

SOBRE A DISTRIBUIÇÃO DO PARÂMETRO DE
CORREÇÃO NUM MODELO LINEAR
REDUZIDO DE APRENDIZAGEM

VILMA MAYUMI TACHIBANA



UNICAMP

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E CIÊNCIA DA COMPUTAÇÃO

CAMPINAS - SÃO PAULO
BRASIL

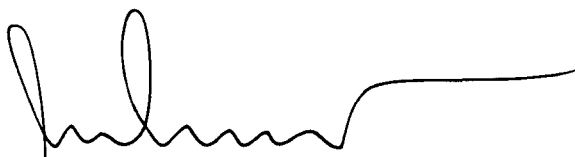
T117s

6093/BC

SOBRE A DISTRIBUIÇÃO DO PARÂMETRO DE CORREÇÃO NUM
MODELO LINEAR REDUZIDO DE APRENDIZAGEM

Este exemplar corresponde a redação ^{final} da tese ^{devidamente corrigida} defendida pela Srta. Vilma Mayumi Tachibana e aprovada pela comissão julgadora.

Campinas, 31 de janeiro de 1984.


P.P. Prof. Dr. Hermann G. Rohrer
Orientador

Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência da Computação, UNICAMP, como requisito parcial para a obtenção do título de Mestre em Estatística.

UNICAMP
BIBLIOTECA CENTRAL

Aos meus pais, vovô

Mi, Sã e New;

Lúcia e vó Katsuko

(que se transformaram

em luzes no decorrer

da tese).

- Professor Hermann G. Hohrer, por ter aceito o encargo da orientação deste trabalho, manifestando paciência, compreensão e competência em todas as fases de elaboração;
- Amiga prof. Regina Célia Moran, pela inestimável ajuda com críticas e sugestões,
- Colegas, professores e alunos do IMECC, pelos bons momentos proporcionados durante estes anos, valorizando as palavras amizade e respeito mútuo,
- funcionários do IMECC, pela eficiência nos serviços prestados, principalmente pela amizade,
- Lourdes, pelo paciente e perfeito trabalho datilográfico e também pela presteza e responsabilidade no atendimento,
- Todos que direta ou indiretamente contribuíram para realização deste trabalho,

Muito obrigada!

INDÍCE

CAPÍTULO	I - Introdução	1
1.1	- Histórico	1
1.2	- Um sistema linear capaz de aprender	2
1.3	- Considerações sobre a construção de um sistema capaz de aprender	8
1.4	- Objetivo do trabalho	12
CAPÍTULO	II - Desenvolvimento Análitico	14
2.1	- Distribuição Exata	14
2.2	- Distribuição da Estatística de Ordem Assintótica.	34
CAPÍTULO	III - Desenvolvimento por Simulação	40
3.1	- Procedimento da Simulação	40
3.2	- Estimação dos Parâmetros de uma Beta (p,q)	45
CAPÍTULO	IV - Resultados	50
CAPÍTULO	V - Conclusão	75
APÊNDICE	I	77
BIBLIOGRAFIA	105

CAPÍTULO I

INTRODUÇÃO

1.1. HISTÓRICO

O estudo matemático de processos de aprendizagem se originou de certos experimentos da escola de psicologia denominada comportamentalismo. A fim de entender melhor os conceitos aqui usados, vamos apresentar um experimento introduzido por Humphreys (1939).

Foi colocada uma lâmpada na extrema direita e outra de igual intensidade na extrema esquerda de um lado de uma mesa. A iluminação dessas lâmpadas era manipulada pelo experimentador ("professor"), fora da vista dos sujeitos ("alunos"). As lâmpadas eram facilmente observadas pelos "alunos" de vários ângulos na sala de experiência. A primeira lâmpada era estímulo inicial e quando acendia indicava o início de um ensaio. A lâmpada esquerda era sempre ligada pelo experimentador e os "alunos" foram instruídos para escreverem em espaços preparados: "X" se eles esperavam que em seguida seria acesa a lâmpada direita e "O" a ausência da luz. Depois de 5 segundos, algumas vezes a lâmpada direita era acesa, indicando se os alunos acertavam ou não a resposta. Por causa disso, o sinal da lâmpada direita é chamado de reforço, e os alunos

tinham que "aprender" a frequência de acendimento da lâmpada - reforço. Após um certo número de ensaios, Humphreys constatou que a frequência de acendimento da lâmpada direita era quase igual a frequência de "X". No sentido restrito em que nós induzimos a palavra "aprender", podemos dizer que os alunos aprendiam.

Em seguida vários autores construíram modelos matemáticos, demonstrando o mesmo comportamento daquele observado por Humphreys. O interesse em uma discussão matemática não está em explicar os fenômenos que podemos observar na "natureza", mas estudar sobre os fenômenos apresentados por máquinas com características de imitar as construções e os comportamentos das preposições desenvolvidas por seres humanos. Estas máquinas são as que fazem: leitura automática de caracteres, classificação de figuras geométricas, algoritmos de classificação, etc.

Agora, temos como objetivo principal resolver a questão de como podemos construir tais algoritmos capazes de aprender.

1.2. UM SISTEMA LINEAR CAPAZ DE APRENDER

DEFINIÇÃO 1.2.1. Um sistema linear capaz de aprender medidas é dado pela quintupla

$$(\Omega, S, g, f, v)$$

onde temos:

Ω : é a classe dos conjuntos do "ambiente" do sistema com uma σ -álgebra S ;

$g: S \rightarrow \mathbb{R}^+$ é a medida de um "evento" $A \in S$ como "apresentação" de um evento no meio do sistema;

$v: S \rightarrow \mathbb{R}^+$ é uma medida de um evento $A \in S$ como "suposição" do sistema sobre um evento;

$f: g \times v \rightarrow v$ é uma função de correção da "memória" (o conjunto das suposições) do sistema após ocorrência de um evento $A \in S$.

Se as medidas na definição 1.1.1 são normalizadas a 1, o sistema é chamado um sistema capaz de aprender probabilidades. E nesse sistema, para S' uma subclasse de S , podemos obter equações funcionais através da propriedade de medida de v , para os seguintes casos:

1. S' é uma subclasse que contém conjuntos disjuntos;
2. S' é uma subclasse que contém conjuntos que são p-independentes.

No primeiro caso, através das equações

$$v(A \cup B) = v(A) + v(B) \implies$$

$$f'(g(A \cup B), v(A \cup B)) = f'(g(A), v(A)) + f'(g(B), v(B))$$

segue a função ajuste aditivo

$$f'(g(A), v(A)) = \frac{\beta g(A) + \alpha v(A)}{\beta + \alpha}, \quad (1.2.1)$$

$$(\beta, \alpha) \neq (0, 0), \quad \beta, \alpha \in \mathbb{R}^+$$

E para o segundo caso, temos através das

$$v(A \cap B) = v(A) \cdot v(B)$$

$$f^*(g(A \cap B), v(A \cap B)) = f^*(g(A), v(A)) \cdot f^*(g(B), v(B))$$

segue a função ajuste multiplicativo

$$f^*(g(A), v(A)) = \frac{(g(A))^\beta (v(A))^\alpha}{\int_{S'} (g(w))^\beta (v(w))^\alpha dw} \quad (1.2.2)$$

$$\text{com } \beta, \alpha \in \mathbb{R}$$

Desde que $c \cdot \log(f^*) = f'$, com $c > 0$, então f^* é uma função monotonicamente crescente de f . Portanto achamos suficiente tratar a função dada por 1.2.1.

Existem, claramente, vários tipos de aprendizagem, associados a sistemas capazes de aprender; um deles é:

O sistema puramente estocástico, definido pela fórmula de recorrência

$$v_N = \beta_N g_N + \alpha_N v_{N-1}, \quad (\beta_N, \alpha_N) \in (0, 1)^2, \quad \forall N \quad (1.2.3)$$

onde N é o índice de tempo, e as distribuições das variáveis aleatórias g_{N+1} , β_{N+1} e α_{N+1} são cópias das distribuições aleatórias g_N , β_N e α_N para $N \geq 1$. Os coeficientes α e β podem ser considerados como parâmetros para a classificação de processos de aprendizagem, isto é, para sequências de ajustamentos ou correções.

Aqui tratamos do caso mais geral, onde em cada etapa de "aprendizagem", os parâmetros α_N e β_N são escolhidos independentemente das distribuições sobre \mathbb{R}^+ com $\mathbb{E}(\alpha)$ e $\mathbb{E}(\beta)$ finitas e $0 \leq \mathbb{V}(\alpha) < \infty$ e $0 < \mathbb{V}(\beta) < \infty$. E também em cada etapa um vetor g_N é escolhido de uma distribuição sobre \mathbb{R}^n com esperanças e variâncias finitas.

Então

a) $\lim_{N \rightarrow \infty} \mathbb{E}(v_N)$ existe e tem o valor

$$\mathbb{E}(v_\infty) = \frac{\mathbb{E} \beta}{1 - \mathbb{E} \alpha} \mathbb{E} g \quad (1.2.4)$$

se e somente se $\mathbb{E} \alpha < 1$.

b) $\lim_{N \rightarrow \infty} \mathbb{V}(v_N)$ existe com

$$\mathbb{V}(v_\infty) = \frac{1}{1 - \mathbb{E} \alpha^2} (\mathbb{V} \beta g + \mathbb{E}^2(v_\infty) \mathbb{V} \alpha) \quad (1.2.5)$$

se e somente se $\mathbb{E} \alpha^2 < 1$.

Agora, queremos uma relação entre as variáveis aleatórias α e β para termos um processo de aprendizagem não-viciado, um processo para qual vale $\mathbb{E}(v_\infty) = \mathbb{E}g$. Achamos que para isso:

$$\mathbb{E}\beta + \mathbb{E}\alpha = 1 \quad (1.2.6)$$

Certamente, a maneira mais fácil de assegurar a validade da equação (1.2.6), consiste em fazer

$$\beta = 1 - \alpha, \quad (1.2.7)$$

perdendo a independência entre α e β . Procedendo dessa maneira, temos o segundo tipo de aprendizagem:

O sistema reduzido, caracterizado pela fórmula de recorrência

$$v_N^* = (1 - \alpha_N)g_N + \alpha_N v_{N-1}^* \quad (1.2.8)$$

onde as distribuições das variáveis aleatórias g_{N+1} e α_{N+1} são cópias das variáveis g_N e α_N , respectivamente, para todos os $N \geq 1$, e as variáveis são p-independentes entre si.

