

# COMPARAÇÃO DE PERFORMANCE DE TESTES DE DISPERSÃO PARA VARIÁVEL NORMAL VIA ESTATÍSTICAS R, S E |S| USADAS NO MONITORAMENTO DE PROCESSOS



Ariane Meneguetti (nani.meneguetti@gmail.com) , Emanuel Pimentel Barbosa (emanuel@ime.unicamp.br)

IMECC - UNICAMP

CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico  
Palavras-chave: Monitoramento de Variabilidade - Amplitude Relativa – Variância Generalizada - Gráficos de Controle - Cornish-Fisher



## Introdução

Neste projeto, buscamos monitorar uma ou mais características de qualidade de um processo quanto a estabilidade de seu parâmetro de dispersão ao longo do tempo. Para tal, se utilizam de estatísticas para gerar gráficos de controle acerca desses parâmetros, visando a detecção de anomalias no processo.

São comparadas as performances de detecção desses gráficos de controle para os casos de uma ou múltiplas características de qualidade, através da Amplitude Amostral R, Desvio Padrão Amostral S e Variância Generalizada Amostral |S|, fazendo uso de medidas de performance como o risco  $\alpha$  de falso alarme e o risco  $\beta$  de não detecção.

## Metodologia

### • Estatística S

Os limites de controle (LC) aproximados, a expressão do risco  $\alpha$  e do risco  $\beta$  exato da estatística S estão mostrados abaixo, onde  $c_4$  é o fator de correção

$$LIC = (c_4 - 3\sqrt{1-c_4^2})\sigma \quad e \quad LSC = (c_4 + 3\sqrt{1-c_4^2})\sigma$$

$$risco \alpha = 1 - \left[ F_{\chi_{n-1}^2}((n-1)(c_4 + 3\sqrt{1-c_4^2})^2) - F_{\chi_{n-1}^2}((n-1)(c_4 - 3\sqrt{1-c_4^2})^2) \right]$$

$$risco \beta = F_{\chi_{n-1}^2} \left( \frac{1}{k^2} \chi_{1-\frac{\alpha}{2}; n-1}^2 \right) - F_{\chi_{n-1}^2} \left( \frac{1}{k^2} \chi_{\frac{\alpha}{2}; n-1}^2 \right)$$

### • Estatística R

Utilizando momentos de  $W=R/\sigma$  para caracterizar os LC da estatística R aproximada normal, onde  $E(W) = d_2$  e  $d.p.(W) = d_3$ , temos os LC 3-sigma para R:

$$LIC = (d_2 - 3d_3)\sigma \quad e \quad LSC = (d_2 + 3d_3)\sigma$$

Usando a relação entre a amplitude relativa e a amplitude estudentizada de Tukey, isso nos permite obter valores para a distribuição de W através de softwares ou tabelas da distribuição de Tukey. Assim, o risco  $\alpha$  e o risco  $\beta$  exato são dados por:

$$risco \alpha = 1 - [F_W(d_2 + 3d_3) - F_W(d_2 - 3d_3)] \quad risco \beta = F_W \left( \frac{1}{k} W_{1-\frac{\alpha}{2}} \right) - F_W \left( \frac{1}{k} W_{\frac{\alpha}{2}} \right)$$

Para a comparação de performance de testes de dispersão das estatísticas R e S, o teste de hipóteses utilizado é:

$$\begin{cases} H_0: \sigma = \sigma_0 \\ H_1: \sigma = \sigma_1 = k\sigma_0 \end{cases}$$

onde  $\sigma_0$  é o desvio padrão de referência, k é o fator de expansão.

