3. Metodologia

Nessa sessão, vamos apresentar a forma como os portfólios usados foram calculados, discutir os desafios do procedimento e apresentar possíveis melhorias ao método atual de cálculo.

Seguindo o método já estabelecido na literatura (Newbould e Poon (1993), Fisher e Lorie (1968), entre outros), calculamos para cada tamanho de portfólio possível diversos portfólios aleatórios, e, para cada um deles, calculamos estatísticas de retorno e risco (no caso mais simples, usamos desvio padrão do retorno como métrica de risco). Agrupando os resultados, calculamos valores de interesse para a métrica de risco escolhida, como a mediana, os valores a um desvio padrão de distância da mediana e os valores críticos para que englobemos 95% dos resultados obtidos. Essas estatísticas devem se aproximar das populacionais, como veremos. A amostragem é necessária pois tentar calcular todos os Portfólios possíveis leva, necessariamente, a um problema computacional. Em um universo de 60 ações, por exemplo, o total de portfólios possíveis (sem que se leve em considerações estruturas de pesos diferentes) é um pouco maior do que $1.1 * 10^{18}$. Calcular retornos e desvios padrões dessas quantidades de objetos não é factível com os recursos de que dispomos. Para que tenhamos maior confiança em nossos dados, repetimos os exercícios para amostras de tamanhos diferentes (isso é, primeiro usamos apenas 50 portfólios por tamanho de portfólio possível, progredindo para 100, 5.000, 50.000 e, por fim, 150.000. Entre cada exercício, os portfólios utilizados não são reciclados nem excluídos; cada amostra é gerada independentemente das anteriores, de forma que os 100 portfólios aleatórios de 18 ações podem ou não englobar os 50 gerados anteriormente). Em todos os casos, há tamanhos de portfólios para os quais é possível se calcular o total da população; quando o tamanho de portfólio atinge o total do mercado, por exemplo, só há um portfólio possível. Similarmente, para portfólios de tamanho igual a 1, a população de portfólios possíveis é igual à quantidade de ativos disponíveis. Sempre que a população de portfólios possíveis for menor do que o tamanho amostral selecionado, ela é utilizada.

O uso de amostragem introduz um problema, no entanto, enunciado já por Markowitz (1952, tradução livre): “[Essa] suposição, de que a lei dos grandes número se aplique a um Portfólio de ativos, não pode ser aceita. Os retornos de ativos são muito intercorrelacionados.

Diversificação não é capaz de eliminar toda a variância”. Apesar de Markowitz destacar a impossibilidade de se eliminar todo o risco, estamos mais interessados em outra consequência da inaplicabilidade da lei dos grandes números para portfólios: não há garantias “automáticas” de que a média do desvio padrão dos retornos amostrais convirja para a populacional. Isso põe sob dúvida o método proposto acima, e amplamente reproduzido.

Contornamos essa questão de forma empírica: para cada tamanho de Portfólio possível, realizamos testes de normalidade sobre os resultados. Por construção, os testes de normalidade clássicos, como o de Shapiro-Wilk, enfrentam problemas para lidar com amostras muito grandes, visto que qualquer divergência em relação aos ideais teóricos de curtose e assimetria são considerados significativos. Isso faz com que o resultado quase sempre seja que amostras não são normalmente distribuídas, mesmo quando a divergência é, para todos os efeitos, irrelevante. Para tratar disso, quando usamos amostras cujos tamanhos fariam com que os testes de normalidade perdessem significado preferimos adotar uma forma de testagem mais permissiva, que permite que ambos os momentos em questão sejam ligeiramente diferentes dos ideais. Especificamente, ao invés de testarmos que excesso de curtose e assimetria em relação à normal sejam iguais a zero, permitimos que ambos sejam menores de que 1 em módulo. Produzimos também gráficos das distribuições empíricas (densidade de probabilidade, probabilidade acumulada e quantil-quantil), que permitem contato mais direto com nossos resultados.

Por fim, como dito, calculamos além dos resultados medianos para cada tamanho de portfólio possível algumas bandas de interesse: mais ou menos um desvio padrão e o intervalo de 95% dos resultados. Isso segue o argumento feito por Newbould e Poon (1993) de que ao escolher um portfólio qualquer, uma investidora não tem garantias de estar selecionando o portfólio médio, o que faz com que os intervalos citados sejam de interesse. Além disso, como estamos lidando com amostras de populações muito grandes, queremos ter garantias de que mesmo que nossas medianas não sejam perfeitamente representativas, nossos resultados são válidos para um espectro amplo de realizações possíveis. Como esse segundo problema não se aplica aos casos em que a totalidade de portfólios possíveis é calculada, e por uma questão de clareza e poluição visual, suprimimos a apresentação dos intervalos nesses casos.

