CÁLCULO DE VALUE AT RISK PARA UMA CARTEIRA DE AÇÕES UTILIZANDO TOOLBOX DO MATLAB

Alexander Bento Melo, Cícero Barroso de Oliveira, Alexandre Coutinho Mateus
Universidade Federal de Uberlândia, Faculdade de Engenharia Elétrica/Engenharia Eletrônica e de Telecomunicações,
alexbentomelo@ufu.br , cicero_alf@hotmail.com, alexandre@ig.com.br

Resumo –Este artigo visa prever o valor de uma carteira de ações fictícia para um período de tempo futuro pré-determinado. Para isso utilizou-se a simulação de Monte Carlo para a previsão assim como a Toolbox do Matlab. Este tipo de análise leva em consideração conhecimentos de estatística e de sistemas de controle para prever perdas e ganhos.

***Palavras-Chave* – Carteira de ações, simulação Monte Carlo, previsão.**

VALUE AT RISK STOCK PORTIFOLIO ANALYSIS USING MATLAB TOOLBOX

***Abstract* –This article shows how to predict the value of a stock portfolio for a predetermined future time. For this we used the Monte Carlo as the Matlab Toolbox. This type of analysis takes into account knowledge of statistics and control systems to predict losses and gains.**

Keywords –Stock Portfolio, Monte Carlo simulation, predict.

I. INTRODUÇÃO

Realizar a análise e previsão do valor futuro de uma carteira de ações assim como efetuar uma boa análise de risco, tem se mostrado uma área interessante de pesquisa dentro da área de controle, dado que técnicas de estatística e também aquela utilizada para predição pode ser empregada para estas análises também. Além disso este tema possui uma grande relevância no atual cenário econômico mundial.

Essa relevância se deve ao fato de que a cada dia que passa e com a piora do cenário econômico e político, os investidores estão mais exigentes no âmbito de minimizar ou extinguir perdas em um dado investimento por quaisquer fatores.

Para atender essas exigências dos investidores, empresas administradoras e seus gerenciadores de riscos buscam softwares e algoritmos capazes de prever o valor de uma dada ação ou carteira de ações com a maior confiabilidade e precisão possíveis.

Para realizar essas previsões, existem variados modelos de simulação. Os principais modelos se dividem em dois tipos: os modelos paramétricos, que se baseiam nos parâmetros de variância e covariância dos valores de ação ou carteira de ações, e os modelos não paramétricos, que pode ser por exemplo, simulação Monte Carlo ou percentil histórico.

O modelo paramétrico parte da premissa que a série de valores segue uma distribuição estatística, com distribuição normal, e partindo desse ponto estima as características utilizadas na análise da carteira[2].

O modelo de Monte Carlo utiliza os parâmetros de valor atual ou anterior e também parâmetros característicos dos valores da ação ou carteira em um período de tempo anterior ao da análise. Já o modelo percentil histórico se baseia nos valores da ação ou carteira de ações num período de tempo para tentar prever estatisticamente o valor dessa ação no futuro.

Existem variados tipos de riscos que influenciam os valores de uma carteira de ações. A proposta deste artigo é prever o valor de uma carteira de ações levando-se em consideração somente o "*Value at Risk* (VAR)", que é o risco que considera somente as flutuações do valor no mercado das ações sem levar em consideração acontecimentos externos como por exemplo problemas relacionados a política.

Para este trabalho realizou-se a comparação dos resultados para uma mesma carteira de ações, mesmo período de tempo futuro, o desempenho na função de previsão do valor da carteira de ações entre uma simulação paramétrica e uma semi paramétrica.

II. TEORIA DE MONTE CARLO

Métodos de inferência são usados para tirar conclusões sobre a população usando informações obtidas a partir de uma amostra. Estimativas pontuais e intervalares para os parâmetros; Testes de hipóteses e Modelagem.

Para obter resultados confiáveis, é necessário conhecer a distribuição da estatística (média, mediana, variância, assimetria, etc.) em estudo. O Método de Monte Carlo é uma saída para fazer inferências quando não se conhece a distribuição do parâmetro de interesse ou quando as suposições de um modelo são violadas.

