

**SCAP: “STATISTICAL CONTROL FOR AUTOCORRELATED PROCESSES” - Módulo AI: ATTRIBUTE INSPECTION**

versão 1.0 - 2003

**UM SOFTWARE NA ÁREA DE ESTATÍSTICA INDUSTRIAL  
INTEGRADO AO SOFTWARE ESTATÍSTICO  
“MINITAB FOR WINDOWS”**

*MANUAL DO USUÁRIO*

*APOIO FINANCEIRO: CNPq e FAPEMIG*

**AUTORES: SUELI APARECIDA MINGOTI (\*)  
JÚLIA PINTO DE CARVALHO (\*\*)**

*AGRADECIMENTO*

**OS AUTORES AGRADECEM AO CNPQ E A FAPEMIG PELO  
APOIO FINANCEIRO QUE POSSIBILITOU A PRODUÇÃO DO  
SOFTWARE “SCAP”**

**(\*) Ph.. D em Estatística - Profa. Adjunta do Departamento de  
Estatística da Universidade Federal de Minas Gerais - UFMG**

**(\*\*) Bacharelanda do Curso de Estatística da UFMG e Bolsista de  
Iniciação Científica da Fapemig durante o desenvolvimento do  
projeto de pesquisa que gerou o SCAP.**

2003

**SCAP: “STATISTICAL CONTROL FOR AUTOCORRELATED PROCESSES” - Módulo AI: ATTRIBUTE INSPECTION**

**versão 1.0 - 2003**

**UM SOFTWARE NA ÁREA DE ESTATÍSTICA INDUSTRIAL  
INTEGRADO AO SOFTWARE ESTATÍSTICO**

**“MINITAB FOR WINDOWS”**

**AUTORES: SUELI APARECIDA MINGOTI (\*)  
JÚLIA PINTO DE CARVALHO (\*\*)**

## **Índice**

<b>Capítulo 1: O Software SCAP - Módulo AI (Attribute Inspection).....</b>	<b>02</b>
<b>Capítulo 2: Os Modelos Binomial Generalizado e Cadeia de Markov no Monitoramento de Processos Autocorrelacionados - Inspeção por Atributos.....</b>	<b>04</b>
<b>Capítulo 3: Instruções e Exemplos de Aplicação do Programa SCAP - Módulo AI.....</b>	<b>24</b>
<b>Referências Bibliográficas.....</b>	<b>66</b>

---

# CAPÍTULO 1

## *O Software SCAP - Módulo AI (Attribute Inspection)*

### 1.0 Introdução

O *software SCAP- módulo AI* foi desenvolvido como parte de um projeto de pesquisa parcialmente financiado pelo *CNPq* e *FAPEMIG*, instituições brasileiras de fomento à pesquisa. Sua estrutura foi concebida de modo a permitir que o usuário execute facilmente a análise estatística de dados que sejam originados de processos industriais autocorrelacionados no que se refere a inspeção de atributos. Este programa funciona como um complemento do popular *software* estatístico “*Minitab for Windows*”. Tal opção pela criação do *SCAP* como um complemento de outro *software* deve-se principalmente à grande difusão, baixo custo e facilidade de uso do “*Minitab for Windows*”. *SCAP* é um programa muito simples de ser usado, funcionando de forma totalmente interativa. As perguntas são feitas ao usuário, passo a passo, em uma seqüência lógica de análise. A utilização do *SCAP* requer do usuário apenas conhecimentos básicos sobre o “*Minitab for Windows*” mas não exige conceitos de programação. Por funcionar a partir do “*Minitab*”, o usuário, além dos recursos disponíveis no *SCAP- módulo AI*, tem a grande vantagem de desfrutar de todos os outros recursos importantes para uma análise estatística, sem precisar sair do *software*. Em termos do espaço ocupado no disco rígido, o *SCAP* é bem econômico. Tecnicamente, para se obter um bom desempenho na utilização do *SCAP, módulo AI*, é necessário que o usuário tenha à sua disposição a versão 11 (ou superior) do *software Minitab for Windows*®. O programa requer no mínimo um computador com processador Pentium II e 32 MB de memória RAM. Cabe salientar que no sistema *Windows 95* (ou superior), a capacidade de armazenamento de dados é limitada apenas pela quantidade de memória disponível. Todo o armazenamento de dados é feito na própria planilha do *Minitab (worksheet)*. Quanto ao espaço necessário para armazenamento dos programas, o programa *SCAP-módulo AI* necessita de 60 Kbytes e o manual em formato .pdf de 494 Kbytes.

No *software SCAP - módulo AI* o usuário poderá estimar a proporção de defeitos (ítems não conformes) do processo, a correlação existente entre as unidades amostrais em relação aos atributos de qualidade que estão sendo avaliados bem como traçar os limites de controle tanto para a proporção de defeituosos quanto para o número de ítems defeituosos do processo. Os

---

modelos estatísticos implementados no *SCAP*, *módulo AI*, são o da Binomial Generalizada (Madsen, 1993) e da Cadeia de Markov (Bhat & Lal, 1990). A estimação do coeficiente de correlação do processo no caso do modelo Binomial Generalizada pode ser feita por meio de três formas: usando o estimador clássico de Madsen (1993), o estimador Não-Paramétrico proposto por Mingoti (2002) e os estimadores Bayesianos propostos por Mingoti (2002). Para a Cadeia de Markov o processo de estimação usado é aquele apresentado em Bhat & Lal (1990). O programa *SCAP* aceita a entrada de dados individuais, isto é, o resultado individual da inspeção de cada item amostrado do processo, ou dados agrupados, isto é, o número total de defeitos observados numa amostra (subgrupo racional) de itens do processo. No caso da Cadeia de Markov somente é permitido a entrada de dados individuais, enquanto que para o modelo Binomial Generalizado se dispõe das duas opções. No Capítulo 2 deste manual, apresentamos uma breve introdução sobre estes dois modelos estatísticos usados no monitoramento, via inspeção de atributos, de processos autocorrelacionados e os estimadores que estão implementados no *SCAP módulo AI*, enquanto que no Capítulo 3 são encontrados exemplos de aplicação.

---

## CAPÍTULO 2

### *Os Modelos Binomial Generalizado e Cadeia de Markov no Monitoramento em Processos Autocorrelacionados - Inspeção por Atributos*

#### 2.0 Introdução

Vários processos são monitorados através da observação do número de itens não-conformes gerados pelo processo ou através da proporção de itens não-conformes. Neste caso, temos a situação clássica do tipo de resposta binária:  $X=0$  ("passa") e  $X=1$  ("não passa"), sendo portanto o "número total de defeituosos" do processo ou de uma amostra do processo visto como tendo uma distribuição Binomial. A metodologia estatística clássica para estas situações é a construção de gráficos de controle cujos limites são determinados através da aproximação da distribuição Binomial pela distribuição Normal, o que pode ser feito sob determinadas condições, ou através do uso da distribuição de Poisson (Montgomery, 1991).

Para processos com um número de defeituosos bem próximo de zero os gráficos de controle fundamentados nesta metodologia clássica têm pouca utilização prática uma vez que devido ao baixo valor da proporção de defeituosos as correções que são efetuadas nos limites de controle via distribuição normal ou via distribuições exatas de Binomial e Poisson criam dois problemas: no primeiro caso a ocorrência de "alarmes falsos" aumenta consideravelmente e no segundo os limites de controle são sempre próximos de zero ou 1, o que inviabiliza o uso prático dos gráficos. Alguns trabalhos interessantes que abordam formas alternativas de monitorar o processo neste caso, são: Goh & Xie (1995), Xie & Goh (1992), Xie *et. al.* (1995), Lai, Xie & Govindaraju (2000). Entre as propostas algumas as mais usadas têm sido:

- (1) o uso da estatística "número de itens conformes testados até a aparição do primeiro item não-conforme", o que leva automaticamente para a aplicação da distribuição geométrica. Esta estatística é conhecida como CCC. Algumas referências importantes são Calvin (1983), Bourke (1991), Nelson (1994), Xie & Goh (1992) e Quesenberry (1995).
- (2) a observação do padrão de comportamento das ocorrências dos itens não-conformes em amostras consecutivas do processo (Goh,1989 ; Lucas, 1989).

---

A teoria clássica de controle de qualidade assume que as unidades amostrais (ou seja os itens produzidos) são independentes em relação ao estado de ser "conforme" ou "não conforme". No entanto, esta suposição nem sempre é válida. Em Alwan e Roberts (1995) é mostrado que a correlação entre unidades amostrais aparece com uma frequência bem maior do que se pensa. Na realidade a correlação entre as unidades amostrais no que se refere a variável resposta em estudo tende a 1 quando o intervalo entre inspeções tende a zero. Deste modo, uma das alternativas sugeridas para tratamento da correlação é um maior espaçamento entre as unidades amostrais (ou grupos racionais) selecionadas para inspeção. No entanto, esta alternativa pode ocasionar grandes perdas financeiras uma vez que pelo fato de se espaçar demais as inspeções demora-se mais para se detectar um problema no processo, ou uma falta de controle no processo.

Outras formas de se tratar a existência de correlação é através do monitoramento do processo utilizando-se modelos estatísticos que permitem a introdução da correlação. A primeira delas refere-se ao trabalho de Bhat & Lal (1990) no qual os autores desenvolveram os limites de controle para monitoramento de atributos de processos correlacionados usando a teoria de Cadeia de Markov (Broadbent,1996), sendo que uma modificação foi sugerida por Lai, Xie & Govindaraju (2000) para processos com baixa fração de defeituosos. Uma outra proposta é aquela sugerida por Lai, Govindaraju & Xie (1998). Neste caso, aos autores sugerem o uso do modelo Binomial Generalizado, ou equivalentemente Binomial Correlacionado de Madsen (1993) para monitorar o processo, sendo que sob este modelo a estatística "*número de itens conformes testados até a aparição do primeiro item não-conforme*" teria uma distribuição geométrica correlacionada. Em seu artigo Lai, Govindaraju & Xie (1998) mostraram os resultados de um estudo sobre o erro do tipo II (ou seja dizer que o processo "está sob controle" quando na realidade "não está") em situações nas quais o modelo probabilístico Binomial Generalizado é usado para o tratamento do número de itens não-conformes do processo. Neste trabalho a correlação entre as unidades amostrais é suposta ser uma constante igual para todas as unidades e o estimador proposto para o coeficiente de correlação é o de Madsen (1993). Também, os autores analisam apenas situações nas quais apenas uma característica de qualidade é avaliada em cada passo e as várias características do processo seriam monitoradas independentemente.

A diferença entre estas duas metodologias é bem clara. No caso de Cadeias de Markov presume-se a inspeção serial das unidades do processo, e portanto temos que manter a

informação sobre o estado de cada item (isto é se é "não conforme" ou "conforme"), na sequência exata em que foi inspecionado. Já no modelo Binomial Generalizado a informação necessária para sua implementação é aquela relacionada com o grupo racional, ou seja, amostras de itens do processo são inspecionadas e em cada amostra apenas o número de "não-conformes" precisa ser guardado.

Em ambas as metodologias, tanto na Cadeia de Markov quanto no modelo Binomial Generalizado, a estimação do coeficiente de correlação tem um papel fundamental. Bhat & Lal (1990) mostram como estimar o coeficiente de correlação no caso de Cadeias de Markov. Lai, Xie & Govindaraju (2000) mostraram propriedades do estimador proposto e a construção de intervalos de confiança para a correlação teórica do processo. Para o modelo Binomial Generalizado tem-se o estimador de correlação proposto por Madsen (1993) e os estimadores não-paramétricos e Bayesianos propostos por Mingoti (2002). Alguns trabalhos interessantes sobre processos autocorrelacionados são: Sampath, Kumar & Rajarshi (1987), MacShane & Turnbull (1991), Stimson & Mastrangelo (1996) e Neves & Mingoti (2001).

## 2.1 Processos Autocorrelacionados: Modelo de Cadeia de Markov

No modelo de Cadeia de Markov (Bhat & Lal, 1990) usado para modelar a inspeção industrial, temos duas probabilidades importantes:

- A : probabilidade de que o item inspecionado seja "não-conforme" quando o item precedente for "conforme"
- B: probabilidade de que o item inspecionado seja "conforme" quando o item precedente for "não conforme".

A inspeção de itens deve ser feita de forma contínua de modo a que se tenha informação necessária para estimação das probabilidades de transição da Cadeia. Basicamente, tem a seguinte situação: Seja,  $\{ Y_n, n = 0, 1, 2, \dots \}$  um Cadeia de Markov com dois estados: 0 (item é conforme) e 1 (item é não conforme) sendo que  $Y_n$  caracteriza o estado do n-ésimo item inspecionado. Neste caso as probabilidades de transição serão dadas por:

$$p_{ij} = Prob[Y_{n+1} = j / Y_n = i] \quad n = 0, 1, 2, \dots; i, j = 0, 1 \quad (1)$$

e a matriz de probabilidades de transição a um passo será:

$$P = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} \quad (1 \geq a, b \geq 0) \quad (2)$$

sendo  $Y_0$  o estado inicial da Cadeia. Neste caso, se considerarmos as probabilidades de transição em  $n$  passos, isto é:

$$p_{ij}^{(n)} = Prob[Y_{n+1} = j / Y_0 = i] \quad n = 1, 2, \dots; i, j = 0, 1 \quad (3)$$

quando  $|1-a-b| < 1$ , as probabilidades limites da Cadeia isto é,

$$\pi_j = \lim(n \rightarrow \infty) p_{ij}^{(n)}, \quad j = 0, 1 \quad (4)$$

são dadas por:

$$\pi_0 = \frac{b}{a+b} \quad e \quad \pi_1 = \frac{a}{a+b} \quad (5)$$

A probabilidade  $\pi_1$  é entendida como a probabilidade de que num grande volume de produção, um item do processo seja defeituoso ( $p$ ), ou seja a fração de defeituosos do processo. Em Bhat & Lal (1990) é mostrado (p. 182) que sob esta condição de equilíbrio da Cadeia a correlação de ordem  $j$ , isto é,

$$\rho_j = corr(Y_n, Y_{n+j})$$

é dada por:

$$\rho_j = [1 - (a + b)]^j = \rho^j, \quad j = 1, 2, \dots; \quad e \quad \rho = 1 - (a + b) \quad (6)$$

Um fato interessante e que vai de encontro ao que se espera em problemas práticos é que a correlação entre unidades amostrais decresce exponencialmente à medida que a distância entre as unidades amostrais aumenta na linha de produção e inspeção.

Para ser uma Cadeia admissível os parâmetros  $p$  e  $\rho$  precisam satisfazer as condições:

$$1 - \min\left[\frac{1}{p}, \frac{1}{1-p}\right] < \rho < 1 \quad (7)$$

ou de outro modo,

$$\max\left[0, \frac{-\rho}{1-\rho}\right] < p < \min\left[1, \frac{1}{1-\rho}\right] \quad (8)$$

o que significa dizer que não se permite situações nas quais a correlação é igual a 1 ou -1, casos nos quais a probabilidade  $p$  de defeituosos seria igual a 1 ou 0. Assim, a matriz de transição  $P$  dada em (1) pode então, ser expressa por:

$$P = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} = \begin{bmatrix} 1 - (p(1-\rho)) & p(1-\rho) \\ (1-p)(1-\rho) & p + \rho(1-p) \end{bmatrix} \quad (9)$$

Por exemplo,  $p = 0,05$  e  $\rho = 0,3$  a matriz de transição será:

$$P = \begin{bmatrix} 0,965 & 0,035 \\ 0,665 & 0,335 \end{bmatrix} \quad (10)$$

Este seria um processo no qual se um item produzido é defeituoso o próximo teria probabilidade igual a 0,335 de também ser defeituoso e 0,665 de ser não defeituoso.