### • Estatística |S|

Fazendo uso dos momentos da estatística |S|, estes foram utilizados para calcular os LC aproximados normal e também os limites com correção de não normalidade (correção quantílica de Cornish-Fisher), temos

$$\text{Versão aproximada normal:} \quad LIC = |\Sigma| (b_1 - 3\sqrt{b_2}) \quad ; \quad LSC = |\Sigma| (b_1 + 3\sqrt{b_2})$$

$$\text{para } b_1 = \frac{1}{(n-1)^p} \prod_{i=1}^p (n-i) \quad ; \quad b_2 = b_1 \left[ \prod_{i=1}^p \frac{(n-i+2)}{(n-1)^p} - b_1 \right]$$

$$risco \alpha = 1 - \left[ F_{\chi_{2n-4}^2}(2(n-1)\sqrt{b_1 + 3b_2^{1/2}}) - F_{\chi_{2n-4}^2}(2(n-1)\sqrt{b_1 - 3b_2^{1/2}}) \right]$$

Quando  $p > 2$ , o risco  $\alpha$  é obtido numericamente, por meio de simulação do |S|.

Método de Cornish-Fisher:

$$q_{|S|^*}(p) \approx q_Z(p) + K_3 \frac{q_Z^2(p) - 1}{6} + K_4 \frac{q_Z^3(p) - 3q_Z(p)}{24} - K_5 \frac{2q_Z^3(p) - 5q_Z(p)}{36}$$

$$\text{onde } |S|^* = \frac{|S| - \mu}{\sigma} \quad , \quad \begin{cases} \mu = \mathbb{E}(|S|) \\ \sigma = \sqrt{Var(|S|)} \end{cases} \quad e \quad Z \equiv N(0, 1)$$

E então, calculamos quantis para o |S| padronizados  $q_{|S|^*}(p)$  e fizemos a transformação inversa  $q_{|S|}(p)$ ; obtendo assim o risco  $\alpha$  sob a correção quantílica de Cornish-Fisher.

