Lp} \right)^{\hat{\gamma}} - 1 \right), \quad (3.4)$$

em que $\hat{\gamma}$ e $\hat{\alpha}$ são os EMV de γ e α . Considerando $\mathbf{Y} = (y_1, \dots, y_M)^\top$, a função de log-verossimilhança dos parâmetros γ e α dada \mathbf{Y} , é :

$$\begin{aligned} \log \mathcal{L}((\gamma, \alpha); \mathbf{Y}) &= \log \prod_{i=\lambda}^M f_{Y_i}(y_i) = \log \prod_{i=\lambda}^M \left(\frac{\alpha}{\gamma(d_{i,\lambda})^c} \right)^{-1} \left(1 + \frac{\gamma}{\alpha} y_i (d_{i,\lambda})^c \right)^{-(1/\gamma+1)} \\ &= - \sum_{i=\lambda}^M \log \left(\frac{\alpha}{(d_{i,\lambda})^c} \right) - \left(\frac{1}{\gamma} + 1 \right) \sum_{i=\lambda}^M \log \left(1 + \frac{\gamma}{\alpha} y_i (d_{i,\lambda})^c \right). \end{aligned}$$

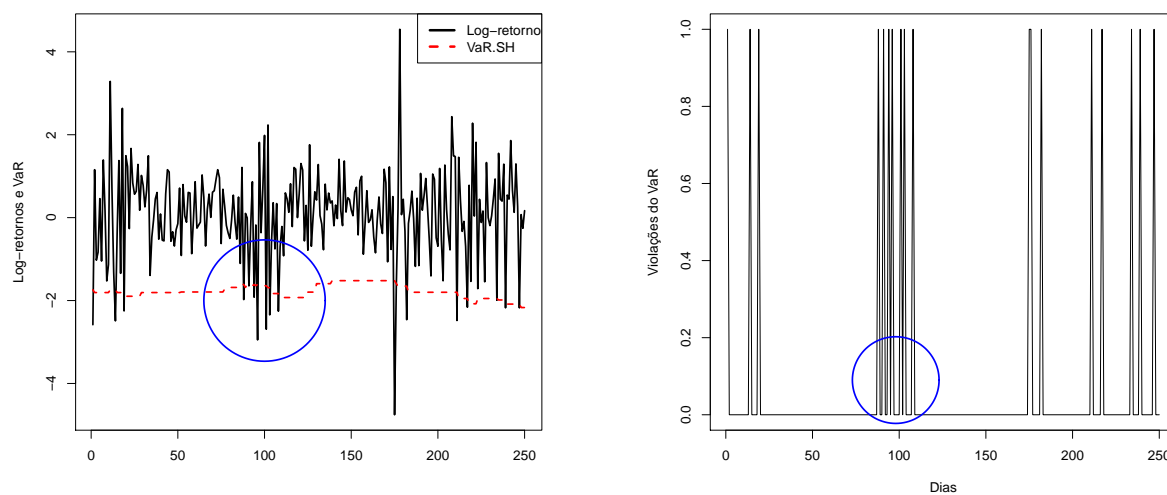
Os parâmetros podem ser obtidos usando otimização numérica por meio do algoritmo Newton-Rapson. O pacote `evt0` no R implementa uma função para obter $\hat{Va}R_{t+1|t}^{DPOT}(\lambda, c)(p)$.

Araújo Santos e Fraga Alves [2013] realizaram um estudo comparativo, no qual foram avaliados diferentes métodos de previsão do VaR. Foi observado que para $c = 0,75$ o método DPOT tem bons resultados, considerando $\lambda = 3$. O limiar μ é escolhido de tal modo que 10% da amostra utilizada para estimar o modelo de previsão seja maior que μ , McNeil e Frey [2000] apresenta um estudo de simulação que recomenda uma escolha similar.

3.7 Agrupamento das Violações do VaR

Ao realizar o processo de avaliação das previsões do VaR, um fenômeno de interesse é o agrupamento das violações do VaR. Em palavras simples, isto acontece quando os tempos em que ocorreram um conjunto de violações são muito próximos. Para ilustrar isto, simulamos 500 retornos diários, utilizando um modelo heterocedástico. Obtemos as previsões do VaR, utilizando o método Simulação Histórica, considerando $ws = 250$, $p = 0.05$, uma amostra inicial e pós-amostra, de tamanhos iguais a 250. Na Figura 3.1a observamos como as violações do VaR acontecem em períodos de alta volatilidade da série, enquanto em períodos calmos não acontecem.

Na Figura 3.1b apresentamos as violações do VaR ao longo do tempo, nesta figura destacamos em um círculo um agrupamento de 7 violações, que ocorreram em um período de 21 dias. As durações entre violações neste exemplo são: 13, 5, **69**, 3, 3, 2, 5, 2, 5, **67**, 1, 6, **29**, 6, 17, 5, 8, 3. Observamos longas durações seguidas de curtas durações associadas a uma sequência de violações que apresentam comportamento de agrupamentos.

(a) Retorno simulado e \widehat{VaR}^{sh} , $p=0,05$.(b) Sequência de indicadores das violações, $p=0,05$.**Figura 3.1:** *Agrupamento das violações do VaR*

Uma definição de eficiência para uma sequência de previsões do VaR, que é apresentada no próximo capítulo, aponta que o agrupamento é um sinal de ineficiência.

Capítulo 4

Testes para avaliação de previsão do VaR

4.1 Introdução

Um dos temas centrais da gestão quantitativa de risco financeiro é o cálculo preciso do VaR. O VaR é uma medida padrão de risco, amplamente utilizada pelas instituições financeiras e reguladoras.

No ponto de vista estatístico, o VaR é um quantil da distribuição condicional dos retornos de um ativo. A estimação do VaR não é trivial, uma vez que a distribuição condicional de retornos é, em geral, desconhecida. Conseqüentemente, alguns métodos populares de previsão do VaR são construídos baseados em suposições irreais, tais como, a suposição de distribuição i.i.d. para os retornos, desconsiderando os fatos estilizados de retornos (agrupamento de volatilidade e retornos variando no tempo). Assim, a avaliação do desempenho da previsão do VaR é de grande importância.

Neste Capítulo, apresentamos a definição de eficiência para intervalos de previsão e os testes de avaliação, os quais são divididos em dois grupos: testes baseados na sequência binária das violações do VaR e testes baseados nas durações entre violações.

4.2 Eficiência dos Intervalos de Previsão

Seja $\mathbf{R}_T = \{r_{t+1}\}_{t=t_0}^T$ uma série de retornos pós-amstral, a sequência de intervalos de previsão é dada por:

$$\mathbf{F}_T = \{(L_{t+1|t}(p), U_{t+1|t}(p))\}_{t=t_0}^T, \quad (4.1)$$

em que $L_{t+1|t}(p)$ e $U_{t+1|t}(p)$ são os limites inferior e superior da previsão para o tempo $t + 1$, feita na origem t , com probabilidade de cobertura igual a p .

Consideremos também a variável indicadora das coberturas, dada na seguinte definição:

Definição 4.1 A variável indicadora I_{t+1} , para um intervalo de previsão

$$(L_{t+1|t}(p), U_{t+1|t}(p)) ,$$

para o tempo $t + 1$ na origem t , é definida como:

$$I_{t+1} = \begin{cases} 1 & , \text{ se } r_{t+1} \in (L_{t+1|t}(p), U_{t+1|t}(p)) \\ 0 & , \text{ caso contrario} \end{cases} .$$

Denotamos por $\mathbf{I}_T = \{I_{t+1}\}_{t=t_0}^T$ a sequência de variáveis indicadoras das coberturas gerada por \mathbf{F}_T e pela pós-amostra \mathbf{R}_T .

A definição geral de eficiência da sequência de intervalos de previsão \mathbf{F}_T dada em (4.1), com respeito à informação passada Ω_t , apresentada em Christoffersen [1998] é dada a seguir:

Definição 4.2 Dizemos que a sequência de intervalos de previsão \mathbf{F}_T dada em (4.1) é eficiente com respeito ao conjunto de informação passada Ω_t , se $E(I_{t+1}|\Omega_t) = p, \forall t$.

Notamos que uma forma equivalente para expressar esta definição de eficiência é a seguinte: dizemos que \mathbf{F}_T é eficiente com respeito à informação passada Ω_t , se

$$P(I_{t+1} = 1|\Omega_t) = p, \forall t. \quad (4.2)$$

Esta definição permite construir testes de avaliação para intervalos de previsão, sem depender de qualquer suposição da distribuição dos retornos. Este fato é importante para a maioria das aplicações em economia, uma vez que, em geral, qualquer tipo de suposição sobre a distribuição dos retornos é altamente questionável e poderia levar a uma incorreta especificação do modelo de previsão do VaR. Além disso, não fazer suposições sobre a distribuição dos retornos também é crucial para a construção de métodos de avaliação.

Christoffersen [1998] propõe considerar a informação passada até o tempo t dada por:

$$\Omega_{1,t} = \{I_t, I_{t-1}, \dots, I_1\}, \quad (4.3)$$

e o seguinte procedimento para testar a eficiência de \mathbf{F}_T com respeito à informação $\Omega_{1,t}$.

Lema 4.1 Testar $E(I_{t+1}|\Omega_{1,t}) = E(I_{t+1}|I_t, I_{t-1}, \dots, I_1) = p, \forall t$, é equivalente a testar se as variáveis da sequência \mathbf{I}_T são independentes e identicamente distribuídas, segundo a distribuição Bernoulli com parâmetro p , ou seja, testar se $I_t \stackrel{i.i.d}{\sim} \text{Bernoulli}(p), \forall t$.

Destacamos que a definição de eficiência (4.2) é geral, no sentido que vale para qualquer tipo de informação passada Ω_t . Em particular, para afirmar que \mathbf{F}_T é eficiente com respeito a $\Omega_{1,t}$, utilizamos a seguinte definição:

Definição 4.3 Dizemos que a sequência de intervalos de previsão \mathbf{F}_T , dada em (4.1), tem uma cobertura condicional correta se $I_t \stackrel{i.i.d}{\sim} \text{Bernoulli}(p), \forall t$.

Uma observação simples, mas muito importante, que precisamos fazer ao lema (4.1) é que testar $E(I_{t+1}|\Omega_{1,t}) = p$, também é equivalente a testar (conjuntamente) se as variáveis da sequência \mathbf{I}_T são independentes e $E(I_{t+1}) = p, \forall t$. Assumindo esta independência, temos que $P(I_{t+1} = 1|\Omega_{1,t}) = P(I_{t+1} = 1) = \pi, \forall t$, para algum $\pi \in (0, 1)$, isto é, a probabilidade de ocorrer uma violação do VaR no instante $t + 1$ não depende das violações passadas. Se a condição $E(I_{t+1}) = p, \forall t$ for válida, dizemos que \mathbf{F}_T tem cobertura não condicional correta. Esta observação é importante, pois frequentemente a avaliação de \mathbf{F}_T é feita testando a independência e a cobertura não condicional correta separadamente, para depois verificar a cobertura condicional correta.

É possível realizar uma avaliação da cobertura condicional da sequência

$$\{(-\infty, VaR_{t+1|t}(p))\}_{t=t_0}^T,$$

considerando a definição (2.1), para o VaR de uma posição comprada, em que

$$p = P(r_{t+1} < VaR_{t+1|t}(p)) = P(r_{t+1} \in (-\infty, VaR_{t+1|t}(p))).$$

Assim, a previsão do VaR pode ser considerada como uma aplicação dos intervalos de previsão do tipo unilateral. Também, considerando que $(L_{t+1|t}(p), U_{t+1|t}(p))$ é equivalente a $(-\infty, U_{t+1|t}(p))$, para $L_{t+1|t}(p) = -\infty$, podemos testar a eficiência dos intervalos de previsão $(-\infty, VaR_{t+1|t}(p))$, utilizando o Lema 4.1, como foi feito para os intervalos bilaterais.

Assim, no decorrer do trabalho, consideramos $\mathbf{F}_T = \{(-\infty, VaR_{t+1|t}(p))\}_{t=t_0}^T$ e I_{t+1} dada em (4.1) como equivalente à indicadora da violação do VaR, apresentada na seção (2.5). As coberturas não condicional e condicional são denotadas por UC e CC, respectivamente (*unconditional* e *conditional coverage*).

Eficiência das previsões do VaR

Até aqui, temos considerado a definição de eficiência de intervalos de previsão com respeito à informação passada $\Omega_{1,t}$. Esta definição é utilizada na avaliação das previsões do VaR, verificando por meio de testes para eficiência da sequência $\mathbf{F}_T = \{(-\infty, VaR_{t+1|t}(p))\}_{t=t_0}^T$. Deste modo, consideraremos que a sequência $\{VaR_{t+1|t}(p)\}_{t=t_0}^T$ é eficiente, se \mathbf{F}_T for eficiente com respeito às informações passadas.

No seguir do trabalho, consideramos que \mathbf{F}_T é eficiente com respeito à Ω_t , se tem uma CC correta com respeito a Ω_t . Em particular, \mathbf{F}_T é eficiente com respeito a $\Omega_{1,t}$, se \mathbf{F}_T tem uma CC correta com respeito a $\Omega_{1,t}$, ou seja $I_t \stackrel{i.i.d}{\sim} \text{Bernoulli}(p), \forall t$. Na literatura, também é de interesse avaliar a UC correta e a independência das variáveis da sequência \mathbf{I}_T . O teste de independência verifica se:

$$E(I_{t+1}|\Omega_{1,t}) = \pi, \forall t, \quad (4.4)$$

ou seja, a sequência \mathbf{F}_T tem um cobertura condicional, com probabilidade de cobertura $\pi \in (0, 1)$.

Em resumo, o Lema 4.1 fornece um procedimento de avaliação da sequência \mathbf{F}_T , o qual pode ser construído a partir de um modelo de séries temporais paramétrico ou não paramétrico, ou a partir de julgamentos. A ideia de testar se $I_t \stackrel{i.i.d}{\sim} \text{Bernoulli}(p), \forall t$, é obter uma indicação de quão perto está a cobertura atual da CC correta .

4.3 Testes Baseados na Sequência \mathbf{I}_T

Dada a sequência de indicadores das violações do VaR, para avaliar a sequência \mathbf{F}_T com respeito à $\Omega_{1,t}$, consideramos as seguintes hipóteses nulas e seus correspondentes testes:

- Para avaliar a hipótese de CC correta com respeito $\Omega_{1,t}$, testamos se $I_t \stackrel{i.i.d}{\sim} \text{Bernoulli}(p), \forall t$.
- Para avaliar a hipótese de independência das variáveis da sequência \mathbf{I}_T , testamos se $I_t \stackrel{i.i.d}{\sim} \text{Bernoulli}(\pi), \forall t$, para $\pi \in (0, 1)$.
- Para avaliar a hipótese de UC correta, assumimos que $I_t \stackrel{i.i.d}{\sim} \text{Bernoulli}(\pi), \forall t$ e testamos $\pi = p$.