Então valem:

$$a) \lim_{N \rightarrow \infty} \mathbb{E}(v_N^*) = \mathbb{E}v_\infty^* = \mathbb{E}g \quad (1.2.9)$$

se e somente se $\mathbb{E}\alpha < 1$.

$$b) \lim_{N \rightarrow \infty} IV(v_N^*) = IV(v_\infty^*) = \frac{IE(1-\alpha)^2}{IE(1-\alpha^2)} IVg \quad (1.2.10)$$

se e somente se $IE\alpha^2 < 1$.

Temos o seguinte corolário que descreve vantagens que um processo reduzido tem sobre puramente estocástico.

COROLÁRIO 1.2.1. Se as variáveis α e g são idênticas as variáveis α e g no processo reduzido, temos

$$IVv_\infty^* \leq IVg \iff IE\alpha \geq IE\alpha^2 \quad (1.2.11)$$

Se $IEg = 0$, então

$$IVv_\infty \leq IVg \iff IE\beta^2 \leq IE(1-\alpha^2) \quad (1.2.12)$$

Se $IEg \neq 0$

$$IVv_\infty \leq IVg \iff \frac{IV\beta + \left[\frac{IE\beta}{1-IE\alpha} \right]^2 IV\alpha}{1 - (IE\beta^2 + IE\alpha^2)} \leq \frac{IVg}{IEg} \quad (1.2.13)$$

e $IE\beta^2 + IE\alpha^2 \leq 1$.

Se $IE\beta^2 + IE\alpha^2 > 1 \implies IVv_\infty \geq IVg$.

A equação (1.2.1) restringe os parâmetros α e β no espaço \mathbb{R}^+ , então para o caso reduzido $\alpha \geq 0$ e $1-\alpha \geq 0$. Nesse caso α é uma variável aleatória definida sobre o intervalo (0,1) cumprindo sempre $\mathbb{E} \alpha \geq \mathbb{E} \alpha^2$.

Então o processo reduzido sempre produz uma distribuição da variável aleatória v_∞ , cuja variância é menor do que aquela variável aleatória a ser aprendida, isto é $\text{IV}(v_\infty^*) \leq \text{IV}g$. As demonstrações e maiores discussões acham-se em Rohrer (1978).

1.3. CONSIDERAÇÕES SOBRE A CONSTRUÇÃO DE UM SISTEMA CAPAZ DE APRENDER.

O nosso sistema "capaz de aprender" está exposto a um ambiente, que produz seqüências de padrões a serem classificados. Ele deveria aprender classificações certas. A probabilidade deverá ser entendida como uma estimativa da chance de que a designação de padrões para as classes, tenha sucesso.

O nosso sistema capaz de aprender (Ω, S, g, f, v) sabe que existe um agrupamento \mathcal{C} sobre Ω , com várias subclasses $C_i \in \mathcal{C}$. E também que qualquer padrão g pertence a uma subclasse C_i . No sistema que está sendo apresentado, as subclasses são descritas por "modelos" $R_j (\hat{=} v_j)$.

No nosso caso específico, trabalhamos sobre um espaço \mathbb{R}^n onde localizam-se as subclasses caracterizadas por hipercubos

com distribuição inicial uniforme para os R_j , e os padrões tem distribuição multinormal. Supomos que o sistema conhece os números das subclasses de Ω em C , mas não conhece o próprio \mathcal{C} . Supomos também que, pelo menos na parte inicial do processo de aprendizagem, o sistema tem um professor ao lado que pode verificar as classificações do sistema de tal forma que quando o sistema errou (isto é, $g \in C_i$ não vale) ele dá a classificação certa (isto é, $g \in C_k$, com $k \neq i$). No outro caso (quando o sistema não errou) ele simplesmente reafirma a classificação dada.

Quanto ao sistema, \mathcal{C} depende apenas do professor, e se ele mudar, o conhecimento aprendido anteriormente sobre \mathcal{C} pode estar sujeito a mudanças. Se as subclasses C_i e C_k tinham representantes v_i e v_k na memória do sistema ($v_i \neq v_k$), e estas classes não são consideradas mais distintas (por exemplo, por uma mudança do professor), queremos então como uma exigência mínima que os representantes v_i e v_k se "aproximem", talvez até coincidir, isto é, que o sistema extinga da sua memória a distinção entre as duas classes. Verbos como "aproximar, coincidir" implicam tacitamente na existência de uma métrica. Explicitamente vamos exigir que o \mathbb{R}^n das "apresentações" e das "suposições" possua a métrica euclidiana, denominada por $\|\cdot\|$. As decisões tomadas pelo sistema são baseadas em medições de distâncias entre "modelos" e os padrões submetidos. Quer dizer é escolhido como certo o modelo R_j que apresentou a menor distância do padrão submetido. No nosso caso específico, calculadas as distâncias D_j , entre pontos

R_j uniformes n -dimensional e o padrão normal n -dimensional, escolhemos como certo o modelo R_k correspondente a menor distância D_k .

E a cada etapa corrigimos o modelo R_k escolhido como certo, a fim de que ele realmente sempre forneça a menor distância D_j .

Então, para que ocorra a convergência do processo de aprendizagem, é necessário que a função correção satisfaça a seguinte condição de Lipschitz:

$$\sup_{v_i^{(N)} \neq v_k^{(N)}} \frac{\|f(g, v_i^{(N)}) - f(g, v_k^{(N)})\|}{\|v_i^{(N)} - v_k^{(N)}\|} \leq 1 \quad (1.3.1)$$

Isto é, f é uma função que diminui distância, pois aplicando-o sobre o sistema reduzido, pelas vantagens citadas anteriormente, a condição fica:

$$\begin{aligned} & \| (1-\alpha_N)g_N + \alpha_N v_i^{(N)} - ((1-\beta_N)g_N + \beta_N v_k^{(N)}) \| \\ & \leq \| v_i^{(N)} - v_k^{(N)} \|, \quad \forall v_i^{(N)} \neq v_k^{(N)} \end{aligned} \quad (1.3.2)$$

Apenas com as condições anteriormente impostas sobre α e β (ou α e $(1-\alpha)$ no processo de aprendizagem reduzido), ou seja, $(\alpha, \beta) \in (0, 1)^2$; não podemos obter a propriedade de adaptação a qualquer mudança no ambiente do sistema. Então para que possamos

obter tal propriedade, devemos impor mais algumas restrições sobre as variáveis α e β .

Então Rohrer (1978) chegou que um operador definido como

$$\beta = 1 - \alpha(g_N, \{v_i^{(N)}\}) = \frac{\|g_N - v_i^{(N)}\|}{\max_{1 \leq j \leq S} \|g_N - v_j^{(N)}\| + a}, \quad a \in \mathbb{R}^+ \quad (1.3.3)$$

cumpra a condição (1.3.1).

A função de aprendizagem reduzido será escrita explicitamente como:

$$v_i^{(N+1)} = \frac{\|g_N - v_i^{(N)}\|}{\max_{1 \leq j \leq S} \|g_N - v_j^{(N)}\| + a} g_N + \left(1 - \frac{\|g_N - v_i^{(N)}\|}{\max_{1 \leq j \leq S} \|g_N - v_j^{(N)}\| + a}\right) v_i^{(N)}, \quad (1.3.4)$$

$$a \geq 0.$$

Segundo o teorema de Iosifescu (1969), esta função de aprendizagem introduz para v uma distribuição assintoticamente normal.

Podemos ver que a forma de (1.3.4) mostra um comportamento que se espera intuitivamente de uma tal função: se o padrão em questão está perto do representante de sua classe, este representante não precisa ser corrigido muito, e no caso contrário, se o padrão se encontrou muito distante, a função assegura que a memória sobre esta classe será mudada drasticamente. Isto quer dizer

que quando cria-se os representantes de uma subclasse (a maneira não importa) e se eles forem mal criados, a função de correção assegura que a aproximação aos representantes "verdadeiros" acontece relativamente rápida, e se forem bem criados, esta mesma função de correção assegura que este estado bom não é deixado.

1.4. OBJETIVO DO NOSSO TRABALHO.

Dado um ambiente, se utilizamos a função de correção β da maneira como foi definida em 1.3.3, ocorre a convergência do processo de aprendizagem.

Então temos como interesse principal determinar a distribuição dessa variável aleatória β , em um caso específico. Nós escolhemos o seguinte ambiente experimental:

- 1) os pontos (padrões) g são criados através de uma distribuição multinormal.
- 2) cada classe é caracterizada por um hipercubo com distribuição uniforme.

No capítulo II, fazemos um desenvolvimento analítico e derivamos a distribuição da variável aleatória β . Devido a intratabilidade da função densidade de probabilidade β , numa forma simples e explícita, recorreremos ao auxílio de um computador. No capítulo III, damos um enfoque paramétrico, simulamos as variáveis β e

estimamos os parâmetros de uma distribuição Beta (p,q) ; a distribuição que propomos para a nossa variável β . Essas variáveis nem sempre podem ser consideradas distribuídas como Beta (p,q) . Então no capítulo IV, verificamos quais as condições necessárias para que β venha a ter distribuição proposta. No capítulo V tiramos as conclusões finais expondo uma conjetura que possibilite um sistema capaz de aprender, não necessitar mais de uma presença constante de um "professor". Anexamos também um Apêndice, com o programa de simulação e um resultado fornecido por este programa.

CAPÍTULO II

DESENVOLVIMENTO ANALÍTICO

2.1. DISTRIBUIÇÃO EXATA.

Sejam X_1, X_2, \dots, X_n variáveis aleatórias independentes tendo a densidade normal

$$f_{X_i}(x_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x_i - \mu_i}{\sigma_i}\right)^2\right\}, \quad -\infty < x_i < \infty \quad (2.1.1)$$

Denotamos $X_i \sim N(\mu_i, \sigma_i)$.

O k -ésimo momento dessa variável pode ser calculado através de

$$\mathbb{E} X_i^k = (M^{(k)}(t) |_{t=0}) / i^k, \quad i^2 = -1 \quad (2.1.2)$$

$t \in \mathbb{R}$, $k > 0$ é um inteiro, onde $M^{(k)}(t) |_{t=0}$ é a k -ésima derivada da sua função característica $M(t) = \exp(\mu_i t i - \frac{\sigma_i^2 t^2}{2})$ no ponto $t = 0$.

Sejam Y_1, Y_2, \dots, Y_n variáveis aleatórias independentes e tendo a densidade uniforme

$$f_{Y_i}(y_i) = \begin{cases} \frac{1}{b_i - a_i}, & a_i \leq y_i \leq b_i \\ & a_i, b_i \in \mathbb{R} \\ 0, & \text{caso contrário} \end{cases} \quad (2.1.3)$$

Denotamos $Y_i \sim U(a_i, b_i)$.

