

Avaliação da Confiabilidade de Itens com Testes Destrutivos - Aplicação da Estimação da Proporção em uma População Finita Amostrada sem Reposição

F. A. A. Coelho e Y. M. Tavares

Diretoria de Sistemas de Armas da Marinha - Rua Primeiro de Março 118, Ed. Barão de Ladário, Centro, Rio de Janeiro RJ

Resumo — Um problema comum na engenharia da qualidade é a estimação do número de itens considerados bons em um lote de produção. Esta análise é conhecida como inspeção por atributos, em que o resultado dos testes é dicotômico (bom ou impróprio para uso). São adotadas normas específicas para definição dos planos de amostragem em inspeções por atributos. Tais normas, entretanto, por questões de ordem prática, limitam-se a alguns valores típicos de níveis de qualidade aceitável e risco. Por outro lado, pode-se usar a distribuição hipergeométrica para fornecer informações sobre a distribuição de probabilidade (fdp) da proporção da amostra. Neste trabalho é apresentada uma nova metodologia de construção dessa fdp, em que o Teorema de Bayes é aplicado de modo a se inferir sobre a proporção da população, quando é conhecida a proporção da amostra, e são realizadas simulações Monte Carlo para validar o método proposto

Palavras-Chave — Estatística, Bayes, Inferência.

I. INTRODUÇÃO

Quando se deseja realizar a inspeção por amostragem de um lote, deve-se decidir pelo melhor plano de amostragem que atenda aos requisitos de economicidade (reduzir o tamanho da amostra) e nível de confiança (aumentar o tamanho da amostra). Para auxiliar na solução deste compromisso estão disponíveis normas [1] [2], as quais permitem definir o tamanho da amostra baseado em parâmetros como Nível de Qualidade Aceitável (NQA) e Qualidade Limite (QL). Estes parâmetros estão relacionados aos erros de decisão tipo I e tipo II, respectivamente, conforme resumido no Quadro 1, onde α é a probabilidade de cometer o erro tipo I (risco do produtor) e β é a probabilidade de cometer o erro tipo II (risco do consumidor).

QUADRO I RESUMO DOS ERROS DE DECISÃO NA INSPEÇÃO POR ATRIBUTOS

	Aceitar o Lote	Rejeitar o Lote
Lote Bom	Decisão Correta	Erro Tipo I (α)
Lote Ruim	Erro Tipo II (β)	Decisão Correta

Coelho, F. A. A., feraac@gmail.com, Tel +55(21) 2104-5937
Tavares, Y. M., yurimtavares@yahoo.com.br.

Fixados valores para esses parâmetros, tem-se determinada uma curva da probabilidade de aceitação do lote, $P(A)$, em função da porcentagem de itens defeituosos no lote, conhecida como Curva Característica de Operação (CCO), conforme ilustrado na Fig. 1.

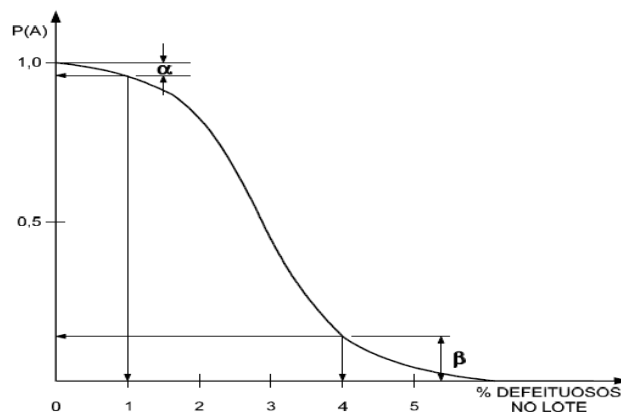


Fig. 1. Curva Característica de Operação típica.

Essas CCO são construídas tomando por base as distribuições de probabilidade Hipergeométrica, Binomial ou Poisson, conforme o plano de amostragem escolhido, ou seja, tamanho do lote, tamanho da amostra, e número aceitável de defeituosos na amostra. As normas disponíveis [1] [2] apresentam curvas traçadas apenas para alguns conjuntos típicos desses parâmetros.