O método originou-se por causa do uso de aleatoriedade e da natureza repetitiva das atividades realizadas em cassinos de Monte Carlo. A roleta era um gerador de números aleatórios. Atualmente termo Monte Carlo é mais geral. É uma técnica baseada no uso de números aleatórios e estatísticas para resolver problemas. Alguns dos usos de Métodos Monte Carlo:

XIV CEEL - ISSN 2178-8308
03 a 07 de Outubro de 2016
Universidade Federal de Uberlândia - UFU
Uberlândia - Minas Gerais - Brasil

- Realizar inferências quando a distribuição da estatística de teste não é conhecida.
- Estimar o desempenho de métodos de inferência quando as suposições paramétricas são violadas.
- Avaliar desempenho de métodos de inferências (poder do teste)

O método basicamente constitui-se em estimar a distribuição de uma estatística extraindo amostras aleatórias de uma população e observar o comportamento da estatística sobre as amostras. Neste caso, o método Monte Carlo é uma abordagem paramétrica porque a amostra é extraída de uma população com distribuição conhecida. A aplicação do método inicia com definição da pseudo-população que é assumida para representar a população real.

III. MÉTODO DE MONTE CARLO

O Método de Monte Carlo é uma técnica amplamente utilizada para análises probabilísticas em sistemas de engenharia [2].

Essa técnica de simulação utiliza recursos computacionais para modelar numericamente sistemas e obter suas análises estatísticas por meio de variáveis de entrada ou saída.

O processo utilizado para prever os valores de cada ação para um período de tempo futuro foi o de Wiener:

$$\frac{\Delta p_t}{p_t} = m.\Delta t + \sigma.\in.\sqrt{\Delta t}$$

Onde p_t é o preço da ação, $\in$ o valor de um número aleatório de uma distribuição normal com média nula e variância unitária, Δt o período de tempo futuro para o qual se deseja fazer a previsão, m o valor médio das ações e σ o desvio padrão.

No entanto, em simulações é mais usual utilizar ln (p_t), e do teorema de Itô, têm-se[3]:

$$d \ln p = (m - \sigma^2)dt + \sigma dz, com\ z = \in.\sqrt{\Delta t}$$

Logo:

$$\ln p_{t+\Delta t} - \ln p_t = \left(m - \frac{\sigma^2}{2}\right)\Delta t + \sigma.\in.\sqrt{\Delta t}$$

Então:

$$p_{t+\Delta t} = p_t.e^{\left[\left(m-\frac{\sigma^2}{2}\right)\Delta t + \sigma.\in.\sqrt{\Delta t}\right]}$$

Assim, os próprios valores da simulação são utilizados para a previsão de um período de tempo futuro maior que um dia.

IV. DESENVOLVIMENTO

A. Carteira de ações

A tarefa inicial para a execução do projeto foi escolher uma carteira de ações para se fazer todo o estudo e simulações.

A carteira de ações escolhidas foi obtida por meio do site de simulação de investimentos “folhainvest.folha.uol.com.br” da Folha de São Paulo, onde após um cadastro pode-se simular a compra e venda de algumas ações.

Incialmente fez-se o cadastro no site e após isso deu-se início à montagem da carteira de ações. Essa carteira foi escolhida observando-se o histórico das ações de forma a garantir que as mesmas não tivessem variações abruptas ao longo do período.

As ações escolhidas para compor a carteira foram: ITAU S.A, OI, EMBRAER e RAIA DROGASIL.

Após a escolha da carteira de ações, iniciou-se o estudo dos valores de cada ação para os meses de janeiro, fevereiro e março e todos esses valores foram inseridos como parâmetro no software para posterior verificação de comportamento.

Utilizando o Matlab obtivemos inicialmente dados importantes sobre as ações escolhidas, como Média, Variância e Desvio Padrão. Estes valores estão sumarizados na Tabela I.

