Medidas De Desempenho De Fundos Considerando Risco De Estimação

William Eid Jr.
FGV-EAESP
weid@fgvsp.br

Ricardo Ratner Rochman
FGV-EAESP
rrochman@fgvsp.br

Marcelo Taddeo
IME-USP
marcelo.taddeo@gmail.com

Medidas de Desempenho de Fundos Considerando Risco de Estimação

Resumo: Neste artigo são apresentadas duas medidas de desempenho de fundos mútuos de investimento baseadas nos índices de Sharpe Generalizado e de Sortino, que são ajustadas para o risco de estimação através de intervalos de confiança gerados por meio de procedimentos de *bootstrap*. O uso das medidas propostas é útil para fundos de investimento que empregam gestão ativa, ou seja, que tenham como objetivo superar um determinado índice de referência. Os índices propostos neste artigo serão comparados com os índices de Sharpe Generalizado e de Sortino originais e com o duplo índice de Sharpe de Vinod e Morey (1999), com uma amostra de 100 fundos do mercado brasileiro no período de 2004, onde serão analisados os resultados de ranqueamento de diferentes fundos conforme seu desempenho. Os resultados dos ranqueamentos e da correlação de Spearman mostram que há grandes mudanças de ranqueamento quando se utiliza as medidas propostas em comparação com as medidas já existentes, principalmente quando se emprega o *boostrap* studentizado.

1. Introdução

Embora os índices de Sharpe (Sharpe, 1966), o de Sharpe Generalizado (Sharpe, 1994) e de Sortino (Sortino e Van der Meer, 1991) não possam ser considerados conceitos recentes na literatura financeira, seu uso é generalizado e ainda objeto de muita pesquisa. Como se sabe, tais índices são medidas de desempenho ajustadas ao risco das carteiras, permitindo a comparação entre as mesmas. No entanto, devido as suas definições também é fato que não é fácil realizar inferências a seu respeito. Uma forma de contornar este problema é utilizar técnicas de *bootstrap* como sugerido por Vinod e Morey (1999a) e Scherer (2004), a partir das quais podemos construir intervalos de confiança para os índices de desempenho ajustados ao risco das carteiras de interesse.

Vinod e Morey (1999b) discutem o problema que o índice de Sharpe possui que é a dificuldade de construir uma medida do seu risco de estimação, e sugerem uma nova medida chamada de o "duplo" índice de Sharpe (*Double Sharpe Ratio*), que é definido como sendo a estimativa do índice de Sharpe dividido pelo desvio padrão da estimativa do índice de Sharpe. Este desvio-padrão da estimativa do índice de Sharpe é obtido através de um processo de *bootstrap*, e o duplo índice de Sharpe considera o risco da carteira e de estimação do índice de desempenho.

Neste artigo são apresentadas duas medidas de desempenho de fundos mútuos de investimento baseadas nos índices de Sharpe Generalizado e de Sortino, medidas de desempenho ajustadas ao risco, que também são ajustadas para o risco de estimação através de intervalos de confiança gerados por meio de *bootstrap*. O uso das medidas propostas é útil para fundos de investimento que empregam gestão ativa, ou seja, que tenham como objetivo superar um determinado índice de referência como, por exemplo, o IBOVESPA ou o CDI.

Os índices propostos neste artigo serão comparados com os índices de Sharpe Generalizado e de Sortino originais e com o duplo índice de Sharpe, com uma amostra de 100 fundos do mercado brasileiro no período de 2004, onde serão analisados os resultados de ranqueamento de diferentes fundos conforme seu desempenho.

2. Conceitos Gerais

2.1. Índice de Sharpe Generalizado (ISG)

Dadas n carteiras, desejamos avaliar a performance de cada uma delas. Para tal, consideremos o vetor dos log-retornos em excesso de cada uma delas, $r_t = (r_{1t},...,r_{nt})$, nos instantes t=1,...,T, sendo o log-retorno em excesso a diferença entre o retorno do fundo de investimento e um índice de referência (benchmark). Assumimos que o vetor dos log-retornos em excesso tem distribuição multivariada com média $\mu = (\mu_1,...,\mu_n)$ e matriz de covariâncias $\Sigma = (\sigma_{ij})_{1 \le i,j \le n}$. O índice de Sharpe Generalizado (Sharpe, 1994) é, então, definido para cada carteira como

$$ISG_i = \frac{\mu_i}{\sigma_i}$$

onde $\sigma_i^2 = \sigma_{ii}$ e i=1,...,n, e é igual a variância dos log-retornos excedentes do fundo de investimento e do índice de referência. Consequentemente, quanto maior o índice de Sharpe Generalizado, melhor a performance da carteira. O ISG deve ser estimado com base nos retornos em excesso das carteiras de interesse, ou seja, o vetor de retornos em excesso esperados μ e a matriz de covariâncias Σ devem ser substituídas, respectivamente, pelos seus estimadores $\hat{\mu}$ e $S = (s_{ij})_{1 \le i,j \le n}$, ou seja, pela média

$$\hat{\mu} \equiv \frac{1}{T}\sum_{t=1}^{T} r_t$$

e pela matriz de covariância amostral

$$S \equiv \frac{1}{T-1}\sum_{t=1}^{T}(r_t - \hat{\mu})(r_t - \hat{\mu})'.$$

Note que ao calcular os índices das carteiras, utilizaremos apenas os termos da diagonal de S, isto é, $\hat{\sigma}_i^2 = s_{ii}$. Deste modo, obtemos o seguinte estimador para o ISG:

$$\hat{ISG}_i = \frac{\hat{\mu}_i}{\hat{\sigma}_i}.$$

Muito embora, o estimador $\hat{ISG}_i$ seja simples de calcular, a presença de $\hat{\sigma}_i$ no denominador dificulta a determinação de intervalos de confiança para pequenas amostras, assim como a determinação de um tamanho amostral mínimo para podermos supor normalidade. Por outro lado, intervalos de confiança são úteis para medir a qualidade do estimador. Numa situação em que o desempenho de duas carteiras A e B estão sendo comparadas através do ISG, não necessariamente aquela com maior índice, A, é a que apresenta melhor performance. De fato, o intervalo de confiança de $\hat{ISG}_A$ pode ser muito maior do que o intervalo de confiança de $\hat{ISG}_B$ a ponto do fato $\hat{ISG}_A > \hat{ISG}_B$ tornar-se pouco relevante. Uma maneira eficiente e simples de se obter tais intervalos de confiança é através de técnicas de *bootstrap*.