O procedimento em uso atualmente pode ser aprimorado de duas formas notáveis. A primeira melhoria, de inspiração teórica, seria fazer variar os momentos iniciais e finais dos Portfólios, como feito em Chong e Phillips (2012), para que não estejamos sujeitos a vieses de tempo (como são calculados hoje, todos os nossos Portfólios são iniciados e finalizados todos nos mesmos momentos). A segunda, de inspiração técnica, é adotar uma forma de *benchmarking* anterior ao cálculo das amostras para que cada conjunto de amostras seja calculado de forma otimizada, em lugar de simplesmente se impor a mesma quantidade máxima de Portfólios em todos os conjuntos. Isso permitiria maior controle sobre o tempo levado para a criação de novos exemplos, permitindo a análise de um universo maior de simulações.

Por fim, todos os procedimentos, códigos e resultados deste trabalho estão disponíveis em https://github.com/ZeVZobaran/ibovPortfoliosSd. Os cálculos, gráficos e análises foram realizados em R, com uso principalmente das bibliotecas da distribuição do *tidyverse*. Este repositório está sujeito a alterações. Os cálculos até agora foram feitos com apenas 37 ações, um subgrupo pequeno do total no mercado brasileiro.

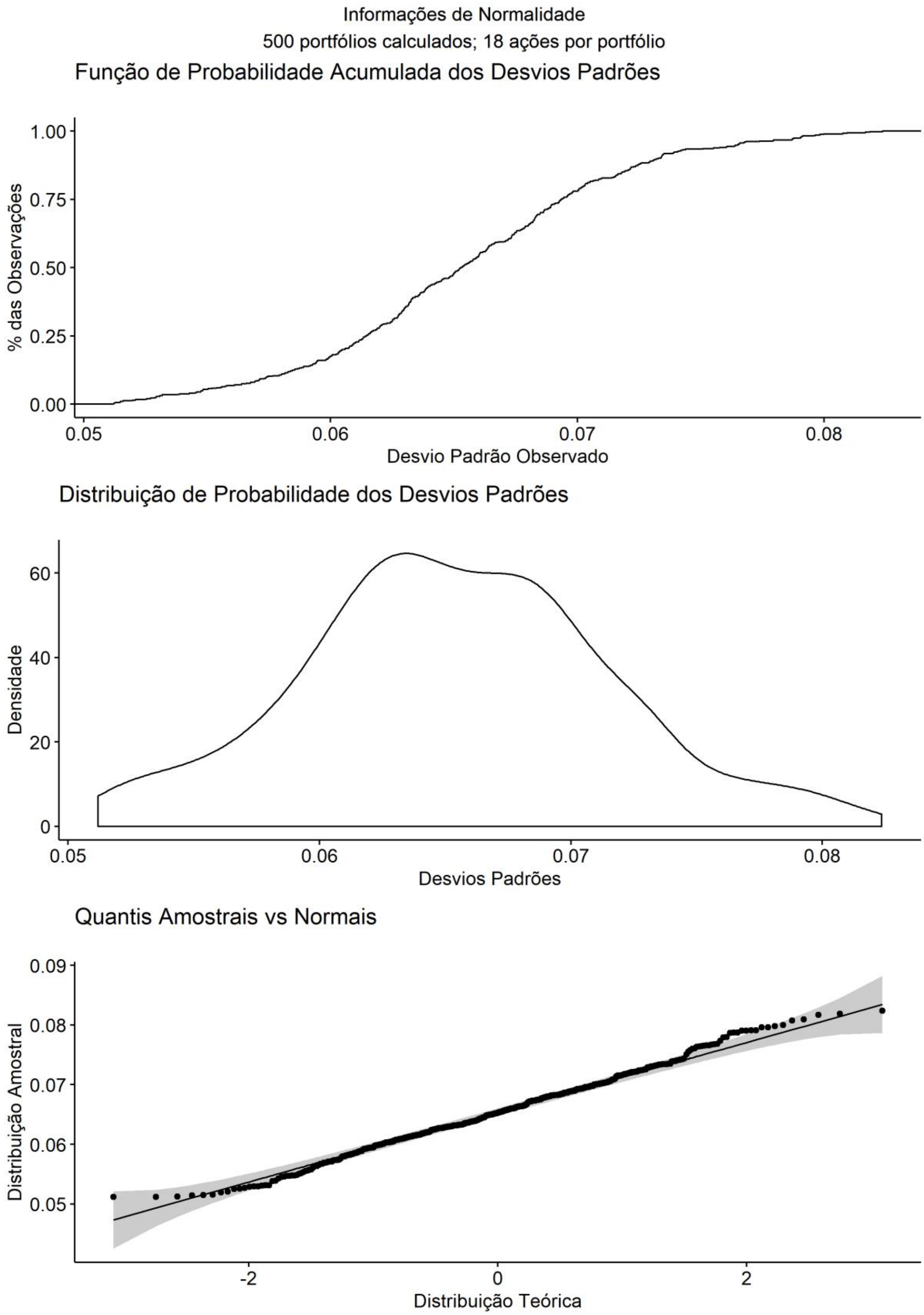
3. Apresentação dos Resultados

3.1 Portfólios *equally-weighted*

O primeiro caso discutido será aquele em que investidores selecionam ações aleatoriamente para compor portfólios em que cada ação tem o mesmo peso. Começamos por este caso principalmente pela relativa simplicidade de se aplicar o nosso algoritmo de cálculo. Como consequência dessa formulação, não fazemos suposições quanto ao retorno dos portfólios em cada amostra, que podem variar livremente. Dado que as ações que compõe o portfólio são selecionadas de forma aleatória e o peso delas é fixo, o retorno do portfólio estará dado de antemão.

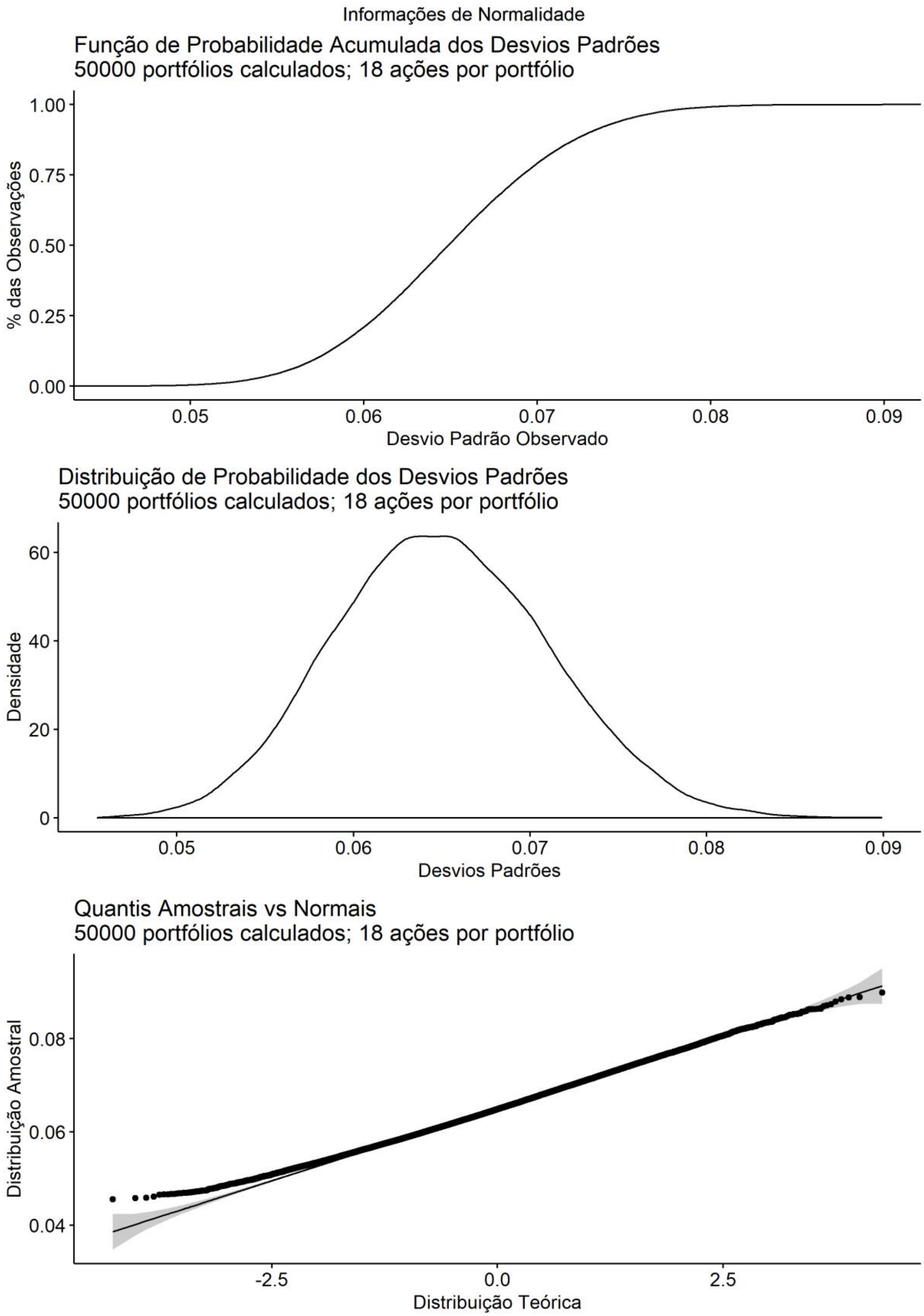
Antes de apresentar os resultados principais, discutiremos a questão de normalidade das amostras apresentada no capítulo anterior. As figuras 1 e 2 apresentam a função de probabilidade acumulada, a densidade de probabilidade e gráfico quantil-quantil dos desvios padrões para o caso de dezoito ações por portfólio. Os resultados da figura 1 foram obtidos a partir de uma amostra de 500 portfólios aleatórios; os da figura 2, a partir de 50.000

