

Medida condicional de *Expected Shortfall* nos mercados latinoamericanos através da teoria do valor extremo[☆]

Elsevier Inc^{a,b}, Rafael Felipe Bressan^{a,1,*}

^a Avenida Madre Benvenuta, 2007 - Santa Mônica Florianópolis - SC 88035-901

^b 360 Park Avenue South, New York

Abstract

This template helps you to create a properly formatted L^AT_EX manuscript.

Keywords: elsarticle.cls, L^AT_EX, Elsevier, template

2010 MSC: 00-01, 99-00

1. Introdução

A medição do risco de mercado. ao qual os portfólios dos investidores está sujeito é objeto de devoção de esforços tanto por parte das instituições e investidores em geral como por parte dos reguladores. Instituições financeiras em todo o mundo, de acordo com suas regulações locais e com os princípios de Basileia (Basel Committee on Banking Supervision - BCBS do Banco de Compensações Internacionais - BIS²) para aquelas que o seguem (o Brasil é um desses países) são obrigadas a reservar uma parcela de seu capital como provisionamento contra flutuações adversas do mercado em seus portfólios de investimento.

Uma importante característica das séries de retornos financeiros é sua alta volatilidade, não constante e tampouco seguindo a distribuição Normal. Assim, eventos extremos, e neste caso estamos interessados em perdas de grande magnitude, acontecem com uma frequência alta demais para serem descartadas como

[☆]Fully documented templates are available in the elsarticle package on CTAN.

^{*}Corresponding author

Email address: rafael.bressan@edu.udesc.br (Rafael Felipe Bressan)

URL: www.elsevier.com (Elsevier Inc)

¹Depto. de Economia/Esag/UDESC

²<http://www.bis.org/bcbs/index.htm?m=3%7C14>

apenas *outliers*, e portanto passaram a atrair a atenção dos participantes do
mercado, entre eles os investidores e também os reguladores. Estas observações
induziram uma gama enorme de estudos, práticos e teóricos, voltados a explicar
o comportamento dos retornos de séries financeiras e modelar de forma adequada
as caudas da distribuição destes retornos. Não somente estes estudos são de
grande relevância para o gerenciamento de risco nas instituições financeiras,
como também são obrigatórios segundo o acordo de Basileia, uma vez que este
requer o cálculo do Valor em Risco - VaR, para então a instituição poder projetar
o seu nível requerido de capital.

De acordo com os princípios de Basileia III, on Banking Supervision (2011,
2013a, 2014), as instituições financeiras supervisionadas pelos Bancos Centrais
devem manter *buffers* de capital contra riscos de mercado, crédito, liquidez,
entre outros. Dentro dos riscos de mercado, as duas formas mais usuais de fazer
a quantificação destes são os métodos de Valor em Risco - VaR e o *Expected
Shortfall* - ES. Este último relacionado ao primeiro, sendo definido como o valor
esperado das perdas que excedem o valor VaR calculado para um determinado
intervalo de confiança.

VaR é um quantil α da distribuição de perdas de um ativo ou portfólio em
um determinado período de tempo, ao passo que ES é o valor esperado das
perdas que excedem VaR, para um mesmo intervalo de confiança α e período.

O método VaR para cálculo de risco de mercado ao qual um portfólio
está sujeito foi primeiramente introduzido através de RiskMetrics (1995), uma
metodologia adotada pelo banco J. P. Morgan. Vem desde então sendo amplamente
adotado pela indústria financeira e largamente estudado pela academia.
Inúmeras variantes do modelo foram propostas e continuam sendo utilizadas com
o passar dos anos. Para o cálculo do VaR é necessária uma suposição acerca da
distribuição dos retornos, e por conseguinte do comportamento da cauda desta.

As variações na metodologia original de estimação do VaR surgem principalmente
em função de críticas a abordagem proposta, a qual inclui a suposição de
retornos independentes e igualmente distribuídos (*iid*), covariâncias constantes
entre os ativos de um portfólio e a distribuição normal dos retornos.

Percentuais
de VaR por
Basileia

Existem pe-
nalidades
regulatórias
para as IF
em que seu
modelo VaR
permite um
número
maior de
perdas do
que seria
estimado
pelo mod-
elo. Verificar
onde nos
princípios de
Basileia

45 Uma das últimas fortes críticas a medida do VaR, e que deu origem a outra métrica conhecida como *Expected Shortfall*, ou simplesmente ES, veio de dois artigos Artzner et al. (1997) e Artzner et al. (1999). Primeiramente foram introduzidas quatro propriedades cunhadas através de axiomas, as quais medidas coerentes de risco deveriam possuir, sendo elas:

- 50
- invariância a translação;
 - sub-aditividade;
 - homogeneidade positiva, e;
 - monotonicidade.

VaR especificamente não possui a propriedade da sub-aditividade para alguns
55 casos. Para contornar este fato, Acerbi & Tasche (2002) propõe o *Expected Shortfall* e comprovam que este é uma medida coerente de risco. Além de ser coerente, o ES possui uma segunda vantagem com relação ao VaR, considerando que o ES nos informa uma medida de tendência central do tamanho das perdas que excedem o valor do quantil VaR. Ou seja, o VaR nos informa apenas que uma
60 proporção α das perdas serão menor que a medida, mas nada nos informa se esta perda extraordinária de fato ocorrer. Mesmo sendo criticado e demonstradamente uma medida não coerente de risco, o VaR continua a ser amplamente utilizado, mesmo que agora em conjunto com o ES.

