

**UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE CIÊNCIAS EXATAS
DEPARTAMENTO DE ESTATÍSTICA
CURSO DE ESTATÍSTICA**

**LUCIANA JUSTINA DA SILVA
RAFAEL CORDEIRO DA ROCHA**

**DELINEAMENTO AMOSTRAL POR CONJUNTOS ORDENADOS
APLICADO AO CONTROLE ESTATÍSTICO DE QUALIDADE**

**CURITIBA
2013**

**LUCIANA JUSTINA DA SILVA
RAFAEL CORDEIRO DA ROCHA**

**DELINEAMENTO AMOSTRAL POR CONJUNTOS ORDENADOS
APLICADO AO CONTROLE ESTATÍSTICO DE QUALIDADE**

Trabalho de conclusão de curso apresentado à disciplina de Laboratório de Estatística, Curso de Graduação em Estatística, do setor de Ciências Exatas da Universidade Federal do Paraná.

Orientador: Cesar Augusto Taconeli.

**CURITIBA
2013**

AGRADECIMENTOS

À Deus primeiramente, por ter nos concedido vida, oportunidade e capacidade para passar por mais esta etapa, sem ele isto não seria possível.

Aos familiares, pelo apoio às nossas escolhas, pelo incentivo em todo período de atividades acadêmicas e pela compreensão nos tempos de ausência.

Aos amigos, pois sempre estiveram torcendo pelo nosso sucesso e ansiosos por esta conquista.

Ao Professor Doutor Cesar Augusto Taconeli, pela disposição e competência ao compartilhar seus conhecimentos e, principalmente, pela maneira brilhante com que conduziu nossos estudos.

Ao Professor Mestre Adilson dos Anjos, pela disponibilidade em aceitar o convite para participar da banca deste trabalho.

RESUMO

Diversos estudos atestam maior eficiência de estimadores baseados em amostragem por conjuntos ordenados (ACO) em relação aos similares via amostragem aleatória simples (AAS). A amostragem por conjuntos ordenados, introduzida por McIntyre (1952), baseia-se na possibilidade de fazer uma ordenação eficiente dos elementos da amostra, antes de sua efetiva mensuração. Na área de controle de qualidade, Muttalak e Al-Sabah (2003) propuseram um série de métodos utilizando este delineamento, tais como *ranked set sampling* (RSS), *median ranked set sampling* (MRSS) e *extreme ranked set sampling* (ERSS) que também são mais eficientes do que AAS. Nesse sentido, no presente trabalho mostramos aplicações e estudamos o desempenho de um estudo por simulação de dados relacionados à qualidade de fibra têxtil, adaptados dos dados originais apresentados em Montgomery (2001), envolvendo variáveis de diâmetro e resistência. A partir destes dados, foram efetuadas simulações com diferentes cenários que foram definidos por diferentes tamanhos de amostras, graus de correlação entre as variáveis e graus de descontrole do processo. E por fim, foram elaborados gráficos de controle, curvas características de operação e comprimentos médios de sequência através de três métodos via ACO para serem comparados com AAS, concluindo que os métodos via ACO são mais eficientes do que via AAS, pois conseguem fazer detecção de processo fora de controle mais rapidamente em praticamente todos os cenários simulados.

Palavras-chave: Amostragem por conjuntos ordenados (ACO), curvas características de operação, comprimento médio de sequência, controle estatístico da qualidade, cartas controle.

LISTA DE FIGURAS

Figura 1 – Representação da seleção de uma amostra por conjuntos ordenados.....	4
Figura 2 – Representação do Gráfico de Controle	8
Figura 3 – Representação da Curva Característica de Operação.....	10
Figura 4 – Resultados da simulação para $n = 5, \rho = 0,4, \delta = 0$	16
Figura 5 – Resultados da simulação para $n = 10, \rho = 0,4, \delta = 0$	17
Figura 6 – Resultados da simulação para $n = 5, \rho = 0,8, \delta = 0$	18
Figura 7 – Resultados da simulação para $n = 10, \rho = 0,8, \delta = 0$	19
Figura 8 – Resultados da simulação para $n = 5, \rho = 0,4, \delta = 0,5\sigma$	20
Figura 9 – Resultados da simulação para $n = 10, \rho = 0,4, \delta = 0,5\sigma$	21
Figura 10 – Resultados da simulação para $n = 5, \rho = 0,8, \delta = 0,5\sigma$	22
Figura 11 – Resultados da simulação para $n = 10, \rho = 0,8, \delta = 0,5\sigma$	23
Figura 12 – Resultados da simulação para $n = 5, \rho = 0,4, \delta = 0,8\sigma$	24
Figura 13 – Resultados da simulação para $n = 10, \rho = 0,4, \delta = 0,8\sigma$	25
Figura 14 – Resultados da simulação para $n = 5, \rho = 0,8, \delta = 0,8\sigma$	26
Figura 15 – Resultados da simulação para $n = 10, \rho = 0,8, \delta = 0,8\sigma$	27
Figura 16 – Curva Característica de Operação com $\rho = 0,4$	28
Figura 17 – Curva Característica de Operação com $\rho = 0,8$	30
Figura 18 – Resultados do Comprimento Médio de Sequência com $\rho = 0,4$	31
Figura 19 – Resultados do Comprimento Médio de Sequência com $\rho = 0,8$	33
Figura 20 – Dados obtidos através de Montgomery (2003).	39

SUMÁRIO

1 INTRODUÇÃO	1
2 REVISÃO DE LITERATURA	3
2.1 DELINEAMENTO AMOSTRAL POR CONJUNTOS ORDENADOS.....	3
2.1.1 Critérios e desempenhos de Ordenação	5
2.2 ESTIMAÇÃO DE PARÂMETROS VIA <i>RANKED SET SAMPLING (RSS)</i>	5
2.3 CONTROLE ESTATÍSTICO DA QUALIDADE.....	7
2.3.1 Gráficos de Controle.....	7
2.3.2 Curva Característica de Operação	9
2.3.3 Comprimento Médio de Sequência.....	10
3 MATERIAIS E MÉTODOS	11
4 RESULTADOS	16
5 CONCLUSÃO	34
REFERÊNCIAS	36
ANEXOS	38
ANEXO A: DADOS	39
ANEXO B: FUNÇÕES E PROGRAMAÇÃO	40
ANEXO C: ELABORAÇÃO DA AMOSTRA	43

1 INTRODUÇÃO

Qualidade, na definição tradicional significa, adequação ao uso. Já na definição moderna, qualidade é inversamente proporcional à variabilidade, ou seja, melhoria da qualidade é a redução da variabilidade nos processos e produtos. O Controle Estatístico do Processo (CEP) contempla um conjunto de ferramentas para a resolução de problemas visando a melhoria e eficiência deste (MONTGOMERY, 2001). Trata-se de um método ou um conjunto de métodos preventivos, baseados na comparação contínua dos resultados de um processo com os padrões/especificações, buscando identificar, a partir de resultados amostrais, as tendências ou variações significativas, a fim de eliminar/controlar tais variações, com o objetivo de reduzi-las cada vez mais.

Delineamento estatístico é o processo de planejar e conduzir um levantamento ou um experimento, incluindo sua implantação, de modo que seja possível recolher dados que possam ser analisados, usando as metodologias estatísticas apropriadas, e que conduzam a conclusões válidas e objetivas.

Para conduzir a aplicação de um Controle Estatístico de Qualidade é essencial a utilização de um delineamento amostral eficaz, com intuito de redução de custos e perdas, melhoria de resultados e potencialização dos ganhos. Para isso, iremos considerar o delineamento por conjuntos ordenados aplicado ao controle estatístico da qualidade, com o objetivo de verificar quais os ganhos reais decorrentes de sua aplicação.

A amostragem por conjuntos ordenados foi introduzida por McIntyre (1952), na estimação da produção média de pastagens. Mais recentemente, diversos aspectos da amostragem em conjuntos ordenados (Ranked Set Sampling - RSS) têm sido estudados por Stokes (1995), Barnett e Moore (1997), Kaur *et al.* (1997), Barnett (1999), Barreto (2001), dentre outros, na busca de estimadores mais eficientes para diferentes parâmetros populacionais.

Neste sentido, o presente trabalho tem como objetivo fazer uma revisão da literatura, elaborar codificação no software R (R: A Language and Environment for Statistical Computing, 2013) para a construção de gráficos de controle baseados em delineamento por conjuntos ordenados, simulação de dados necessários para elaboração das cartas de controle e avaliação de desempenho através de

comparativos baseados em gráficos e/ou resultados do delineamento em relação ao delineamento usual, baseado em amostras aleatórias simples.

2 REVISÃO DE LITERATURA

2.1 DELINEAMENTO AMOSTRAL POR CONJUNTOS ORDENADOS

A busca por delineamentos amostrais mais eficientes têm sido estudada, dando-se grande ênfase a situações em que a obtenção de dados pode envolver altos custos de mensuração ou mesmo restrições ao acesso de resultados (Barnett, 1999; Barreto, 2001). Neste contexto, a amostragem por conjuntos ordenados, introduzida por McIntyre (1952), baseia-se na possibilidade de fazer uma ordenação eficiente dos elementos da amostra, antes de sua efetiva mensuração.

O delineamento de amostragem por conjuntos ordenados é definido da seguinte forma:

1. Seleção de n amostras aleatórias simples com k elementos cada;
2. Ordenação de cada amostra através de julgamento pessoal ou profissional, ou por alguma variável concomitante ou por outro método não dispendioso, sem que de fato tenham sido observados os valores da variável de interesse;
3. Mensuração de um elemento de cada amostra quanto à variável de interesse: mede-se aquele que foi julgado como sendo o menor valor da variável de interesse na primeira amostra; na amostra subsequente mede-se a unidade amostral julgado com a segunda menor medida e assim sucessivamente, até a n -ésima amostra, em que a observação julgada com a n -ésima menor, ou seja, a maior é de fato mensurada;
4. Repetição deste procedimento quantas m vezes forem desejáveis.

Na Figura 1, temos a ilustração do esquema de seleção de uma amostra por conjuntos ordenados de tamanho $n=3$, com três replicações ($m=3$). Supõe-se que a variável de interesse é a quantidade de líquido dentro da garrafa e o critério utilizado na sua ordenação é classificar o conjunto de garrafas em relação ao nível de líquido avaliado por inspeção visual. As garrafas foram selecionadas ao acaso e separadas de forma aleatória em três conjuntos com três garrafas cada. Desde o primeiro conjunto ordenado, seleciona-se a garrafa que julga-se ter o menor conteúdo vazio entre as três garrafas. A partir do segundo conjunto de três garrafas,

seleciona-se a garrafa que julga-se ter a segunda menor parte vazia. Finalmente, no último conjunto seleciona-se a garrafa com a maior parte vazia. Repete-se este ciclo por três vezes.

Cada uma das amostras inicialmente selecionadas, encontra-se disposta na horizontal, estando assinaladas as unidades amostrais que irão compor a amostra final, sendo as únicas de fato mensuradas quanto à variável de interesse.

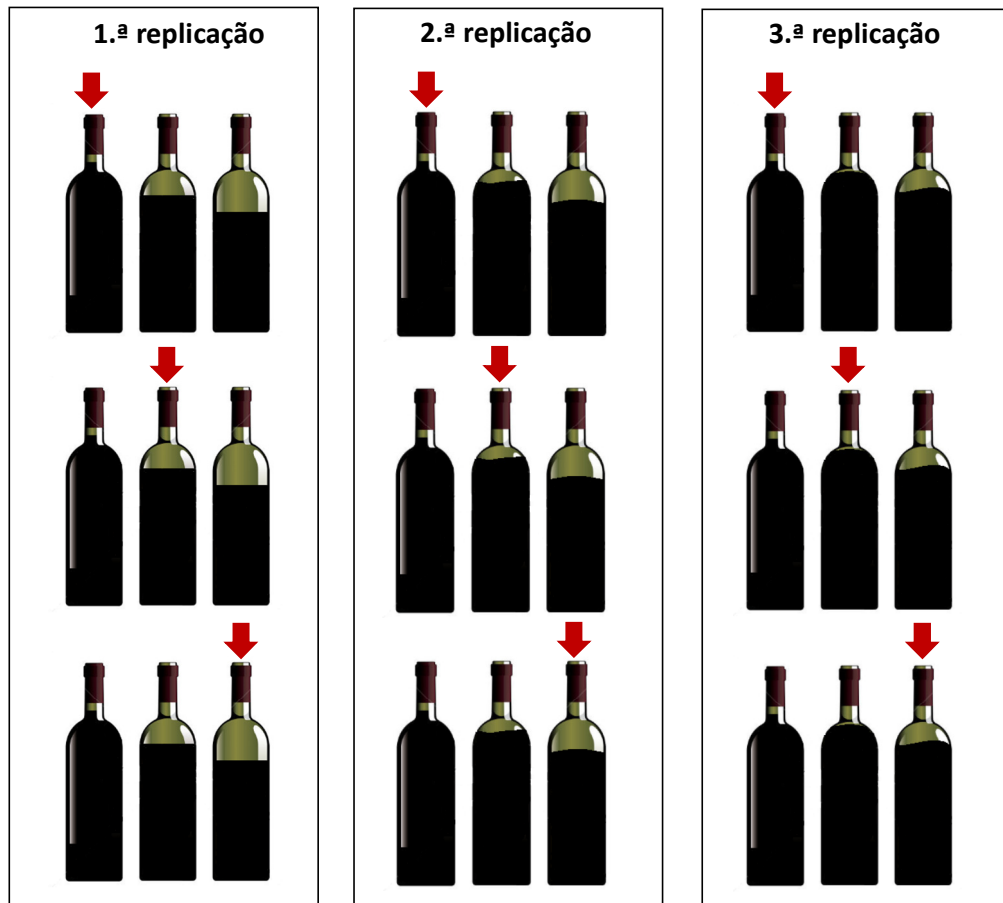


Figura 1 – Representação da seleção de uma amostra por conjuntos ordenados
Fonte: Os autores (2013)

Como exemplo de ordenação por variável concomitante, temos em Taconeli (2005), como ilustração, o estudo das alturas de árvores em uma região densa, o qual pode dificultar e limitar o tamanho amostral, devido à dificuldade em se alcançar o topo das árvores e à impossibilidade de inferir os tamanhos a partir de suas sombras. Pode-se, no entanto, ordenar de maneira simples e eficiente amostras de árvores segundo suas alturas com base no diâmetro de seus troncos à altura do peito, caso o diâmetro (variável concomitante), seja fortemente correlacionado com a altura.

2.1.1 Critérios e desempenhos de Ordenação

Conforme Taconeli (2005), o critério de ordenação das amostras pode se basear em uma opinião subjetiva, fornecida pelo próprio pesquisador, ou em alguma variável concomitante, que seja de fácil mensuração e apresente uma forte correlação com a variável de interesse.

Podemos classificar o procedimento de amostras por conjuntos ordenados de duas formas de acordo com o desempenho na ordenação das unidades amostrais:

- ordenação perfeita: ocorre quando o método utilizado não apresenta erros, ou seja, as observações que se julgou de tal ordem de fato são necessariamente.
- ordenação imperfeita: ocorre quando a ordenação das unidades amostrais pode apresentar erros.

2.2 ESTIMAÇÃO DE PARÂMETROS VIA RANKED SET SAMPLING (RSS)

Foram utilizados três métodos para estimação dos parâmetros do processo e determinação dos limites de controle via *RSS* e um método via Amostragem Aleatória Simples (AAS), baseados em diferentes estimadores da média e da variância apresentados na sequência. A seguir, temos a notação utilizada para o delineamento via *RSS*:

$$X_{(i)j}; i=1, 2, \dots, n; j=1, 2, \dots, m$$

onde $X_{(i)j}$ refere-se à observação que foi julgada como estatística de ordem i , no j -ésimo ciclo.

Takahasi & Wakimoto (1968) propuseram como estimador não viciado da média populacional a média da amostra por conjuntos ordenados (ACO), como apresentado em:

$$\bar{\bar{X}} = \sum_{j=1}^m \sum_{i=1}^n \frac{X_{(i)j}}{(mn)}, \quad (2.2.1)$$

sendo $X_{(i)j}$ a estatística que se julga ser a i -ésima estatística de ordem numa amostra por conjuntos ordenados do j -ésimo ciclo ou repetição do processo, m o número de repetições do procedimento e n o tamanho de cada amostra aleatória simples extraída.

