

FABIANA DE MATOS MOURA

AVALIAÇÃO BAYESIANA DE UM SISTEMA DE INSPEÇÃO
DICOTÔMICO

Dissertação apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais como requisito parcial à obtenção do título de Mestre em Estatística.

Orientador: Prof. Roberto da Costa Quinino

BELO HORIZONTE, JUNHO 2007

Agradecimentos

A realização deste trabalho tornou-se possível graças à contribuição de instituições e pessoas das quais gostaria de expressar meus sinceros agradecimentos:

- À UFMG, que inclui especialmente todo o Departamento de Estatística, pela oportunidade de obtenção de mais qualificação profissional;
- Ao CNPq, pela concessão de bolsa de estudo para a realização do curso;
- Ao mestre e orientador Roberto Quinino, pela atenção e confiança despendidas, pela sabedoria na orientação e pelos ensinamentos que levarei para a vida;
- Ao colega e grande profissional Anderson Laécio, pela disponibilidade e presteza nas artes computacionais. Às professoras Arminda, Emília e Linda Lee Ho, componentes da banca examinadora, por suas contribuições e atenção dispensadas para a conclusão deste trabalho;
- Ao Leo, meu grande companheiro, pelo amor, pelo otimismo, pelos momentos de descontração e por todo suporte oferecido;
- Aos meus pais Luiza e Juarez, aos meus irmãos Polly e Marcelo, aos amigos, colegas e familiares e a Deus, que também tornaram mais esta conquista possível.

Avaliação Bayesiana de um Sistema de Inspeção Dicotômico

Fabiana de Matos Moura Roberto da Costa Quinino
Orientador

Departamento de Estatística - ICEX - UFMG
31270-901 - Belo Horizonte - MG - Brasil

Junho - 2007

Resumo

A avaliação da conformidade de itens é um procedimento importante dentro do controle de qualidade por atributos. Entretanto, o sistema que realiza a classificação pode não estar isento de cometer erros de classificação. Dois tipos de erros são possíveis: o primeiro, denotado por erro tipo I, ocorre quando um item conforme é classificado como não-conforme; e o segundo, denotado por erro tipo II, ocorre quando um item não-conforme é classificado como conforme. Neste trabalho, um método Bayesiano é utilizado para avaliar a eficiência de um sistema de inspeção, baseado na realização de classificações repetidas dos itens inspecionados. Supõe-se que a classificação real dos itens avaliados seja desconhecida.

Palavras-Chave: Atributos, Erros de classificação, Classificações repetidas, Análise bayesiana, Sistema de inspeção dicotômico.

1 Introdução

Um sistema de controle, visto de forma ampla, pode ser considerado como um programa planejado e documentado compreendendo o pessoal, máquinas, equipamentos, instrumentos, procedimentos e padrões. Tudo com o objetivo de proporcionar uma melhoria contínua da qualidade. A inspeção direta de um produto ou serviço é uma atividade necessária dentro do sistema de controle e, para tanto, é fundamental que o sistema de inspeção a ser empregado seja eficiente. Desta forma, evita-se que sejam tomadas decisões equivocadas tais como: utilizar itens não-conformes ou descartar itens conformes.

Quando a inspeção baseia-se em atributos, isto é, inspeção na qual uma ou mais características da qualidade são analisadas, a unidade de produto ou serviço correspondente, que parte de um lote, batelada ou população, pode ser classificada simplesmente como conforme ou não em relação a uma dada especificação. Baseando-se nas frequências de conformidade observadas, determina-se a necessidade ou não de interferência no processo produtivo.

Um dos grandes problemas em estabelecer um sistema de inspeção para produtos classificados como conforme ou não-conforme está na falta de informações sobre a eficiência do sistema de classificação utilizado. Dois tipos de erros podem ocorrer: o primeiro, conhecido como erro tipo I, ocorre quando um item conforme é classificado como não-conforme; e o segundo, denotado por erro tipo II, ocorre quando um item não-conforme é classificado como conforme.

Os efeitos de se ignorar erros de classificação foram pioneiramente estudados por Bross (1954), que mostrou que estimações clássicas baseadas em proporções são viciadas. Em termos práticos, a estimativa da proporção de interesse poderia estar comprometida e decisões poderiam ser tomadas equivocadamente. Outros autores como Johnson e Kotz (1988), Johnson *et al.* (1991), Greenberg e Stokes (1995), Evans *et al.* (1996) e, Gustafson (2003) enfatizaram que os erros de classificação não podem ser negligenciados, pois eles podem comprometer seriamente o processo de inspeção e tomada de decisão. Burke *et al.*

(1995), Gramopadhye *et al.* (1995) e, Kemp e Kemp (1988) também relatam exemplos de erros de diagnóstico que influenciam criticamente o processo de avaliação. Neste sentido, é evidente que o controle de qualidade por atributos depende consideravelmente da eficiência do sistema de inspeção adotado.

Em geral, as metodologias existentes requerem o conhecimento do estado real dos itens avaliados para viabilizar a análise dos erros tipo I e tipo II. Barry e Christopher (1981), Johnson e Funke (1980) e Sylla (2002) apresentaram métodos para julgar o sistema de inspeção quando a classificação real dos itens é conhecida. Quando não há tal conhecimento, Johnson *et al.* (1991) mostraram que não é possível estimar os erros de classificação sem a utilização de informações adicionais. Gaba e Winkler (1992) e Joseph *et al.* (1995) propuseram métodos bayesianos para incorporar os erros de classificação. Funções de densidade marginal *a posteriori* dos erros foram obtidas, tornando possível julgar o sistema de inspeção. No entanto, tais métodos dependem fortemente da especificação de uma distribuição *a priori* para os parâmetros, que necessita ser muito informativa para obtenção de uma distribuição *a posteriori* capaz de gerar decisões confiáveis.

Obter distribuições *a priori* muito informativas pode não ser possível, porque a necessidade de julgar o sistema de inspeção tem como principal suposição alta incerteza sobre a magnitude dos erros de classificação cometidos, contradizendo a especificação de uma distribuição *a priori* muito informativa.

Quinino *et al.* (2005), em uma abordagem bayesiana para a avaliação de inspetores em classificações dicotômicas, utilizaram distribuições *a priori* pouco informativas e uma função de verossimilhança incompleta. Classificaram cada item avaliado m vezes e consideraram o estado real do item avaliado como conforme se a maioria de suas classificações fosse conforme. Com a reconstrução do estado real dos itens, procederam uma análise bayesiana usual. No entanto, incorporaram uma fonte de variação indesejada, uma vez que a reconstrução do estado real dos itens inspecionados pode apresentar equívocos.

A proposta deste artigo é apresentar um método bayesiano que utilize a função de verossimilhança completa - que não necessita da reconstrução do estado real dos itens

avaliados - sem a necessidade de distribuições *a priori* muito informativas - pequena variabilidade. O objetivo principal é estimar os erros de classificação e, conseqüentemente, julgar o sistema de inspeção dicotômico como qualificado ou não-qualificado. Para tanto, usaremos classificações repetidas, que podem constituir-se em fonte adicional de informação e, conseqüentemente, diminuir a necessidade de informações *a priori* muito informativas. Além disto, discutimos também o tamanho amostral e o número de classificações repetidas necessários para garantir uma “precisão” *a posteriori* pré-fixada.

Na seção 2 apresentamos o esquema utilizado para incorporar as classificações repetidas, bem como a determinação da função de verossimilhança. A análise bayesiana dos erros e a determinação do tamanho amostral são apresentadas respectivamente nas seções 3 e 4. Um exemplo numérico é exposto na seção 5 e as considerações finais na seção 6.

2 Modelo Probabilístico e Função de Verossimilhança

Considere que uma amostra aleatória de n itens seja retirada de um processo de produção e que cada item seja julgado como conforme ou não-conforme, sendo p a probabilidade de conformidade. O número de itens conformes observados (X) possui distribuição Binomial com parâmetros (n, p) , ou simbolicamente, $X \sim Bin(n, p)$. Contudo, a existência de erros de classificação no sistema implica numa modificação dessa função de probabilidade. Seja e_1 a probabilidade de que um item conforme seja erroneamente classificado como não-conforme, e seja e_2 a probabilidade de que um item não-conforme seja classificado como conforme. Então, a probabilidade de que um item seja classificado como conforme é dada por $q = p(1 - e_1) + (1 - p)e_2$. Conseqüentemente, a variável aleatória X terá distribuição binomial com parâmetro q ao invés de p .

Suponha agora que cada item da amostra aleatória de tamanho n seja classificado m vezes, independentemente, como conforme ou não-conforme. Seja C_{ij} ($i = 1, 2, \dots, n$; $j = 1, 2, \dots, m$) a variável aleatória Bernoulli(q) correspondente à j -ésima classificação do i -

ésimo item. Desta forma, $C_{12} = 1$ significa que o primeiro item da amostra foi classificado como conforme em sua segunda classificação. Seja C_i o total de classificações conformes do i -ésimo item após as m classificações, isto é, $C_i = \sum_{j=1}^m C_{ij}$. A Tabela 1 apresenta a descrição desse procedimento de classificação para uma melhor visualização.

Tabela 1: Classificações repetidas de n itens m vezes cada

Item	Resultado da classificação (C_{ij})					Classificações Conformes C_i
	1	2	3	\dots	m	
1	C_{11}	C_{12}	C_{13}	\dots	C_{1m}	C_1
2	C_{21}	C_{22}	C_{23}	\dots	C_{2m}	C_2
3	C_{31}	C_{32}	C_{33}	\dots	C_{3m}	C_3
\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
n	C_{n1}	C_{n2}	C_{n3}	\dots	C_{nm}	C_n

A probabilidade de observarmos $C = (C_1; C_2; \dots; C_n)$ é dada por:

$$P(C = c) = \prod_{i=1}^n \left[p \binom{m}{c_i} (1 - e_1)^{c_i} (e_1)^{m-c_i} + (1 - p) \binom{m}{c_i} (e_2)^{c_i} (1 - e_2)^{m-c_i} \right] \quad (1)$$

Seja $n_k = \sum_{i=1}^n I_{\{c_i=k\}}$, $k = 0, 1, \dots, m$ a frequência de cada um dos resultados possíveis de C_i . Assim, expressão (1) pode ser reescrita como:

$$P(C = c) = \prod_{k=0}^m \left[p \binom{m}{k} (1 - e_1)^k (e_1)^{m-k} + (1 - p) \binom{m}{k} (e_2)^k (1 - e_2)^{m-k} \right]^{n_k} \quad (2)$$

Desenvolvendo a expressão (2), pelo uso do Binômio de Newton e adotando $\eta = (n_0; n_1; \dots; n_m)$, podemos expressar a função de verossimilhança por:

$$L(\eta|p, e_1, e_2, m) \propto \prod_{k=0}^m \sum_{l_k=0}^{n_k} \binom{n_k}{l_k} [p(1 - e_1)^k (e_1)^{m-k}]^{l_k} [(1 - p)(e_2)^k (1 - e_2)^{m-k}]^{n_k - l_k} \quad (3)$$

A expressão (3) pode ainda ser reescrita e expressa como:

$$L(\eta|p, e_1, e_2, m) \propto \sum_{l_0=0}^{n_0} \sum_{l_1=0}^{n_1} \cdots \sum_{l_m=0}^{n_m} \binom{n_0}{l_0} \binom{n_1}{l_1} \cdots \binom{n_m}{l_m} p^{x_1} (1-p)^{x_2} e_1^{x_3} (1-e_1)^{x_4} e_2^{x_5} (1-e_2)^{x_6} \quad (4)$$

em que

$$\begin{aligned} x_1 &= \sum_{k=0}^m l_k; & x_2 &= n - \sum_{k=0}^m l_k; \\ x_3 &= \sum_{k=0}^m (m-k)l_k; & x_4 &= \sum_{k=0}^m kl_k; \\ x_5 &= \sum_{k=0}^m k(n_k - l_k); & x_6 &= \sum_{k=0}^m (m-k)(n_k - l_k); \\ n_k &= 0, 1, \dots, n. \end{aligned}$$

3 Análise Bayesiana dos Erros de Classificação

Considere uma distribuição *a priori* conjunta de (p, e_1, e_2) dada por:

$$f(p, e_1, e_2) = f_\beta(p|\alpha, \beta) f_\beta(e_1|\alpha_1, \beta_1) f_\beta(e_2|\alpha_2, \beta_2) \quad (5)$$

em que $f_\beta(a|b, c)$ é uma função densidade Beta para a variável aleatória a com parâmetros b e c . Distribuições Beta são amplamente utilizadas em modelos bayesianos para descrever informações sobre proporções [Gupta & Nadarajah (2004)]. Neste artigo, consideramos que as variáveis aleatórias (e_1, e_2, p) são mutuamente independentes *a priori*. Além disso, consideramos que a informação *a priori* sobre p é menos informativa do que para os erros de classificação. Tal suposição é razoável, uma vez que os itens a serem avaliados podem possuir múltiplas origens e conseqüentemente, diferentes padrões de conformidade. Já para os erros, é razoável admitir que possa existir uma informação maior, ainda que superficial, uma vez que o desempenho de classificar itens como conforme e não-conforme

independe da proporção de conformidade.