2.2. Índice de Sortino

Além do Índice de Sharpe e suas variações como o ISG, outras medidas de desempenho ajustado ao risco existem como aquelas que utilizam o conceito de *downside deviation* com relação a um ponto de referência ou taxa de retorno mínimo aceitável, discutidas no artigo de Plantinga e De Groot (2001). O conceito de *downside deviation* (ou *downside risk*) diz respeito à parcela do risco com a qual o investidor está efetivamente preocupado, ou seja, o risco de que o retorno proporcionado por um ativo esteja abaixo do retorno mínimo aceitável. Uma medida baseada em *downside risk* é o índice de Sortino de Sortino e Van der Meer (1991), estimado por

$$Sortino = \frac{E[R] - R_{ma}}{\delta}$$

onde R_{ma} é a taxa de retorno mínimo aceitável, como o retorno diário do índice de referência do fundo, e δ é a medida de *downside risk* com respeito ao retorno mínimo aceitável. O valor de δ é dado por

$$\delta = \left[\frac{\sum_{R=-\infty}^{\tau} (\tau - R)^2}{N_\tau} \right]^{1/2}, \qquad (\tau - R) > 0$$

onde τ é o retorno mínimo aceitável (*target*), R é o retorno diário de um fundo e N_τ é o número de dias nos quais o retorno do fundo foi inferior ao *target* diário.

É importante destacar as diferenças existentes entre a utilização do desvio padrão (σ) e do *downside risk* (δ) como medidas de risco. Enquanto o desvio padrão dos retornos excedentes, utilizado no índice de Sharpe generalizado, mede a volatilidade dos retornos do ativo tanto acima quanto abaixo de uma taxa de referência, o *downside risk* preocupa-se apenas com a volatilidade indesejada, aquela que sujeita o investidor à perdas em seu patrimônio. Em outras palavras, o *downside risk* está associado ao cálculo da semi-variância. A medida de *downside risk*, por outro lado, funciona independentemente da distribuição de freqüência dos retornos. Conforme nota-se da expressão de δ, leva-se em conta apenas a parte da distribuição dos retornos que encontra-se abaixo do retorno mínimo aceitável.

2.3. Bootstrapping

Sob certas circunstâncias a inferência de uma dada estatística torna-se problemática devido a dificuldade em se determinar sua distribuição. Este problema é ainda mais desafiador quando se lida com amostras pequenas para as quais os resultados assintóticos não se aplicam. Em alguns outros casos, a própria análise assintótica pode ser problemática. Técnicas como *bootstrapping* ou *jacknife* fornecem alternativas para a realização de inferências, como a determinação de intervalos de confiança ou a realização de testes de hipóteses, em amostras destes tipos. Ambas as técnicas baseiam-se em simulação e, assim como na técnica de monte-carlo, geramos computacionalmente pseudo amostras aleatória (tais amostras são obtidas através de algum gerador de números aleatórios e, consequentemente, tem-se que a aleatoriedade delas é apenas “aproximada”) ou réplicas grandes o suficiente para realizarmos as inferências desejadas. A diferença entre o *bootstrap* ou o *jacknife* e o monte-carlo reside no fato em que, nos dois primeiros, é utilizada a função distribuição acumulada estimada a partir dos dados, enquanto que, no caso da técnica de Monte-Carlo, esta distribuição é conhecida a priori.

O foco deste artigo é o uso do *bootstrap* para a determinação de intervalos de confiança dos índices de Sharpe Generalizado e de Sortino baseados nos retornos de fundos mútuos de investimento. De fato, como já foi ressaltado acima, a presença do estimador do desvio padrão dos retornos ou da medida de *downside risk* no denominador destes índices dificulta a determinação das distribuições dos seus estimadores e, consequentemente, a inferência dos mesmos. Embora o *bootstrap* tenha sido introduzido por Efron (1979) basicamente como uma técnica não-paramétrica para a inferência de estatísticas para as quais os métodos usuais dificilmente se aplicariam, sua idéia evoluiu e uma série de variantes surgiram na literatura. Utilizaremos aqui duas delas, a saber, o *bootstrap* simples, o qual aqui será tratado simplesmente por *bootstrap,* e o *bootstrap* “studentizado”, ou *bootstrap-t,* que é um ajuste do *bootstrap* simples de modo a tornar os intervalos de confiança mais precisos.