Tabela I: Sumário de Avaliação Preliminar das Ações Utilizadas

	ITAU S.A			
	Quantidade de Amostras	Média	Desvio Padrão	Variância
Fevereiro	29	6.7003	0.1214	0.0147
Março	31	8.2219	0.3780	0.1429
Abril	30	8.3460	0.2909	0.0846
	OI			
	Quantidade de Amostras	Média	Desvio Padrão	Variância
Fevereiro	29	1.6628	0.1933	0.0374
Março	31	1.2426	0.0698	0.0049
Abril	30	1.0190	0.0718	0.0052
	EMBRAER			
	Quantidade de Amostras	Média	Desvio Padrão	Variância
Fevereiro	29	28.8417	0.9893	0.9787
Março	31	23.6687	1.6513	2.7269
Abril	30	22.0790	0.6769	0.4582
	RAIA DROGASIL			
	Quantidade de Amostras	Média	Desvio Padrão	Variância
Fevereiro	29	44.2928	2.4052	5.7849
Março	31	49.2039	2.4386	5.9466
Abril	30	53.7037	1.4675	2.1535

B. Modelagem no Matlab

A Figura 1 ilustra os dados utilizados, especificamente, os gráficos mostram os movimentos dos preços relativos de cada ação. O nível inicial de cada índice foi normalizado para a unidade para facilitar a comparação de desempenho relativo em relação ao máximo no período de cada ação que será utilizado posteriormente.

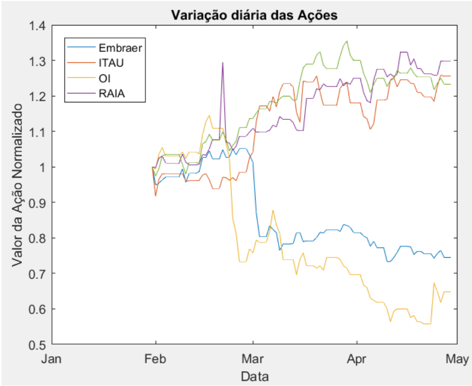

Fig. 1. Variação diária das ações que compõe a carteira.

Para modelar, caracterizarmos individualmente a distribuição de retorno de cada ação. Embora a distribuição de cada série possa ser caracterizada parametricamente, isto será útil para ajustar um modelo semi-paramétrico usando uma distribuição com caudas como a Pareto generalizadas. Isto usa teoria dos valores extremos para melhor caracterizar o comportamento em cada cauda.

No Matlab, utilizamos os "*tollbox Statistics*" e "*Machine Learning Toolbox(TM)*" que suportam dois tipos de distribuições de probabilidade relacionadas com a *"Extreme Value Theory (EVT)"*, uma ferramenta estatística para modelar o comportamento de cauda gorda de dados financeiros, tais como os retornos de ativos, abaixo seguem as descrições destes dois tipos:

- "*Generalized Extrem Value (GEV)*": usa uma técnica de modelagem conhecida como "*block maxima*" ou "*minima method*". Esta abordagem, divide o conjunto de dados históricos em um conjunto de sub-intervalos, ou blocos, e a maior ou menor amostra em cada bloco é guardada e posteriormente aproximada por uma GEV.
- "*Generalized Pareto (GP)*": utiliza uma técnica de modelagem conhecida como o "*distribution of exceedances*" ou o método conhecido como "*peaks over threshold*". Esta abordagem classifica um conjunto de dados históricos, e encaixa o montante pelo qual essas observações excedem um limite especificado para uma distribuição GP.

Na primeira análise realizada neste trabalho foi utilizada a distribuição de Pareto, que é amplamente usada em aplicações de gestão de risco que é um dos objetivos deste trabalho. Suponha que desejamos criar uma descrição estatística completa da distribuição de probabilidade de retornos diários de ativos de qualquer uma das ações estudadas. Assumindo que esta descrição é fornecida por uma distribuição semi-paramétrica, onde o comportamento assintótico em cada cauda é caracterizado por uma distribuição de Pareto.

A "*cumulative distribution function* (CDF)" e a inversa CDF irão capturar a volatilidade dos retornos simulados como parte do termo de difusão da *"Stochastic Differential Equation (SDE)"*. O retorno médio de cada índice é regido pela taxa sem risco e incorporado no termo derivativo da SDE. A Figura 2 abaixo mostra os gráficos com os retornos centralizados para cada ação.