O k -ésimo momento da variável aleatória Y_i é definido como

$$\mathbb{E} Y_i^{(k)} = \frac{b_i^{k+1} - a_i^{k+1}}{(k+1)(b_i - a_i)}, \quad k > 0 \text{ inteiro.} \quad (2.1.4)$$

Sejam W_1, W_2, \dots, W_n variáveis aleatórias, onde W_i são definidas como o quadrado da diferença entre as variáveis aleatórias X_i e Y_i . Isto é,

$$W_i = (X_i - Y_i)^2 = Z_i^2.$$

Como primeiro passo, determinamos a função densidade da variável aleatória Z_i . Por convolução temos que:

$$f_{Z_i}(z_i) = \int_{-\infty}^{\infty} f_{Y_i}(x_i - z_i) f_{X_i}(x_i) dx_i, \quad (2.1.5)$$

para $-\infty < x_i < \infty$ e $a_i \leq x_i - z_i \leq b_i$.

Substituindo as densidades das variáveis X_i e Y_i na função acima, temos:

$$\begin{aligned}
 f_{Z_i}(z_i) &= \int_{a_i+z_i}^{b_i+z_i} f_{Y_i}(x_i - z_i) f_{X_i}(x_i) dx_i = \\
 &= \int_{a_i+z_i}^{b_i+z_i} \frac{1}{b_i - a_i} \frac{1}{\sqrt{2\pi} \sigma_i} e^{-\frac{1}{2} \left\{ \frac{x_i - \mu_i}{\sigma_i} \right\}^2} dx_i = \\
 &= \frac{1}{b_i - a_i} \frac{1}{\sqrt{2\pi}} \int_{(a_i+z_i - \mu_i)/\sigma_i}^{(b_i+z_i - \mu_i)/\sigma_i} e^{-t^2/2} dt.
 \end{aligned}$$

Portanto, Z_1, Z_2, \dots, Z_n é uma sequência de variáveis aleatórias independentes com a função densidade

$$f_{Z_i}(z_i) = \frac{1}{b_i - a_i} \frac{1}{\sqrt{2\pi}} \int_{(a_i+z_i - \mu_i)/\sigma_i}^{(b_i+z_i - \mu_i)/\sigma_i} e^{-t^2/2} dt, \quad (2.1.6)$$

$$-\infty < z_i < \infty.$$

Tendo a densidade de Z_i , queremos obter a densidade da variável aleatória W_i , onde $W_i = Z_i^2$.

A função densidade de W_i , para os valores de $w_i \geq 0$ é

$$\begin{aligned}
 f_{W_i}(w_i) &= \frac{1}{2\sqrt{w_i}} \{f_{Z_i}(\sqrt{w_i}) + f_{Z_i}(-\sqrt{w_i})\} = \\
 &= \frac{1}{2\sqrt{w_i}} \frac{1}{b_i - a_i} \frac{1}{\sqrt{2\pi}} \int_{(a_i - \mu_i + \sqrt{w_i})/\sigma_i}^{(b_i - \mu_i + \sqrt{w_i})/\sigma_i} e^{-t^2/2} dt + \\
 &+ \int_{(a_i - \mu_i - \sqrt{w_i})/\sigma_i}^{(b_i - \mu_i - \sqrt{w_i})/\sigma_i} e^{-t^2/2} dt. \tag{2.1.7}
 \end{aligned}$$

A média e a variância da variável aleatória w_i foram calculadas e são respectivamente

$$\mathbb{E}(W_i) = \mathbb{E}(X_i - Y_i)^2 = \mathbb{E}X_i^2 - 2\mathbb{E}(X_i)\mathbb{E}(Y_i) + \mathbb{E}(Y_i^2).$$

Utilizando os resultados (2.1.2) e (2.1.4), temos que

$$\mathbb{E}(W_i) = \sigma_i^2 + \mu_i^2 - \mu_i(a_i + b_i) + \frac{b_i^2 + a_i b_i + a_i^2}{3} \tag{2.1.8}$$

$$\begin{aligned}
IV(W_i) &= IV(X_i - Y_i)^2 = \mathbb{E}(X_i - Y_i)^4 - \mathbb{E}^2(X_i - Y_i)^2 = \\
&= \mathbb{E}(X_i^4) - 4\mathbb{E}(X_i)\mathbb{E}(Y_i^3) + 6\mathbb{E}(X_i^2)\mathbb{E}(Y_i^2) - \\
&\quad - 4\mathbb{E}(X_i^3)\mathbb{E}(Y_i) + \mathbb{E}(Y_i^4) - \mathbb{E}^2(X_i - Y_i)^2.
\end{aligned}$$

Novamente, utilizando os resultados (2.1.2) e (2.1.4) mais os resultados (2.1.8) temos que

$$\begin{aligned}
IV(W_i) &= 4\mu_i^2\sigma_i^2 + 2\sigma_i^4 - \frac{1}{3}\mu_i b_i^3 + \frac{1}{3}\mu_i a_i b_i^2 + \frac{1}{3}\mu_i a_i^2 b_i \\
&\quad - \frac{1}{3}\mu_i a_i^3 + \frac{4}{3}\sigma_i^2 b_i^2 + \frac{4}{3}\sigma_i^2 a_i b_i + \frac{4}{3}\sigma_i^2 a_i^2 + \frac{1}{3}\mu_i^2 b_i^2 \\
&\quad - \frac{2}{3}\mu_i^2 a_i b_i + \frac{1}{3}\mu_i^2 a_i^2 - 4\sigma_i^2 \mu_i a_i - 4\sigma_i^2 \mu_i b_i + \frac{4}{45}b_i^4 \\
&\quad - \frac{1}{45}b_i^3 a_i - \frac{2}{15}a_i^2 b_i^2 - \frac{1}{45}a_i^3 b_i + \frac{4}{45}a_i^4 \tag{2.1.9}
\end{aligned}$$

Agora, queremos estudar uma distribuição que meça a distância das variáveis normais e uniformes, seja esta medida D_n a distância euclidiana. $V_n = \sum_{i=1}^n W_i = \sum_{i=1}^n (X_i - Y_i)^2$, W_i independentes.

Então

$$D_n = \sqrt{V_n} = \sqrt{\sum_{i=1}^n W_i} = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2} \quad (2.1.10)$$

LEMA 2.1.1. A distribuição da variável aleatória V_n , soma das diferenças ao quadrado entre variáveis normais e uniformes é a seguinte:

$$f_{V_n}(v_n) = \begin{cases} \frac{1}{2^2} \pi^{-n/2} \prod_{i=1}^n (b_i - a_i)^{-1} K_{(n)} & , \\ \text{para } v_n \geq v_{n-1} \geq \dots \geq v_1 > 0 & (2.1.11) \\ 0 & , \text{ caso contrário} \end{cases}$$

onde

$$K_{(n)} = \begin{cases} \int_0^{v_n} K_{(n)}(v_n - v_{n-1}) K_{(n-1)} dv_{n-1} & , \text{ para } n \geq 2 \\ K^{(1)}(v_1) & , \text{ para } n = 1. \end{cases}$$

e

$$K^{(i)}(v_i) = \frac{1}{\sqrt{v_i}} \left[\int_{(a_i - \mu_i + \sqrt{v_i})/\sigma_i}^{(b_i - \mu_i + \sqrt{v_i})/\sigma_i} e^{-t^2/t} dt + \int_{(a_i - \mu_i - \sqrt{v_i})/\sigma_i}^{(b_i - \mu_i - \sqrt{v_i})/\sigma_i} e^{-t^2/2} dt \right]$$

PROVA:

Para $n = 1$

$V_1 = W_1 = (X_1 - Y_1)^2$, temos a densidade por (2.1.7)

$$f_{V_1}(v_1) = \begin{cases} \frac{1}{2\sqrt{v_1}} \frac{1}{b_1 - a_1} \frac{1}{\sqrt{2\pi}} \left\{ \int_{(a_1 - \mu_1 + \sqrt{v_1})/\sigma_1}^{(b_1 - \mu_1 + \sqrt{v_1})/\sigma_1} e^{-t^2/2} dt + \int_{(a_1 - \mu_1 - \sqrt{v_1})/\sigma_1}^{(b_1 - \mu_1 - \sqrt{v_1})/\sigma_1} e^{-t^2/2} dt \right\}, & \text{para } v_1 > 0 \\ 0, & \text{caso contrário} \end{cases} \quad (2.1.12)$$

Esta função densidade pode também ser escrita como:

$$f_{V_1}(v_1) = \frac{1}{2^{3/2} (b_1 - a_1)} \cdot \frac{1}{\pi^{1/2}} K_{(1)} \quad (2.1.13)$$

onde

$$K_{(1)} = K^{(1)}(v_1) = \frac{1}{\sqrt{v_1}} \left[\int_{(a_1 - \mu_1 + \sqrt{v_1})/\sigma_1}^{(b_1 - \mu_1 + \sqrt{v_1})/\sigma_1} e^{-t^2/2} dt + \int_{(a_1 - \mu_1 - \sqrt{v_1})/\sigma_1}^{(b_1 - \mu_1 - \sqrt{v_1})/\sigma_1} e^{-t^2/2} dt \right]$$

Para $n = 2$

$V_2 = W_1 + W_2$, W_1 e W_2 são independentes.