Stokes (1980) sugere um estimador assintoticamente não-viciado para a variância populacional, comprovando maior eficiência deste estimador mesmo na presença de erros de ordenação, conforme apresentando em 2.2.2.

$$\hat{\sigma}^2 = \sum_{j=1}^m \sum_{i=1}^n \frac{(X_{(i)j} - \bar{X})^2}{mn - 1}. \quad (2.2.2)$$

Por outro lado, Yu *et al.* (1999) propuseram uma versão não-viciada de $\hat{\sigma}^2$ na estimação da variância populacional, dada por

$$\hat{\sigma}_{MS}^2 = \left(mn - 1 + \sum_{i=1}^n \frac{\alpha_{(i:n)}^2}{n} \right)^{-1} \sum_{j=1}^m \sum_{i=1}^n (X_{(i)j} - \bar{X})^2 \quad (2.2.3)$$

no qual $\alpha_{(i:n)}$ é a esperança da i -ésima estatística de ordem de uma amostra aleatória simples de tamanho n de uma distribuição normal padronizada. Nota-se que a diferença entre $\hat{\sigma}^2$ e $\hat{\sigma}_{MS}^2$ está no denominador.

Um estimador da variância de \bar{X} foi proposto por Muttlak e Al-Sabah (2003)

$$\hat{\sigma}_{\bar{X}}^2 = \frac{1}{n} \hat{\sigma}^2 - \frac{1}{n^2} \sum_{i=1}^n (\bar{X}_{(i)} - \bar{X})^2, \quad (2.2.4)$$

sendo $\bar{X}_{(i)}$, a média das observações julgadas como i -ésima estatística de ordem ($i=1,2,\dots,n$), ou usando o estimador não-viciado de $\hat{\sigma}$:

$$\hat{\sigma}_{\bar{X}MS}^2 = \frac{1}{n} \hat{\sigma}_{MS}^2 - \frac{1}{n^2} \sum_{i=1}^n (\bar{X}_{(i)} - \bar{X})^2. \quad (2.2.5)$$

O BLUE (*best linear unbiased linear estimator*) via RSS para a média populacional, consiste na ponderação da amostra, atribuindo diferentes pesos às unidades amostrais. Como estimador da média da população utilizando RSS, Sinha *et al* (1996) sugere:

$$\bar{X}_{BLUE} = \sum_{i=1}^n \frac{\bar{X}_{(i)}}{\frac{v_{(i:n)}}{1}}, \quad (2.2.6)$$

no qual $v_{(i:n)}$ é a variância da i -ésima estatística de ordem de uma amostra aleatória simples de tamanho n da distribuição normal padronizada e $\bar{X}_{(i)}$ é a média todas as observações que apareceram na posição i , $i=1, 2, \dots, n$. Neste caso, considera-se o seguinte estimador para $\hat{\sigma}^2$:

$$\hat{\sigma}_{BLUE}^2 = \sum_{j=1}^m \sum_{i=1}^n \frac{(X_{(i)j} - \bar{X}_{BLUE})^2}{mn - 1}, \quad (2.2.7)$$

para o qual os autores apresentam o seguinte estimador para $\sigma_{\bar{X}_{BLUE}}^2$:

$$\hat{\sigma}_{\bar{X}_{BLUE}}^2 = \frac{\sigma^2}{\sum \frac{1}{v_{(i:n)}}}. \quad (2.2.8)$$

2.3 CONTROLE ESTATÍSTICO DA QUALIDADE

No controle estatístico de qualidade existem vários métodos que podem ser utilizados para monitorar um processo. Para o presente trabalho serão utilizados gráficos de controle e curvas características de operação.

2.3.1 Gráficos de Controle

Segundo Montgomery (2001), em qualquer processo de produção uma certa quantidade de variabilidade sempre existirá. Para o controle desta variabilidade temos o gráfico de controle que consiste no acompanhamento de um processo ao longo do tempo. A formação básica deste gráfico é composta de uma linha média no gráfico acompanhada de uma linha superior e uma linha inferior, chamadas de limite superior de controle (LSC) e limite inferior de controle (LIC), respectivamente, e também são dispostos no gráfico a estatística avaliadas em cada amostra por meio de pontos. Os limites são construídos baseados em critérios estatísticos com base na distribuição da estatística usada para monitoramento do processo e tem como objetivo verificar se o processo está sob controle.

Para a detecção de um descontrole no processo, além dos limites superior e inferior controle, temos outras que permitem identificar um processo fora de controle. Para este estudo específico, levaremos em consideração apenas os limites inferiores e superiores. Os limites de controle são escolhidos de modo que, se o

processo está sob controle, praticamente todos os pontos amostrais estarão entre eles. Contando que os pontos estejam entre os limites de controle, o processo é considerado sob controle, não necessitando intervenção no processo. No entanto, um ponto que caia fora dos limites de controle é interpretado como evidência de que o processo está fora de controle e uma possível intervenção se faz necessária. É comum unir os pontos amostrais no gráfico de controle por segmentos de reta, de modo a facilitar a visualização da evolução da sequência de pontos ao longo do tempo. Na Figura 2, temos a representação deste tipo de gráfico.

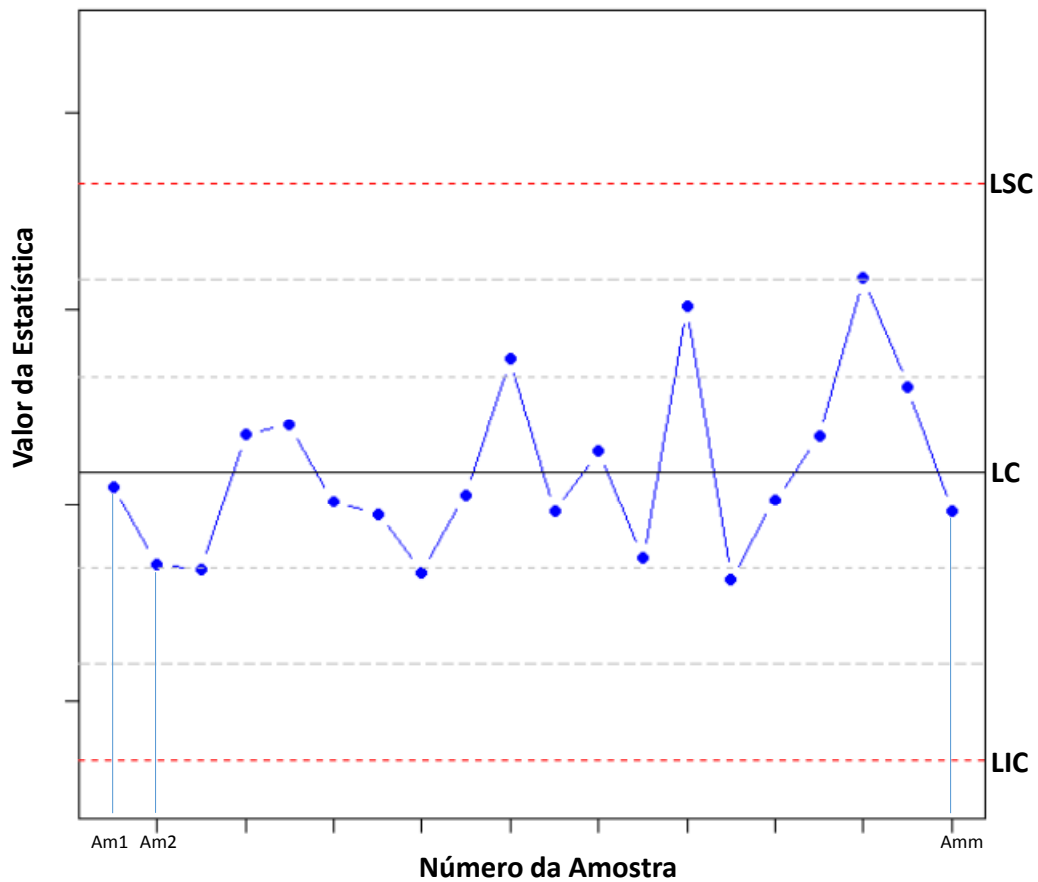


Figura 2 – Representação do Gráfico de Controle
Fonte: Os autores (2013)

Conforme Godoy (2009), o gráfico de controle pode ser construído da seguinte maneira:

- Primeiramente, são realizadas amostragens para estimação dos parâmetros do processo sob controle, medindo-se a característica que se quer controlar em cada amostra;

- Determinam-se, a partir das amostras, estimativas para os parâmetros de interesse;
- A seguir, os limites de controle são calculados. Usualmente, as linhas de controle ficam numa distância de três desvios padrão da estatística usada para monitoramento da média ou alvo do processo.

Como regra geral temos que se um ponto encontra-se fora dos limites de controle, há a indicação de que o processo está fora de controle. Porém existem regras suplementares, que podem ser usadas para aumentar a sensibilidade dos gráficos de controle a uma pequena mudança no processo, de modo a responder mais rapidamente a uma causa atribuível. Como exemplo, temos algumas Regras da WESTERN ELECTRIC COMPANY (1956) que podem concluir se o processo está fora de controle, tais como: um ponto localizado fora da área de controle, dois em três pontos consecutivos se localizados além dos limites de alerta 2-sigma, quatro em cinco pontos consecutivos localizados a uma distância de um sigma ou mais da linha central ou oito pontos consecutivos localizados de um mesmo lado da linha central (NELSON, 1984).

Porém, regras sensibilizantes precisam ser usadas com considerável cuidado, uma vez que o número excessivo de falsos alarmes pode ser danoso para um programa de controle estatístico do processo eficaz, além de tornar o processo de decisão ainda mais complicado, perdendo a simplicidade implícita do gráfico de controle.

2.3.2 Curva Característica de Operação

A curva característica de operação (CCO) consiste na representação da probabilidade de não detecção de um descontrole em função do quanto o processo alterou em termos do parâmetro que está sendo controlado, servindo como base para avaliar a capacidade de detecção do descontrole do processo e também como uma ferramenta de planejamento para se definir o tamanho da amostra. Uma representação geral de uma curva característica de operação é apresentada na Figura 3.

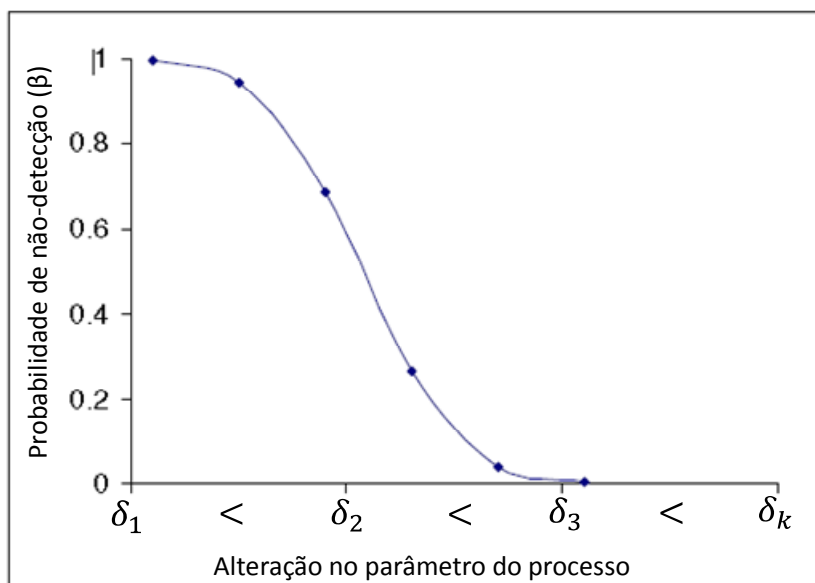


Figura 3 – Representação da Curva Característica de Operação
Fonte: Os autores (2013)

A figura 3 mostra com uma Curva Característica de Operação geralmente se comporta. A CCO ideal se encontra completamente na horizontal, com probabilidade de aceitação $P(a)=1$

2.3.3 Comprimento Médio de Sequência

De acordo com Montgomery (2001), outra maneira de avaliar as decisões relativas a tamanho da amostra e frequência da amostragem é através do comprimento médio da sequência (CMS) do gráfico de controle. Essencialmente, o CMS é o número médio de pontos que devem ser marcados antes que um ponto indique uma condição fora de controle. Se as observações do processo são não-correlacionadas, então, para qualquer gráfico de controle, o CMS pode ser calculado por:

$$CMS = \frac{1}{p} \quad (2.3.3.1)$$

onde p é a probabilidade de que qualquer ponto exceda os limites de controle. Esse coeficiente pode ser usado para avaliar o desempenho do gráfico de controle.

3 MATERIAIS E MÉTODOS

Os dados utilizados neste trabalho foram adaptados de Montgomery (2001) conforme verifica-se no Anexo A, os quais configuram uma situação potencial em que pode ser aplicado o delineamento amostral por conjuntos ordenados. Os dados tratam da qualidade de fibra têxtil sendo fornecidas duas variáveis importantes que devem ser controladas: diâmetro e resistência.

No estudo o interesse é mensurar a resistência da fibra têxtil. Mesmo obtendo o diâmetro da fibra através de aparelhos de medição, para obtenção da informação de resistência faz-se necessário exercer uma determinada força até o seu rompimento e, como consequência temos a destruição da amostra, tornando o custo elevado. Assim, conclui-se que o delineamento amostral por conjuntos ordenados é bastante apropriado para este processo tanto pela eficiência como redução de custo na obtenção das amostras.

Para o desenvolvimento do estudo foram gerados conjuntos de dados com as mesmas características dos dados do Montgomery, ou seja, um conjunto de dados com duas variáveis, diâmetro e resistência, com suas respectivas médias, variância e covariância.

Os cenários de simulação foram definidos com base nos dados originais gerando valores de uma distribuição normal bivariada, com vetor de médias (μ_D, μ_R) e matriz de variâncias e covariâncias $\begin{bmatrix} \sigma_D^2 & \sigma_{RD} \\ \sigma_{DR} & \sigma_R^2 \end{bmatrix}$, nos quais os parâmetros foram estimados com base nos dados simulados, convenientemente, conforme objetivo do estudo.

As estimativas usadas para estes parâmetros foram a média do diâmetro e resistência $\mu_D=1,06$ polegadas e $\mu_R=115,59$ psi, a variância sendo $\sigma_D=1,23$ psi² e $\sigma_R=0,83$ polegadas² e a covariância entre elas de $\sigma_{DR} = 0,79$.

Determinou-se que a variável concomitante seria diâmetro e a variável de interesse a resistência. Foram gerados dados configurando diferentes cenários, referente a diferentes combinações de ρ , a correlação existente entre a variável de interesse e a variável utilizada para ordenação, δ , a alteração na média da variável de interesse em unidades de seu desvio padrão ($\mu' = \mu + \delta\sigma$) e n o tamanho da amostra. Para cada cenário, foram simulados 30 amostras por conjuntos ordenados

para o processo operando sob controle ($\mu' = \mu$), ou seja, para realização do cálculos dos limites foram utilizados somente as 30 amostras iniciais. Após foram simuladas mais 20 amostras adicionais tomando-se o processo operando sob controle (em um dos cenários) e fora de controle ($\mu' \neq \mu$) nos demais.

Para estes diferentes cenários foram definidos os seguintes parâmetros:

- Tamanho da amostra: $n = 5$ e $n = 10$ elementos;
- Grau de correlação das variáveis: $\rho = 0.4$ e $\rho = 0.8$;
- Grau de descontrole do processo: $\delta = 0$; $\delta = 0,5\sigma$ e $\delta = 0,8\sigma$.

Utilizando o conjunto de dados simulados de acordo com os parâmetros citados em MONTGOMERY, foi realizada a extração das amostras de duas maneiras: uma para Amostra Aleatória Simples e outra para Amostra por Conjuntos Ordenados.