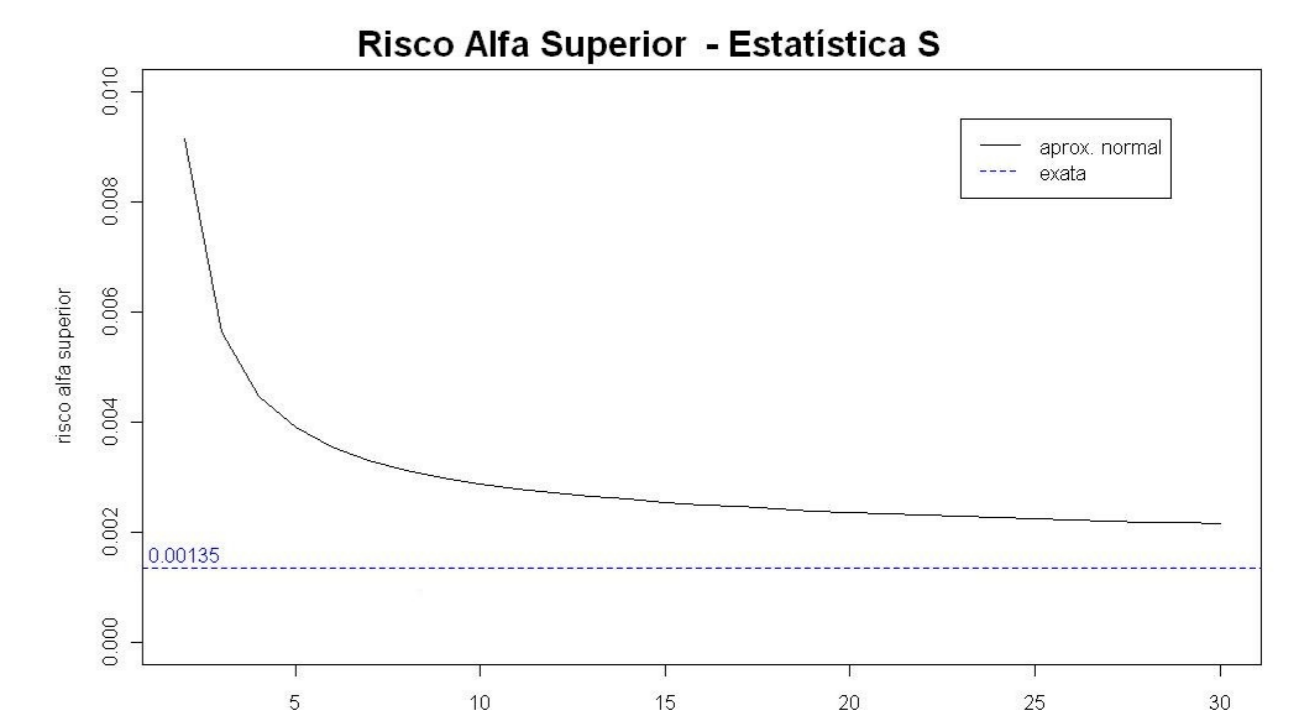
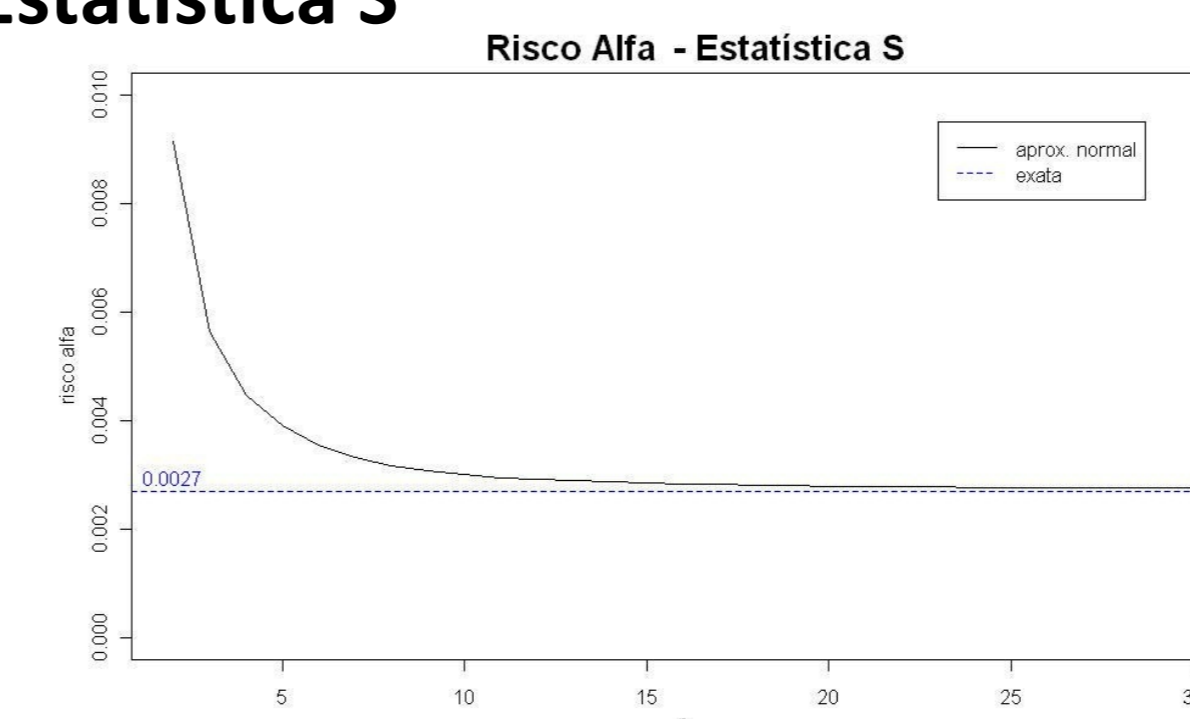
$$q_{|S|}(p) = |\Sigma| \left[ b_1 + \sqrt{b_2} q_{|S|^*}(p) \right]$$

$$risco \alpha = 1 - \left[ F_{\chi_{2n-4}^2}(2(n-1)\sqrt{CF_2}) - F_{\chi_{2n-4}^2}(2(n-1)\sqrt{CF_1}) \right]$$