A função densidade da variável aleatória V_2 é por convolução:

$$\begin{aligned}
 f_{V_2}(v_2) &= \int_{-\infty}^{\infty} f_{W_2}(v_2 - v_1) f_{V_1}(v_1) dv_1, \quad 0 \leq v_2 - v_1 < \infty \implies v_1 < v_2 \\
 &= \int_0^{v_2} f_{W_2}(v_2 - v_1) f_{V_1}(v_1) dv_1 = \int_0^{v_2} \frac{1}{2\sqrt{v_2 - v_1}} \frac{1}{b_2 - a_2} \frac{1}{\sqrt{2\pi}} \times \\
 &\quad \times \left\{ \int_{(a_2 - \mu_2 + \sqrt{v_2 - v_1})/\sigma_2}^{(b_2 - \mu_2 + \sqrt{v_2 - v_1})/\sigma_2} e^{-t^2/2} dt + \int_{(a_2 - \mu_2 - \sqrt{v_2 - v_1})/\sigma_2}^{(b_2 - \mu_2 - \sqrt{v_2 - v_1})/\sigma_2} e^{-t^2/2} dt \right\} \times \\
 &\quad \times \frac{1}{2\sqrt{v_1}} \frac{1}{b_1 - a_1} \frac{1}{\sqrt{2\pi}} \left\{ \int_{(a_1 - \mu_1 + \sqrt{v_1})/\sigma_1}^{(b_1 - \mu_1 + \sqrt{v_1})/\sigma_1} e^{-t^2/2} dt + \right. \\
 &\quad \left. + \int_{(a_1 - \mu_1 - \sqrt{v_1})/\sigma_1}^{(b_1 - \mu_1 - \sqrt{v_1})/\sigma_1} e^{-t^2/2} dt \right\} dv_1.
 \end{aligned}$$

Portanto, a função densidade da variável aleatória V_2 é

$$\begin{aligned}
 f_{V_2}(v_2) = & \left\{ \frac{1}{(b_2 - a_2)(b_1 - a_1) 2^2 (2\pi)} \int_0^{v_2} \frac{1}{\sqrt{v_2 - v_1}} \left\{ \int_{(a_2 - \mu_2 + \sqrt{v_2 - v_1})/\sigma_2}^{(b_2 - \mu_2 + \sqrt{v_2 - v_1})/\sigma_2} e^{-t^2/2} dt \right. \right. \\
 & + \left. \int_{(a_2 - \mu_2 - \sqrt{v_2 - v_1})/\sigma_2}^{(b_2 - \mu_2 - \sqrt{v_2 - v_1})/\sigma_2} e^{-t^2/2} dt \right\} \frac{1}{\sqrt{v_1}} \left\{ \int_{(a_1 - \mu_1 + \sqrt{v_1})/\sigma_1}^{(b_1 - \mu_1 + \sqrt{v_1})/\sigma_1} e^{-t^2/2} dt + \right. \\
 & \left. \left. + \int_{(a_1 - \mu_1 - \sqrt{v_1})/\sigma_1}^{(b_1 - \mu_1 - \sqrt{v_1})/\sigma_1} e^{-t^2/2} dt \right\} dv_1, \text{ para } v_2 \geq v_1 > 0 \right. \\
 & \left. 0, \text{ caso contrário.} \right.
 \end{aligned} \tag{2.1.14}$$

Esta função (2.1.14) também pode ser escrita da seguinte maneira

$$f_{V_2}(v_2) = \frac{1}{2^3 (b_2 - a_2)(b_1 - a_1)} \pi^{-1} \cdot K(2) \tag{2.1.15}$$

onde

$$K(2) = \int_0^{v_2} K^{(2)}(v_2 - v_1) K^{(1)} dv_1$$

e

$$K^{(2)}(v_2 - v_1) = \frac{1}{\sqrt{v_2 - v_1}} \left[\int_{(a_2 - \mu_2 + \sqrt{v_2 - v_1})/\sigma_2}^{(b_2 - \mu_2 + \sqrt{v_2 - v_1})/\sigma_2} e^{-t^2/2} dt + \right. \\ \left. + \int_{(a_2 - \mu_2 - \sqrt{v_2 - v_1})/\sigma_2}^{(b_2 - \mu_2 - \sqrt{v_2 - v_1})/\sigma_2} e^{-t^2/2} dt \right].$$

Suponhamos que temos a validade da função para $n-1$, isto é,

$$V_{n-1} = W_1 + W_2 + \dots + W_{n-1}.$$

$$f_{V_{n-1}}(v_{n-1}) = \begin{cases} \frac{1}{2} \pi^{-\frac{(n-1)}{2}} \prod_{i=1}^{n-1} (b_i - a_i)^{-1} K_{(n-1)}, & \text{para } v_{n-1} \geq v_{n-2} \geq \dots \geq v_1 > 0 \\ 0, & \text{caso contrário.} \end{cases} \quad (2.1.16)$$

onde

$$K_{(n-1)} = \int_0^{v_{n-1}} K^{(n-1)}(v_{n-1} - v_{n-2}) K_{(n-2)} dv_{n-2}.$$

e

$$K^{(n-1)}(v_{n-1}-v_{n-2}) = \frac{1}{\sqrt{v_{n-1}-v_{n-2}}} \left[\int \frac{(b_{n-1}-\mu_{n-1}+\sqrt{v_{n-1}-v_{n-2}})/\sigma_{n-1}}{(a_{n-1}-\mu_{n-1}+\sqrt{v_{n-1}-v_{n-2}})/\sigma_{n-1}} e^{-t^2/2} dt \right. \\ \left. + \int \frac{(b_{n-1}-\mu_{n-1}-\sqrt{v_{n-1}-v_{n-2}})/\sigma_{n-1}}{(a_{n-1}-\mu_{n-1}-\sqrt{v_{n-1}-v_{n-2}})/\sigma_{n-1}} e^{-t^2/2} dt \right]$$

Então para n , temos $V_n = W_1 + W_2 + \dots + W_{n-1} + W_n =$
 $= V_{n-1} + W_n$.

A função densidade da variável V_n é

$$f_{V_n}(v_n) = \int_{-\infty}^{\infty} f_{W_n}(v_n-v_{n-1}) f_{V_{n-1}}(v_{n-1}) dv_{n-1}, \quad 0 \leq v_n - v_{n-1} < \infty \\ \Rightarrow v_{n-1} < v_n$$

$$= \int_0^{v_n} f_{W_n}(v_n-v_{n-1}) f_{V_{n-1}}(v_{n-1}) dv_{n-1} =$$

$$= \int_0^{v_n} \frac{1}{2\sqrt{v_n-v_{n-1}}} \frac{1}{b_n-a_n} \frac{1}{\sqrt{2\pi}} \left\{ \frac{(b_n-\mu_n+\sqrt{v_n-v_{n-1}})/\sigma_n}{(a_n-\mu_n+\sqrt{v_n-v_{n-1}})/\sigma_n} e^{-t^2/2} dt \right.$$

$$+ \int \left(\frac{(b_n - \mu_n - \sqrt{v_n - v_{n-1}})/\sigma_n}{(a_n - \mu_n - \sqrt{v_n - v_{n-1}})/\sigma_n} e^{-t^2/2} dt \right) \frac{1}{2 \frac{3(n-1)}{2}} \pi^{\frac{-(n-1)}{2}} \prod_{i=1}^{n-1} (b_i - a_i)^{-1}$$

$$\times K_{(n-1)} dv_{n-1} =$$

$$= \frac{1}{2 \frac{3n}{2}} \prod_{i=1}^n (b_i - a_i)^{-1} \pi^{-n/2} \int_0^{v_n} \frac{1}{\sqrt{v_n - v_{n-1}}} \left[\int \frac{(b_n - \mu_n + \sqrt{v_n - v_{n-1}})/\sigma_n}{(a_n - \mu_n + \sqrt{v_n - v_{n-1}})/\sigma_n} e^{-t^2/2} dt \right.$$

$$\left. + \int \frac{(b_n - \mu_n - \sqrt{v_n - v_{n-1}})/\sigma_n}{(a_n - \mu_n - \sqrt{v_n - v_{n-1}})/\sigma_n} e^{-t^2/2} dt \right] K_{(n-1)} dv_{n-1} =$$

$$= \frac{1}{2 \frac{3n}{2}} \prod_{i=1}^n (b_i - a_i)^{-1} \pi^{-n/2} K_{(n)} \quad (2.1.16)$$

onde

$$K_{(n)} = \int_0^{v_n} K^{(n)}(v_n - v_{n-1}) K_{(n-1)} dv_{n-1}$$

e

$$K^{(n)}(v_n - v_{n-1}) = \frac{1}{\sqrt{v_n - v_{n-1}}} \left[\int_{(a_n - \mu_n + \sqrt{v_n - v_{n-1}})/\sigma_n}^{(b_n - \mu_n + \sqrt{v_n + v_{n-1}})/\sigma_n} e^{-t^2/2} dt + \right. \\ \left. + \int_{(a_n - \mu_n - \sqrt{v_n - v_{n-1}})/\sigma_n}^{(b_n - \mu_n - \sqrt{v_n + v_{n-1}})/\sigma_n} e^{-t^2/2} dt \right]$$

Com isto, a fórmula de recorrência está provada.

Pensando em D_n (2.1.10) como $D_n = \sqrt{V_n}$, a densidade desta variável é obtida:

$$f_{D_n}(d_n) = \begin{cases} 2d_n f_{V_n}(d_n^2) & , d_n \geq 0 \\ 0 & , \text{ caso contrário} \end{cases} \quad (2.1.17)$$

\therefore

$$f_{D_n}(d_n) = \frac{d_n}{2^{n-1}} (2\pi)^{-n/2} \prod_{i=1}^n (b_i - a_i)^{-1} K^{(n)} \quad (2.1.18)$$

onde

$$K^{(n)} = \begin{cases} K^{(1)}(d_1^2) & , \text{ para } n = 1 \\ 2 \int_0^{d_n^2} K^{(n)}(d_n^2 - d_{n-1}^2) K^{(n-1)} d_{n-1} dd_{n-1} & \text{ para } n \geq 2 \end{cases}$$

e

$$K^{(i)}(d_i^2) = \frac{1}{d_i} \left[\int_{(a_i - \mu_i + d_i)/\sigma_i}^{(b_i - \mu_i + d_i)/\sigma_i} e^{-t^2/2} dt + \int_{(a_i - \mu_i - d_i)/\sigma_i}^{(b_i - \mu_i - d_i)/\sigma_i} e^{-t^2/2} dt \right],$$

$$\text{para } d_n^2 \geq d_{n-1}^2 \geq \dots \geq d_1^2 > 0.$$

O desenvolvimento anterior foi feito para calcular a distância de um ponto normal n-dimensional em relação a um ponto contido em uma das classes uniforme n-dimensional. Mas é de nosso interesse calcular a distância desse ponto, em relação a vários pontos uniformes n-dimensionais. Onde cada ponto uniforme esta contido num hipercubo que representa uma classe. A seguir estendemos para o caso de multiclasse.