4.3.1 Teste do Tipo Bernoulli

Além da definição geral de eficiência para intervalos de previsão, Christoffersen [1998] propõe um procedimento para testar a hipótese de CC correta com respeito a $\Omega_{1,t}$. Isto pode ser feito convenientemente utilizando o TRGV. A seguir apresentamos testes do tipo TRVG para UC correta, independência e um procedimento que combina os dois testes, para formar um teste completo para CC.

a) Teste Kupiec de Cobertura não Condicional

Assumindo que a sequência de indicadores $\mathbf{I}_T = \{I_{t+1}\}_{t=t_0}^T$ são independentes com distribuição Bernoulli(π). O teste da UC correta procura verificar se $E(I_t) = p, \forall t$. As hipóteses propostas são:

$$\begin{cases} H_0 : \pi = p \\ H_1 : \pi \neq p \end{cases} .$$

Considerando $\Theta_0 = \{\pi = p\}$ e $\Theta = \{\pi : 0 < \pi < 1\}$. A função de verossimilhança sob H_0 é dada por:

$$\mathcal{L}(p; \mathbf{I}_T) = \prod_{t=t_0}^T (1-p)^{1-i_{t+1}} \times p^{i_{t+1}} = (1-p)^{n_0} \times p^{n_1},$$

em que i_{t+1} é o valor observado de I_{t+1} e n_0 e n_1 são as quantidades de zeros e uns na sequência \mathbf{I}_T , respectivamente.

A função de verossimilhança sob H_1 é dada por:

$$\mathcal{L}(\pi; \mathbf{I}_T) = \prod_{t=t_0}^T (1-\pi)^{1-i_{t+1}} \times \pi^{i_{t+1}} = (1-\pi)^{n_0} \times \pi^{n_1}.$$

Logo, a estatística do *TRVG* para a *UC* correta, denotada por TB_{uc} , é dada por:

$$TB_{uc} = -2(\log \mathcal{L}(p; \mathbf{I}_T) - \log \mathcal{L}(\hat{\pi}; \mathbf{I}_T)) \xrightarrow{D} \chi_1^2,$$

em que $\hat{\pi} = \frac{n_1}{n_0+n_1} = \frac{\sum_{t=t_0}^T I_{t+1}}{n_0+n_1}$ é o EMV de π .

Este procedimento, também conhecido como o teste de Kupiec, somente avalia a correta proporção de violações, sem considerar a possível dependência entre elas. O seguinte procedimento propõe testar um tipo particular de dependência.

b) Teste Markov de Independência

Consideremos que $\{I_{t+1}|I_t\}_{t=t_0}^T$ é uma cadeia de Markov de primeira ordem, com matriz de probabilidade de transição dada por:

$$\Pi_1 = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix},$$

em que $\pi_{ij} = P(I_{t+1} = j | I_t = i)$, para $i, j \in \{0, 1\}$.

A função de verossimilhança aproximada deste processo é dada por:

$$\begin{aligned} \mathcal{L}(\Pi_1; \mathbf{I}_T) &= \prod_{t=t_0}^T P(I_{t+1} = j | I_t = i) \\ &= \prod_{\{t: I_t=0\}} P(I_{t+1} = j | I_t = 0) \times \prod_{\{t: I_t=1\}} P(I_{t+1} = j | I_t = 1) \\ &= \pi_{01}^{n_{01}} (1 - \pi_{01})^{n_{00}} \pi_{11}^{n_{11}} (1 - \pi_{11})^{n_{10}}, \end{aligned}$$

em que n_{ij} são as quantidades de observações com valor i seguido por j na sequência \mathbf{I}_T . O EMV para Π_1 é dado por:

$$\hat{\Pi}_1 = \begin{bmatrix} \frac{n_{00}}{n_{00}+n_{01}} & \frac{n_{01}}{n_{00}+n_{01}} \\ \frac{n_{10}}{n_{10}+n_{11}} & \frac{n_{11}}{n_{10}+n_{11}} \end{bmatrix}.$$

Se as variáveis da sequência \mathbf{I}_T forem independentes, então temos que

$$P(I_{t+1} = j | I_t = i) = P(I_{t+1} = j) = \pi_j.$$

Assim, testar as hipóteses

$$\begin{cases} H_0 : I_t \stackrel{i.i.d}{\sim} \text{Bernoulli}(\pi), \forall t \\ H_1 : I_{t+1}|I_t \text{ é uma cadeia de Markov de primeira ordem} \end{cases},$$

é equivalente a testar:

$$\left\{ \begin{array}{l} H_0 : \Pi_1^* = \begin{bmatrix} 1 - \pi_1 & \pi_1 \\ 1 - \pi_1 & \pi_1 \end{bmatrix} \\ H_1 : \Pi_1 = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix} \end{array} \right. .$$

Considerando $\Theta_0 = \{\pi_1 : 0 < \pi_1 < 1\}$ e $\Theta = \{(\pi_{10}, \pi_{01}, \pi_{11}, \pi_{00}) : 0 < \pi_{ij} < 1\}$ para $i, j \in \{0, 1\}$. A função de verossimilhança sob H_0 é dada por:

$$\mathcal{L}(\Pi_1^*; \mathbf{I}_T) = (1 - \pi_1)^{n_{00} + n_{10}} \pi_1^{n_{01} + n_{11}}.$$

Sob a hipótese de independência, os EMV de π_1 e Π_1^* são

$$\hat{\pi}_1 = \frac{n_{01} + n_{11}}{n_{00} + n_{10} + n_{01} + n_{11}}$$

e

$$\hat{\Pi}_1^* = \begin{bmatrix} 1 - \hat{\pi}_1 & \hat{\pi}_1 \\ 1 - \hat{\pi}_1 & \hat{\pi}_1 \end{bmatrix} ,$$

respectivamente.

A estatística do TRVG para a Independência, denotada por TB_{ind} , é dada por:

$$TB_{ind} = -2 \left(\log \mathcal{L}(\hat{\Pi}_1^*; \mathbf{I}_T) - \log \mathcal{L}(\hat{\Pi}_1; \mathbf{I}_T) \right) \xrightarrow{D} \chi_1^2.$$

c) Teste Markov de Cobertura Condicional

Os testes para a UC correta e independência das variáveis da sequência \mathbf{I}_T são combinados em um único procedimento para formar um teste completo CC correta com respeito a $\Omega_{1,t}$, que considera as seguintes hipóteses.

$$\left\{ \begin{array}{l} H_0 : \Pi_1^* = \begin{bmatrix} 1 - p & p \\ 1 - p & p \end{bmatrix} \\ H_1 : \Pi_1 = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix} \end{array} \right. .$$

Considerando $\Theta_0 = \{\pi_1 = p\}$ e $\Theta = \{(\pi_{10}, \pi_{01}, \pi_{11}, \pi_{00}) : 0 < \pi_{ij} < 1\}$ para $i, j \in \{0, 1\}$. A estatística do TRVG para a correta CC com respeito a $\Omega_{1,t}$, denotada por TB_{cc} , é dada por:

$$TB_{cc} = -2 \left(\log \mathcal{L}(\Pi_1^*; \mathbf{I}_T) - \log \mathcal{L}(\hat{\Pi}_1; \mathbf{I}_T) \right) \xrightarrow{D} \chi_2^2.$$

Notemos que esta estatística também pode ser calculada por $TB_{cc} = TB_{uc} + TB_{ind}$.

4.3.2 Ampliação da Informação Passada

Christoffersen [1998] também propõe ampliar o conjunto de informação passada $\Omega_{1,t}$, para avaliar o efeito de outras covariáveis (além das indicadoras da cobertura) na probabilidade de cobertura dos intervalos de previsão. Consideremos a informação

$$\Omega_t^* = \{M_t, \dots, M_1\},$$

em que M_t é um vetor de q covariáveis, que contem o retorno r_t e I_t . Consideremos o seguinte modelo para $E(I_{t+1} | \Omega_t^*)$:

$$E(I_{t+1} | \Omega_t^*) = \beta_0 + \beta^\top \rho(M_t), \quad (4.5)$$

em que $\rho : \mathcal{R}^q \rightarrow \mathcal{R}^f$, β^\top um vetor f -dimensional.

O seguinte lema estabelece o procedimento para testar a eficiência de F_T com respeito a Ω_t^* , dado o modelo em (4.5).

Lema 4.2 *Testar a hipótese nula de eficiência de F_T com respeito à informação Ω_t^* , dada por $E(I_{t+1}|\Omega_t^*) = p, \forall t$, contra a alternativa $E(I_{t+1}|\Omega_t^*) = \beta_0 + \beta^\top \rho(M_t)$, é equivalente a testar $\beta_0 = p$ e $\beta = \mathbf{0}_f$ versus $\beta_0 \neq p$ ou $\beta \neq \mathbf{0}_f$, em que $\mathbf{0}_f$ é um vetor de zeros f -dimensional.*

Este procedimento, pode ser considerado como um teste conjunto de independência da informação passada ($\beta = \mathbf{0}_f$) e de UC correta ($\beta_0 = p$).

Para a construção de testes que avaliem o efeito de covariáveis na probabilidade de $I_{t+1} = 1$, para nossos propósitos é suficiente considerar a expansão da informação passada $\Omega_{1,t}$, dada por:

$$\Omega_{1,t}^* = \{I_t, I_{t-1}, \dots, r_t, r_{t-1}, \dots\} \quad (4.6)$$

e também a seguinte definição:

Definição 4.4 *Dizemos que uma sequência de intervalos de previsão F_T , gerada pelas previsões do VaR, tem cobertura condicional correta com respeito a $\Omega_{1,t}^*$, dada em (4.6), se:*

$$E(I_{t+1}|\Omega_{1,t}^*) = p, \forall t,$$

que é equivalente a dizer

$$P(I_{t+1} = 1|\Omega_{1,t}^*) = p, \forall t.$$

4.3.3 Teste de Quantis Dinâmicos

Para testar a CC correta com respeito a $\Omega_{1,t}^*$, dada em (4.6), Engle e Manganelli [1999] propõe trabalhar com a seguinte variável aleatória:

$$Hit_{t+1} = I_{t+1} - p.$$

O procedimento procura verificar a dependência de Hit_{t+1} com seus valores passados e a previsão do $VaR_{t+1|t}(p)$, considerando a seguinte regressão artificial:

$$Hit_{t+1} = \delta_0 + \sum_{i=0}^q \delta_{i+1} Hit_{t-i} + \delta_{q+2} VaR_{t+1|t}(p) + u_{t+1},$$

em que:

$$u_{t+1} = \begin{cases} -p & , \text{ com probabilidade } 1 - p \\ 1 - p & , \text{ com probabilidade } p \end{cases}.$$

No modelo é considerada a dependência linear de Hit_{t+1} com as Hit_{t-i} e o $VaR_{t+1|t}(p)$ para $q = 0, \dots, t - 1$. O modelo também permite incluir uma variável *dummy* anual para avaliar sua possível correlação com Hit_{t+1} .

Consideremos a seguinte forma matricial do modelo:

$$Hit_{t+1} = \mathbf{X}^\top \boldsymbol{\delta} + \mathbf{u}_{t+1},$$

em que $\mathbf{X} = (1, Hit_t, \dots, Hit_{t-q}, VaR_{t+1|t}(p))^\top$ e $\boldsymbol{\delta} = (\delta_0, \dots, \delta_{q+2})^\top$.

Notemos que:

$$E(I_{t+1}|\Omega_{1,t}^*) = E(Hit_{t+1} + p|\Omega_{1,t}^*) = E(Hit_{t+1}|\Omega_{1,t}^*) + p, \forall t,$$

Deste modo, F_T tem uma CC correta com respeito a $\Omega_{1,t}^*$, se e somente se

$$E(Hit_{t+1}|\Omega_{1,t}^*) = 0, \forall t.$$

Analogamente ao lema (4.2), testar a hipótese de CC correta, dada por $E(Hit_{t+1}|\Omega_{1,t}^*) = 0, \forall t$ contra a alternativa $E(Hit_{t+1}|\Omega_{1,t}^*) = \delta_0 + \sum_{i=1}^q \delta_{i+1}Hit_{t-i} + \delta_{q+2}VaR_{t+1|t}(p)$, é equivalente a testar as seguintes hipóteses:

$$\begin{cases} H_0 : \boldsymbol{\delta} = 0 \\ H_1 : \boldsymbol{\delta} \neq 0 \end{cases}.$$

Estimamos $\boldsymbol{\delta}$ via mínimos quadrados e sob as condições de regularidades apresentadas por Engle e Manganelli [1999], a distribuição assintótica de $\hat{\boldsymbol{\delta}}$, sob H_0 , é dada por:

$$\hat{\boldsymbol{\delta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{H}it_{t+1} \xrightarrow{D} N(\mathbf{0}, p(1-p)(\mathbf{X}^\top \mathbf{X})^{-1}).$$

A estatística para o teste do tipo Wald é:

$$[\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}^0]^\top [Var(\hat{\boldsymbol{\delta}})]^{-1} [\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}^0],$$

em que $\boldsymbol{\delta}^0$ é dado por H_0 e $Var(\hat{\boldsymbol{\delta}})$ é a variância assintótica do estimador de $\boldsymbol{\delta}$. Finalmente a estatística para testar a CC correta, denominada por estatística de Quantis Dinâmicos (DQ), é dada por:

$$\begin{aligned} DQ &= \hat{\boldsymbol{\delta}}^\top [p(1-p)(\mathbf{X}^\top \mathbf{X})^{-1}]^{-1} \hat{\boldsymbol{\delta}} \\ &= \frac{\hat{\boldsymbol{\delta}}^\top [\mathbf{X}^\top \mathbf{X}] \hat{\boldsymbol{\delta}}}{p(1-p)} \xrightarrow{D} \chi_{q+2}^2. \end{aligned}$$

4.3.4 Testes do Tipo CAViaR

Berkowitz *et al.* [2011] continua a abordagem do teste de quantis dinâmicos, considerando para u_t uma distribuição logística e propõe ajustar um modelo logístico para modelar $E(I_{t+1}|\Omega_{1,t}^*)$. Os procedimentos são do tipo TRVG e denominados como testes do tipo CAViaR.

Para a implementação dos testes, consideramos um modelo linear generalizado que considera a sequência $\{I_{t+1}\}_{t=t_0}^T$, tais que:

$$I_{t+1}|\Omega_{1,t}^* \sim \text{Bernoulli}(\pi_{t+1}) \quad (4.7)$$

e a parte sistemática do modelo dada por:

$$\log \left(\frac{P(I_{t+1} = 1|\Omega_{1,t}^*)}{1 - P(I_{t+1} = 1|\Omega_{1,t}^*)} \right) = \log \left(\frac{\pi_{t+1}}{1 - \pi_{t+1}} \right) = \eta_{t+1}, \quad (4.8)$$

em que η_{t+1} é o preditor linear dado por $\eta_{t+1} = \delta_0 + \delta_1 I_t + \delta_2 VaR_{t+1|t}(p)$.