Mais recentemente o Comitê de Supervisão Bancária de Basileia tem se pro-
65 posto a adotar o *Expected Shortfall* como medida de risco de mercado. on Banking Supervision (2013b). O Comitê cita a grande importância da escolha da medida de risco e sua calibração, e portanto estas são relevantes para as decisões de política do Banco. Entre as dificuldades encontradas pelo VaR estão mais notadamente sua inabilidade em estimar o "risco de cauda" da distribuição de
70 perdas, uma vez que VaR não leva em conta a distribuição das perdas acima do valor de corte.

Desta forma, foi decidido que o ES seria a medida de risco favorita para a abordagem pelo banco chamada de modelos internos. Ou seja, os bancos

supervisionados devem utilizar o ES para o cálculo do risco de mercado a que
75 estão sujeitos em seus modelos internos. O comitê também se decidiu por um
intervalo de confiança de 97,5% para o ES, em contraposição a 99% para o
VaR. O comitê espera que esta abordagem para o cálculo da medida de risco de
mercado trará benefícios se comparada a antiga abordagem pelo Var, entre elas
um modelo com resultados mais estáveis e menor sensibilidade a observações
80 extremas (*outliers*).

Teoria do valor extremo. EVT da sigla em inglês, é um ramo da estatística que
lida diretamente com eventos raros, extremos. Seu objetivo é modelar eventos
que se distanciam muito da mediana de uma distribuição. Justamente por esta
característica, a EVT está sendo utilizada para modelar riscos que possuem
85 distribuição com caudas longas, um dos fatos estilizados que apresentamos sobre
retornos de ativos financeiros.

Ao utilizarmos a EVT, e mais especificamente o método conhecido como
peaks over threshold – POT, estamos interessados em modelar apenas a parte
da cauda da distribuição das perdas de um ativo financeiro maiores que um
90 determinado valor de limiar u . É da modelagem desta cauda, portanto, que
faremos as estimativas de risco VaR_α e ES_α .

2. O método *peaks over threshold*

Consideremos uma amostra de uma variável aleatória (v.a.) cujas observações
sejam independentes e igualmente distribuídas (i.i.d.), $L_i, i \in \mathbb{N}$, que represente
95 as perdas financeiras de um determinado ativo.

A EVT está interessada em investigar o comportamento da distribuição dos
máximos desta v.a. dados por $M_n = \max(L_1, \dots, L_n)$ para vários valores de n
e a medida que $n \rightarrow \infty$. A sequência M_n é chamada de máximos em bloco e
é possível demonstrar que a única distribuição para a qual M_n converge com
100 $n \rightarrow \infty$ é a distribuição de valores extremos generalizada (GEV, da sigla em
inglês).

Definição 1 (MDA). Domínio de máxima atração, se a equação (??) é válida para uma H não-degenerada, então se diz que $F \in MDA(H)$, F pertence ao domínio de máxima atração de H .

Teorema 2 (Fisher-Tippett, Gnedenko). Se $F \in MDA(H)$ para alguma H não-degenerada, então H deve ser uma distribuição do tipo de valores extremos generalizada – GEV.

O teorema 2 foi estabelecido através de três artigos, Fisher & Tippett (1928), Gnedenko (1941, 1943).

Definição 3 (GEV). Distribuição de valores extremos generalizada, é definida por sua p.d.f (função densidade de probabilidades) a qual é dada por:

$$H_{\xi}(x) = \begin{cases} \exp(-(1 + \xi x)^{-\frac{1}{\xi}}), & \xi \neq 0, \\ \exp(-e^{-x}), & \xi = 0, \end{cases} \quad (1)$$

O parâmetro ξ é conhecido como o parâmetro de forma da distribuição e dependendo deste valor teremos diferentes tipos de distribuição (casos particulares da GEV). Quando $\xi = 0$ a distribuição resultante é uma Gumbel, quando $\xi > 0$ uma Fréchet surge, e por fim quando $\xi < 0$ temos uma Weibull.

Para as aplicações financeiras não necessitamos calcular a qual MDA pertencem nossas distribuições contínuas, bastando saber que basicamente todas as distribuições de utilidade prática estão contidas em $MDA(H_{\xi})$ para algum valor de ξ (McNeil et al., 2015, p. 139).

2.1. Excessos acima de um limiar

O método conhecido como POT, para calcular a função de distribuição dos valores que excedem um determinado limiar de um conjunto de dados vem sendo empregado no mundo financeiro para ajustar as caudas das distribuições de retornos, ou perdas, dos ativos. Este método é preferido a teoria clássica de valores extremos (e.g. máximos em bloco), pois, desperdiça uma quantidade

menor de dados da série original. Qualquer valor que exceda o limiar pré-determinado é considerado na distribuição dos excessos. Esta distribuição dos valores da série que estão acima de um determinado limiar u é definida como:

Definição 4 (Distribuição dos excessos). Seja X uma variável aleatória com função de distribuição c.d.f F . A distribuição dos excessos sobre um limiar u tem a seguinte função de distribuição:

$$F_u(x) = P(X - u \leq x | X > u) = \frac{F(x + u) - F(u)}{1 - F(u)} \quad (2)$$

para $0 \leq x < x_F - u$, onde $x_F \leq \infty$ é o limite direito da distribuição F .