Quando o coeficiente de correlação é igual a zero temos a situação de ensaios de Bernoulli independentes. O modelo de Markov permite que haja correlações negativas, no entanto não permite que haja correlação igual a 1 ou -1.

Para outros detalhes sobre Cadeias de Markov sugerimos a leitura de Karlin e Taylor (1975).

### 2.1.1 Estimação dos Parâmetros do Modelo de Cadeia de Markov

Os parâmetros do processo devem ser estimados quando o mesmo está sob controle. Estimadores de máxima verossimilhança de  $p$  e  $\rho$  são dados por (Bhat & Lal, 1990) como segue:

$$\hat{p}_{ij} = \frac{n_{ij}}{n_{i0} + n_{i1}}, \quad i, j=0,1$$

$$\hat{p} = \frac{\hat{p}_{01}}{\hat{p}_{01} + \hat{p}_{10}} \quad (11)$$

$$\hat{\rho} = 1 - (\hat{p}_{01} + \hat{p}_{10})$$

A distribuição de  $\hat{\rho}$  é aproximadamente normal com variância dada por (Lai, Xie & Govindaraju, 2000):

$$Var(\hat{\rho}) = \frac{1}{n} (a^2(1-a) + b^2(1-b)) \frac{(a+b)}{ab} = \frac{1}{n} \left[ \frac{(1-\rho) [p(1-p)(1-3\rho) + \rho]}{p(1-p)} \right] \quad (12)$$

### 2.1.2 Limites de Controle na Cadeia de Markov: Aproximação Normal

Considere a variável  $X_n$ , número de defeituosos observados em  $n$  itens inspecionados. A distribuição de probabilidades exata de  $X_n$  é um pouco complicada como pode ser observado em Bhat & Lal (1990; 1988). Quando a Cadeia de Markov tem a estrutura dada como na seção 2.1 então, pode ser mostrado que a esperança matemática e a variância de  $X_n$  são dadas por:

$$E(X_n) = n p \tag{13}$$

$$Var(X_n) = n p(1-p) + 2 p(1-p) \frac{\rho}{1-\rho} \left( n - \frac{1-\rho^n}{1-\rho} \right)$$

Usando a aproximação normal, o que é viável para o caso da Cadeia de Markov em questão, os limites de controle para a fração de defeituosos do processo seriam dados respectivamente por:

$$LIC = p - \frac{k}{n} k \sqrt{Var(X_n)} \quad ; \quad LSC = p + \frac{k}{n} k \sqrt{Var(X_n)} \quad ; \quad LM = n p \tag{14}$$

onde *LIC* e *LSC* representam respectivamente os limites inferior e superior de controle e *LM* o limite médio.

Se tivermos estimativas de  $p$  e  $\rho$  temos automaticamente os valores dos limites de controle do processo. Isto de um certo modo é interessante dado que o usuário de controle de qualidade, em geral, já está habituado com os gráficos de controle de Shewhart (Montgomery, 1991) que têm como base a distribuição normal para construção dos limites de controle em caso de processos não correlacionados. Assim, mesmo na presença de correlação o usuário, ao optar pela modelagem via Cadeia de Markov, ainda teria a opção de se manter dentro da classe da distribuição normal.

## 2.2 Processos Autocorrelacionados: Modelo Binomial Generalizado

Suponha que uma amostra de  $n$  itens do processo tenha sido inspecionada. Para cada item na amostra defina a variável:

$$X_i = \begin{cases} 1 & \text{se o } i\text{-ésimo item for não conforme} \\ 0 & \text{se o } i\text{-ésimo item for conforme} \end{cases}$$

Suponha que os itens sejam correlacionados da seguinte forma:

$$\text{corr}(X_i, X_j) = \rho, \forall i \neq j, \rho > 0 \quad (15)$$

Seja  $Z_n$  o número total de itens não conformes observados na amostra de tamanho  $n$  do processo. Se considerarmos que a sequência  $\{X_i, i=1,2,\dots\}$  é permutável (exchangeable) (Berger, 1985 pg.105), a distribuição de probabilidades exata da variável  $Z_n$  é Binomial Generalizada proposta por Madsen (1993), isto é,

$$P[Z_n = k] = \begin{cases} \rho(1-p) + (1-\rho)(1-p)^n, & \text{se } k = 0 \\ (1-\rho) \binom{n}{k} p^k (1-p)^{n-k}, & \text{se } 1 \leq k \leq n-1 \\ \rho p + (1-\rho)p^n, & \text{se } k = n \end{cases} \quad (16)$$

Esta distribuição se transforma na Binomial clássica quando o coeficiente de correlação é igual a zero. Pode ser mostrado que a esperança e variância de  $Z_n$  são dadas por:

$$\begin{aligned} E(Z_n) &= np \\ \text{Var}(Z_n) &= np(1-p) + n(n-1)p(1-p)\rho = np(1-p)[1 + (n-1)\rho] \end{aligned} \quad (17)$$

Um modelo como este permite também a observação do processo via definição da variável  $Y$ : o número total de inspeções feitas até que se encontre o primeiro item não conforme, condicionado ao estado inicial  $X_0 = 1$ . Sob o modelo Binomial Generalizado de Madsen, a distribuição de probabilidades exata de  $Y$  é uma distribuição Geométrica correlacionada, isto é,

$$P[Y = i] = \begin{cases} \rho + (1-\rho)p, & \text{se } i = 1 \\ (1-\rho)p(1-p)^{i-1}, & \text{se } i \geq 2 \end{cases} \quad (18)$$

que se reduz a uma distribuição Geométrica clássica quando a correlação é igual a zero. Esta variável permite a construção dos gráficos do tipo  $RL_1$  (comprimento da corrida entre dois itens não conformes sucessivos) quando a inspeção é feita 100% e dos gráficos  $RL_2$  que trabalha com a variável  $W$  que é a soma das duas  $RL_1$  mais recentes (Bourke,1991). Neste caso, tem-se que:

$$E(Y) = \rho + \frac{(1-\rho)}{p} \quad \text{e} \quad E(W) = 2 E(Y) \quad (19)$$

O interessante do modelo Binomial Generalizado é a não necessidade da inspeção contínua do processo ou a necessidade de se guardar o resultado da inspeção de itens individuais, o que é preciso para a implementação da Cadeia de Markov. No entanto, pelo fato de se ter uma estrutura de correlação como na equação (15) não é possível aproximar a distribuição dada em (16) pela distribuição Normal. Deste modo, qualquer cálculo de probabilidades exigirá o uso da distribuição exata (dada em (16)) ou da aproximação de Poisson, que será dada por:

$$\begin{aligned} P[Z_n = 0] &= \rho + (1-\rho) e^{-\lambda} \\ P[Z_n = k] &= (1-\rho) \frac{e^{-\lambda} \lambda^k}{k!}, \quad k \in \{1, 2, 3, \dots\} \end{aligned} \quad (20)$$

quando  $n \rightarrow \infty$  e  $p \rightarrow 0$ , de modo que  $np = \lambda$  permaneça constante.

### 2.2.1 Estimação Clássica dos Parâmetros do Modelo Binomial Generalizado

Suponha que tenhamos  $m$  amostras aleatórias de tamanho  $n$  da variável  $Z_n$ . Vamos denotar os valores observados de  $Z_n$  para cada amostra  $i$  por  $Z_{n,i}$ ,  $i=1,2,\dots,m$ ,  $m>2$ . Então, o estimador do coeficiente de correlação  $\rho$  sugerido por Madsen (1993) é dado por:

$$\hat{\rho} = \frac{[s^2 - np(1-p)]}{[n(n-1)p(1-p)]} \quad (21)$$

$$\text{onde } s^2 = \frac{\sum_{i=1}^m (Z_{n,i} - \bar{Z}_n)^2}{m} \quad \text{sendo } \bar{Z}_n = \frac{\sum_{i=1}^m Z_{n,i}}{m}$$

Este estimador é obtido pelo método dos momentos (Garthwaite *et. al.*,1995). Caso a fração de não conformes  $p$  não esteja previamente especificada, ela poderá ser estimada usando a fração amostral de "não conformes" observadas dentre as  $m$  amostras observadas isto é,  $\bar{Z}_n$ .

A proposta original de Madsen sugere que se use a variância amostral  $s^2$  calculada com o divisor  $m$ . No entanto, uma outra alternativa é utilizar-se o estimador não viciado cujo divisor é  $(m-1)$ . Este estimador, o qual chamamos de Madsen modificado se mostrou melhor que a proposta original (21) em algumas simulações que foram feitas.

Se as  $m$  amostras não tiverem o mesmo tamanho, o estimador do coeficiente de correlação terá que ser redefinido de acordo com Madsen (1993), ou então, calculado utilizando-se uma prática comum em gráficos de controle para dados não correlacionados, isto é,  $\hat{\rho}$  seria definido como:

$$\hat{\rho} = \frac{[s^2 - \bar{n} p(1-p)]}{[\bar{n}(\bar{n}-1)p(1-p)]} \quad (22)$$

$$\text{onde } s^2 = \frac{\sum_{i=1}^m (Z_{n,i} - \bar{Z}_n)^2}{m} \quad \text{sendo } \bar{Z}_n = \frac{\sum_{i=1}^m Z_{n,i}}{m}, \quad \bar{n} = \frac{\sum_{i=1}^m n_i}{m}$$

onde  $n_i$  representa o número de ítems inspecionados na  $i$ -ésima amostra.

É importante observar que caso tenhamos uma amostra contendo  $n$  valores iguais a estimativa do coeficiente de correlação da forma como definido em (21) será indefinido fato não observado no artigo original de Madsen (1993) e nem no artigo de Lai, Govindaraju & Xie (1998). É natural pensarmos que se todos os valores amostrais são iguais teríamos uma associação perfeita. Portanto, nestes casos decidimos estabelecer a estimativa  $\hat{\rho}=1$ . Além disso,

pelo fato do estimador de Madsen depender da variância amostral  $s^2$  ele será muito influenciado por valores amostrais extremos (*outliers*). Estudos simulados mostram que a qualidade do estimador de Madsen depende da região em que o verdadeiro valor de  $\rho$  se encontra, podendo resultar em estimativas com erros bem acentuados em algumas situações. Caso o pesquisador tenha alguma informação prévia confiável sobre  $\rho$  sugerimos que ele procure utilizar os estimadores Bayesianos que serão descritos na seção 2.2.2. Caso o tamanho da amostra seja maior ou igual a 100, o pesquisador poderá utilizar o estimador clássico Multinomial que será apresentado na seção 2.2.2.1 obtendo resultados mais precisos que o estimador de Madsen.

## 2.2.2 Estimadores Alternativos para o Coeficiente de Correlação do Modelo Binomial Generalizado

A alternativa de estimação de Madsen para o coeficiente de correlação é uma das poucas existentes na literatura para modelos Binomiais correlacionados que é de fácil utilização. No entanto, Mingoti (2002) propôs outros estimadores tanto para o caso em que temos acesso apenas a informação final de quantos itens conformes e não conformes apareceram na amostra quanto para o caso nos quais se tem acesso à sequência de "0" e "1" observada na inspeção de produtos, o que neste manual é chamado de observações individuais.

### 2.2.2.1 Estimador Multinomial (Mingoti, 2002)

Suponha que tenhamos  $m$  amostras independentes da distribuição Binomial Generalizada  $\text{Bin}(n, p, \rho)$ . Seja  $Y_k$  a variável aleatória definida como o "número de amostras que caem na classe  $k$ , ou seja que apresentam  $k$  itens "não conformes". Então,  $Y_k$  tem uma distribuição Multinomial com parâmetros  $p_k$ , dados segundo o modelo da Binomial Generalizado. Seja  $f_k$  o número de amostras, dentre as  $m$  observadas, que apresentam o valor  $k$ , isto é, apresentam  $k$  itens não conformes,  $k=0,1,\dots,n, m>2$

Então, podemos estimar os parâmetros da distribuição Multinomial pelo método de máxima verossimilhança (Garthwaite *et. al.*, 1995) da seguinte forma:

$$\hat{p}_k = \frac{f_k}{m} \quad (23)$$

Considerando-se que:

$$\sum_{k=1}^{n-1} \hat{p}_k \approx (1-\rho) \sum_{k=1}^{n-1} \binom{n}{k} p^k (1-p)^{n-k} = (1-\rho) \{1 - [(1-p)^n + p^n]\} \quad (24)$$

tem-se,

$$\left[ \sum_{k=1}^{n-1} \hat{p}_k \right] \cdot \left[ \frac{1}{1 - \{1 - [(1-p)^n + p^n]\}} \right] = (1-\rho) \quad (25)$$

e assim uma estimativa para o coeficiente de correlação seria:

$$\hat{\rho}_I = 1 - \frac{\sum_{k=1}^{n-1} \hat{p}_k}{\{1 - [(1-p)^n + p^n]\}} \quad (26)$$

Caso o valor do parâmetro  $p$  não esteja pré-especificado à priori ele poderá ser estimado pela fração amostral de itens não conformes, isto é,

$$\hat{p} = \frac{\text{númerototal de itens não conformes}}{n \times m} \quad (27)$$

Uma forma alternativa de se expressar  $\hat{\rho}_I$  e que é mais fácil de ser implementada computacionalmente, é dada por:

$$\hat{\rho}_I = \frac{\hat{p}_0 + \hat{p}_n - (1-p)^n - p^n}{\{1 - [(1-p)^n + p^n]\}}, \quad 0 < p < 1 \quad (28)$$

Para valores de  $\hat{p}$  iguais a zero ou 1 o coeficiente de correlação  $\hat{\rho}_I$  será definido como igual a 1. Se  $\hat{p}_0$  e  $\hat{p}_n$  forem iguais a zero simultaneamente,  $\hat{\rho}_I$  será definido como zero, assim como em qualquer situação onde o valor de (28) for negativo. Estudos simulados mostram que o estimador Multinomial é uma boa alternativa para amostras de tamanho grande ( $n \geq 100$ ) superando o estimador de Madsen em termos de precisão.