De (4.8), consideramos o seguinte modelo para $E(I_{t+1}|\Omega_{1,t}^*)$:

$$\begin{aligned} E(I_{t+1}|\Omega_{1,t}^*) &= P(I_{t+1} = 1|\Omega_{1,t}^*) \\ &= \frac{e^{\eta_{t+1}}}{1 + e^{\eta_{t+1}}} = \frac{e^{\delta_0 + \delta_1 I_t + \delta_2 VaR_{t+1|t}(p)}}{1 + e^{\delta_0 + \delta_1 I_t + \delta_2 VaR_{t+1|t}(p)}}, \end{aligned} \quad (4.9)$$

A função de probabilidade de $I_{t+1}|\Omega_{1,t}^*$ em função de η_{t+1} é dada por:

$$\begin{aligned} P(I_{t+1} = i_{t+1}|\Omega_{1,t}^*) &= \pi_{t+1}^{i_{t+1}} (1 - \pi_{t+1})^{1-i_{t+1}} \\ &= \left(\frac{e^{\eta_{t+1}}}{1 + e^{\eta_{t+1}}} \right)^{i_{t+1}} \left(\frac{1}{1 + e^{\eta_{t+1}}} \right)^{1-i_{t+1}}. \end{aligned}$$

Aplicando logaritmo obtemos:

$$\begin{aligned} \log P(I_{t+1} = i_{t+1} | \Omega_{1,t}^*) &= i_{t+1} \log \left(\frac{e^{\eta_{t+1}}}{1 + e^{\eta_{t+1}}} \right) - (1 - i_{t+1}) \log(1 + e^{\eta_{t+1}}) \\ &= i_{t+1} \log(e^{\eta_{t+1}}) - i_{t+1} \log(1 + e^{\eta_{t+1}}) - \log(1 + e^{\eta_{t+1}}) + i_{t+1} \log(1 + e^{\eta_{t+1}}) \\ &= i_{t+1} \eta_{t+1} - \log(1 + e^{\eta_{t+1}}). \end{aligned}$$

Dado que

$$P(I_{t_0+1}, I_{t_0+1}, \dots, I_T, I_{T+1}) = \prod_{t=t_0}^T P(I_{t+1} | I_t, I_{t-1}, \dots) = \prod_{t=t_0}^T P(I_{t+1} | \Omega_{1,t}^*),$$

pela especificação de $P(I_{t+1} | \Omega_{1,t}^*)$ dada em (4.9), a função de log-verossimilhança de $\boldsymbol{\delta}^* = (\delta_0, \delta_1, \delta_2)$ correspondente à amostra observada \mathbf{I}_T , é dada por:

$$\log \mathcal{L}_{\logit}(\boldsymbol{\delta}^*; \mathbf{I}_T) = \sum_{t=t_0}^T \log P(I_{t+1} = i_{t+1} | \Omega_{1,t}^*) = \sum_{t=t_0}^T i_{t+1} \eta_{t+1} - \sum_{t=t_0}^T \log(1 + e^{\eta_{t+1}}).$$

Podemos testar a UC correta, independência e a CC correta, formulando apropriadas hipóteses sobre $E(I_{t+1} | \Omega_{1,t}^*)$ dado em (4.9).

a) Teste CAViaR de Cobertura não Condicional

Assumindo que $\delta_1 = \delta_2 = 0$, ou seja $E(I_{t+1} | \Omega_{1,t}^*)$ não depende de I_t e $\text{VaR}_{t+1|t}(p)$, temos que $E(I_{t+1}) = e^{\delta_0} / (1 + e^{\delta_0})$. O procedimento procura testar as hipóteses

$$\begin{cases} H_0 : E(I_{t+1}) = p \\ H_1 : E(I_{t+1}) \neq p \end{cases},$$

que são equivalentes às hipóteses:

$$\begin{cases} H_0 : \delta_0 = \log \left(\frac{p}{1-p} \right) \\ H_1 : \delta_0 \neq \log \left(\frac{p}{1-p} \right) \end{cases}.$$

Considerando $\Theta_0 = \{\delta_0 : \delta_0 = \log(p/(1-p))\}$ e $\Theta = \{\delta_0 : \delta_0 \in \mathcal{R}\}$. Se $\delta_1 = \delta_2 = 0$, a função de log-verossimilhança de δ_0 é definida como:

$$\log \mathcal{L}_{\logit}(\delta_0; \mathbf{I}_T) = \delta_0 \sum_{t=t_0}^T i_{t+1} - \sum_{t=t_0}^T \log(1 + e^{\eta_{t+1}}) = \delta_0 \sum_{t=t_0}^T i_{t+1} - T_A \log(1 + e^{\delta_0}),$$

em que $T_A = T - t_0 + 1$ é o número de indicadoras das violações em \mathbf{I}_T .

O EMV de δ_0 sob H_1 , denotado por $\hat{\delta}_0$ é

$$\hat{\delta}_0 = -\log \left(\frac{T_A}{\sum_{t=t_0}^T I_{t+1}} - 1 \right). \quad (4.10)$$

Como

$$\frac{d^2}{d\delta_0^2} \log \mathcal{L}_{\logit}(\delta_0; \mathbf{I}_T) = -\frac{T_A e^{\delta_0}}{(1 + e^{\delta_0})^2},$$

temos que

$$\frac{d^2}{d\delta_0^2} \log \mathcal{L}_{\logit_0}(\hat{\delta}_0; \mathbf{I}_T) < 0.$$

Logo, podemos afirmar que $\log \mathcal{L}_{\logit_0}(\hat{\delta}_0; \mathbf{I}_T)$ atinge o máximo em $\hat{\delta}_0$.

A estatística do TRVG para a UC correta, denotada por $CAViaR_{uc}$, é dada por:

$$CAViaR_{uc} = -2 \left(\log \mathcal{L}_{\logit_0}(\delta_0; \mathbf{I}_T) - \log \mathcal{L}_{\logit_0}(\hat{\delta}_0; \mathbf{I}_T) \right) \xrightarrow{D} \chi_1^2,$$

em que $\delta_0 = \log \left(\frac{p}{1-p} \right)$.

b) Teste CAViaR de Independência

Sob esta hipótese de independência, a probabilidade de ocorrer uma violação no tempo $t+1$ não depende da violação ao instante t e $VaR_{t+1|t}(p) \forall t$. O procedimento procura testar as hipóteses:

$$\begin{cases} H_0 : E(I_{t+1} | \Omega_{1,t}^*) = \pi, \forall t \\ H_1 : E(I_{t+1} | \Omega_{1,t}^*) = e^{\eta_{t+1}} / (1 + e^{\eta_{t+1}}) \end{cases},$$

para $\eta_{t+1} = \delta_0 + \delta_1 I_t + \delta_2 VaR_{t+1|t}(p)$, que são equivalente às hipóteses propostas por [Araújo Santos e Fraga Alves \[2012\]](#):

$$\begin{cases} H_0 : \delta_1 = 0 \text{ e } \delta_2 = 0 \\ H_1 : \delta_1 \neq 0 \text{ ou } \delta_2 \neq 0. \end{cases}$$

Considerando $\Theta_0 = \{\delta_0 : \delta_0 \in \mathcal{R}\}$ e $\Theta = \{\delta^* : \delta^* \in \mathcal{R}^3\}$. Neste caso, sob H_0 a função de log-verossimilhança é maximizada por $\hat{\delta}_0$, dado em (4.10).

Sob a H_1 , o EMV de δ^* , denotado por $\hat{\delta}^* = (\hat{\delta}_0, \hat{\delta}_1, \hat{\delta}_2)$, não tem solução analítica explícita, mas pode ser calculado mediante otimização numérica. A estatística do TRVG para a independência, denotada por $CAViaR_{ind}$, é dada por:

$$CAViaR_{ind} = -2 \left(\log \mathcal{L}_{\logit_0}(\hat{\delta}_0; \mathbf{I}_T) - \log \mathcal{L}_{\logit_0}(\hat{\delta}^*; \mathbf{I}_T) \right) \xrightarrow{D} \chi_2^2.$$

c) Teste CAViaR de Cobertura Condicional

Para testar a hipótese de CC correta, o seguinte procedimento avalia conjuntamente a UC correta e independência, considerando as seguintes hipóteses:

$$\begin{cases} H_0 : E(I_{t+1} | \Omega_{1,t}^*) = p, \forall t \\ H_1 : (I_{t+1} | \Omega_{1,t}^*) = e^{\eta_{t+1}} / (1 + e^{\eta_{t+1}}) \end{cases},$$

para $\eta_{t+1} = \delta_0 + \delta_1 I_t + \delta_2 VaR_{t+1|t}(p)$, que são equivalentes às seguintes hipóteses, propostas por [Berkowitz et al. \[2011\]](#).

$$\begin{cases} H_0 : \delta_0 = \log \left(\frac{p}{1-p} \right) \text{ e } \delta_1 = 0 \text{ e } \delta_2 = 0 \\ H_1 : \delta_0 \neq \log \left(\frac{p}{1-p} \right) \text{ ou } \delta_1 \neq 0 \text{ ou } \delta_2 \neq 0. \end{cases}$$

Considerando $\Theta_0 = \{\delta_0 : \delta_0 = \log(p/(1-p))\}$ e $\Theta = \{\delta^* : \delta^* \in \mathcal{R}^3\}$. A estatística do TRVG para a CC correta, denotada por $CAViaR_{cc}$, é dada por:

$$CAViaR_{cc} = -2 \left(\log \mathcal{L}_{\logit_0}(\delta_0; \mathbf{I}_T) - \log \mathcal{L}_{\logit_0}(\hat{\delta}^{1*}; \mathbf{I}_T) \right) \xrightarrow{D} \chi_3^2,$$

em que $\delta_0 = \log \left(\frac{p}{1-p} \right)$. Notemos que esta estatística pode ser calculada por $CAViaR_{cc} = CAViaR_{uc} + CAViaR_{ind}$.

4.4 Testes Baseados em Durações

Os seguintes procedimentos avaliam a eficiência de F_T através da sequência de durações entre violações do VaR, dada por $\{D_i\}_{i=1}^{N(T)}$. No decorrer do trabalho consideramos a seguinte definição para a i -ésima duração, denotada por D_i .

Definição 4.5 *Depois da $(i-1)$ -ésima violação do VaR, definimos a duração D_i como o tempo até ocorrer a i -ésima violação, isto é*

$$D_i = t_i - t_{i-1},$$

em que t_i é o instante em que ocorre a i -ésima violação.

A função de probabilidade de D_i apresentada em Berkowitz *et al.* [2011], assumindo que $I_t \stackrel{i.i.d}{\sim} \text{Bernoulli}(\pi), \forall t$, é dada por:

$$\begin{aligned} P(D_i = d) &= P(I_{t_{i-1}+d} = 1, I_{t_{i-1}+d-1} = 0, \dots, I_{t_{i-1}+1} = 0) \\ &= \pi(1 - \pi)^{d-1}, \forall i. \end{aligned}$$

Portanto, $D_i \sim \text{Geométrica}(\pi), \forall i$ (com $i = 1, \dots, N(T)$). Segundo este resultado, é possível construir testes equivalentes aqueles baseados na sequência I_T , utilizando as durações entre violações do VaR. Nesta abordagem, são consideradas os seguintes procedimentos de teste.

- Para testar a CC correta com respeito a $\Omega_{1,t}$, testamos $D_i \stackrel{i.i.d}{\sim} \text{Geométrica}(p) \forall i$.
- Para testar a independência, dada por $I_t \stackrel{i.i.d}{\sim} \text{Bernoulli}(\pi), \forall t$, testamos $D_i \stackrel{i.i.d}{\sim} \text{Geométrica}(\pi) \forall i$, para $\pi \in (0, 1)$.
- Para testar a UC correta, assumimos $I_t \stackrel{i.i.d}{\sim} \text{Bernoulli}(\pi)$, e testamos $E(D_i) = 1/p \forall i$.

O uso destas durações também tem uma motivação prática. Christoffersen e Pelletier [2004] destacam que em presença de agrupamento das violações do VaR, descritas na seção (3.7), a maioria das durações obtidas devem ser pequenas (durante períodos de alta volatilidade da série) e algumas longas (em períodos calmos). Assim, as durações apresentam evidência contra a hipótese de independência das variáveis indicadoras em I_T quando isso acontece.

4.4.1 Testes Baseados na Distribuição Weibull Discreta

A distribuição Geométrica de D_i é caracterizada pela seguinte função de risco:

$$\begin{aligned} \lambda_d^i &= P(D_i = d | D_i \geq d) = P(D_i = d) / P(D_i \geq d) \\ &= P(I_{t_{i-1}+d} = 1 | I_{t_{i-1}+d-1} = 0, \dots, I_{t_{i-1}+1} = 0, I_{t_{i-1}} = 1). \end{aligned} \quad (4.11)$$

Pela definição da i -ésima duração, $d = t_i - t_{i-1}$, deste modo $t_i = t_{i-1} + d$ é o tempo em que ocorreu a i -ésima violação do VaR. Assim, λ_d^i mede a probabilidade de ocorrer i -ésima violação do VaR no tempo t_i , dado que não ocorreram violações nos $d - 1$ períodos passados. Ou alternativamente, a probabilidade de ocorrer a i -ésima violação depois de d períodos.

Se F_T tivesse uma CC correta com respeito a $\Omega_{1,t}$, então a função de risco λ_d^i deve ser constante igual a p . Por outro lado, se a previsão do VaR não for eficiente com respeito a $\Omega_{1,t}$, a função de risco não é mais constante. A probabilidade de $I_{t+1} = 1$ poderia depender da sequência passada I_t, I_{t-1}, \dots . Berkowitz *et al.* [2011] propõe o Teste Geométrico que permite testar conjuntamente a dependência temporal da sequência de violações I_T e a UC correta, especificando a seguinte função de risco para as durações:

$$\lambda_d^i = \lambda_{WD}(d) = \pi d^{b-1}, \quad (4.12)$$

que corresponde à distribuição Weibull Discreta com parâmetros $0 \leq \pi < 1$ e $0 < b \leq 1$, especificada em Stein e Dattero [1984]. Esta variável assume valores em $\{1, 2, 3, \dots\}$ e será denotada por $W.D(\pi, b)$. Esta função de risco especifica a dependencia temporal das violações da seguinte forma: a probabilidade de acontecer a violação no tempo t_i , depende do tempo sem ocorrer violações depois da $(i-1)$ -ésima violação, ou seja depende de sua duração. Omitimos o índice i , pois não interessamos a ordem, apenas o valor da duração.

Assumindo que a função de risco das durações é dada por $\lambda_{WD}(d)$, Berkowitz *et al.* [2011] propõe o teste Geométrico de CC correta, que pode ser descomposto em teste de UC correta e de independência da duração. A hipótese de independência da duração é dada por $b = 1$, ou seja a duração não tem memoria e a função de risco é constante. Assumindo esta independência, a hipótese de UC correta é $\pi = p$. A hipótese de CC correta no teste Geométrico é dada por $\pi = p$ e $b = 1$. Como hipótese alternativa, a probabilidade de acontecer uma violação no tempo t_i é grande, se a duração nesse tempo for curta, se a duração for longa, então a probabilidade de acontecer uma violação é baixa, ou seja, a hipótese alternativa é $\pi \neq p$ ou $b < 1$.

Para a construção destes procedimentos, utilizamos um TRVG, para o qual precisamos de uma função de log-verossimilhança com censura na primeira e na última duração.

a) Função de log-verossimilhança com durações censuradas

Para implementar os testes baseados em durações, transformamos a sequência $I_{t_o+1}, \dots, I_{T+1}$ em as durações $D_1, \dots, D_{N(T)}$, com um tipo particular de censura (para simplificar a notação consideramos $N(T) = N$). Consideremos a sequência $C_1, \dots, C_{N(T)}$, em que:

$$C_i = \begin{cases} 1 & , \text{ se duração } D_i \text{ é censurada} \\ 0 & , \text{ caso contrario} \end{cases} .$$

Se a sequência I_T começa com um 0, então D_i é o número de dias até a primeira violação e consideramos $C_1 = 1$, pois D_1 é censurada a esquerda. Se $I_{T+1} = 0$, D_N é o número de dias após a última violação, assim $C_N = 1$, pois D_N é censurada à direita. Somente são censuradas a primeira e última duração, as restantes são calculadas segundo a definição (4.5).