Ou seja, a função distribuição dos excessos sobre um limiar u é a probabilidade condicional que um valor X retirado dos dados subtraído de u (o excesso) seja menor que um dado quantil x , sabendo-se que X é maior que u . Uma importante distribuição que surge na modelagem dos excessos sobre um limiar é a distribuição de pareto generalizada – GPD, que segue.

Definição 5 (GPD). Distribuição de Pareto Generalizada é definida por sua função de distribuição:

$$G_{\xi, \beta(u)}(X) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\beta(u)}\right)^{-\frac{1}{\xi}}, & \xi \neq 0, \\ 1 - \exp\left(-\frac{x}{\beta(u)}\right), & \xi = 0, \end{cases} \quad (3)$$

onde $\beta > 0$, e $x \geq 0$ quando $\xi \geq 0$ ou $0 \leq x \leq -\beta/\xi$ quando $\xi < 0$.

Os parâmetros ξ e β são conhecidos respectivamente como parâmetros de forma e escala da distribuição. Na figura 1 abaixo, são mostradas três parametrizações para a função de distribuição acumulada (c.d.f) e para a densidade de probabilidades (p.d.f) de GPD com parâmetro ξ iguais a -0,5, 0 e 0,5 enquanto que o parâmetro de escala β é mantido constante e igual a 1. Perceba como para $\xi < 0$ a p.d.f tem um limite direito que é dado por $-\beta/\xi$ a partir do qual os valores de $g(x)$ são zero.

A distribuição de Pareto generalizada tem papel fundamental na teoria de

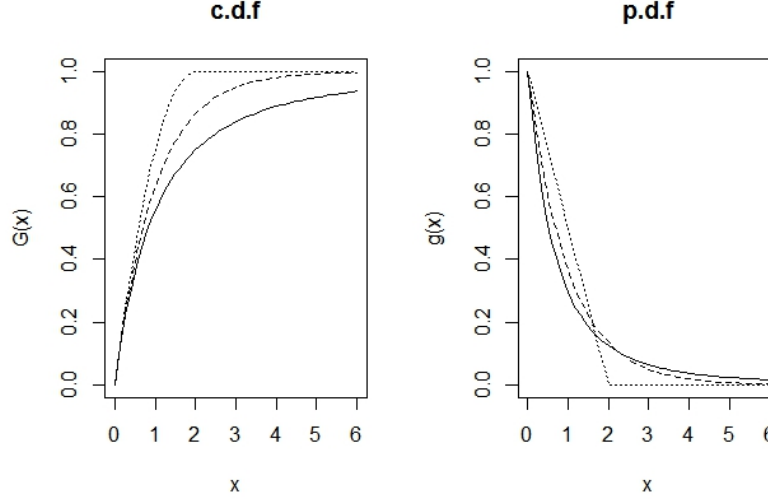


Figure 1: Três parametrizações para uma GPD. A linha sólida corresponde a $\xi = 0,5$, a linha pontilhada a $\xi = -0,5$ e a linha tracejada a $\xi = 0$.

valor extremo em função do teorema de Pickands-Balkema-de Haan (Pickands
150 (1975) e Balkema & de Haan (1974)) conforme abaixo:

Teorema 6 (Pickands-Balkema-de Haan). Pode ser encontrada uma função $\beta(u)$ tal que:

$$\lim_{u \rightarrow x_F} \sup_{0 \leq x < x_F - u} |F_u(x) - G_{\xi, \beta(u)}(x)| = 0$$

se e somente se $F \in MDA(H_\xi)$ para $\xi \in \mathbb{R}$.

O que este teorema nos diz é que para distribuições as quais os máximos
em bloco normalizados convergem para uma GEV (na forma da equação (1)),
então a distribuição dos excessos acima de um limiar destas mesmas distribuições
155 convergem para uma GPD, dado um valor de limiar u adequado.

Suposição 7. Seja F a distribuição de perdas com limite direito x_F , assuma
que para um valor limiar alto o suficiente u nós temos que $F_u(x) = G_{\xi, \beta}(x)$,
onde $F_u(x)$ denota a distribuição dos excessos de x em relação ao valor de limiar
 u , para $0 \leq x < x_F - u$, $\xi \in \mathbb{R}$ e $\beta > 0$.

160 Esta é uma suposição, uma vez que a distribuição dos excessos não segue exatamente uma GPD, mas apenas tende a esta distribuição dado um limiar u alto e uma amostra de dados grande o suficiente.

Dada a parametrização de uma GPD, é interessante sabermos o valor esperado desta distribuição, uma vez que esta medida de valor central nos fornece
165 importante informação sobre a quantidade de risco que estamos buscando medir, assim como a informação de que a própria distribuição foi ajustada aos dados de forma satisfatória, como será demonstrado adiante.

Definição 8 (função média dos excessos). A função média dos excessos de uma variável aleatória X com média finita é dada por:

$$e(u) = E(X - u | X > u) \quad (4)$$

170 Ou seja, a equação (4) representa o valor esperado da função de distribuição dos excessos dada pela Definição 4. Ela representa a média de F_u como uma função do limiar u . Esta função por vezes também é conhecida como função média de vida residual (*mean residual life function*), sendo encontrada esta denominação em alguns pacotes de *software* estatísticos.