Para realizar a seleção da amostra por conjuntos ordenados, o procedimento efetuou-se o seguinte procedimento: considerando uma amostra de tamanho n , foram simulados n elementos, ou seja, n medidas de diâmetro (variável concomitante) com sua respectiva resistência (variável de interesse). Ordenamos estes valores de n pela variável diâmetro do menor valor para o maior, após a ordenação selecionamos aquele com o menor diâmetro, e apanhamos o seu valor de resistência relacionado. Esta resistência é o primeiro elemento da primeira amostra. Em seguida retiramos mais n elementos do conjunto de dados, observando que não há reposição em todo o processo, uma nova ordenação é realizada nestes elementos e após a ordenação selecionamos a segunda menor medida do diâmetro e coletamos a sua resistência relacionada, que corresponderá ao segundo elemento da primeira amostra. Este procedimento é realizado até obter os n elementos necessários para a primeira amostra. Todo este processo foi repetido 50 vezes, até a obtenção das 50 amostras que foram analisadas nos cenários construídos.

Para obtenção da amostra aleatória simples, foi retirado do mesmo conjunto de dados os n elementos necessários para as 50 amostras de forma aleatória e sem reposição.

Na realização desses delineamentos, tanto via ACO como AAS, foi elaborada uma função no software R para geração das amostras, a qual está disponível no Anexo C, sendo utilizada uma semente fixa para elaboração dos resultados.

Com os dados disponíveis, determinou-se que para desenvolvimento dos gráficos de controle utilizar-se-iam os seguintes métodos para determinação dos limites de controle, extraídos de Muttlak e Al-Sabah (2003).

Método 1: Baseado no estimador da média populacional de Takahasi e Wakimoto (1968) e no estimador assintoticamente não-viciado da variância de Stokes (1980), temos:

$$LSC = \bar{X} + 3\hat{\sigma}_{\bar{X}}$$

$$LC = \bar{X}$$

$$LIC = \bar{X} - 3\hat{\sigma}_{\bar{X}}$$

sendo

$$\bar{X} = \sum_{j=1}^m \sum_{i=1}^n \frac{X_{(i)j}}{(mn)},$$

$$\hat{\sigma}_{\bar{X}} = \left[\frac{1}{n} \hat{\sigma}^2 - \frac{1}{n^2} \sum_{i=1}^n (\bar{X}_{(i)} - \bar{X})^2 \right]^{\frac{1}{2}},$$

e

$$\hat{\sigma}^2 = \sum_{j=1}^m \sum_{i=1}^n \frac{(X_{(i)j} - \bar{X})^2}{mn - 1}.$$

Método 2: De acordo com a média populacional e a variância do BLUE via RSS proposto por Sinha *et al* (1996), temos:

$$LSC = \bar{X} + 3\hat{\sigma}_{\bar{X}_{BLUE}}$$

$$LC = \bar{X}$$

$$LIC = \bar{X} - 3\hat{\sigma}_{\bar{X}_{BLUE}}$$

sendo

$$\bar{X}_{BLUE} = \sum_{i=1}^n \frac{\frac{\bar{X}_{(i)}}{v_{(i:n)}}}{\frac{1}{v_{(i:n)}}},$$

onde

$$\hat{\sigma}_{\bar{X}_{BLUE}} = \left(\frac{\hat{\sigma}^2}{\sum \frac{1}{v_{(i:n)}}} \right)^{\frac{1}{2}},$$

e

$$\hat{\sigma}_{BLUE}^2 = \sum_{j=1}^m \sum_{i=1}^n \frac{(X_{(i)j} - \bar{X}_{BLUE})^2}{mn - 1}.$$

Método 3: Baseado no estimador não-viciado para a variância sugerido por Yu *et al.* (1999), temos:

$$LSC = \bar{X} + 3\hat{\sigma}_{\bar{X}_{MS}}$$

$$LC = \bar{X}$$

$$LIC = \bar{X} - 3\hat{\sigma}_{\bar{X}_{MS}}$$

onde

$$\hat{\sigma}_{\bar{X}_{MS}}^2 = \frac{1}{n} \hat{\sigma}_{MS}^2 - \frac{1}{n^2} \sum_{i=1}^n (\bar{X}_{(i)} - \bar{X})^2,$$

e

$$\hat{\sigma}_{MS}^2 = \left(mn - 1 + \sum_{i=1}^n \frac{\alpha_{(i:n)}^2}{n} \right)^{-1}.$$

Método 4: Conforme Montgomery (2001), em relação a amostragem aleatória simples, temos:

$$LSC = \bar{X} + \frac{3\bar{S}}{c_4\sqrt{n}}$$

$$LC = \bar{X}$$

$$LIC = \bar{X} - \frac{3\bar{S}}{c_4\sqrt{n}}$$

onde

$$\bar{X} = \sum_{i=1}^n \frac{X_i}{n},$$

$$S_j = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}}; \text{ para } j = 1, \dots, m$$

$$\bar{S} = \frac{1}{m} \sum_{j=1}^m S_j \text{ e}$$

$$c_4 \cong \frac{4(n-1)}{4n-3}.$$

Todos os métodos foram utilizados para comparações quanto ao desempenho na detecção de descontrole e na não produção de falsos alarmes. Estes resultados serão vistos através das simulações e por meio dos gráficos de controle, curvas características de operação e comprimento médio de sequência.

Em todos os cenários simulados as amostras via AAS e ACO são diferentes, pois os delineamentos em si são diferentes, porém os cenários simulados são os mesmos, ou seja, possuem a mesma variabilidade, mesmo tamanho de amostra e mesmo deslocamento na média, para serem utilizados com fins comparativos.

4 RESULTADOS

Na sequência são apresentados os gráficos dos cenários resultantes das simulações dos métodos citados no Capítulo 3, referente a Materiais e Métodos. Foram utilizadas 30 amostras sob controle e adicionadas 20 amostras sem alteração nos limites de controle originais, mas também gerados com o processo sob controle.

Cenário 1: Parâmetros da simulação: $n = 5$, $\rho = 0,4$, $\delta = 0$

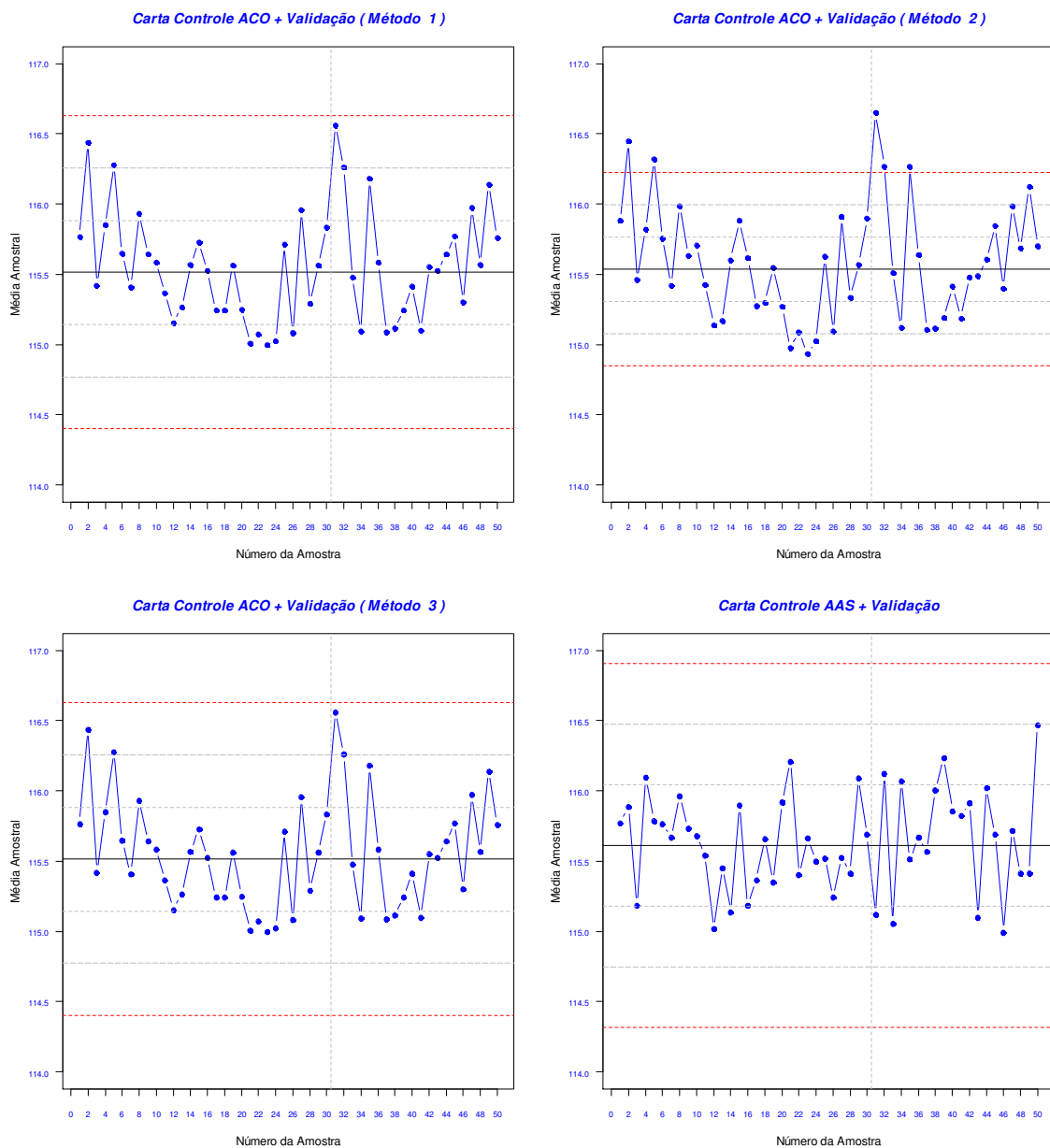


Figura 4 – Resultados da simulação para $n = 5$, $\rho = 0,4$, $\delta = 0$
Fonte: Os autores (2013)

Nos Figura 4 identificamos descontrole já início do processo via BLUE enquanto nos demais métodos não há indicativos de descontrole. A seguir verifica-se que os métodos 1 e 3 tem comportamentos similares, mas não chegam a apresentar descontrole do processo, o mesmo é verificado no método via AAS.

Cenário 2: Parâmetros da simulação: $n = 10$, $\rho = 0,4$, $\delta = 0$

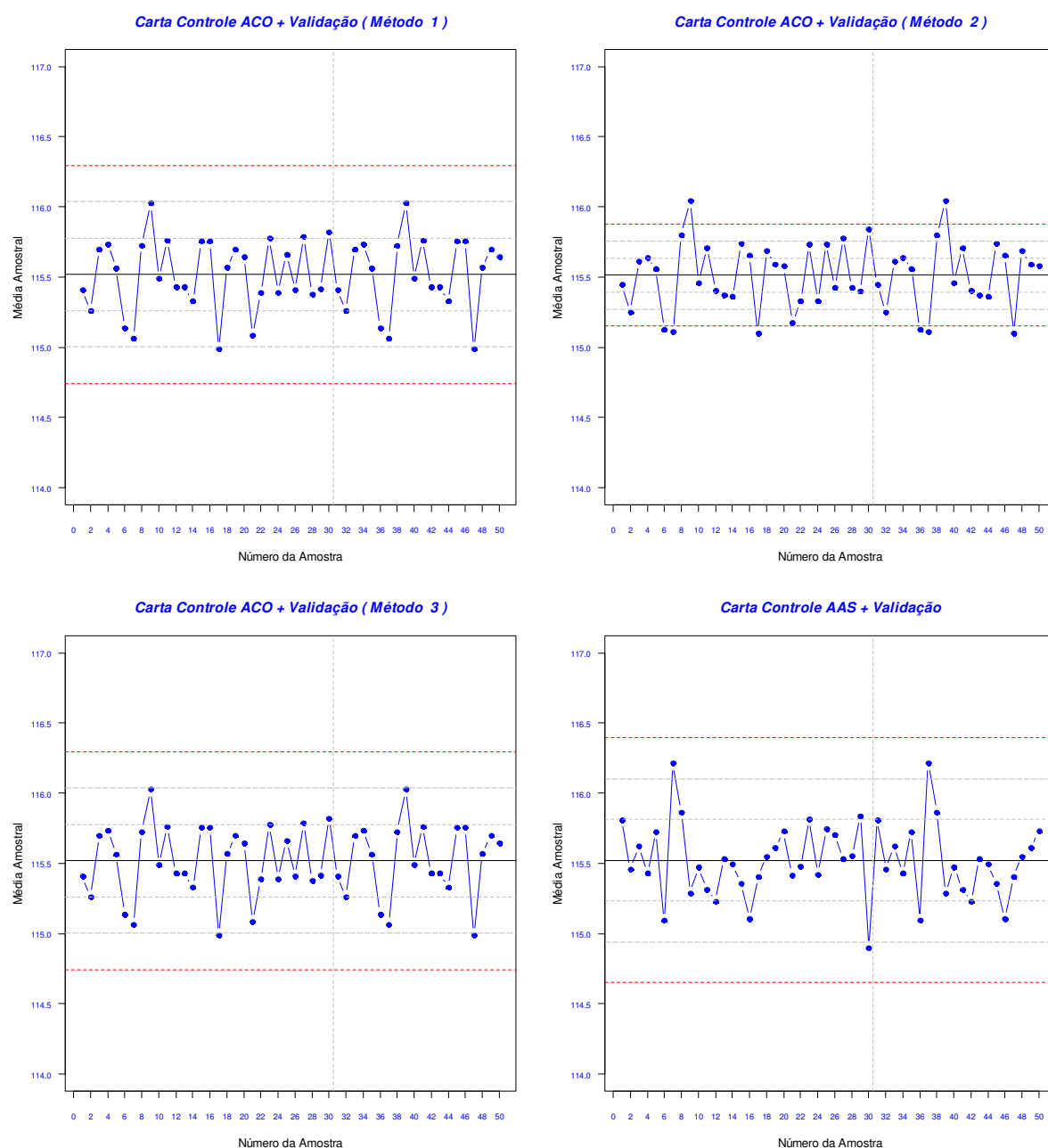


Figura 5 – Resultados da simulação para $n = 10$, $\rho = 0,4$, $\delta = 0$
Fonte: Os autores (2013)

Na Figura 5 verifica-se também que os métodos 1 e 3 via ACO são bastante similares, indicando processo sob controle. O mesmo é possível verificar

através do método via AAS. No entanto, o método via BLUE ainda no processo sob controle identifica pontos fora de controle, configurando novamente falsos alarmes.