Se $D_i \sim W.D(\pi, b)$, a função de probabilidade e de sobrevivência são obtidas utilizando λ_d^i , segundo as propriedades (2.5) e (2.6) da seção (2.6).

$$\begin{aligned} f_{WD}(d) &= P(D_i = d) = \lambda_{WD}(d) \prod_{j=1}^{d-1} \lambda_{WD}(j) \\ &= \begin{cases} \pi d^{b-1} \prod_{j=1}^{d-1} (1 - \pi j^{b-1}) & , \text{ se } d \in \{2, 3, \dots\} \\ \pi & , \text{ se } d = 1. \end{cases} \end{aligned}$$

e

$$S_{WD}(d) = P(D_i > d) = \prod_{j=1}^d (1 - \pi j^{b-1}).$$

Se a duração D_i com valor observado igual a d_i não é censurada, contribui à função de verossimilhança com $f_{WD}(d_i)$. Se a duração é censurada, só conhecemos que D_i será maior que d_i , sua contribuição é dada por $S_{WD}(d_i)$.

Assumindo que os tempos de censura são independentes das durações e que a censura é não informativa, a função de log-verossimilhança do parâmetro $\theta = (\pi, b)$, correspondente à amostra $\mathbf{D}^* = ((d_1, c_1), \dots, (d_N, c_N))^\top$, é dada por:

$$\log \mathcal{L}_{WD}(\theta; \mathbf{D}^*) = c_1 \log S_{WD}(d_1) + (1 - c_1) \log f_{WD}(d_1) \quad (4.13)$$

$$+ \sum_{i=2}^{N-1} \log f_{WD}(d_i) + c_N \log S_{WD}(d_N) + (1 - c_N) \log f_{WD}(d_N).$$

Os seguintes procedimentos, utilizam a função de risco $\lambda(d)_{W.D}$ para construir os testes de avaliação de \mathbf{F}_T com respeito a $\Omega_{1,t}$.

b) Teste Weibull Discreto de Cobertura não Condicional

Assumindo que $D_i \stackrel{i.i.d}{\sim} \text{Geométrica}(\pi)$, $\forall i$ (ou seja $b = 1$), o procedimento testa a hipótese nula $\lambda(d)_{W.D} = p$, $\forall i, d$, contra a alternativa $\lambda(d)_{W.D} \neq p \forall i, d$. As hipóteses do teste são:

$$\begin{cases} H_0 : \pi = p \\ H_1 : \pi \neq p \end{cases}.$$

Considerando $\Theta_0 = \{\pi : \pi = p\}$ e $\Theta = \{\pi : 0 \leq \pi < 1\}$. Sob H_1 o EMV de π é

$$\hat{\pi} = \left(\frac{N - C_1 - C_N}{\sum_{i=1}^N D_i} \right). \quad (4.14)$$

Seja $\theta^0 = (p, 1)$ e $\hat{\theta} = (\hat{\pi}, 1)$, a estatística TRVG para a UC correta, denotada por WD_{uc} , é dada por:

$$WD_{uc} = -2 \left(\log \mathcal{L}_{W.D}(\theta^0; \mathbf{D}^*) - \log \mathcal{L}_{W.D}(\hat{\theta}; \mathbf{D}^*) \right).$$

c) Teste Weibull Discreto de Independência da Duração

Este procedimento testa a hipótese nula $\lambda(d)_{W.D} = \pi \forall i, d$, ou seja, que a probabilidade de ocorrer uma violação não depende de sua duração. São consideradas as seguintes hipóteses:

$$\begin{cases} H_0 : b = 1 \\ H_1 : b \neq 1, \end{cases}$$

em que $\Theta_0 = \{(\pi, b) : 0 \leq \pi < 1, b = 1\}$ e $\Theta = \{(\pi, b) : 0 \leq \pi < 1, 0 < b \leq 1\}$.

Seja $\hat{\theta}^0 = (\hat{\pi}, 1)$, para π dado em (4.14), e $\hat{\theta} = (\hat{\pi}, \hat{b})$ é obtido via otimização numérica. A estatística do TRVG para a independência, denotada por WD_{ind} , é dada por:

$$WD_{ind} = -2 \left(\log \mathcal{L}_{W.D}(\hat{\pi}; \mathbf{D}^*) - \log \mathcal{L}_{W.D}(\hat{\theta}; \mathbf{D}^*) \right).$$

d) Teste Weibull Discreto de Cobertura Condicional

Para testar a hipótese de CC correta, o teste Geométrico, verifica conjuntamente as hipóteses de UC correta e independência da duração, considerando as seguintes hipóteses:

$$\begin{cases} H_0 : \pi = p \text{ e } b = 1 \\ H_1 : \pi \neq p \text{ ou } b \neq 1 \end{cases}.$$

Considerando $\Theta_0 = \{(\pi, b) = (p, 1)\}$ e $\Theta = \{(\pi, b) : 0 \leq \pi < 1, 0 < b \leq 1\}$. Seja $\theta^0 = (p, 1)$ e $\hat{\theta} = (\hat{\pi}, \hat{b})$, a estatística do TRVG para a CC correta, denotada por WD_{cc} , é dada por:

$$WD_{cc} = -2 \left(\log \mathcal{L}_{W.D}(\theta^0; \mathbf{D}^*) - \log \mathcal{L}_{W.D}(\hat{\theta}; \mathbf{D}^*) \right).$$

A abordagem dos testes de avaliação utilizando as durações, foi introduzido por [Christoffersen e Pelletier \[2004\]](#), que consideraram uma função de risco continua com parâmetros $0 \leq \pi < 1$ e $0 < b \leq 1$. Esta escolha tem a seguinte motivação: a correspondente distribuição Weibull, gera um excessivo número de curtas e algumas longas durações, que corresponde a durações geradas por agrupamentos das violações do VaR.

4.4.2 Testes usando Modelos de Regressão Weibull Discretos

A seguir, apresentamos a construção dos testes baseados em modelos de regressão discretos, utilizando um modelo paramétrico para a função de risco λ_d^i , definida em (4.11), condicionada a um vetor de covariáveis $\mathbf{x}_i(d)$. Isto permitirá estender o teste Geométrico, considerando a influência de covariáveis na probabilidade de acontecer uma violação. A construção deste modelo, apresentado em Kalbfleisch e Prentice [2011], utiliza uma função de risco *basal*, que denotamos por λ_d . Para este tipo de modelo, a construção da função da log-verossimilhança é descrita em Fahrmeir L [2001].

Posteriormente, apresentaremos algumas opções para modelar a função de risco das durações que incluem a previsão do *Var* como covariável. Deste modo, é possível construir procedimentos baseados nas durações, para testar a CC correta com respeito a $\Omega_{1,t}^*$.

a) Modelo paramétrico para a função de risco dependente das covariáveis

Seja $\mathbf{x}_i(d) = (x_{i1}, \dots, x_{id})^\top$ o vetor de covariáveis observado, não influenciado pelos tempos de falha (do tipo externa). Consideramos a x_{ij} como o elemento do vetor $\mathbf{x}_i(d)$, no tempo em que $D_i = j$, para $j = 1, \dots, d$. A função de risco de D_i , dependente do vetor $\mathbf{x}_i(d)$, é definida por:

$$\lambda_d^i = \lambda^i(d|\mathbf{x}_i(d)) = P(D_i = d | D \geq d; \mathbf{x}_i(d)). \quad (4.15)$$

O principal objetivo desta abordagem é estudar a possível dependência entre λ_d^i e a informação contida no vetor $\mathbf{x}_i(d)$. Para modelar λ_d^i , consideremos a função de ligação $g(\cdot)$ e o preditor η_{id} , que depende do tempo d tal que $\lambda_d^i = g(\eta_{id})$. Em particular, podemos considerar o preditor linear :

$$\eta_{id} = \beta_{0d} + z_{id}\beta, \quad (4.16)$$

em que β_{0d} é o efeito basal dependente do tempo d , z_{id} é um elemento do vetor de covariáveis $\mathbf{z}_i(d) = (z_{i1}, \dots, z_{id})^\top$, construído a partir do vetor de covariáveis original $\mathbf{x}_i(d)$ e β é o efeito da covariável z_{id} . Os modelos para λ_d^i são obtidos segundo a escolha de $g(\cdot)$.

De modo geral, consideramos:

$$h(\lambda_d^i) = h(\lambda_d) + z_{id}\beta, \quad (4.17)$$

em que λ_d é uma função de risco discreta no tempo d , quando $z_{id} = 0$, e $h(\cdot)$ uma função monótona crescente e diferenciável, tal que $h : [0, 1] \rightarrow \mathcal{R}$ e $h(0) = -\infty$. Algumas possíveis escolhas para h são:

$$h(u) = \log(-\log(1 - u)), \quad (4.18)$$

$$h(u) = \log(u), \quad (4.19)$$

$$h(u) = \log(u/(1 - u)) \quad (4.20)$$

e

$$h(u) = \Phi^{-1}(u), \quad (4.21)$$

em que Φ é a função de distribuição acumulada da normal padrão. Considerando para os quatro casos que $g : \mathcal{R} \rightarrow [0, 1]$, tal que $g(u) = h^{-1}(u)$, o modelo geral para λ_d^i pode ser expresso como:

$$\lambda_d^i = g(\eta_{id}) = g(h(\lambda_d) + z_{id}\beta). \quad (4.22)$$

Os modelos particulares para λ_d^i são obtidos selecionando uma função de risco λ_d e a função de ligação $g(\cdot)$, associada à função $h(\cdot)$. Para as funções $h(\cdot)$ em (4.18), (4.19), (4.20) e (4.21) as funções de ligação respectivas são: complementar *log-log*, $g(u) = 1 - \exp\{-\exp(u)\}$; exponencial, $g(u) = \exp(u)$; logit, $g(u) = \exp(u)\{1 + \exp(u)\}^{-1}$; e probit, $g(u) = \Phi(u)$. Em todos os casos, o efeito basal é dado por $\beta_{0d} = h(\lambda_d)$. Note que, se $\beta = \mathbf{0}$, então $\lambda_d^i = \lambda_d, \forall i$.

Assumindo a censura dada na construção de (4.13), a contribuição da i -ésima duração, D_i , à função de verossimilhança, é dada por:

$$L_i = P(D_i = d_i | \mathbf{x}_i(d))^{1-c_i} P(D_i > d_i | \mathbf{x}_i(d))^{c_i}.$$

Logo, pelas propriedades (2.5) e (2.6), temos que:

$$\begin{aligned} L_i &= \left(\lambda_{d_i}^i \right)^{1-c_i} \prod_{j=1}^{d_i-1} (1 - \lambda_j^i)^{1-c_i} \prod_{j=1}^{d_i} (1 - \lambda_j^i)^{c_i} \\ &= \left(\lambda_{d_i}^i \right)^{1-c_i} (1 - \lambda_{d_i}^i)^{c_i} \prod_{j=1}^{d_i-1} (1 - \lambda_j^i). \end{aligned}$$

Assim, a função de verossimilhança de $\theta = (\pi, b, \beta)$, correspondete à amostra $\mathbf{D}^* = ((d_1, C_1, \mathbf{x}_1(d_1)), \dots, (d_N, C_N, \mathbf{x}_N(d_N)))^\top$, é dada por:

$$\mathcal{L}_{WD}(\theta; \mathbf{D}^*) = \prod_{i=1}^N \left(\left(\lambda_{d_i}^i \right)^{1-c_i} (1 - \lambda_{d_i}^i)^{c_i} \prod_{j=1}^{d_i-1} (1 - \lambda_j^i) \right),$$

e a respectiva função de log-verossimilhança tem a forma:

$$\begin{aligned} \log \mathcal{L}_{WD}(\theta; \mathbf{D}^*) &= \sum_{i=1}^N \left((1 - c_i) \log(\lambda_{d_i}^i) + c_i \log(1 - \lambda_{d_i}^i) \right) \\ &\quad + \sum_{i=1}^N \sum_{j=1}^{d_i-1} \log(1 - \lambda_j^i) \\ &= \sum_{i=1}^N \log(\lambda_{d_i}^i) + \sum_{i=1}^N c_i \log \left(\frac{1 - \lambda_{d_i}^i}{\lambda_{d_i}^i} \right) + \sum_{i=1}^N \sum_{j=1}^{d_i-1} \log(1 - \lambda_j^i) \\ &= \sum_{i=1}^N \log(\lambda_{d_i}^i) + \sum_{i=1}^N \sum_{j=1}^{d_i-1} \log(1 - \lambda_j^i) \\ &\quad + c_1 \log \left(\frac{1 - \lambda_{d_1}^1}{\lambda_{d_1}^1} \right) + c_N \log \left(\frac{1 - \lambda_{d_N}^N}{\lambda_{d_N}^N} \right), \end{aligned} \tag{4.23}$$

em que $\sum_{j=1}^{d_i-1} \log(1 - \lambda_j^i) = 0$ se $d_i = 1$.

b) Modelo paramétrico para a função de risco dependente das durações e da previsão do VaR

Lai [2013] mostra que a função $\lambda(k)$, $k = \{1, 2, \dots\}$, é uma função de risco discreta se e somente se $0 \leq \lambda(k) \leq 1$ e $\sum_{i=1}^{+\infty} \lambda(i) = +\infty$. Além disso, são apresentadas formas de construir uma distribuição de tempos de falha discretos a partir de uma distribuição contínua. Duas dessas formas são: discretização da função de distribuição acumulada e discretização da função de risco.

Segundo Rinne [2010], a função de risco $\lambda_{WD}(d)$, apresentada em (4.12), para a distribuição Weibull discreta do Tipo II, é obtida por meio da discretização da função de risco continua $\lambda_W(d) = a^b d^{b-1}$, considerando o parâmetro $\pi = a^b b$, como uma combinação dos parâmetros de escala e forma a e b . Desta forma, obtemos

$$\lambda_{WD}(d) = \begin{cases} \pi d^{b-1} & , \text{ para } d = 1, 2, \dots, m \\ 0 & , \text{ para } d = 0 \text{ ou } d > m \end{cases} ,$$

em que $\pi > 0$, $b > 0$ e m é o valor truncado dado por

$$m = \begin{cases} \lceil \pi^{-1/(b-1)} \rceil & , \text{ se } b > 1 \\ +\infty & , \text{ se } 0 < b \leq 1 \end{cases} .$$

Um caso particular para o modelo da função de risco das durações (4.22), dependente da previsão do VaR, é dada por:

$$\lambda_d^i = g(h(\lambda_{WD}(d)) + \beta VaR_{t_{i-1}+d}), \quad (4.24)$$

considerando $\lambda_{WD}(d)$ como a função de risco basal λ_d e $\mathbf{z}_i(d) = (VaR_{t_{i-1}+1}, \dots, VaR_{t_{i-1}+d})^\top$. Por simplicidade de notação, a previsão do VaR no tempo $D_i = j$ é dada por $VaR_{t_{i-1}+j} = VaR_{t_{i-1}+j|t_{i-1}+j-1}(p)$. Neste caso, x_{ij} representa um elemento de $\mathbf{x}_i(d)$, que contém a informação necessária para obter $VaR_{t_{i-1}+j}$, ou seja os retornos até o tempo $t_{i-1} + j - 1$.