Seja a variável aleatória (X_1, X_2, \dots, X_n) com uma distribuição normal n-dimensional. Então a sua densidade $f(x_1, x_2, \dots, x_n)$ é da forma:

$$f_{\mathbf{X}}(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} \sqrt{|\mathbf{V}|}} \exp\left[-\frac{1}{2|\mathbf{V}|} \sum_{j,k=1}^n |\mathbf{V}_{jk}| (x_j - \mu_j)(x_k - \mu_k)\right], \quad (2.1.19)$$

$$-\infty < x_i < \infty,$$

onde $|\mathbf{V}| \neq 0$ é o determinante da matriz \mathbf{V} de momentos de segunda ordem σ_{jk} definidos por:

$$\mathbf{V} = \begin{vmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{vmatrix} \quad \text{e } \sigma_{ik} \text{ é a covariância} \\ \text{de } X_i \text{ e } X_k.$$

e $|\mathbf{V}_{jk}|$ é o complemento algébrico do termo σ_{ik} no determinante da matriz \mathbf{V} .

No problema apresentado aqui, supomos que as variáveis aleatórias X_1, X_2, \dots, X_n são independentes, portanto, $\sigma_{ik} = 0$ ($i \neq k$), isto é, X_1, X_2, \dots, X_n são não correlacionadas e $|\mathbf{V}|$ toma o seu valor máximo $|\mathbf{V}| = \sigma_1^2 \sigma_2^2 \dots \sigma_n^2$.

E a função densidade (2.1.19) fica como:

$$f_{\mathbf{X}}(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} \prod_{i=1}^n \sigma_i} e^{(-1/2 \sum_{i=1}^n (\frac{x_i - \mu_i}{\sigma_i})^2)} \quad (2.1.20)$$

para $-\infty < x_i < \infty$.

Seja $\mathbf{Y}^j = (y_1^j, y_2^j, \dots, y_n^j)$, $j=1, \dots, k$, indicando as k classes; um vetor aleatório formado por variáveis aleatórias uniformes independentes com densidade:

$$f_{Y^j}(y_1^j, y_2^j, \dots, y_n^j) = \begin{cases} \frac{1}{\prod_{i=1}^n (b_i^j - a_i^j)}, & a_i^j \leq y_i^j \leq b_i^j \\ 0 & , \text{ caso contrário} \end{cases} \quad (2.1.21)$$

Para a primeira apresentação de um vetor \mathbf{X} temos então o vetor aleatório \mathbf{D}^j definido como

$$D_n^j = \left(\sum_{m=1}^n (y_m^j - x_m)^2 \right)^{1/2}, \quad j=1, 2, \dots, k \quad (2.1.22)$$

E para a variável aleatória $D_n^j = \sqrt{V_n^j} = \sqrt{\sum_{i=1}^n W_i^j}$, a função densidade é

$$f_{D_n^j}(d_n^j) = \begin{cases} \frac{d_n^j}{2^{n-1}} (2\pi)^{-n/2} \prod_{i=1}^n (b_i^j - a_i^j)^{-1} \cdot \frac{1}{(n, d_n^j)}, & \\ \text{para } (d_n^j)^2 \geq (d_{n-1}^j)^2 \geq \dots \geq (d_1^j)^2 > 0 & \\ 0 & , \text{ caso contrário} \end{cases} \quad (2.1.23)$$

onde

$$K_{(n, d_n^j)}^j = \begin{cases} K^{(1)}(d_1^2) & , \text{ para } n = 1 \\ 2 \int_0^{(d_n^j)^2} K^{(n)} [(d_n^j)^2 - (d_{n-1}^j)^2] K_{(n-1, d_{n-1}^j)}^j d_{n-1}^j d_{n-1}^j & , \text{ para } n \geq 2 \end{cases}$$

e

$$K^{(i)}((d_i^j)^2) = \frac{1}{d_n^j} \left[\int_{(a_i^j - \mu_i + d_i^j)/\sigma_i}^{(b_i^j - \mu_i + d_i^j)/\sigma_i} e^{t^2/2} dt + \int_{(a_i^j - \mu_i - d_i^j)/\sigma_i}^{(b_i^j - \mu_i - d_i^j)/\sigma_i} e^{t^2/2} dt \right].$$

Afinal lembramos que é de nosso interesse encontrar a distribuição de β , uma variável aleatória que denominaremos $BETA^j$ e é definida como:

$$BETA^j = \frac{D_n^j}{\max_{1 \leq i \leq k} D_n^i}, \quad j = \text{um valor entre } 1 \text{ e } k \quad (2.1.24)$$

onde D_n^j é a variável aleatória cuja função densidade foi dada em (2.1.23).

Seja $(D_n^1, D_n^2, \dots, D_n^k)$ um vetor k -dimensional e seja $(d_n^1, d_n^2, \dots, d_n^k)$ valores assumidos por $(D_n^1, D_n^2, \dots, D_n^k)$. Denotamos $(d_n^{(1)}, d_n^{(2)}, \dots, d_n^{(k)})$ a estatística de ordem dos valores d_n^j que a variável aleatória $D_n^{(j)}$ pode assumir.

Então a variável aleatória $BETA^j$ pode ser escrita como

$$BETA^j = \frac{D_n^{(j)}}{D_n^{(k)}}, \quad j=1, 2, \dots, k. \quad (2.1.25)$$

A função densidade conjunta da i -ésima e da j -ésima estatística de ordem $D_n^{(i)}$ e $D_n^{(j)}$ ($1 \leq i \leq j \leq k$) é

$$\begin{aligned} g_{ij}(d_n^{(i)}, d_n^{(j)}) &= \frac{k!}{(i-1)!(j-i-1)!(k-j)!} [F_{D_n^i}^i(d_n^{(i)})]^{i-1} \times \\ &\times [F_{D_n^j}^j(d_n^{(j)}) - F_{D_n^i}^i(d_n^{(i)})]^{j-i-1} [1 - F_{D_n^j}^j(d_n^{(j)})]^{k-j} \times \\ &\times f_{D_n^i}^i(d_n^{(i)}) f_{D_n^j}^j(d_n^{(j)}), \quad \text{para } d_n^{(i)} < d_n^{(j)} \end{aligned} \quad (2.1.26)$$

A função densidade da variável aleatória para $0 \leq d_n^{(k)} < \infty$ é:

$$\begin{aligned}
f_{\text{BETA}}^j(\text{beta}^j) &= \int_{-\infty}^{\infty} g_{ik}(\text{beta}^j, d_n^{(k)}) |d_n^{(k)}| d d_n^{(k)} = \\
&= \int_0^{\infty} \frac{k!}{(i-1)!(k-i-1)!} [F_{D_n^j}(\text{beta}^j, d_n^{(k)})]^{i-1} [F_{D_n^k}(d_n^{(k)}) \\
&\quad - F_{D_n^j}(\text{beta}^j, d_n^{(k)})]^{k-i-1} \cdot f_{D_n^j}(\text{beta}^j, d_n^{(k)}) \cdot \\
&\quad \cdot f_{D_n^k}(d_n^{(k)}) d_n^{(k)} d d_n^{(k)} \tag{2.1.27}
\end{aligned}$$

Fazendo as substituições com resultados obtidos em (2.1.23)

temos:

$$\begin{aligned}
f_{\text{BETA}}^j(\text{beta}^j) &= \left\{ \int_0^{\infty} \frac{k!}{(i-1)!(k-i-1)!} \int_{-\infty}^{\text{beta}^j d_n^{(k)}} \frac{t_n^j}{2^{n-1}} (2\pi)^{-n/2} \times \right. \\
&\quad \times \prod_{i=1}^n (b_i^j - a_i^j)^{-1} K_{(n, t_n^j)}^j d t_n^j]^{i-1} \times \\
&\quad \left. \times \int_{-\infty}^{d_n^{(k)}} \frac{t_n^k}{2^{n-1}} (2\pi)^{-n/2} \prod_{i=1}^n (b_i^k - a_i^k)^{-1} K_{(n, t_n^k)}^k d t_n^k \right. \\
&\tag{2.1.28}
\end{aligned}$$

$$\left. \begin{aligned}
& - \int_{-\infty}^{\beta_n^j} \frac{\beta_n^j d_n^{(k)}}{2^{n-1}} (2\pi)^{-n/2} \prod_{i=1}^n (b_i^j - a_i^j)^{-1} K_n^j(n, t_n^j) dt_n^j \times \\
& \times \frac{\beta_n^j d_n^{(k)}}{2^{n-1}} (2\pi)^{-n/2} \prod_{i=1}^n (b_i^j - a_i^j)^{-1} K_n^j(n, \beta_n^j d_n^{(k)}) \times \\
& \times \frac{d_n^{(k)}}{2^{n-1}} (2\pi)^{-n/2} \prod_{i=1}^n (b_i^k - a_i^k)^{-1} K_n^k(n, d_n^{(k)}) d_n^{(k)} dd_n^{(k)}, \\
& (d_n^j)^2 \geq (d_{n-1}^j)^2 \geq \dots \geq (d_1^j)^2 > 0 \\
& 0, \quad \text{caso contrário}
\end{aligned} \right\}$$

onde

$$K_n^j(n, h_n^m) = \int_0^{(h_n^m)^2} 2K^{(n)} [(h_n^m)^2 - (h_{n-1}^m)^2] K^{(n-1)}(n-1, h_{n-1}^m) (h_{n-1}^m) d(h_{n-1}^m)$$

e

$$K^{(i)}((h_i^m)^2) = \frac{1}{h_i^m} \left[\int_{(a_i^m - \mu_i + h_i^m)/\sigma_i}^{(b_i^m - \mu_i + h_i^m)/\sigma_i} e^{-t^2/2} dt + \int_{(a_i^m - \mu_i - h_i^m)/\sigma_i}^{(b_i^m - \mu_i - h_i^m)/\sigma_i} e^{-t^2/2} dt \right]$$

que descreve a distribuição exata da variável aleatória β .

2.2. DISTRIBUIÇÃO DA ESTATÍSTICA DE ORDEM ASSINTÓTICA.

Para podermos tratar não somente o primeiro caso de classificação de um vetor X , mais também os seguintes para $\ell = 2, 3, \dots, N$, supomos agora que para um k grande, $D_{n\ell}^j$ tem aproximadamente uma distribuição normal, com média μ_ℓ e variância σ_ℓ .