Para deduzir as distribuições *a priori* dos erros de classificação e da proporção de conformidade, utilizamos o método do histograma. Para tanto, devemos considerar uma partição do espaço paramétrico de interesse em k intervalos. Assim, pergunta-se ao especialista sobre qual a probabilidade que ele atribui ao parâmetro em estudo de pertencer a cada um dos intervalos. Com base nesta informação, é possível construir um histograma e avaliar qual distribuição Beta que melhor se ajusta. Maiores detalhes podem ser obtidos em Paulino *et al.* (2003). O'Hagan (1998) sugere usar *feedback* para obter a distribuição *a priori* final, isto é, depois de obtido o histograma e adaptada uma distribuição, apresentar o resultado obtido ao especialista para que ele averigüe se está ou não de acordo com suas crenças. Um processo iterativo permitirá obter uma distribuição final consoante com os conhecimentos subjetivos do especialista.

A densidade conjunta *a posteriori* de (p, e_1, e_2) , expressa por (6), é obtida multiplicando-se a distribuição *a priori* (5) pela verossimilhança (4) e normalizando como requerido pelo teorema de Bayes [Winkler (2003) e Gelman (1995)].

$$f(p, e_1, e_2 | \eta, m) = \sum_{l_0=0}^{n_0} \sum_{l_1=0}^{n_1} \dots \sum_{l_m=0}^{n_m} \omega_{l_0 l_1 \dots l_m} f_\beta(p | s_1, s_2) f_\beta(e_1 | s_3, s_4) f_\beta(e_2 | s_5, s_6) \quad (6)$$

em que

$$s_1 = \alpha + x_1; \quad s_2 = \beta + x_2;$$

$$s_3 = \alpha_1 + x_3; \quad s_4 = \beta_1 + x_4;$$

$$s_5 = \alpha_2 + x_5; \quad s_6 = \beta_2 + x_6;$$

$$\omega_{l_0 l_1 \dots l_m} = \frac{h_{l_0 l_1 \dots l_m}}{\sum_{l_0=0}^{n_0} \sum_{l_1=0}^{n_1} \dots \sum_{l_m=0}^{n_m} h_{l_0 l_1 \dots l_m}};$$

$$h_{l_0 l_1 \dots l_m} = \binom{n_0}{l_0} \binom{n_1}{l_1} \times \dots \times \binom{n_m}{l_m} B(s_1, s_2) B(s_3, s_4) B(s_5, s_6).$$

Os valores de $B(x, y)$ indicam uma função Beta definida em (x, y) e $f_\beta(z|x, y)$ denota uma função densidade Beta para z com parâmetros x e y . Como o objetivo primordial do artigo é fazer inferências sobre e_1 e e_2 , obtemos sua função densidade conjunta marginal através da integração de (6) em relação a p . O resultado é expresso por (7).

$$f(e_1, e_2|\eta, m) = \sum_{l_0=0}^{n_0} \sum_{l_1=0}^{n_1} \dots \sum_{l_m=0}^{n_m} \omega_{l_0 l_1 \dots l_m} f_\beta(e_1|s_3, s_4) f_\beta(e_2|s_5, s_6); \quad 0 < e_1, e_2 < 1. \quad (7)$$

As densidades marginais de e_1 e e_2 são expressas respectivamente por:

$$f(e_1|\eta, m) = \sum_{l_0=0}^{n_0} \sum_{l_1=0}^{n_1} \dots \sum_{l_m=0}^{n_m} \omega_{l_0 l_1 \dots l_m} f_\beta(e_1|s_3, s_4); \quad 0 < e_1 < 1 \quad (8)$$

$$f(e_2|\eta, m) = \sum_{l_0=0}^{n_0} \sum_{l_1=0}^{n_1} \dots \sum_{l_m=0}^{n_m} \omega_{l_0 l_1 \dots l_m} f_\beta(e_2|s_5, s_6); \quad 0 < e_2 < 1. \quad (9)$$

A partir de (8) e (9) temos as informações necessárias para a avaliação dos erros de classificação e conseqüentemente da eficiência do sistema de inspeção. Neste artigo, considera-se que um sistema de inspeção é qualificado se $P(e_1 < a_1) > \alpha_1$ e $P(e_2 < a_2) > \alpha_2$. Os valores de (a_1, α_1) e (a_2, α_2) são definidos pela empresa ou por uma norma específica de procedimentos. Em geral, em sistemas de controle de qualidade, adotam-se valores superiores a 0,8 para (α_1, α_2) e inferiores a 0,15 para (a_1, a_2) [Johnson *et al.* (1991) e Taguchi *et al.* (2004)].

A metodologia aqui discutida permite que a empresa também possa avaliar a proporção de conformidade durante a avaliação do sistema de inspeção. Para tanto, calculamos a distribuição *a posteriori* de p integrando a expressão (6) em relação a e_1 e e_2 . Tal distribuição pode ser expressa pela expressão (10) e possui toda a informação da proporção de conformidade para um processo de tomada de decisão.

$$f(p|\eta, m) = \sum_{l_0=0}^{n_0} \sum_{l_1=0}^{n_1} \dots \sum_{l_m=0}^{n_m} \omega_{l_0 l_1 \dots l_m} f_\beta(p|s_1, s_2); \quad 0 < p < 1. \quad (10)$$

4 Tamanho Amostral

No tópico anterior, o experimento a ser realizado para a determinação das distribuições *a posteriori* de e_1 , e_2 e p depende do tamanho amostral (n) e do número de classificações repetidas independentes (m) por item avaliado. Quanto maiores os valores de n e m menor será a dispersão *a posteriori* de e_1 , e_2 e p . Assim, é de fundamental importância a determinação desses valores de tal forma a garantir uma determinada precisão sem no entanto, inviabilizar operacionalmente o estudo. Para tanto, utilizamos o Critério da Cobertura Média, em inglês *Average Coverage Criterion (ACC)*, descrito em Rahme *et al.* (2000). Explicaremos os resultados para e_1 , uma vez que são similares para e_2 .

Devemos obter os menores valores possíveis de m e n de tal forma que o intervalo de credibilidade, de comprimento $2l$, $0 < l < 1/2$ e simétrico em torno da média *a posteriori* de $e_1(\hat{e}_1)$, tenha pelo menos probabilidade $1 - \alpha$ (cobertura). Assim, desejamos encontrar os valores mínimos de n e m tal que:

$$\int_{\hat{e}_1-l}^{\hat{e}_1+l} f(e_1|\eta, n, m) de_1 \geq 1 - \alpha \quad (11)$$

O problema decorrente do uso da expressão (11) para o cálculo do tamanho amostral é que ela depende do resultado experimental ainda não realizado. Uma alternativa é escolher sempre o resultado experimental (de todos possíveis para n e m fixos) que gere a maior cobertura possível. Assim, o aumento sucessivo dos valores de m e n acabaria por convergir para um par (m, n) menor possível que satisfizesse tanto a cobertura quanto a amplitude do intervalo de credibilidade. Entretanto, tal abordagem conservativa é muito rigorosa, uma vez que determinados resultados experimentais são pouco prováveis. Uma solução mais razoável seria trabalhar com uma média ponderada de todos os resultados experimentais possíveis e escolher os valores mínimos de (n, m) que atendam a expressão (11), isto é, a abordagem ACC expressa por (12).

$$\int_{\phi} \left\{ \int_{\hat{e}_1-l}^{\hat{e}_1+l} f(e_1|\eta, n, m) de_1 \right\} k(\eta) d\eta \geq 1 - \alpha \quad (12)$$

em que $\eta \in \phi$ e ϕ representa o espaço de todos os resultados experimentais possíveis que podem ser obtidos. A função $k(\cdot)$ representa a probabilidade de um particular resultado experimental e é calculada por meio da expressão (1).

Com o aumento de n e m , o número de resultados experimentais possíveis cresce, tornando a pesquisa através da expressão (12) computacionalmente intensiva, o que exige um programa eficiente e otimizado. Observa-se que, em termos operacionais, devemos calcular o tamanho amostral (n) e o número de classificações repetidas independentes (m) para cada uma das variáveis aleatórias e_1 , e_2 e p . O tamanho amostral final será então o maior entre os três cálculos. Desenvolvemos uma macro no *software* Matlab para determinar n e m que satisfaçam conjuntamente as variáveis aleatórias e_1 , e_2 e p . Observe ainda que a amplitude (2l) do intervalo de credibilidade e a respectiva probabilidade de cobertura (que denominaremos $1 - \alpha$) dependem das necessidades do pesquisador e, portanto devem ser especificadas por ele antes da realização do experimento.

Um problema decorrente do uso da expressão (12) é a possibilidade de termos múltiplas soluções. Podem existir múltiplos pares (n, m) que satisfaçam tanto a amplitude como a cobertura pré-fixada pelo pesquisador. Um determinado par pode apresentar um n grande, mas um m pequeno contradizendo com outra alternativa que poderia apresentar um n menor e um m maior. Uma alternativa para resolver tal problema seria definirmos um custo para a avaliação e classificação de um item e outro custo para repetidas classificações. Seja c_1 o custo para seleção e classificação de um item uma única vez e, c_2 o custo para cada classificação adicional. Assim, entre os pares que satisfazem à amplitude e a cobertura pré-fixadas, escolheremos aquele que apresentar o menor custo possível. O custo (C) para uma determinada alternativa (n, m) é calculado pela expressão (13).

$$C = nc_1 + (m - 1)nc_2 \quad (13)$$

5 Exemplo Numérico

Considere um experimento para avaliar um sistema de inspeção (e_1, e_2) e estimar a proporção de conformidade (p) em que o estado real dos itens inspecionados seja desconhecido. Suponha que a aplicação da metodologia do histograma para a obtenção das distribuições *a priori* dos erros e da proporção de conformidade resultou em uma distribuição $Beta(2; 10)$ para os erros de classificação e $Beta(1, 5; 1)$ para proporção de conformidade. As Figuras 1 e 2 mostram as distribuições *a priori* para uma melhor visualização. A Figura 1 indica que para a proporção de conformidade temos uma grande ausência de informação. Já a Figura 2 indica que para os erros de classificação a informação é ruim, mas permite concluir que o especialista possui uma forte convicção de que tais erros devem estar entre 0 e 0,5.