Para descrever o *bootstrap,* consideremos uma amostra aleatória $X_1,...,X_n \sim F$, onde F é uma distribuição de probabilidade sobre a reta real. Defina $X = (X_1,...,X_n)$, e suponha que queremos determinar intervalos de confiança para o estimador $\hat{\theta} = \hat{\theta}(X)$ de θ. Realizar o *bootstrap* (não-paramétrico) significa sortear *N* pseudo amostras $X^* = (X_1^*,...,X_N^*)$ de acordo com a distribuição de probabilidades empírica $\hat{F}$ (ou seja, cada pseudo amostra X_i^* é obtida sorteando-se independentemente e com reposição termos do conjunto $X = (X_1,...,X_n)$ de acordo com a distribuição $\hat{F}$) baseada na amostra *X* para, em seguida, calcular o valor da estatística $\hat{\theta}$ para cada uma destas pseudo amostras obtendo, consequentemente, um conjunto de réplicas $\hat{\theta}^* = (\hat{\theta}_1^*,...,\hat{\theta}_N^*)$ de $\hat{\theta}$. Tomando *N* suficientemente grande, podemos utilizar $\hat{\theta}^*$ para construir intervalos de confiança de nível α para θ com extremos dados pelos quantis empíricos de ordens α e $1-\alpha$, respectivamente, e denotados por $\hat{\theta}_\alpha^*$ e $\hat{\theta}_{1-\alpha}^*$. O algoritmo *bootstrap* simples para determinação de intervalos de confiança de nível α para θ é descrito abaixo:

- **1º passo**: selecionar *N* pseudo amostras de tamanho *n* independentes e com reposição de $X = (X_1,...,X_n)$ de acordo com a distribuição empírica $\hat{F}$, isto é, $X^* = (X_1^*,...,X_N^*)$;
- **2º passo**: para cada pseudo amostra, calcular a réplica $\hat{\theta}_i^*$ para $i = 1,...,N$;
- **3º passo**: tomar como extremos do intervalo de confiança de nível α os quantis empíricos de ordens α e $1-\alpha$, respectivamente.

É possível ajustar o intervalo de confiança acima de modo a torná-lo mais confiável no sentido de conter o verdadeiro valor de θ subtraindo $\hat{\theta}$ dos valores obtidos através do *bootstrap* simples e dividindo o resultado pelo seus respectivos desvios-padrões (pseudo-) amostrais (ou desvios *bootstrap*), isto é, “studentizando” tais valores, daí nomearmos este procedimento de *bootstrap studentizado*, ou mais simplesmente, *bootstrap-t*. O procedimento abaixo descreve o algoritmo para determinação de intervalos de confiança para θ de acordo com o *bootstrap-t*:

- **1º passo**: gerar *N* pseudo amostras de tamanho *n* independentes e com reposição de acordo com a distribuição empírica $\hat{F}$, isto é, $X^* = (X_1^*,...,X_N^*)$;

- **2o passo:** para cada pseudo amostra, calcular a réplica $\hat{\theta}_i^*$ para $i = 1,...,N$;
- **3º passo**: "studentizar"as réplicas $\hat{\theta}_i^*$, obtendo os valores $\hat{t}_i = \dfrac{\hat{\theta}_i^* - \hat{\theta}}{\hat{D}P_i}$, para $i = 1,...,N$, onde $\hat{D}P_i$ é o valor estimado do desvio padrão de $\hat{\theta}_i^*$ com base na pseudo amostra X_i^*;
- **4º passo**: ordenar a sequência $\hat{t}_i$;
- **5º passo**: obter os quantis empíricos de ordens α e $1-\alpha$ da amostra *bootstrap* "studentizada", ou seja, determinar os valores $\hat{t}^{(\gamma)}$ tal que $\dfrac{\#\{\hat{t}_i \le \hat{t}^{(\gamma)}\}}{N} = \alpha$, para $\gamma \in \{\alpha, 1-\alpha\}$.
- **6º passo**: o intervalo de confiança *bootstrap-t* é dado por $\left(\hat{\theta} - \hat{t}^{(1-\alpha)}\hat{D}P, \hat{\theta} - \hat{t}^{(\alpha)}\hat{D}P\right)$, onde $\hat{D}P$ é o desvio padrão estimado de $\hat{\theta}$ com base na amostra original.

Observações. com relação ao algoritmo *bootstrap-t*:

- Para determinar $\hat{t}^{(\gamma)}$ com $\gamma \in \{\alpha, 1-\alpha\}$ utilizamos a seguinte regra:
  i) $\hat{t}^{(\alpha)}$ é o elemento na posição $N\alpha$ na sequência ordenada $\left(\hat{t}_{(1)},...,\hat{t}_{(N)}\right)$;
  ii) $\hat{t}^{(\alpha)}$ é o elemento na posição $N(1-\alpha)$ na sequência ordenada $\left(\hat{t}_{(1)},...,\hat{t}_{(N)}\right)$;
- Note que necessitamos estimar o desvio padrão, $\hat{D}P_i$, de $\hat{\theta}_i^*$ com base na pseudo amostra X_i^*. Em alguns casos, pode haver uma fórmula explícita para $\hat{D}P_i$ e, nesses casos, basta inserir nesta fórmula os valores da pseudo amostra X_i^*. No entanto, em grande parte dos casos, tal fórmula não é acessível e nestes casos, podemos estimar o desvio padrão utlizando o *bootstrap*, ou seja, tomamos N_i réplicas X_{ij}^* de X_i^* de modo que $\hat{D}P_i = \sqrt{\dfrac{1}{N_i - 1}\sum_{j=1}^{N_i}\left(\hat{\theta}_{ij}^* - \hat{\mu}_i\right)^2}$ onde $\hat{\theta}_{ij}^*$ é o estimador *bootstrap* calculado de acordo com a j-ésima pseudo amostra, X_{ij}^*, e $\hat{\mu}_i = \dfrac{1}{N_i}\sum_{j=1}^{N_i}\hat{\theta}_{ij}^*$ é a média *bootstrap* de $\hat{\theta}_i^*$. A experiência mostra que 50 réplicas X_{ij}^* são suficientes para que se obtenha um bom estimador do desvio padrão (Efron e Tibshirani, 1993).