175 Para uma variável distribuída na forma de uma GPD, o parâmetro de escala é uma função linear em u dado por $\beta(u) = \beta + \xi u$, (Embrechts et al., 1997, Teorema 3.4.13(e)). Utilizando-se deste fato e da equação (??) chegamos ao cálculo da função média dos excessos para uma GPD, dada por:

$$e(u) = \frac{\beta + \xi u}{1 - \xi} \quad (5)$$

onde $0 \leq u < \infty$ se $0 \leq \xi < 1$ e $0 \leq u \leq -\beta/\xi$ se $\xi < 0$. É possível observar que
180 de fato a função média dos excessos em uma GPD é linear em u . Esta é uma característica importante de uma GPD e que nos auxilia a escolher um valor adequado do limiar u de tal forma que a Suposição 7 faça sentido.

Lema 9. Sob a Suposição 7 segue que $F_v(x) = G_{\xi, \beta + \xi(v-u)}(x)$ para qualquer valor limiar $v \geq u$.

185 Logo, a distribuição dos excessos sobre limiares mais altos que u , também segue uma GPD com o mesmo parâmetro de forma ξ e parâmetro de escala que cresce linearmente com este limiar mais alto v . Se $\xi < 1$, a média desta nova GPD converge e é dada por:

$$e(v) = \frac{\beta + \xi(v - u)}{1 - \xi} = \frac{\xi v}{1 - \xi} + \frac{\beta - \xi u}{1 - \xi} \quad (6)$$

Esta é a função média dos excessos sobre limiares mais altos, e está definida
190 para $u \leq v < \infty$ se $0 \leq \xi < 1$ e, $u \leq v \leq u - \beta/\xi$ se $\xi < 0$.

Esta função é muito útil para calcularmos o ES_α (*expected shortfall*), considerando que VaR_α nada mais é que um quantil superior ao limiar u escolhido.

3. Modelando caudas e medidas de risco associadas

Através da modelagem da cauda da distribuição F de perdas por uma
195 GPD, como feito na seção anterior, é possível calcularmos as medidas de riscos VaR_α e ES_α para esta distribuição de perdas em função dos parâmetros da GPD estimada e também fazendo uso da distribuição empírica de F .

Sob a Suposição 7 nós temos que a cauda da distribuição F , $\bar{F}(x)$, para $x \geq u$ é dada por:

$$\begin{aligned} \bar{F}(x) &= P(X > u)P(X > x|X > u) \\ &= \bar{F}(u)P(X - u > x - u|X > u) \\ &= \bar{F}(u)\bar{F}_u(x - u) \\ &= \bar{F}(u) \left(1 + \xi \frac{x - u}{\beta}\right)^{-1/\xi} \end{aligned} \quad (7)$$

200 Da qual se soubéssemos exatamente a distribuição F teríamos um modelo analítico para as probabilidades de perdas na cauda da distribuição. Aqui x são os valores a serem observados das perdas, e portanto $x - u$ são as perdas em excesso ao limiar.

O que fizemos através da equação (7) foi efetivamente separar a distribuição F , ou melhor, sua cauda, em duas partes. A primeira parte, para valores menores
205 u , não foi modelado analiticamente e portanto utilizamos a distribuição empírica das perdas, aqui representada por sua cauda $\bar{F}(u)$, que nada mais é que o número observado de excessos de u sobre o número total de observações da amostra.

210 A segunda parte é justamente a modelagem através de uma GPD com parâmetros ξ e β dado o limiar u . Por esta modelagem paramétrica podemos conhecer as probabilidades de cauda para valores de x maiores que u .

O quantil α é a inversa da função distribuição e nos retorna o valor para o qual um percentual α de observações da amostra é menor ou igual. Assim sendo,
215 VaR_α nada mais é que um quantil alto para o qual determinamos que $\alpha\%$ das perdas devem ser menores ou iguais a este valor.

Como a equação (7) fornece a probabilidade de cauda, então esta é igual a $1 - \alpha$ para um valor de $\alpha \geq F(u)$. Fazendo $\bar{F}(x) = 1 - \alpha$ na equação (7) o valor de x representará VaR_α e nos basta manipular esta equação até isolarmos VaR_α
220 como função de $\bar{F}(u)$, α e dos parâmetros da GPD ξ e β . Que nos garante a equação abaixo:

$$VaR_\alpha = q_\alpha(F) = u + \frac{\beta}{\xi} \left[\left(\frac{1 - \alpha}{\bar{F}(u)} \right)^{-\xi} - 1 \right] \quad (8)$$

A medida ES_α pode ser entendida como a média das perdas que excedem o valor dado por VaR_α . Como o próprio VaR_α é um quantil acima do valor de limiar u , ES_α é dado pelo valor do VaR_α somado a função média dos excessos
225 dada pela equação (6) fazendo $v = VaR_\alpha$. Esta média é convergente para valores de $\xi < 1$ conforme já demonstrado. Ou seja, $ES_\alpha = VaR_\alpha + e(VaR_\alpha)$. A qual nos rende de forma mais geral:

$$ES_\alpha = \frac{VaR_\alpha}{1 - \xi} + \frac{\beta - \xi u}{1 - \xi} \quad (9)$$

Portanto, ambas medidas de risco VaR_α e ES_α , para distribuições de perdas

que tiveram suas caudas modeladas através de uma GPD da forma $G_{\xi,\beta(u)}$ com
 230 $\xi < 1$ e $\beta > 0$, podem ser calculadas respectivamente através das equações
 dadas em (8) e (9). As estimativas destas medidas de risco serão encontradas
 através das estimativas dos parâmetros da GPD, assim como do limiar utilizado
 e de uma medida empírica de $\bar{F}(u)$ que será o número de excessos verificados
 sobre o total de amostras. É claro que, ao adotarmos esta estimativa para $\bar{F}(u)$
 235 estamos implicitamente supondo que o número de amostras na série de perdas é
 significativa, assim como o número de excessos verificados. Daí a importância de
 se utilizar um valor u adequado, conforme explicitado na seção 2.1.