Cenário 3: Parâmetros da simulação: $n = 5$, $\rho = 0,8$, $\delta = 0$

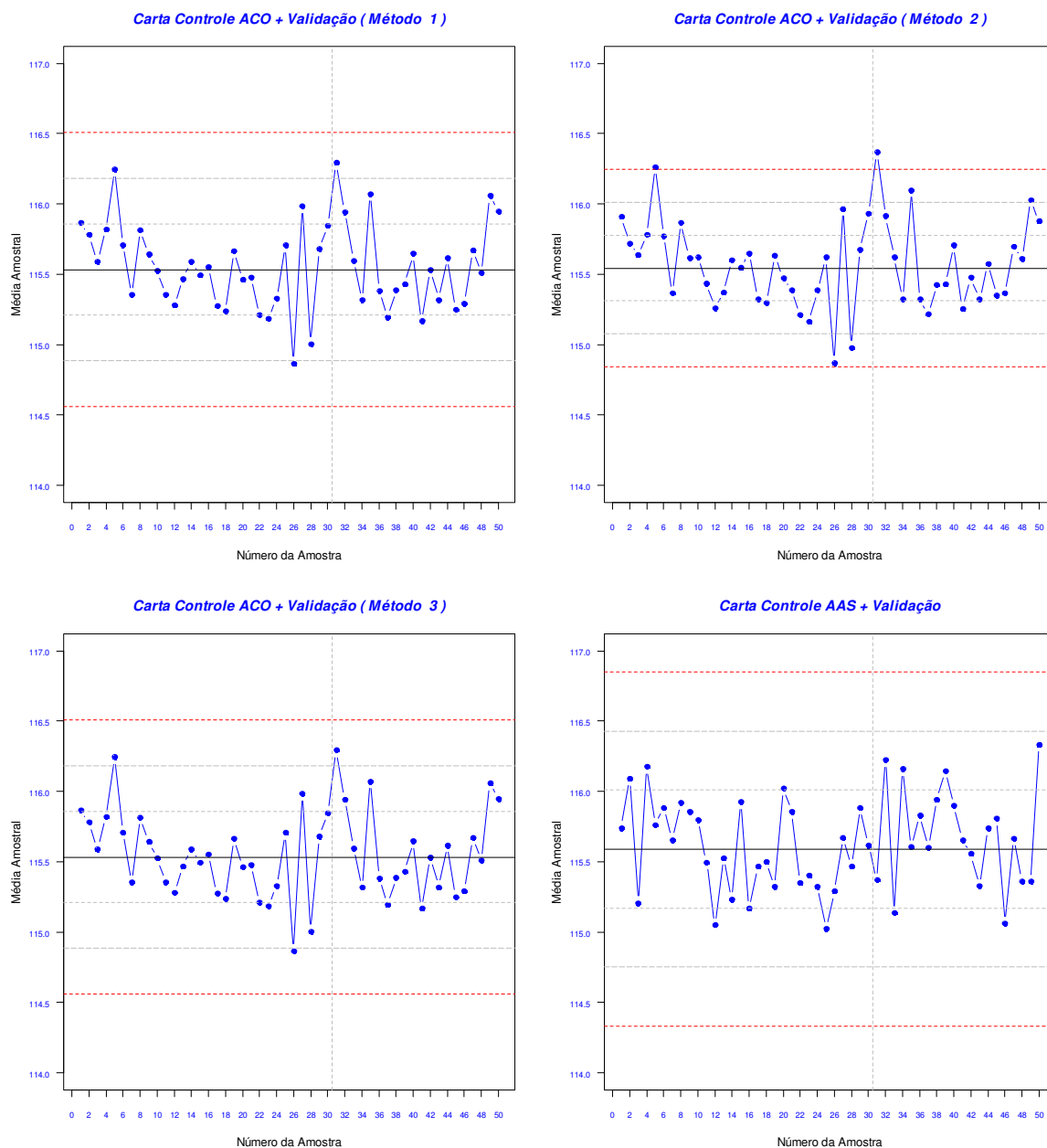


Figura 6 – Resultados da simulação para $n = 5$, $\rho = 0,8$, $\delta = 0$
Fonte: Os autores (2013)

Na Figura 6, verificamos através do método BLUE duas amostras fora de controle, configurando falsos alarmes, e as demais mantem-se abaixo de 3σ . Nos demais métodos não houve detecção de pontos fora de controle.

Cenário 4: Parâmetros da simulação: $n = 10$, $\rho = 0,8$, $\delta = 0$

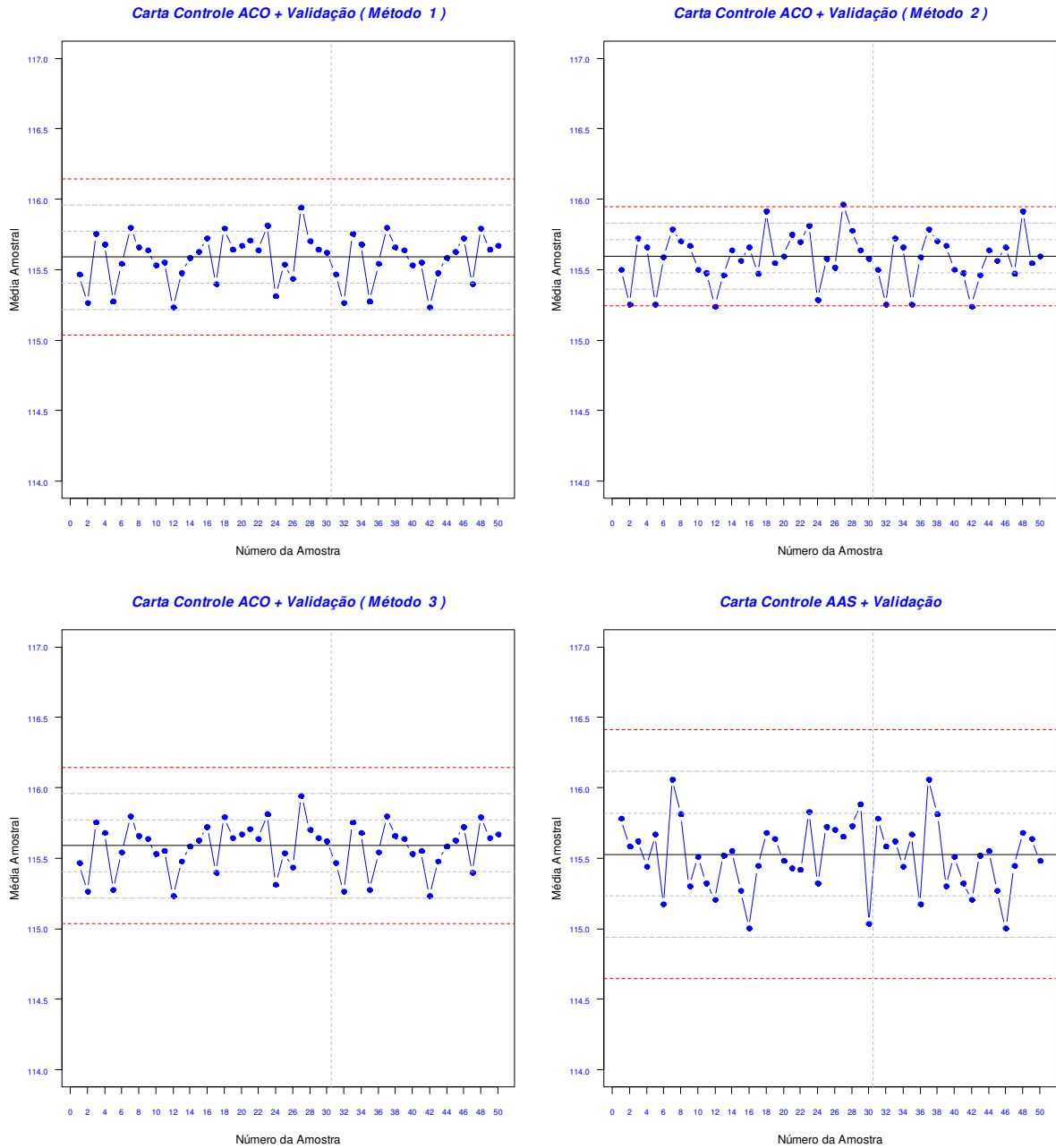


Figura 7 – Resultados da simulação para $n = 10$, $\rho = 0,8$, $\delta = 0$
 Fonte: Os autores (2013)

Na Figura 7 temos um comportamento similar entre os métodos 1, 3 e via AAS com todos os pontos entre os limites de controle. No método 2 identificamos duas amostras ainda no processo de desenvolvimento dos limites muito próximas ao limite 3σ , indicando incorretamente um possível descontrole.

Cenário 5: Parâmetros da simulação: $n = 5$, $\rho = 0,4$, $\delta = 0,5\sigma$

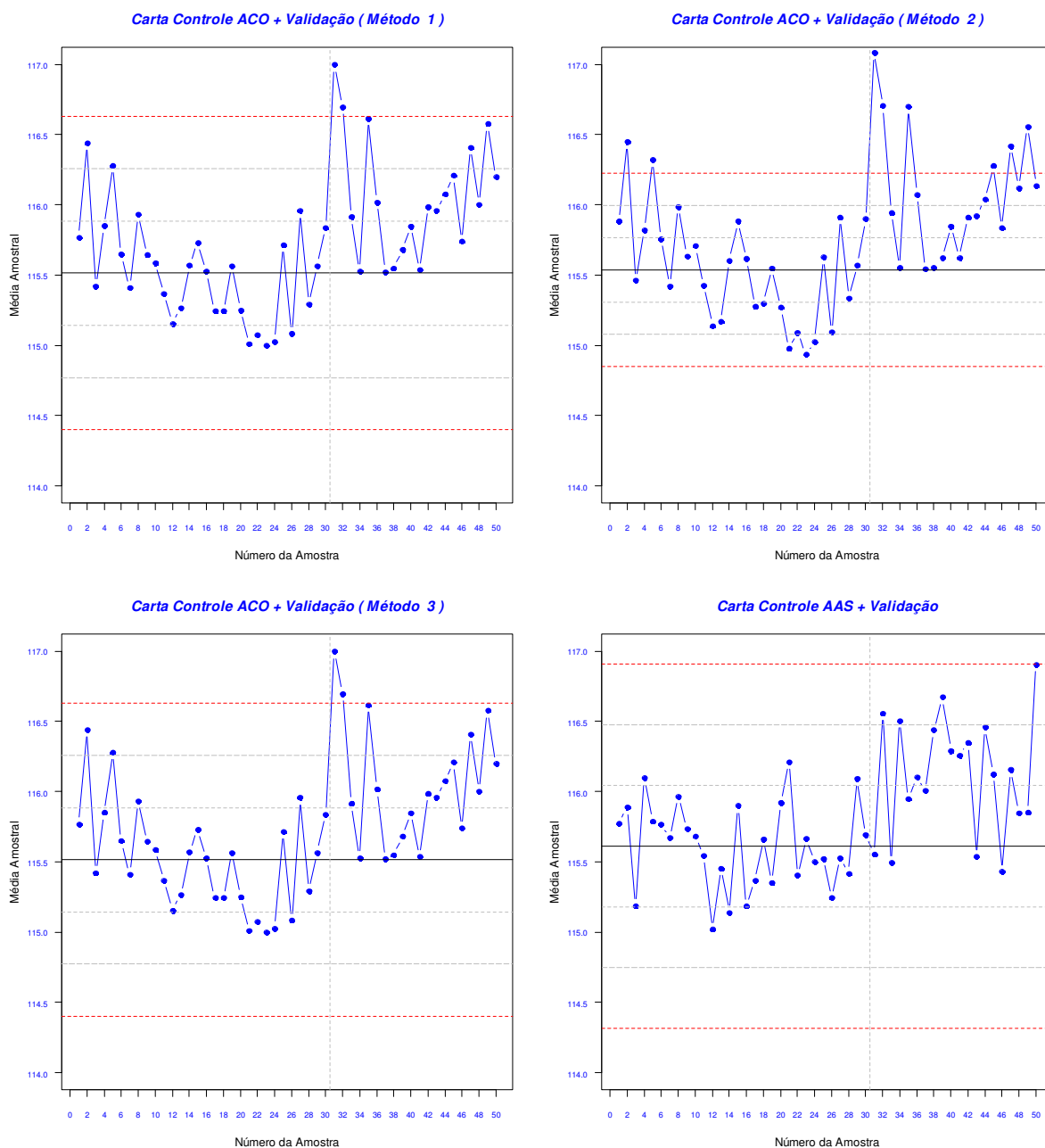


Figura 8 – Resultados da simulação para $n = 5$, $\rho = 0,4$, $\delta = 0,5\sigma$
 Fonte: Os autores (2013)

Na Figura 8, adicionando um descontrole de $0,5\sigma$ e uma correlação menor, os métodos 1 e 3 via ACO apresentaram uma amostra fora dos limites, já sinalizando o descontrole, enquanto no método 2 (BLUE) aponta diversas amostras fora de controle, possivelmente evidenciando falsos alarmes. O método via AAS não acusa amostras fora de controle.

Cenário 6: Parâmetros da simulação: $n = 10$, $\rho = 0,4$, $\delta = 0,5\sigma$

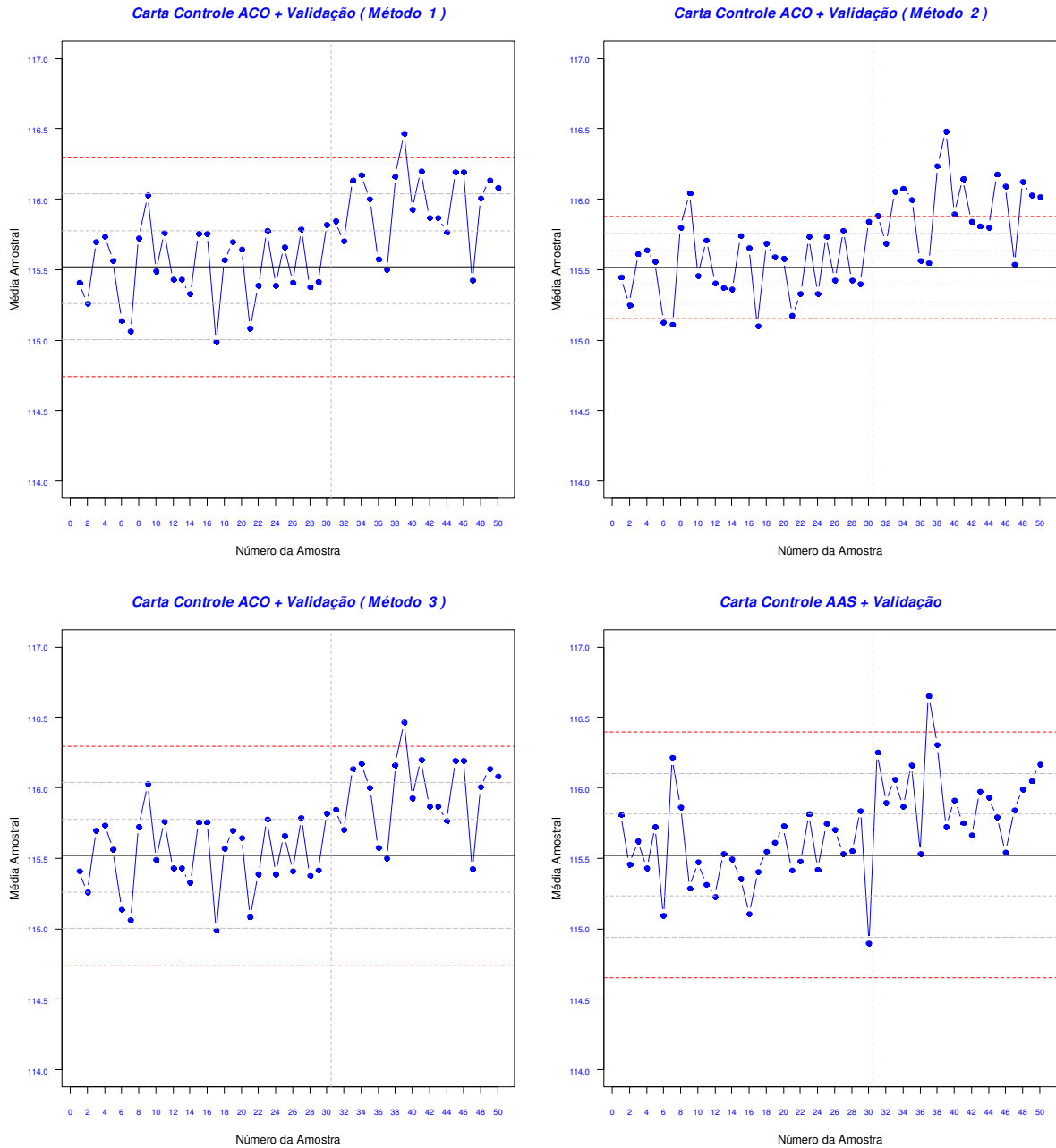


Figura 9 – Resultados da simulação para $n = 10$, $\rho = 0,4$, $\delta = 0,5\sigma$
 Fonte: Os autores (2013)

Ao adicionar $0,5\sigma$ à média do processo, verificamos que todos os processos produziram pontos acima de 3σ , indicando o descontrole. Verificou-se ainda que no método via AAS identificou mais rapidamente o descontrole.