### 2.2.2.2 Estimador Bayesiano I (Mingoti, 2002)

Estimadores Bayesianos (Berger, 1985) podem ser construídos se considerarmos que  $\rho$  é uma variável aleatória e tem uma distribuição à priori. Considerando ainda o modelo Binomial Generalizado (16) vamos supor que  $\rho$  tenha uma distribuição "à priori" Beta  $(\alpha, \beta)$ ,  $\alpha > 0$ ,  $\beta > 0$ , isto é,

$$g(\rho) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \rho^{\alpha-1} (1-\rho)^{\beta-1}, \quad 0 < \rho < 1 \quad (29)$$

sendo a função densidade de  $Z_n$  dado  $p$  e  $\rho$  será:

$$P[Z_n = k / p, \rho] = \begin{cases} \rho(1-p) + (1-\rho)(1-p)^n, & \text{se } k = 0 \\ (1-\rho) \binom{n}{k} p^k (1-p)^{n-k}, & \text{se } 1 \leq k \leq n-1 \\ \rho p + (1-\rho)p^n, & \text{se } k = n \end{cases} \quad (30)$$

Calculando-se então, a esperança condicional de  $\rho$  de acordo com a distribuição à posteriori de  $\rho$  dado  $Z_n$  e  $p$ , obtemos os seguintes resultados:

$$\begin{aligned}
E[\rho / Z_n=0] &= \frac{\alpha [(1-p)(\alpha+1) + \beta(1-p)^n]}{[\alpha(1-p) + \beta(1-p)^n](\alpha + \beta + 1)} \\
E[\rho / Z_n=k] &= \frac{\alpha}{\alpha + \beta + 1}, \quad 1 \leq k \leq n-1 \\
E[\rho / Z_n=n] &= \frac{\alpha [p(\alpha+1) + \beta p^n]}{(\alpha + \beta + 1)[\alpha p + p^n \beta]}
\end{aligned} \tag{31}$$

Neste caso, podemos usar a esperança condicional de  $\rho$  dado o valor observado de  $Z_n$  como estimador para o coeficiente de correlação do processo. Caso o valor de  $p$  não esteja pré-especificado ele poderá ser estimado pela fração amostral observada de itens não conformes.

Os valores dos parâmetros  $(\alpha, \beta)$  precisam ser determinados à priori pelo pesquisador. Isto poderá ser feito de acordo com o conhecimento prévio que se tem do comportamento do processo. Uma distribuição possível é a Uniforme no intervalo  $(0,1)$  que ocorre quando se escolhe  $\alpha = 1$  e  $\beta = 1$  e que seria então uma distribuição "flat" para  $\rho$ , ou não informativa. Neste caso, o estimador Bayesiano de  $\rho$  simplificaria para:

$$\begin{aligned}
E[\rho / Z_n=0] &= \frac{[2(1-p) + (1-p)^n]}{[(1-p) + (1-p)^n](3)} \\
E[\rho / Z_n=k] &= \frac{1}{3}, \quad 1 \leq k \leq n-1 \\
E[\rho / Z_n=n] &= \frac{[2p + p^n]}{(3)[p + p^n]}
\end{aligned} \tag{32}$$

O estimador Bayesiano I foi desenvolvido com base no conhecimento de apenas uma amostra da Binomial Generalizada (16). Se tivermos  $m > 2$  amostras independentes da Binomial Generalizada então, podemos estimar  $\rho$  através do procedimento (31) para cada amostra de tamanho  $n$  e no final podemos fazer a média das estimativas encontradas, tendo-se assim uma estimativa final para o coeficiente de correlação do processo aproveitando-se a informação de todas as  $m$  amostras.

---

O interessante deste estimador é que basta uma amostra da Binomial Generalizada para que se possa ter uma estimativa de  $\rho$ , algo que não é possível quando se utiliza o estimador de Madsen (1993). No entanto, o pesquisador precisa especificar os valores dos parâmetros  $(\alpha, \beta)$ . Uma alternativa seria estimar-se os valores de  $(\alpha, \beta)$  a partir de amostras anteriores para as quais se tenha estimado o coeficiente de correlação  $\rho$  por algum método dando assim, um aspecto Bayesiano Empírico ao estimador definido em (31). No entanto, se o pesquisador tem alguma noção à priori da região em que o valor de  $\rho$  se encontra, por exemplo digamos no intervalo  $[c, d]$ ,  $c > 0$ ,  $d < 1$ ,  $c < d$ , ele poderá especificar uma distribuição Beta à priori simétrica em  $[c, d]$  especificando os parâmetros  $(\alpha, \beta)$  convenientemente. Através de simulações mostra-se que este procedimento resulta em melhores estimativas que aquelas produzidas pelo estimador clássico de Madsen. No entanto, caso o pesquisador especifique erradamente a região à priori na qual espera que  $\rho$  pertença, a estimativa produzida pelo estimador Bayesiano I poderá perder qualidade em relação ao estimador de Madsen. A perda em qualidade depende do grau de erro envolvido na especificação à priori para  $\rho$ . Assim, recomendamos que o pesquisador somente use o estimador Bayesiano I que envolve a distribuição priori Beta quando tenha à disposição alguma informação anterior confiável sobre o comportamento de  $\rho$ . No Gráfico 1 mostramos alguns exemplos de distribuições Beta para que o leitor possa ter uma noção das formas matemáticas da distribuição conforme os valores de  $(\alpha, \beta)$  se modificam.

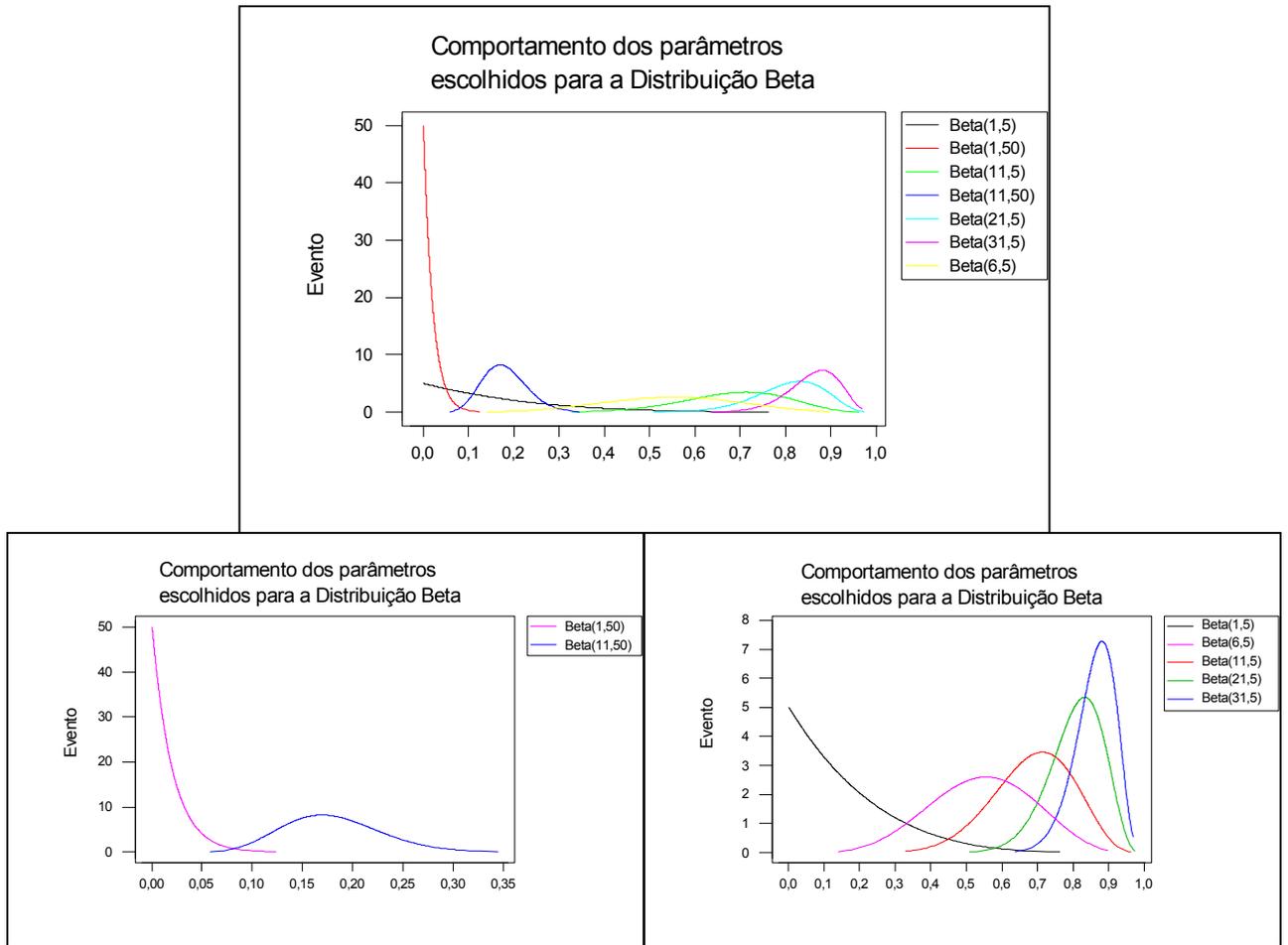


Gráfico 1: Exemplos do comportamento da distribuição Beta para vários valores dos parâmetros ( $\alpha, \beta$ )

### 2.2.2.3 Estimador Bayesiano II (Mingoti, 2002)

Suponha agora que a distribuição à priori para  $\rho$  seja uma Uniforme no intervalo de  $(a, b)$ ,  $a \geq 0$ ,  $b \leq 1$ , isto é,

$$g(\rho) = \frac{1}{(b-a)}, \quad a \leq \rho \leq b \quad (33)$$

Neste caso, calculando-se então a esperança condicional de  $\rho$  de acordo com a distribuição à posteriori de  $\rho$  dado  $Z_n$  e  $p$ , obtemos os seguintes resultados:

$$E[\rho / Z_n = 0] = \frac{(\frac{b^3 - a^3}{3})[(1-p) - (1-p)^n] + (1-p)^n (\frac{b^2 - a^2}{2})}{(\frac{b^2 - a^2}{2})[(1-p) - (1-p)^n] + (b-a)(1-p)^n}$$

$$E[\rho / Z_n = k] = \frac{[(\frac{b^2 - a^2}{2}) - (\frac{b^3 - a^3}{3})]}{[(b-a) - (\frac{b^2 - a^2}{2})]}, \quad 1 \leq k \leq n-1 \quad (34)$$

$$E[\rho / Z_n = n] = \frac{(\frac{b^3 - a^3}{3})[p - p^n] + (\frac{b^2 - a^2}{2})p^n}{(\frac{b^2 - a^2}{2})[p - p^n] + [(b-a)p^n]}$$

e o parâmetro  $p$  poderia ser estimado pela fração amostral de defeituosos na amostra. Se tivermos  $m > 2$  amostras independentes da Binomial Generalizada (16) podemos estimar  $\rho$  através do procedimento (34) para cada amostra de tamanho  $n$  e no final podemos fazer a média das estimativas encontradas, tendo-se assim uma estimativa final para o coeficiente de correlação do processo aproveitando-se a informação das  $m$  amostras.

Novamente o interessante deste estimador é que basta uma amostra da Binomial Generalizada para que se possa ter uma estimativa de  $\rho$ . No entanto, o pesquisador precisa especificar previamente os valores dos parâmetros  $a$  e  $b$ . Do mesmo modo que no caso da

distribuição Beta, uma alternativa seria estimar-se os valores de  $(a,b)$  a partir de amostras anteriores para as quais se tenha estimado o coeficiente de correlação  $\rho$  por algum método dando assim, um aspecto Bayesiano Empírico ao estimador definido em (34).

Ao contrário do estimador Bayesiano I, o estimador Bayesiano II exige menos conhecimento prévio do pesquisador em relação ao comportamento de  $\rho$ . Basta que o pesquisador tenha alguma noção à priori da região em que o valor de  $\rho$  se encontra para que possa aperfeiçoar as estimativas de  $\rho$  que serão produzidas pelas amostras do processo. Através de simulações mostra-se que este procedimento resulta em melhores estimativas que aquelas produzidas pelo estimador clássico de Madsen. No entanto, caso o pesquisador especifique erradamente a região à priori na qual espera que  $\rho$  pertença, a estimativa produzida pelo estimador Bayesiano II poderá perder em qualidade em relação ao estimador de Madsen. A perda em qualidade depende do grau de erro envolvido na especificação da região à priori para  $\rho$ .

#### 2.2.2.4 Estimador Multinomial para o caso de Observações Sequenciais

Um estimador alternativo para o caso em que o usuário tem as observações individuais da sequência de "0" e "1", considerando-se que os dados das amostras agrupadas vêm de uma Binomial Generalizada é o que vamos descrever a seguir.

Suponha que uma amostra de  $n$  itens do processo tenha sido inspecionada. Para cada item na amostra defina a variável:

$$X_i = \begin{cases} 1 & \text{se o } i\text{-ésimo item for não conforme} \\ 0 & \text{se o } i\text{-ésimo item for conforme} \end{cases}$$

Suponha que os itens sejam correlacionados da seguinte descrita em (15). Seja  $Z_n$  o número total de itens não conformes observados na amostra de tamanho  $n$  do processo. Considerando-se o modelo Binomial Generalizado descrito em (16) temos:

$$\begin{aligned}
P[X_i=1, X_{i+1}=1] &= \rho p + (1-\rho) p^2 \\
P[X_i=1, X_{i+1}=0] &= (1-\rho) p(1-p) \\
P[X_i=0, X_{i+1}=0] &= \rho(1-p) + (1-p)^2(1-\rho) \\
P[X_i=0, X_{i+1}=1] &= (1-\rho) p(1-p)
\end{aligned}
\tag{35}$$

Se temos  $n$  ensaios de  $X_i$  podemos formar pares  $(X_i, X_{i+1})$  na sequência em que os itens foram inspecionados,  $i=1, 2, \dots, n-1$ . O número total de pares que podem ser formados é igual a  $(n-1)$ . Sejam  $f_{10}$  e  $f_{01}$  a frequência amostral observada de pares do tipo  $(X_i = 1, X_{i+1} = 0)$  e  $(X_i = 0, X_{i+1} = 1)$  respectivamente. Então, usando o mesmo procedimento descrito na seção 2.2.2.1 teríamos o seguinte estimador para  $\rho$  :

$$\hat{\rho} = 1 - \frac{f_{10} + f_{01}}{2(n-1)p(1-p)}
\tag{36}$$

Caso o valor de  $p$  não esteja pré-especificado ele poderá ser estimado pela fração amostral observada de itens não conformes.

No caso em que o pesquisador tem à disposição  $m$  amostras independentes da variável  $Z_n$  e para cada amostra tem os valores individuais de "0" e "1" na sequência observada, ele poderá para cada amostra individualmente estimar  $\rho$  de acordo com o procedimento em (36) e tomar então a média de todas as  $m$  estimativas calculadas tendo-se assim, um estimador final para  $\rho$ . Este seria também um estimador alternativo para o coeficiente de correlação de Cadeia de Markov no caso em que a correlação é considerada maior que zero.