A experiência demonstra que o intervalo de confiança obtido a partir do *bootstrap-t* é inerentemente mais preciso do que aquele obtido via o *bootstrap* simples. De fato, para estatísticas que são aproximadamente normais, o *bootstrap* simples implica em uma precisão de primeira ordem, ou seja, $\Pr(\theta \le \theta_\alpha^*) = \alpha + O(n^{-1/2})$, enquanto que o *bootstrap-t* implica em uma precisão de segunda ordem, ou seja, $\Pr(\theta \le \theta_\alpha^*) = \alpha + O(n^{-1})$. Em outras palavras, o intervalo de confiança via *bootstrap-t* tem maior probabilidade de conter o verdadeiro valor de θ do que intervalo de confiança obtido via bootstrap simples. Davinson e Hinkley (1997), discute detalhadamente sobre estes intervalos de confiança e a comparação entre ambos.

Uma outra possibilidade, não explorada neste artigo, é o *bootstrap* paramétrico cujo procedimento é similar ao *bootstrap* simples, mas com a diferença de que agora a distribuição dos dados é conhecida, exceto por um número finito de parâmetros que devem ser estimados com base na amostra. Uma vez determinada a distribuição, podemos sortear N pseudo amostras e para cada uma delas, calcular o valor da estatística de interesse, e assim realizar as inferências desejadas exatamente como nas alternativas de *bootstrap* apresentadas anteriormente.

3. Medidas de Desempenho Ajustadas aos Riscos da Carteira e Estimação

Nesta seção introduzimos os índices de Sharpe e de Sortino ajustados. Tal ajuste consiste em se tomar a razão entre as estimativas destes índices pelos comprimentos de seus respectivos intervalos de confiança. Note que esta nova medida reduz as distorções citadas na introdução deste artigo e que se referem a possibilidade de um ranqueamento infeliz dos fundos quando nos baseamos exclusivamente nas estimativas dos índices de Sharpe ou de Sortino sem considerar a dispersão de ambos. De fato, ao se dividir tais estimativas pelos comprimentos de seus intervalos de confiança, estamos favorecendo os fundos cujos índices de Sharpe ou de Sortino apresentam menor disperção.

3.1. Índice de Sharpe Generalizado Ajustado (ISGA(α))

O índice de Sharpe Generalizado Ajustado (ISGA(α)) além de considerar o retorno excedente do fundo em relação a um índice referência sobre o desvio-padrão dos retornos diários excedentes do fundo em relação ao índice de referência, também divide este resultado pelo intervalo de confiança com nível de significância α para ajustar o ISG ao risco de estimação. Este intervalo de confiança é obtido por meio de um procedimento de *bootstrap*. Dessa forma o ISGA(α) é definido pela fórmula:

$$ISGA(\alpha) = \frac{\hat{I}SG}{|ICISG(\alpha)|}$$

onde ÎSG é o índice de Sharpe Generalizado estimado pelo *bootstrap*, ICISG(α) é o intervalo de confiança do índice de Sharpe Generalizado estimado pelo *bootstrap* com nível de significância igual a α e $|I|$ representa o comprimento do intervalo I .

Com esta medida está considerando-se a relação do retorno do fundo excedente ao índice referência sobre o risco adicional ao índice de referência assumido, e o risco de estimação que reduzirá o valor da medida do ISG quanto maior for o intervalo de confiança da estimativa do índice de Sharpe Generalizado.

3.2. Índice de Sortino Ajustado (ISOA(α))

O índice de Sortino Ajustado (ISOA(α)) relaciona o índice de Sortino estimado, que divide o retorno do fundo excedente ao índice referência sobre o risco de perda em relação ao índice de referência (*downside risk*), com o risco de estimação desse índice estimado pelo intervalo de confiança da estimativa do índice de Sortino. O ISOA(α) é definido pela fórmula:

$$ISOA(\alpha) = \frac{\hat{I}SO}{|ICISO(\alpha)|}$$

onde ÎSO é o índice de Sortino estimado pelo *bootstrap*, e ICISO(α) é o intervalo de confiança do índice de Sortino estimado pelo *bootstrap* com nível de significância igual a α e $|I|$ representa o comprimento do intervalo I .

As medidas propostas devem afetar o ranqueamento dos fundos de investimento antes realizado somente com o ISG e o índice Sortino devido aos intervalos de confiança, que “penalizam” os índices estimados pelas fórmulas tradicionais. Além disso, as estimativas do ISG e do índice de Sortino são obtidas por meio de uma distribuição de probabilidades construída por um procedimento de bootstrap, e não apenas do emprego da fórmula do respectivo índice.

4. Metodologia e Dados

4.1. Amostra

A amostra foi composta por 100 fundos de investimento com índice de Sharpe Generalizado e Sortino maiores que zero (estes índices não são aplicáveis para ranqueamento se forem negativos), não-exclusivos relacionados no banco de dados SI-Anbid da ANBID (Associação Nacional de Bancos de Investimento), com dados sobre cotas diárias dos fundos no ano de 2004. Os fundos selecionados na amostra foram subdivididos por categorias de risco e gestão ativa ou passiva, conforme classificação da ANBID. Os fundos de gestão passiva, cujo objetivo é acompanhar (e não superar) o desempenho do seu índice de referência (*benchmark*), foram eliminados, pois neste trabalho vamos nos ater aos fundos cujo objetivo seja superar o índice de referência. Também foram excluídos da amostra os fundos de previdência, de privatização, capital protegido e off-shore por não possuírem um índice de referência claro ou por possuírem objetivos diferentes de superar a sua referência.