Figura 1.



Informações de Normalidade
500 portfólios calculados; 18 ações por portfólio
Função de Probabilidade Acumulada dos Desvios Padrões
% das Observações
1.00
0.75
0.50
0.25
0.00
0.05
0.06
0.07
0.08
Desvio Padrão Observado
Distribuição de Probabilidade dos Desvios Padrões
Densidade
60
40
20
0
0.05
0.06
0.07
0.08
Desvios Padrões
Quantis Amostrais vs Normais
Distribuição Amostral
0.09
0.08
0.07
0.06
0.05
-2
0
2
Distribuição Teórica

Figura 2.


Informações de Normalidade
Função de Probabilidade Acumulada dos Desvios Padrões
50000 portfólios calculados; 18 ações por portfólio
% das Observações
1.00
0.75
0.50
0.25
0.00
0.05
0.06
0.07
0.08
0.09
Desvio Padrão Observado
Distribuição de Probabilidade dos Desvios Padrões
50000 portfólios calculados; 18 ações por portfólio
Densidade
60
40
20
0
0.05
0.06
0.07
0.08
0.09
Desvios Padrões
Quantis Amostrais vs Normais
50000 portfólios calculados; 18 ações por portfólio
Distribuição Amostral
0.08
0.06
0.04
-2.5
0.0
2.5
Distribuição Teórica

Para todos os casos calculados, os mesmos gráficos estão disponíveis em https://github.com/ZeVZobaran/ibovPortfoliosSd.

Os resultados apresentados endossam a nossa opção por rejeitar o uso de testes de normalidade clássicos, que para ambas as amostras apresentadas rejeitam a hipótese de normalidade. Além disso, a progressão da figura 1 para a 2, em que pode-se notar que com um aumento do tamanho da amostra a distribuição dos desvios padrões aproxima-se de uma distribuição normal mesmo sem que tenhamos garantia da validade da lei dos grandes números, é promissora para a hipótese de normalidade e, portanto, de validez do uso de amostras para representar a população. Igualmente promissor é o fato de que, apesar das curvas ficaram mais suaves com o aumento da amostra, o espaço que elas ocupam continua similar; não parece que o aumento da amostra transforme desvio padrões que eram *outliers* em observações usuais.

O comportamento aparentemente anormal da cauda inferior das distribuições, especialmente perceptível no gráfico quantil-quantil da figura 2, tampouco é surpreendente. Desvios padrões de portfólios possuem um limite inferior maior do que zero (o que ficará ainda mais claro quando observarmos os gráficos de convergência de desvio padrão, a seguir), e é de se esperar que, no inicio da distribuição, eles sejam mais concentrados em valores ligeiramente mais altos do que uma normal estabeleceria. A diferença, no entanto, é apenas marginal, e o nosso maior interesse está na sessão superior de desvios padrões, que representam os portfólios de alto risco, de forma que essa questão não é problemática.