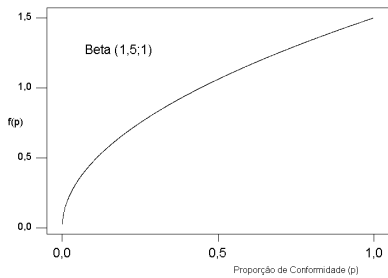


Figura 1: Distribuição *a priori* da Proporção de Conformidade

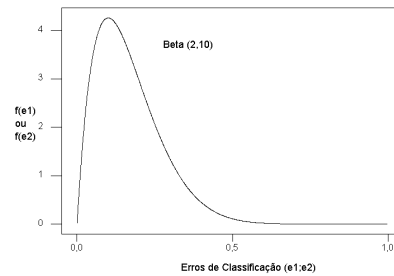


Figura 2: Distribuição *a priori* dos Erros de Classificação

Após obtermos as distribuições *a priori*, é necessário determinar os valores do tamanho amostral (n) e do número de classificações repetidas independentes (m) . Suponha que a empresa determine que deseja trabalhar com a menor amostra que garanta uma cobertura de $(1 - \alpha) \cdot 100\% = 95\%$ para um intervalo de amplitude $(2l) = 0,10$ em torno da média amostral *a posteriori* para e_1, e_2 e p . Também é de entendimento da empresa que amostras superiores a 10000 itens e classificações repetidas superiores a 13 tornam o procedimento inviável tanto em termos de tempo como em termos econômicos. Após avaliações econômicas, o sistema contábil da empresa concluiu que $c_1 = \$3,00$ e $c_2 = \$1,00$

e que o sistema de inspeção será julgado qualificado se $P(e_1 < a_1 = 0,13) > \alpha_1 = 0,95$ e $P(e_2 < a_2 = 0,11) > \alpha_2 = 0,95$. Utilizando a macro desenvolvida no *software* Matlab [Anexo A], encontramos que o tamanho amostral ótimo será $n = 528$ e $m = 7$. A Figura 3 mostra que, para $m \leq 2$ e $n \leq 10000$, não é possível garantir uma cobertura de 0,95 para um intervalo de amplitude 0,1 em torno da média *a posteriori* conjuntamente para e_1 , e_2 e p . Observa-se também a existência de múltiplas soluções.

Utilizando os custos $c_1 = \$3,00$ e $c_2 = \$1,00$, podemos concluir que a combinação ($n = 528; m = 7$) resulta em um custo de \$4752, constituindo-se no ótimo em termos econômicos, ver Tabela 2.

Tabela 2: Custo C para tamanho amostral (n) e classificações repetidas (m)

n	m	C
2854	3	14270
1138	4	6828
804	5	5628
605	6	4840
528	7	4752
481	8	4810
441	9	4851
421	10	5052
410	11	5330
400	12	5600
394	13	5910

A Figura 4 apresenta uma visualização dos pares de n e m versus custo. Observa-se uma minimização acentuada do custo para m entre seis e nove, sendo seu valor ótimo igual a sete.

Com a definição do tamanho amostral e do número m de classificações repetidas, procedeu-se um experimento simulado com erros de classificação iguais a 0,10 e proporção de conformidade igual a 0,85. A magnitude dos valores escolhidos foram baseados em

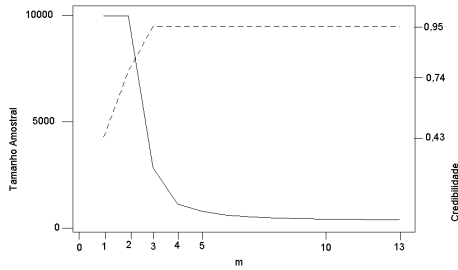


Figura 3: Tamanho amostral (n) versus cobertura versus classificações repetidas (m)

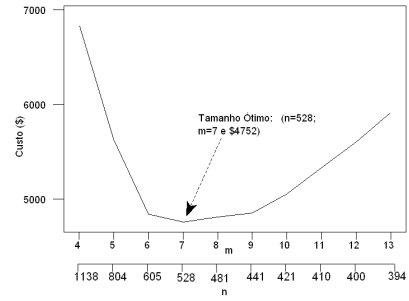


Figura 4: Custo (C) versus tamanho amostral (n) e classificações repetidas (m)

Johnson *et al.* (1991). Foi simulado a inspeção de 528 itens, sendo cada item classificado independentemente sete vezes. Os resultados do experimento são apresentados no Anexo B e um síntese na tabela 3.

Tabela 3: Síntese dos resultados das classificações repetidas de 528 itens

Classificações conformes	Número de itens
0	41
1	34
2	11
3	2
4	9
5	61
6	159
7	211

Para o cálculo das distribuições *a posteriori* de e_1 , e_2 e p , foi desenvolvida uma outra macro no *software* Matlab [Anexo C]. As Figuras 5, 6 e 7 mostram respectivamente as distribuições *a posteriori* da proporção de conformidade, do erro de classificação e_1 e erro de classificação e_2 . Uma estimativa bayesiana (média *a posteriori*) para proporção de conformidade foi de 0,848 com desvio padrão 0,028. Para os erros de classificação, obtivemos respectivamente médias *a posteriori* 0,126 e 0,103 e desvio padrão *a posteriori*

0,026 e 0,029. Para a avaliação do sistema de inspeção, obtivemos $P(e_1 < a_1 = 0,13) = 0,56 < \alpha_1 = 0,95$ e $P(e_2 < a_2 = 0,11) = 0,59 < \alpha_2 = 0,95$. Desta forma, o sistema de inspeção avaliado foi julgado inadequado tanto para erro tipo I como para o erro tipo II.

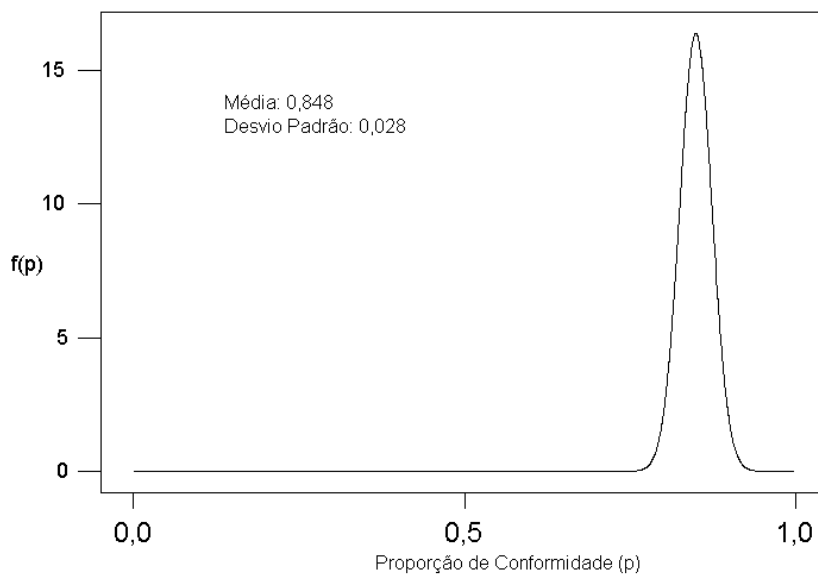


Figura 5: Distribuição *a posteriori* da Proporção de Conformidade

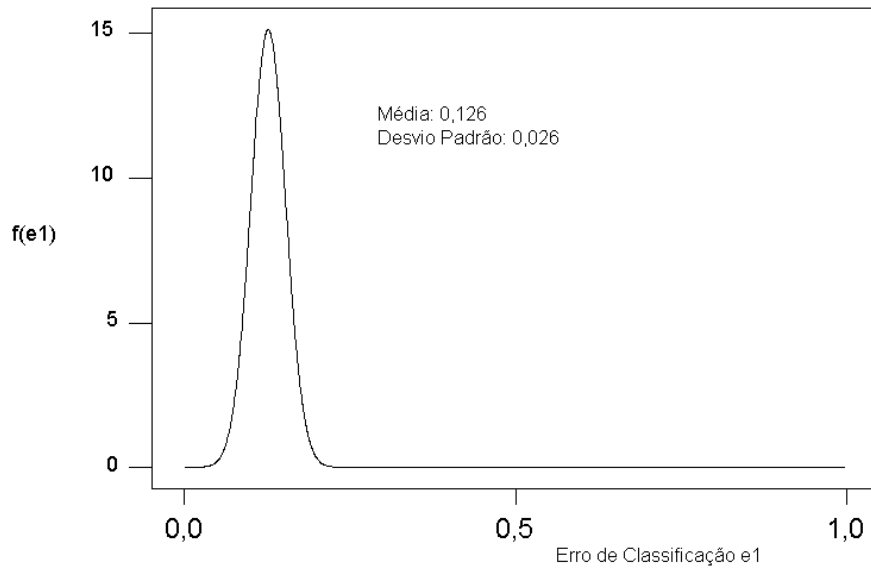


Figura 6: Distribuição *a posteriori* do Erro de Classificação e_1

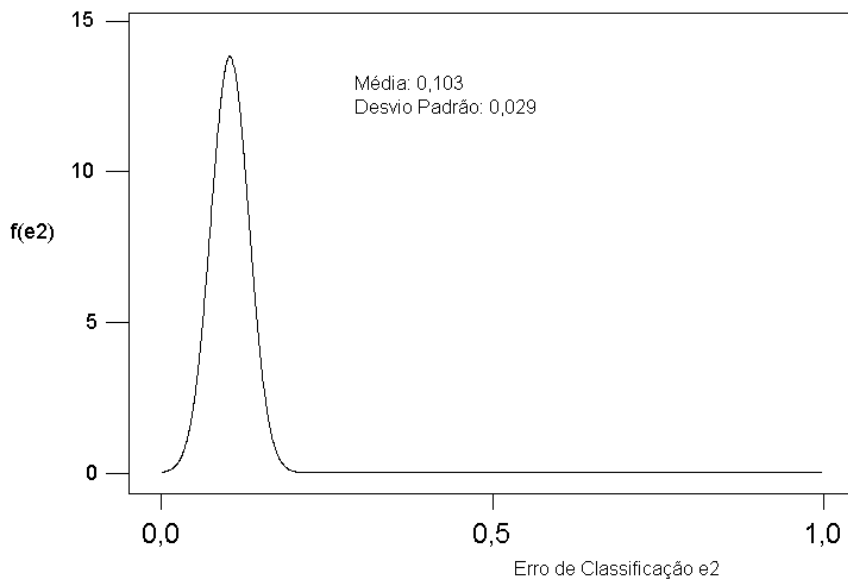


Figura 7: Distribuição *a posteriori* do Erro de Classificação e_2

6 Considerações Finais

No controle de qualidade por atributos, é crucial que o sistema de inspeção, que classifica itens manufaturados como conforme ou não-conforme, proporcione resultados confiáveis. No entanto, observa-se que sistemas de inspeção não estão livres de cometer erros. Quando o estado real dos itens a serem inspecionados é conhecido, a tarefa de se avaliar a eficiência do sistema de inspeção torna-se direta por meio da quantificação do número de classificações repetidas equivocadas. No caso em que o estado real dos itens não é conhecido, propomos a realização de classificações repetidas de cada item e a utilização de uma função de verossimilhança completa em conjunto com a análise bayesiana. Também é fundamental determinar o número ótimo de itens a serem inspecionados, assim como o total de classificações repetidas, baseados nos critérios de custo total e no ACC. Como forma de facilitar a implementação nas empresas, recomendamos as macros desenvolvidas no *software* Matlab [Anexos A e C] que permitem uma rápida aplicação da metodologia aqui discutida.

Referências Bibliográficas

Barry, & Christopher, B. (1981). Bayesian evaluation of inspectors in sensory tests. *Journal of Quality Technology*, 13(2), 120-124.

Bross, I. (1954). Misclassification in 2X2 tables. *Biometrics*, 10, 478-486.

Burke, J. R., Davis, R. D., Kaminsky, F. C., & Roberts, A. E. P. (1995). The effect of inspector errors on the true fraction non-conforming: An industrial experiment. *Quality Engineering*, 7(3), 543-550.

Evans, M., Guttman, I., Haitovsky, Y., & Swartz, T. (1996). Bayesian analysis of binary data subject to misclassification. In D. Berry, K. Chaloner, & J. Geweke (Eds.), *Bayesian analysis in statistics and econometrics: Essays in honor of Arnold Zellner* (pp. 66-77). New York: North Holland.

Gaba, A., & Winkler, R. L. (1992). Implications of errors in survey data: A Bayesian model. *Management Science*, 38(7), 913-925.