Com as especificações de h em (4.18), (4.19), (4.20) e (4.21), obtemos as seguintes equações para λ_d^i :

$$\lambda_d^i = 1 - (1 - \lambda_{WD}(d))^{\exp\{\beta VaR_{t_{i-1}+d}\}}, \quad (4.25)$$

$$\lambda_d^i = \lambda_{WD}(d) e^{\beta VaR_{t_{i-1}+d}}, \quad (4.26)$$

$$\lambda_d^i = \frac{e^{h^*(\lambda_{WD}(d)) + \beta VaR_{t_{i-1}+d}}}{1 + e^{h^*(\lambda_{WD}(d)) + \beta VaR_{t_{i-1}+d}}}, \quad (4.27)$$

$$\lambda_d^i = \Phi(\Phi^{-1}(\lambda_{WD}(d)) + \beta VaR_{t_{i-1}+d}), \quad (4.28)$$

em que $h^*(u) = \log(u/(1-u))$ em (4.27).

Segundo o modelo (4.26), considerado em Pelletier e Wei [2014], a função de risco das durações dependente da previsão do VaR pode ser expressada como:

$$\begin{aligned} \lambda_d^i &= P(I_{t_{i-1}+d} = 1 | I_{t_{i-1}+d-1} = 0, \dots, I_{t_{i-1}+1} = 0, I_{t_{i-1}} = 1, \mathbf{x}_i(d)) \\ &= \pi d^{b-1} e^{\beta VaR_{t_{i-1}+d}}, \text{ para } \beta \geq 0. \end{aligned}$$

Baseados nesta especificação de λ_d^i , a probabilidade de ocorrer uma violação no tempo t_i depende de sua duração e da previsão do VaR feita no instante $t_i - 1$. Utilizando este modelo, podemos considerar os seguintes testes de avaliação.

- De independência da previsão do VaR, testando $\beta = 0$.
- De independência da duração e da previsão do VaR, testando $b = 1$ e $\beta = 0$.
- De CC correta, testando conjuntamente $\pi = p$, $b = 1$ e $\beta = 0$.

c) Teste Weibull Discreto de Independência do VaR

O procedimento testa se λ_d^i , depende da previsão do VaR, considerando as seguintes hipóteses.

$$\begin{cases} H_0 : \beta = 0 \\ H_1 : \beta > 0 \end{cases} .$$

Seja $\Theta_0 = \{(\pi, b, \beta) : 0 \leq \pi < 1, 0 < b \leq 1, \beta = 0\}$ e $\Theta = \{(\pi, b, \beta) : 0 \leq \pi < 1, 0 < b \leq 1, \beta \geq 0\}$. A estatística do TRVG para a independência da previsão do VaR, denotada por WD_{Vind} , é dada por:

$$WD_{Vind} = -2 \left(\log \mathcal{L}_{WD}(\hat{\theta}^0; \mathbf{D}) - \log \mathcal{L}_{WD}(\hat{\theta}; \mathbf{D}) \right),$$

em que os estimadores de θ sob H_0 e H_1 , $\hat{\theta}^0 = (\hat{\pi}, \hat{b})$ e $\hat{\theta} = (\hat{\pi}, \hat{b}, \hat{\beta})$, respectivamente, são obtidos usando otimização numérica.

d) Teste Weibull Discreto de Independência da duração e da previsão do VaR

É possível construir um procedimento análogo ao teste CAViaR de independência, para testar se a probabilidade de ocorrer as violações depende de sua duração e da previsão do VaR, considerando as seguintes hipóteses:

$$\begin{cases} H_0 : b = 1 \text{ e } \beta = 0 \\ H_1 : b \neq 1 \text{ ou } \beta > 0 \end{cases}.$$

Seja $\Theta_0 = \{(\pi, b, \beta) : 0 \leq \pi < 1, b = 1, \beta = 0\}$ e $\Theta = \{(\pi, b, \beta) : 0 \leq \pi < 1, 0 < b \leq 1, \beta \geq 0\}$. A estatística do TRVG, denotada por WD_{DVaR} , é dada por:

$$WD_{DVaR} = -2 \left(\log \mathcal{L}_{WD}(\hat{\theta}^0; \mathbf{D}) - \log \mathcal{L}_{WD}(\hat{\theta}; \mathbf{D}) \right),$$

em que, sob H_0 , o EMV de θ^0 é $\hat{\theta}^0 = (N - C_1 - C_N) / \sum_{i=1}^N D_i$ e sob H_1 , $\hat{\theta} = (\hat{\pi}, \hat{b}, \hat{\beta})$ que é obtido utilizando otimização numérica. Denominamos o procedimento como teste de independência Duração-VaR.

e) Teste de Cobertura Condicional Geométrico-VaR

Em [Pelletier e Wei \[2014\]](#) o procedimento é denominado como teste Geométrico-VaR, pois combina o teste Geométrico e o teste de independência da previsão do VaR, considerando as seguintes hipóteses:

$$\begin{cases} H_0 : \pi = p \text{ e } b = 1 \text{ e } \beta = 0 \\ H_1 : \pi \neq p \text{ ou } b \neq 1 \text{ ou } \beta > 0 \end{cases},$$

em que $\Theta_0 = \{(\pi, b, \beta) = (p, 1, 0)\}$ e $\Theta = \{(\pi, b, \beta) : 0 \leq \pi < 1, 0 < b \leq 1, \beta \geq 0\}$. A estatística do TRVG para testar a CC correta, denotada por WD_{GV} , é dada por:

$$WD_{GV} = -2 \left(\log \mathcal{L}_{WD}(\theta^0; \mathbf{D}) - \log \mathcal{L}_{WD}(\hat{\theta}; \mathbf{D}) \right),$$

em que sob H_0 , $\theta^0 = (p, 1, 0)$ e sob H_1 , $\hat{\theta} = (\hat{\pi}, \hat{b}, \hat{\beta})$ é obtido via otimização numérica.

Neste trabalho estamos interessados em comparar os desempenhos dos modelos (4.26), (4.27) e (4.28) através do poder empírico via simulação, em os dois últimos casos $\beta \in \mathcal{R}$. A implementação dos testes é feita no *software* R.

Os testes que consideram uma distribuição Weibull contínua ou discreta para D_i tem a seguinte dificuldade: os valores dos parâmetros propostos em H_0 , por exemplo $b = 1$ e $\beta = 0$, estão na fronteira do espaço paramétrico Θ , o qual pode afetar a distribuição assintótica da estatística do TRVG. Além disso, os dados que temos são binários (\mathbf{I}_T) e a distribuição desta estatística não é contínua. Para contornar estes problemas é sugerida uma técnica de Monte Carlo para obter o *p-valor*.

4.4.3 Testes baseados no Método dos Momentos Generalizado (MMG)

A abordagem do MMG para testar a CC correta com respeito a $\Omega_{1,t}$ é apresentado em [Candelon et al. \[2010\]](#), utilizando os polinômios ortogonais da distribuição Geométrica. A principal motivação desta abordagem é o seguinte resultado: se $M_j(\cdot)$ é o polinômio ortogonal correspondente à variável aleatória Y , então $E(M_j(Y)) = 0$ $j \in \mathcal{N}$, ver [Bontemps \[2006\]](#).

O $(j + 1)$ -ésimo polinômio ortogonal associada à distribuição Geométrica, é definido como:

Definição 4.6 *O polinômio ortogonal associado à distribuição Geométrica, com probabilidade de sucesso π , é definido pela seguinte relação recursiva:*

$$M_{j+1}(d, \pi) = \frac{(1 - \pi)(2j + 1) + \pi(j - d + 1)}{(j + 1)\sqrt{1 - \pi}} M_j(d; \pi) - \left(\frac{j}{j + 1} \right) M_{j-1}(d, \pi), \quad (4.29)$$

para $\forall d \in \mathcal{N}^*$ e $j \in \mathcal{N}$, com $M_{-1}(d, \pi) = 0$ e $M_0(d, \pi) = 1$.

Seja D_1, \dots, D_N uma sequência de N durações entre violações do VaR. Se $D_i \stackrel{i.i.d}{\sim} \text{Geométrica}(\pi) \forall i$, então temos:

$$E(M_j(D_i; \pi)) = 0, \forall j \in \mathcal{N} \quad \forall D_i \in \mathcal{N}^*, \quad (4.30)$$

que são as condições de momentos utilizadas na construção do teste. Na prática trabalhamos com $j \leq k$, em que k é um número inteiro arbitrário.

a) Teste do tipo MMG de Cobertura Condicional

Sob a suposição de CC correta, $D_i \stackrel{i.i.d}{\sim} \text{Geométrica}(\pi) \forall i$, então a hipótese nula de CC correta pode ser escrita como:

$$H_0 : E(M_j(D_i; p)) = 0, \quad j = \{1, \dots, k\},$$

em que k é o número de polinômios ortogonais. A estatística do teste, denotada por $J_{cc}(k)$, é dada na seguinte proposição:

Proposição 4.1 *Assumindo que o processo das durações $\{D_i : i \geq 1\}$ é estacionário e ergódico. Sob a hipótese de Cobertura Condicional, temos*

$$J_{cc}(k) = \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbf{M}(D_i, p) \right)^\top \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbf{M}(D_i, p) \right) \xrightarrow[N \rightarrow +\infty]{D} \chi_k^2,$$

em que $\mathbf{M}(D_i, p)$ denota um vetor cujos componentes são os polinômios ortogonais $M_j(D_i; p)$, para $j = \{1, \dots, k\}$.

A demonstração decorre do lemma 4.4, apresentado em Hansen [1982].

b) Teste do tipo MMG de Cobertura não Condicional

A propriedade de UC correta estabelece que $P(I_{t+1} = 1) = p, \forall t$. Assim, assumindo que $D_i \sim \text{Geométrica}(\pi) \forall i$, \mathbf{F}_T tem uma UC correta se $\pi = p$ ou $E(D_i) = 1/p, \forall i$.

Na abordagem MMG, hipótese de UC correta pode ser expressada por:

$$H_0 : E(M_1(D_i; p)) = 0,$$

como $M_1(D_i; p) = (1 - pD_i)/\sqrt{1 - p}$, a condição

$$E(M_1(D_i; p)) = \frac{1 - pE(D_i)}{\sqrt{1 - p}} = 0,$$

que é equivalente a $E(D_i) = 1/p, \forall i$.

A estatística para o teste, denotada por J_{uc} , é obtida como um caso particular da estatística $J_{cc}(k)$, quando é considerado somente o primeiro polinômio ortogonal, ou seja $\mathbf{M}(D_i, p) = M_1(D_i, p)$. Assim J_{uc} é equivalente a $J_{cc}(1)$, dada por:

$$J_{uc} = \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N M_1(D_i; p) \right)^2 \xrightarrow[N \rightarrow +\infty]{D} \chi_1^2.$$

c) Teste do tipo MMG de Independência

A propriedade de independência afirma que as violações do VaR deveriam ser distribuídas independentemente, com a mesma taxa de cobertura. Sob esta hipótese temos que $D_i \sim \text{Geométrica}(\pi)$, $\forall i$, em que π não necessariamente igual p . Nesta abordagem, a hipótese de independência pode ser expressada por:

$$H_0 : E(M_j(D_i; \pi)) = 0, \forall j = 1, \dots, k .$$

A estatística $J_{ind}(k)$, para testar esta hipótese, é dada por:

$$J_{ind}(k) = \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbf{M}(D_i, \pi) \right)^\top \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbf{M}(D_i, \pi) \right) \xrightarrow[N \rightarrow +\infty]{D} \chi_k^2,$$

em que $\mathbf{M}(D_i, \pi)$ denota o vetor cujos componentes são os polinômios ortogonais $M_j(D_i, \pi)$, $\forall i = 1, \dots, k$. Contudo, neste caso, o verdadeiro valor π é desconhecido, conseqüentemente, a estatística de teste de independência deve ser baseada em polinômios ortogonais que dependem do parâmetro estimado, ou seja, utilizamos $M_j(D_i, \hat{\pi})$ em lugar de $M_j(D_i, \pi)$, em que $\hat{\pi}$ é um estimador consistente de π , com taxa de convergência igual a $1/\sqrt{N}$. A substituição de π por seu estimador pode mudar a distribuição assintótica da estatística $J_{ind}(k)$. [Candelon et al. \[2010\]](#) indica que com esta modificação, a estatística de independência pode ser da forma:

$$J_{ind}(k) = \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbf{M}(D_i, \hat{\pi}) \right)^\top \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbf{M}(D_i, \hat{\pi}) \right) \xrightarrow[N \rightarrow +\infty]{D} \chi_{k-1}^2,$$

neste trabalho consideramos $\hat{\pi}$ como o EMV de π dado por $\hat{\pi} = T^{-1} \sum_{i=1}^T I_t$.

No presente trabalho, consideramos um número de polinômios ortogonais igual a 3 e 5. A estatística para testar a CC correta, para $k = 5$ é dada por:

$$\begin{aligned} J_{cc}(5) &= \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbf{M}(D_i, p) \right)^\top \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbf{M}(D_i, p) \right) \\ &= \sum_{j=1}^5 \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N M_j(D_i, p) \right)^2 \xrightarrow[N \rightarrow +\infty]{D} \chi_5^2, \end{aligned}$$

em que $\mathbf{M}(D_i, p)$ denota o vetor cujos componentes são $M_j(D_i, p)$, para $j = \{1, \dots, 5\}$ dado por:

$$\mathbf{M}(D_i, p) = (M_1(D_i; p), \dots, M_5(D_i; p))^\top.$$

Também, a estatística para o teste de independência é dada por:

$$J_{ind}(5) = \sum_{j=1}^5 \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N M_j(D_i, \hat{\pi}) \right)^2 \xrightarrow[N \rightarrow +\infty]{D} \chi_4^2.$$

4.5 Considerações finais

Há outros trabalhos não considerados aqui, mas que devem ser mencionados, por exemplo, [Dumitrescu et al. \[2012\]](#), que estende o teste de Quantis Dinâmico para estudar a dependência não linear entre a probabilidade das violações do VaR e variáveis explicativas e o trabalho apresentado em [Araújo Santos e Fraga Alves \[2012\]](#) que propõe um teste de nível α^* para testar a hipótese de independência. Além dos testes, existem outras opções para avaliar a previsão do VaR, como a abordagem do teste de densidade de previsão apresentado em [Berkowitz \[2001\]](#).

Capítulo 5

Simulações e Aplicações

5.1 Introdução

Neste capítulo avaliamos o desempenho de alguns testes apresentados no Capítulo 4. Observamos a partir dos estudos de simulação, o comportamento de cada teste através do poder empírico. Os desempenhos das previsões do VaR geradas pelos métodos POT e DPOT também são avaliadas utilizando os retornos diários do Ibovespa.

Na literatura, o procedimento comumente empregado para calcular o poder empírico é:

- Gerar dados de um processo heterocedástico .
- Utilizar o método de Simulação Histórica (SH) para fazer as previsões do VaR. Essas previsões não são eficientes devido aos agrupamentos de violações causados por a heterocedasticidade do modelo.
- Sob a hipótese de eficiência, o poder empírico é calculado pelas frequências de rejeições da hipótese de eficiência destas previsões em um determinado número de replicas, segundo o p-valor de Monte Carlo.