### 2.2.2.5 Estimador Não-Paramétrico para o caso de Observações Sequenciais

Considere novamente a situação na qual temos uma amostra de  $n$  itens do processo e para cada item tenhamos guardado a informação de  $X_i$ . Neste caso, podemos adaptar o coeficiente não paramétrico de concordância de Kendall (Sprent, 1989) para construir um índice

que venha a medir a correlação entre os ensaios. Fazendo-se isto tem-se o seguinte estimador para  $\rho$  :

$$\hat{\rho}_k = \frac{(f_{00} + f_{11}) - [(n-1) - f_{00} + n_{11}]}{n-1} = \frac{2[(f_{00} + f_{11})] - (n-1)}{n-1} \quad (37)$$

onde  $f_{11}$  e  $f_{00}$  representam a frequência amostral de pares do tipo  $(X_i = 1, X_{i+1} = 1)$  e  $(X_i = 0, X_{i+1} = 0)$  respectivamente. Este estimador tem a vantagem de não depender do conhecimento da distribuição original de  $X_i$  ou de  $Z_n$  podendo então ser utilizado tanto para dados gerados via Cadeia de Markov quanto para dados gerados via distribuição Binomial Generalizada, além de situações mais gerais.

### 2.2.3 Limites de Controle no Modelo Binomial Genralizado

A introdução do coeficiente de correlação no modelo Binomial gera uma distribuição que não pode ser aproximada pela distribuição normal quando o tamanho amostral  $n$  é grande. Neste caso, a determinação dos limites inferior e superior de controle, tanto para a proporção de itens defeituosos ( $p$ ) quanto para o número de defeituosos do processo ( $np$ ), precisa ser feita através da distribuição exata. Isto significa dizer que se o usuário deseja construir um intervalo de  $(1-\alpha)100\%$  de confiança para  $p$  ou  $np$  será necessário encontrar as constantes  $c_1$  e  $c_2$  tais que:

$$\sum_{i=0}^{c_1} P[Z_n=i] = \alpha_1 ; \sum_{i=c_2}^n P[Z_n=i] = \alpha_2 ; \alpha_1 + \alpha_2 = \alpha \quad , \quad 0 < \alpha < 1 ; 0 < \alpha_j < 1 , j = 1,2.$$

---

## CAPÍTULO 3

### *Instruções e Exemplos de Aplicação do Programa SCAP - Módulo AI*

#### 3.0 Introdução

Neste capítulo apresentamos instruções gerais para o usuário do *SCAP - módulo AI (Attribute Inspection)*. A macro que implementa os estimadores descritos no Capítulo 2 é chamada de **scapai.mac** (Mingoti & Carvalho, 2002) e funciona dentro do sistema operacional do *software* estatístico *Minitab for Windows*.

Para utilização deste programa espera-se que o usuário tenha a versão 11, ou mais recente, do *software Minitab for Windows*. Os exemplos aqui apresentados foram gerados na versão 13.

Para que o usuário possa usufruir mais adequadamente dos recursos disponíveis no módulo *AI* do *SCAP* é essencial que tenha lido o capítulo 2 deste manual, no qual a teoria estatística relacionada com os modelos Binomial Generalizado e Cadeia de Markov é apresentada, assim como os estimadores do coeficiente de correlação construídos via Estatística Clássica Paramétrica, Estatística Bayesiana e Estatística Não-Paramétrica.

O objetivo deste capítulo é mostrar ao usuário o processamento da macro *scapai.mac* em termos da entrada e saída de dados e as várias opções disponíveis para a estimação da proporção de itens "não-conformes" (defeituosos) do processo bem como o respectivo coeficiente de correlação entre as unidades amostrais no que se refere à variável resposta.

Estando o usuário na janela de seção do *Minitab (session window)* este deverá chamar a macro **scapai.mac** usando o comando (supondo que a macro esteja salva no drive "a" do computador):

```
%a: \scapai.mac
```

A partir daí, várias perguntas serão feitas ao usuário para que as estimações correspondentes possam ser processadas. Nas seções a seguir apresentamos a formatação de dados e alguns exemplos de utilização.

### 3.1 Formatação dos Dados

O módulo *AI* do programa *SCAP* no qual está implementado a macro *scapai.mac* aceita dados em termos de *número total de defeitos* ou *dados individuais*. Entende-se como dados individuais o resultado observado em cada item inspecionado, definindo-se como “0” o resultado para o item “conforme” (não defeituoso) e “1” para o item “não-conforme” (defeituoso).

Antes de processar a macro *scapai.mac* o usuário deverá definir o formato em que os dados serão apresentados. Independentemente de qual for a escolha do usuário, os dados deverão estar dispostos no *Worksheet do Minitab (planilha de dados)*, que é uma janela própria para o fornecimento, visualização e edição dos dados a serem analisados. Cada variável (característica de qualidade) deverá ter seus resultados amostrais em uma planilha de dados separada.

#### 3.1.1 Dados Apresentados pelo “Total de Defeitos”.

A macro *scapai.mac* aceita analisar amostras com tamanhos iguais ou diferentes.

Amostras de tamanhos iguais: Neste caso a primeira coluna (C1) do *Worksheet do Minitab* deve ser utilizada para numerar as amostras (subgrupos) de 1 a “m”, sendo “m” o número total de amostras a serem analisadas. A segunda coluna (C2) deverá ser utilizada para informar o número de itens defeituosos que foram observados em cada subgrupo, o qual pode variar de 1 a “n”, sendo “n” o número total de itens observados em cada amostra.

↓	C1	C2
1	1	[1,n]
2	2	[1,n]
3	3	[1,n]
	•	[1,n]
	•	[1,n]
m	m	[1,n]

↓	C1	C2	C3
1	1	[1,n1]	n1
2	2	[1,n2]	n2
3	3	[1,n3]	n3
	•	•	•
	•	•	•
	•	•	•
m	m	[1,nm]	nm

Amostras de tamanhos diferentes: Analogamente ao caso anterior, a primeira coluna deve ser preenchida com valores sequenciais a partir de 1 até “m” visando numerar as amostras. A segunda coluna (C2) também deve conter o número de itens defeituosos observados em cada amostra, sendo o intervalo de valores possíveis  $\{[1,n_i], \forall i=1,\dots,m\}$  para cada célula. A coluna (C3) deverá conter o tamanho  $n_i$  de cada amostra,  $i = 1, 2, \dots, m$ .

**Observação:** Para análise de amostras de tamanhos diferentes, os dados somente podem ser apresentados pelo *número total de defeitos*.

### 3.1.2 Dados Apresentados “Individualmente”.

Neste formato (que não é aplicável a dados constituídos por amostras de tamanhos diferentes), deve ser fornecido a *seqüência* em que os dados de cada amostra (subgrupo) foram obtidos.

Cada coluna deverá conter uma seqüência constituída de “zeros” e “uns” e com a ordem *EXATA* em que os dados foram produzidos. Logo, o *Worksheet do Minitab* deverá conter uma matriz com “*m*” colunas (cada coluna representando uma amostra/subgrupo) e “*n*” linhas, cada linha

↓	C1	C2		Cm
1	0 ou 1	0 ou 1		0 ou 1
2	0 ou 1	0 ou 1		0 ou 1
3	0 ou 1	0 ou 1		0 ou 1
	•	•	• • •	•
	•	•		•
	•	•		•
n	0 ou 1	0 ou 1		0 ou 1

representando o resultado observado da inspeção de um item da amostra, sendo que o valor 0 (zero) representa um item não defeituoso ("conforme ou aceitável") e o valor 1 (um) representa um item defeituoso ("não conforme ou defeituoso").

Após a leitura dos dados o usuário poderá proceder a estimação dos parâmetros “*p*” (probabilidade do processo gerar um item não conforme), e “*ρ*” (coeficiente de correlação do processo) através das metodologias de Estatística Paramétrica Clássica, Não-Paramétrica e Bayesiana, como descrito no capítulo 2 deste manual.

## 3.2 Fornecendo Informações Preliminares

A partir do momento em que o usuário processar a macro **scapai.mac** a mesma passará a interagir com ele de uma forma extremamente simples. Para fazer a análise dos dados basta seguir os passos indicados e estar sempre atento às mensagens fornecidas pela macro.

### 3.2.1 Informando o Tamanho da Amostra

Tendo os dados no *Worksheet* do *Minitab* o usuário deverá responder as seguintes questões:

**Quantos são os subgrupos analisados (amostras)?**

**DATA>** ⇐ Neste campo entre com o número de subgrupos (amostras) a serem analisadas.

**Os subgrupos (amostras) possuem o mesmo tamanho (Y/N)?**

**DATA>** ⇐ Responda “Y” para resposta afirmativa e “N” para resposta negativa.

Caso a resposta à questão anterior seja positiva, será formulada a seguinte pergunta:

**Qual o tamanho de cada subgrupo (amostra)?**

**DATA>** ⇐ Entre com o valor respectivo ao tamanho de cada subgrupo (amostra).

A seguinte mensagem será apresentada:

**Para amostras de tamanhos diferentes, a única formatação aceitável é aquela onde os dados são apresentados pelo total de defeitos.**

**Deseja continuar (Y/N)?**

**DATA>** ⇐ Responda “Y” para resposta afirmativa e “N” para resposta negativa.

Caso o usuário responda negativamente, a macro terminará seu processamento; e caso responda positivamente, a macro continuará seu processamento com a seguinte seqüência de questões para o usuário:

**Para cada subgrupo (amostra), seus dados estão apresentados em:**

**1. Total de defeitos;**

**2. Dados individuais.**

**Digite 1 ou 2.**

**DATA>** ⇐ Neste campo entre com o código respectivo a formatação desejada.

Caso a formatação escolhida seja a “*dados individuais*”, será emitida a seguinte mensagem:

**Para dados apresentados individualmente, cada coluna deve conter os dados de uma amostra. Deseja continuar (Y/N)?**

**DATA>**  $\Leftarrow$  Responda “Y” para resposta afirmativa e “N” para resposta negativa.

Caso o usuário responda negativamente, a macro terminará o seu processamento; e caso o usuário responda afirmativamente ele poderá estimar a probabilidade de “defeito”, ou seja do processo gerar um item “não conforme”, e o coeficiente de correlação do processo como mostrado nas seções 3.4 e 3.5 a seguir.

### **3.3 Estimando a Probabilidade de Defeito “p”**

Após a definição da entrada de dados a macro passará a formular perguntas sobre o que o usuário desejará estimar. As primeiras questões referem-se a estimação da presença de itens “não-conformes” do processo, como mostrado a seguir.

**Você deseja estimar a probabilidade de defeito “p” (Y/N)?**

**DATA>**  $\Leftarrow$  Responda “Y” para resposta afirmativa e “N” para resposta negativa.

Caso o usuário responda negativamente:

**Forneça o valor pré-determinado (à priori) de “p”:**

**DATA>**  $\Leftarrow$  Forneça o valor de “p” (o valor de p deverá estar no intervalo [0;1]).

Caso a usuário responda afirmativamente, a macro continuará da seguinte maneira:

- I. Se as amostras analisadas tiverem tamanhos diferentes, será emitida a seguinte mensagem de aviso:

**A probabilidade de defeito “p” será estimada com base na análise da fração de defeituosos na amostra.**

- II. Se as amostras analisadas tiverem tamanhos iguais, a macro formulará a seguinte questão:

Qual é o método de estimação de "p"?

1. Fração de defeituosos na amostra;
2. Cadeia de Markov (somente para dados individuais).

Digite 1 ou 2.

DATA> ⇐ Neste campo entre com o código respectivo ao método desejado.

### 3.3.1 Estimação de "p" por *Fração de Defeituosos na Amostra*.

Nesta seção mostraremos alguns exemplos para estimação da fração de defeituosos "p" considerando-se amostras de tamanhos iguais e diferentes.

#### 3.3.1.1 Estimação de "p" para Amostras de Tamanhos Iguais.

*Exemplo 1:* Para exemplificar, considere o conjunto de dados apresentado no Quadro 1 no qual se tem  $m=10$  amostras (subgrupos) de tamanho  $n=50$ . Para cada amostra tem-se o número de itens "não conformes" observado.

Quadro 1: Dados do exemplo 1.

Amostra	NºDefeituosos
1	0
2	2
3	25
4	10
5	0
6	2
7	5
8	0
9	3
10	3

Neste caso, teremos o seguinte processamento da macro:

```

MTB > %scapai
Executing from file: scapai.MAC

Quantos são os subgrupos analisados (amostras)?
DATA> 10

Os subgrupos (amostras) possuem o mesmo tamanho (Y/N)?
Y

Qual o tamanho de cada subgrupo (amostra)?
DATA> 50

Para cada subgrupo (amostra), seus dados estão apresentados em:
1. Total de defeitos;
2. Dados individuais.
Digite 1 ou 2.
DATA> 1

Você deseja estimar a probabilidade de defeito "p" (Y/N)?
Y

Qual é o método de estimação de "p"?
1. Fração de defeituosos na amostra;
2. Cadeia de Markov (somente para dados individuais).
Digite 1 ou 2.
DATA> 1

Data Display

p      0,100000

```

Portanto a estimativa de “p” será igual a 0,10, que nada mais é do que a fração amostral de itens "não-conformes" observada.

### 3.3.1.2 Estimação de “p” para Amostras de Tamanhos Diferentes.

Quando o usuário tiver amostras de tamanhos diferentes, o único método possível implementado neste programa para estimação da probabilidade de defeito é o da “*fração de defeituosos na amostra*”.

Quando isso ocorrer, será emitida na janela *Session* do *Minitab* a seguinte mensagem:

**A probabilidade de defeito "p" será estimada com base na análise da fração de defeituosos na amostra.**

*Exemplo 2:* Para exemplificar, considere o conjunto de dados apresentado no Quadro 2 no qual se tem  $m=80$  amostras de tamanhos diferentes. Para cada amostra tem-se o número de itens "não conformes" (nº de defeituosos) observado.

*Quadro 2: Dados do exemplo 2.*

Amostra	Número de Defeituosos	Tamanho da amostra	Amostra	Número de Defeituosos	Tamanho da amostra
1	2	466	41	6	366
2	1	413	42	4	385
3	6	495	43	7	306
4	18	501	44	7	434
5	6	558	45	10	501
6	7	407	46	7	505
7	9	486	47	12	656
8	10	548	48	16	599
9	8	469	49	8	508
10	7	405	50	9	532
11	6	541	51	9	506
12	11	426	52	12	311
13	6	444	53	5	435
14	10	505	54	9	419
15	14	544	55	9	425
16	15	589	56	6	442
17	16	585	57	5	401
18	18	579	58	7	410
19	17	603	59	8	410
20	13	564	60	4	320
21	13	657	61	2	309
22	12	775	62	2	371
23	10	749	63	2	343
24	12	977	64	4	331
25	6	606	65	3	342
26	12	711	66	9	302
27	13	728	67	11	299
28	13	672	68	8	296
29	20	700	69	6	237
30	14	528	70	7	335
31	11	503	71	10	369
32	14	683	72	17	376
33	20	424	73	7	323
34	10	539	74	7	264
35	5	421	75	7	250
36	11	267	76	9	289
37	6	340	77	9	301
38	15	414	78	7	252
39	11	413	79	10	271
40	9	362			
79	10	271			

Neste caso, teremos o seguinte processamento da macro:

```
MTB > %scapai
Executing from file: scapai.MAC

Quantos são os subgrupos analisados (amostras)?
DATA> 79

Os subgrupos (amostras) possuem o mesmo tamanho (Y/N)?
n

Para amostras de tamanhos diferentes, a única formatação aceitável é
aquela onde os dados são apresentados pelo total de defeitos.
Deseja continuar (Y/N)?
Y

Você deseja estimar a probabilidade de defeito "p" (Y/N)?
Y

A probabilidade de defeito "p" será estimada com base na análise da
fração de defeituosos na amostra.

Data Display

p      0,0201981
```

Portanto, a estimativa de “p” será de 0,0201981

### 3.3.2 Estimação de “p” por *Cadeia de Markov*.

Este método é aplicável somente quando os dados forem apresentados individualmente. Portanto, se o usuário tiver apresentado seus dados com formatação “total de defeitos” e escolher este método para estimar a probabilidade de defeito “p”, será emitida a seguinte mensagem de erro:

**!!! ERRO !!! Este método pode ser aplicado somente em dados individuais.**

com o posterior abortamento término do processamento da macro.