Desta forma obtivemos a seguinte amostra classificada por categorias da ANBID:

Classificação ANBID	**Patrimônio Líquido em 31/12/2004 (R$)**	**Número de Fundos**	**Patrimônio Líquido Médio (R$)**
Ações IBOVESPA Ativo	1.003.157.943,51	24	41.798.247,65
Ações IBOVESPA Ativo Com Alavancagem	1.511.364.732,99	12	125.947.061,08
Ações IBX Ativo	740.492.279,48	7	105.784.611,35
Ações Outros	880.767.510,48	11	80.069.773,68
Ações Outros Com Alavancagem	32.864.795,27	3	10.954.931,76
Balanceados	11.606.112,33	1	11.606.112,33
Multimercados Com RV	35.321.693,24	1	35.321.693,24
Multimercados Com RV Com Alavancagem	6.062.830.805,80	12	505.235.900,48
Multimercados Sem RV	5.403.830.051,74	8	675.478.756,47
Multimercados Sem RV Com Alavancagem	306.225.699,24	4	76.556.424,81
Renda Fixa	27.292.948.001,37	6	4.548.824.666,90
Renda Fixa Crédito	6.749.909.609,63	7	964.272.801,38
Renda Fixa Multi-Índices	659.998.581,22	4	164.999.645,31
Total	**50.691.317.816,30**	**100**	**506.913.178,16**

Tabela 1. Classificação e patrimônio dos fundos da amostra

Para cada categoria de fundos associamos um determinado índice de referência, que será utilizado para o cálculo dos log-retornos excedentes, e uma categoria de risco que será empregada para ranqueamento dos fundos, que descrevemos na tabela abaixo:

Categoria ANBID	Índice de Referência	Categoria de Risco
Ações IBOVESPA Ativo Com Alavancagem	Ibovespa Médio	Ações
Ações IBOVESPA Ativo	Ibovespa Médio	Ações
Ações IBX Ativo	IBX Médio	Ações
Ações Outros Com Alavancagem	IBX Médio	Ações
Ações Outros	IBX Médio	Ações
Balanceados	CDI	Multimercados
Multimercados Com RV Com Alavancagem	CDI	Multimercados
Multimercados Sem RV Com Alavancagem	CDI	Multimercados
Multimercados Com RV	CDI	Multimercados
Multimercados Sem RV	CDI	Multimercados
Renda Fixa	CDI	Renda Fixa
Renda Fixa Crédito	CDI	Renda Fixa
Renda Fixa Multi-Índices	CDI	Renda Fixa

Tabela 2. Classificação e índices de referência dos fundos da amostra

De posse das cotas diárias dos fundos, e da informação se elas são de abertura ou de fechamento proveniente do SI-Anbid, calcula-se o log-retorno excedente diário de cada fundo da seguinte forma:

$$RE_t = \ln\left(\frac{Cota_t}{Cota_{t-1}}\right) - \ln\left(\frac{Benchmark_t}{Benchmark_{t-1}}\right)$$

onde RE_t é o log-retorno excedente diário de um fundo qualquer, $Cota_t$ é o valor da cota do fundo na data t, e $Benchmark_t$ é o valor do índice de referência do fundo na data t. São estes retornos excedentes que serão utilizados como dados base da metodologia apresentada a seguir.

4.2. Metodologia

Para comparar as diferenças entre os índices de Sharpe Generalizado e de Sortino, e duplo Sharpe Generalizado com as medidas propostas (ISGA(α) e ISOA(α)) são realizados os seguintes passos:

a) São calculados os log-retornos excedentes diários de cada fundo (RE_t) para as observações de um dos anos da amostra;
b) Com base nos log-retornos do item (a) são estimados o ISG e de Sortino por meio das fórmulas definidas originalmente pelos autores destes índices de desempenho;
c) Utiliza-se o procedimento de *bootstrap* simples, com os log-retornos do item (a), para estimar a distribuição de probabilidades do ISG e de Sortino. Neste bootstrap, para cada fundo de investimento, são geradas 1000 amostras pseudo-aleatórias, de tamanho 50 cada, ou seja, contendo 50 log-retornos excedentes diários retirados dos log-retornos do anos de 2004, que contém 251 observações. Os índices de desempenho são estimados com cada uma das amostras de 50 observações;
d) Com as distribuições de probabilidades do item (c) e para o nível de significância de 90% são estimados intervalos de confiança para os índices de desempenho, e os índices propriamente ditos;
e) Os resultados do item (d) são empregados para estimar as medidas propostas ISGA(α) e ISOA(α);

f) É realizado o *bootstrap-t* com os log-retornos do item (a), para estimar a distribuição de probabilidades do ISG e de Sortino, e se repete os passos dos itens (d) ao (e) acima com o resultado deste *bootstrap-t*. Neste *bootstrap*, para cada fundo de investimento, são geradas 1000 amostras pseudo-aleatórias, de tamanho 50 cada. Para cada amostra de 50 log-retornos, usada para estimar os índices de desempenho, é estimado o desvio-padrão de cada índice estimado, por meio de um *bootstrap* simples com 50 amostras de tamanho 50, gerada a partir da amostra de 50 log-retornos utilizados para estimar os índices de desempenho;
g) Para o ano de 2004, e cada uma das medidas de desempenho estimadas é realizado um ranqueamento dos fundos por decis;
h) São comparados as diferenças de ranqueamento dos fundos em um ano conforme cada medida de desempenho, considerando os decis do item (h), e estimada a correlação de ranqueamento de Spearman entre as medidas apresentadas.

Os resultados da aplicação da metodologia na amostra de fundos são analisados na próxima seção.