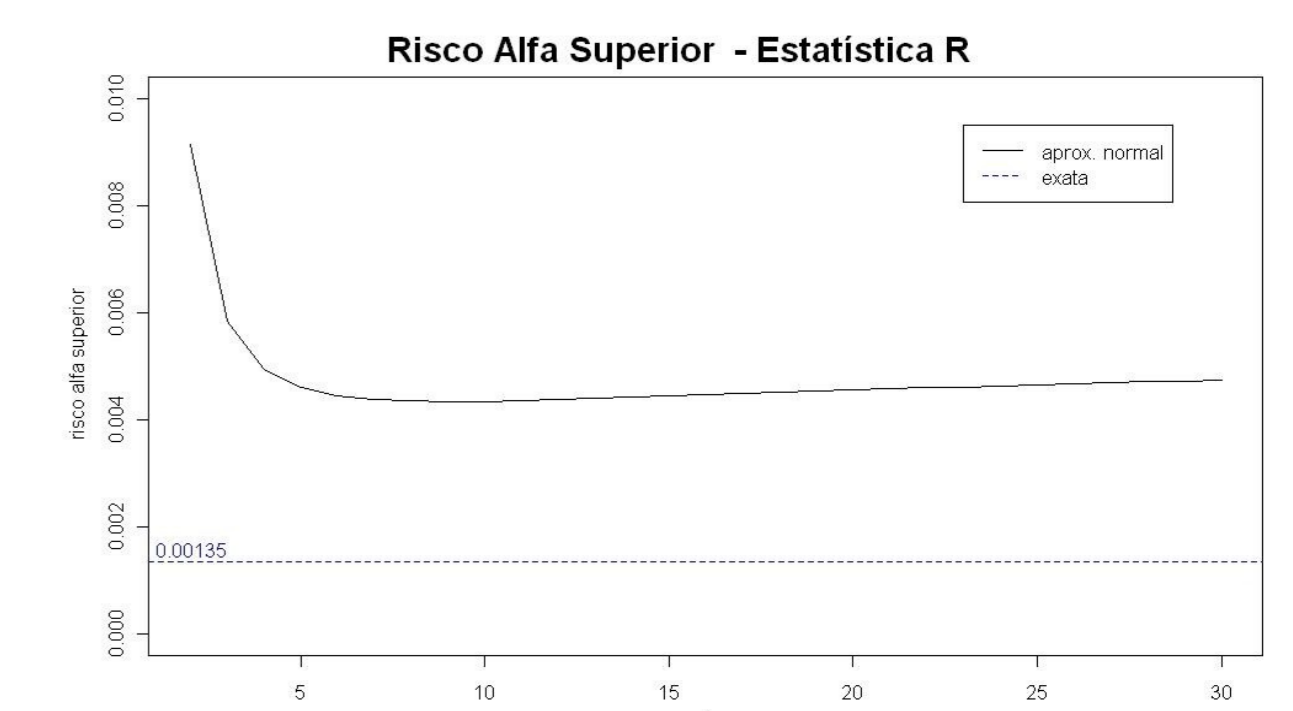
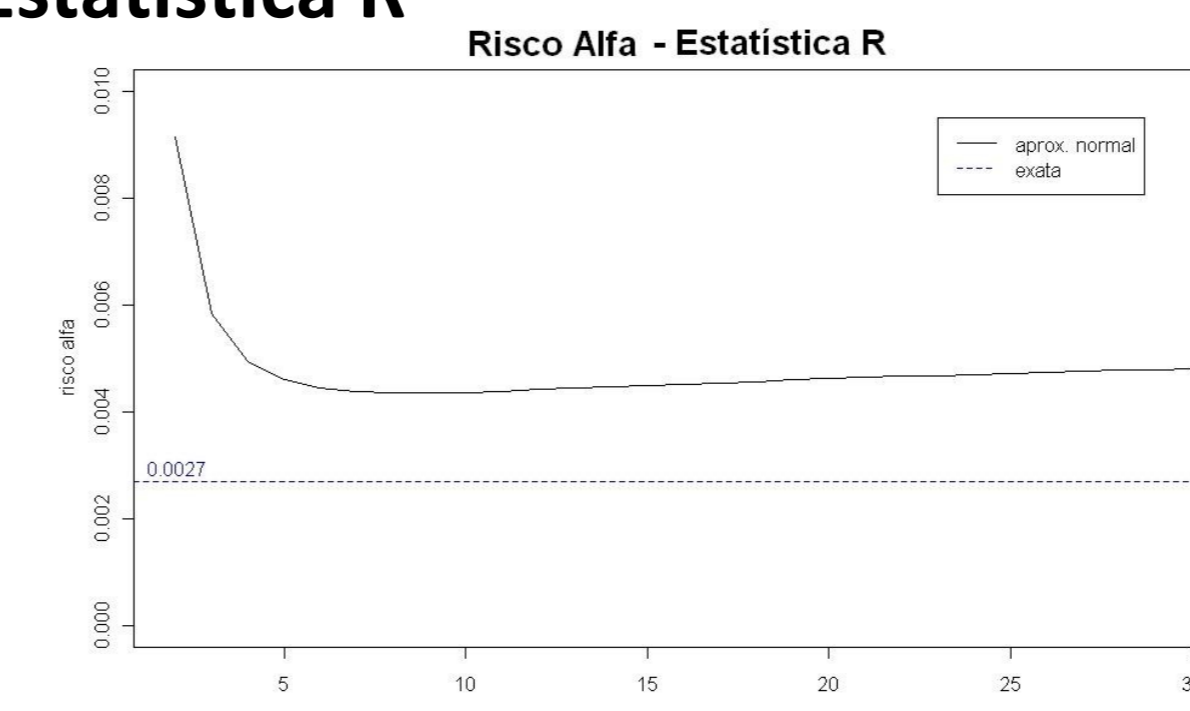
As comparações foram feitas da seguinte forma: |S| exata versus aproximada normal versus com correção quantílica de Cornish-Fisher (caso  $p=2$  e 3).