Cenário 7: Parâmetros da simulação: $n = 5$, $\rho = 0,8$, $\delta = 0,5\sigma$

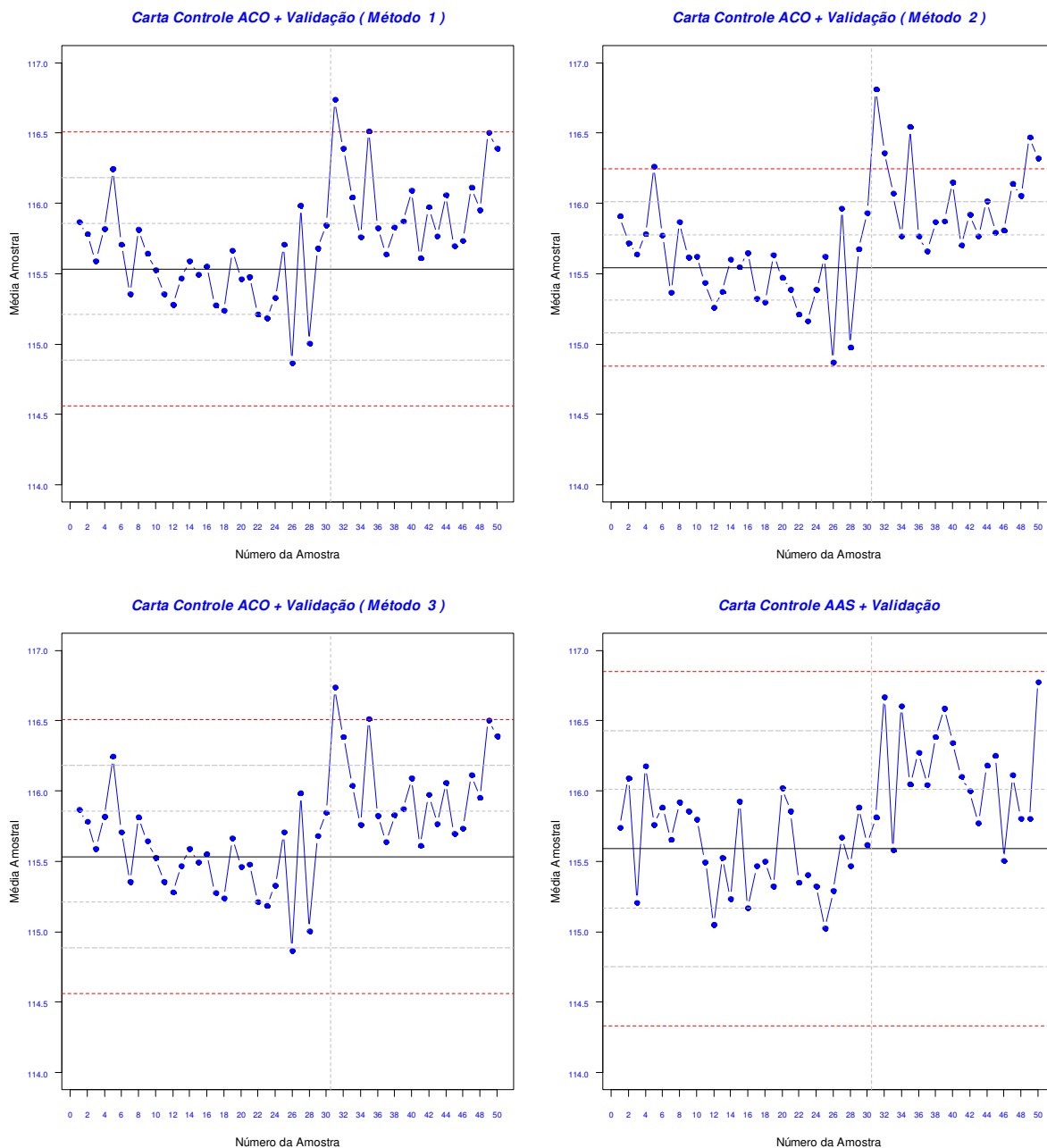


Figura 10 – Resultados da simulação para $n = 5$, $\rho = 0,8$, $\delta = 0,5\sigma$
 Fonte: Os autores (2013)

Na Figura 10 verificamos que todos os métodos via ACO identificaram processo fora de controle logo na 1ª amostra. No método 1 e 3, apenas um ponto ultrapassa o limite 3σ enquanto os demais permanecem no limite. Através do BLUE identificamos mais amostras fora dos limites de controle. No método via AAS não identificamos o processo fora de controle.

Cenário 8: Parâmetros da simulação: $n = 10$, $\rho = 0,8$, $\delta = 0,5\sigma$

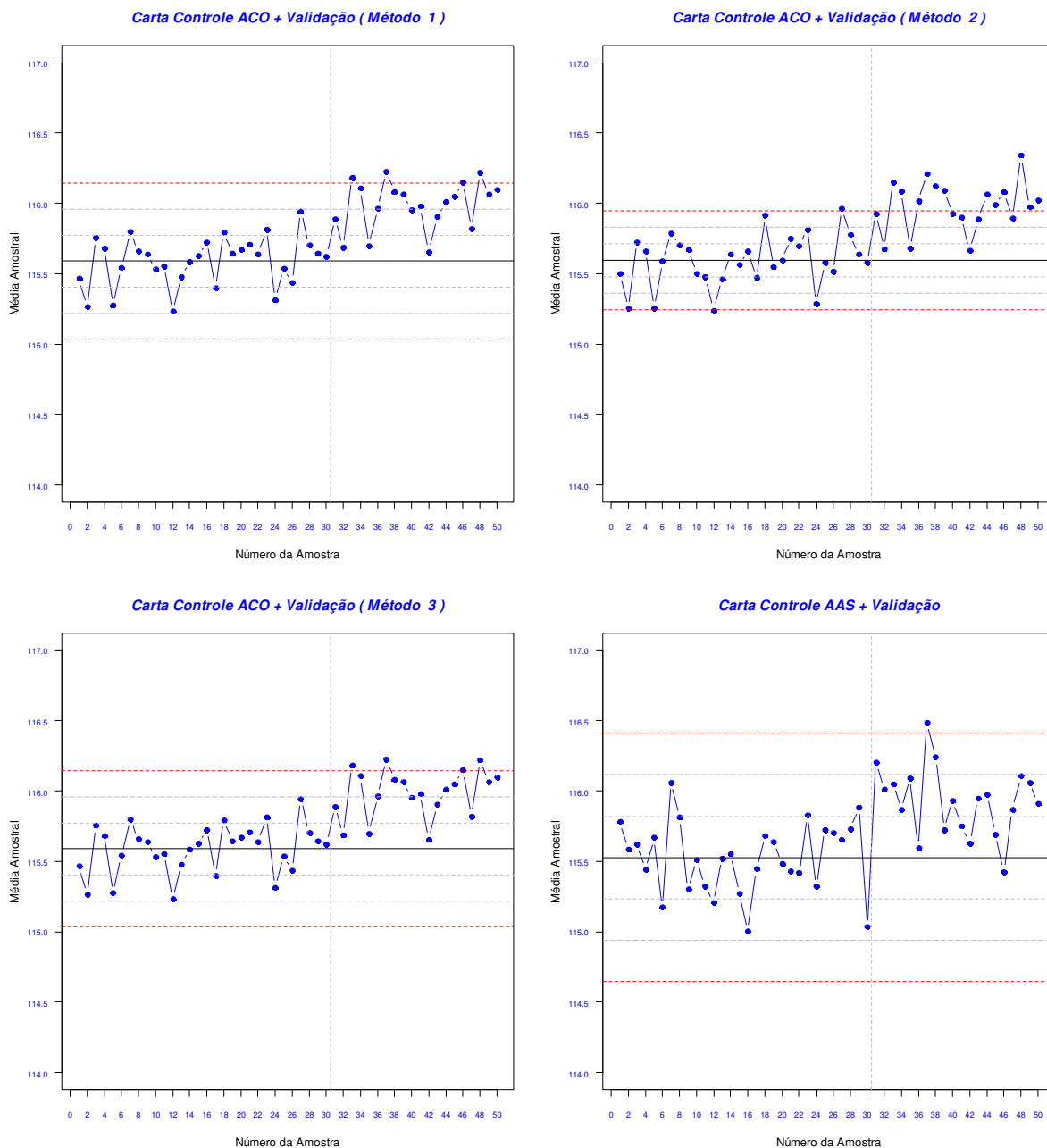


Figura 11 – Resultados da simulação para $n = 10$, $\rho = 0,8$, $\delta = 0,5\sigma$
 Fonte: Os autores (2013)

Na Figura 11 os três métodos via ACO identificam processo fora de controle a partir do 3º amostra pós descontrole, enquanto o método via AAS só faz a mesma identificação a partir da 7º amostra. Neste cenário tivemos um melhor desempenho dos métodos via ACO do que na Figura 9, no qual tínhamos o mesmo

tamanho de amostra, porém uma correlação maior entre as duas variáveis, que passou de $\rho = 0,4$ para $\rho = 0,8$.

Cenário 9: Parâmetros da simulação: $n = 5$, $\rho = 0,4$, $\delta = 0,8\sigma$

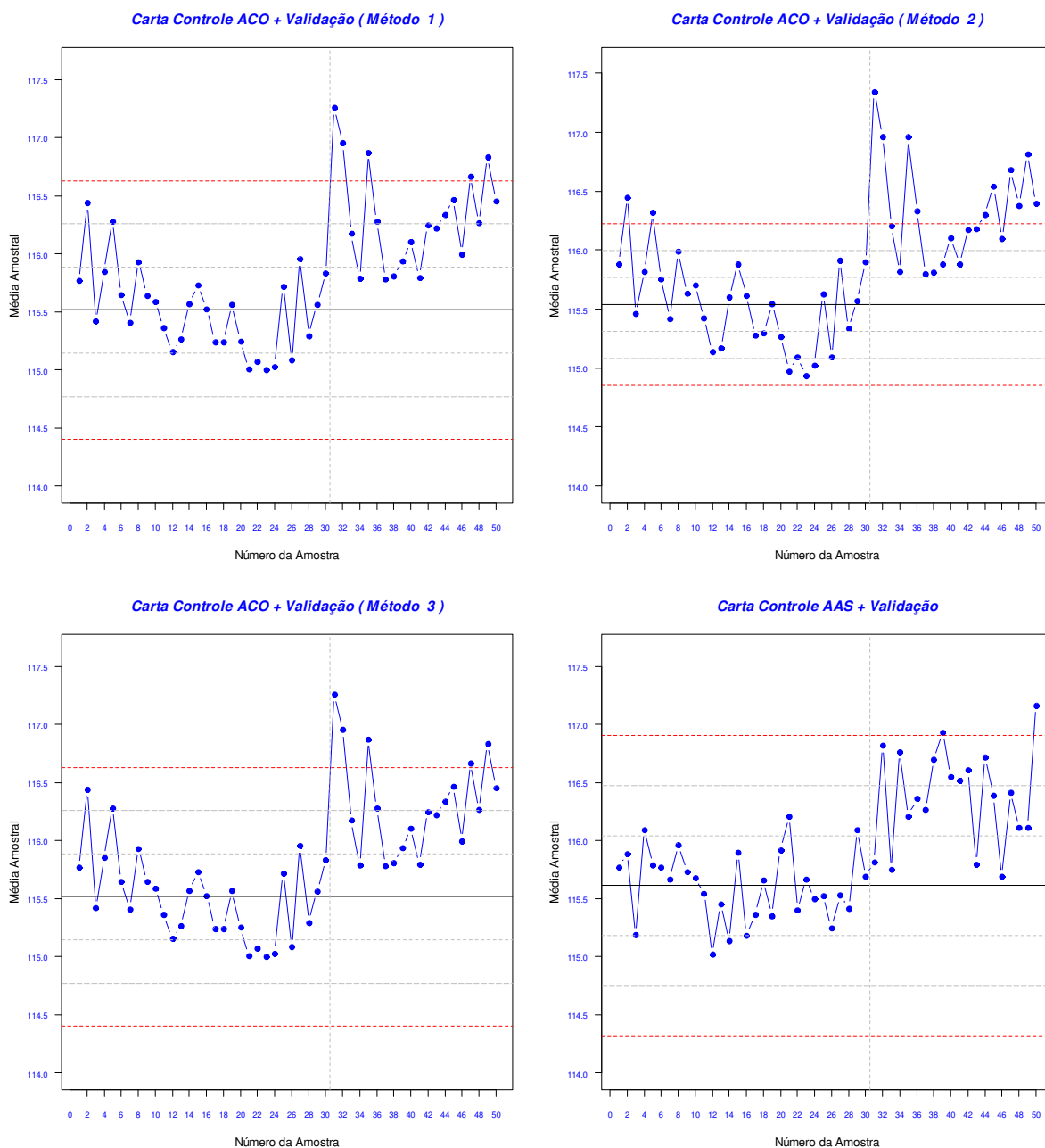


Figura 12 – Resultados da simulação para $n = 5$, $\rho = 0,4$, $\delta = 0,8\sigma$
Fonte: Os autores (2013)

Mantendo a alteração na média, mas com uma correlação menor, é possível identificar o descontrole do processo nos primeiros pontos nos métodos via ACO. No método via AAS apenas o último ponto identificou processo em descontrole, evidenciando a eficiência dos métodos via ACO.

Cenário 10: Parâmetros da simulação: $n = 10$, $\rho = 0,4$, $\delta = 0,8\sigma$

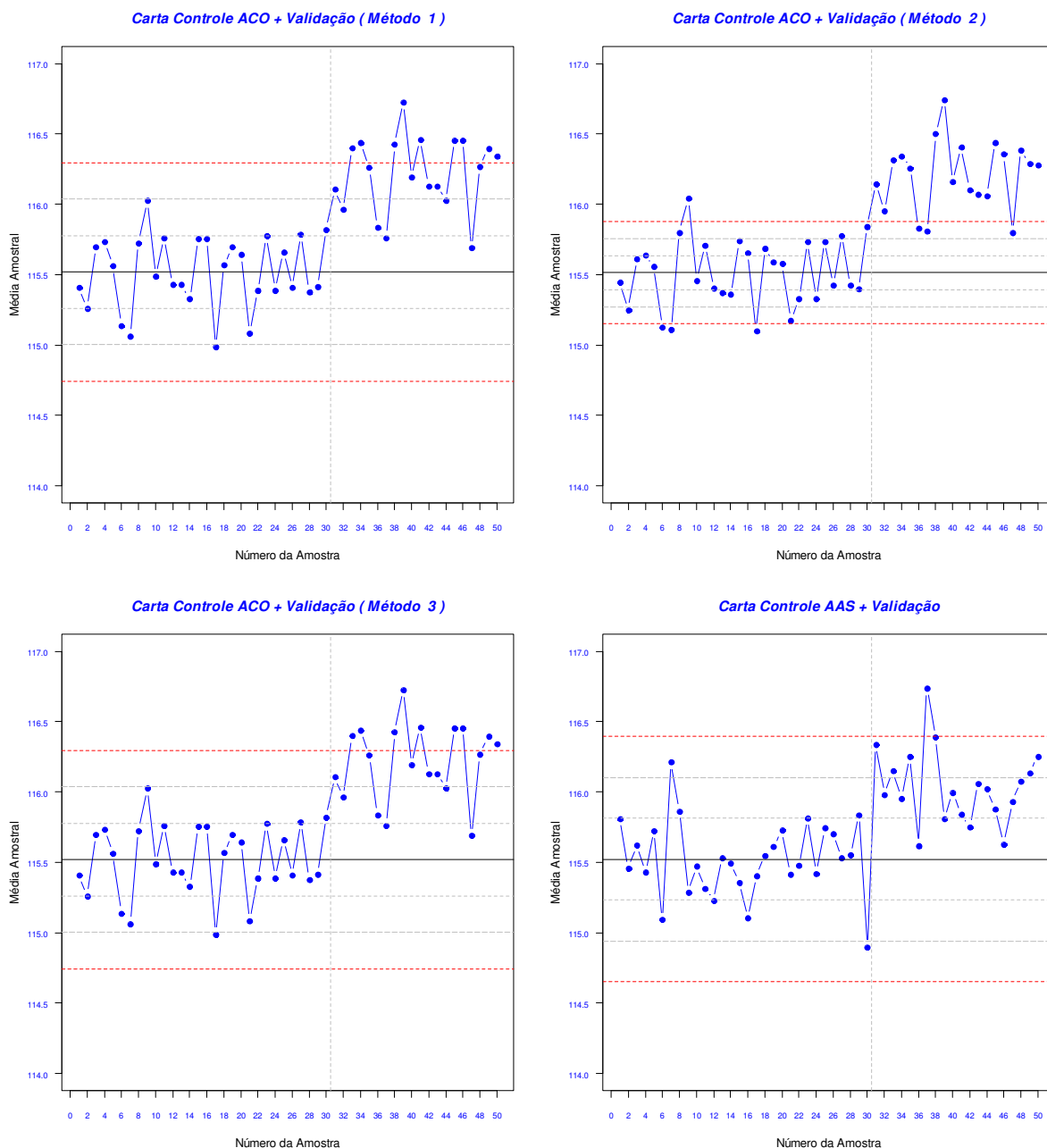


Figura 13 – Resultados da simulação para $n = 10$, $\rho = 0,4$, $\delta = 0,8\sigma$
Fonte: Os autores (2013)

Para uma alteração na média de $0,8\sigma$, verificamos que o método 1 e 3 via ACO tiveram comportamento similar, identificando descontrole do processo a partir da 3ª amostra. O método 2 (BLUE) verifica descontrole no início do processo, o que pode ser indicativo de falsos alarmes, demonstrando que o método está instável. No

método via AAS também identificamos processo fora de controle, porém somente a partir da 9ª amostra.