Neste estudo de simulação comparativa, estamos interessados em avaliar os métodos para testar as hipóteses de independência e cobertura condicional. Para o primeiro caso, consideramos os testes MMG, CAViaR e o Duração-VaR. Para o teste de cobertura condicional, os testes MMG, CAViaR e Geométrico-VaR são empregados. O objetivo principal é comparar o poder empírico obtido, utilizando três tipos de processos. Mencionamos também que todas as simulações e aplicações são feitas via *software* R.

5.2 Estudo de Simulação Comparativa

Neste estudo de simulação utilizamos processos do tipo heteroscedástico para gerar os dados. A seguir apresentamos os processos geradores dos dados (PGD) que consideramos neste trabalho.

O PGD 1 é gerado pelo modelo *APARCH*(1,1):

$$r_{t+1} = \sigma_{t+1}\epsilon_{t+1},$$
$$\sigma_{t+1}^\delta = w + \alpha(|r_t| - \gamma r_t)^\delta + \beta\sigma_t^\delta,$$

em que ϵ_t segue uma distribuição *t*-Student assimétrica com v graus de liberdade e parâmetro de assimetria dado por ϑ . Os parâmetros são selecionados como em [Araújo Santos e Fraga Alves \[2012\]](#), em que $w = 0,03$, $\alpha = 0,086$, $\gamma = 0,64$, $\beta = 0,91$, $\delta = 1,15$, $\vartheta = 0,88$ e $v = 10$.

O PGD 2 é dado pelo $GARCH(1,1)$ - $t(v)$ com efeitos de assimetria, dado por :

$$r_{t+1} = \sigma_{t+1} \sqrt{\frac{v-2}{v}} \epsilon_{t+1},$$

$$\sigma_{t+1}^2 = w + \alpha \sigma_t^2 \left(\sqrt{\frac{v-2}{v}} \epsilon_t - \theta \right) + \beta \sigma_t^2,$$

em que ϵ_{t+1} tem distribuição t-Student com v graus de liberdade. Os parâmetros são selecionados como em Candelon *et al.* [2010], em que $\alpha = 0,1$, $\theta = 0,5$, $\beta = 0,85$, $w = 3,9683e^{-6}$ e $v = 8$.

O PGD 3 é gerado pelo modelo $AR(3)$ - $TGARCH(1,1)$:

$$r_{t+1} = \phi_1 r_t + a_{t+1},$$

$$a_{t+1} = \sigma_{t+1} \epsilon_{t+1},$$

$$\sigma_{t+1}^2 = w + \alpha a_t^2 + \gamma a_t^2 d_t + \beta \sigma_t^2,$$

em que $\epsilon_{t+1} \sim N(0,1)$. Os parâmetros são selecionados como em Morettin [2008], em que $\phi = -0,051$, $w = 0,00013$, $\alpha = 0,044$, $\gamma = 0,063$ e $\beta = 0,910$. Neste caso $d_t = 1$ se $a_t < 0$ e $d_t = 0$ no caso contrário.

Nas simulações que apresentamos, são considerados os tamanhos pós-amostrais $\mathcal{P} = 250, 500, 750$ e 1000. As previsões do VaR são feitas empregando os métodos descritos na seção (2.5), considerando um horizonte $h = 1$, $ws = 250$ e 500, e as probabilidades de cobertura $p = 0,01$ e 0,05.

O critério de seleção das amostras é o seguinte: para os testes do tipo MMG e $CAViaR$ precisamos de pelo menos 2 violações do VaR, para os testes Duração-VaR e Geométrico-VaR precisamos de pelo menos três durações.

5.2.1 Probabilidade do Erro de Tipo I e Poder Empírico

A seguir apresentamos o poder empírico dos testes via simulação, adotando o p-valor de Monte Carlo. Para justificar que o uso do p-valor de Monte Carlo é mais adequado comparando com a região crítica estabelecida pela região assintótica da estatística do teste para calcular o poder empírico, a probabilidade do erro de tipo I é calculada nas simulações, utilizando os valores críticos assintóticos.

Probabilidade do erro de tipo I empírica

Se a região de rejeição, obtida utilizando a distribuição assintótica da estatística do teste, for correta (ou precisa ou exata) para os tamanhos de amostras que consideramos, a frequência de rejeição sobre um número grande de amostras simuladas sob H_0 , deveria ser próxima à probabilidade do erro de tipo I ($\alpha^* = 0,1$) fixada (por exemplo, α^*). Para o teste de independência Duração-VaR e Geométrico-VaR, consideramos a distribuição qui-quadrado com 2 e 3 graus de liberdade, respectivamente.

Com o objetivo de verificar se a região crítica é corretamente especificada ou não, via simulação, em cada replica geramos sequências i.i.d. de variáveis aleatórias com distribuição Bernoulli de parâmetro p . Adicionalmente, para os testes que utilizam a previsão do VaR como covariável, são calculadas previsões do VaR utilizando o método de simulação histórica, a partir da PGD 2, independentes da sequência Bernoulli. Utilizando 5000 réplicas e $p = 0,05$, a frequência de rejeições representa a proporção de estatísticas que caíram na região de rejeição.

A probabilidade empírica do erro de tipo I para os testes de independência está apresentada na Tabela 5.1, considerando o PGD 2 e $ws = 250$ e $ws = 500$.

Tabela 5.1: Probabilidade empírica do erro de tipo I, para os testes de independência. PGD 2, $\alpha^* = 0,1$, $p = 0,05$.

	$ws = 250$				$ws = 500$			
	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$
MMG_3	0,001	0,002	0,007	0,009	0,012	0,026	0,038	0,039
MMG_5	0,001	0,002	0,006	0,007	0,010	0,018	0,026	0,025
$CAViaR$	0,072	0,064	0,058	0,058	0,085	0,103	0,126	0,120
$DVaR_0$	0,010	0,009	0,005	0,009	0,006	0,008	0,012	0,010
$DVaR_1$	0,010	0,009	0,005	0,009	0,006	0,009	0,012	0,010
$DVaR_2$	0,010	0,009	0,005	0,009	0,006	0,008	0,012	0,010
$FAR1$	0,277	0,039	0,004	2×10^{-4}	0	0	0	0
$FAR2$	0,282	0,040	0,004	2×10^{-4}	0	0	0	0

Nota: Para os testes do tipo Duração-VaR, denotado por $DVaR_i$, os subíndices $i = 0, 1, 2$ designam o uso da ligações exponencial, logito e probito, respectivamente. Denotamos por MMG_k o teste do tipo MMG, utilizando k polinômios ortogonais. Reportamos a frequência de amostras rejeitadas FAR1 para os testes MMG e CAViaR e FAR2 para o teste Duração-VaR.

Tabela 5.2: Probabilidade empírica do erro de tipo I, para os testes de cobertura condicional. PGD 2, $\alpha^* = 0,1$, $p = 0,05$.

	$ws = 250$				$ws = 500$			
	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$
MMG_3	0,049	0,067	0,071	0,073	0,058	0,062	0,068	0,068
MMG_5	0,042	0,051	0,053	0,058	0,047	0,047	0,049	0,055
$CAViaR$	0,083	0,100	0,099	0,058	0,084	0,103	0,108	0,114
$GVaR_0$	0,026	0,030	0,025	0,123	0,027	0,028	0,023	0,025
$GVaR_1$	0,026	0,030	0,025	0,028	0,027	0,028	0,023	0,025
$GVaR_2$	0,027	0,030	0,025	0,029	0,028	0,028	0,023	0,025
$FAR1$	0,157	0,032	0,003	1×10^{-4}	0	0	0	0
$FAR2$	0,271	0,038	0,004	1×10^{-4}	0	0	0	0

Nota: Para os testes do tipo Geométrico-VaR, denotado por $GVaR_i$, os subíndices $i = 0, 1, 2$ designam o uso da ligações exponencial, logito e probito, respectivamente. Denotamos por MMG_k o teste do tipo MMG, utilizando k polinômios ortogonais. Reportamos a frequência de amostras rejeitadas FAR1 para os testes MMG e CAViaR e FAR2 para o teste Geométrico-VaR.

Pela Tabela 5.1, podemos observar que os testes do tipo MMG e Duração-VaR apresentam uma probabilidade estimada menor que α^* . O teste do tipo CAViaR em alguns casos é menor e em outros maior. As probabilidades empíricas do erro de tipo I, considerando o PGD 2 e $p = 0,05$ dos testes de cobertura condicional apresentadas na Tabela 5.2. Observamos um comportamento similar ao caso anterior.

Os resultados apresentados nas Tabelas 5.1 e 5.2 estão convergentes com as conclusões obtidas na literatura, que para os tamanhos amostrais considerados, o uso dos valores críticos assintóticos não é adequado para calcular o poder empírico dos testes, se desejamos ter controlada a probabilidade do erro do tipo I. Os autores propõem utilizar a técnica de Monte Carlo, apresentada em Dufour [2006], para obter o p-valor e decidir sobre a rejeição de H_0 , para maiores detalhes ver o Apêndice A.2.

Poder empírico

Os resultados da simulação para obter o poder empírico dos testes de independência são apresentados nas Tabelas 5.3 e 5.4, e o poder dos testes de cobertura condicional nas Tabelas 5.5 à 5.8.

Em todos os casos, foi utilizado o p-valor de Monte Carlo para decidir sobre a rejeição de H_0 .

Para obter o poder empírico via simulação, assumimos que os corretos processos geradores dos dados são o PGD 1, o PGD 2 e o PGD 3. Como em outros estudos na literatura, utilizamos o método de Simulação Histórica, para obter as previsões do VaR e com elas a sequência \mathbf{I}_T . Estas previsões geram agrupamento das violações, devido aos processos heteroscedásticos simulados, indicando uma forte evidência para rejeitar a hipótese de independência, e consequentemente da hipótese de CC correta. O poder empírico é calculado utilizando a frequência de rejeições da respectiva H_0 em 5000 réplicas.

Tabela 5.3: Poder empírico para os testes de independência. PGD 1, $\alpha^* = 0,1$, $ws = 250$.

	$p = 0,01$				$p = 0,05$			
	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$
MMG_3	0,363	0,543	0,716	0,840	0,582	0,859	0,956	0,981
MMG_5	0,444	0,616	0,752	0,858	0,571	0,860	0,948	0,980
$CAViaR$	0,539	0,640	0,713	0,771	0,547	0,710	0,798	0,858
$DVaR_0$	0,495	0,732	0,868	0,942	0,664	0,918	0,981	0,993
$DVaR_1$	0,505	0,741	0,869	0,943	0,673	0,919	0,981	0,993
$DVaR_2$	0,582	0,773	0,874	0,946	0,477	0,789	0,922	0,967
$FAR1$	0,216	0,011	0	0	0,013	0	0	0
$FAR2$	0,219	0,011	0	0	0,014	0	0	0

Tabela 5.4: Poder empírico para os testes de independência. PGD 2, $\alpha^* = 0,1$, $p = 0,05$.

	$ws = 250$				$ws = 500$			
	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$
MMG_3	0,487	0,236	0,755	0,836	0,349	0,581	0,745	0,830
MMG_5	0,308	0,296	0,747	0,827	0,451	0,577	0,752	0,828
$CAViaR$	0,258	0,484	0,676	0,757	0,419	0,486	0,579	0,653
$DVaR_0$	0,506	0,413	0,846	0,866	0,282	0,616	0,791	0,863
$DVaR_1$	0,470	0,413	0,845	0,865	0,282	0,616	0,791	0,863
$DVaR_2$	0,494	0,417	0,851	0,871	0,284	0,617	0,792	0,864
$FAR1$	0,002	0,003	0	0	0,004	0	0	0
$FAR2$	0,002	0,003	0	0	0,004	0	0	0

Nota: Para os testes do tipo Duração-VaR, denotado por $DVaR_i$, os subíndices $i = 0,1,2$ designam o uso da ligações exponencial, logito e probito, respectivamente. Denotamos por MMG_k o teste do tipo MMG, utilizando k polinômios ortogonais. Reportamos a frequência de amostras rejeitadas FAR1 para os testes MMG e CAViaR e FAR2 para o teste Duração-VaR.

O poder empírico dos testes de independência são apresentados na Tabela 5.3, considerando o PGD 1 e $ws = 250$. Tanto para $p = 0,01$ quanto para $p = 0,05$, todos os testes apresentaram poderes mais altos para os casos com $\mathcal{P} = 750$ e $\mathcal{P} = 1000$. Os testes de Duração-VaR possuem maiores poderes para $\mathcal{P} \geq 500$ comparando com outros testes, mostrando que é mais eficiente.

Os resultados de poder empírico considerando o PGD 2 para $p = 0,05$, são apresentados na Tabela 5.4. Em geral, a Tabela 5.4 mostra que o grupo de testes Duração-VaR foi mais poderoso. O alto poder empírico depende do tamanho da pós-amostra \mathcal{P} e não do tamanho da amostra móvel ws .

Com respeito ao poder dos testes de cobertura condicional na Tabela 5.5 apresentamos os resultados, considerando o PGD 1 e $ws = 250$. Neste caso, os testes do tipo Geométrico-VaR apresentam maiores poderes para $\mathcal{P} \geq 500$, tanto para $p = 0,01$ quanto para $p = 0,05$. Além

disso, neste caso, o teste CAViaR também apresenta bons desempenhos, ao contrário do teste de independência.

Tabela 5.5: Poder empírico para os testes de cobertura condicional. PGD 1, $\alpha^* = 0,1$, $ws = 250$.

	$p = 0,01$				$p = 0,05$			
	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$
MMG_3	0,379	0,619	0,685	0,780	0,687	0,866	0,935	0,971
MMG_5	0,389	0,661	0,741	0,827	0,688	0,898	0,930	0,968
$CAViaR$	0,599	0,732	0,786	0,871	0,691	0,832	0,804	0,864
$GVaR_0$	0,491	0,729	0,845	0,940	0,765	0,932	0,974	0,993
$GVaR_1$	0,478	0,732	0,852	0,941	0,749	0,934	0,974	0,992
$GVaR_2$	0,525	0,772	0,860	0,944	0,796	0,940	0,975	0,992
$FAR1$	$4,2 \times 10^{-3}$	0,002	0	0	0,013	0	0	0
$FAR2$	$5,2 \times 10^{-3}$	0,002	0	0	0,014	0	0	0

Nota: Para os testes do tipo Geométrico-VaR, denotado por $GVaR_i$, os subíndices $i = 0,1,2$ designam o uso da ligações exponencial, logito e probito, respectivamente. Denotamos por MMG_k o teste do tipo MMG, utilizando k polinômios ortogonais. Reportamos a frequência de amostras rejeitadas FAR1 para os testes MMG e CAViaR e FAR2 para o teste Geométrico-VaR.

Tabela 5.6: Poder empírico para os testes de cobertura condicional. PGD 2, $\alpha^* = 0,1$, $p = 0,05$.