5. Análise dos Resultados

Para cada fundo da amostra foram estimadas as seguintes medidas: o ISG original (como formulado originalmente pelo seus autores), o índice de Sortino original, o ISG e o Sortino médio (obtido das médias das 1000 amostras do bootstrap), o duplo ISG e Sortino (resultado dos índices médios sobre os desvios-padrão estimados dos índices), as medidas propostas ISGA e ISOA estimadas com intervalo de confiança de 90% e com os procedimentos de bootstrap simples e studentizado.

Com os resultados obtidos das medidas, foi realizado um ranqueamento dos fundos da amostra para cada uma das medidas e os fundos foram divididos em decis, estando os melhores fundos, segundo a medida utilizada, no primeiro decil e os piores no último decil. Tanto para o índice de Sharpe como de Sortino, e suas variantes, quanto maior e positivo for a medida melhor foi o desempenho do fundo.

Os resultados dos ranqueamentos do fundos pelas medidas forma tabulados e comparados nas tabelas a seguir:

Núm. de fundos	**Decis do ISG Médio**									
Decis do ISG Original	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	9	1								
2	1	6	3							
3		3	6	1						
4			1	7	2					
5				2	4	2		1		1
6					4	4	1	1		
7						3	5	1		1
8						1	3	2	2	2
9								4	5	1
10							1	1	3	5
Total de fundos	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**

Tabela 3. Comparação ranqueamento ISG Original vs. ISG Médio

A Tabela 3 exibe a comparação dos ranqueamentos dos fundos da amostra realizados com o ISG original e o ISG médio proveniente do *bootstrap*. A tabela pode ser interpretada da seguinte forma: do total de 10 fundos que se tem em cada decil, em particular do quarto decil do ranqueamento realizado pelo ISG original, um fundo foi para o decil 3 (subiu de posição) segundo o ranqueamento pelo ISG médio, 7 fundos se mantiveram no decil 4, e 2 fundos foram deslocados (caíram de posição) par o decil 5. Pode-se notar que os fundos mudam consideravelmente de posição no ranqueamento pelo ISG médio em relação ao ISG original já a partir do segundo decil do ranqueamento pelo ISG original, indicando que a inferência do ISG através de bootstrap já interfere nas decisões que o investidor tomará com base no ranqueamento realizado. No entanto, não há praticamente diferença de ranqueamento dos 10 melhores fundos conforme o ISG original.

Núm. de fundos	**Decis do Duplo ISG**									
Decis do ISG Original	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	9	1								
2	1	7	2							
3		2	7	1						
4			1	8	1					
5				1	6	1		1		1
6					3	6	1			
7						2	6	1		1
8						1	3	2	2	2
9								4	5	1
10								2	3	5
Total de fundos	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**

Tabela 4. Comparação ranqueamento ISG Original vs. Duplo ISG

A Tabela 4 exibe a comparação dos ranqueamentos dos fundos da amostra realizados com o ISG original e o Duplo ISG de Vinod e Morey (1999). Os resultados destes dois ranqueamentos comparados é similar ao da Tabela 3, que apresenta a comparação do ISG original como o ISG médio. Pode-se concluir que o resultado do ranqueamento pelo ISG médio é similar ao do realizado pelo Duplo ISG, como pode-se notar na Tabela 5 abaixo:

Núm. de fundos	**Decis do ISG Médio**									
Decis do Duplo ISG	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	9	1								
2	1	8	1							
3		1	9							
4				8	2					
5				2	6	2				
6					2	7	1			
7						1	8	1		
8							1	9		
9									10	
10										10
Total de fundos	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**

Tabela 5. Comparação ranqueamento Duplo ISG vs. ISG Médio

O duplo ISG corrige o ISG médio para o risco de estimação, mas aparentemente esta correção não gerou grandes mudanças no ranqueamento como se vê na Tabela 5.

Na Tabela 6 compara-se o ranqueamento pelo ISG original com o realizado pela medida proposta neste artigo, o ISGA com, neste caso, estimado com 90% de confiança através do *bootstrap* simples. O ISGA considera o risco de estimação, assim como o duplo ISG, e por isso afeta consideravelmente as posições dos fundos nos ranqueamentos, como pode ser visto na Tabela 6, a seguir.

Núm. de fundos	Decis do ISGA Bootstrap Simples									
Decis do ISG Original	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	9	1								
2	1	6	3							
3		3	6	1						
4			1	7	2					
5				2	5	1		1		1
6					3	6	1			
7						2	6	1		1
8						1	3	2	2	2
9								4	5	1
10								2	3	5
Total de fundos	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**

Tabela 6. Comparação ranqueamento ISG Original vs. ISGA com Bootstrap Simples

Percebe-se as comparações do ISG original com o duplo ISG, e do ISG original com o ISGA com *bootstrap* simples não diferem muito entre si, o que pode ser constatado na Tabela 7 abaixo, onde se vê que os fundos praticamente permanecem nos mesmos decis que estavam seja pelo ranqueamento pelo duplo ISG, seja pelo ISGA com *bootstrap* simples.

Núm. de fundos	Decis do ISGA Bootstrap Simples									
Decis do Duplo ISG	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	10									
2		9	1							
3		1	9							
4				9	1					
5				1	9					
6						10				
7							10			
8								9	1	
9								1	9	
10										10
Total de fundos	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**

Tabela 7. Comparação ranqueamento do Duplo ISG Original vs. ISGA com Boostrap Simples

A comparação dos ranqueamentos pelo ISG original com a medida proposta ISGA estimado pelo bootstrap studentizado, com nível de confiança 90%, mostra uma grande diferença nas posições dos fundos nos decis, e isto já começa ser percebido nos fundos do primeiro decil conforme o ranqueamento pelo ISG original, como se vê na Tabela 8:

Núm. de fundos	Decis do ISGA Bootstrap Studentizado									
Decis do ISG Original	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	7	3								
2	3	4	3							
3		3	5	2						
4			2	6	2					
5				2	5		1	1		1
6					3	6		1		
7						3	6			1
8						1	3	3	1	2
9								4	5	1
10								1	4	5
Total de fundos	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**

Tabela 8. Comparação ranqueamento ISG Original vs. ISGA om Bootstrap Studentizado

Na Tabela 9 tem-se a comparação dos ranqueamentos pelo duplo ISG e o ISGA estimado com o bootstrap studentizado, onde se percebe várias mudanças nos ranqueamentos apesar das duas medidas considerarem o risco de estimação.