## Resultados

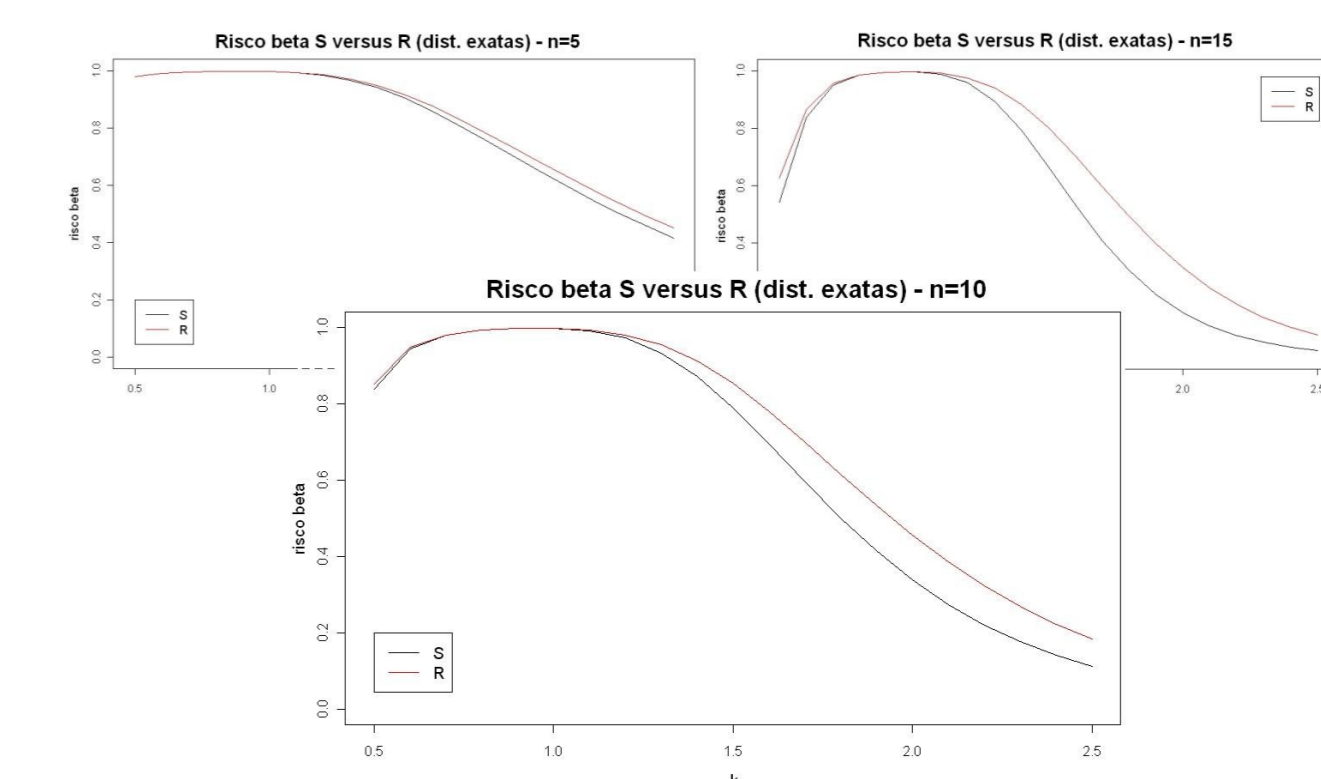
### • Estatística S



### • Estatística