Exemplo 3: Considere o seguinte conjunto de dados apresentado no Quadro 3 no qual se tem  $m=5$  subgrupos constituídos de  $n=60$  observações (itens inspecionados), porém com a seguinte formatação:

Quadro 3: Dados do exemplo 3.

1	0	1	1	0
0	0	0	1	0
0	1	0	1	0
0	0	0	1	0
0	0	1	0	0
1	0	1	1	0
0	0	1	0	0
0	0	1	0	0
0	0	0	1	1
1	0	0	1	1
0	0	0	1	1
0	0	0	1	0
0	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	1	0
1	0	0	0	0
0	1	1	0	1
0	1	1	1	0
0	1	1	0	0
1	0	0	0	1
0	0	0	0	1
0	0	0	1	0
1	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	1	0	0
1	0	0	0	0
0	1	0	1	1
0	0	0	1	0
0	0	0	0	0
0	1	0	1	0
1	0	1	0	0
1	0	0	1	1
0	0	0	1	1
0	0	0	0	0
0	0	0	0	0
0	1	0	1	0
0	0	0	1	0
0	0	0	1	1
0	0	0	1	1
0	0	1	1	0
0	0	1	0	0
1	0	1	0	1
1	1	0	1	0
1	1	0	0	0
1	0	0	1	1
0	0	0	1	0
0	0	0	1	0

*cont. Quadro 3*

0	0	0	0	0
1	0	0	0	0
0	0	0	0	1
1	0	0	1	1
0	0	0	0	1
0	0	0	1	0
0	0	1	0	0
0	0	0	1	0
1	0	0	0	1
1	1	0	1	0
0	0	0	0	1
0	0	1	0	0

Neste caso, teremos o seguinte processamento da macro:

```
MTB > %scapai
Executing from file: scapai.MAC

Quantos são os subgrupos analisados (amostras)?
DATA> 5

Os subgrupos (amostras) possuem o mesmo tamanho (Y/N)?
Y

Qual o tamanho de cada subgrupo (amostra)?
DATA> 60

Para cada subgrupo (amostra), seus dados estão apresentados em:
1. Total de defeitos;
2. Dados individuais.
Digite 1 ou 2.
DATA> 2

Para dados apresentados individualmente, cada coluna deve conter os
dados de uma amostra. Deseja continuar (Y/N)?
Y

Você deseja estimar a probabilidade de defeito "p" (Y/N)?
Y

Qual é o método de estimação de "p"?
1. Fração de defeituosos na amostra;
2. Cadeia de Markov (somente para dados individuais).
Digite 1 ou 2.
DATA> 2

Data Display

p      0,303670
```

Portanto a estimativa de “p” será de 0,303670.

**Exemplo 4:** Considere o seguinte conjunto de dados apresentado no Quadro 4 no qual se tem  $m=1$  subgrupo constituídos de  $n=208$  observações (itens inspecionados), porém com a seguinte formatação:

*Quadro 4: Dados do exemplo 4.*

*Erros de leitura em Discos Rígidos de Microcomputadores (número de não-conformes lendo as linhas de cima para baixo e as colunas da esquerda para a direita).*

1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0
0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1
0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
0	1	0	0	0	0	0	0	0	0	1	0	1	0	1	0	1	1	1	0
0	0	0	0	0	0	1	1	0	0	1	0	1	1	0	1	1	1	0	0
1	0	1	1	0	0	0	1	0	1	1	1	1	1	1	1	1	1	1	0
1	0	1	1	1	1	1	1	1	1	0	1	1	1	0	0	1	1	0	0
1	1	1	0	1	0	1	0	0	0	1	0	0	1	0	1	0	1	1	0
1	0	1	0	0	1	0	0	1	0	1	1	0	1	0	0	0	1	0	1
0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	0	0	0

Fonte: LAI,C.D.; XIE,M. & GOVINDARAJU,K.(2000). Study of a Markov model for a high-quality dependent process, *Journal of Applied Statistics*, 27(4), pp. 465.

**Observação:** Na janela *Worksheet* do *Minitab for Windows*, esses dados deverão estar dispostos somente em uma coluna, havendo assim somente uma amostra para análise.

Neste caso, teremos o seguinte processamento da macro:

```
MTB > %scapai
Executing from file: scapai.MAC

Quantos são os subgrupos analisados (amostras)?
DATA> 1

Os subgrupos (amostras) possuem o mesmo tamanho (Y/N)?
Y

Qual o tamanho de cada subgrupo (amostra)?
DATA> 208

Para cada subgrupo (amostra), seus dados estão apresentados em:
1. Total de defeitos;
2. Dados individuais.
Digite 1 ou 2.
DATA> 2

Para dados apresentados individualmente, cada coluna deve conter os
dados de uma amostra. Deseja continuar (Y/N)?
Y
```

```
Você deseja estimar a probabilidade de defeito "p" (Y/N)?
```

```
Y
```

```
Qual é o método de estimação de "p"?
```

1. Fração de defeituosos na amostra;
2. Cadeia de Markov (somente para dados individuais).

```
Digite 1 ou 2.
```

```
DATA> 2
```

#### **Data Display**

```
p      0,419380
```

Portanto a estimativa de “p” será de 0,419380.

### **3.4 Estimando o Coeficiente de Correlação “ $\rho$ ”**

O módulo *AI* do programa *SCAP* fornece ao usuário várias alternativas para a estimação do coeficiente de correlação do processo. As escolhas deverão ser consistentes com o modelo de entrada de dados que o usuário optou. Alguns estimadores: *Cadeia de Markov*, *Mingoti II e Não-Paramétrico* são próprios para dados individuais, enquanto que o *Método dos Momentos (Madsen)*, *Multinomial (Mingoti I)* e os métodos *Bayesiano I e II (Mingoti III e IV)* são próprios para dados agrupados do modelo *Binomial Generalizado*, embora também estimem o coeficiente de correlação “ $\rho$ ” com dados apresentados individualmente.

Para estimação do coeficiente de correlação, as seguintes perguntas terão que ser respondidas pelo usuário:

**Você deseja estimar o valor do coeficiente de correlação (Y/N)?**

**DATA>**  $\Leftarrow$  Responda “Y” para resposta afirmativa e “N” para resposta negativa.

Caso o usuário responda negativamente, a seguinte pergunta aparecerá na tela:

**Forneça o valor pré-determinado (à priori) do coeficiente de correlação:**

**DATA>**  $\Leftarrow$  Forneça o valor de “ $\rho$ ” (“ $\rho$ ” deverá estar no intervalo  $[-1;1]$ , e ser consistente com a formulação do modelo Binomial Generalizado ou com a Cadeia de Markov dependendo do caso).

Caso a usuário responda afirmativamente e as amostras analisadas tenham tamanhos iguais, a seguinte pergunta aparecerá na tela:

Qual é o método de estimação do coeficiente de correlação?

1. Método dos momentos;
2. Método de Cadeia de Markov (somente para dados individuais);
3. Método Não-Paramétrico (somente para dados individuais);
4. Método Multinomial (Mingoti I);
5. Método Mingoti II (somente para dados individuais);
6. Método Bayesiano I (Mingoti III);
7. Método Bayesiano II (Mingoti IV).

Digite 1,2,3,4,5,6 ou 7.

DATA> ⇐ Neste campo entre com o código respectivo ao método desejado.

A estimação se dará de acordo com a escolha do usuário desde que o estimador escolhido seja compatível com o formato de entrada dos dados utilizado.

Se amostras analisadas tiverem tamanhos diferentes, será emitida a seguinte mensagem de aviso:

O coeficiente de correlação "rô", será estimado com base na teoria do Método dos momentos (Madsen).

### 3.4.1 Estimação do Coeficiente de Correlação para Dados Apresentados pelo "Total de Defeitos".

#### 3.4.1.1 Estimação do Coeficiente de Correlação pelo "Método dos Momentos (Madsen, 1993)".

##### A. Estimação do Coeficiente de Correlação para Amostras de Tamanhos Iguais.

Caso o usuário opte pela opção 1 (*Método dos Momentos*) para a estimação do coeficiente de correlação, durante o processamento da macro será emitida a seguinte mensagem de aviso:

!!! OBSERVAÇÃO !!! Considerando o método escolhido, será estimado:  
"rô1"= com variância obtida pelo divisor (m-1), e  
"rô2"= com variância obtida pelo divisor(m) (proposto por Madsen);

---

e posteriormente os dois resultados estimados pelo programa serão apresentados na janela *Session do Minitab* da seguinte forma:

**Data Display**

$r\hat{\theta}1$   $\Leftarrow$  Valor obtido pela macro para o primeiro estimador da correlação.

**Data Display**

$r\hat{\theta}2$   $\Leftarrow$  Valor obtido pela macro para o segundo estimador da correlação.

e finalmente, o usuário deverá definir qual será o estimador que será adotado nas etapas subseqüentes para construção do gráfico de controle, caso seja de seu interesse respondendo a seguinte questão:

**Considerando os estimadores do coeficiente de correlação obtidos, qual você deseja adotar nas etapas seguintes?**

1. " $r\hat{\theta}1$ " (Madsen Modificado)

2. " $r\hat{\theta}2$ " (Proposto por Madsen)?

Digite (1 ou 2)

**DATA>**  $\Leftarrow$  Neste campo entre com o código respectivo ao estimador do coeficiente de correlação desejado.

O resultado final será emitido na janela *Session do Minitab*:

**Data Display**

$r\hat{\theta}$   $\Leftarrow$  Valor estimado pela macro e escolhido pelo usuário para uso na construção do gráfico de controle.

*Exemplo 5:* Considere o conjunto de dados apresentado anteriormente no Quadro 1, seção 3.3.1.1. Neste caso, teremos o seguinte processamento da macro:

```
MTB > %scapai
Executing from file: scapai.MAC

Quantos são os subgrupos analisados (amostras)?
DATA> 10

Os subgrupos (amostras) possuem o mesmo tamanho (Y/N)?
Y

Qual o tamanho de cada subgrupo (amostra)?
DATA> 50

Para cada subgrupo (amostra), seus dados estão apresentados em:
1. Total de defeitos;
2. Dados individuais.
Digite 1 ou 2.
DATA> 1

Você deseja estimar a probabilidade de defeito "p" (Y/N)?
Y

Qual é o método de estimação de "p"?
1. Fração de defeituosos na amostra;
2. Cadeia de Markov (somente para dados individuais).
Digite 1 ou 2.
DATA> 1

Data Display

p      0,100000

Você deseja estimar o valor do coeficiente de correlação (Y/N)?
Y

Qual é o método de estimação do coeficiente de correlação?
1. Método dos momentos;
2. Método de Cadeia de Markov (somente para dados individuais);
3. Método Não-Paramétrico (somente para dados individuais);
4. Método Multinomial (Mingoti I);
5. Método Mingoti II (somente para dados individuais);
6. Método Bayesiano I (Mingoti III);
7. Método Bayesiano II (Mingoti IV).
Digite 1,2,3,4,5,6 ou 7.
DATA> 1

!!! OBSERVAÇÃO !!! Considerando o método escolhido, será estimado:
"rô1"= com variância obtida pelo divisor (m-1), e
"rô2"= com variância obtida pelo divisor(m) (proposto por Madsen);

Data Display

rô1    0,244646
```

### Data Display

rô2 0,218141

Considerando os estimadores do coeficiente de correlação obtidos, qual você deseja adotar nas etapas seguintes?

- 1."rô1" (Madsen Modificado);
- 2."rô2" (Proposto por Madsen)?

Digite (1 ou 2)

DATA> 1

### Data Display

ro 0,244646

Portanto, as estimativas do coeficiente de correlação serão 0,218141 e 0,244646 mas apenas 0,244646 será utilizada na construção de gráficos de controle caso o usuário venha a optar por este procedimento nas etapas subsequentes.

## B. Estimação do Coeficiente de Correlação para Amostras de Tamanhos Diferentes.

Quando o usuário tiver amostras de tamanhos diferentes, o único método possível para estimação do coeficiente de correlação implementado neste programa é o “*Método dos Momentos (Madsen)*”. Quando isso ocorrer, será emitida na janela *Session* do *Minitab* a seguinte mensagem:

**O coeficiente de correlação "rô", será estimado com base na teoria do Método dos Momentos (Madsen, 1993).**

Exemplo 6: Considere o conjunto de dados apresentado anteriormente no Quadro 2, seção 3.3.1.2. Neste caso, teremos o seguinte processamento da macro:

```

MTB > %scapai
Executing from file: scapai.MAC

Quantos são os subgrupos analisados (amostras)?
DATA> 80
      1 rows read.

Os subgrupos (amostras) possuem o mesmo tamanho (Y/N)?
n

Para amostras de tamanhos diferentes, a única formatação aceitável
é aquela onde os dados são apresentados pelo total de defeitos.
Deseja continuar (Y/N)?
y

Você deseja estimar a probabilidade de defeito "p" (Y/N)?
y

A probabilidade de defeito "p" será estimada com base na análise
da fração de defeituosos na amostra.

Data Display

p      0,0201981

Você deseja estimar o valor do coeficiente de correlação (Y/N)?
y

O coeficiente de correlação "rô", será estimado com base na teoria
do Método dos momentos (Madsen).

!!! OBSERVAÇÃO !!! Considerando o método escolhido, será estimado:
"rô1"= com variância obtida pelo divisor (m-1), e
"rô2"= com variância obtida pelo divisor(m) (proposto por Madsen);

Data Display

rô1    0,00228346

Data Display

rô2    0,00222698

Considerando os estimadores do coeficiente de correlação obtidos,
qual você deseja adotar nas etapas seguintes?
1."rô1" (Madsen Modificado);
2."rô2" (Proposto por Madsen)?
Digite (1 ou 2)
DATA> 1

Data Display

ro     0,00222698

```

Portanto, a estimativa do coeficiente de correlação do processo 0,00222698, será o valor utilizado na construção de gráficos de controle caso o usuário venha a optar por este procedimento posteriormente.