As estimativas de medidas de risco desenvolvidas nesta seção se qualificam
 como medidas incondicionais, no sentido que elas não dependem do estado atual
 240 das coisas, mas sim de todo o histórico de eventos de forma uniforme. Em outras
 palavras, VaR_α e ES_α derivados a partir das equações (8) e (9) são medidas
 históricas de risco associado ao ativo em análise, entretanto, não levam em
 consideração se nos eventos mais recentes a volatilidade das perdas pode ser
 diferente do valor histórico.

245 De fato, uma das características marcantes das perdas (ou retornos, como
 o leitor preferir) dos ativos financeiros é o chamado *clustering* de volatilidade,
 onde grandes volatilidades (retornos positivos ou negativos) têm tendência a
 ficarem próximas ao longo da linha temporal. Em geral estas aglomerações de
 volatilidades surgem a partir da autocorrelação destas, ou seja, a volatilidade
 250 em um período t é dependente das volatilidades verificadas em períodos ante-
 riores. Um modelo bastante encontrado na literatura que busca modelar estas
 dependências é o modelo ARCH e suas variantes como GARCH, propostos por
 Engle (1982) e Bollerslev (1986) respectivamente.

Assim, ao passo que as estimativas de risco desenvolvidas nesta seção são
 255 valiosas para prazos mais longos, ainda é necessário desenvolver um modelo
 que lide com o fato das autocorrelações de volatilidades e portanto, que nossa
 variável aleatória não é independente e igualmente distribuída ao longo do tempo.
 O modelo proposto por McNeil & Frey (2000) pode ser utilizado para encontrar
 as medidas de risco VaR_α e ES_α condicionais que desejamos, ainda dentro da

Incluir so-
bre Smith e
 $\bar{F}(u)$

260 metodologia de *peaks over threshold* - POT.

4. Modelo ARMA-GARCH-EVT

Ativos financeiros possuem características de autocorrelação, senão em seus retornos propriamente ditos, ao menos em suas volatilidades ou variações absolutas. Ou seja, dada uma grande variação no momento t é de se esperar
265 novamente uma grande variação, não necessariamente na mesma direção daquela anterior, para o momento $t+1$ e posteriores. Desta forma, medidas de risco incondicionais, conforme aquelas derivadas na Seção 3 podem ser adequadas somente para horizontes temporais mais longos, pois implicitamente tomam em consideração os fatos mais recentes com o mesmo valor de predição que fatos
270 mais longínquos.

Assim, nesta subseção trabalharemos com um modelo semelhante ao proposto por McNeil & Frey (2000) os quais fazem uma adequação dos retornos dos ativos a um modelo AR-GARCH e posteriormente tratam os erros desta modelagem como *iid* e portanto, a metodologia de POT pode ser aplicada.

275 Primeiramente iremos estabelecer um modelo ARMA-GARCH para as perdas do ativo subjacente. Se denotarmos L_t como sendo a perda observada no período t , μ_t e σ_t são respectivamente a média e o desvio padrão condicionais e mensuráveis através do conjunto de informações disponíveis em $t-1$ e seja Z_t inovações *iid* com média zero e desvio padrão unitário, então temos que:

$$L_t = \mu_t + \sigma_t Z_t \quad (10)$$

280 Seja $F_L(l)$ a distribuição marginal de L_t , então $F_{L_{t+1}|\mathcal{G}_t}(l)$ é a distribuição preditiva da perda para o próximo período, onde \mathcal{G}_t é o conjunto de informações disponíveis no período t , incluindo-o. Portanto, para o cálculo das medidas condicionais de risco estamos interessados em um quantil α na cauda de $F_{L_{t+1}|\mathcal{G}_t}(l)$. Este quantil α , que será o nosso VaR_α , é o ínfimo l tal que o valor da distribuição
285 preditiva seja maior ou igual a α . Ao passo que o valor condicional do *ES* será o

valor esperado das perdas previstas que sejam maiores que VaR para o mesmo intervalo de confiança. Ou seja:

$$VaR_{\alpha}^t = \inf\{l \in \mathbb{R} : F_{L_{t+1}|\mathcal{G}_t}(l) \geq \alpha\}, \quad (11)$$

$$ES_{\alpha}^t = E[L_{t+1} | L_{t+1} > VaR_{\alpha}^t] \quad (12)$$

Considerando que nossa distribuição de perdas é dada pela equação (10) e sabendo das propriedades de variáveis aleatórias e do operador de expectância,
 290 as equações dadas acima subsumem a:

$$VaR_{\alpha}^t = \mu_{t+1} + \sigma_{t+1} z_{\alpha}, \quad (13)$$

$$ES_{\alpha}^t = \mu_{t+1} + \sigma_{t+1} E[Z | Z > z_{\alpha}] \quad (14)$$

onde z_{α} é o quantil α das inovações Z .