R

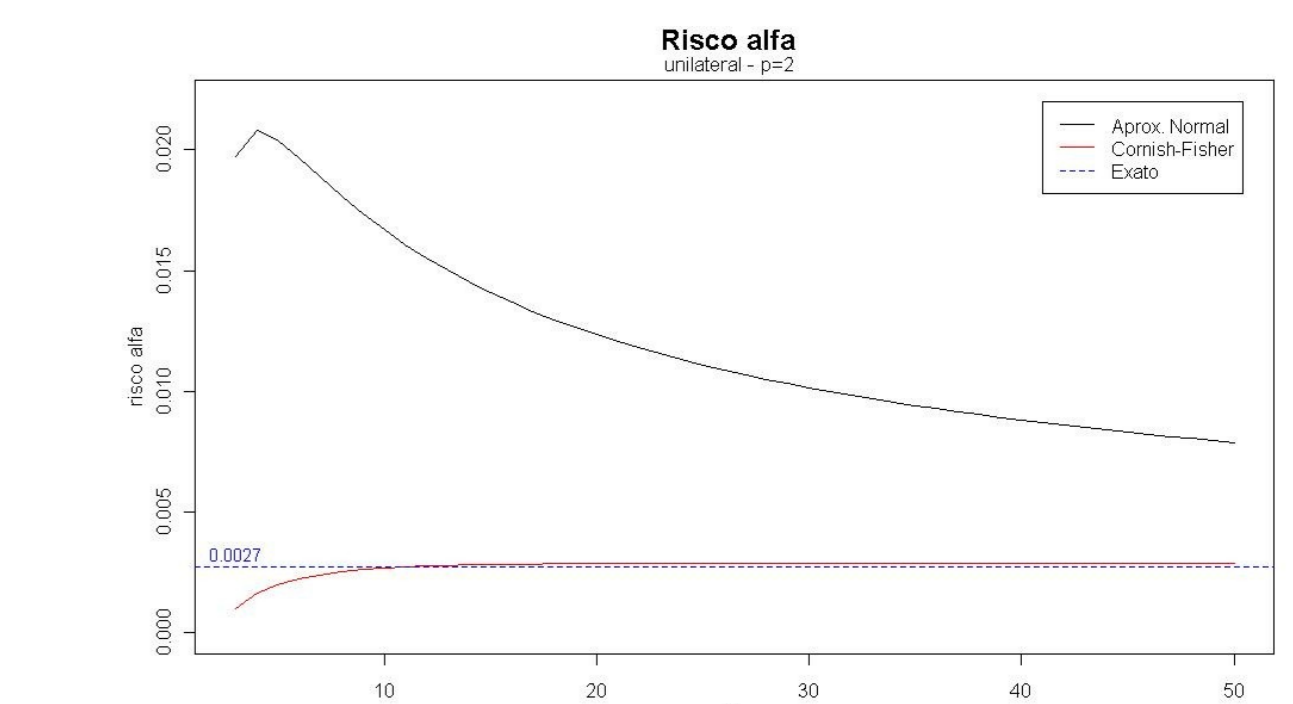
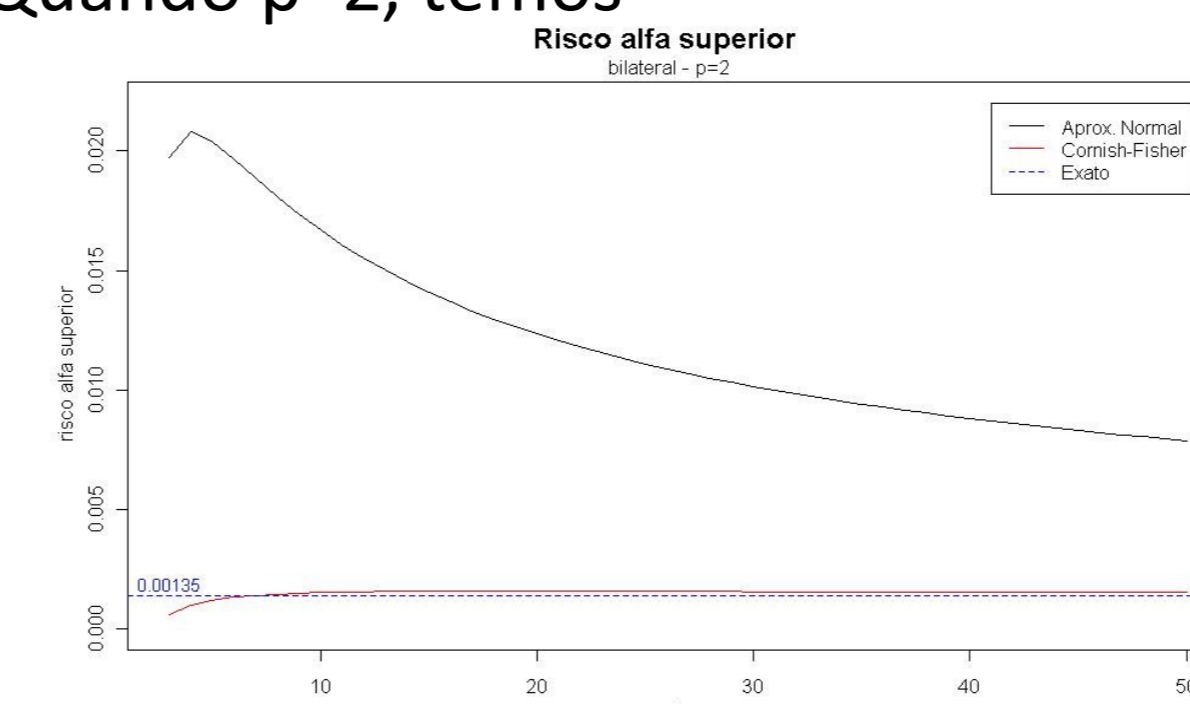