### 3.4.1.1 Estimação do Coeficiente de Correlação pelo “Método Multinomial (Mingoti I)”.

Neste método, somente o valor estimado pela macro será emitido na janela *Session do Minitab*:

#### Data Display

ro ← Valor estimado pela macro

*Exemplo 7:* Considere o conjunto de dados apresentado anteriormente no Quadro 1, seção 3.3.1.1. Neste caso, teremos o seguinte processamento da macro:

```
MTB > %scapai
Executing from file: scapai.MAC

Quantos são os subgrupos analisados (amostras)?
DATA> 10

Os subgrupos (amostras) possuem o mesmo tamanho (Y/N)?
Y

Qual o tamanho de cada subgrupo (amostra)?
DATA> 50

Para cada subgrupo (amostra), seus dados estão apresentados em:
1. Total de defeitos;
2. Dados individuais.
Digite 1 ou 2.
DATA> 1

Você deseja estimar a probabilidade de defeito "p" (Y/N)?
Y

Qual é o método de estimação de "p"?
1. Fração de defeituosos na amostra;
2. Cadeia de Markov (somente para dados individuais).
Digite 1 ou 2.
DATA> 1
```

### Data Display

p 0,100000

Você deseja estimar o valor do coeficiente de correlação (Y/N)?

Y

Qual é o método de estimação do coeficiente de correlação?

1. Método dos momentos;
2. Método de Cadeia de Markov (somente para dados individuais);
3. Método Não-Paramétrico (somente para dados individuais);
4. Método Multinomial (Mingoti I);
5. Método Mingoti II (somente para dados individuais);
6. Método Mingoti III (Método Bayesiano I);
7. Método Mingoti IV (Método Bayesiano II).

Digite 1,2,3,4,5,6 ou 7.

DATA> 4

### Data Display

ro 0,296374

Portanto, a estimativa do coeficiente de correlação será de 0,296374.

#### 3.4.1.2. Estimação do Coeficiente de Correlação pelo “Método Bayesiano I (Mingoti III)”.

Para estimar o coeficiente de correlação por este método o usuário deverá fornecer os parâmetros "alfa" e "beta" da distribuição Beta na qual este estimador é fundamentado (ver Capítulo 2 deste manual).

Será emitida na janela *Session do Minitab* a seguinte mensagem:

**Considerando que o coeficiente de correlação possui uma distribuição Beta com parâmetros alfa e beta, forneça à priori:**

**O valor da constante alfa (alfa > 0):**

**DATA>** ← Neste campo entre com o valor respectivo à constante alpha.

**O valor da constante beta (beta > 0):**

**DATA>** ← Neste campo entre com o valor respectivo à constante beta.

Caso o usuário não tenha pré-fixado a probabilidade de defeito anteriormente, o coeficiente de correlação poderá ser estimado usando-se a probabilidade de defeito estimada pela *fração de defeituosos na amostra* considerando-se todas as amostras

conjuntamente, ou também pela *fração de defeituosos na amostra*, porém considerando-se um estimador da probabilidade de defeito para cada amostra individualmente e compondo-se as  $m$  estimativas de " $p$ " de acordo com o estimador Bayesiano proposto para o coeficiente de correlação (ver capítulos 1 e 2 deste manual).

A seguinte pergunta aparecerá na tela:

**Considerando que a probabilidade de defeito não foi anteriormente pré-fixada, você deseja que a mesma seja obtida:**

1. **Analisando cada amostra individualmente;**
2. **Analisando todas as amostras conjuntamente.**

**Digite 1 ou 2.**

**DATA>** ⇐ Neste campo entre com o código respectivo ao método desejado.

Finalmente será emitido o resultado final na janela *Session do Minitab*:

### **Data Display**

**ro** ⇐ Valor estimado pela macro e escolhido pelo usuário.

Exemplo 8: Considere o conjunto de dados apresentado anteriormente no Quadro 1, seção 3.3.1.1. Neste exemplo, vamos considerar os parâmetros  $\alpha=\beta=1$ , e vamos analisar as amostras individualmente. Assim, teremos o seguinte processamento da macro:

```
MTB > %scapai
Executing from file: scapai.MAC

Quantos são os subgrupos analisados (amostras)?
DATA> 10

Os subgrupos (amostras) possuem o mesmo tamanho (Y/N)?
Y

Qual o tamanho de cada subgrupo (amostra)?
DATA> 50

Para cada subgrupo (amostra), seus dados estão apresentados em:
1. Total de defeitos;
2. Dados individuais.
Digite 1 ou 2.
DATA> 1

Você deseja estimar a probabilidade de defeito "p" (Y/N)?
Y
```

```
Qual é o método de estimação de "p"?
1. Fração de defeituosos na amostra;
2. Cadeia de Markov (somente para dados individuais).
Digite 1 ou 2.
DATA> 1
```

### Data Display

```
p      0,100000
```

```
Você deseja estimar o valor do coeficiente de correlação (Y/N)?
Y
```

```
Qual é o método de estimação do coeficiente de correlação?
1. Método dos momentos;
2. Método de Cadeia de Markov (somente para dados individuais);
3. Método Não-Paramétrico (somente para dados individuais);
4. Método Multinomial (Mingoti I);
5. Método Mingoti II (somente para dados individuais);
6. Método Mingoti III (Método Bayesiano I);
7. Método Mingoti IV (Método Bayesiano II).
Digite 1,2,3,4,5,6 ou 7.
DATA> 6
```

```
Considerando que o coeficiente de correlação possui uma distribuição
Beta com parâmetros alfa e beta, forneça à priori:
O valor da constante alfa (alfa > 0):
DATA> 1
```

```
O valor da constante beta (beta > 0):
DATA> 1
```

```
Considerando que a probabilidade de defeito não foi anteriormente
pré-fixada, você deseja que a mesma seja obtida:
1. Analisando cada amostra individualmente;
2. Analisando todas as amostras conjuntamente.
Digite 1 ou 2.
DATA> 1
```

### Data Display

```
ro     0,383333
```

Portanto, a estimativa do coeficiente de correlação por este método será igual a 0,38333.

### 3.4.1.3 Estimação do Coeficiente de Correlação pelo “Método Bayesiano II (Mingoti IV)”.

Para estimar o coeficiente de correlação por este método o usuário deverá fornecer os parâmetros "a" e "b" da distribuição Uniforme na qual este estimador é fundamentado (ver capítulo 2 deste manual).

Ao escolher este método para estimar o coeficiente de correlação, será emitida na janela *Session do Minitab* a seguinte mensagem:

```
Considerando que o coeficiente de correlação possui uma distribuição Uniforme com parâmetros a e b, forneça à priori:
```

```
O valor da constante a ( $0 \leq a \leq 1$ ):
```

```
DATA> ← Neste campo entre com o valor respectivo à constante "a".
```

```
O valor da constante b ( $0 \leq b \leq 1$ )
```

```
DATA> ← Neste campo entre com o valor respectivo à constante "b".
```

Finalmente será emitido o resultado final na janela *Session do Minitab*:

#### Data Display

```
ro ← Valor estimado pela macro e escolhido pelo usuário.
```

*Exemplo 9:* Considere o conjunto de dados apresentado anteriormente no Quadro 1, seção 3.3.1.2. Neste exemplo, vamos considerar os parâmetros da Uniforme iguais a  $a=0$  e  $b=0,5$ . Neste caso, teremos o seguinte processamento da macro:

```
MTB > %scapai
Executing from file: scapai.MAC

Quantos são os subgrupos analisados (amostras)?
DATA> 10

Os subgrupos (amostras) possuem o mesmo tamanho (Y/N)?
Y

Qual o tamanho de cada subgrupo (amostra)?
DATA> 50
```

Para cada subgrupo (amostra), seus dados estão apresentados em:

1. Total de defeitos;
2. Dados individuais.

Digite 1 ou 2.

DATA> 1

Você deseja estimar a probabilidade de defeito "p" (Y/N)?

Y

Qual é o método de estimação de "p"?

1. Fração de defeituosos na amostra;
2. Cadeia de Markov (somente para dados individuais).

Digite 1 ou 2.

DATA> 1

### Data Display

p 0,100000

Você deseja estimar o valor do coeficiente de correlação (Y/N)?

Y

Qual é o método de estimação do coeficiente de correlação?

1. Método dos momentos;
2. Método de Cadeia de Markov (somente para dados individuais);
3. Método Não-Paramétrico (somente para dados individuais);
4. Método Multinomial (Mingoti I);
5. Método Mingoti II (somente para dados individuais);
6. Método Mingoti III (Método Bayesiano I);
7. Método Mingoti IV (Método Bayesiano II).

Digite 1,2,3,4,5,6 ou 7.

DATA> 7

Considerando que o coeficiente de correlação possui uma distribuição Uniforme com parâmetros a e b, forneça à priori:

O valor da constante a ( $0 \leq a \leq 1$ ):

DATA> 0

O valor da constante b ( $0 \leq b \leq 1$ )

DATA> .5

### Data Display

ro 0,254993

Portanto, a estimativa do coeficiente de correlação por este método será igual a 0,254993.

O interessante na análise dos dados Quadro 1 foram gerados a partir de uma Binomial Generalizada com  $n=50$ , coeficiente de correlação igual a 0,25 e probabilidade p

de defeito igual a 0.10. Esta informação é relevante no momento de entendermos o comportamento dos estimadores construídos fundamentados na distribuição Binomial Generalizada. O que se percebe neste pequeno exemplo, é que o estimador de Madsen modificado apresenta melhores resultados que o estimador original de Madsen apresentado por ele em seu artigo de 1993. Percebe-se também que o estimador Multinomial está superestimando um pouco o verdadeiro valor da correlação, e que os estimadores Bayesianos serão úteis se o investigador tiver alguma informação do processo que pode ser utilizada na escolha dos parâmetros da distribuição a priori assumida para o coeficiente de correlação. Por exemplo, a escolha de  $\alpha$  e  $\beta$  iguais a 1 é equivalente ao modelo uniforme no intervalo  $a=0$  e  $b=1$ . Ou seja, tem-se uma distribuição a priori pouco informativa sobre o comportamento do coeficiente de correlação. No entanto, quando se utiliza uma informação mais refinada como é o caso no qual escolhemos a uniforme no intervalo  $a=0$  e  $b=0,5$  o resultado da estimação melhora consideravelmente, pois a estimativa obtida é praticamente igual ao valor verdadeiro da correlação teórica do processo. Este é o papel claro da Estatística Bayesiana, ou seja, incorporar informações adicionais sobre os parâmetros do processo (vistos como variáveis aleatórias) para melhorar as estimativas finais.

### **3.4.2 Estimação do Coeficiente de Correlação para "*Dados Apresentados Individualmente*".**

Os métodos apresentados a seguir são aplicáveis somente em dados apresentados individualmente. Portanto, caso o usuário tenha apresentado seus dados com formatação "*total de defeitos*", a escolha de um destes métodos para estimar o coeficiente de correlação " $\rho$ ", fará com que seja emitida a seguinte mensagem de erro:

**!!! ERRO !!! Este método pode ser aplicado somente em dados individuais.**

com posterior término do processamento da macro.

### 3.4.2.1 Estimação do Coeficiente de Correlação pelo “Método de Cadeia de Markov”.

O método de *Cadeia de Markov* limita a probabilidade de defeito “p”, em função do valor do coeficiente de correlação “ρ” e vice-versa (Bhat & Lal, 1990). Caso um dos valores de “p” ou de “ρ” não estejam definidos dentro da região de definição da Cadeia e o usuário deseje utilizar o método de *Cadeia de Markov*, será emitida a seguinte mensagem de erro:

```
!!!ERRO!!! Um dos estimadores determinados pelo usuário não pertencem ao intervalo determinado pelo método de Cadeia de Markov  $\max\{0, -r_0/(1-r_0)\} < p < \min\{1, 1/(1-r_0)\}$  para o estimador de p e  $1 - \min\{1/p, 1/(1-p)\} < r_0 < 1$  para o estimador de correlação.
```

Neste caso, a macro terminará o seu processamento.

Embora a Cadeia de Markov não seja definida no caso em o coeficiente de correlação teórico é igual a 1, caso alguma amostra venha a apresentar todos os seus valores iguais a "0" (zero) ou todos os seus valores iguais a "1" (um), para esta amostra o coeficiente de correlação amostral será considerado como sendo igual a 1, e a seguinte mensagem de aviso será emitida:

```
!!!OBSERVAÇÃO!!! Uma das suas amostras possuem dados que indicam  $r_0=1$  e "p"=0, porém o método de Cadeia de Markov não é válido para dados onde  $r_0=1$ .
```

e finalmente será emitido o resultado final na janela *Session do Minitab*:

#### Data Display

**ro**    ← Valor estimado pela macro e escolhido pelo usuário.

Exemplo 10: Considere o conjunto de dados apresentado anteriormente no Quadro 3, seção 3.3.2. Neste caso, teremos o seguinte processamento da macro:

```

MTB > %scapai
Executing from file: scapai.MAC

Quantos são os subgrupos analisados (amostras)?
DATA> 5

Os subgrupos (amostras) possuem o mesmo tamanho (Y/N)?
Y

Qual o tamanho de cada subgrupo (amostra)?
DATA> 60

Para cada subgrupo (amostra), seus dados estão apresentados em:
1. Total de defeitos;
2. Dados individuais.
Digite 1 ou 2.
DATA> 2

Para dados apresentados individualmente, cada coluna deve conter os
dados de uma amostra. Deseja continuar (Y/N)?
Y

Você deseja estimar a probabilidade de defeito "p" (Y/N)?
Y

Qual é o método de estimação de "p"?
1. Fração de defeituosos na amostra;
2. Cadeia de Markov (somente para dados individuais).
Digite 1 ou 2.
DATA> 2

Data Display
p      0,303670

Você deseja estimar o valor do coeficiente de correlação (Y/N)?
Y

Qual é o método de estimação do coeficiente de correlação?
1. Método dos momentos;
2. Método de Cadeia de Markov (somente para dados individuais);
3. Método Não-Paramétrico (somente para dados individuais);
4. Método Multinomial (Mingoti I);
5. Método Mingoti II (somente para dados individuais);
6. Método Bayesiano I (Mingoti III);
7. Método Bayesiano II (Mingoti IV).
Digite 1,2,3,4,5,6 ou 7.
DATA> 2

Data Display
ro     0,137912

```

Portanto, a estimativa do coeficiente de correlação por este método será igual a 0,137912.

Exemplo 11: Considere o conjunto de dados apresentado anteriormente no Quadro 4. seção

3.3.2. Neste caso, teremos o seguinte processamento da macro:

```
MTB > % scapai
Executing from file: scapai.MAC

Quantos são os subgrupos analisados (amostras)?
DATA> 1

Os subgrupos (amostras) possuem o mesmo tamanho (Y/N)?
Y

Qual o tamanho de cada subgrupo (amostra)?
DATA> 208

Para cada subgrupo (amostra), seus dados estão apresentados em:
1. Total de defeitos;
2. Dados individuais.
Digite 1 ou 2.
DATA> 2

Para dados apresentados individualmente, cada coluna deve conter os
dados de uma amostra. Deseja continuar (Y/N)?
Y

Você deseja estimar a probabilidade de defeito "p" (Y/N)?
Y

Qual é o método de estimação de "p"?
1. Fração de defeituosos na amostra;
2. Cadeia de Markov (somente para dados individuais).
Digite 1 ou 2.
DATA> 2

Data Display
p      0,419380

Você deseja estimar o valor do coeficiente de correlação (Y/N)?
Y

Qual é o método de estimação do coeficiente de correlação?
1. Método dos momentos;
2. Método de Cadeia de Markov (somente para dados individuais);
3. Método Não-Paramétrico (somente para dados individuais);
4. Método Multinomial (Mingoti I);
5. Método Mingoti II (somente para dados individuais);
6. Método Bayesiano I (Mingoti III);
7. Método Bayesiano II (Mingoti IV).
Digite 1,2,3,4,5,6 ou 7.
DATA> 2

Data Display
ro     0,158422
```

Portanto, os dados do *Quadro 4* levariam a uma estimativa de probabilidade de defeito “p” igual a 0,419380 e do coeficiente de correlação “ $\rho$ ” igual a 0,158422.