Na área militar, às vezes é necessário realizar testes destrutivos para avaliar a confiabilidade de um item cuja vida útil estabelecida foi expirada, com o intuito de tomar decisão quanto a sua revalidação (extensão da vida útil). Nesses casos, a população é finita, a amostragem é sem reposição e o tamanho da amostra deve ser reduzido em relação ao tamanho do lote (população), de modo a preservar um maior número de itens remanescentes. Desse modo, a distribuição Hipergeométrica pode ser aplicada para estimar o número de itens defeituosos encontrados na amostra, conhecido o número de itens defeituosos na população [3].

Este trabalho aborda esse problema de outra forma: estimar-se a proporção de itens bons na população baseado no resultado do experimento (número de itens bons encontrados na amostra). Para tal, é apresentado um novo método para construir a função distribuição de probabilidade

de itens bons na população, donde se pode obter estatísticas diversas e realizar inferências com os níveis de confiança pretendidos.

II. DESENVOLVIMENTO MATEMÁTICO

Seja um lote de N itens sob avaliação, que contém M itens defeituosos. Realiza-se então uma amostragem sem reposição de n itens para ensaio destrutivo, e observam-se x sucessos (itens defeituosos encontrados nessa amostra). A função distribuição de probabilidade da proporção de itens defeituosos na amostra, $p=x/n$, é uma distribuição Hipergeométrica dada por (1) [3].

$$P(p = x | n) = \begin{cases} \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}, & x = 0, 1, \dots, \min\{n, M\} \\ 0, & \text{caso contrário} \end{cases} \quad (1)$$

A Equação (1) nos dá a probabilidade da proporção, p , em uma amostra, quando se conhece a proporção da população, ou M , ou seja $P(p | M)$.

Interessa-nos, entretanto, saber qual a probabilidade de M , dada a proporção da amostra, ou seja $P(M | p)$. Usando o teorema de Bayes obtém-se:

$$P(M | p) = \frac{P(M)P(p | M)}{\sum_{M=0}^N P(M)P(p | M)} \quad (2)$$

Como M pode assumir qualquer valor entre 0 e N , a probabilidade de ocorrência de um valor em particular, em princípio, é $P(M)=1/(N+1)$. Substituindo este resultado em (2) resulta

$$P(M | p) = \frac{\frac{1}{N+1} P(p | M)}{\sum_{M=0}^N \frac{1}{N+1} P(p | M)} = \frac{\frac{1}{N+1} P(p | M)}{\frac{1}{N+1} \sum_{M=0}^N P(p | M)} \Rightarrow P(M | p) = \frac{P(p | M)}{\sum_{M=0}^N P(p | M)} \quad (3)$$

Perceba que, realizado o experimento e encontrados x sucessos, a probabilidade de $M < x$ é nula, bem como a probabilidade de $M > N-(n-x)$. Então (3) pode ser reescrita como

$$P(M | p) = \frac{P(p | M)}{\sum_{M=x}^{N-(n-x)} P(p | M)} \quad (4)$$

Agora substituindo (1) em (4) e eliminando alguns termos resulta:

$$P(M | p) = \begin{cases} \frac{\binom{M}{x} \binom{N-M}{n-x}}{\sum_{M=x}^{N-(n-x)} \binom{M}{x} \binom{N-M}{n-x}}, & M = x, x+1, \dots, N-(n-x) \\ 0, & \text{caso contrário} \end{cases} \quad (5)$$

Lembrando que x é o número de sucessos observados na amostra de n elementos retirados sem reposição, então, $x = pn$.