	$ws = 250$				$ws = 500$			
	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$
MMG_3	0,427	0,549	0,637	0,743	0,459	0,599	0,696	0,777
MMG_5	0,427	0,550	0,633	0,745	0,462	0,600	0,695	0,772
$CAViaR$	0,503	0,560	0,652	0,719	0,461	0,538	0,566	0,630
$GVaR_0$	0,430	0,572	0,703	0,793	0,477	0,621	0,752	0,824
$GVaR_1$	0,429	0,567	0,695	0,787	0,477	0,619	0,750	0,820
$GVaR_2$	0,430	0,570	0,693	0,790	0,476	0,620	0,748	0,823
$FAR1$	0,008	0,003	0	0	$3,8 \times 10^{-3}$	0	0	0
$FAR2$	0,008	0,003	0	0	$3,8 \times 10^{-3}$	0	0	0

Nota: Para os testes do tipo Geométrico-VaR, denotado por $GVaR_i$, os subíndices $i = 0,1,2$ designam o uso da ligações exponencial, logito e probito, respectivamente. Denotamos por MMG_k o teste do tipo MMG, utilizando k polinômios ortogonais. Reportamos a frequência de amostras rejeitadas FAR1 para os testes MMG e CAViaR e FAR2 para o teste Geométrico-VaR.

O poder dos testes de cobertura condicional, considerando o PGD 2 e $p = 0,05$, é apresentado na Tabela 5.6. Mesmos comportamentos ao teste de independência, o alto poder empírico depende do tamanho da pós- amostra \mathcal{P} e não do tamanho da amostra móvel ws . Os testes do tipo Geométrico-VaR apresentam melhores desempenhos para $\mathcal{P} \geq 750$. Para $\mathcal{P} \leq 500$, todos os testes apresentaram desempenhos semelhantes.

Na Tabela 5.7 observamos os resultados quando é considerado o PGD 3 e $p = 0,01$. Para $ws = 250$ o teste CAViaR foi mais poderoso para $\mathcal{P} \leq 750$ e para $\mathcal{P} = 1000$ o teste $GVaR_2$. Para $ws = 500$, o teste $GVaR_2$ foi mais poderoso para $\mathcal{P} \geq 750$. Considerando o mesmo PGD, na Tabela 5.8 observamos os resultados de poder empírico quando $p = 0,05$. Para $ws = 250$, o teste $GVaR_2$ foi mais poderoso para $\mathcal{P} \geq 500$. Para $ws = 500$, o teste MMG_3 foi mais poderoso para $\mathcal{P} \leq 750$ e para $\mathcal{P} = 1000$ o teste $GVaR_2$.

Dos resultados obtidos, podemos concluir que o grupo do testes Duração-VaR e Geométrico-VaR apresentam melhores desempenhos para o teste de independência e também de CC, respectivamente, para tamanhos pós-amostrais $\mathcal{P} \geq 750$. Os testes do tipo MMG e CAViaR apresentam um poder empírico levemente menor.

Tabela 5.7: Poder empírico para os testes de cobertura condicional. PGD 3, $\alpha^* = 0,1$, $p = 0,01$.

	$ws = 250$				$ws = 500$			
	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$
MMG_3	0,525	0,598	0,697	0,780	0,492	0,599	0,688	0,776
MMG_5	0,548	0,628	0,740	0,814	0,540	0,648	0,741	0,806
$CAViaR$	0,662	0,710	0,785	0,849	0,570	0,666	0,713	0,762
$GVaR_0$	0,575	0,563	0,765	0,841	0,594	0,607	0,773	0,840
$GVaR_1$	0,574	0,562	0,765	0,841	0,593	0,608	0,773	0,840
$GVaR_2$	0,584	0,588	0,779	0,857	0,597	0,624	0,781	0,850
$FAR1$	0,004	2×10^{-4}	0	0	$2,6 \times 10^{-3}$	$1,4 \times 10^{-3}$	2×10^{-4}	0
$FAR2$	0,202	0,009	2×10^{-4}	0	0,371	0,089	0,011	0,001

Nota: Para os testes do tipo Geométrico-VaR, denotado por $GVaR_i$, os subíndices $i = 0, 1, 2$ designam o uso da ligações exponencial, logito e probito, respectivamente. Denotamos por MMG_k o teste do tipo MMG, utilizando k polinômios ortogonais. Reportamos a frequência de amostras rejeitadas FAR1 para os testes MMG e CAViaR e FAR2 para o teste Geométrico-VaR.

Tabela 5.8: Poder empírico para os testes de cobertura condicional. PGD 3, $\alpha^* = 0,1$, $p = 0,05$.

	$ws = 250$				$ws = 500$			
	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$	$\mathcal{P} = 250$	$\mathcal{P} = 500$	$\mathcal{P} = 750$	$\mathcal{P} = 1000$
MMG_3	0,610	0,780	0,877	0,939	0,659	0,849	0,924	0,960
MMG_5	0,603	0,763	0,864	0,928	0,645	0,843	0,913	0,956
$CAViaR$	0,630	0,702	0,752	0,814	0,633	0,737	0,762	0,794
$GVaR_0$	0,594	0,788	0,892	0,944	0,652	0,827	0,921	0,961
$GVaR_1$	0,593	0,787	0,891	0,944	0,652	0,827	0,921	0,961
$GVaR_2$	0,604	0,810	0,906	0,951	0,657	0,836	0,924	0,965
$FAR1$	0,009	0	0	0	0,04	0,001	0	0
$FAR2$	4×10^{-4}	0	0	0	0,002	0	0	0

Nota: Para os testes do tipo Geométrico-VaR, denotado por $GVaR_i$, os subíndices $i = 0, 1, 2$ designam o uso da ligações exponencial, logito e probito, respectivamente. Denotamos por MMG_k o teste do tipo MMG, utilizando k polinômios ortogonais. Reportamos a frequência de amostras rejeitadas FAR1 para os testes MMG e CAViaR e FAR2 para o teste Geométrico-VaR.

5.3 Aplicações

Nesta seção apresentamos a avaliação das previsões do VaR para uma posição comprada de horizonte $h = 1$, utilizando os métodos POT e DPOT apresentados nas seções (3.5) e (3.6), respectivamente. Nosso interesse é avaliar a UC correta, a independência e a CC correta. Consideramos uma probabilidade de cobertura $p = 0,05$, $\mathcal{P} = 1000$ e para a amostra móvel, cujos tamanhos são $ws = 500$ e 1000 .

A série utilizada para obter os log-retornos são os valores do índice da IBOVESPA, a pós-amostra é obtida usando os valores do índice no período de 21/07/98 a 07/08/02. Para $ws = 500$ a amostra inicial é obtida usando os valores no período de 15/07/96 a 20/07/98 e para $ws = 1000$ no período de 04/07/94 a 20/07/98.

Nas Figuras 5.1.a e 5.1.b observamos a série do índice da Ibovespa e seus log- retornos durante o período total considerado (04/07/94-07/08/02).

A Tabela 5.9 apresenta as estatísticas descritivas dos log-retornos.

Para obter a sequência de previsões do VaR, utilizando os métodos POT e DPOT, consideramos as sugestões feitas por Araújo Santos e Fraga Alves [2013], em que o limiar μ é escolhido de tal modo que 10% dos retornos da amostra utilizada para estimar o modelo sejam maiores que μ . No caso

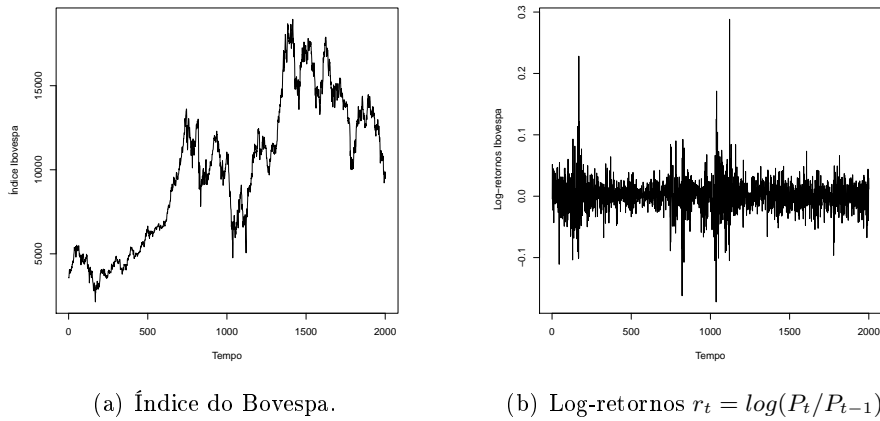


Figura 5.1: Gráfico do Ibovespa e os log-retornos no período de 04/07/1994 a 07/08/02 .

Tabela 5.9: Estatísticas dos log-retornos do índice da Ibovespa

Média	0,0005
Mediana	0,0012
Desvio Padrão	0,0274
Assimetria	0,5885
Curtose	14,7740
1º Quartil	-0,0128
3º Quartil	0,0153
Mínimo	-0,1723
Máximo	0,2882

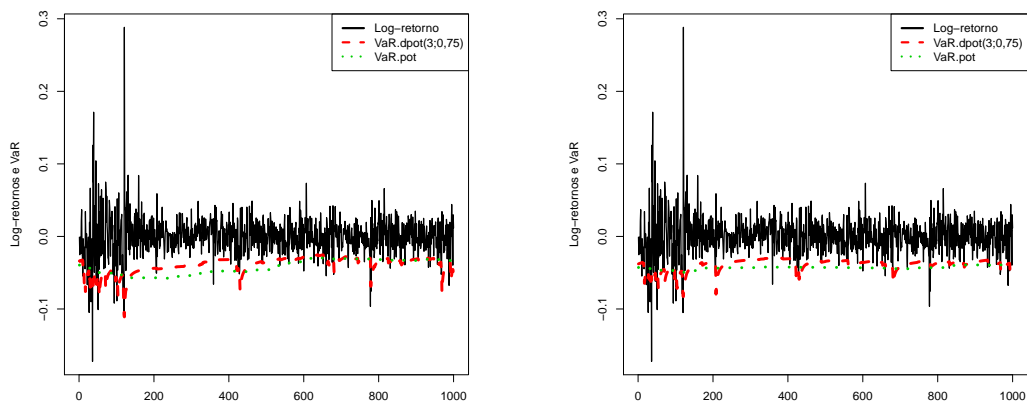


Figura 5.2: Pós-amostra e previsões do VaR utilizando os métodos POT e DPOT .

particular do modelo DPOT, consideramos $\lambda = 3$ e $c = 0.75$. Na Figura 5.2.a e 5.2.b observamos a série pós-amostral e as correspondentes previsões do VaR, utilizando $ws = 500$ e 1000 .

Para avaliar a UC correta, utilizamos o teste de Kupiec. Para avaliar a independência, utilizamos os testes do tipo MMG, CAViaR e Duração-VaR. Para a CC correta, consideramos os testes do tipo MMG, CAViaR e Geométrico-VaR. Em todos os casos, obtemos o p-valor de Monte Carlo e um nível de significância $\alpha^* = 0,05$.

Tabela 5.10: Avaliação das previsões do VaR gerada pelos métodos POT e DPOT.

	$ws = 500$		$ws = 1000$	
	POT	DPOT	POT	DPOT
Frequência de violações	0,048	0,047	0,037	0,047
N_1	47	48	37	47
N_2	49	48	38	48
Teste de Cobertura não Condicional				
Kupiec p-valor	0,829	0,725	0,198	0,769
Testes de Independência				
MMG_3 p-valor	<0,001	0,011	0,001	0,017
CAViaR p-valor	0,003	0,201	<0,001	0,144
$DVaR_0$ p-valor	<0,001	<0,001	<0,001	<0,001
Testes de Cobertura Condicional				
MMG_3 p-valor	<0,001	0,020	0,002	0,034
CAViaR p-valor	0,321	0,317	0,205	0,248
$GVaR_0$ p-valor	0,002	0,063	<0,001	0,026

Nota: Reportamos p-valor de Monte Carlo para cada teste. Os testes $GVaR_0$ e $DVaR_0$ corresponde ao teste Duração-Var e Geométrico-VaR com ligação exponencial.

Com respeito à hipótese de CC correta para $ws = 500$ e 1000 , o teste MMG_3 rejeita a hipótese nula em todos os casos, o teste CAViaR não rejeita H_0 em ambos os casos. O teste Geométrico-VaR não rejeita H_0 para o método DPOT, somente no caso em que $ws = 500$.

As previsões do VaR obtidas são comparadas com a pós-amostra para obter a sequência I_T e as correspondentes durações. Na Tabela 5.10 apresentamos as frequências de violações do VaR, o número de durações sem e com censura (N_1 e N_2), e os correspondentes p-valores, via simulação de Monte Carlo, dos testes.

Tabela 5.11: Tabela de rejeições para $\alpha^* = 0,05$.

Teste	Hipótese UC	$ws=500$		$ws=1000$	
		POT	DPOT	POT	DPOT
Kupiec	Não rejeita	Não rejeita	Não rejeita	Não rejeita	Não rejeita
Hipótese Independência					
MMG_3	Rejeita	Rejeita	Rejeita	Rejeita	Rejeita
CAViaR	Rejeita	Não Rejeita	Rejeita	Não Rejeita	Rejeita
Duração- VaR	Rejeita	Rejeita	Rejeita	Rejeita	Rejeita
Hipótese CC					
MMG_3	Rejeita	Rejeita	Rejeita	Rejeita	Rejeita
CAViaR	Não rejeita	Não rejeita	Não rejeita	Não rejeita	Não rejeita
Geométrico-VaR	Rejeita	Não rejeita	Rejeita	Rejeita	Rejeita

A Tabela 5.11 resume os resultados da avaliação da sequência de previsão do VaR. O teste

Kupiec não rejeita a UC correta em todos os casos. Com respeito à hipótese de independência, considerando $ws = 500$ e 1000 , o único teste que não rejeita H_0 é teste do tipo CAViaR para o método DPOT.

Como os resultados das simulações mostraram que o teste da classe Duração-VaR e Geométrico-VaR é mais poderoso para $\mathcal{P} \geq 500$, sugerimos que as conclusões deste análise devem ser feitas de acordo com os resultados do teste $DVaR_0$ e $GVaR_0$.

Além disso, observamos diferenças nas avaliações do teste CAViaR de independência e a correta cobertura condicional. Da avaliação separada: as previsões tem uma UC correta, mas os agrupamentos das violações do VaR acontecidas no período de alta volatilidade inicial forneceram uma forte evidência em contra a hipóteses de independência. Também podemos concluir que as considerações feitas sobre o método DPOT para a escolha do limiar μ e λ feitas no estudo anterior, não foram satisfatórias.

Capítulo 6

Conclusões e Trabalhos Futuros

6.1 Conclusões

O objetivo deste trabalho foi comparar os desempenhos dos testes de avaliação da previsão do VaR através de um estudo de simulação comparativa. Além disso, avaliamos um novo método de previsão do VaR, o qual é aplicado nos retornos diários do Ibovespa.

Os resultados da probabilidade empírica do erro de tipo I evidenciam que, para os tamanhos amostrais considerados, o uso dos valores críticos assintóticos para calcular o poder empírico não é adequado se desejamos controlar a probabilidade do erro do tipo I, portanto, procedimentos como a técnica de Monte Carlo para obter o p-valor são recomendados.

Para testes de independência, tanto para PGD 1 quanto para PGD 2, todos os testes apresentaram poderes altos para os casos em que $\mathcal{P} = 750$ e $\mathcal{P} = 1000$. Destacamos que os testes de Duração-VaR possuem maiores poderes para $\mathcal{P} \geq 500$, comparando com outros testes, mostrando um melhor desempenho. Com respeito ao teste de cobertura condicional para os PGD 1, PGD 2 e PGD 3, o teste Geométrico-VaR apresenta um maior poder empírico para $\mathcal{P} \geq 500$.