### 3.4.2.2 Estimação do Coeficiente de Correlação pelo “Método Não-Paramétrico”.

Neste método, somente o valor estimado pela macro será apresentado na janela *Session do Minitab*:

Exemplo 12: Considere o conjunto de dados apresentado no Quadro 3, seção 3.3.2. Neste caso, teremos o seguinte processamento da macro:

```
MTB > % scapai
Executing from file: scapai.MAC

Quantos são os subgrupos analisados (amostras)?
DATA> 5

Os subgrupos (amostras) possuem o mesmo tamanho (Y/N)?
Y

Qual o tamanho de cada subgrupo (amostra)?
DATA> 60

Para cada subgrupo (amostra), seus dados estão apresentados em:
1. Total de defeitos;
2. Dados individuais.
Digite 1 ou 2.
DATA> 2

Para dados apresentados individualmente, cada coluna deve conter os
dados de uma amostra. Deseja continuar (Y/N)?
Y

Você deseja estimar a probabilidade de defeito "p" (Y/N)?
Y

Qual é o método de estimação de "p"?
1. Fração de defeituosos na amostra;
2. Cadeia de Markov (somente para dados individuais).
Digite 1 ou 2.
DATA> 2

Data Display

p      0,303670

Você deseja estimar o valor do coeficiente de correlação (Y/N)?
Y

Qual é o método de estimação do coeficiente de correlação?
1. Método dos momentos;
```

```
2. Método de Cadeia de Markov (somente para dados individuais);
3. Método Não-Paramétrico (somente para dados individuais);
4. Método Multinomial (Mingoti I);
5. Método Mingoti II (somente para dados individuais);
6. Método Bayesiano I (Mingoti III);
7. Método Bayesiano II (Mingoti IV).
Digite 1,2,3,4,5,6 ou 7.
DATA> 3
```

#### **Data Display**

```
ro      0,294915
```

Portanto, os dados do Quadro 3 levariam a uma estimativa de probabilidade de defeito “p” igual a 0,303670 e do coeficiente de correlação “ $\rho$ ” igual a 0,294915 pelo método Não-Paramétrico.

#### **3.4.2.3 Estimação do Coeficiente de Correlação pelo “Método Mingoti II”.**

Neste método, somente o valor estimado pela macro será apresentado na janela *Session do Minitab*:

#### **Data Display**

```
ro      ⇐ Valor estimado pela macro
```

*Exemplo 13:* Considere novamente o conjunto de dados apresentado no Quadro 3, seção 3.3.2. O processamento da macro se dará da seguinte forma:

```

MTB > % scapai
Executing from file: scapai.MAC

Quantos são os subgrupos analisados (amostras)?
DATA> 5

Os subgrupos (amostras) possuem o mesmo tamanho (Y/N)?
Y

Qual o tamanho de cada subgrupo (amostra)?
DATA> 60

Para cada subgrupo (amostra), seus dados estão apresentados em:
1. Total de defeitos;
2. Dados individuais.
Digite 1 ou 2.
DATA> 2

Para dados apresentados individualmente, cada coluna deve conter os
dados de uma amostra. Deseja continuar (Y/N)?
Y

Você deseja estimar a probabilidade de defeito "p" (Y/N)?
Y

Qual é o método de estimação de "p"?
1. Fração de defeituosos na amostra;
2. Cadeia de Markov (somente para dados individuais).
Digite 1 ou 2.
DATA> 2

Data Display

p      0,303670

Você deseja estimar o valor do coeficiente de correlação (Y/N)?
Y

Qual é o método de estimação do coeficiente de correlação?
1. Método dos momentos;
2. Método de Cadeia de Markov (somente para dados individuais);
3. Método Não-Paramétrico (somente para dados individuais);
4. Método Multinomial (Mingoti I);
5. Método Mingoti II (somente para dados individuais);
6. Método Bayesiano I (Mingoti III);
7. Método Bayesiano II (Mingoti IV).
Digite 1,2,3,4,5,6 ou 7.
DATA> 5

Data Display

ro     0,166387

```

### 3.5 Construindo o Gráfico (Carta) de Controle de *SHEWHART*.

Nesta seção, vamos mostrar como o usuário poderá construir gráficos de controle para o número total de defeitos e para a probabilidade de gerar itens defeituosos.

O programa *SCAP*, módulo *AI*, permite que o usuário construa gráficos de controle de *Shewhart* considerando a distribuição Binomial Generalizada Exata e a aproximação Normal no caso de Cadeias de Markov. No caso em que o coeficiente de correlação é igual a zero, o modelo da Binomial Generalizada torna-se o modelo Binomial clássico comum e neste caso os limites dos gráficos de controle são construídos através da aproximação Normal.

Para obter os gráficos de controle o usuário deverá responder às seguintes questões.

**Você deseja traçar um gráfico de controle (Y/N)?**

**DATA>** ← Responda “Y” para resposta afirmativa e “N” para resposta negativa.

Caso a usuário responda afirmativamente:

**Você deseja traçar o gráfico usando qual distribuição?**

1. Binomial Generalizada (distribuição exata);
2. Cadeia de Markov (aproximação Normal).

**Digite 1 ou 2.**

**DATA>** ← Neste campo entre com o código respectivo ao método desejado.

Caso o usuário responda negativamente a macro será finalizada.

#### 3.5.1 Obtenção do Gráfico de Controle pela “*Distribuição Binomial Generalizada*”.

A *Distribuição Binomial Generalizada* limita o nível de significância  $\alpha$ , e o nível de confiança  $(1-\alpha)(100)\%$  que poderá ser utilizado para o cálculo dos limites de controle do processo (limites de confiança),  $0 < \alpha < 1$ . Para maiores detalhes técnicos sobre este fato, ver capítulo 2 deste manual. Com base no valor inferior assumido por  $\alpha$  que é calculado pela macro, o usuário deverá escolher qual o nível de confiança que deseja para construção dos limites e posteriormente do gráfico (carta) de controle.

Para a construção do gráfico de controle o usuário deverá responder à questão:

nível de significância mínimo

Qual o nível de significância (em %)?

DATA> ⇐ Neste campo entre com o valor respectivo do nível de significância.

e posteriormente, o usuário terá que decidir sobre qual formatação é desejada do gráfico, respondendo as questões:

Qual o gráfico de controle desejado?

1. Para o número de defeitos (np);
2. Para a proporção de defeitos (p);
3. Um para o número e outro para a proporção de defeitos.

Digite 1,2 ou 3.

DATA> ⇐ Neste campo entre com o valor respectivo à constante beta.

Sempre serão emitidos na janela *Session do Minitab* os valores obtidos para os limites do intervalo (gráfico) isto é, os limites médio e os limites inferior e superior (LCL;UCL), construídos de acordo com o nível de significância estabelecido pelo usuário.

Exemplo 14: Considere o conjunto de dados apresentado no Quadro 1 seção 3.3.1.1, e o modelo *Binomial Generalizado*. A probabilidade de defeito e o coeficiente de correlação foram anteriormente estimados pela *fração de defeituosos na amostra* e pelo *método dos Momentos*, fornecendo uma estimativa de “p” igual a 0,1 e uma estimativa de “p” igual a 0,244646 (ver exemplo 5). Considerando-se esses valores estimados, o gráfico de controle seria traçado através do seguinte procedimento:

```

Você deseja traçar um gráfico de controle (Y/N)?
Y
Você deseja traçar o gráfico usando qual distribuição?
1. Binomial Generalizada (distribuição exata);
2. Cadeia de Markov (aproximação Normal).
Digite 1 ou 2.
DATA> 1

O nível de significância mínimo (em porcentagem) admitido para esta
distribuição é de:

Data Display

nível de significância mínimo 39,2653

Qual o nível de significância deseja para construir o gráfico(em %)?

DATA> 40

Qual o gráfico de controle desejado?
1. Para o número de defeitos (np);
2. Para a proporção de defeitos (p);
3. Um para o número e outro para a proporção de defeitos.
Digite 1,2 ou 3.
DATA> 3

Data Display

media      5,00000

Data Display

LCL      0

Data Display

UCL      11,0000

NP Chart: N°Defeituosos

Data Display

média das proporções      0,100000

Data Display

LCL      0

Data Display

UCL      0,220000

```

Neste caso, obteríamos os gráficos de controle para o número de defeituosos e para a proporção de defeituosos do processo como apresentados na Figura 1.

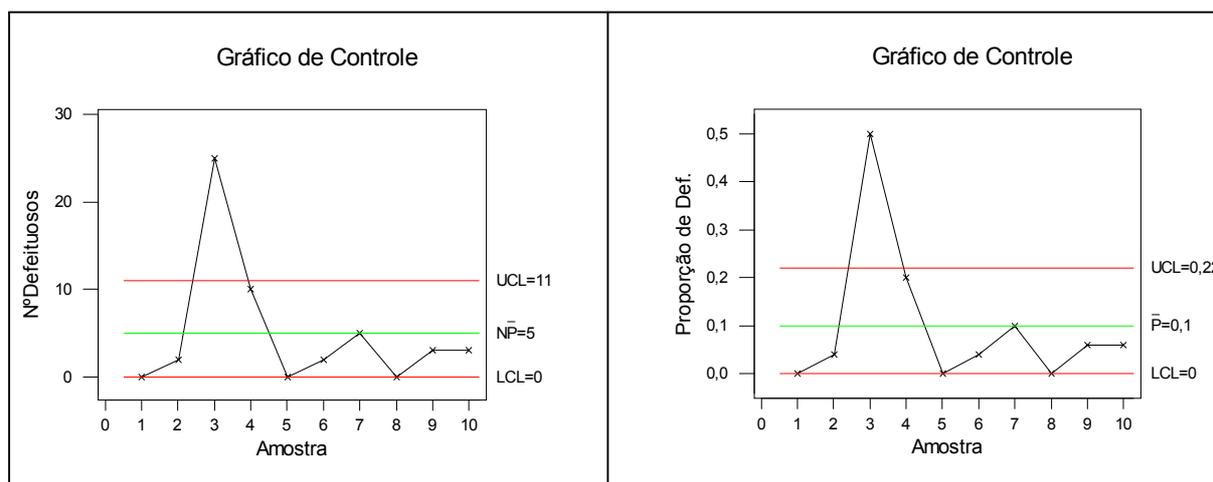


Figura 1: Gráfico de controle para o número de defeitos e proporção de defeitos referente ao conjunto de dados do Quadro 1 - Binomial Generalizada.

Caso o coeficiente de correlação “ $\rho$ ” seja igual a zero, os gráficos serão obtidos através da distribuição Binomial clássica usando-se a aproximação normal. Neste caso, não haverá limitação no nível de significância mínimo que o usuário poderá escolher para construir os gráficos de controle. O usuário poderá construir gráficos de controle com níveis de confiança pré-especificados. O programa permite a escolha dos seguintes níveis de confiança: 68,3, 95,4 e 99,7%. Para determinação do gráfico de controle o usuário deverá então, responder a seguinte mensagem:

**Qual o nível de confiança do gráfico de controle?**

1. 68,3% (1 desvio padrão na Normal);
2. 95,4% (2 desvios padrões na Normal);
3. 99,7% (3 desvios padrões na Normal);

**Digite 1,2 ou 3.**

**DATA>** ⇐ Neste campo entre com o valor respectivo ao nível de significância.

Posteriormente será emitida a seguinte mensagem de aviso na janela *Session do Minitab*:

**Considerando que o valor do coeficiente de correlação assumido é igual a zero, o gráfico será obtido com base na teoria da Distribuição Binomial Clássica via aproximação pela distribuição Normal.**

**Exemplo 15:** Considere novamente o conjunto de dados apresentado no Quadro 2 seção 3.3.1.2, e o modelo *Binomial Generalizado*. A probabilidade de defeito e o coeficiente de correlação foram anteriormente estimados pela fração de defeituosos na amostra e pelo método dos Momentos, fornecendo uma estimativa de “p” igual a 0,0201981 e uma estimativa de “ρ” igual a 0,00222698 (ver exemplo 6). Considerando-se esses valores estimados, o gráfico de controle seria traçado através do seguinte procedimento:

```
Você deseja traçar um gráfico de controle (Y/N)?
y
Você deseja traçar o gráfico usando qual distribuição?
1. Binomial Generalizada (distribuição exata);
2. Cadeia de Markov (aproximação Normal).
Digite 1 ou 2.
DATA> 1
O nível de significância mínimo (em porcentagem) admitido para esta
distribuição é de:
Data Display
nível de significância mínimo: 0,436400
Qual o nível de significância (em %)?
DATA> .5
Qual o gráfico de controle desejado?
1. Para o número de defeitos (np);
2. Para a proporção de defeitos (p);
3. Um para o número e outro para a proporção de defeitos.
Digite 1,2 ou 3.
DATA> 3
Data Display
media 9,29114
Data Display
LCL 1,00000
Data Display
UCL 21,0000
NP Chart: NºDefeituosos
Data Display
média das proporções 0,0201981
```

## Data Display

LCL 0,00217391

## Data Display

UCL 0,0456522

Neste caso, obteríamos os gráficos de controle para o número de defeituosos e para a proporção de defeituosos do processo como apresentados na Figura 2.

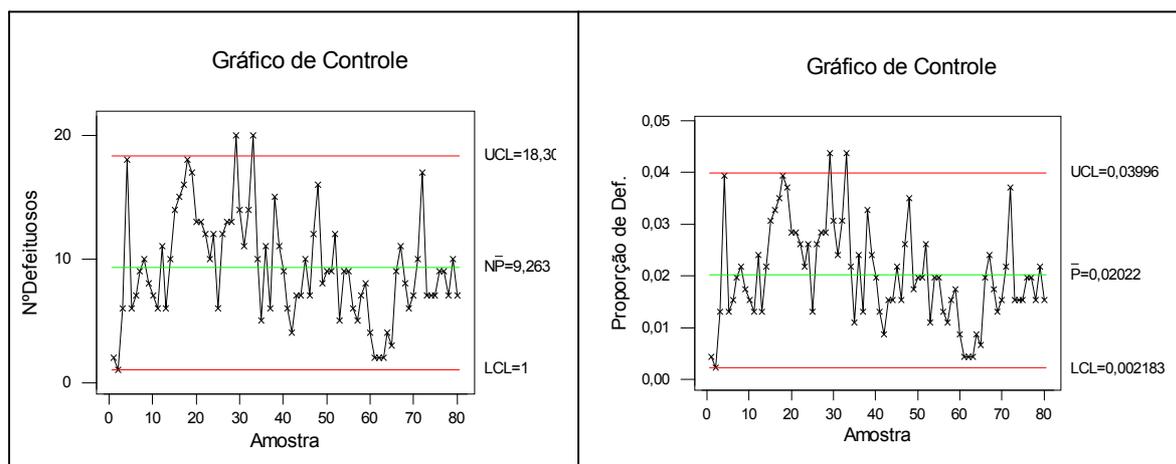


Figura 2: Gráfico de controle para o número de defeitos e proporção de defeitos referente ao conjunto de dados do Quadro 2 - Binomial Generalizada.