Com esses resultados, podemos ter confiança de que as curvas de convergência apresentadas nas figuras 3 e 4 são válidas. Apresentamos os resultados para amostras de 500 e 50.000 portfólios respectivamente. Os resultados para outros tamanhos de amostra estão disponíveis em https://github.com/ZeVZobaran/ibovPortfoliosSd.

Figura 3.

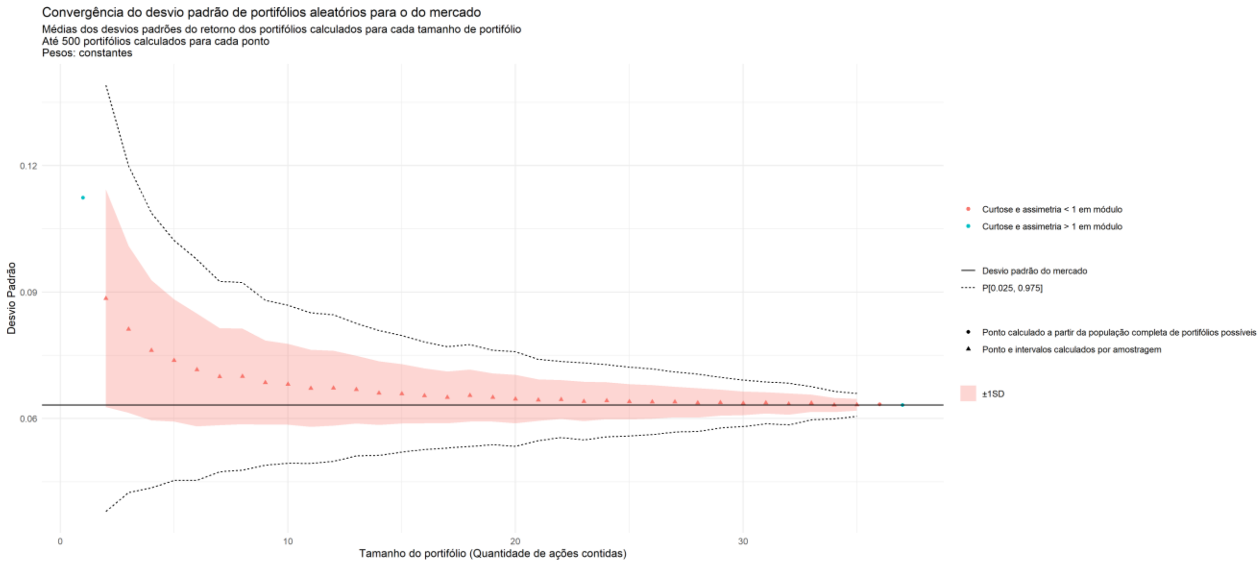
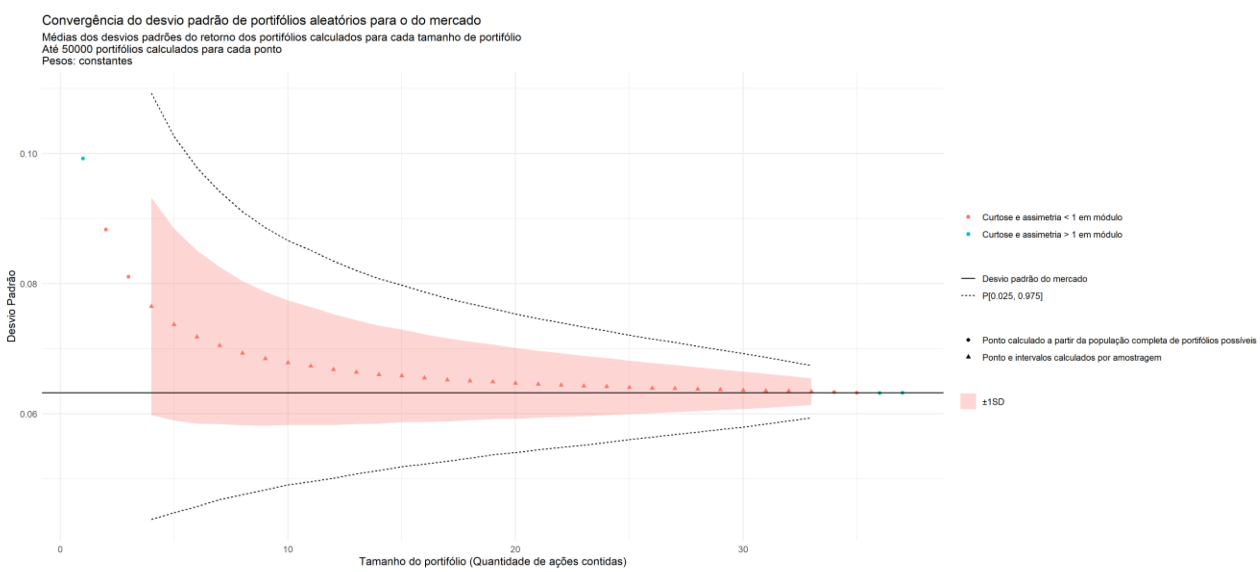