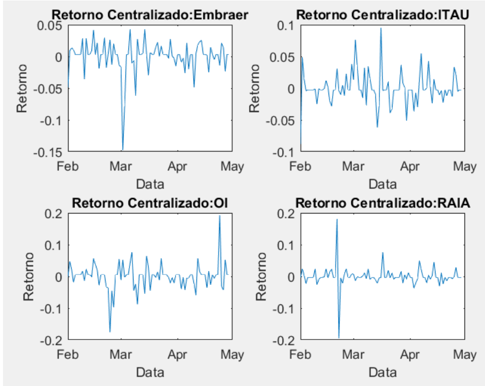

Fig. 2. Retornos centralizados para cada ação.

Após isso usando os retornos centrados, estimaremos o CDF de cada índice com uma através de "*Gaussian Kernel"*. Isso suaviza as estimativas do CDF, eliminando o padrão escadaria das amostras obtidas. Embora estimativas não paramétricas são bem adequadas para o interior da distribuição, onde a maior parte dos dados é encontrada, eles tendem a ter uma baixa performance quando aplicado às caudas superiores e inferiores. Para melhor estimar as caudas da distribuição, aplicaremos a EVT aos retornos que se enquadram em cada cauda.

Especificamente, encontraremos os limiares superior e inferior de tal modo que 10% dos rendimentos está reservado para cada cauda. Em seguida, ajustaremos o valor pelo qual os retornos extremos em cada cauda caem além do limite associado a uma distribuição de Pareto.

O resultado permite a interpolação dentro do interior da CDF e a extrapolação (avaliação de função) em cada cauda. A extrapolação permite estimar os quantis fora do registro histórico, que é de valor inestimável para as aplicações de gestão de risco.

Agora que as três regiões distintas da distribuição por partes foram estimadas, o gráfico resultante para a ação da Embraer é apresentado na Figura 3.

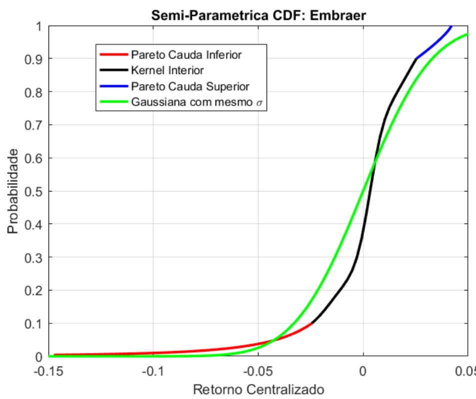

Fig. 3. CDF Semi-Paramétrica para a ação da Embraer.

As regiões de cauda inferior e superior, exibidas em vermelho e azul, respectivamente, são adequadas para extrapolação, enquanto o interior, em preto, é adequado para a interpolação.

O Toolbox de Estatística e "*Machine Learning"* do Matlab incluem algumas funcionalidades que nos permitem calibrar e simular as t cópulas e as cópulas Gaussianas.

Usando o índice de retornos diários, estimaremos os parâmetros das t cópulas e da cópula Gaussiana através da função | **copulafit** |, do Matlab.
Apesar da calibração da matriz de correlação linear de uma cópula Gaussian ser simples, a calibração de uma t cópula não é. Por esta razão, utilizaremos o Toolbox de Estatística e "*Machine Learning"* para realizar esta tarefa.
Após estimar os parâmetros da copula, simularemos conjuntamente as variáveis dependentes utilizando a função | **copularnd** |.
Então, através da extrapolação das caudas de Pareto e interpolando o interior alisado, transformaremos as variáveis uniformes obtidas em retornos diários centralizados através da CDF inversa de cada índice. Estes retornos centrados simulados são consistentes com os obtidos a partir do conjunto de dados históricos. Os retornos são assumidos por serem independente no tempo, mas em qualquer ponto no tempo possuem a dependência e de correlação induzida pela cópula.
A Figura 4 mostra o resultado da análise para as ações da Embraer e do Itau.