III. ESTUDO DE CASO

Considere-se um lote de cinquenta itens ($N=50$), cuja validade de dez anos se expirou. Com o intuito de se revalidar o lote por um certo período, foram realizados ensaios em uma amostra de cinco itens ($n=5$) e verificado que apenas um encontrava-se com os parâmetros operacionais fora da especificação ($x=1$). Aplicando estes dados em (5) obtém-se a seguinte função de probabilidade de M .

$$P(M | p) = \begin{cases} \frac{\binom{M}{1} \binom{50-M}{4}}{\sum_{M=1}^{46} \binom{M}{1} \binom{50-M}{4}}, & M = 1, 2, \dots, 46 \\ 0, & \text{caso contrário} \end{cases}$$

$$P(M | p) = \begin{cases} \frac{M(50-M)(49-M)(48-M)(47-M)}{\sum_{M=1}^{46} M(50-M)(49-M)(48-M)(47-M)}, & M = 1, 2, \dots, 46 \\ 0, & \text{caso contrário} \end{cases}$$

Os valores das probabilidades estão listados na Tabela I, enquanto que as Figs. 2 a 4 apresentam graficamente a Função Distribuição de Probabilidade (FDP), a Função de Distribuição Acumulada (FDA) e a Curva Característica de Operação (dada por $CCO=1-FDA$), respectivamente.

TABELA I VALORES DAS PROBABILIDADE DE M (VALORES NÃO APRESENTADOS SÃO NULOS)

M	1	2	3	4	5
$P(M p)$	0,0117	0,0216	0,0297	0,0362	0,0414
M	6	7	8	9	10
$P(M p)$	0,0452	0,0480	0,0497	0,0506	0,0507
M	11	12	13	14	15
$P(M p)$	0,0502	0,0492	0,0477	0,0458	0,0436
M	16	17	18	19	20
$P(M p)$	0,0412	0,0386	0,0359	0,0332	0,0304
M	21	22	23	24	25
$P(M p)$	0,0277	0,0250	0,0224	0,0199	0,0177
M	26	27	28	29	30
$P(M p)$	0,0153	0,0133	0,0114	0,0096	0,0081
M	31	32	33	34	35
$P(M p)$	0,0067	0,0054	0,0044	0,0034	0,0026
M	36	37	38	39	40
$P(M p)$	0,002	0,002	0,001	0,001	0,001
M	41	42	43	44	45
$P(M p)$	0,000	0,000	0,000	0,000	0,000

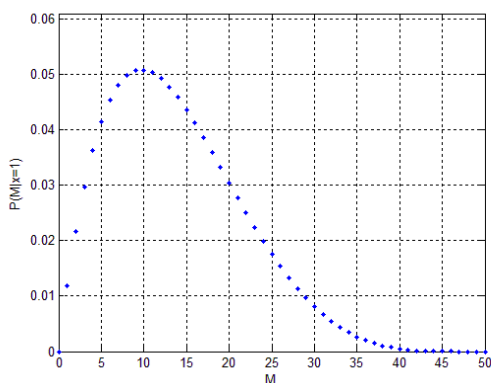


Fig. 2. Função Distribuição de Probabilidade.

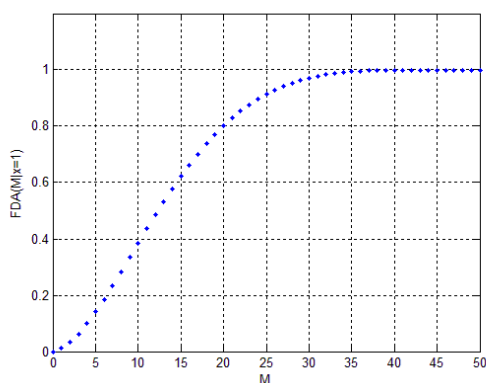


Fig. 3. Função Distribuição Acumulada.

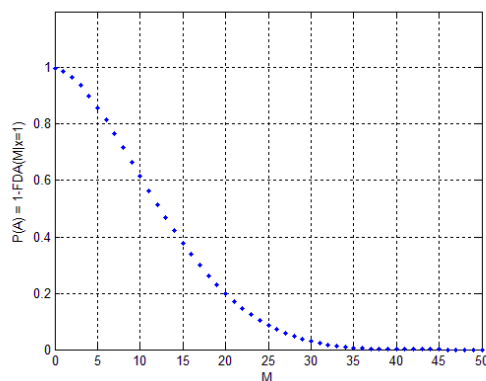


Fig. 4. Curva Característica de Operação.