Gelman, A., Carlin, J., Stern, H., & Rubin, D. (1995). *Bayesian data analysis*. New York: Chapman & Hall.

Gramopadhye, A. K., Pajjuri, V., Kimbler, D. L., & Ferrell, W. G. (1995). Compensating for inspection errors in attribute inspection. *Quality Engineering*, 8(2), 311-322.

Greenberg, B. S., & Stokes, S. L. (1995). Repetitive testing in the presence of inspection errors. *Technometrics*, 37(1), 102-111.

Gupta, A. K., & Nadarajah, S. (2004). *Handbook of Beta Distribution and Its Applications*. New York: Marcel Dekker.

Gustafson, P. (2003). *Measurement Error and Misclassification in Statistics and Epide-*

miology: Impacts and Bayesian Adjustments. New York: Chapman & Hall.

Johnson, N. L., & Kotz, S. (1988). Estimation from binomial data with classifiers of known and unknown imperfections. *Naval Research Logistics*, 35, 147-156.

Johnson, N. L., Kotz, S., & Wu, X. (1991). *Inspection errors for attributes in quality control*. London: Chapman & Hall.

Johnson, S. L., & Funke, D. J. (1980). An analysis of human reliability measures in visual inspection. *Journal of Quality Technology*, 21(2), 71-74.

Joseph, L., Gyorkos, T. W., & Coupal, L. (1995). Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. *American Journal of Epidemiology*, 141(3), 263-272.

Kemp, A. W., & Kemp, C. D. (1988). A simple inspection scheme for two types of defect. *Journal of the Operational Research Society*, 39(3), 311-315.

O'Hagan, A. (1998). Eliciting expert beliefs in substantial practical applications. *The Statistician*, 47, 21-35.

Paulino, C. D., Turkman, M. A. A., & Murteira, B. (2003). *Estatística Bayesiana*. Fundação Calouste Gulbenkian, Lisboa.

Quinino, R.C., Ho, L.L., & Trindade, A. L.G., (2005). Bayesian judgment of a dichotomous inspection system when the true state of an evaluated item is unknown. *Computers & Industrial Engineering*, 49, 591-599.

Rahme, E., Joseph, L., & Gyorkos, T. W. (2000). Bayesian sample size determination for estimating binomial parameters from data subject to misclassification. *Applied Statistics*, 49 (1), 119-128.

Sylla, C. (2002). Experimental investigation of human and machine-vision arrangements in inspection tasks. *Control Engineering Practice*, 10, 347-361.

Taguchi, G., Chowdhury, S., & Wu, Y. (2004). *Taguchi's Quality Engineering Handbook*. New Jersey: John Wiley & Sons, Inc.

Winkler, R. L. Bayesian Inference and Decisions. (2003) 2a. ed. London: *Probabilistic Publishing*, (p 384).

Anexo A

Macro desenvolvida no *software* Matlab para a determinação do tamanho amostral n e m . Ao executar o programa, mensagens serão geradas tornando-o auto explicativo. Uma cópia do programa pode ser obtida com o orientador deste trabalho por meio do endereço eletrônico roberto@est.ufmg.br.

```
clc; clear all; tic;

global PRIORIp PRIORIE1 PRIORIE2 NMAX L COBERTURA nM ICONT

PRIORIp=[1 1 1];
PRIORIE1=[1 2 10];
PRIORIE2=[1 2 10];

NMAX=10000;
M=10:15;
L=0.10;
COBERTURA=0.95;
namostras=2000;
ntriplas=2000;

nM = length(M);
TAMANHO=zeros(nM,3);

nmin=inf;
nmax=-inf;

ICONT=0;
status(1,ICONT,nM,toc,'Obtendo o tamanho da amostra',0,[]);

for im=1:nM

    filename=['dados_m' num2str(im) '.txt'];

    m=M(im);

    [vn,vacc,n,acc,stdev]=bissecacao(m,namostras,ntriplas);

    TAMANHO(im,1)=m;
    TAMANHO(im,2)=n;
    TAMANHO(im,3)=acc;

    if im==1
        texto=['m=' num2str(m,6) ' '; N=' num2str(n,6) '];
        ACC = ' num2str(acc,4) ' ; std = ' num2str(stdev,4)];
    else
```

```

        texto=strvcat(texto, ['m=' num2str(m,6) ' ; N=' num2str(n,6) ' ;
            ACC = ' num2str(acc,4) ' ; std = ' num2str(stdev,4)']);
    end

    ICONT=ICONT+1;
    status(1,ICONT,nM,toc,'Obtendo o tamanho da amostra...',0,[]);

    if min(vn)<nmin
        nmin=min(vn);
    end
    if max(vn)>nmax
        nmax=max(vn);
    end

    SAIDA=[vn vacc];
    save(filename,'-ascii','-tabs','SAIDA');
    save espacio

end

grafico;

function y=graficoi(x,y,i)
estilo={'-';'-.';':';'--'};
simbolo={'.';'*';'s';'o';'v';'d';'x';''};
cor={'r';'b';'g';'y';'k'};
limite=7*4*5;
temp=mod(i,limite);
if temp==0
    temp=limite;
end
cont=0;
for i1=1:4
    for i2=1:7
        for i3=1:5
            cont=cont+1;
            if cont==temp
                formato=[(char(cor(i3))),(char(simbolo(i2))),(char(estilo(i1)))];
                hold on;
                plot(x,y,formato);
                box on;
                set(gcf,'Color',[1,1,1]);
                drawnow;
                return
            end
        end
    end
end
end
end
end

```

```

function [prob, stdev, probzeros]=avgcov(n,m,namostras,ntriplas)

global COBERTURA L
global PRIORIp PRIORIE1 PRIORIE2
global LNK0

if PRIORIp(1)==0
    NEWP=unifrnd(PRIORIp(2),PRIORIp(3),namostras,1);
else
    NEWP=betarnd(PRIORIp(2),PRIORIp(3),namostras,1);
end

if PRIORIE1(1)==0
    NEWE1=unifrnd(PRIORIE1(2),PRIORIE1(3),namostras,1);
else
    NEWE1=betarnd(PRIORIE1(2),PRIORIE1(3),namostras,1);
end

if PRIORIE2(1)==0
    NEWE2=unifrnd(PRIORIE2(2),PRIORIE2(3),namostras,1);
else
    NEWE2=betarnd(PRIORIE2(2),PRIORIE2(3),namostras,1);
end

coveragex=zeros(1,namostras);
coveragevalido=ones(1,namostras);
mx=zeros(1,namostras);

nzeros=0;

minp=0;
maxp=0;
mine1=0;
maxe1=0;
mine2=0;
maxe2=0;

bins=(0:m)-0.5;

for ia=1:namostras

    newp=NEWP(ia);
    newe1=NEWE1(ia);
    newe2=NEWE2(ia);

    amostra=geraamostra(n,m,newp,newe1,newe2);
    subamostra=transpose(amostra);

```



```

if m>1
    vetorc=sum(subamostra);
else
    vetorc=subamostra;
end

newvetorc=contar(vetorc,m);

if PRIORIp(1)==0
    P=unifrnd(PRIORIp(2),PRIORIp(3),ntriplas/10,1);
else
    P=betarnd(PRIORIp(2),PRIORIp(3),ntriplas,1);
end

if PRIORIE1(1)==0
    E1=unifrnd(PRIORIE1(2),PRIORIE1(3),ntriplas/10,1);
else
    E1=betarnd(PRIORIE1(2),PRIORIE1(3),ntriplas,1);
end

if PRIORIE2(1)==0
    E2=unifrnd(PRIORIE2(2),PRIORIE2(3),ntriplas/10,1);
else
    E2=betarnd(PRIORIE2(2),PRIORIE2(3),ntriplas,1);
end

LNK0=0;
logweight=funcaol(newvetorc,P,E1,E2,m);
LNK0=max(logweight);

logweight=logweight-LNK0;

indice=find((logweight-max(logweight))>-5);

minp=min(P(indice));
maxp=max(P(indice));
mine1=min(E1(indice));
maxe1=max(E1(indice));
mine2=min(E2(indice));
maxe2=max(E2(indice));

if PRIORIp(1)==0
    P=unifrnd(minp,maxp,ntriplas,1);
end

if PRIORIE1(1)==0

```

```

        E1=unifrnd(mine1,maxe1,ntriplas,1);
    end

    if PRIORIE2(1)==0
        E2=unifrnd(mine2,maxe2,ntriplas,1);
    end

    logweight=funcaol(newvetorc,P,E1,E2,m);
    weight=exp(logweight);

    nzeros=nzeros+sum(weight==0);
    sw=sum(weight);

    if sw>0
        postpi=weight/sw;
    else
        postpi=zeros(1,ntriplas);
        coveragevalido(ia)=0;
    end

    phat=postpi*P;
    indice=find((P>=(phat-L/2))&(P<=(phat+L/2)));
    coveragex(ia)=sum(postpi(indice));
end

minp=minp/namostras;
maxp=maxp/namostras;
mine1=mine1/namostras;
maxe1=maxe1/namostras;
mine2=mine2/namostras;
maxe2=maxe2/namostras;

indice=find(coveragevalido);
prob=mean(coveragex(indice))-COBERTURA;
stdev=std(coveragex(indice));
probzeros=nzeros/(namostras*ntriplas);

function y=contar(amostra,m)

y=zeros(1,m+1);
for i=0:m
    y(i+1)=sum(amostra==i);
end

function y = funcaol(vetorc,p,e1,e2,m)

global LNK0
n=length(p);

```

```

PR1= repmat(1-e1, [1 m+1]);
PR2= repmat(e2, [1 m+1]);
P= repmat(p, [1 m+1]);

x=0:1:m;
X= repmat(x, [n 1]);
M= repmat(m, [n m+1]);

F= repmat(vetorc, [n 1]);

BINO1=mybinopdf(X,M,PR1);
BINO2=mybinopdf(X,M,PR2);
TERMO1=P.*BINO1+(1-P).*BINO2;
TERMO1=log(TERMO1);
TERMO2=TERMO1.*F;
TERMO3=sum(TERMO2,2);
y=transpose(TERMO3)-LNK0;

function [vn,vacc,n,acc,stdev]=bissecao(m,namostras,ntriplas)

global NMAX L COBERTURA ICONT nM

na=2;
[fa,stdeva]=avgcov(na,m,namostras,ntriplas);

nb=NMAX;
[fb,stdevb]=avgcov(nb,m,namostras,ntriplas);

nm=ceil((na+nb)/2);
[fm,stdevm]=avgcov(nm,m,namostras,ntriplas);

status(1,ICONT,nM,toc,'Obtendo o tamanho da amostra...',0,[na nm nb]);

vn=[na;nb;nm];
vacc=[fa;fb;fm];

if fa*fm<=0
    fb=fm;
    nb=nm;
    stdevb=stdevm;
else
    fa=fm;
    na=nm;
    stdeva=stdevm;
end

nm=ceil((na+nb)/2);

```

```

[fm,stdevm]=avgcov(nm,m,namostras,ntriplas);

vn=[vn;nm];
vacc=[vacc;fm];

status(1,ICONT,nM,toc,'Obtendo o tamanho da amostra...',0,[na nm nb]);

while (abs(nb-na)-1e-6)>1

    if fa*fm<=0
        fb=fm;
        nb=nm;
        stdevb=stdevm;
    else
        fa=fm;
        na=nm;
        stdeva=stdevm;
    end

    nm=ceil((na+nb)/2);
    fm=avgcov(nm,m,namostras,ntriplas);

    vn=[vn;nm];
    vacc=[vacc;fm];

    status(1,ICONT,nM,toc,'Obtendo o tamanho da amostra...',0,[na nm nb]);
end

[vn,indice]=sort(vn);
vacc=vacc(indice)+COBERTURA;

if fa>0
    acc=fa+COBERTURA;
    n=na;
    stdev=stdeva;
else
    acc=fb+COBERTURA;
    n=nb;
    stdev=stdevb;
end
function y=geraamostra(n,m,p,e1,e2)

real=binornd(1,p,n,1);
conforme=sum(real);
naoconforme=n-conforme;

amostra=[];
if conforme>0
    amostra=binornd(1,1-e1,conforme,m);