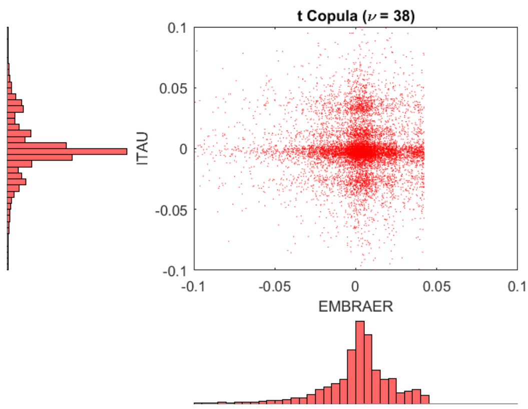

Fig. 4. Retorno Centralizado utilizando o método da t cópula.

A Figura 5 mostra a simulação dos retornos centralizados usando a Gaussiana.

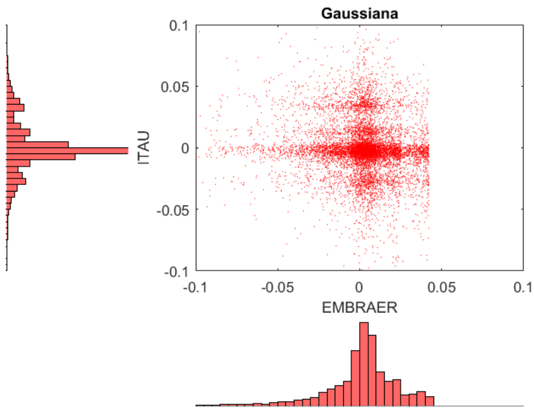

Fig. 5. Retorno Centralizado utilizando o método da Cópula Gaussiana.

Examinando as Figuras 4 e 5, há uma forte semelhança entre os histogramas em miniatura nos eixos correspondentes de cada figura. Esta semelhança não é mera coincidência.

Ambas as cópulas simulam variáveis aleatórias uniformes, que são então transformadas para retornos diários centrados pela CDF inversa da distribuição de cada índice. Portanto, os retornos simulados de um determinado índice são identicamente distribuídos independentemente da cópula.
No entanto, o gráfico de dispersão de cada figura indica a estrutura de dependência associada com a cópula dada, e em contraste com o mostrado nos histogramas, os gráficos de dispersão são distintos.
Mais uma vez, a cópula define uma estrutura de dependência, independentemente das suas margens e, portanto, oferece muitos recursos não limitadas à calibração sozinho.
Para referência, simularemos os retornos centrados utilizando a distribuição Gaussiana, que fundamenta o modelo de movimento browniano tradicional, a Figura 6 mostra este resultado.

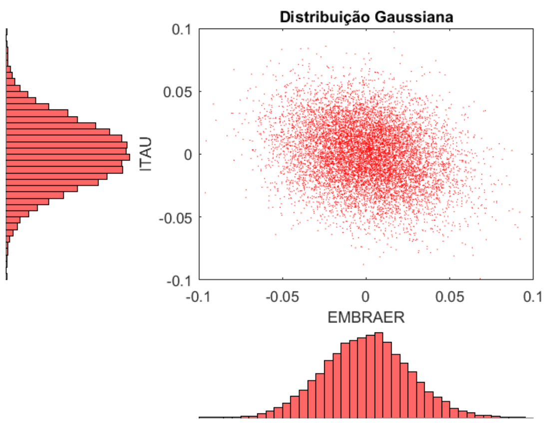

Fig. 6. Distribuição Gaussiana.

Tendo simulado os retornos de cada ação e a partir daí formado um portfólio global, a partir do qual obtivemos o máximo ganho e a máxima perda, como também o VaR para vários níveis de confidência, observando-se que calculamos para um horizonte de risco de 1 mês. A Figura 7 mostra a CDF para este retorno.