Cenário 11: Parâmetros da simulação: $n = 5$, $\rho = 0,8$, $\delta = 0,8\sigma$

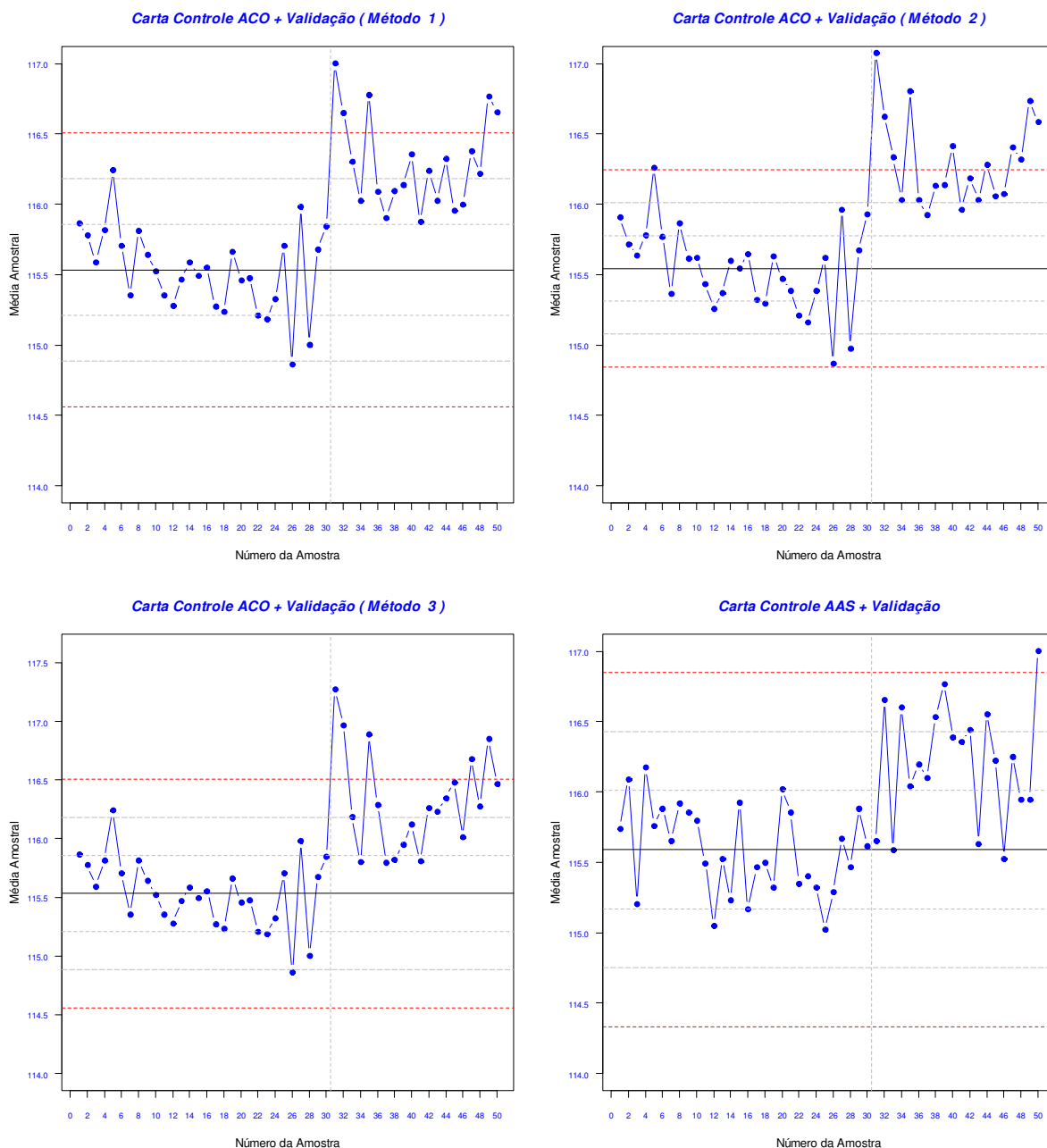


Figura 14 – Resultados da simulação para $n = 5$, $\rho = 0,8$, $\delta = 0,8\sigma$
Fonte: Os autores (2013)

Ainda com alteração de $0,8\sigma$ na média, mas para $n = 5$, identificamos em nos métodos via ACO indicativo de descontrole logo na 1ª amostra pós descontrole. No método via AAS este descontrole só é identificado na última amostra pós descontrole.

Cenário 12: Parâmetros da simulação: $n = 10$, $\rho = 0,8$, $\delta = 0,8\sigma$

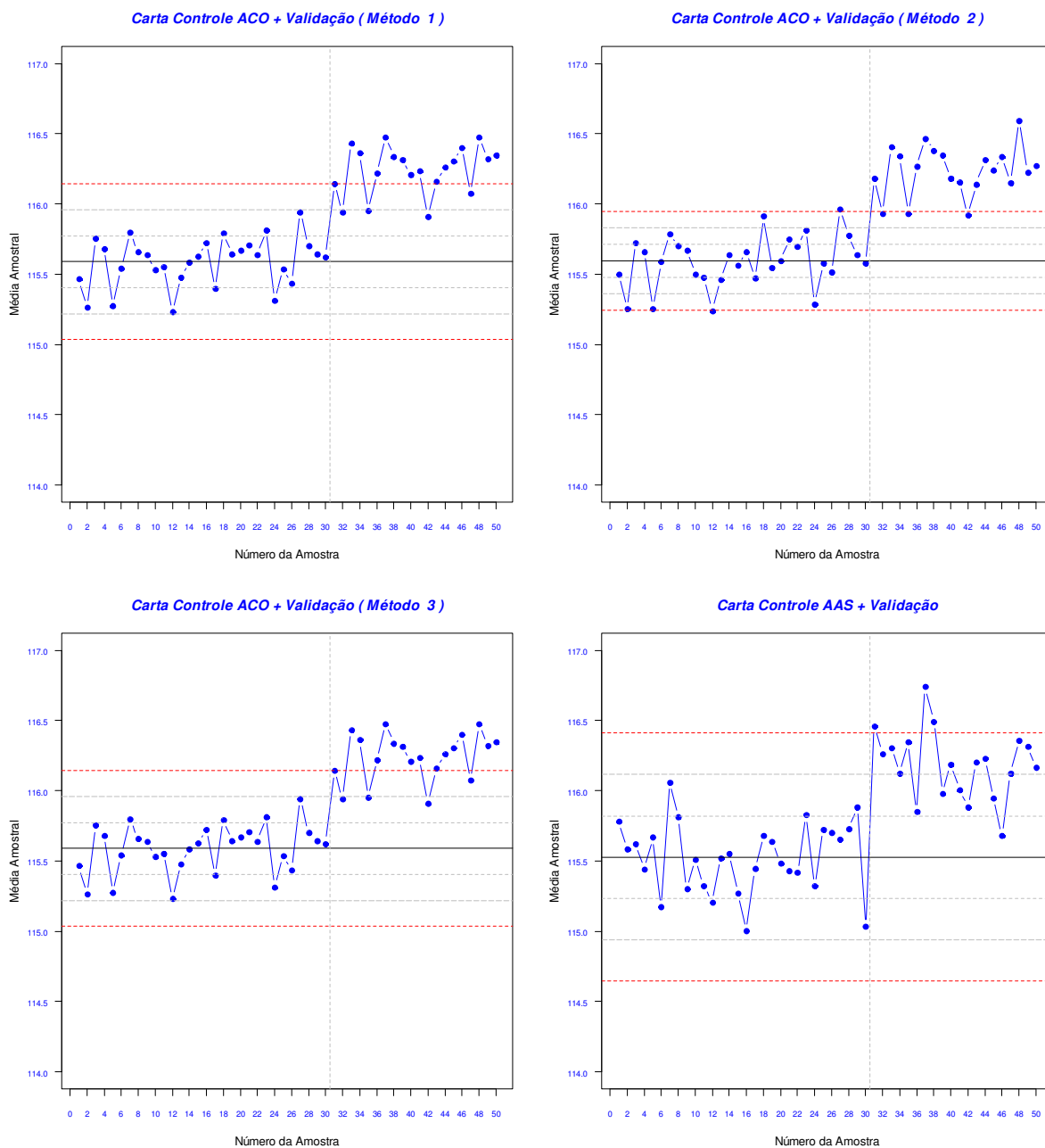


Figura 15 – Resultados da simulação para $n = 10$, $\rho = 0,8$, $\delta = 0,8\sigma$
 Fonte: Os autores (2013)

Na Figura 15 com alteração na média de $0,8\sigma$, verificamos que todos os métodos identificaram processo fora de controle na primeira amostra pós descontrole.

A seguir temos os resultados apresentados pelas Curvas Características de Operação e Comprimentos Médios de Sequência.

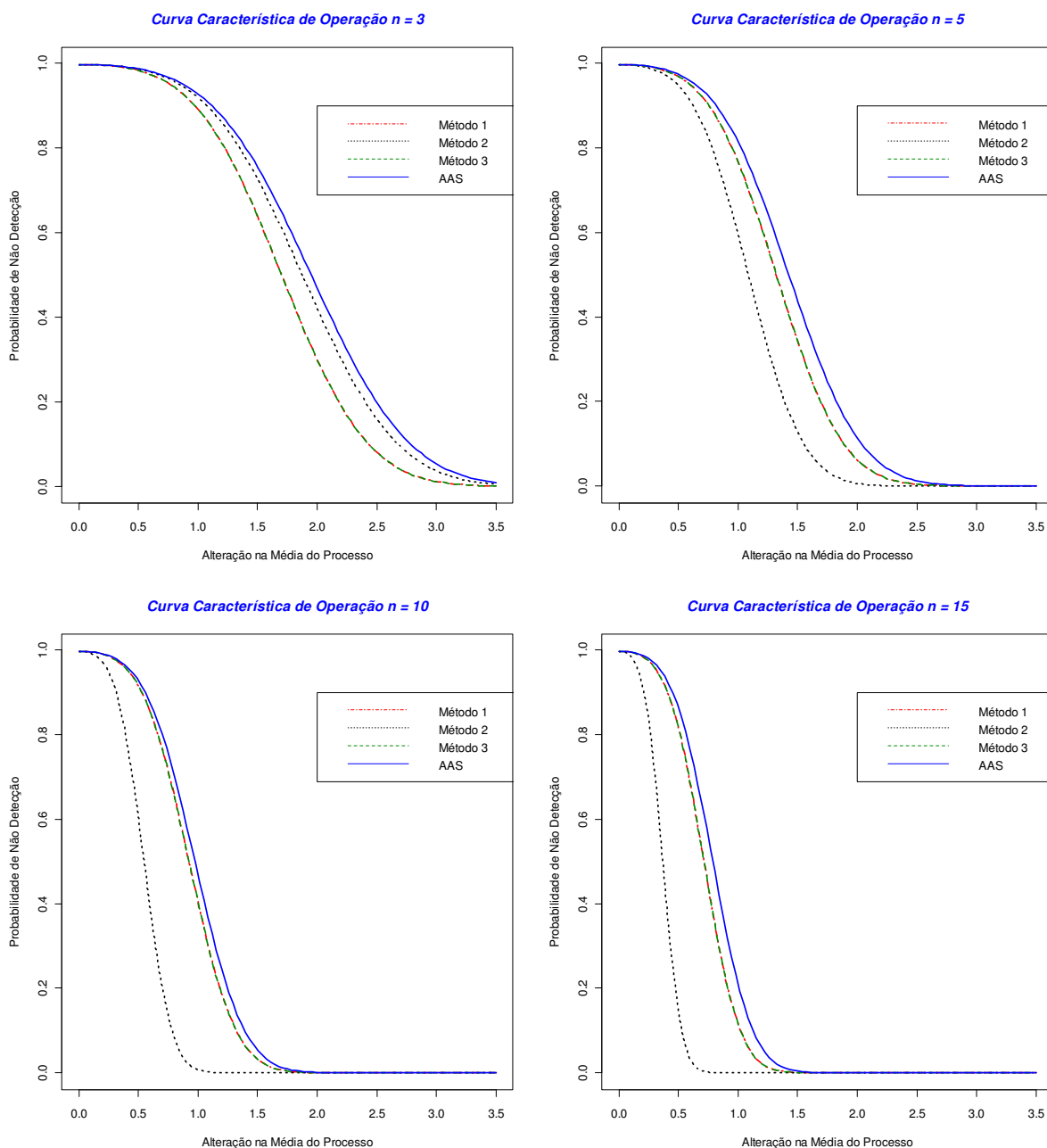


Figura 16 – Curva Característica de Operação com $\rho = 0,4$
Fonte: Os autores (2013)

Na Figura 16, temos as Curvas Características de Operação nas quais conclui-se que com $n = 3$ e $\rho = 0,4$, verifica-se um melhor desempenho nos métodos 1 e 3 que ficaram sobrepostos. Para $n = 5$, $n = 10$ e $n = 15$, temos que o método 2 apresenta-se mais eficaz, ou seja, tem a detecção mais rápida. Logo após,

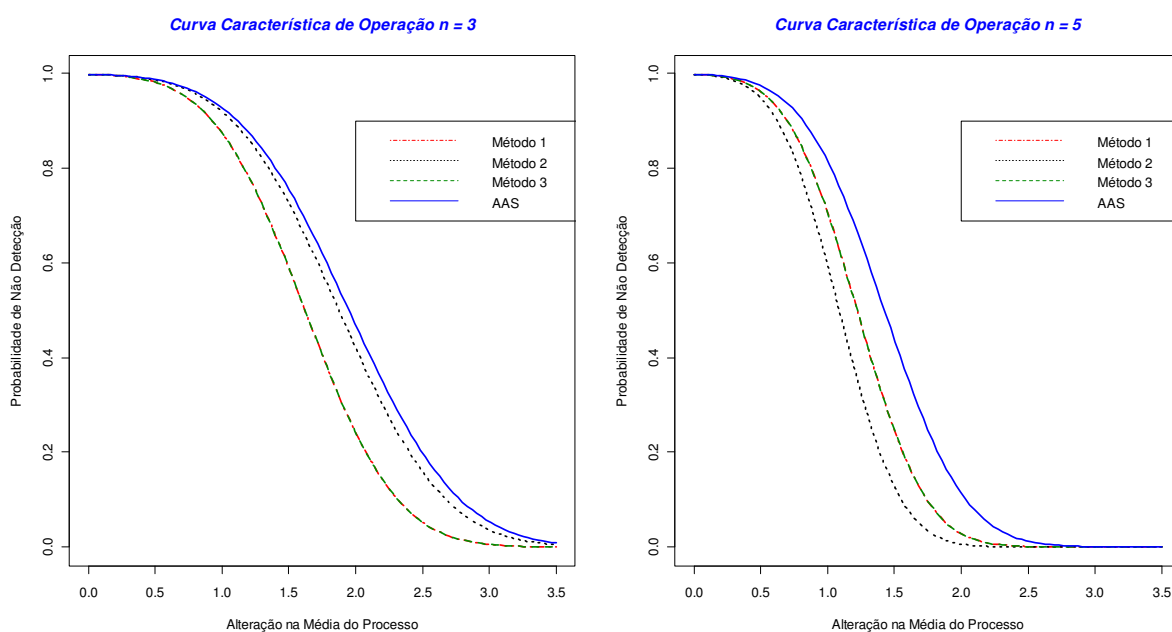
os métodos 1 e 3 via ACO apresentam-se mais eficientes. Verifica-se que para todos os tamanhos de n os métodos via ACO tem melhores resultados em relação ao método via AAS. Na tabela a seguir, verifica-se resultados pontuais relacionados a Figura 16, para melhor comparativo entre os métodos.