```

```

end

if naoconforme>0
    amostra=[amostra;binornd(1,e2,naoconforme,m)];
end

y=amostra;

function y = mybinopdf(x,n,p)

if nargin < 3,
    error('stats:binopdf:TooFewInputs','Requires three input arguments');
end

y = zeros(size(x));

if ~isfloat(x)
    x = double(x);
end
if ~isfloat(n)
    n = double(n);
end
k = find(x >= 0 & x == round(x) & x <= n);
if any(k)
    t = (p(k)==0);
    if any(t)
        kt = k(t);
        y(kt) = (x(kt)==0);
        k(t) = [];
    end
    t = (p(k)==1);
    if any(t)
        kt = k(t);
        y(kt) = (x(kt)==n(kt));
        k(t) = [];
    end
end
end

if any(k)
    nk = gammaln(n(k) + 1) - gammaln(x(k) + 1) - gammaln(n(k) - x(k) + 1);
    lny = nk + x(k).*log( p(k)) + (n(k) - x(k)).*log(1 - p(k));
    y(k) = exp(lny);
end

k1 = find(n < 0 | p < 0 | p > 1 | round(n) ~= n);
if any(k1)
    y(k1) = NaN;
end
figure;

```

```

fonte=8;
MINIMO=1;
MAXIMO=0;
for im=1:nM
    filename=['dados_m' num2str(im) '.txt'];
    dados=load(filename);
    graficoi(dados(:,1),dados(:,2),im);
    if max(dados(:,2))>MAXIMO
        MAXIMO=max(dados(:,2));
    end
    if min(dados(:,2))<MINIMO
        MINIMO=min(dados(:,2));
    end
end

MINIMO=floor(MINIMO*10)/10;
MAXIMO=ceil(MAXIMO*10)/10;

titulo=['num. de amostras = ' num2str(namostras,6) ' ';
        num triplas = ' num2str(ntriplas,6), ' ; l=' num2str(L,4)];

hold on;
plot([nmin nmax],[COBERTURA COBERTURA],'k:');
ylim([MINIMO MAXIMO]);
xlim([2 NMAX]);
novotexto=strvcat(texto, ['1-\alpha=' num2str(COBERTURA*100,3) '%']);
drawnow;
h=legend(novotexto);
set(h,'Location','SouthEast','FontSize',fonte-1);
title(titulo);
xlabel('N');
ylabel('ACC p');

set(gca,'XScale','log');
drawnow;
h=legend(novotexto);
set(h,'Location','SouthEast','FontSize',fonte-1);
print('-dpng','-r120','saida_log');
print('-dpdf','-r120','saida_log.pdf');
close(gcf);

function y=status(inicio,i,casos,tempo,texto,correcao,vec)

mmax=casos;
mmin=inicio;
m=i;

progressbar=['|'];

```

```

i=i-inicio+1;
casos=casos-inicio+1;
inicio=1;

passo=floor(casos*0.03);
if passo==0;
    passo=1;
end

if i==1 | i==inicio | i==casos | mod(i,passo)==0
else
    return
end

if nargin==5
    correcao=0;
    vec=[];
end

if nargin==6
    vec=[];
end

if correcao==0
    tempoestimado=tempo/60*(casos-i)/(i+1e-6);
else
    total=0.5*mmax*(mmax+1)-0.5*mmin*(mmin-1);
    atual=0.5*m*(m+1)-0.5*mmin*(mmin-1);
    tempoestimado=tempo/60*(total-atual)/(atual);
end

clc;
for kkk=1:floor(i/casos*100/5)
    progressbar=[progressbar,'*'];
end
for kkk=(floor(i/casos*100/5)+1):20
    progressbar=[progressbar,' '];
end
progressbar=[progressbar,'|'];

fprintf('%60s\n', '*****');
fprintf('%60s\n', 'Progress');
fprintf('%60s\n', '*****');
fprintf('%60s\n', texto);
fprintf('%2s\n', ' ');
fprintf('%24s\t %5.0f\t %2s\t %5.0f\n', 'current iteration = ', i, 'of', casos);
fprintf('%24s\t %5.0f %2s\n', 'Progress = ', i/casos*100, ' %');
fprintf('%24s %22s %3s\n', '0%', progressbar, '100%');
fprintf('%2s\n', ' ');

```

```

fprintf('%40s\t %4.6f\n', 'elapsed time (min)           = ', tempo/60);
fprintf('%2s\n', ' ');
fprintf('%40s\t %4.6f\n', 'time remaining (min)         = ', tempoestimado);
fprintf('%2s\n', ' ');
if length(vec)>0
    fprintf('%24s\t %5.0f\n', 'Li N = ', vec(1));
    fprintf('%24s\t %5.0f\n', 'Lm N = ', vec(2));
    fprintf('%24s\t %5.0f\n', 'Ls N = ', vec(3));
end
fprintf('%60s\n', ' *****');
drawnow;

```


Anexo B

Resultados do experimento de classificações repetidas de 528 itens, sete vezes cada.

Item	Resultado da Classificação Repetida (C_{ij})						
	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1
3	1	1	0	0	1	1	1
4	1	0	1	1	1	1	1
5	0	0	0	0	0	0	0
6	1	1	1	1	0	1	0
7	0	0	0	0	0	0	0
8	1	1	1	1	1	1	1
9	1	0	1	1	1	1	1
10	1	1	1	1	1	1	0
11	1	1	1	1	1	1	0
12	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1
14	0	0	0	1	0	0	0
15	0	1	1	1	0	1	0
16	1	1	0	1	1	1	1
17	1	0	0	0	0	0	0
18	0	1	0	1	0	0	0
19	1	1	1	1	1	1	0
20	1	0	0	0	0	0	0
21	1	1	1	1	1	1	1
22	0	1	0	0	0	0	0
23	1	1	1	1	1	1	1
24	1	1	1	1	1	1	1
25	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1
27	1	0	1	1	1	0	1
28	1	1	1	1	1	1	1
29	1	1	1	1	1	1	1
30	1	1	1	1	1	1	1
31	1	1	0	1	0	1	1
32	0	0	0	0	0	0	1
33	1	1	1	0	1	1	1
34	1	0	0	1	1	1	1
35	1	1	1	1	1	1	0
36	1	1	1	1	1	1	1
37	0	0	0	0	0	0	0
38	1	1	1	1	1	1	1
39	1	0	1	1	1	1	1
40	1	1	1	1	1	1	1
41	1	0	1	1	1	1	1
42	0	1	1	1	1	1	1
43	1	1	1	1	1	1	0
44	1	1	1	1	1	1	1
45	1	1	1	0	1	1	1
46	1	1	1	1	1	1	0
47	0	0	0	0	0	0	0
48	1	1	1	1	0	1	1
49	0	1	1	1	1	1	1

50	1	1	1	1	1	1	1
51	0	0	0	0	0	0	0
52	1	1	1	0	1	1	1
53	0	1	1	1	1	1	0
54	0	0	1	0	0	0	0
55	0	0	0	0	0	0	0
56	1	1	1	1	1	1	1
57	1	1	1	1	1	1	1
58	1	1	1	1	1	1	1
59	1	1	1	1	1	1	1
60	1	1	1	1	1	1	1
61	1	1	1	1	1	1	1
62	1	1	1	1	1	1	1
63	1	1	1	1	1	1	1
64	1	1	1	1	1	1	0
65	1	1	0	1	1	1	1
66	1	0	1	0	0	0	0
67	1	0	1	1	1	1	1
68	1	1	1	1	1	1	1
69	1	1	1	1	1	1	1
70	1	1	1	1	1	1	1
71	0	0	1	0	0	0	0
72	0	1	1	1	1	1	1
73	1	1	1	0	0	1	1
74	1	1	1	1	1	0	0
75	1	1	1	1	0	1	0
76	1	1	1	1	1	1	1
77	1	1	1	1	0	1	1
78	1	1	1	1	1	0	1
79	1	1	1	1	1	1	1
80	1	1	1	1	1	1	1
81	1	0	1	1	1	1	1
82	1	1	1	1	1	1	1
83	1	1	1	1	1	1	1
84	0	0	0	0	0	0	0
85	1	1	1	1	1	1	1
86	0	0	0	0	0	0	0
87	0	0	0	0	0	0	1
88	1	1	1	0	1	1	1
89	1	1	1	1	1	1	0
90	0	0	0	0	0	0	0
91	1	1	1	1	1	0	1
92	0	1	0	0	1	0	0
93	0	1	1	1	1	1	0
94	0	0	0	0	0	0	0
95	1	1	1	1	1	1	1
96	1	1	1	1	1	1	1
97	1	1	1	0	1	1	1
98	1	1	1	1	1	1	1
99	1	1	0	1	1	1	0
100	1	1	1	1	1	1	1
101	1	1	1	1	1	1	1
102	1	1	1	1	1	1	1
103	1	1	1	1	1	1	1

104	1	1	1	1	1	1	1
105	1	1	1	1	1	1	1
106	1	1	1	1	0	1	1
107	1	1	1	1	1	1	1
108	0	0	1	0	0	0	0
109	1	1	1	1	1	1	1
110	1	1	1	1	0	0	0
111	1	1	1	0	1	1	1
112	1	1	1	1	1	1	1
113	1	1	1	1	1	1	1
114	1	1	0	1	1	0	1
115	1	1	1	1	1	1	1
116	0	0	0	0	0	0	1
117	1	1	1	1	1	1	1
118	1	1	1	1	1	0	1
119	0	1	1	0	1	1	1
120	1	1	1	0	0	0	1
121	1	1	1	1	1	1	1
122	0	0	0	0	0	0	0
123	1	1	1	1	1	1	1
124	1	1	1	1	1	1	1
125	1	1	1	1	1	0	1
126	1	1	1	1	1	0	1
127	1	1	1	1	1	1	1
128	1	1	1	1	0	1	1
129	1	1	1	1	1	1	1
130	1	1	1	1	1	0	1
131	1	1	1	1	1	1	1
132	0	0	0	0	0	1	0
133	1	0	1	1	0	1	1
134	0	1	1	1	1	1	0
135	1	1	1	1	1	1	1
136	1	1	0	0	1	0	1
137	0	0	0	0	0	0	0
138	1	1	1	1	1	1	0
139	1	1	1	1	1	1	1
140	0	0	1	1	1	1	1
141	0	0	0	0	0	0	1
142	1	1	1	1	1	1	1
143	1	1	1	1	1	1	1
144	1	1	1	0	0	1	1
145	0	0	0	0	0	0	0
146	0	0	0	0	1	0	0
147	0	0	0	0	0	0	0
148	1	1	0	1	1	1	1
149	1	1	1	0	1	1	1
150	1	1	1	1	1	0	1
151	1	1	0	1	1	1	1
152	1	0	1	1	1	1	1
153	0	0	0	0	0	0	0
154	1	1	0	1	1	1	1
155	1	0	1	1	1	1	1
156	1	1	1	1	1	1	1
157	1	1	1	1	1	1	1