Núm. de fundos	Decis do ISGA Bootstrap Studentizado									
Decis do Duplo ISG	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	6	4								
2	4	2	3	1						
3		4	6							
4			1	6	3					
5				3	5	1	1			
6					2	7	1			
7						2	7	1		
8							1	8	1	
9								1	9	
10										10
Total de fundos	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**

Tabela 9. Comparação ranqueamento do Duplo ISG vs. ISGA com Boostrap Studentizado

A correlação de ranqueamento de Spearman (*Spearman´s rank correlation*) foi estimada para as medidas de desepenho relacionadas com o índice de Sharpe Generalizado, e é apresentada na tabela abaixo:

Correlação de Ranqueamento Spearman	ISG Original	ISG Médio	Duplo ISG	ISGA Bootstrap Simples	ISGA Bootstrap Studentizado
ISG Original	1,000	0,945	0,948	0,949	0,942
ISG Médio	0,945	1,000	0,998	0,998	0,987
Duplo ISG	0,948	0,998	1,000	1,000	0,986
ISGA Bootstrap Simples	0,949	0,998	1,000	1,000	0,985
ISGA Bootstrap Studentizado	0,942	0,987	0,986	0,985	1,000

Tabela 10. Matriz de correlação de ranqueamento de Spearman para medidas de ISG

A matriz na Tabela 10 mostra que apesar da existência de diferenças de ranqueamento entre os decis das várias medidas, todas possuem entre si uma alta correlação, indicando juntamente com as análises anteriores que existem mudanças nos ranqueamentos entre as medidas, mas estas mudanças são pequenas, ou seja, os deslocamentos dos fundos nos ranqueamentos são pequenos quando comparados entre si.

Foi realizado também o mesmo estudo com o índice de Sortino, que lida com o conceito de *downside risk*, as análises anteriores foram refeitas para os resultados que são apresentados nas tabelas abaixo.

Núm. de fundos	Decis do Sortino Duplo										Decis do Sortino Médio									
Decis do Sortino Original	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	6	2				1		1			9	1								
2	4	4	1	1							1	8		1						
3		4	5			1						1	5	4						
4			3	6						1			2	3	5					
5			1	3	4	2							2	1	2	5				
6					2	4	3	1							1	3	4	2		
7					3	2	2	3					1		2	1	3	3		
8							2	5	3								2	3	4	1
9					1		1		5	3				1				2	4	3
10							2		2	6						1	1		2	6
Total de fundos	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**

Tabela 11. Comparação do ranqueamento do Sortino original com o Sortino Duplo e Sortino Médio

Os resultados obtidos com o ranqueamento do Sortino original com o Sortino médio obtido pelo *bootstrap* simples, e com o Sortino duplo (versão do duplo Sharpe para o índice de Sortino) mostram que considerar o risco de estimação, como faz o Sortino duplo, causa grandes variações nos ranqueamento devido as mudanças dos fundos nos decis.

Núm. de fundos	Decis do ISOA com Bootstrap Simples										Decis do ISOA com Bootstrap Studentizado									
Decis do Sortino Original	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	7	1			1			1			2		1			3		1	1	2
2	3	5	1	1							4	2	2				1			1
3		3	6			1					3	2	3					1		1
4		1	2	5	1					1	1	3	1	1	2			1		1
5			1	2	5	2						2	3	2	2					1
6				1		4	4	1				1		4	2	1		2		
7				1	2	1	3	3						3	3	2		1	1	
8							2	3	4	1					1	3	4	1	1	
9					1		1	1	4	3						1	5	1	2	1
10						2		1	2	5								2	5	3
Total de fundos	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**

Tabela 12. Comparação do ranqueamento com o Sortino original com o ISOA estimado com boostraps

Na Tabela 12 tem-se os ranqueamentos do Sortino original com o ISOA estimado com nível de confiança de 90% e pelos métodos de *bootstrap* simples e studentizado. Assim como ocorreu com as análises sobre o ISG, as mudanças de ranqueamento do ISOA com *bootstrap* simples foram similares às ocorridas como o Sortino Duplo. No entanto, o ranqueamento com o ISOA com *bootstrap* studentizado apresenta grandes variações em comparação com as outras medidas, inclusive com as outras medidas que ajustam o desempenho dos fundos para o risco de estimação, como é exibido na Tabela 13 abaixo.

Núm. de fundos	Decis do ISOA com Bootstrap Simples										Decis do ISOA com Bootstrap Studentizado									
Decis do Sortino Duplo	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	8	2									5		2			2		1		
2	2	6	2								3	2	2			1	1		1	
3		2	7	1							1	5	2		1			1		
4			1	7	2						1	1	1	3	2			1		1
5				2	6	2						2	2		1	2		1	2	
6					2	6	2						1	4				2		3
7						2	8							1	3	2	2		2	
8								8	2					2	3	1	1	1	1	1
9								2	7	1						2	6	1	1	
10									1	9								2	3	5
Total de fundos	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**

Tabela 13. Comparação do ranqueamento com o Sortino Duplo com o ISOA estimado com boostraps

Ao realizar a análise da correlação de ranqueamento de Spearman percebe-se que as medidas possuem alta correlação entre si, com exceção da medida ISOA estimada com *bootstrap* studentizado que apresentou baixas correlações com as medidas que não ajustam para o risco de estimação, e correlações um pouco mais elevadas com as outras medidas que ajustam ao risco de estimação, no caso o Sortino Duplo e o ISOA estimada com *bootstrap* simples. Uma das possíveis explicações para este efeito é o uso da semi-variância na medida desenvolvida por Sortino e Van der Meer (1991), e o fato do bootstrap studentizado utilizar intervalos de confiança que têm maior probabilidade de conter o valor verdadeiro da variável estimada como colocado por Davinson e Hinkley (1997).