A partir dessas curvas, diversas informações podem ser retiradas. Por exemplo, a partir da CCO, estabelecendo-se um erro do consumidor, $\beta=5\%$ (nível de confiança de 95%), é possível afirmar que existem menos de 28 itens defeituosos na amostra ($M < 28$), ou que a confiabilidade do lote é $R = (50 - 28)/50 = 44\%$.

IV. SIMULAÇÃO MONTE CARLO

Para validar o método proposto, foi realizada uma simulação Monte Carlo. Para a simulação foram criados 51 lotes de cinquenta itens cada ($N=50$). Cada lote era identificado pelo número de defeituosos que continha. Assim o lote 0 possuía zero itens defeituosos, o lote 1 possuía um item defeituoso e assim sucessivamente.

Para cada lote, foram amostrados cinco itens ($n=5$) e repetido o experimento cem mil vezes (100.000 histórias) com reposição. Os resultados foram armazenados em uma matriz com 51 linhas (valores possíveis de M) e 6 colunas (valores possíveis de x). Cada linha correspondia a cada um dos 51 lotes, e cada coluna ao número de defeituosos encontrados nas amostras. Assim, ao final da simulação, dividindo-se cada item da matriz pelo somatório de sua coluna correspondente, obtém-se a probabilidade de se encontrar x itens defeituosos quando o lote possuir M itens defeituosos. O fluxograma usado para realizar a simulação pode ser visto na Fig. 5, e o script de Matlab utilizado está listado no apêndice.

Os resultados encontrados no histórico de amostras obtidas foram utilizados para montar os gráficos de probabilidade de se encontrar x itens defeituosos exibidos na Fig. 6.

A Fig. 7 mostra uma comparação entre a FDP obtida no estudo de caso da Seção 3 (linha azul formada pela ligação dos pontos teóricos calculados), com o resultado da simulação com $x=1$ (pontos em vermelho), validando o procedimento sugerido neste trabalho.

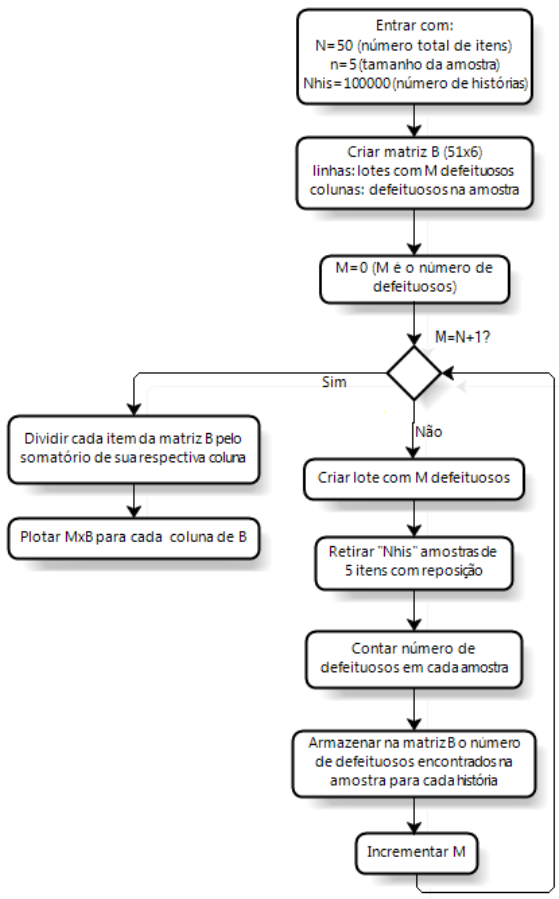


Fig. 5. Fluxograma utilizado na simulação.