Em aplicações à séries financeiras, as previsões do VaR obtidas pelo método POT geram, frequentemente, agrupamento das violações. O método DPOT foi introduzido com o objetivo de superar esta deficiência. Os testes de avaliação foram uma ferramenta útil para verificar se as considerações feitas sobre o limiar μ e o número de excessos passados λ foram adequadas para a implementação do método DPOT. Com respeito à avaliação das previsões geradas pelos métodos POT e DPOT, utilizando os testes que tiveram um melhor desempenho nas simulações, rejeitamos a CC correta, excluindo o caso das previsões do método DPOT, para $ws = 250$. Mostramos também, como a avaliação separada foi útil para determinar a possível causa da rejeição da hipótese de CC correta em três casos: a rejeição da hipótese de independência. Nesta aplicação, os tamanhos escolhidos para as amostras móveis $ws = 500$ e $ws = 1000$ levaram a conclusões diferentes em relação à eficiência das previsões.

6.2 Trabalhos futuros

As conclusões deste trabalho são limitadas, uma vez que elas foram feitas baseadas em simulações de alguns cenários. Para um trabalho futuro pretendemos considerar outros valores da amostra móvel ws , do tamanho da pós-amostra e avaliar também previsões do VaR de uma posição vendida, com o objetivo de obter conclusões mais consistentes.

Apêndice A

Apêndice

A.1 Domínio Máximo de Atração

Dizemos que a variável aleatória X (ou a função de distribuição F de X , ou a distribuição de X) pertence ao domínio máximo de atração de uma distribuição do valor extremo H , se existem as constantes $c_n > 0$ e $d_n \in R$ tal que $\frac{X-d_n}{c_n} \xrightarrow[n \rightarrow +\infty]{D} H$.

A.2 P-valor de Monte Carlo

A técnica do teste de Monte Carlo para calcular o p-valor, apresentada em Dufour [2006], é implementada da seguinte forma. Geramos R independentes realizações dos dados sob H_0 . Para cada teste considerado, obtemos a estatística do teste LR_i , para a i -ésima realização dos dados, para $i = 1, \dots, R$. Seja a LR_0 a estatística obtida da amostra original.

Como estamos trabalhando com uma distribuição discreta, existe a probabilidade não nula de que LR_0 ser igual a LR_i . Para quebrar os empates, geramos, independentemente, uma variável $U_i \sim \mathcal{U}(0, 1)$, $i = 1, \dots, R$, para cada estatística LR_i . Então, o p-valor de Monte Carlo é obtido por:

$$\hat{p}_R(LR_0) = \frac{R \hat{G}(LR_0) + 1}{R + 1},$$

para

$$\hat{G}(LR_0) = 1 - \frac{1}{R} \sum_{i=1}^R I_{\{LR_i \leq LR_0\}} + \frac{1}{R} \sum_{i=1}^R I_{\{LR_i = LR_0\}} I_{\{U_i = U_0\}},$$

em que I_A é a variável indicadora do evento A . Rejeitamos a hipótese nula se $\hat{p}_R(LR_0) \leq \alpha^*$, para um determinado nível de significância α^* .

A.3 Programação em R: Teste Geométrico VaR

```
#####  
####  PROGRAMAÇÃO LOG-VEROSIMILHANÇA do modelo  ####  
####  DE REGRESSÃO WEIBULL DISCRETO                ####  
#####
```

```
### Funções g e h: modelo logaritmo linear ###  
h0<- function(u) {out<- log(u); return(out) }  
g0<- function(u) {out<- exp(u); return(out) }
```

```

#### Funções g e h: modelo logístico #####
g1<-function(u){out<- exp(u)/( 1+exp(u) ); return(out)}
h1<- function(u ){out<- log(u/(1-u)); return(out)}

#### Funções g e h: Modelo Probit  ####
g3<-function(u){out<- pnorm(u, mean = 0, sd = 1); return(out)}
h3<- function(u ){out<- qnorm(u, mean = 0, sd = 1); return(out)}

#### Função de Risco Dattero & Stein (1984)

lambda<- function(para,t){
pi<- para[1]
b<- para[2]
vetor<- rep(0,length(t))
for(s in 1:length(vetor)){vetor[s]<- pi*(t[s])^(b-1)}
return(vetor) }

#### Função de Logverossimilância negativa
#### Não inclui as funções complementares.
nLVCox <- function (theta, D,Indcen,Hit, VAR,lig){

if(lig=="0"){g<- g0 ; h<- h0}
if(lig=="1"){g<- g1 ; h<- h1}
if(lig=="3"){g<- g3 ; h<- h3}

pi <- theta[1]
b <- theta[2]
Beta <- theta[3]

V <- abs(VAR) # Co-variáveis positivas
N <- length(D)
# Tempos t_i + d das violações usando função tvarcen
tci <- tvarcen(Hit,D, Indcen)

# VaRj são os VaR_{ti + d}
VaRj <- rep (0, length(tci ))
for (i in 1: length ( tci )) { VaRj[i]<- varselect( V, tci[i])}

vector1 <- rep(0, N)
vector2 <- rep(0, N)

for(i in 1:N){

vector1[i]<- log( g( h( lambda(c(pi,b),D[i]) ) - Beta*VaRj[i] ) )
tor<- tdorigcen(Hit,D,Indcen)

j <- seq(1, D[i]-1)
var<- 0
for(t in 1:length(j) ){var[t]<- varselect(V,tor[i]+ t) }

if(D[i]!="1"){ vector2[i] <- sum(log(1- g( h( lambda(c(pi,b),j) )

```

```

- Beta*var ))} else {vector2[i]= 0 }

}

nll <- - ( sum(vector1)
+ Indcen[1]*log( (1- g ( h(lambda(c(pi,b),D[1]))
- Beta*VaRj[1]) ) /g ( h(lambda(c(pi,b),D[1])) - Beta*VaRj[1]) )
+ Indcen[N]*log( (1- g (h(lambda(c(pi,b),D[N]))
- Beta*VaRj[N]))/g (h(lambda(c(pi,b),D[N])) - Beta*VaRj[N]))
+ sum( vector2) )

return(nll)
}

{Teste Geométrico-VaR}

GeoVaR <- function(conf.level ,p , outsample, VaR,lig,star )
{
cd_am <- Durcen (outsample ,VaR)
D <-cd_am$D
c <-cd_am$C
N<-length(D)
hit = ifelse(outsample < VaR, 1, 0)

#### Otimiza nLVCox #####

if(lig=="0" ){ LOWER<- c(0.001,0,0) ; UPPER<- c(0.999,1, 5) } # log()
if(lig=="1"){ LOWER<- c(0.001,-3) ;UPPER<- c(0.999,1, 5)} # logit
if(lig=="3" ){ LOWER<- c(0.001,0,-5) ; UPPER<- c(0.999,1, -5) } # probito

outnLVCox <- try( optim(par=star , nLVCox , D=D, Indcen=c, Hit=hit,
VAR=VaR,lig=lig, method= "L-BFGS-B",
lower = LOWER, upper = UPPER ), silent=TRUE)

if ( class(outnLVCox ) == "try-error"){ l.h1 <- NA ; Beta <- NA }
else{l.h1 <- - outnLVCox$value ;Beta<- outnLVCox$par[3] }

l.h0 <- - nLVGeo2(p,D,c)

RV_gv <- -2*( l.h0 - l.h1 )
return ( list(RV= RV_gv, Beta=Beta) )

}

### Nota: A programação do teste Duração-VaR é similar.

```


Referências Bibliográficas

- Araújo Santos e Fraga Alves(2012)** P Araújo Santos e MI Fraga Alves. A new class of independence tests for interval forecasts evaluation. *Computational Statistics & Data Analysis*, 56(11):3366–3380. Citado na pág. 25, 34, 35
- Araújo Santos e Fraga Alves(2013)** P Araújo Santos e MI Fraga Alves. Forecasting value-at-risk with a duration-based pot method. *Mathematics and Computers in Simulation*, 94:295–309. Citado na pág. 2, 11, 15, 40
- Baillie e Bollerslev(1992)** Richard T Baillie e Tim Bollerslev. Prediction in dynamic models with time-dependent conditional variances. *Journal of Econometrics*, 52(1):91–113. Citado na pág. 1
- Barone-Adesi et al.(2002)** Giovanni Barone-Adesi, Kostas Giannopoulos e Les Vosper. Backtesting derivative portfolios with filtered historical simulation (fhs). *European Financial Management*, 8(1):31–58. Citado na pág. 1
- Berkowitz(2001)** Jeremy Berkowitz. Testing density forecasts, with applications to risk management. *Journal of Business & Economic Statistics*, 19(4):465–474. Citado na pág. 34
- Berkowitz et al.(2011)** Jeremy Berkowitz, Peter Christoffersen e Denis Pelletier. Evaluating value-at-risk models with desk-level data. *Management Science*, 57(12):2213–2227. Citado na pág. 2, 23, 25, 26, 27
- Bickel e Doksum(1977)** PJ Bickel e K Doksum. *Mathematical Statistics: Basic Ideas and Selected Topics*. San Francisco: HoldenDay. Citado na pág. 8
- Bontemps(2006)** Christian Bontemps. Moment-based tests for discrete distributions. *Toulouse School of Economics: Working Paper*. Citado na pág. 32
- Candelon et al.(2010)** Bertrand Candelon, Gilbert Colletaz, Christophe Hurlin e Sessi Tokpavi. Backtesting value-at-risk: a gmm duration-based test. *Journal of Financial Econometrics*, página nbq025. Citado na pág. 2, 32, 34, 36
- Christoffersen e Pelletier(2004)** Peter Christoffersen e Denis Pelletier. Backtesting value-at-risk: A duration-based approach. *Journal of Financial Econometrics*, 2(1):84–108. Citado na pág. 2, 26, 28
- Christoffersen(1998)** Peter F Christoffersen. Evaluating interval forecasts. *International economic review*, páginas 841–862. Citado na pág. 2, 17, 18, 19, 21
- Danielsson et al.(1998)** Jon Danielsson, Philipp Hartmann e Casper de Vries. The cost of conservatism. *Risk*, 11(1):101–103. Citado na pág. 1
- Danielsson et al.(2000)** Jón Danielsson, Yuji Morimoto e Nihon Ginkō. *Forecasting extreme financial risk: a critical analysis of practical methods for the Japanese market*. Institute for Monetary and Economic Studies, Bank of Japan. Citado na pág. 2

- Dowd(2007)** Kevin Dowd. *Measuring market risk*. John Wiley & Sons. Citado na pág. [1](#)
- Dufour(2006)** Jean-Marie Dufour. Monte carlo tests with nuisance parameters: A general approach to finite-sample inference and nonstandard asymptotics. *Journal of Econometrics*, 133(2):443–477. Citado na pág. [37](#), [47](#)
- Dumitrescu et al.(2012)** Elena-Ivona Dumitrescu, Christophe Hurlin e Vinson Pham. Backtesting value-at-risk: From dynamic quantile to dynamic binary tests. *Finance*, 33(1):79–112. Citado na pág. [34](#)
- Embrechts et al.(1997)** Paul Embrechts, Claudia Klüppelberg e Thomas Mikosch. *Modelling extremal events: for insurance and finance*, volume 33. Springer. Citado na pág. [14](#)
- Engle e Manganelli(2004)** Robert F Engle e Simone Manganelli. Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4):367–381. Citado na pág. [1](#)
- Engle e Manganelli(1999)** Robert F Engle e Simone Manganelli. Caviar: Conditional autoregressive value at risk by regression quantiles. Relatório técnico, University of California, San Diego. Citado na pág. [22](#), [23](#)
- Fahrmeir L(2001)** Tutz G Fahrmeir L. *Multivariate Statistical Modelling Based on Generalized Linear Models*. Springer Series in Statistics. Citado na pág. [29](#)
- Galdi e Pereira(2007)** Fernando Caio Galdi e Leonel Molero Pereira. Value at risk (var) using volatility forecasting models: Ewma, garch and stochastic volatility. *Brazilian Business Review*, 4(1):74–94. Citado na pág. [1](#)
- Haas(2005)** Markus Haas. Improved duration-based backtesting of value-at-risk. *Journal of Risk*, 8(2):17–38. Citado na pág. [2](#)
- Hansen(1982)** Lars Peter Hansen. Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, páginas 1029–1054. Citado na pág. [33](#)
- Kalbfleisch e Prentice(2011)** John D Kalbfleisch e Ross L Prentice. *The statistical analysis of failure time data*, volume 360. John Wiley & Sons. Citado na pág. [9](#), [29](#)
- Kupiec(1995)** Paul H Kupiec. Techniques for verifying the accuracy of risk measurement models. *THE J. OF DERIVATIVES*, 3(2). Citado na pág. [2](#)
- Lai(2013)** CD Lai. Issues concerning constructions of discrete lifetime models. *QUALITY TECHNOLOGY AND QUANTITATIVE MANAGEMENT*, 10(2):251–262. Citado na pág. [30](#)
- McNees e Fine(1995)** S McNees e LK Fine. Forecast uncertainty: can it be measured. *Federal reserve Bank of New York Discussion Paper*. Citado na pág. [1](#)
- McNeil e Frey(2000)** Alexander J McNeil e Rüdiger Frey. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of empirical finance*, 7(3):271–300. Citado na pág. [15](#)
- McNeil et al.(1998)** Alexander J McNeil, Alexander J McNeil e Alexander J McNeil. *Calculating quantile risk measures for financial return series using extreme value theory*. Departement Mathematik, Eidgenössische Technische Hochschule Zürich. Citado na pág. [1](#)
- Morettin(2008)** Pedro Alberto Morettin. *Econometria financeira: um curso em séries temporais financeiras*. Edgard Blücher. Citado na pág. [5](#), [13](#), [36](#)

- Morgan(1996)** JP Morgan. *Riskmetrics: technical document*. Morgan Guaranty Trust Company of New York. Citado na pág. [1](#)
- Pelletier e Wei(2014)** Denis Pelletier e Wei Wei. The geometric-var backtesting method. Relatório técnico, Working Paper, 28p. Citado na pág. [2](#), [31](#), [32](#)
- Rinne(2010)** Horst Rinne. *The Weibull distribution: a handbook*. CRC Press. Citado na pág. [30](#)
- Sen e Singer(1994)** Pranab K Sen e Julio M Singer. *Large sample methods in statistics: An introduction with applications*, volume 25. CRC Press. Citado na pág. [8](#)
- Smith(1987)** Richard L Smith. Estimating tails of probability distributions. *The annals of Statistics*, páginas 1174–1207. Citado na pág. [14](#)
- Stein e Dattero(1984)** William E Stein e Ronald Dattero. A new discrete weibull distribution. *IEEE transactions on reliability*, 33(2):196–197. Citado na pág. [27](#)
- Tashman(2000)** Leonard J Tashman. Out-of-sample tests of forecasting accuracy: an analysis and review. *International Journal of Forecasting*, 16(4):437–450. Citado na pág. [8](#)
- Tsay(2005)** Ruey S Tsay. *Analysis of financial time series*, volume 543. John Wiley & Sons. Citado na pág. [14](#)