158	1	1	1	1	1	1	1
159	1	1	1	1	1	0	1
160	0	0	0	0	0	0	1
161	1	1	1	1	1	1	0
162	1	1	1	1	1	1	1
163	1	0	1	1	1	1	1
164	0	0	1	1	0	0	0
165	1	1	0	1	1	1	1
166	0	1	1	1	1	1	0
167	1	1	1	1	1	1	1
168	1	1	1	1	1	1	1
169	1	1	1	1	1	1	1
170	0	1	1	1	1	1	1
171	1	1	0	1	1	1	1
172	1	1	1	1	1	1	1
173	0	0	0	0	0	0	1
174	0	0	0	0	1	0	0
175	1	1	1	1	1	1	1
176	0	1	0	0	0	0	0
177	1	1	0	1	1	1	1
178	1	1	1	1	1	1	1
179	1	1	1	1	1	0	1
180	1	1	1	1	1	1	1
181	1	1	1	1	1	1	1
182	1	1	1	1	1	1	1
183	1	1	0	0	0	0	0
184	1	1	1	1	1	0	1
185	1	1	1	1	1	1	1
186	1	0	0	0	0	0	0
187	1	1	1	0	1	1	1
188	0	1	0	0	0	0	0
189	1	1	1	1	1	0	1
190	1	1	1	1	0	1	1
191	1	0	1	1	1	1	1
192	1	1	1	1	1	0	1
193	1	1	1	1	1	1	1
194	1	1	0	1	0	1	1
195	1	1	1	1	1	1	1
196	1	1	1	1	1	1	1
197	1	0	1	1	1	1	1
198	1	1	1	1	1	1	1
199	1	1	1	1	1	1	1
200	1	1	1	1	1	1	1
201	1	1	1	1	1	1	1
202	1	1	1	1	1	1	1
203	1	1	1	0	1	1	1
204	1	1	1	1	1	1	1
205	1	1	1	1	1	1	1
206	1	1	1	1	0	1	1
207	0	1	0	0	0	0	1
208	1	1	1	1	1	1	1
209	1	1	1	1	1	1	1
210	0	1	1	1	1	1	1
211	1	1	1	1	1	0	0

212	0	1	1	1	1	1	1
213	0	1	1	0	1	1	1
214	0	0	0	0	1	1	0
215	1	1	1	1	1	0	1
216	1	1	1	1	1	1	1
217	1	1	1	1	1	1	1
218	1	1	1	1	1	1	1
219	1	1	1	1	1	1	1
220	1	1	1	1	1	1	0
221	0	1	0	0	0	0	0
222	1	1	1	1	1	1	1
223	0	1	1	1	1	0	1
224	0	1	0	0	0	0	0
225	1	1	1	1	1	1	1
226	1	1	1	1	1	1	1
227	1	1	1	1	1	1	1
228	1	1	1	1	1	1	1
229	1	1	1	1	1	1	1
230	1	1	0	0	1	1	1
231	1	1	1	0	1	1	1
232	1	1	1	1	1	1	1
233	1	1	1	1	1	1	0
234	0	0	0	0	0	0	1
235	1	1	1	1	1	1	1
236	0	0	0	0	0	0	0
237	1	1	1	1	1	1	1
238	1	1	1	0	1	1	1
239	1	1	1	1	1	1	1
240	1	1	0	1	1	1	1
241	1	1	1	1	1	1	1
242	1	1	1	1	1	1	1
243	1	1	1	1	1	1	1
244	0	0	0	0	0	0	0
245	1	1	0	0	1	1	1
246	1	1	1	1	1	1	1
247	1	1	1	1	1	1	1
248	1	1	1	1	0	1	0
249	1	1	1	1	1	1	1
250	1	1	0	1	1	1	1
251	1	1	1	1	1	1	1
252	1	1	1	1	0	1	1
253	1	1	0	1	1	1	1
254	0	0	1	0	1	0	0
255	1	1	1	1	1	1	1
256	1	1	1	1	0	1	0
257	0	0	1	1	0	1	1
258	1	1	1	1	1	1	1
259	1	0	1	1	1	1	1
260	0	0	0	0	0	0	0
261	1	1	1	1	1	1	1
262	1	1	1	1	1	0	1
263	1	1	1	1	1	1	1
264	1	1	1	1	1	1	0
265	0	1	1	1	1	1	1

266	1	1	0	1	1	1	1
267	1	1	1	1	1	1	1
268	1	1	1	1	1	1	1
269	1	1	1	1	1	1	1
270	1	0	1	1	1	1	1
271	1	1	1	1	1	1	1
272	0	0	1	0	1	1	1
273	1	0	1	1	1	1	1
274	1	1	1	1	1	1	1
275	1	1	1	1	1	1	1
276	1	1	1	1	1	1	1
277	1	1	1	1	1	1	1
278	1	0	1	1	1	1	1
279	0	0	0	0	0	0	0
280	0	0	1	0	0	0	1
281	0	1	1	1	1	1	1
282	1	1	1	1	1	1	1
283	1	1	1	0	1	1	1
284	1	1	1	1	1	1	1
285	1	1	1	1	1	1	1
286	0	0	0	0	0	0	0
287	1	1	1	0	1	1	1
288	1	1	1	1	1	0	1
289	1	0	1	1	1	0	1
290	1	1	1	1	1	1	1
291	0	0	0	0	0	0	1
292	1	1	1	1	1	1	1
293	1	1	1	1	1	0	1
294	1	1	1	1	1	1	1
295	1	1	1	1	1	1	1
296	1	1	0	1	1	1	1
297	0	0	0	0	0	0	0
298	0	0	0	0	0	0	0
299	1	0	1	0	0	1	1
300	1	1	1	1	1	1	1
301	1	0	1	1	1	1	1
302	1	1	1	1	1	1	1
303	1	1	1	1	1	1	1
304	1	1	0	1	0	1	1
305	1	1	1	1	1	1	1
306	1	1	0	1	1	1	1
307	1	1	1	1	0	0	1
308	1	1	0	1	1	0	1
309	1	1	1	1	1	1	1
310	1	1	1	1	1	1	1
311	1	0	1	1	1	1	1
312	1	1	1	0	1	1	0
313	1	1	1	1	1	0	1
314	0	0	0	0	0	0	0
315	1	1	0	1	1	1	1
316	1	1	1	1	1	1	1
317	1	1	1	1	1	1	1
318	0	1	1	0	0	1	0
319	1	1	1	1	1	1	1

320	0	1	1	1	1	1	1
321	1	1	1	1	1	1	1
322	1	1	1	1	1	1	1
323	1	1	1	1	1	1	0
324	1	1	1	1	1	1	1
325	1	1	1	0	1	1	1
326	1	0	1	1	1	1	1
327	1	0	0	1	1	1	1
328	1	0	1	1	1	1	1
329	1	0	1	1	1	1	1
330	1	1	1	0	1	1	1
331	1	1	1	1	1	1	1
332	1	1	1	1	1	1	1
333	1	1	1	1	1	1	0
334	1	1	1	1	1	1	1
335	1	0	1	1	1	1	1
336	1	1	1	1	1	0	1
337	1	1	1	1	1	1	1
338	1	0	0	0	0	0	0
339	0	1	1	1	1	1	0
340	1	1	1	1	1	1	1
341	0	0	1	0	0	0	0
342	0	1	1	1	1	0	1
343	1	1	1	1	1	1	1
344	1	1	1	1	0	1	1
345	1	1	1	1	1	0	1
346	1	1	1	1	1	1	1
347	0	1	1	1	1	0	1
348	1	1	1	1	1	1	1
349	1	1	1	1	1	1	1
350	1	1	1	1	1	1	1
351	1	1	1	1	1	1	1
352	1	1	1	1	1	1	1
353	1	1	1	1	1	1	0
354	0	0	0	0	0	0	0
355	0	1	1	1	1	1	1
356	0	0	0	0	0	0	0
357	1	0	1	1	1	1	1
358	1	1	1	1	1	1	1
359	0	1	0	0	0	0	0
360	1	1	1	1	1	0	1
361	1	1	1	1	1	1	1
362	0	0	0	0	0	0	0
363	1	1	1	0	1	1	1
364	0	0	1	0	0	0	0
365	1	1	0	1	0	1	1
366	1	1	1	1	1	1	1
367	1	1	1	1	1	1	0
368	1	1	1	1	0	1	1
369	1	1	1	1	1	1	1
370	1	0	1	1	1	1	1
371	1	1	0	0	1	1	1
372	1	1	0	1	1	1	1
373	1	1	1	1	0	1	1

374	1	0	1	1	0	1	1
375	0	1	1	1	1	1	1
376	1	0	1	1	1	1	1
377	1	1	1	1	1	1	1
378	1	1	1	1	0	1	1
379	1	1	0	1	0	1	1
380	0	1	1	1	1	1	1
381	0	0	1	0	0	0	0
382	0	0	0	0	0	0	0
383	0	0	0	0	0	0	0
384	1	0	0	0	0	0	0
385	1	1	1	0	1	1	1
386	1	0	1	1	1	1	1
387	1	0	1	1	0	0	0
388	1	1	1	1	1	1	1
389	1	1	1	1	1	1	1
390	0	0	0	0	0	0	0
391	1	1	1	1	1	1	1
392	1	1	1	1	1	1	1
393	0	1	1	0	1	1	0
394	1	1	1	1	1	1	1
395	0	1	1	1	0	1	1
396	1	1	1	1	1	1	1
397	1	1	1	0	0	1	0
398	1	1	0	1	1	1	1
399	0	0	0	0	0	0	0
400	1	1	0	1	1	1	1
401	0	0	0	0	0	0	0
402	0	0	1	0	0	0	0
403	1	1	1	1	1	1	1
404	1	1	1	1	1	1	1
405	1	1	1	1	1	1	1
406	1	1	1	1	0	1	1
407	0	0	0	0	0	0	0
408	0	0	0	0	0	0	1
409	1	1	1	0	1	1	1
410	0	1	1	0	1	1	1
411	1	1	1	1	1	1	1
412	1	1	1	1	1	1	1
413	1	1	0	1	1	1	1
414	1	1	1	1	0	1	1
415	1	1	1	0	1	1	1
416	1	1	1	0	1	1	0
417	1	1	1	1	1	1	1
418	0	0	0	0	0	0	0
419	1	1	1	0	1	0	1
420	1	1	1	1	1	1	1
421	1	1	1	1	1	1	1
422	1	1	1	0	1	1	1
423	1	1	1	1	1	1	1
424	0	0	1	1	1	1	1
425	1	1	1	1	1	1	1
426	1	1	1	1	1	1	1
427	1	1	1	0	1	0	1

428	0	0	0	0	0	0	0
429	0	0	0	0	0	1	0
430	0	0	0	0	0	0	0
431	1	1	0	1	0	1	1
432	1	1	1	0	1	1	1
433	1	1	1	1	1	1	1
434	1	1	1	1	1	1	1
435	1	1	1	1	1	1	1
436	1	1	1	1	1	1	0
437	1	1	1	1	1	1	1
438	1	1	1	1	0	0	1
439	1	1	1	0	1	1	0
440	1	1	1	1	1	1	1
441	1	1	1	1	1	1	1
442	1	1	0	1	1	1	1
443	1	1	1	1	0	1	1
444	1	1	1	1	1	1	1
445	0	0	0	0	0	0	0
446	1	0	1	1	1	1	1
447	0	0	1	1	1	1	1
448	0	0	0	0	0	0	1
449	1	1	1	1	1	1	1
450	1	1	1	1	1	0	1
451	1	1	1	1	1	1	1
452	1	1	1	1	1	0	1
453	1	1	1	0	1	1	0
454	0	0	1	1	1	1	1
455	1	1	1	1	1	1	1
456	1	1	1	1	1	1	1
457	1	1	1	1	1	1	1
458	0	1	1	1	1	1	1
459	1	1	1	1	1	1	1
460	1	1	1	1	1	1	1
461	1	1	1	0	1	1	1
462	1	1	0	1	1	0	1
463	0	0	0	0	0	0	0
464	1	1	0	1	1	1	0
465	1	1	1	1	0	1	1
466	0	1	1	1	1	1	1
467	1	1	1	1	1	1	0
468	1	1	1	1	1	1	1
469	1	0	1	1	1	1	1
470	1	1	1	0	1	1	1
471	1	1	1	1	1	1	1
472	0	1	1	1	1	1	1
473	1	1	1	1	0	1	0
474	0	1	1	1	1	1	1
475	0	0	0	0	0	0	0
476	1	1	1	1	1	0	1
477	0	1	0	1	1	1	1
478	1	1	1	1	1	1	1
479	1	1	1	0	1	1	0
480	0	0	0	0	0	0	0
481	0	1	0	1	1	1	1