Tabela 1 – Resultados das simulações da Curva Característica de Operação com $\rho = 0,4$

Curva Característica de Operação				
n	ACO			AAS
	Método 1	Método 2	Método 3	Método 4
n = 03	0,983	0,986	0,983	0,987
n = 05	0,969	0,947	0,969	0,974
n = 10	0,916	0,603	0,916	0,928
n = 15	0,816	0,145	0,816	0,864

Fonte: OS autores (2013).

A Tabela 1 corresponde a uma correlação igual a 0,4 e alteração na média do processo de $0,5\sigma$. Verificamos que a probabilidade de não detecção é maior através da amostragem aleatória simples para qualquer tamanho de amostra, ou seja, a detecção pelos métodos via ACO é mais eficiente que AAS.



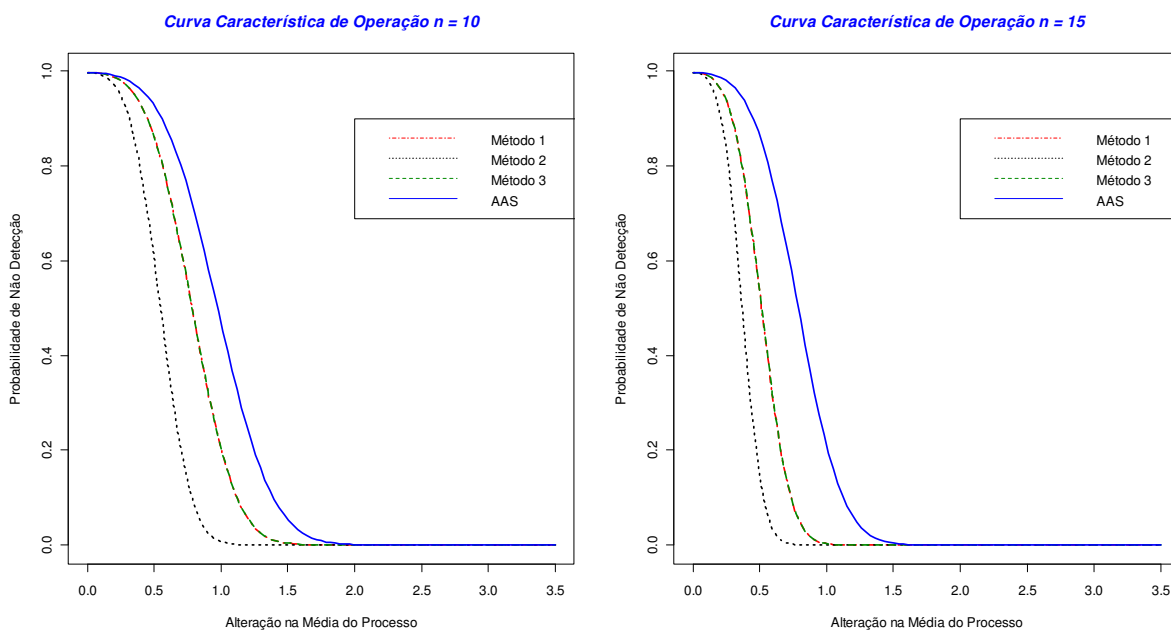


Figura 17 – Curva Característica de Operação com $\rho = 0,8$
 Fonte: Os autores (2013)

Na Figura 17, temos o mesmo desempenho das Curvas características de Operação da Figura 16, nos quais os métodos via ACO demonstraram ser mais eficientes do que o método via AAS. Na Tabela 2 temos uma melhor visualização destes resultados.

Tabela 2 – Resultados das simulações da Curva Característica de Operação com $\rho = 0,8$

Curva Característica de Operação				
n	ACO			AAS
	MÉTODO 1	MÉTODO 2	MÉTODO 3	MÉTODO 4
3	0,981	0,986	0,981	0,987
5	0,962	0,947	0,962	0,974
10	0,862	0,603	0,862	0,928
15	0,532	0,145	0,532	0,864

Fonte: OS autores (2013).

Para uma correlação maior, esse detecção é ainda mais eficiente, tendo em vista que os métodos via ACO tem uma melhor performance em relação ao método via AAS. Mantendo as mesmas características de baixa correlação, o método 2 também não apresenta mudanças na detecção de descontrole quando se altera a correlação, porém verifica-se um comportamento nada similar aos demais métodos.

A seguir temos os gráficos referente ao Comprimento Médio de Sequência.

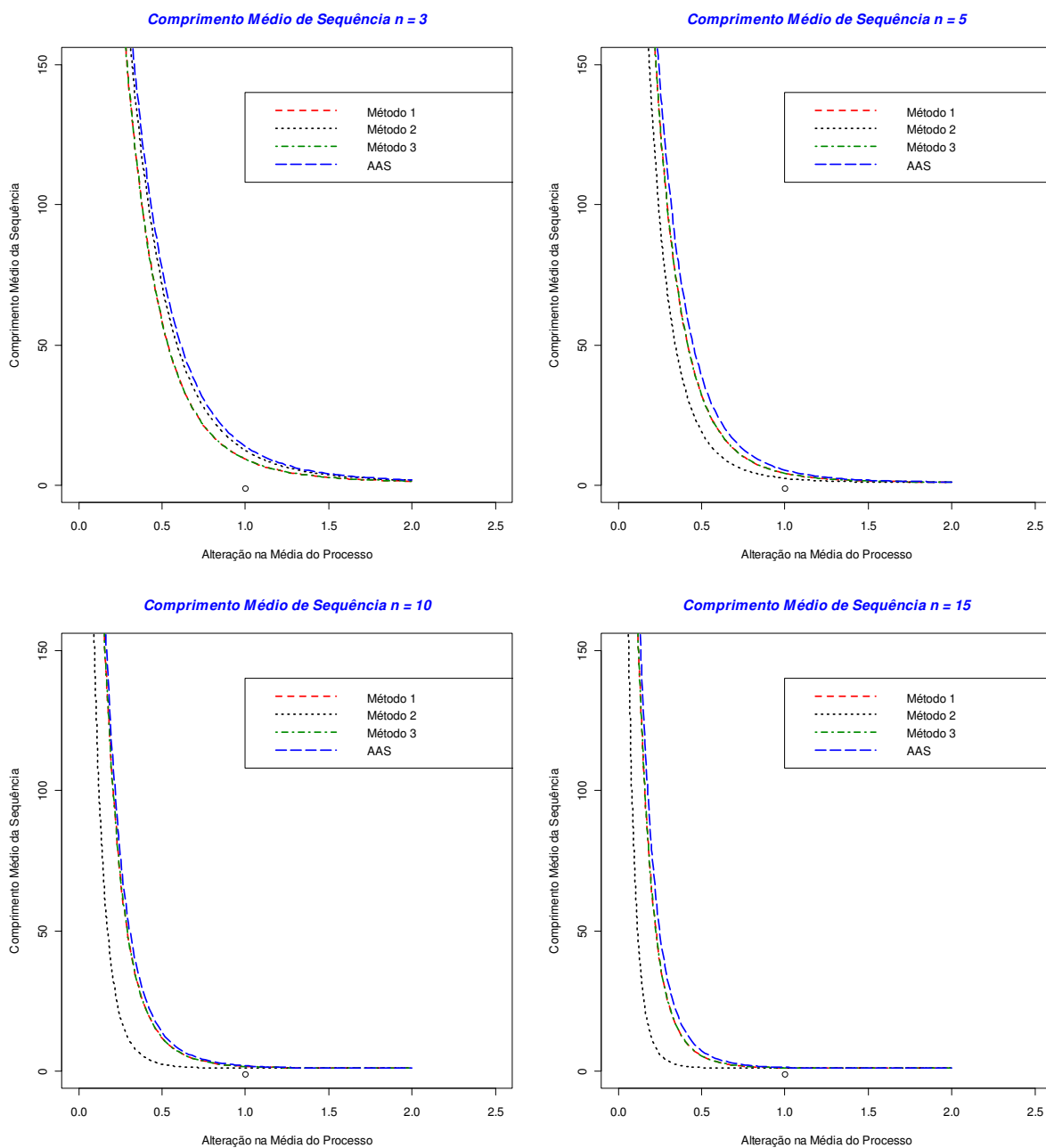


Figura 18 – Resultados do Comprimento Médio de Sequência com $\rho = 0,4$
 Fonte: Os autores (2013)

No Figura 18, verifica-se que o número esperado de amostras necessárias para se detectar um descontrolo. Podemos identificar que o método 2 (BLUE) necessita de um número de amostra menor que os demais para detecção de descontrolo. Porém, para $n = 3$ temos os métodos 1 e 3 via ACO são mais eficientes, enquanto para todos os tamanhos de n temos como pior desempenho o método via AAS.

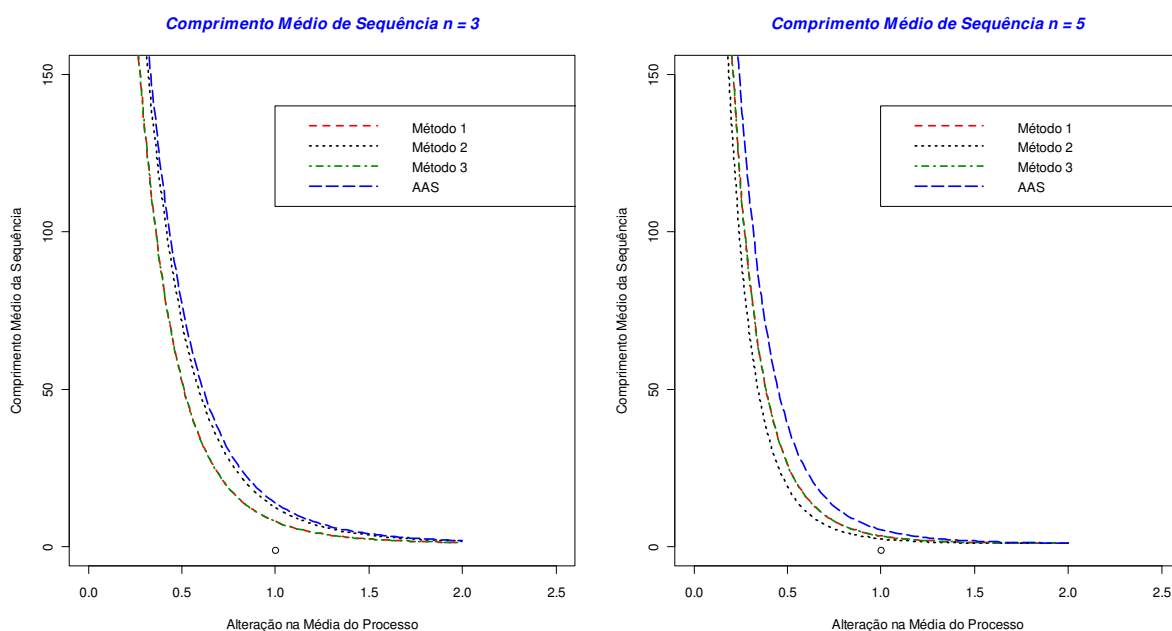
Para auxiliar na comparação dos métodos, temos a Tabela 4. Nela verifica-se que quanto menor o valor, menor é a quantidade de amostras necessárias para detecção de um descontrolado, após ocorrer a alteração na média do processo.

Através da Tabela 4, com correlação igual a 0,4, identificamos que os métodos via ACO precisam de uma quantidade de amostras menor para identificar um descontrolado no processo quando comparados com o método via AAS. Como exemplo, temos o processo contendo 10 elementos na amostra: para os métodos 1 e 3 via ACO faz-se necessário aproximadamente 11 amostras para detecção de descontrolado no processo, enquanto para o método 4, que trata-se de delineamento via AAS necessita de aproximadamente 14 amostras para identificação de processo fora de controle.

Tabela 3 – Resultados das simulações do Comprimento Médio de Sequência com $\rho = 0,4$

Comprimento Médio de Sequência				
n	ACO			AAS
	Método 1	Método 2	Método 3	Método 4
n = 03	58,416	71,544	58,412	77,227
n = 05	32,126	19,004	32,126	38,836
n = 10	11,888	2,518	11,888	13,902
n = 15	5,435	1,168	5,435	7,341

Fonte: OS autores (2013).



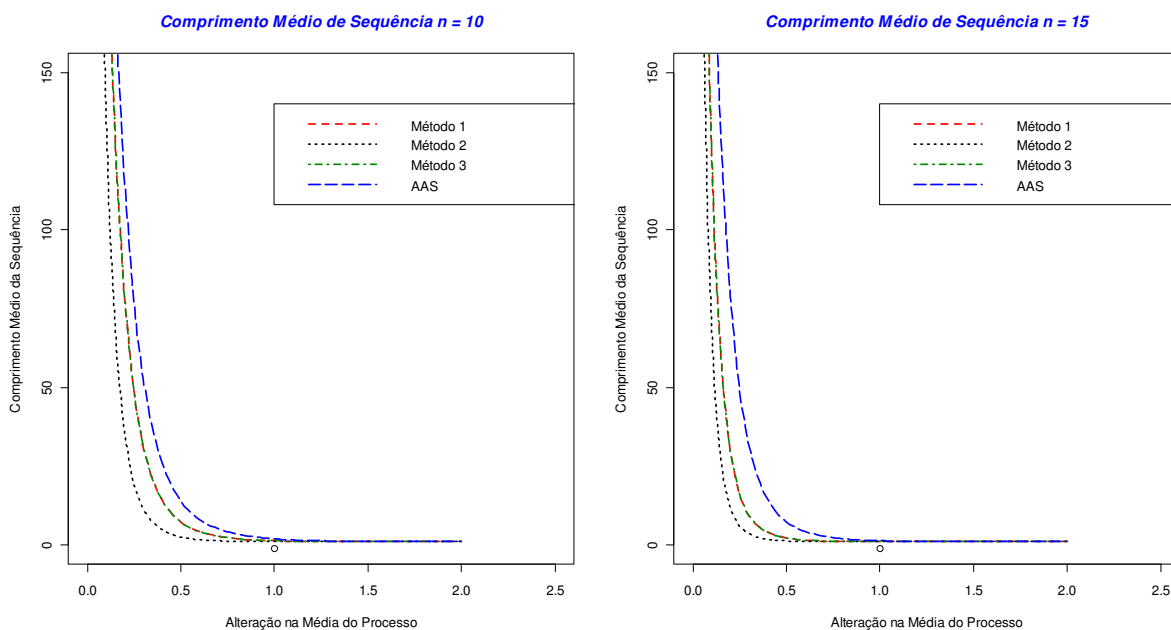


Figura 19 – Resultados do Comprimento Médio de Sequência com $\rho = 0,8$
 Fonte: Os autores (2013)

Na Figura 19, temos o mesmo desempenho do Comprimento Médio de Sequência da Figura 18, nos quais os métodos via ACO demonstraram ser mais eficientes do que o método via AAS. Na Tabela 4, temos os resultados obtidos através dos gráficos.

Tabela 4 – Resultados das simulações do Comprimento Médio de Sequência com $\rho = 0,4$

Comprimento Médio de Sequência				
ACO				AAS
n	Método 1	Método 2	Método 3	Método 4
n = 03	52,498	71,544	52,483	77,227
n = 05	26,138	19,004	26,131	38,836
n = 10	7,229	2,518	7,228	13,902
n = 15	2,136	1,169	2,136	7,341

Fonte: OS autores (2013).