Figura 4.



Mais uma vez, a semelhança entre as duas figuras reforça nossa confiança na validade dos resultados. Em particular, o fato de a mediana do desvio padrão amostral para o caso de duas ações por portifólio (figura 3) estar muito próxima à mediana do desvio padrão populacional para esse mesmo caso (figura 4) é reconfortante. Como antes, o principal efeito do aumento da amostra parece ser o de suavizar a trajetória das curvas, tanto das medianas quanto das bandas de +- um desvio padrão e de 95% dos resultados obtidos. Observamos também que o nosso teste de normalidade adaptado é satisfeito em quase todos os pontos, com apenas uma exceção importante na figura 4 (o penúltimo). Não normalidade no caso de uma ação por portfólio não é importante, visto que são poucas as observações possíveis (apenas uma para cada ação; 37 no universo utilizado), e é esperado no caso do portfólio de mercado, pois como só há uma observação possível não se pode definir uma distribuição.

As medianas dos desvios padrões dos retornos de cada tamanho de portfólio possível descrevem uma curva monotonicamente decrescente na qual os ganhos marginais até aproximadamente cinco ações são significativos. A partir desse ponto, e apesar de ainda se haver ganhos absolutos importantes por realizar, não é mais visualmente claro que um aumento de portfólio seja justificado na margem. Essa discussão será feita em maior profundidade no capítulo 5, quando compararemos ganhos com custos por diversificação, e utilizaremos métodos estatísticos para definir a existência de ganhos marginais e absolutos, e se a extinção dos primeiros justifica que se ignore os segundos. Essa discussão é sugerida por Elton e Grueber (1977), mas eles pouco se aprofundam no tema. Interessantemente, o desvio padrão amostral mediano converge para o de mercado antes de se atingir o portifólio de mercado pleno. Isso pode se dever ao tamanho limitado do universo de ações de que estamos tratando, mas pode indicar também que o retorno histórico de certos ativos é linearmente dependente dos retornos do resto do mercado, o que justificaria o comportamento visto e indicaria alto grau de correlação entre certas empresas ou grupos de empresas na bolsa brasileira.

O comportamento das bandas superiores é similar ao das medianas e em linha com o esperado. As bandas inferiores, no entanto, apresentam resultados interessantes. O desvio padrão dos retornos dos portfólios cuja volatilidade (usados como sinônimos) é um desvio padrão menor do que a volatilidade mediana (apresentados como o limiar inferior da linha vermelha) é próximo de 0.06, um pouco menor do que a de mercado, independente ao tamanho de portfólio selecionado. A consistência desse resultado demonstra a dificuldade de se construir um portfólio aleatório com volatilidade realmente baixa. Uma interpretação possível para a aderência desse nível é que ele seja o resultado da soma do desvio padrão de mercado (o risco sistêmico) e a média dos riscos idiossincráticos de todas as ações disponíveis. Testar essa hipótese requereria o cálculo desse risco para cada ação, no entanto.

O outro limiar inferior, que marca o ponto em que apenas 2.5% dos desvios padrões estão abaixo da linha, se comporta de forma diferente, e mais parecida com o comportamento da mediana, embora espelhada. Se o resultado anterior denota a dificuldade da construção de portfólios pouco diversificados e de baixo risco, este aponta para a sua possibilidade. Uma gestora excepcional, ou excepcionalmente sortuda, poderia selecionar um portfólio nesta região, obtendo resultados superiores ao mercado mesmo sem uma carteira extensa. Ressaltamos que como as únicas restrições à formação de portfólios neste exercício foram o seu tamanho e a distribuição de pesos entre ativos (que deve ser constante), é possível que esses portfólios "excepcionais" apresentem retornos baixos.