Agora nos falta escolher um processo que modele nossa série temporal dada em (10), ou seja, precisamos especificar o comportamento de μ_t e σ_t . Por suposição do modelo, especificamos que o comportamento destas variáveis é dependente de
 295 acontecimentos passados, contidos no conjunto de informações \mathcal{G}_{t-1} . Dentre os diversos modelos já propostos para estimar médias e volatilidades condicionais, está o simples porém efetivo modelo GARCH(1,1) para a volatilidade condicional e o modelo AR(1) para a média condicional. Uma extensão destes modelos básicos para outros mais complexos pode ser facilmente obtida e é vasta na
 300 literatura, como exemplo modelos GARCH-M, *Threshold* GARCH, EGARCH, etc. e modelos do tipo ARMA, ARIMA, entre outros encontrados em Tsay (2010).

Neste trabalho, visando aplicar a teoria do valor extremo para o cálculo das medidas condicionais de risco, não faremos maiores assunções acerca da distribuição das inovações, como por exemplo assumir uma determinada dis-
 305 tribuição (e.g. Normal ou t de Student), mas está implícito que esta pertence ao MDA de uma GEV e portanto a distribuição de seus excessos sobre um limiar

Implementar
com modelo
EGARCH

segue aproximadamente uma GPD.

Dadas estas considerações, o modelo adotado segue um formato em dois estágios para ser implementado, como segue.

- 310 • Ajustar um modelo AR-GARCH para os dados de perdas, sem fazer suposições sobre a distribuição de Z , porém, utilizando um método de pseudo máxima verossimilhança (PML). Deste modelo tiramos as estimativas de μ_{t+1} e σ_{t+1} e portanto, calculamos as inovações implícitas resultantes através da equação (10).
- 315 • Consideramos estas inovações calculadas como sendo as realizações da variável aleatória Z , a qual por suposição pode ter sua cauda ajustada a uma GPD utilizando o método descrito na seção 3. Encontraremos por fim o valor de z_α e $E[Z|Z > z_\alpha]$, com os quais finalizamos os cálculos de nossas medidas condicionais de risco dadas em (13) e (14).

320 Nosso modelo completo para as medidas de risco VaR_α e ES_α condicionais dada a distribuição de perdas L_t de um ativo será, portanto:

$$L_t = \mu_t + \epsilon_t \quad (15)$$

$$\mu_t = \mu + \phi_1 \mu_{t-1} \quad (16)$$

$$\epsilon_t = \sigma_t Z_t \quad (17)$$

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (18)$$

$$Z_t \sim \mathcal{D}(0, 1) \text{ e } \mathcal{D} \in MDA(H_\xi) \quad (19)$$

com as equações (13) e (14) nos fornecendo os valores das medidas VaR_α e ES_α respectivamente, quando utilizadas em conjunto com aquelas dadas por (8) e (9).

- 325 O termo de erro ϵ_t será uma série estritamente estacionária se $\beta_1 + \alpha_1 < 1$, o que garante que a variância de nosso modelo (σ_t^2) será finita, apesar de variável no tempo.

Este modelo é ajustado utilizando-se o método de pseudo máxima verossimilhança, o que significa que, apesar de não adotarmos a distribuição normal para reger as inovações Z_t , a utilizamos apenas para estimar os parâmetros do modelo. É demonstrado na literatura, (Gouriéroux, 1997, Capítulo 4), que o método PML fornece estimadores consistentes e assintoticamente normais, assim podemos utilizar tal método para obter nossas estimativas de parâmetros e erros padrão, mesmo não aceitando a condição de normalidade das inovações.

Exemplo 10 (Risco condicional da Apple). Como exemplo tomaremos os dados de perdas diárias da Apple Inc. (*ticker* AAPL) compreendidas no período entre 02/01/2009 a 16/09/2016, 1.941 observações.

Seguindo o procedimento em 2 estágios, primeiramente iremos ajustar a série de perdas diárias da Apple ao modelo AR(1)-GARCH(1,1) descrito acima, considerando a distribuição condicional como uma normal.

A figura 2 abaixo nos mostra alguns resultados desta regressão. Seguindo os gráficos da esquerda para a direita, acima e depois abaixo, temos no primeiro gráfico as perdas observadas envoltas em uma banda composta por $\pm 2\sigma$. No segundo gráfico temos a autocorrelação das próprias observações, ou seja, das perdas. Nota-se que estatisticamente não se pode afirmar que estas sejam correlacionadas e portanto, seria de se esperar que o valor de ϕ_1 em nosso modelo não seja significativo. No terceiro gráfico temos as autocorrelações dos quadrados das perdas, e neste caso sim, de acordo com os fatos estilizados de ativos financeiros, estes são altamente correlacionados.

Enquanto os três primeiros gráficos nos trazem informações da série de perdas original, a sequência na parte de baixo da figura 2 mostra informações acerca do modelo AR-GARCH ajustado. O primeiro deles é um gráfico do tipo QQ (quantil-quantil) muito utilizado para verificar a distribuição de variáveis aleatórias. Nele temos uma distribuição de referência, neste caso a Normal e, se a variável de interesse segue esta distribuição os pontos no gráfico QQ devem ficar todos sobre a linha contínua. Podemos perceber claramente que ao nos aproximarmos das caudas, os resíduos padronizados, z_t , se desviam

consideravelmente da normalidade, apresentando caudas muito mais longas que uma distribuição normal. Para o nosso modelo isso era esperado, uma vez que

360 sabemos que as inovações em processos estocásticos de ativos financeiros não seguem uma normal. São os excessos acima de um limiar destes resíduos que, no segundo passo de nossa modelagem, ajustaremos a uma distribuição do tipo GPD.