Para uma correlação de 0,8 ainda fica evidente a efetividade dos métodos via ACO comparado com o método via AAS. Identificamos que com uma correlação maior o desempenho dos métodos 1 e 3 tem um ganho ainda mais significativo. Para os métodos 2 e 4 não tiveram alterações devido ao aumento na correlação.

5 CONCLUSÃO

O delineamento amostral por conjuntos ordenados tem se mostrado mais precisos para diversos parâmetros do que a amostragem aleatória simples na estimação de diversos parâmetros, principalmente nos casos em que não é possível extrair uma amostra numerosa, sendo indicado quando os elementos amostrais podem ser ordenados de maneira simples e eficiente antes de sua efetiva mensuração.

No controle estatístico de qualidade existem vários métodos que podem ser utilizados para monitorar um processo: desde a estatística descritiva até métodos mais específicos como gráficos de controle, curva característica de operação e comprimento médio de sequência. Através do gráfico de controle é possível detectar se um processo está sob controle. Já através da curva característica de operação é utilizado como base para avaliar a capacidade de detecção do descontrole do processo e o comprimento médio de sequência é utilizado para avaliar o desempenho do gráfico de controle quanto ao número esperado de amostras até se ter um alarme.

Neste trabalho, através de um estudo por simulação e por meio de programação de funções, foram considerados quatro métodos de construção de gráficos de controle, com diferentes cenários simulados que foram definidos pelo tamanho da amostra, grau de correlação das variáveis e grau de descontrole do processo.

Após a elaboração dos gráficos obtivemos como principais resultados que para todos os tamanhos de n observou-se um ganho em termos de probabilidade de detecção e de comprimento médio de sequência dos métodos baseados em ACO em relação ao método via AAS.

Ressalta-se que o método via BLUE foi implementado rigorosamente de acordo com a proposta apresentada por Muttlak e Al-Sabah (2003), mostrando-se excessivamente sensível, produzindo elevadas taxas de falsos alarmes, ainda que, nas situações em que o processo operava sob controle.

Finalmente, sugerimos a aplicação deste delineamento em problemas locais. Propõe-se a divulgação desta metodologia e o uso em problemas de controle de qualidade reais, além da aplicação de outros métodos que utilizam este

delineamento, tais como: *median ranked set sampling* (MRSS) e *extreme ranked set sampling* (ERSS).

REFERÊNCIAS

BARNETT, V. Ranked set sampling design for environmental investigations. **Environ. Ecol. Stat.**, London, v.6, p.59-74, 1999.

BARNETT, V.; MOORE, K. **Best linear unbiased estimates in ranked set sampling with particular reference to imperfect ordering.** J. Appl. Stat., Abingdon, v.24, p.697-710, 1997.

BARRETO, M.C.M. **Planejamentos eficientes em pesquisa do Meio Ambiente usando amostragem em conjuntos ordenados.** Rev. Mat. Estat., São Paulo, v.19, p.71-84, 2001.

CONTROLE Estatístico do Processo. Disponível em: <<http://www.portalaction.com.br/>>. Acesso em: 24 abr. 2013.

GODOY, Adelize Leite. **Tutorial: Ferramentas da Qualidade.** CEDET – Centro de Desenvolvimento Profissional e Tecnológico. Disponível em: <www.cedet.com.br>. Acesso em: 24 abr. 2013.

KAUR, A.; PATIL, G.P.; TAILLIE, C. Unequal allocation models for ranked set sampling with skew distributions. **Biometrics**, Washington, v.53, p.123-30, 1997.

MUTTAK, Hassen; AL-SABAH, Walid. Statistical quality control based on ranked set sampling. **Journal of Applied Statistics**, 30:9, p. 1055-1078, 2003.

MCINTYRE, G.A **A method for unbiased selective sampling, using ranked sets.** Australian J. Agric. Res., Collingwood, v.3, p.385-90, 1952.

MONTGOMERY, D. C. **Introdução ao Controle Estatístico da Qualidade.** São Paulo: LTC, 3ª edição, 2001.

NELSON, L. S. The Shewhart Control Chart – Tests for Special Causes. **Journal of Quality Technology**, v.16, 1984.

Orientação para Normalização de Trabalhos Acadêmicos. Disponível em: <<http://www.portal.ufpr.br/normalizacao.html>>. Acesso em: 07 mai. 2013.

R DEVELOPMENT CORE TEAM. R: A Language and Environment for Statistical Computing. Disponível em: <<http://cran.r-project.org/>>. Acesso em: 24 abr. 2013.

STOKES, S.L. Estimation of variance using judgment ordered ranked set sampling. **Biometrics**, v. 36, p. 35-42, 1980.

SINHA, Arun K. On some recent developments in ranked set sampling. **Bulletin of informatics and cybernetics**, v. 37, p137-160, 2005.

Sinha, B. K. and Purkayastha, S.. On some aspects of ranked set sampling for estimation of normal and exponential parameters. **Statistics & Decision**, v. 14, p. 223– 240, 1996.

TACONELI, Cesar Augusto. Reamostragem Bootstrap em Amostragem por Conjuntos Ordenados e Intervalos de Confiança não-Paramétricos para a Média. São Carlos: UFSCar, 2005.

Takahasi, K. and Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering, **Annals of the Institute of Statistical Mathematics**, v. 20, p.1–31, 1968.

YU, P. L. H.; LAM, K.; SINHA, B. K. Estimation of normal variance based on balanced and unbalanced ranked set samples. **Environmental and ecological statistics**, v. 6, p. 23-46, 1999.

ANEXOS

ANEXO A: DADOS

Dados de diâmetro e resistência de fibra têxtil, extraídas de Montgomery (2003).

Tamanho da Amostra K	Médias Amostrais		Variâncias e Covariâncias			Estatísticas do Gráfico de Controle	
	Força de Resistência	Diâmetro	S^2_{1k}	S^2_{2k}	S_{12k}	T^2_k	$ S_k $
1	115,25	1,04	1,25	0,87	0,80	2,16	0,45
2	115,91	1,06	1,26	0,85	0,81	2,14	0,41
3	115,05	1,09	1,30	0,90	0,82	6,77	0,50
4	116,21	1,05	1,02	0,85	0,81	8,29	0,21
5	115,90	1,07	1,16	0,73	0,80	1,89	0,21
6	115,55	1,06	1,01	0,80	0,76	0,03	0,23
7	114,98	1,05	1,25	0,78	0,75	7,54	0,41
8	115,25	1,10	1,40	0,83	0,80	3,01	0,52
9	116,15	1,09	1,19	0,87	0,83	5,92	0,35
10	115,92	1,05	1,17	0,86	0,95	2,41	0,10
11	115,75	0,99	1,45	0,79	0,78	1,13	0,54
12	114,90	1,06	1,24	0,82	0,81	9,96	0,36
13	116,01	1,05	1,26	0,55	0,72	3,86	0,17
14	115,83	1,07	1,17	0,76	0,75	1,11	0,33
15	115,29	1,11	1,23	0,89	0,82	2,56	0,42
16	115,63	1,04	1,24	0,91	0,83	0,70	0,19
17	115,47	1,03	1,20	0,95	0,70	0,19	0,65
18	115,58	1,05	1,18	0,83	0,79	0,00	0,36
19	115,72	1,06	1,31	0,89	0,76	0,35	0,59
20	115,40	1,04	1,29	0,85	0,68	0,62	0,63

Figura 20 – Dados obtidos através de Montgomery (2003).

Fonte: Montgomery (2003).

ANEXO B: FUNÇÕES E PROGRAMAÇÃO

Função desenvolvida através do software R para que o usuário possa gerar o gráfico da carta de controle via ACO, podendo colocar uma amostra para validação, gráfico de curvas características e comprimento médio de sequência, além de disponibilizar a média, variância e variância amostral conforme método selecionado:

```
funcao=function(data,method,validation)
{
# Usuário necessita entrar com os dados finais, ou seja, com os dados ordenados contendo nas
linhas as amostras e nas colunas os elementos da amostra:

#Validação dos Métodos informados

if (method!=1)
{
  if (method!=2)
  {
    if (method!=3)
      stop("Método não informado ou inesistente")
  }
}

#Elementos auxiliares para a função

n=ncol(data)           #Número de Elementos da Amostra
l=nrow(data)           #Número de Amostras
mu=c()                 #Vetor para calculo da estatística de ordem
vi=c()                 #Vetor para calculo da estatística de ordem

#Estatística de Ordem

for(i in 1:n)
{
  f2=function(y,n,i)
  y*(factorial(n)/(factorial(i-1)*factorial(n-i)))*((pnorm(y))**(i-1))*((1-pnorm(y))**(n-i))*dnorm(y)
  mu_aux=integrate(f2,-Inf,Inf,n,i)$value
  mu[i]=integrate(f2,-Inf,Inf,n,i)$value
  f3=function(y,n,i)
  (y**2)*(factorial(n)/(factorial(i-1)*factorial(n-i)))*((pnorm(y))**(i-1))*((1-
pnorm(y))**(n-i))*dnorm(y)
  s=integrate(f3,-Inf,Inf,n,i)$value
  vi[i]=s-mu_aux**2
}

#Cálculo para método 1

if (method==1)
{
  Media=mean(data)           #Calculando a Média para o método 1
  Variancia=var(as.numeric(data)) #Calculando a Variância para o método 1
  Variancia_Amostral=(1/n)*Variancia-(1/n^2)*sum((apply(data,2,mean)-Media)^2)
  #Calculando a Variância Amostral para o método 1
}
```

```

S=sqrt(Variancia) #Desvio Padrão para o método 1
Sbarra=sqrt(Variancia_Amostral) #Desvio Padrão Amostral para o método 1
a=apply(data,1,mean) #Pontos a serem plotados para o método 1
}

```

#Cálculo para método 2

```

if (method==2)
{
Media=sum(apply(data,2,mean)/vi)/sum(1/vi) #Calculando a Média para o método 2
Variancia = (1/(l*n-1))*sum((data-Media)^2) #Calculando a Variância para o método 2
Variancia_Amostral= Variancia/sum(1/vi) #Calculando a Variância Amostral para o
método 2
S=sqrt(Variancia) #Desvio Padrão para o método 2
Sbarra=sqrt(Variancia_Amostral) #Desvio Padrão Amostral para o método 2
a=data%*%(1/vi)/sum(1/vi) #Pontos a serem plotados para o método 2
}

```

#Cálculo para método 3

```

if (method==3)
{
Media = 1/(l*n)*sum(data) #Calculando a Média para o método 3
Variancia=(n*l-1+sum((mu/n)^2))^(-1)*sum((data-Media)^2) #Calculando a Variância para o método 3
Variancia_Amostral=(1/n)*Variancia-(1/n^2)*sum((apply(data,2,mean)-Media)^2) #Calculando a Variância Amostral para o método 3
S=sqrt(Variancia) #Desvio Padrão para o método 3
Sbarra=sqrt(Variancia_Amostral) #Desvio Padrão Amostral para o método 3
a=apply(data,1,mean) #Pontos a serem plotados para o método 3
}

```

#Calculando os Limites

```

s1=Media+1*sqrt(Variancia_Amostral) #Calculo 1º limite superior
s2=Media+2*sqrt(Variancia_Amostral) #Calculo 2º limite superior
s3=Media+3*sqrt(Variancia_Amostral) #Calculo 3º limite superior
i1 =Media-1*sqrt(Variancia_Amostral) #Calculo 1º limite inferior
i2 =Media-2*sqrt(Variancia_Amostral) #Calculo 2º limite inferior
i3 =Media-3*sqrt(Variancia_Amostral) #Calculo 3º limite inferior

```

#Verificação se o usuário identificou data para validação

```

if (missing(validation))
grafico = a
else
grafico = c(a,apply(validation,1,mean))

```

#Calculando limites das margens para o gráfico

```

minimo=min(c(grafico,i3))-2*(sqrt(Variancia_Amostral)) #Margem Inferior
maximo=2*(sqrt(Variancia_Amostral))+max(c(grafico,s3)) #Margem Superior
extensao=length(grafico)
cenario=length(a)

```

#Gerando Gráficos

```

par(mfrow=c(2,2))
plot(grafico,type="b",ylim=c(minimo,maximo),xlab='Número de Amostras', ylab='Média Amostral',col = "blue", bg = "blue", pch = 21,axes=FALSE)

```

```

box()
title(main=paste("Carta Controle ACO - Método ",method),col.main="blue",cex.main=1.2,font.main= 4)
axis(2, col.axis="blue", las=1)
axis(1, col.axis="blue", las=1,at = seq(0, extensao, by = 2))

```

#Inserindo as Linhas de Limite

```

abline(h=Media)           #Inserindo a linha Média
abline(h=s1, col=8, lty=2) #Inserindo linha do 1º limite superior
abline(h=s2, col=8, lty=5) #Inserindo linha do 2º limite superior
abline(h=s3, col=2, lty=2) #Inserindo linha do 3º limite superior
abline(h=i1, col=8, lty=2) #Inserindo linha do 1º limite inferior
abline(h=i2, col=8, lty=5) #Inserindo linha do 2º limite inferior
abline(h=i3, col=2, lty=2) #Inserindo linha do 3º limite inferior

```

#Em caso de amostra de validação inserir linha vertical e texto identificando a validação

```

if (!missing(validation))
{
  abline(v=cenario+0.5, col=8, lty=2)
  w=cenario+ncol(validation)/2
  z=minimo+sqrt(Variância_Amostral)
  text(w,z,"Validação",col="red")
}

```

#Curva Característica

```

c4=(4*(1-1))/((4*1)-3)
k1=function(k) 1-(pnorm(3-k*S/Sbarra,lower.tail=F)+pnorm(-3-k*S/Sbarra))
curve(k1,0,3.5, col="red",xlab='Alteração na Média do Processo', ylab='Probabilidade de Não
Detecção', lty=4, , lwd=2)
title(main=paste("Curva Característica de Operação n =",n),col.main="blue",cex.main=1.2,font.main=
4)

```

#Comprimento Médio de Sequência

```

k1=function(k) 1/(pnorm(3-k*S/Sbarra,lower.tail=F)+pnorm(-3-k*S/Sbarra))
plot(-1,type="b",ylim=c(0,150),xlim=c(0,1.5),xlab='Alteração na Média do Processo',ylab=
'Comprimento Médio da Sequência')
curve(k1,0,1.5, add=TRUE, col="red", lty=2, , lwd=2)
title(main=paste("Comprimento Médio de Sequência n =",n),col.main="blue",cex.main=1.2,font.main=
4)

```

#lista de variáveis disponibilizada

```

List (
  'Media'           = Media,
  'Variância'       = Variância,
  'Variância Amostral' = Variância_Amostral
)
}
funcao(data=X,method=Y,validation=Z)

```


ANEXO C: ELABORAÇÃO DA AMOSTRA

Função desenvolvida no software R para gerar as amostras para elaboração dos cenários:

```

amostra=function(Elementos,Amostras)
{
  require(MASS)          #Pacote
  set.seed(0729)
  Sigma=matrix(c(1.23,0.79,0.79,0.83),2,2)
  qtde=Elementos^2*Amostras
  base=mvrnorm(n=qtde,mu=c(1.06,115.59),Sigma)
  Amostra=matrix(ncol=2,)[-1,]
  a=1
  b=Elementos
  for(t in 1:Amostras)
  {
    Repositorio=matrix(ncol=2,nrow=Elementos)
    for(i in 1:Elementos)
    {
      aux=base[a:b,]
      Repositorio[i,2]=aux[i,2]
      colnames(aux) <- c("A","B")
      aux2 = aux[order(aux[,"A"]),]
      Repositorio[i,1]=aux2[i,2]
      a=a+Elementos
      b=b+Elementos
    }
    Amostra=rbind(Amostra,Repositorio)
    ACO=matrix(Amostra[,1],ncol=Elementos,byrow=T) #Amostra via ACO
    AAS=matrix(Amostra[,2],ncol=Elementos,byrow=T) #Amostra via ACO
  }
  list(
    'ACO'=ACO,
    'AAS'=AAS,
    'Base'=base,
    'qtde'=qtde)
}
c=amostra(Elementos=10,Amostras=30)

```