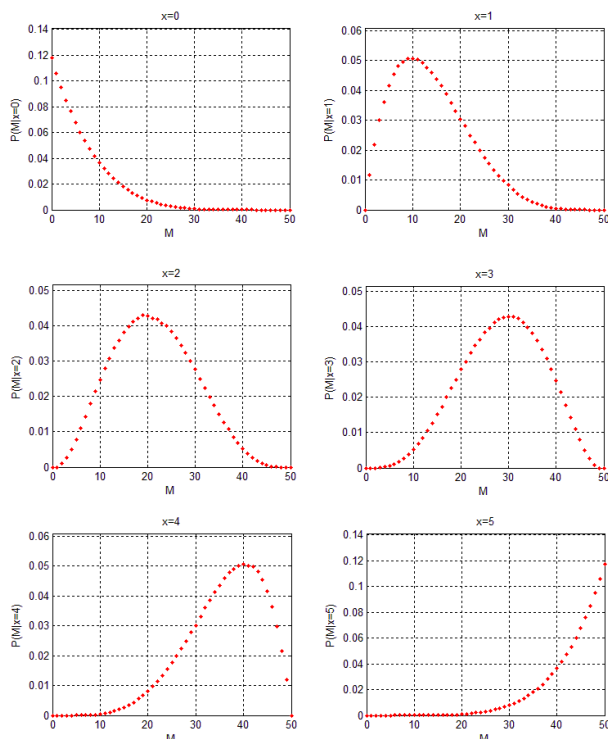


Fig. 6. Gráficos dos resultados da simulação.

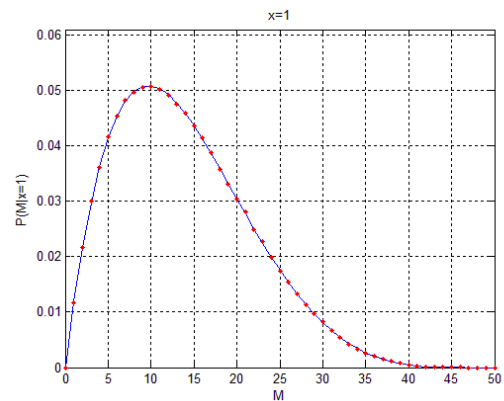


Fig. 7. Comparação entre a FDP obtida pelo método proposto (linha azul formada pela ligação dos pontos teóricos calculados) e a obtida em simulação (pontos em vermelho).

V. CONCLUSÃO E CONSIDERAÇÕES FINAIS

Foi aplicada a distribuição hipergeométrica ao problema da inspeção por atributos, de modo a se fazer inferência sobre a proporção da amostra. Para tal o Teorema de Bayes é aplicado no intuito de obter a proporção da população, quando é conhecida a proporção da amostra.

O método proposto foi validado comparando-se a Função Distribuição de Probabilidade obtida com os resultados de Simulações Monte Carlo realizadas.

APÊNDICE - SCRIPT DE MATLAB USADO PARA SIMULAÇÃO MONTE CARLO

```

%Artigo SIGE 2012
%SCRIPT de simulação MONTE CARLO
%-----
%Tarefa:
%0: Para M=0 até 50
%1: Montar lote contendo M itens defeituosos
%2: Retirar Nhis amostras de 5 itens (sem reposição) e contar
o número de
%defeituosos na amostra
%3: Armazenar os seis números encontrados em uma matriz
B(M=51,x=6)
%4: fim do loop
%5: somar o número de resultados das colunas de B
%6: dividir cada elemento de B pelo resultado acima da sua
respectiva
%coluna
%7: montar gráficos das colunas de B
%Onde:
%N: número total de itens
%M: número de itens defeituosos em N
%n: número de itens amostrados
%x: número de itens defeituosos na amostra "n"
%-----
% versão de 30/08/12

clear all %limpa todas as variáveis
clc %limpa prompt de comando
close all %fecha todas as janelas
  