Correlação de Ranqueamento Spearman	Sortino Original	Sortino Médio	Duplo Sortino	ISOA Bootstrap Simples	ISOA Bootstrap Studentizado
Sortino Original	1,000	0,926	0,870	0,874	0,493
Sortino Médio	0,926	1,000	0,875	0,903	0,326
Duplo Sortino	0,870	0,875	1,000	0,993	0,628
ISOA Bootstrap Simples	0,874	0,903	0,993	1,000	0,564
ISOA Bootstrap Studentizado	0,493	0,326	0,628	0,564	1,000

Tabela 14. Matriz de correlação de ranqueamento de Spearman para medidas de Sortino

Os resultados obtidos de cada uma das medidas apresentadas, e o banco de dados de log-retornos excedentes ao índice de referência foram excluídas deste artigo, mas estão disponíveis com os autores para quem estiver interessado.

6. Comentários e Pesquisa Futura

Neste trabalho foram apresentadas duas novas medidas de avaliação de desempenho, baseadas nos índices de Sharpe Generalizado de de Sortino, que relacionam retorno excedente do fundo de investimento com o risco do fundo e seu índice de referência, e ajustam este valor ao risco de estimação. O ajuste ao risco de estimação é realizado através do comprimento do intervalo de confiança da estimativa do ISG ou de Sortino, que é obtida por meio de um método de bootstrap, sendo que aqui empregados dois tipos diferentes de *bootstrap*.

Com as análises dos ranqueamentos dos fundos pela várias medidas pode-se concluir que não há grandes alterações nos ranqueamentos se utilizarmos o as medidas originais ou as medidas originais estimadas por *bootstrap* simples. Assim, devido ao tempo e custo computacional não compensaria fazer um *bootstrap* se não se deseja considerar o risco de estimação da medida, mas sim a própria medida original, resultado confirmado pela correlação de ranqueamento de Spearman.

Já considerando o ajuste das medidas de desempenho originais pelo risco de estimação, tem-se que as medidas propostas índice de Sharpe Generalizado Ajustado (ISGA) e índice de Sortino Ajustado (ISOA) estimadas com *bootstrap* studentizado levam a ranqueamentos de fundos bastante diferentes dos demais, em destaque os resultados da correlação de Spearman mostram que há grandes mudanças de ranqueamento quando utiliza-se o ISOA em comparação às outras medidas, mas no caso do ISGA com *boostrap* studentizado não leva o ranqueamento a mudanças tão grandes quanto o ISGA estimado com *bootstrap* simples. O custo computacional do bootstrap studentizado é muito grande, e demanda computadores de grande velocidade de capacidade, e é justificado se o risco de estimação ou a volatilidade do fundo forem muito altos. Os raqueamentos das medidas propostas também pode variar em função do nível de confiança escolhido para estimá-las.

Este trabalho apresenta as medidas ISOA e ISGA que ajustam a relação retorno e risco do fundo ao risco de estimação da própria medida de desempenho, e mostra que as medidas afetam o ranqueamento de fundos e conseqüentemente a decisão de alocação de recursos do investidor. A partir deste trabalho uma extensão natural é a avaliação da persistência e consistência dos resultados dos fundos de investimento utilizando as medidas ISOA e ISGA como referência de desempenho.

7. Bibliografia

- DAVISON, A.C.; HINKLEY, D.V. *Bootstrap Methods and Their Application*. Cambridge University Press. 1997.
- EFRON, B; TIBSHIRANI, R. J. *Na Introduction to the Bootstrap.* New York. Chapman & Hall. 1993.
- EFRON, B. Bootstrap methods: Another look at the Jacknife. *Annals of Statistics*, v.7, p.1-26, 1979.
- PLANTINGA, A.; DE GROOT, S. Risk-adjusted performance measures and implied risk-attitudes. Working Paper, 2001. Disponível em <http://ssrn.com/abstract=289193>. Acesso em: 02 dez. 2004.
- SCHERER, B. An alternative route to performance hypothesis testing. *Journal of Asset Management*, v.5, n.1, june 2004.
- SHARPE, W.F. Mutual fund performance. *The Journal of Business*, v.2, n.1, p.119-138, jan. 1966.

- SHARPE, W.F. “The Sharpe ratio”. *Journal of Portfolio Management*, v.21, p.49-58, fall, 1994.
- SORTINO, F.A.; VAN DER MEER, R. The Dutch triangle. *Journal of Portfolio Management*, v.18, p.27-31, summer 1991.
- VINOD, H.D.; MOREY, M.R. Confidence intervals and hypothesis testing for the Sharpe and Treynor performance measures: A bootstrap approach. In: ABU-MOSTAFA, Y.S.; LEBARON, B.; LO, A.W.; WEIGEND, A.S. (eds.). *Computational Finance 1999*. Cambridge: The MIT Press, p.25-39, 1999a.
- VINOD, H.D.; MOREY, M.R. A “Double” Sharpe ratio. Working Paper, 1999b. Disponível em <http://ssrn.com/abstract=168748>. Acesso em: 02 dez. 2004.