### • S versus R



### • Estatística |S|

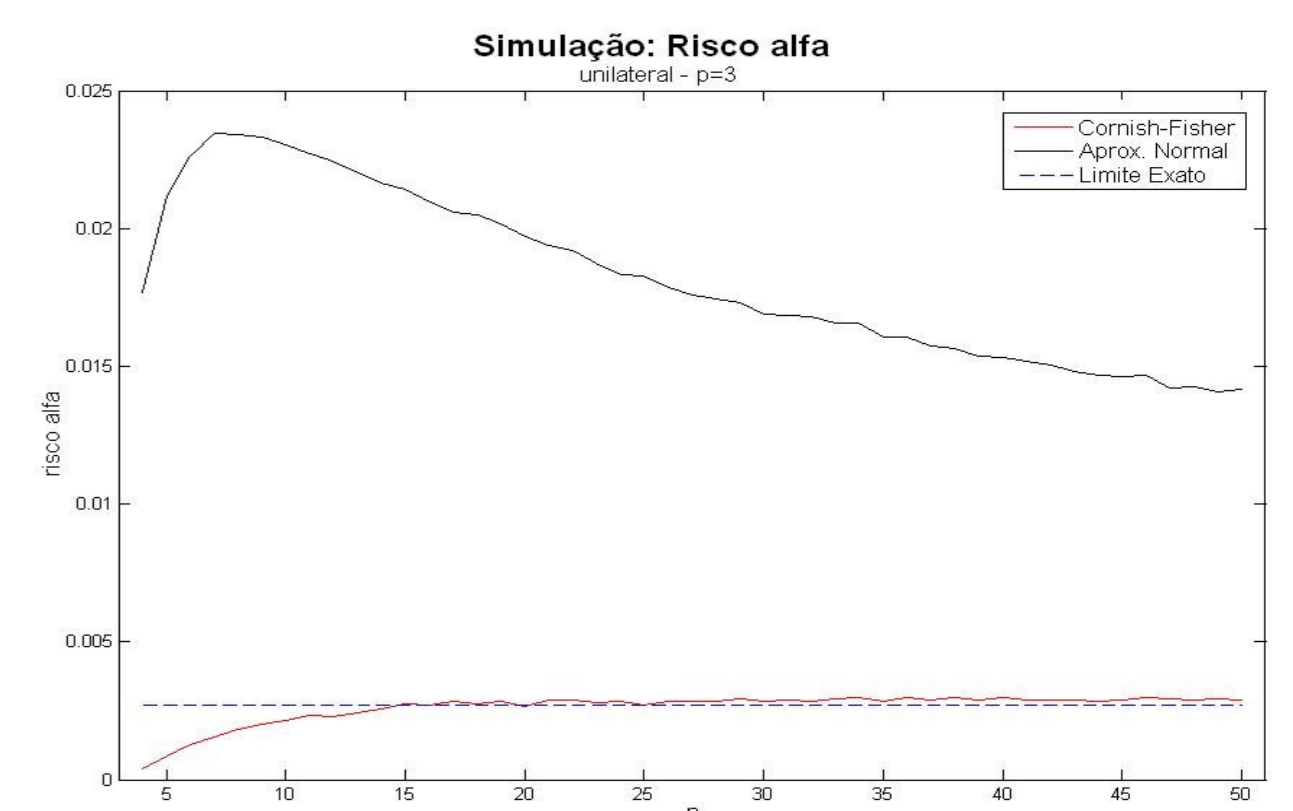
Quando  $p=2$ , temos



E quando  $p=3$ ,

Tabela 1: Quantis (unilateral)

n	Exato(sim.)	Ap. Normal	CF
4	2.68949	1.03849	5.50433
5	2.89698	1.23081	4.28237
6	2.82795	1.29575	3.54973
7	2.68866	1.31462	3.11760
8	2.53757	1.31384	2.81575
9	2.42403	1.30405	2.59090
10	2.29517	1.29008	2.41570
15	1.89068	1.21041	1.90241
20	1.64740	1.14353	1.64311
30	1.38473	1.04938	1.37043



## Conclusões

A estatística S tem uma aproximação normal aceitável para amostras não muito pequenas (digamos  $n > 5$ ). No caso da estatística R, sua aprox. normal é inaceitável (não converge para a normal) e é expressivo o aumento da prob. de falso alarme.

Ao compararmos S versus R no caso exato, é aceitável a troca do uso de uma estatística pela outra quando  $n < 10$ .

A estatística |S| (casos  $p=2$  e 3) sob aproximação normal traz um aumento significativo da prob. de falso alarme, gerando enorme aumento desnecessário de custo e/ou descrédito do sistema de monitoramento e controle.

O método de Cornish-Fisher tem excelente aprox. para a verdadeira distribuição do |S|, resolvendo significativamente, e de modo simples, o problema do falso alarme, que até então só se tinha uma solução simples para  $p=2$ . Pelo fato da distribuição exata do |S| quando  $p \geq 3$  não ter uma forma muito simples, pode-se utilizar a aprox. quantílica de Cornish-Fisher proposta neste trabalho.