```

```

%-----

%Entrada de Dados
N=50;      %N: tamanho do lote a ser avaliado
n=5;       %n: tamanho da amostra
Nhis=100000; %número de histórias

%Declaração de variáveis
B=zeros(N+1,n+1); %B:matriz para guardar resultados das
                  %histórias

%Principal

B(1,1)=Nhis; %para M=0 (nenhum item com defeito), só
             %haverá x=0
for i=1:1:N %loop para criar diversos lotes com M
             %itens defeituosos diferentes
    lote=zeros(1,N); %lote criado para cada M
    for j=1:1:i
        lote(1,j)=1; %criação do lote com i defeituosos:
                     %valor "0" é normal, valor "1" é defeituoso
    end
    x=zeros(1,n+1); %vetor que guarda o número de
                   %defeitos para cada história
    for k=1:1:Nhis %história
        def=0;
        flagr=false; %flag de controle para não repetir números
                     %em r
        while flagr==false %controle para não repetir números
            r=randi([1,50],1,n); %gera um vetor de n números
                                %aleatórios de 1 a 50
                                %r é o vetor que contém os índices para
                                %buscar itens no "lote"
            r=sort(r); %coloca os elementos do vetor em
                     %ordem crescente
            flagr=true; %inicia teste
            for q=1:1:(n-1) %testar se não há índice
                           %(número) repetido
                if r(1,q)==r(1,q+1)
                    flagr=false; %se houver número repetido,
                                %"flagr" volta para "false"
                    break;
                end
            end
        end
        for t=1:1:n %contar número de itens defeituosos na
                   %amostra: x
            if lote(1,r(1,t))==1 %r aponta item em "lote" e
                                %verifica-se se o mesmo é defeituoso
                def=def+1;
            end
        end
        x(1,def+1)=x(1,def+1)+1; %incrementa posição em x
                                %referente ao número de defeituosos
    end
    B(i+1,:)=x(1,:); %guarda proporção de número de
                    %defeituosos em Nhis na
                    %posição correspondente a M

%defeituosos no lote N
end

Bx=sum(B); %soma todas as colunas de B
for u=1:1:n+1 %divide cada elemento de B pelo
              %somatório de sua coluna
    B(:,u)=B(:,u)./Bx(1,u);
end

%Exibir resultados
subplot(2,3,1)
plot(0:1:N,B(:,1),'r.')
axis([0 N 0
1.2*max(B(:,1))]);xlabel('M');ylabel('P(M|x=0)');title('x=0');
grid on

subplot(2,3,2)
plot(0:1:N,B(:,2),'r.')
axis([0 N 0
1.2*max(B(:,2))]);xlabel('M');ylabel('P(M|x=1)');title('x=1');
grid on

subplot(2,3,3)
plot(0:1:N,B(:,3),'r.')
axis([0 N 0
1.2*max(B(:,3))]);xlabel('M');ylabel('P(M|x=2)');title('x=2');
grid on

subplot(2,3,4)
plot(0:1:N,B(:,4),'r.')
axis([0 N 0
1.2*max(B(:,4))]);xlabel('M');ylabel('P(M|x=3)');title('x=3');
grid on

subplot(2,3,5)
plot(0:1:N,B(:,5),'r.')
axis([0 N 0
1.2*max(B(:,5))]);xlabel('M');ylabel('P(M|x=4)');title('x=4');
grid on

subplot(2,3,6)
plot(0:1:N,B(:,6),'r.')
axis([0 N 0
1.2*max(B(:,6))]);xlabel('M');ylabel('P(M|x=5)');title('x=5');
grid on

% FIM -----

REFERÊNCIAS

[1] UNITED STATES OF AMERICA. Department of Defense, MIL-STD-1916 "Test Method Standard. DoD preferred Methods for acceptance of product", Abril 1996.
[2] ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, NBR 5426 "Planos de amostragem e procedimentos na inspeção por atributos", Janeiro 1985.
[3] CANCHO V. G., "Noções de Estatística e Probabilidade", Universidade Federal de Ouro Preto, 2004.

```