482	1	1	0	1	1	1	1
483	1	1	1	1	1	1	1
484	1	1	1	1	1	1	1
485	0	0	0	0	0	0	0
486	1	1	1	1	1	1	0
487	1	1	1	1	1	1	1
488	1	1	1	1	1	1	1
489	0	0	1	1	0	0	0
490	1	1	1	0	1	1	1
491	1	1	1	1	1	1	0
492	0	1	0	0	0	0	0
493	1	1	1	1	0	1	1
494	1	1	1	1	1	1	1
495	1	1	1	0	1	1	1
496	1	1	1	1	1	1	1
497	1	1	1	1	1	1	1
498	1	1	1	1	0	1	1
499	1	1	1	1	1	1	1
500	1	1	1	1	1	1	0
501	1	1	1	1	1	1	1
502	1	1	1	1	1	1	1
503	1	1	1	1	1	1	0
504	1	1	1	1	0	1	1
505	1	1	0	1	1	0	1
506	0	0	0	0	0	0	0
507	1	1	1	1	1	1	1
508	1	1	1	1	1	1	1
509	1	1	1	1	1	1	1
510	1	1	1	1	1	1	0
511	0	0	1	0	1	0	0
512	1	0	1	1	1	1	0
513	1	1	1	1	1	1	1
514	1	1	1	1	1	0	1
515	1	1	1	1	1	1	1
516	1	1	1	1	1	1	1
517	1	1	1	0	1	1	1
518	1	1	1	1	1	1	1
519	1	1	1	1	1	1	1
520	0	0	1	1	1	1	1
521	1	1	1	1	1	1	0
522	1	1	1	1	1	1	1
523	1	1	1	1	1	1	1
524	1	1	1	1	1	0	0
525	1	1	1	1	1	1	1
526	1	1	1	0	1	1	1
527	1	1	1	1	1	0	1
528	1	1	1	1	1	1	1

Anexo C

Macro desenvolvida no *software* Matlab para a determinação das distribuições *à posteriori* dos erros de classificação e da proporção de interesse. Ao executar o programa, mensagens serão geradas tornando-o auto explicativo. Uma cópia do programa pode ser obtida com o orientador deste trabalho por meio do endereço eletrônico `roberto@est.ufmg.br`.

```
clc;
clear all;
tic;
global pteorico e1teorico e2teorico L
warning off MATLAB:log:logOfZero

pteorico=0.85;
e1teorico=0.1;
e2teorico=0.1;
semente=1;

L=0.10;

pmin=0;
pmax=1;

e1min=0;
e1max=1;

e2min=0;
e2max=1;
N=[500];
M=[7];

nptos=501;

amostra=geraamostra(max(N),max(M),pteorico,e1teorico,e2teorico,semente);

numN=length(N);
numM=length(M);

aK=zeros(numN,numM);
VICIOP=aK;
VICIOE1=aK;
VICIOE2=aK;
VARP=aK;
VARE1=aK;
VARE2=aK;

icont=0;
```

```

status(1,icont,numN*numM,toc,'Obtendo posterioris...');
for in=1:length(N)
    na=N(in);
    for im=1:length(M)
        ma=M(im);
        subamostra=transpose(amostra(1:na,1:ma));
        if ma>1
            vetorc=sum(subamostra);
        else
            vetorc=subamostra;
        end
        bins=(0:ma)-0.5;
        newvetorc=hist(vetorc,bins);

        [K,vicio,var]=geraposteriori(newvetorc,na,ma,nptos,pmin,
                                    pmax,e1min,e1max,e2min,e2max);

        aK(in,im)=K;
        VICIOP(in,im)=vicio(1);
        VICIOE1(in,im)=vicio(2);
        VICIOE2(in,im)=vicio(3);
        VARP(in,im)=var(1);
        VARE1(in,im)=var(2);
        VARE2(in,im)=var(3);
        icont=icont+1;
        status(1,icont,numN*numM,toc,'Obtendo posterioris...');
    end
end

gvicio

gvar

gvarlog

function y=graficoi(x,y,i)
estilo={'-','-.',':','--'};
simbolo={'*','s','o','v','d','x',''};
cor={'r','b','g','y','k'};
limite=7*4*5;
temp=mod(i,limite);
if temp==0
    temp=limite;
end
cont=0;
for i1=1:4
    for i2=1:7
        for i3=1:5
            cont=cont+1;

```

```

        if cont==temp
            formato=[(char(cor(i3))), (char(simbolo(i2))), (char(estilo(i1)))];
            hold on;
            plot(x,y,formato);
            box on;
            set(gcf, 'Color', [1,1,1]);
            return
        end
    end
end
end
end

```

```
function y = intele2(e1,e2)
```

```

global vetorc m P LNK0
vsize=[length(e1) length(e2)];
lmax=max(vsize);

```

```

vp=transpose(ones(1,lmax).*P);
ve1=transpose(ones(1,lmax).*e1);
ve2=transpose(ones(1,lmax).*e2);

```

```

% calculando
PR1= repmat(1-ve1, [1 m+1]);
PR2= repmat(ve2, [1 m+1]);
PP= repmat(vp, [1 m+1]);

```

```

x=0:1:m;
X= repmat(x, [lmax 1]);
M= repmat(m, [lmax m+1]);

```

```
F= repmat(vetorc, [lmax 1]);
```

```

BIN01=mybinopdf(X,M,PR1);
BIN02=mybinopdf(X,M,PR2);
TERMO1=PP.*BIN01+(1-PP).*BIN02;
TERMO1=log(TERMO1);
TERMO2=TERMO1.*F;
TERMO3=sum(TERMO2,2)-LNK0;
y=transpose(exp(TERMO3));

```

```
function y = lintpe1e2(p,e1,e2)
```

```

global vetorc m LNK0
vsize=[length(p) length(e1) length(e2)];
lmax=max(vsize);

```

```
vp=transpose(ones(1,lmax).*p);
```

```

ve1=transpose(ones(1,lmax).*e1);
ve2=transpose(ones(1,lmax).*e2);

PR1= repmat(1-ve1,[1 m+1]);
PR2= repmat(ve2,[1 m+1]);
PP= repmat(vp,[1 m+1]);

x=0:1:m;
X= repmat(x,[lmax 1]);
M= repmat(m,[lmax m+1]);

F= repmat(vetorc,[lmax 1]);

BIN01=mybinopdf(X,M,PR1);
BIN02=mybinopdf(X,M,PR2);
TERMO1=PP.*BIN01+(1-PP).*BIN02;
TERMO1=log(TERMO1);
TERMO2=TERMO1.*F;
TERMO3=sum(TERMO2,2)-LNK0;
y=transpose(TERMO3);

function [vicio,variancia] = salvarfigura(vx,vy,valor teorico,textoeixox,
                                         textoeixoy,titulo,filename,iflag)

global L

nptos=1001;
xmin=min(vx);
xmax=max(vx);

newvx=linspace(xmin,xmax,nptos);
newvy=interp1(vx,vy,newvx,'spline');

if iflag==1
    figure;
    set(gcf,'Color',[1 1 1]);
end

subplot(2,2,iflag);

integral0=trapz(newvx,newvy);
if abs(integral0-1)>1e-4
    newvy=newvy/integral0;
end

integral=trapz(newvx,newvy);

plot(vx,vy/integral0,'r. ');
xlim([min(vx) max(vx)]);

```

```

hold on
plot(newvx,newvy,'m-');

media=trapz(newvx,newvx.*newvy);
variancia=trapz(newvx,newvy.*((newvx-media).^2));
[valor,indicemoda]=max(newvy);
moda=newvx(indicemoda);
vicio=moda-valorteorico;

delta=newvx(2)-newvx(1);

m=round(L/delta)+1;

vech=1:(nptos-m+1);
vech=vech-1;
vecv=transpose(1:m);

I1= repmat(vecv,[1 nptos-m+1]);
I2= repmat(vech,[m 1]);

I=I1+I2;

hpds=sum(newvy(I))/sum(newvy);
[hpd indice]=max(hpds);

li=I(1,indice);
ls=I(m,indice);

hpd=trapz(newvx(li:ls),newvy(li:ls));
li=newvx(li);
ls=newvx(ls);

[ minimo, indice]=min(abs(newvx-media));
m=round((L/2)/delta);
imin=indice(1)-m;

if imin<1
    imin=1;
end

imax=indice(1)+m;
if imax>nptos
    imax=nptos;
end

accmedia=trapz(newvx(imin:imax),newvy(imin:imax));

if (indicemoda+m)>nptos

```

```

        indicemoda=nptos-m;
end

if (indicemoda-m)<1
    indicemoda=m+1;
end

imin=indicemoda-m;
imax=indicemoda+m;

accmoda=trapz(newvx(imin:imax),newvy(imin:imax));

texto='';
texto=strvcat(texto, ['\mu = ' num2str(media,6) ' (ACC [' num2str(L,4) ']= '
                    num2str(accmedia,6) ')']);
texto=strvcat(texto, ['moda = ' num2str(modas,6) ' (ACC [' num2str(L,4) ']= '
                    num2str(accmoda,6) ')']);
texto=strvcat(texto, ['HPD = [' num2str(li,6) '];' num2str(ls,6) ' ] = '
                    num2str(hpd,6) ')']);
texto=strvcat(texto, ['vicio (moda) = ' num2str(vicio,6)]);
texto=strvcat(texto, ['\sigma^2 = ' num2str(variancia,6)]);
texto=strvcat(texto, ['integral = ' num2str(integral,6) '('
                    num2str(integral0,6) ')']);

eixox=get(gca,'Xlim');
eixoy=get(gca,'Ylim');
text(eixox(1)+(eixox(2)-eixox(1))*0.05,eixoy(2)-(eixoy(2)-eixoy(1))*0.05,
     texto,'VerticalAlignment','top','FontSize',9)
drawnow;

ylim([0 max(eixoy)]);

xlabel(textoeixox);
ylabel(textoeixoy);
title(titulo);

drawnow;
if iflag==3
    print('-dpng','-r120',filename);
    saveas(gcf,[filename '.fig']);
    close(gcf);
    drawnow;
end

function y=geraamostra(n,m,p,e1,e2,semente)
rand('state', semente);
randn('state', semente);

```



```

real=binornd(1,p,n,1);
amostra=zeros(n,m);

for i=1:n
    if real(i)==0
        class=binornd(1,e2,1,m);
    else
        class=binornd(1,1-e1,1,m);
    end
    amostra(i,:)=class;
end
y=amostra;
save amostra.txt amostra -ascii

figure;
subplot(2,2,1);
for i=1:numM
    if i==1
        texto=['m=' num2str(M(i),3)];
    else
        texto=strvcat(texto, ['m=' num2str(M(i),3)]);
    end
    graficoi(N,transpose(abs(VARP(:,i))),i);
end
h=legend(texto,'Location','Best');
set(h,'FontSize',6);
xlabel('n');
ylabel('\sigma^2');
title('\sigma^2 de p');

SAIDA=[M;VARP];
SAIDA=[[0;N'] SAIDA];
save('varp.txt','-ascii','-tabs','SAIDA');

subplot(2,2,2);
for i=1:numM
    if i==1
        texto=['m=' num2str(M(i),3)];
    else
        texto=strvcat(texto, ['m=' num2str(M(i),3)]);
    end
    graficoi(N,transpose(abs(VARE1(:,i))),i);
end
h=legend(texto,'Location','Best');
set(h,'FontSize',6);
xlabel('n');
ylabel('\sigma^2');
title('\sigma^2 de e_1');

```

```

SAIDA=[M;VARE1];
SAIDA=[[0;N'] SAIDA];
save('vare1.txt','-ascii','-tabs','SAIDA');

% Fazendo o gráfico da variancia de e2
subplot(2,2,3);
for i=1:numM
    if i==1
        texto=['m=' num2str(M(i),3)];
    else
        texto=strvcat(texto, ['m=' num2str(M(i),3)]);
    end
    graficoi(N,transpose(abs(VARE2(:,i))),i);
end
h=legend(texto,'Location','Best');
set(h,'FontSize',6);
xlabel('n');
ylabel('\sigma^2');
title('\sigma^2 de e_2');

SAIDA=[M;VARE2];
SAIDA=[[0;N'] SAIDA];
save('vare2.txt','-ascii','-tabs','SAIDA');

drawnow;
print('-dpng','-r120','variancia');
close(gcf);
drawnow;

function y = intpe2(p,e1)

global vetorc m E2 LNK0
vsize=[length(p) length(e1)];
lmax=max(vsize);

vp=transpose(ones(1,lmax).*p);
ve1=transpose(ones(1,lmax).*e1);
ve2=transpose(ones(1,lmax).*E2);

PR1= repmat(1-ve1,[1 m+1]);
PR2= repmat(ve2,[1 m+1]);
PP= repmat(vp,[1 m+1]);

x=0:1:m;
X= repmat(x,[lmax 1]);
M= repmat(m,[lmax m+1]);

F= repmat(vetorc,[lmax 1]);