### 3.5.2 Obtenção do Gráfico de Controle por “Cadeia de Markov”.

A metodologia de *Cadeia de Markov* permite ao usuário escolher qual o nível de significância que deseja para construção dos limites do gráfico (carta) de controle. O gráfico é construído de acordo com a aproximação Normal (Bhat e Lal, 1990). O usuário poderá construir gráficos de controle de 68,3, 95,4 ou 99,7% de confiança. Para a determinação do gráfico será emitida na janela *Session do Minitab* a seguinte mensagem:

**Qual o nível de confiança do gráfico de controle?**

1. 68,3% (1 desvio padrão na Normal);
2. 95,4% (2 desvios padrões na Normal);
3. 99,7% (3 desvios padrões na Normal);

Digite 1,2 ou 3.

**DATA>** ⇐ Neste campo entre com o valor respectivo ao nível de significância.

Exemplo 16: Considere o conjunto de dados apresentado no Quadro 3 seção 3.3.1.1, e o modelo considerado de Cadeia de Markov. A probabilidade de defeito e o coeficiente de correlação foram anteriormente estimados por Cadeia de Markov, fornecendo uma estimativa de “p” igual a 0,303670 e uma estimativa de “ $\rho$ ” igual a 0,137912 (ver exemplo 10). Considerando-se esses valores estimados, o gráfico de controle seria traçado através do seguinte procedimento:

```
Você deseja traçar um gráfico de controle (Y/N)?
y
Você deseja traçar o gráfico usando qual distribuição?
1. Binomial Generalizada (distribuição exata);
2. Cadeia de Markov (aproximação Normal).
Digite 1 ou 2.
DATA> 2
Qual o gráfico de controle desejado?
1. Para o número de defeitos (np);
2. Para a proporção de defeitos (p);
3. Um para o número e outro para a proporção de defeitos.
Digite 1,2 ou 3.
DATA> 3
Qual o nível de confiança do gráfico de controle?
1. 68,3% (1 desvio padrão na Normal);
2. 95,4% (2 desvios padrões na Normal);
3. 99,7% (3 desvios padrões na Normal);
Digite 1,2 ou 3.
DATA> 3
Data Display
media      18,6000
Data Display
var        12,6873
Data Display
LCL        7,91424
Data Display
UCL        29,2858
Data Display
média das proporções      0,310000
```

### Data Display

var 12,6873

### Data Display

LCL 0,125574

### Data Display

UCL 0,481766

Neste caso, obteríamos os gráficos de controle para o número de defeituosos e para a proporção de defeituosos do processo apresentados na Figura 3. É importante observar que o programa *SCAP - módulo AI* permite que o usuário refaça sua análise usando outros estimadores ou faça outros gráficos de controle sem que este tenha que sair do programa principal e recomeçar todo o processamento novamente.

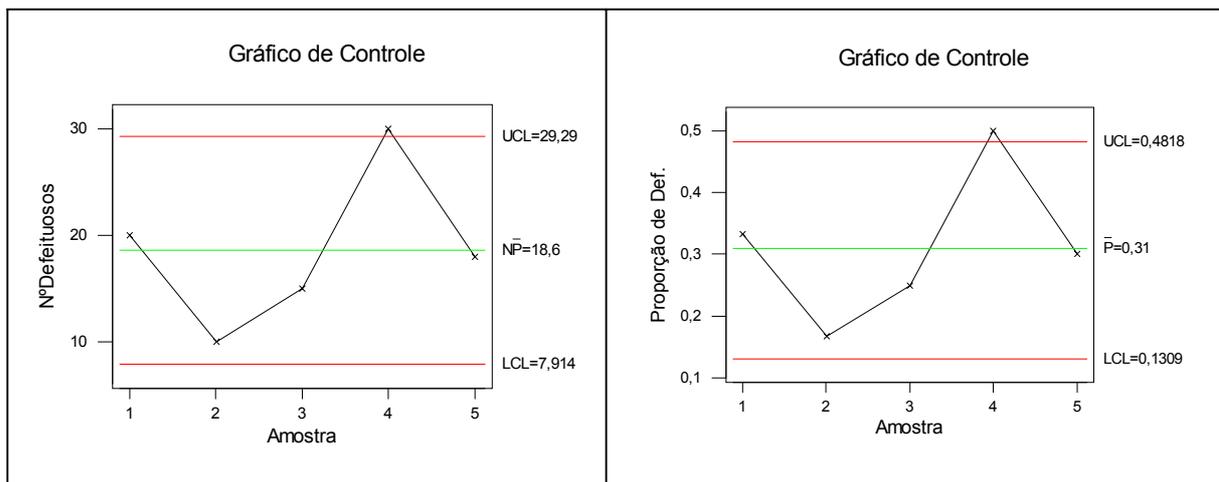


Figura 3: Gráfico de controle para o número de defeitos e proporção de defeitos referente ao conjunto de dados do Quadro 3 - Cadeia de Markov.

Exemplo 17: Considere o conjunto de dados apresentado no Quadro 4 seção 3.3.2. e o modelo de Cadeia de Markov. A probabilidade de defeito e o coeficiente de correlação foram anteriormente estimados por *Cadeia de Markov*, fornecendo uma estimativa de “p” igual a 0,419380 e uma estimativa de “ $\rho$ ” igual a 0,158422 (ver exemplo 11). Considerando esses valores estimados, o gráfico de controle seria traçado através do seguinte procedimento:

Você deseja traçar um gráfico de controle (Y/N)?

Y

Você deseja traçar o gráfico usando qual distribuição?

1. Binomial Generalizada (distribuição exata);
2. Cadeia de Markov (aproximação Normal).

Digite 1 ou 2.

DATA> 2

Qual o gráfico de controle desejado?

1. Para o número de defeitos (np);
2. Para a proporção de defeitos (p);
3. Um para o número e outro para a proporção de defeitos.

Digite 1,2 ou 3.

DATA> 3

Qual o nível de confiança do gráfico de controle?

1. 68,3% (1 desvio padrão na Normal);
2. 95,4% (2 desvios padrões na Normal);
3. 99,7% (3 desvios padrões na Normal);

Digite 1,2 ou 3.

DATA> 3

#### Data Display

media 88,0000

#### Data Display

var 50,6481

#### Data Display

LCL 66,6498

#### Data Display

UCL 109,350

#### Data Display

média das proporções 0,423077

#### Data Display

var 50.6481

#### Data Display

LCL 0,316735

#### Data Display

UCL 0,522026

Você deseja reestimar algum parâmetro ou traçar novamente uma carta de controle (Y/N)?

n

Neste caso, obteríamos os gráficos de controle para o número de defeituosos e para a proporção de defeituosos do processo como apresentados na Figura 4.

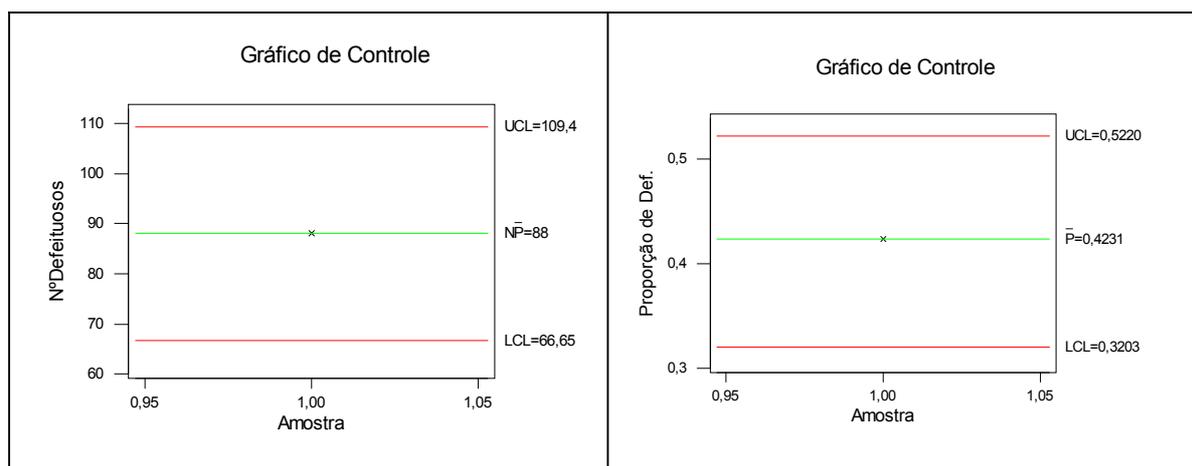


Figura 4: Gráfico de controle para o número de defeitos e proporção de defeitos referente ao conjunto de dados do Quadro 4 - Cadeia de Markov

É importante notar que neste exemplo há apenas uma amostra sendo avaliada e portanto, o gráfico de controle apresenta os limites de controle mas apenas um valor amostral do número de defeituosos é grafado.

### 3.6 Considerações Finais

Neste capítulo apresentamos exemplos de aplicação do *software SCAP módulo AI*. O usuário tem à sua disposição vários estimadores para os parâmetros  $p$  e  $\rho$  do processo. Caberá a ele fazer a opção pelo método de estimação que deseja utilizar respeitando a forma natural em que os dados foram digitados na planilha de dados do *Minitab for Windows*. Os modelos de *Cadeia de Markov* se aplicam apenas para o caso de dados individuais enquanto que o modelo *Binomial Generalizado* se aplica para ambas as situações. No entanto, este modelo só é válido para processos nos quais o coeficiente de correlação é maior ou igual zero enquanto que o modelo de *Cadeia de Markov* permite que o coeficiente de correlação seja negativo, dependendo do valor assumido pela probabilidade de defeitos  $p$  do processo. Os estimadores Bayesianos para o coeficiente de correlação no caso da distribuição Binomial Generalizada permite que o usuário utilize seu conhecimento prévio do processo de produção para escolha dos parâmetros  $\alpha$  e  $\beta$  (alpha e Beta) no caso do estimador Mingoti III e dos parâmetros  $a$  e  $b$ , no caso do estimador Mingoti IV e caso este conhecimento prévio seja confiável eles produzirão estimativas melhores que os estimadores clássicos. Além disso, vale destacar que caso o usuário tenha à disposição apenas uma amostra de tamanho  $n$  ele poderá estimar  $\rho$  se optar pelos estimadores Bayesianos. No entanto, para utilizar o estimador de Madsen necessitará de um número maior de amostras. É importante salientar a boa qualidade do estimador Multinomial para amostras de tamanhos maiores ou iguais a 100, faixa na qual obtém-se estimativas mais precisas que as de Madsen.

Finalmente, para amostras onde se tem a informação do estado individual de cada item inspecionado, o estimador Não-Paramétrico pode constituir uma boa alternativa para estimação de  $\rho$  visto que além de produzir boas estimativas ele também independe do conhecimento da distribuição original que está gerando os resultados amostrais.

## Referências Bibliográficas

- Altham, P. M. E. (1978) Two generalizations of the binomial distribution, *Applied Statistics*, 27, pp. 162-167.
- Berger, J. O. (1985) *Statistical decision theory and bayesian analysis*. 2<sup>nd</sup>, New York: Springer Verlag.
- Bhat, U.N., Lal, R. (1990) Attribute control charts for Markov dependent production processes, *IIE Transactions*, 22, 2, 181-188.
- Bourke, P. D. (1991) Detecting a shift in fraction of nonconforming using run-length control charts with 100% inspection, *Journal of Quality Technology*, 23, pp. 225-328.
- Broadbent, S. R. (1958) The inspection of a Markov process. *Journal of Royal Statistical Society, Series B*, 20, 111-119.
- Calvin, T. W. (1983) Quality control techniques for zero-defects, *IEEE Transactions on Components, Hybrids and Manufacturing Technology*, CHMT-6, pp. 323-328.
- Garthwaite, P. H., Jolliffe, I. T., Jones, B. (1995) *Statistical inference*. New York: Prentice Hall.
- Goh, T. N. (1989) A charting technique for control of low nonconforming production, *International Journal of Quality and Reliability Management*, 4, pp. 53-62.
- Goh, T. N. e Xie, M. (1995) Statistical process control for low nonconforming processes, *International Journal of Reliability Quality and Safety Engineering*, 2, pp. 15-22.
- Karlin, S., Taylor, H. (1975) *A first course in stochastic processes*. New York: Academic Press.
- Lai, C. D. , Govindaraju, K. e Xie, M. (2000) Study of a Markov model for a high-quality dependent process, *Journal of Applied Statistics*, 27, 4, pp. 461-473.
- Lai, C. D. , Govindaraju, K. e Xie, M. (1998) Effects of correlation on fraction nonconforming statistical process control procedures, *Journal of Applied Statistics*, 25, 4, pp. 535-543.
- Lucas, J. M. (1989) Control schemes for low count levels, *Journal of Quality Technology*, 21, pp. 199-201.
- Madsen, R. W. (1993) Generalized binomial distributions, *Communications in Statistics: Theory and Methods*, 22, pp. 3065-3086.
- MacShane, L. M. e Turnbull, B. W. (1991) Probability limits on outgoing quality for continuous sampling plans, *Technometrics*, 33, pp. 393-404.
- Nelson, L. S. (1994) A control chart for parts-per-million nonconforming itens. *Journal of Quality Technology*, 19, pp. 239-240.

- Mingoti, S. A., Carvalho, J.P. (2002) O programa SCAP para monitoramento de processos autocorrelacionados e inspeção por atributos. *Anais da III Jornada Regional de Estatística e II Semana da Estatística*, Universidade Estadual de Maringá, 124-132.
- Mingoti, S. A., Carvalho, J.P. (2001) Monitoramento de processos autocorrelacionados: usando o software Minitab no controle via análise de atributos. *Caderno de Resumos da 9a. Escola de Séries Temporais e Econometria (9a. ESTE)*, Belo Horizonte: ABE/Depto. de Estatística-UFGM, p.119.
- Montgomery, D. C. (1991) *Introduction to statistical quality control*. New York: John Wiley.
- Bhat, U. N., Lal, R. (1988) Number of success in Markov trials. *Advanced Applied Probability*, 20, 677-680.
- Sampath Kumar, V. S. e Rajarshi, M. B. (1987) Continuous sampling plans for markov-dependent production process, *Naval Research Logistics*, 34, pp. 629-644.
- Sprenst, P. (1989) *Applied Nonparametric statistical methods*. New York: Chapman Hall.
- Stimson, W. A. e Mastrangelo, C. M. (1996) Monitoring serially dependent process with attributes data, *Journal of Quality Technology*, 28, pp. 279-288.
- Xie, M. e Goh, T. N. (1992) Some procedures for decision making om high-yield processes, *Quality and Reliability Engineering International*, 8, pp. 355-360.
- Xie, W. , Xie, M. e Goh, T. N. (1995) Control charts for processes subject to random shocks, *Quality and Reliability Engineering International*, 11, pp. 355-360.