Então a função densidade de Beta_ℓ^j é:

$$\begin{aligned}
 f_{\text{BETA}_\ell^j}(\text{beta}_\ell^j) &= \int_0^\infty \frac{k!}{(i-1)!(k-i-1)!} \left[\int_{-\infty}^{\text{beta}_{n\ell}^j(k)} \frac{1}{\sqrt{2\pi} \sigma_\ell} e^{-((t_{n\ell}^j - \mu_\ell)/\sigma_\ell)^2/2} dt_{n\ell}^j \right]^{i-1} \\
 &\times \left[\int_{-\infty}^{d_{n\ell}^{(k)}} \frac{1}{\sqrt{2\pi} \sigma_\ell} e^{-((t_{n\ell}^j - \mu_\ell)/\sigma_\ell)^2/2} dt_{n\ell}^j - \int_{-\infty}^{\text{beta}_{n\ell}^j(k)} \frac{1}{\sqrt{2\pi} \sigma_\ell} e^{-((t_{n\ell}^j - \mu_\ell)/\sigma_\ell)^2/2} dt_{n\ell}^j \right]^{k-i-1} \\
 &\times \frac{1}{\sqrt{2\pi} \sigma_\ell} e^{-((\text{beta}_{n\ell}^j(k) - \mu_\ell)/\sigma_\ell)^2/2} \frac{1}{\sqrt{2\pi} \sigma_\ell} e^{-((d_{n\ell}^{(k)} - \mu_\ell)/\sigma_\ell)^2/2} d_{n\ell}^{(k)} dd_{n\ell}^{(k)} = \\
 &= \frac{k!}{(i-1)!(k-i-1)!} \int_0^\infty \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(\text{beta}_{n\ell}^j(k) - \mu_\ell)/\sigma_\ell} e^{-t^2/2} dt \right]^{i-1} \\
 &\times \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(d_{n\ell}^{(k)} - \mu_\ell)/\sigma_\ell} e^{-t^2/2} dt - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(\text{beta}_{n\ell}^j(k) - \mu_\ell)/\sigma_\ell} e^{-t^2/2} dt \right]^{k-i-1}
 \end{aligned}$$

$$\times \frac{1}{\sqrt{2\pi} \sigma_\ell} e^{-((\text{beta}_{\ell n\ell}^{j_d(k)} - \mu_\ell)/\sigma_\ell)^2/2} \frac{1}{\sqrt{2\pi} \sigma_\ell} e^{-((d_{n\ell}^{(k)} - \mu_\ell)/\sigma_\ell)^2/2} d_{n\ell}^{(k)} dd_{n\ell}^{(k)} =$$

$$= \frac{k!}{(i-1)!(k-i-1)!} \int_0^\infty \left[\frac{1}{\sqrt{2\pi}} \int_{(\mu_\ell - \text{beta}_{\ell n\ell}^{j_d(k)})/\sigma_\ell}^\infty e^{-t^2/2} dt \right]^{i-1}$$

$$\times \left[\frac{1}{\sqrt{2\pi}} \int_{(\mu_\ell - d_{n\ell}^{(k)})/\sigma_\ell}^\infty e^{-t^2/2} dt - \frac{1}{\sqrt{2\pi}} \int_{(\mu_\ell - \text{beta}_{\ell n\ell}^{j_d(k)})/\sigma_\ell}^\infty e^{-t^2/2} dt \right]^{k-i-1}$$

$$\times \frac{1}{\sqrt{2\pi} \sigma_\ell} e^{-((\text{beta}_{\ell n\ell}^{j_d(k)} - \mu_\ell)/\sigma_\ell)^2/2} \frac{1}{\sqrt{2\pi} \sigma_\ell} e^{-((d_{n\ell}^{(k)} - \mu_\ell)/\sigma_\ell)^2/2} d_{n\ell}^{(k)} dd_{n\ell}^{(k)}$$

(2.2.1)

fazendo a mudança $\frac{1}{\sqrt{2}} t = s$, temos que

$$= \frac{k!}{(i-1)!(k-i-1)!} \int_0^\infty \left[\frac{1}{\sqrt{\pi}} \int_{(\mu_\ell - \text{beta}_{\ell n\ell}^{j_d(k)})/(\sqrt{2} \sigma_\ell)}^{+\infty} e^{-s^2} ds \right]^{i-1}$$

$$\times \left[\frac{1}{\sqrt{\pi}} \int_{(\mu_\ell - d_{n\ell}^{(k)})/(\sqrt{2} \sigma_\ell)}^\infty e^{-s^2} ds - \frac{1}{\sqrt{\pi}} \int_{(\mu_\ell - \text{beta}_{\ell n\ell}^{j_d(k)})/(\sqrt{2} \sigma_\ell)}^\infty e^{-s^2} ds \right]^{k-i-1}$$

$$\begin{aligned}
& \times \frac{1}{\sqrt{2\pi} \sigma_\ell} e^{-((\text{beta}_{\ell \text{nl}}^{\text{j}_d(k)} - \mu_\ell)/\sigma_\ell)^2/2} \frac{1}{\sqrt{2\pi} \sigma_\ell} e^{-((d_{\text{nl}}^{(k)} - \mu_\ell)/\sigma_\ell)^2/2} d_{\text{nl}}^{(k)} dd_{\text{nl}}^{(k)} = \\
& = \frac{k!}{(i-1)!(k-i-1)!} \int_0^\infty \left[\frac{1}{2} \operatorname{erfc}\left(\frac{\mu_\ell - \text{beta}_{\ell \text{nl}}^{\text{j}_d(k)}}{\sqrt{2} \sigma_\ell}\right) \right]^{i-1} \times \\
& \times \left[\frac{1}{2} \operatorname{erfc}\left(\frac{\mu_\ell - d_{\text{nl}}^{(k)}}{\sqrt{2} \sigma_\ell}\right) - \frac{1}{2} \operatorname{erfc}\left(\frac{\mu_\ell - \text{beta}_{\ell \text{nl}}^{\text{j}_d(k)}}{\sqrt{2} \sigma_\ell}\right) \right]^{k-i-1} \\
& \times \frac{1}{\sqrt{2\pi} \sigma_\ell} e^{-((\text{beta}_{\ell \text{nl}}^{\text{j}_d(k)} - \mu_\ell)/\sigma_\ell)^2/2} \frac{1}{\sqrt{2\pi} \sigma_\ell} e^{-((d_{\text{nl}}^{(k)} - \mu_\ell)/\sigma_\ell)^2/2} d_{\text{nl}}^{(k)} dd_{\text{nl}}^{(k)} = \\
& = \frac{k!}{(i-1)!(k-i-1)!} \int_0^\infty \left[\frac{1}{2} - \frac{1}{\sqrt{\pi}} \operatorname{Erf}\left(\frac{\mu_\ell - \text{beta}_{\ell \text{nl}}^{\text{j}_d(k)}}{\sqrt{2} \sigma_\ell}\right) \right]^{i-1} \\
& \times \left[\frac{1}{2} - \frac{1}{\sqrt{\pi}} \operatorname{Erf}\left(\frac{\mu_\ell - d_{\text{nl}}^{(k)}}{\sqrt{2} \sigma_\ell}\right) - \frac{1}{2} + \frac{1}{\sqrt{\pi}} \operatorname{Erf}\left(\frac{\mu_\ell - \text{beta}_{\ell \text{nl}}^{\text{j}_d(k)}}{\sqrt{2} \sigma_\ell}\right) \right]^{k-i-1} \times \\
& \times \frac{1}{\sqrt{2\pi} \sigma_\ell} e^{-((\text{beta}_{\ell \text{nl}}^{\text{j}_d(k)} - \mu_\ell)/\sigma_\ell)^2/2} \frac{1}{\sqrt{2\pi} \sigma_\ell} e^{-((d_{\text{nl}}^{(k)} - \mu_\ell)/\sigma_\ell)^2/2} d_{\text{nl}}^{(k)} dd_{\text{nl}}^{(k)} .
\end{aligned}$$

(2.2.2)

Portanto, a função densidade de $BETA^j$ pode ser escrita como:

$$\begin{aligned}
 f_{BETA^j_\ell}(\beta_\ell^j) = & \frac{k!}{(i-1)!(k-i-1)! \sigma_\ell^2 2\pi} \int_0^\infty \left[\frac{1}{2} + \frac{1}{\sqrt{\pi}} \left(\frac{\beta_\ell^j d_{n\ell}^{(k)} - \mu_\ell}{\sqrt{2} \sigma_\ell} \right) \right] \times \\
 & \times {}_1F_1\left(\frac{1}{2}; \frac{3}{2}, -\left(\frac{\mu_\ell - \beta_\ell^j d_{n\ell}^{(k)}}{\sqrt{2} \sigma_\ell}\right)^2\right)^{i-1} \times \\
 & \times \left[\frac{1}{\sqrt{\pi}} \left(\frac{d_{n\ell}^{(k)} - \mu_\ell}{\sqrt{2} \sigma_\ell} \right) {}_1F_1\left(\frac{1}{2}; \frac{3}{2}, -\left(\frac{\mu_\ell - d_{n\ell}^{(k)}}{\sqrt{2} \sigma_\ell}\right)^2\right) + \right. \\
 & \left. + \frac{1}{\sqrt{\pi}} \left(\frac{\mu_\ell - \beta_\ell^j d_{n\ell}^{(k)}}{\sqrt{2} \sigma_\ell} \right) {}_1F_1\left(\frac{1}{2}; \frac{3}{2}, -\left(\frac{\mu_\ell - \beta_\ell^j d_{n\ell}^{(k)}}{\sqrt{2} \sigma_\ell}\right)^2\right) \right]^{k-i-1} \\
 & {}_1F_1\left(c; c, -\left(\frac{\beta_\ell^j d_{n\ell}^{(k)} - \mu_\ell}{\sigma_\ell \sqrt{2}}\right)^2\right) {}_1F_1\left(c; c, -\left(\frac{d_{n\ell}^{(k)} - \mu_\ell}{\sqrt{2} \sigma_\ell}\right)^2\right) \\
 & \times d_{n\ell}^{(k)} d_{n\ell}^{(k)} \tag{2.2.3} \\
 & \text{para } c \neq 0 \quad \text{e} \quad 0 \leq d_{n\ell}^{(k)} < \infty \\
 & 0 \quad , \quad \text{caso contrário.}
 \end{aligned}$$

onde ${}_1F_1$ é a função hipergeométrica definida em Luke (1969) como:

$${}_1F_1(a; c, z) = \sum_{k=0}^{\infty} \frac{(a)_k z^k}{(c)_k k!} \quad (2.2.4)$$

a e c quaisquer valores
exceto $c = 0$.

$$e \quad (a)_k = a(a+1) \dots (a+k-1),$$

$$(a)_0 = 1,$$

$$e^z = {}_1F_1(a; a, z).$$

Erf e Erfc são respectivamente função erro e função erro complementar definidas como:

$$\text{Erf}(z) = \int_0^z e^{-t^2} dt = \frac{\sqrt{\pi}}{2} \text{erf}(z) \quad (2.2.5)$$

$$\text{Erfc}(z) = \int_z^{\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{2} \text{erfc}(z) \quad (2.2.6)$$

$$\text{erfc}(z) + \text{erf}(z) = 1.$$

A função erro pode ser escrita como uma função hipergeométrica do seguinte modo:

$$\text{Erf}(z) = z {}_1F_1\left(\frac{1}{2}; \frac{3}{2}, -z^2\right) \quad (2.2.7)$$

Chegamos a uma distribuição exata da variável aleatória β na subseção 2.1. E como vemos por (2.1.28), ela envolve fórmulas recursivas, que mesmo quando conhecemos os fatores e os termos substituindo na fórmula, ainda teremos resultados de formas complexas.