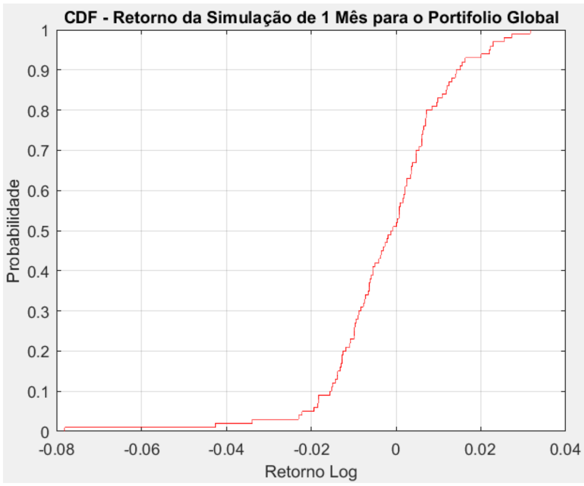

Fig. 7. CDF – Retorno da Simulação de Risco para 1 mês.

Os valores de VAR encontrados são apresentados abaixo:

- Máxima Perda Simulada: 7.8044%

- Máximo Ganho Simulado: 3.1608%
- VaR Simulado 90%: -1.5389%
- VaR Simulado 95%: -2.0718%
- VaR Simulado 99%: -6.0308%

V. CONCLUSÕES

A utilização do Toolbox do Matlab para o desenvolvimento deste artigo foi de suma importância dado a facilidade de se manipular as operações estatísticas e a precisão dos resultados obtidos.

A simulação foi executada mais de 100 vezes para um espaço de tempo de 10, 15, 20 e 30 dias. Isso significa que o VaR foi calculado entorno de 300 vezes para cada horizonte de tempo e os resultados obtidos mostraram uma fidelidade muito interessante em relação aos resultados reais.

Como próximos passos para este trabalho, serão analisadas outras carteiras de ação, para que se possa comparar os resultados com os obtidos até o momento, assim como serão utilizados métodos não paramétricos para se comparar também.

REFERÊNCIAS

[1] Teramossi, L. M; Perez, V. M. Análise de Risco de uma carteira de ações no Mercado Financeiro, *São José dos Campos*, 113 f, Trabalho de Conclusão de Curso (Graduação), Instituto Tecnológico de Aeronáutica, 2008.

[2] Cruse, T. A. *Reliability-Based Mechanical Design*, Marcel Dekker, Nova York, 1997.

[3] Dvorsak, P. Ferramenta para Análise de Fatores de Risco de Crédito para swaps, *São José dos Campos*, Trabalho de Conclusão de Curso (Graduação), Instituto Tecnológico de Aeronáutica, 2003.

[4] Bouye, E., V. Durrleman, A. Nikeghbali, G. Riboulet, and Roncalli, T. "Copulas for Finance: A Reading Guide and Some Applications." Groupe de Rech. Oper., Credit Lyonnais, Paris,2000.

[5] Embrechts, P., A. McNeil, and D. Straumann. "Correlation and Dependence in Risk Management: Properties and Pitfalls". Risk Management:_Value At Risk and Beyond. Cambridge: Cambridge University Press, 1999, pp. 176; 223.

[6] Glosten, L. R., R. Jagannathan, and D. E. Runkle. "On the Relation between Expected Value and the Volatility of the Nominal Excess Return on Stocks."The Journal of Finance. Vol. 48, 1993, pp. 1779; 1801.

[7] McNeil, A. and R. Frey. "Estimation of Tail Related Risk Measure for Heteroscedastic Financial Time Series: An Extreme Value Approach."Journal of Empirical Finance. Vol. 7, 2000, pp. 271; 300.

[8] Nystrom, K. and J. Skoglund. "Univariate Extreme Value Theory, GARCH and Measures of Risk Technical report, Swedbank, Group Financial Risk Contr ol, S-105 34 Stockholm, Sweden 2002. Available at: <http://gloriamundi.com/UploadFile/20102/knjsug.pdf>.

[9] Nystrom, K. and J. Skoglund. "A Framework for Scenario-Based Risk Management." Available at SSRN: http://ssrn.com/abstract=2246672.

[10] Roncalli, T., A. Durrleman, and A. Nikeghbali. "Which Copula Is the Right One?" Groupe de Rech. Oper., Credit Lyonnais, Paris,2000.

[11] Mashal, R. and A. Zeevi. "Beyond Correlation: Extreme Co-movements between Financial Assets." Columbia University, New York, 2002.