Ainda assim é possível observar que a filtragem dos dados promovida pelo

365 ajuste AR-GARCH nos serviu para descorrelacionar serialmente os resíduos padronizados, conforme demonstram os dois gráficos seguintes, de forma a ser possível utilizar a técnica de valor extremo a estes resíduos.

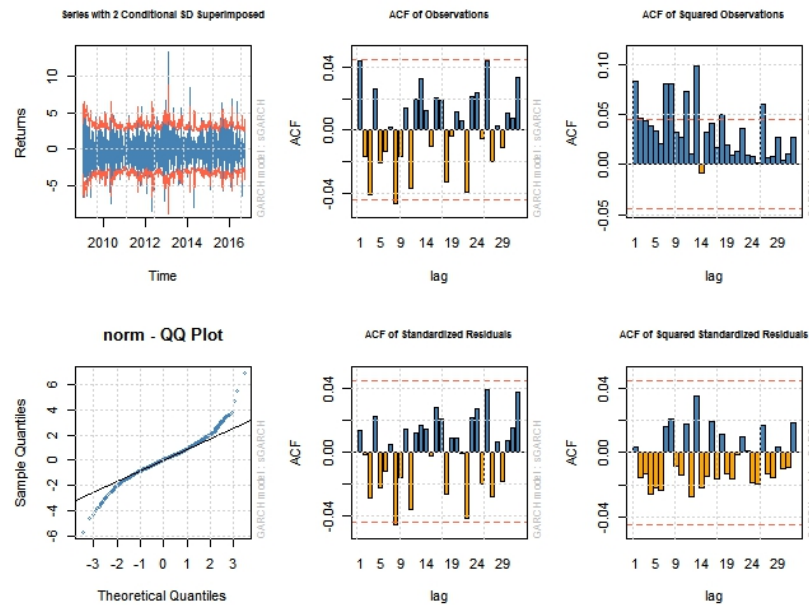


Figure 2: Ajuste de APPL ao modelo AR(1)-GARCH(1,1). Autocorrelação nos quadrados das perdas e distribuição dos resíduos padronizados com caudas longas.

A tabela 1 abaixo traz as informações sobre o modelo ajustado. Os valores estimados dos parâmetros, assim como seus erros padrões, estatística t e valores

370 p . Percebe-se que de fato o modelo se ajustou aos dados, considerando que todos os parâmetros estimados são significativos com ao menos 95% de confiança com

exceção do parâmetro ϕ_1 o qual, pelo gráfico de autocorrelação, já esperávamos fosse pouco significativo.

Ajustado este primeiro modelo, que é o equivalente a filtrar os dados originais de perda com um modelo AR-GARCH de modo que os resíduos padronizados
 375 agora sim sejam as realizações de uma variável aleatória *iid*, então podemos aplicar a técnica de modelagem da cauda destes resíduos conforme a teoria de valor extremo. Cabe salientar que é fundamental fazer esta filtragem nos dados, uma vez que a POT supõe dados independentes e identicamente distribuídos, caso
 380 contrário não teríamos a convergência dada no teorema 6 de Pickands-Balkema-de Haan.

Table 1: Parâmetros estimados para o modelo AR-GARCH de AAPL

	Estimativa	Erro Padrão	Valor t	$\Pr(> t)$
μ	-0.1683	0.0381	-4.4138	0.0000
ϕ_1	0.0400	0.0246	1.6282	0.1035
ω	0.2107	0.0514	4.1011	0.0000
α_1	0.0876	0.0187	4.6804	0.0000
β_1	0.8445	0.0301	28.0768	0.0000

Portanto, aplicaremos agora a modelagem POT para os resíduos padronizados, os quais sabemos não serem normalmente distribuídos e que possuem considerável curtose em excesso (vide gráfico QQ da figura 2) e encontraremos os valores
 385 de z_α e $E[Z|Z > z_\alpha]$ necessários para computar os valores de $VarR_\alpha$ e ES_α das equações (13) e (14).

Os valores de z_α e $E[Z|Z > z_\alpha]$ foram calculados para $\alpha = 0,99$ e são apresentados na tabela 2 abaixo juntamente com os limites inferior e superior de seu intervalo de confiança a 95%. Note que $E[Z|Z > z_\alpha]$ na tabela em questão foi chamado de $s_{.99}$.