A fim de se evitar essas fórmulas recursivas, usamos o argumento de que a variável distância tem distribuição normal assintoticamente e obtivemos o resultado (2.2.3). Mas, ainda este caminho simplificado nos levou a utilização de recursos como as funções especiais e o resultado permaneceu difícil de manusear e segundo Rathie e Rohrer (1978), até seu desenvolvimento numérico é de extrema complexidade.

Para contornar o problema da intratabilidade das distribuições resultantes do modelo (padrão normal, classe uniforme) assumido em (2.1.28), propusemos um enfoque paramétrico; trata-se de simular a variável β e comparar a distribuição observada com uma distribuição familiar encontrando as condições necessárias para um bom ajuste.

CAPÍTULO III

DESENVOLVIMENTO POR SIMULAÇÃO

3.1. PROCEDIMENTO DA SIMULAÇÃO.

A simulação envolveu as seguintes etapas:

1. Partimos de um espaço \mathbb{R}^n ($n \geq 1$), sobre ele uma normal n-variado com parâmetros μ e Σ identificada como a dos padrões. Espaço este contendo k ($k \geq 1$) classes C_k inicialmente uniformes com os parâmetros $((a_1^k, b_1^k), \dots, (a_n^k, b_n^k))$.
2. Para iniciar a simulação, em cada classe C_j , escolhemos aleatoriamente segundo a distribuição $U(a_j, b_j)$ um ponto $Y_j^1 = (Y_1^{j1}, Y_2^{j1}, \dots, Y_n^{j1})$. Denominaremos este ponto de "referente" R_j^1 da classe C_j .
3. Geramos aleatoriamente $\mathbf{X}_m = (X_{m1}, X_{m2}, \dots, X_{mn})$, um vetor com distribuição normal (μ, Σ) que envolve os seguintes passos:
 - 3.1. Queremos um vetor de dimensão n , mas inicialmente geramos (z_1, z_2) um par de números de $(0, 1)^2$ com distribuição uniforme $U(0, 1)$. Fazemos:

$$y_j = 2z_j - 1, \quad j = 1, 2 \quad (3.1.1)$$

$$s = Y_1^2 + Y_2^2 \quad (3.1.2)$$

$$\text{Para } s \leq 1, \quad z_{m\ell} = y_j \cdot \sqrt{\frac{-2 \ln s}{s}}, \quad \ell = \ell + j, \quad j = 1, 2, \quad \ell = 0, \dots, \lfloor \frac{n}{2} \rfloor \quad (3.1.3)$$

Isto é, geramos dois números normalmente distribuídos e repetimos esse processo de geração até obtermos no total n números, ou seja $\ell = \lfloor \frac{n}{2} \rfloor$. Juntamos todos esses números no vetor $\mathbf{z}_m = (z_{m1}, z_{m2}, \dots, z_{mn})$ que deste modo tem distribuição $N(\mathbf{0}, I)$. Fazendo a nossa matriz de covariância I fica fácil gerar amostras normalmente distribuídas, porque as variáveis são independentes e identicamente distribuídas com variância 1.

3.2. Para obtermos \mathbf{x}_m , um vetor com distribuição normal com média μ e matriz de covariância Σ , usamos a seguinte transformação (ver por exemplo Fukunaga (1972))

$$\mathbf{x}_m = \mathbf{Q}\mathbf{L}^{1/2}\mathbf{z}_m + \mu, \quad \text{onde } \mathbf{Q} = \text{matriz de autovetores de } \Sigma \quad (3.1.4)$$

$\mathbf{L} = \text{matriz diagonal de autovalores } \lambda_m \text{ de } \Sigma.$

4. Nesta etapa, verificamos se o ponto gerado pertence a alguma das classes C_j . Se o ponto não pertencer a nenhuma classe, voltamos a etapa 3 e procedemos até localizarmos \mathbf{x}_m , um ponto que pertence a alguma das k classes.

5. Calculamos as distâncias do ponto gerado com relação a todas as k classes, ou seja, as distâncias do vetor \mathbf{X}_i do referente $\mathbf{R}_j^m = (Y_1^{jm}, \dots, Y_n^{jm})$, $m \geq 1$

$$D_m^j = \left(\sum_{i=1}^n (Y_i^{jm} - X_{mi})^2 \right)^{1/2}, \quad j=1,2,\dots,k \quad (3.1.5)$$

Diremos que $\mathbf{X}_m \in C_j$ se a distância euclidiana entre \mathbf{X}_m e \mathbf{R}_j^m for a mínima entre todas as distâncias de \mathbf{X}_m e \mathbf{R}_j^m , $j=1,\dots,k$. E verificamos sobre o acerto, se realmente o ponto pertence a classe que resultou na menor distância do referente com o ponto. Senão, $j = j'$, onde j é índice da classe correta.

6. Como o referente tem que ser aprendido, efetuamos uma correção sobre ele a fim de que numa próxima etapa produza a menor distância do ponto aleatório, se realmente \mathbf{X}_m pertence a esta classe C_j . Então

$$\mathbf{R}_j^{m+1} = \beta_m \mathbf{X}_m + (1 - \beta_m) \mathbf{R}_j^m, \quad (3.1.6)$$

onde

$$\beta_m = \frac{D_m^j}{\max_{1 \leq j \leq k} D_m^j} \quad (3.1.7)$$

7. Voltamos a etapa 3 e repetimos esse processo até $m=N$, o tamanho da nossa amostra. Portanto obtemos o vetor amostral $(\beta_1, \beta_2, \dots, \beta_N)$.

8. A fim de testarmos o ajuste dos valores obtidos com o modelo Beta (p,q) proposto, utilizamos o teste χ^2 .

8.1. Fazemos o histograma da variável β_i , dividindo em r intervalos. Calculamos a frequência absoluta $f_a(\ell)$ das variáveis β_i em cada intervalo $(L(\ell), S(\ell)]$, $\ell = 1, \dots, r$.

$$f_a(\ell) = \sum_{i=1}^N \chi_{A(\ell)}(\beta_i) , \quad (3.1.8)$$

onde $A(\ell) = [L(\ell), S(\ell)]$, (obviamente $A(\ell) \subset (0,1) \forall \ell$)

$$\chi_{A(\ell)}(\beta_i) = \begin{cases} 1 & \text{se } \beta_i \in A(\ell) \\ 0 & \text{se } \beta_i \notin A(\ell). \end{cases}$$

8.2. Calculamos a frequência estimada $f_e(\ell)$, segundo a distribuição Beta (p,q) que propusemos.

8.2.1. Estimamos os parâmetros p e q da distribuição Beta, pelo método de máxima verossimilhança que descreveremos na subseção 3.2.

8.2.2. Calculamos a probabilidade

$$\hat{p}(\ell) = \int_{L(\ell)}^{S(\ell)} \frac{\Gamma(\hat{p})\Gamma(\hat{q})}{\Gamma(\hat{p} + \hat{q})} t^{\hat{p}-1} (1-t)^{\hat{q}-1} dt \quad (3.1.9)$$

Onde a integral

$$I(\ell) = \int_{L(\ell)}^{S(\ell)} t^{\hat{p}-1} (1-t)^{\hat{q}-1} dt \quad (3.1.10)$$

é resolvida usando o método de Simpson com 60 subintervalos. Como a função tem derivadas contínuas de quarta ordem, segundo Bakhavalov (1975), a regra de Simpson converge para o verdadeiro valor da integral, no pior caso com velocidade k^{-4} , onde k é o número de subintervalos utilizados.

8.2.3. A frequência estimada é obtida por:

$$f_e(\ell) = N\hat{p}(\ell) = N \times \frac{I(\ell)}{\sum_{i=1}^r \int_{L(i)}^{S(i)} t^{\hat{p}-1} (1-t)^{\hat{q}-1} dt}, \quad \ell = 1, \dots, r \quad (3.1.11)$$

8.3. Estimados os parâmetros p e q através das amostras, queremos verificar se a característica β_ℓ tem uma distribuição $\beta(\hat{p}, \hat{q})$.

Então finalmente aplicamos o teste χ^2 ,

$$\chi^2 = \sum_{\ell=1}^r \frac{(f_a(\ell) - f_e(\ell))^2}{f_e(\ell)} \quad (3.1.12)$$

O modelo Beta(p,q) será considerado satisfatório se houver evidência de ajuste, ou seja, se $\chi_{r-3}^2 < \chi_{\alpha, r-3}^2$ para α pré-fixado.

9. Voltamos a etapa 8.2.2 para outros valores dos parâmetros, a fim de verificar se o método que utilizamos para estimação dos parâmetros deu bom resultado. Então, fazemos \hat{p} e \hat{q} variarem até 2% em torno de seus valores. Tomamos

$$\hat{p}' = \hat{p} \pm 0,005 s \hat{p} \quad (3.1.13)$$

$$, s = 1, \dots, 4$$

$$\hat{q}' = \hat{q} \pm 0,005 s \hat{q} \quad (3.1.14)$$

3.2. ESTIMAÇÃO DOS PARÂMETROS DE UMA BETA (p,q).

Para testarmos o ajuste dos valores obtidos com o modelo Beta (p,q), temos que estimar os parâmetros desta distribuição.

Seja β uma variável aleatória, cuja distribuição é Beta com parâmetros p e q. Sua densidade é definida como:

$$f_{\beta}(b) = \begin{cases} \frac{b^{p-1} (1-b)^{q-1}}{B(p,q)} & , 0 < b < 1 \\ 0 & , \text{ caso contrário} \end{cases} \quad (3.2.1)$$

Usando o fato que

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \quad , \quad (3.2.2)$$

podemos escrever essa função densidade como

$$f_{\beta}(b) = \begin{cases} \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} b^{p-1} (1-b)^{q-1} & , 0 < b < 1 \\ 0 & , \text{ caso contrário} \end{cases} \quad (3.2.3)$$

onde

$$\Gamma(s) = \int_0^{\infty} x^{s-1} e^{-x} dx \quad , \quad s > 0 \quad (3.2.4)$$

A nossa hipótese não especificou os valores dos parâmetros p e q . Então usamos o seguinte método desenvolvido por Beckman e Tietjen (1978) para estimar o valor desses parâmetros.