```

```

BIN01=mybinopdf(X,M,PR1);
BIN02=mybinopdf(X,M,PR2);
TERMO1=PP.*BIN01+(1-PP).*BIN02;
TERMO1=log(TERMO1);
TERMO2=TERMO1.*F;
TERMO3=sum(TERMO2,2)-LNK0;
y=transpose(exp(TERMO3));

function [LNK0,vicio,var]=geraposteriori(experimento,nexp,mexp,nptos,pmin,
                                         pmax,e1min,e1max,e2min,e2max)

global vetorc m P E1 E2 pteorico e1teorico e2teorico KO LNK0

vicio=zeros(1,3);
var=zeros(1,3);

vetorc=experimento;
m=mexp;
delta=1e-4;

filename=['n=' num2str(nexp) ',m=' num2str(m,3)];

LNK0=0;
vp=unifrnd(pmin,pmax,1,500);
ve1=unifrnd(e1min,e1max,1,500);
ve2=unifrnd(e2min,e2max,1,500);

LNK0=(max(lintpe1e2(vp,ve1,ve2)))
LNK0=log(mean(exp(lintpe1e2(vp,ve1,ve2))))+LNK0

vp=linspace(pmin,pmax,nptos);
ve1=linspace(e1min,e1max,nptos);
ve2=linspace(e2min,e2max,nptos);

vp=linspace(pmin,pmax,nptos);
postp=[];
for i=1:nptos
    P = vp(i);
    postp(i)=dblquad(@inte1e2,e1min,e1max,e2min,e2max,0.01);
end
integral=trapz(vp,postp)
LNK0=LNK0-log(1/integral)
vp=linspace(pmin,pmax,nptos);
postp=[];
for i=1:nptos
    P = vp(i);
    postp(i)=dblquad(@inte1e2,e1min,e1max,e2min,e2max,delta);
end
[viciop,varp]=salvarfigura(vp,postp,pteorico,'p','posteriori p',

```

```

                                filename,filename,1);
vicio(1)=viciop;
var(1)=varp;

ve1=linspace(e1min,e1max,nptos);
poste1=[];
for i=1:nptos
    E1 = ve1(i);
    poste1(i)=dblquad(@intpe2,pmin,pmax,e2min,e2max,delta);
end
[vicioe1,vare1]=salvarfigura(ve1,poste1,e1teorico,'e_1','posteriori e_1',
                                filename,filename,2);
vicio(2)=vicioe1;
var(2)=vare1;

ve2=linspace(e2min,e2max,nptos);
poste2=[];
for i=1:nptos
    E2 = ve2(i);
    poste2(i)=dblquad(@intpe1,pmin,pmax,e1min,e1max,delta);
end

[vicioe2,vare2]=salvarfigura(ve2,poste2,e2teorico,'e_2','posteriori e_2',
                                filename,filename,3);
vicio(3)=vicioe2;
var(3)=vare2;

SAIDA=[vp;postp;ve1;poste1;ve2;poste2];
SAIDA=transpose(SAIDA);
filename=[filename '.txt'];
save(filename,'-ascii','-tabs','SAIDA');

figure;
subplot(2,2,1);
for i=1:numM
    if i==1
        texto=['m=' num2str(M(i),3)];
    else
        texto=strvcat(texto, ['m=' num2str(M(i),3)]);
    end
    graficoi(N,transpose(abs(VARP(:,i))),i);
end
set(gca,'YScale','log')
h=legend(texto,'Location','Best');
set(h,'FontSize',6);
xlabel('n');
ylabel('\sigma^2');
title('\sigma^2 de p');

```

```

subplot(2,2,2);
for i=1:numM
    if i==1
        texto=['m=' num2str(M(i),3)];
    else
        texto=strvcat(texto, ['m=' num2str(M(i),3)]);
    end
    graficoi(N,transpose(abs(VARE1(:,i))),i);
end
set(gca,'YScale','log')
h=legend(texto,'Location','Best');
set(h,'FontSize',6);
xlabel('n');
ylabel('\sigma^2');
title('\sigma^2 de e_1');

subplot(2,2,3);
for i=1:numM
    if i==1
        texto=['m=' num2str(M(i),3)];
    else
        texto=strvcat(texto, ['m=' num2str(M(i),3)]);
    end
    graficoi(N,transpose(abs(VARE2(:,i))),i);
end
set(gca,'YScale','log')
h=legend(texto,'Location','Best');
set(h,'FontSize',6);
xlabel('n');
ylabel('\sigma^2');
title('\sigma^2 de e_2');

drawnow;
print('-dpng','-r120','logvariancia');
close(gcf);
drawnow;

function y = intpe1e2(p,e1,e2)

global vetorc m K0
vsize=[length(p) length(e1) length(e2)];
lmax=max(vsize);

vp=transpose(ones(1,lmax).*p);
ve1=transpose(ones(1,lmax).*e1);
ve2=transpose(ones(1,lmax).*e2);

PR1= repmat(1-ve1,[1 m+1]);
PR2= repmat(ve2,[1 m+1]);

```

```

PP= repmat(vp, [1 m+1]);

x=0:1:m;
X=repmat(x, [lmax 1]);
M=repmat(m, [lmax m+1]);

F=repmat(vetorc, [lmax 1]);

BINO1=mybinopdf(X,M,PR1);
BINO2=mybinopdf(X,M,PR2);
TERMO1=PP.*BINO1+(1-PP).*BINO2;
TERMO2=TERMO1.^F;
TERMO3=prod(TERMO2,2)/K0;
y=transpose(TERMO3);

function y = mybinopdf(x,n,p)

if nargin < 3,
    error('stats:binopdf:TooFewInputs','Requires three input arguments');
end
y = zeros(size(x));

if ~isfloat(x)
    x = double(x);
end
if ~isfloat(n)
    n = double(n);
end
k = find(x >= 0 & x == round(x) & x <= n);
if any(k)
    t = (p(k)==0);
    if any(t)
        kt = k(t);
        y(kt) = (x(kt)==0);
        k(t) = [];
    end
    t = (p(k)==1);
    if any(t)
        kt = k(t);
        y(kt) = (x(kt)==n(kt));
        k(t) = [];
    end
end
end

if any(k)
    nk = gammaln(n(k) + 1) - gammaln(x(k) + 1) - gammaln(n(k) - x(k) + 1);
    lny = nk + x(k).*log( p(k)) + (n(k) - x(k)).*log(1 - p(k));
    y(k) = exp(lny);
end
end

```

```

k1 = find(n < 0 | p < 0 | p > 1 | round(n) ~= n);
if any(k1)
    y(k1) = NaN;
end

function y=status(inicio,i,casos,tempo,texto,correcao)

mmax=casos;
mmin=inicio;
m=i;

progressbar=['|'];
i=i-inicio+1;
casos=casos-inicio+1;
inicio=1;

passo=floor(casos*0.03);
if passo==0;
    passo=1;
end

if i==1 | i==inicio | i==casos | mod(i,passo)==0

else
    return
end

if nargin==5
    correcao=0;
end

if correcao==0
    tempoestimado=tempo/60*(casos-i)/(i+1e-6);
else
    total=0.5*mmax*(mmax+1)-0.5*mmin*(mmin-1);
    atual=0.5*m*(m+1)-0.5*mmin*(mmin-1);
    tempoestimado=tempo/60*(total-atual)/(atual);
end

clc;
for kkk=1:floor(i/casos*100/5)
    progressbar=[progressbar,'*'];
end
for kkk=(floor(i/casos*100/5)+1):20
    progressbar=[progressbar,' '];
end
progressbar=[progressbar,'|'];

```

```

fprintf('%60s\n', '*****');
fprintf('%60s\n', 'Progress');
fprintf('%60s\n', '*****');
fprintf('%60s\n', texto);
fprintf('%2s\n', ' ');
fprintf('%24s\t %5.0f\t %2s\t %5.0f\n', 'current iteration = ', i, 'of', casos);
fprintf('%24s\t %5.0f %2s\n', 'Progress = ', i/casos*100, ' %');
fprintf('%24s %22s %3s\n', '0%', progressbar, '100%');
fprintf('%2s\n', ' ');
fprintf('%40s\t %4.6f\n', 'elapsed time (min) = ', tempo/60);
fprintf('%2s\n', ' ');
fprintf('%40s\t %4.6f\n', 'time remaining (min) = ', tempoestimado);
fprintf('%2s\n', ' ');
fprintf('%60s\n', '*****');
drawnow;

figure;
subplot(2,2,1);
for i=1:numM
    if i==1
        texto=['m=' num2str(M(i),3)];
    else
        texto=strvcat(texto, ['m=' num2str(M(i),3)]);
    end
    graficoi(N, transpose(abs(VICIOP(:,i))), i);
end
h=legend(texto, 'Location', 'Best');
set(h, 'FontSize', 6);
xlabel('n');
ylabel('abs(vicio)');
title('vicio de p');

SAIDA=[M;VICIOP];
SAIDA=[[0;N'] SAIDA];
save('viciop.txt', '-ascii', '-tabs', 'SAIDA');

subplot(2,2,2);
for i=1:numM
    if i==1
        texto=['m=' num2str(M(i),3)];
    else
        texto=strvcat(texto, ['m=' num2str(M(i),3)]);
    end
    graficoi(N, transpose(abs(VICIOE1(:,i))), i);
end
h=legend(texto, 'Location', 'Best');
set(h, 'FontSize', 6);
xlabel('n');
ylabel('abs(vicio)');

```



```

title('vicio de e_1');

SAIDA=[M;VICIOE1];
SAIDA=[[0;N'] SAIDA];
save('vicioe1.txt','-ascii','-tabs','SAIDA');

subplot(2,2,3);
for i=1:numM
    if i==1
        texto=['m=' num2str(M(i),3)];
    else
        texto=strvcat(texto, ['m=' num2str(M(i),3)]);
    end
    graficoi(N,transpose(abs(VICIOE2(:,i))),i);
end
h=legend(texto,'Location','Best');
set(h,'FontSize',6);
xlabel('n');
ylabel('abs(vicio)');
title('vicio de e_2');

```

```

SAIDA=[M;VICIOE2];
SAIDA=[[0;N'] SAIDA];
save('vicioe2.txt','-ascii','-tabs','SAIDA');

```

```

drawnow;
print('-dpng','-r120','graf_vicio');
close(gcf);
drawnow;

```

```

function y = intpe2(p,e2)

```

```

global vetorc m E1 LNK0
vsize=[length(p) length(e2)];
lmax=max(vsize);

```

```

vp=transpose(ones(1,lmax).*p);
ve1=transpose(ones(1,lmax).*E1);
ve2=transpose(ones(1,lmax).*e2);

```

```

PR1= repmat(1-ve1,[1 m+1]);
PR2= repmat(ve2,[1 m+1]);
PP= repmat(vp,[1 m+1]);
x=0:1:m;
X= repmat(x,[lmax 1]);
M= repmat(m,[lmax m+1]);

```

```
F= repmat(vetorc, [lmax 1]);  
  
BIN01=mybinopdf(X,M,PR1);  
BIN02=mybinopdf(X,M,PR2);  
TERM01=PP.*BIN01+(1-PP).*BIN02;  
TERM01=log(TERM01);  
TERM02=TERM01.*F;  
TERM03=sum(TERM02,2)-LNK0;  
y=transpose(exp(TERM03));
```