Table 2: Valores de $z_{.99}$ e $s_{.99}$ encontrados e seus respectivos intervalos de confiança a 95%

	Inf	Estimativa	Sup
$z_{.99}$	2.40	2.61	2.89
$s_{.99}$	3.11	3.55	4.90

390

A figura 3 traz algumas informações acerca do ajuste feito a uma GPD dos

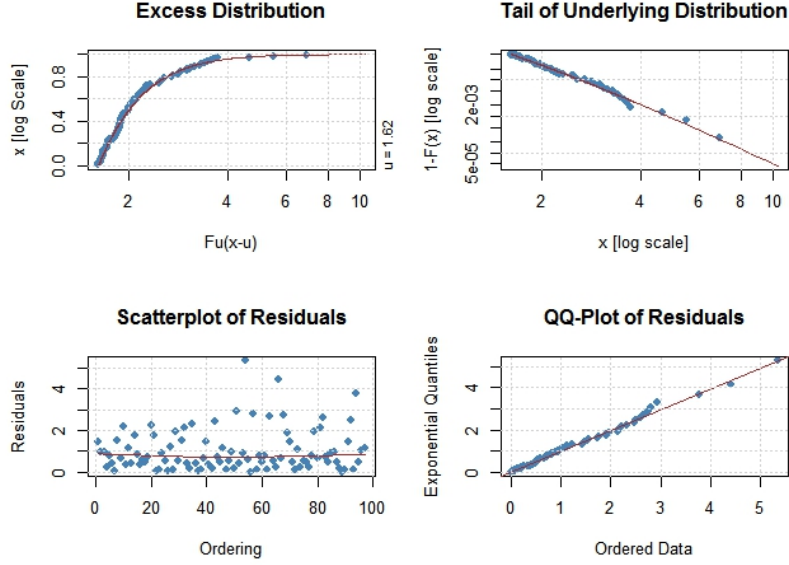


Figure 3: Ajuste dos resíduos padronizados a um modelo POT. Distribuição dos excessos acima do limiar $u = 1,62$.

resíduos padronizados obtidos. Foi escolhido como limiar u o valor de 1,62 que é o quantil de z_t a 0,95. Esse valor foi ratificado através da análise qualitativa do gráfico de vida média residual (MRL plot) conforme encontrado na definição

395 8. Com a escolha deste limiar, obtivemos 97 observações de perdas em excesso a este valor. Quantidade razoável e que nos permite calcular os parâmetros da distribuição GPD.

Pode-se verificar através da figura 3 que os ajustes feitos foram bastante satisfatórios, principalmente com a GPD capturando muito bem a curtose em

400 excesso na cauda da distribuição dos resíduos padronizados (*tail of underlying distribution*) e os quantis dos próprios resíduos deste ajuste ficarem juntos dos valores teóricos esperados (*QQ plot of residuals*).

Possuímos todos os parâmetros de nosso modelo híbrido AR-GARCH-POT para então calcularmos os valores das medidas de risco condicional dadas pelas

405 equações (13) e (14). Com o modelo AR-GARCH executado no período t obtemos as estimativas de μ_{t+1} e de σ_{t+1} e através do modelo POT obtemos os

valores de z_α e $E[Z|Z > z_\alpha]$, no caso do nosso exemplo escolhemos $\alpha = 0,99$.

Toda esta metodologia pode ser aplicada período a período, mantendo uma memória fixa das últimas n observações de perdas e que resultará em valores calculados para VaR_α^{t+1} e ES_α^{t+1} , os quais poderão ser comparados com a
410 realização futura de perda.

5. Conclusão

References

- Acerbi, C., & Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, 26, 1487–1503.
415
- Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1997). Thinking Coherently. *RISK*, 10, 68–71.
- Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, 9, 203–228. doi:10.1111/1467-9965.00068.
- 420 Balkema, A. A., & de Haan, L. (1974). Residual Life Time at Great Age. *The Annals of Probability*, 2, 792–804. doi:10.1214/aop/1176996548.
- on Banking Supervision, B. C. (2011). Basel III: A global regulatory framework for more resilient banks and banking systems. URL: <http://www.bis.org/publ/bcbs189.pdf>.
- 425 on Banking Supervision, B. C. (2013a). Basel III: The liquidity coverage ratio and liquidity risk monitoring tools. URL: <http://www.bis.org/publ/bcbs238.pdf>.
- on Banking Supervision, B. C. (2013b). Fundamental review of the trading book: A revised market risk framework. URL: <http://www.bis.org/publ/bcbs265.pdf>.
430
- on Banking Supervision, B. C. (2014). Basel III: The net stable funding ratio. URL: <http://www.bis.org/bcbs/publ/d295.pdf>.

- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307–327. doi:10.1016/0304-4076(86)90063-1.
- 435 Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer.
- Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50, 987. doi:10.2307/1912773.
- 440 Fisher, R. A., & Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, 24, 180–190.
- Gnedenko, B. V. (1941). Limit theorems for the maximal term of a variational series. *Comptes Rendus (Doklady) de L'Académie des Sciences de l'URSS*, 32, 7–9.
- 445 Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics*, 44, 423–453.
- Gouriéroux, C. (1997). *ARCH Models and Financial Applications*. Springer Series in Statistics. New York, NY: Springer New York.
- 450 URL: <http://link.springer.com/10.1007/978-1-4612-1860-9>. doi:10.1007/978-1-4612-1860-9.
- McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7, 271–300. doi:10.1016/S0927-5398(00)00012-8.
- 455 McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: concepts, techniques and tool*. (Revised ed ed.). Princeton University Press.
- Pickands, J. (1975). Statistical Inference Using Extreme Order Statistics. *Annals of Statistics*, 3, 119–131. doi:10.1214/aos/1176343003.

- RiskMetrics (1995). *Technical Document - 3rd Edition*. Technical Report J.P
460 Morgan Guaranty Trust Company.
- Tsay, R. S. (2010). *Analysis of financial time series*. (3rd ed.). Wiley.