Dada uma amostra $\beta_1, \beta_2, \dots, \beta_n$ de variáveis aleatórias

independentes de uma distribuição $B(p,q)$, as equações de máxima verossimilhança para os estimadores \hat{p} e \hat{q} são as seguintes:

$$\psi(\hat{p}) - \psi(\hat{p} + \hat{q}) = \ln G1 \quad (3.2.5)$$

$$\psi(\hat{q}) - \psi(\hat{p} + \hat{q}) = \ln G2$$

onde

$$G1 = \prod_{i=1}^n b_i^{1/n} \quad e \quad G2 = \prod_{i=1}^n (1-b_i)^{1/n}$$

$$e \quad \psi(z) = \frac{d \ln \Gamma(z)}{dz} \quad : \text{ função Psi.}$$

Resolvemos a primeira equação em (3.2.5) para $\psi(\hat{p})$

$$\psi(\hat{p}) = \ln G1 + \psi(\hat{p} + \hat{q}) \quad (3.2.6)$$

Da segunda equação em (3.2.5)

$$\psi(\hat{p} + \hat{q}) = \psi(\hat{q}) - \ln G2 \quad (3.2.7)$$

Substituindo essa expressão em (3.2.6), obtemos

$$\psi(\hat{p}) = \ln G1 - \ln G2 + \psi(\hat{q}), \text{ então}$$

$$\hat{p} = \psi^{-1}(\ln G1 - \ln G2 + \psi(\hat{q})) \quad (3.2.8)$$

Substituímos esse valor de \hat{p} na segunda equação de (3.2.5), para obtermos uma equação com uma única variável:

$$\psi(\hat{q}) - \psi\{\psi^{-1}(\ln G1 - \ln G2 + \psi(\hat{q}))\} + \hat{q} - \ln G2 = 0 \quad (3.2.9)$$

Para resolver a equação (3.2.9) para \hat{q} , usamos um método baseado no método da secante para encontrar a raiz. Este método requer que especifiquemos um intervalo (A,B) dentro da qual a raiz está localizada.

A solução requer que estejamos capacitados para estimar ambos $\psi(z)$ e $\psi^{-1}(z)$.

Estimar $\psi^{-1}(z)$ é equivalente a encontrar a raiz c da equação $\psi(c) - z = 0$, e isso é realizada pela técnica de resolver raízes discutida acima.

Para calcular aproximadamente $\psi(z)$, usamos a seguinte aproximação

$$\psi(z) \sim \ln z - \frac{1}{2z} - \frac{1}{12z^2} + \frac{1}{120z^4} - \frac{1}{256z^6}, \text{ para } z \geq 3 \quad (3.2.10)$$

e

$$\psi(z) = \psi(z+1) - \frac{1}{z}, \text{ para } z < 3. \quad (3.2.11)$$

Essas aproximações segundo Beckman e Tietjen (1978) dão uma precisão de 6 casas decimais.

Poderíamos ter estimado esses parâmetros por outros métodos mais simples, como por exemplo usando método dos momentos. Porém, nesse caso não poderíamos garantir a validade do teste χ^2 . Pelo contrário, se os parâmetros p e q são estimados pelo método de máxima verossimilhança, a distribuição da estatística χ^2 definida em etapa 8.3 da subseção 3.1, quando $n \rightarrow \infty$ tende para a distribuição χ^2 com $k-3$ graus de liberdade (Fisz (1963)).

CAPÍTULO IV

RESULTADOS

Em seguida apresentamos os resultados obtidos e como prosseguimos para encontrar as condições sob as quais a variável β_i tem distribuição $B(p,q)$.

Quando trabalhamos no espaço com dimensão $n=2$ e tomamos o número de classes $k=5$, obtivemos o seguinte:

Características das classes	Acertamento	Parâmetros	χ^2 Observado	χ^2 para 5%	Graus de Liberdade
classes pequenas e separadas.	Certos = 2000 Errados = 0	$\hat{p} = 3.26$ $\hat{q} = 35.37$	91.207	32.671	21
aumentado as classes, ainda sem interseção	Certos=1982 Errados = 18	$\hat{p} = 3.12$ $\hat{q} = 18.37$	53.083	31.410	20
aumentando mais as classes com alguns pontos comuns	Certos = 1848 Errados = 152	$\hat{p} = 2.81$ $\hat{q} = 10.83$	47.891	33.924	22
aumentando mais as classes com vários pontos comuns	Certos = 1610 Errados = 390	$\hat{p} = 2.68$ $\hat{q} = 8.40$	40.285	33.924	22
grandes classes interceptando uma na outra.	Certos = 1331 Errados = 669	$\hat{p} = 2.16$ $\hat{q} = 5.43$	13.407	28.869	18

TABELA 4.1.

Para $n = 3, k = 5$

Características das classes	Acertamento	Parâmetros	χ^2 Observado	χ^2 para 5%	Graus de Liberdade
classes pequenas e separadas	Certos = 2000 Errado = 0	$\hat{p} = 5.42$ $\hat{q} = 40.89$	38.490	28.869	18
aumentando um pouco, ainda, sem intersecção	Certos = 1998 Errados = 2	$\hat{p} = 4.39$ $\hat{q} = 22.67$	38.620	27.587	17
aumentado mais as classes com alguns pontos comuns	Certos = 1943 Errados = 57	$\hat{p} = 3.97$ $\hat{q} = 17.00$	27.219	26.296	16
aumentando mais as classes com vários pontos comuns	Certos = 1786 Errados = 214	$\hat{p} = 3.59$ $\hat{q} = 13.34$	43.860	26.296	16
várias classes interceptando uma na outra	Certos = 15.77 Errados = 423	$\hat{p} = 2.06$ $\hat{q} = 3.98$	214.017	30.144	19

TABELA 4.2.

Para $n = 5, k = 5$

Características das classes	Acertamento	Parâmetros	χ^2 Observado	χ^2 para 5%	Graus de Liberdade
classes pequenas e separadas	Certos = 1991 Errados = 9	$\hat{p} = 9.10$ $\hat{q} = 42.02$	23.512	27.587	17
classes pequenas com alguns pontos em comum	Certos = 1916 Errados = 84	$\hat{p} = 7.36$ $\hat{q} = 28.20$	21.967	27.587	17

aumentando mais as clas ses com vá- rios pontos em comum	Certos = 1792 Errados = 208	$\hat{p} = 6.72$ $\hat{q} = 22.78$	31.010	27.587	17
aumentando mais as classes	Certos = 1397 Errados = 603	$\hat{p} = 5.05$ $\hat{q} = 14.57$	67.409	26.296	16
classes gran des intercep tando uma na outra	Certos = 902 Errados = 1098	$\hat{p} = 4.92$ $\hat{q} = 12.16$	111.941	23.685	14

TABELA 4.3.

Os resultados acima pela tabela (4.1) parecia nos permitir conjecturar o seguinte: a medida que fomos aumentando o tamanho das classes, obtivemos valores menores de χ^2 . Até chegarmos a não ter nenhuma evidência em rejeitar que a nossa variável β_i tem distribuição Beta com parâmetros $\hat{p} = 2.16$ e $\hat{q} = 5.43$, quando tomamos as nossas 5 classes bem grandes, interceptando uma na outra. Já a análise das tabelas 4.2 e 4.3 nos levaria a uma conclusão exatamente oposta a anterior, pois obtivemos valores de χ^2 pequenos quando tomamos as 5 classes pequenas e separadas.

Então tivemos que deixar esses resultados de lado e partir em outras buscas, pois eles não nos levam a uma conjectura coerente.

Visto que a distribuição multivariada em k dimensões tem uma densidade constante sobre elipsóides da forma

$$(x - \mu)' \Sigma^{-1} (x - \mu) = c^2, \quad (4.1)$$

c^2 sendo uma constante. Então concentramos as nossas classes nesta elipsóide.

Iniciamos com pontos normais bivariados que formam uma região de concentração de forma elíptica com a direção dos eixos dadas por autovetores da matriz de covariância.

Obtivemos resultados semelhantes para diferentes vetores média e matrizes de covariância. A fim de ilustrarmos as conclusões obtidas, tomamos como exemplo uma das matrizes por nós trabalhada.

Temos $n = 2$, o vetor média $\mu = (-10, 80)$ e a matriz de co

variância $\Sigma = \begin{pmatrix} 2412.77 & -1780.40 \\ -1780.40 & 2681.17 \end{pmatrix}$

Tomamos $k = 10$ classes (retângulos) que são:

CLASSES	EIXO 1	EIXO 2
1	(-50, -30)	(120, 140)
2	(-80, -60)	(90, 115)
3	(-30, -10)	(90, 115)
4	(-60, -30)	(55, 75)
5	(-52, -27)	(22, 42)
6	(-20, 20)	(50, 75)
7	(-10, 25)	(24, 42)
8	(35, 60)	(50, 75)
9	(8, 30)	(80, 102)
10	(0, 30)	(110, 130)

Com os valores estimados $\hat{p} = 2.88$ e $\hat{q} = 23.31$, obtivemos a seguinte tabela:

INTERVALOS	FREQUÊNCIA ABSOLUTA	FREQUÊNCIA ESTIMADA
.0000 — .0162	33	20.7
.0162 — .0323	89	93.6
.0323 — .0485	140	166.0
.0485 — .0647	189	214.0
.0647 — .0808	231	234.0
.0808 — .0970	262	231.0
.0970 — .1132	226	212.0
.1132 — .1294	192	185.0
.1294 — .1455	154	155.0
.1455 — .1617	156	125.0
.1617 — .1779	90	98.4
.1779 — .1940	72	75.3
.1940 — .2102	39	56.4
.2102 — .2264	30	41.3
.2264 — .2425	25	29.6
.2425 — .2587	29	20.9
.2587 — .2749	15	14.4
.2749 — .2910	8	9.8
.2910 — .3072	7	6.6
.3072 — .3234	6	4.3
.3234 — .3396	1	2.8

.3896 — .3557	2	1.8
.3557 — .3719	0	1.1
.3719 — .3881	1	.7
.3881 — .4042	4	.4

TABELA 4.4.

E o χ^2 observado foi 42.293 com 17 graus de liberdade

$$\chi_{5\%}^2 = 27.587$$

Graficamente, as 10 regiões estão localizadas da seguinte

forma:

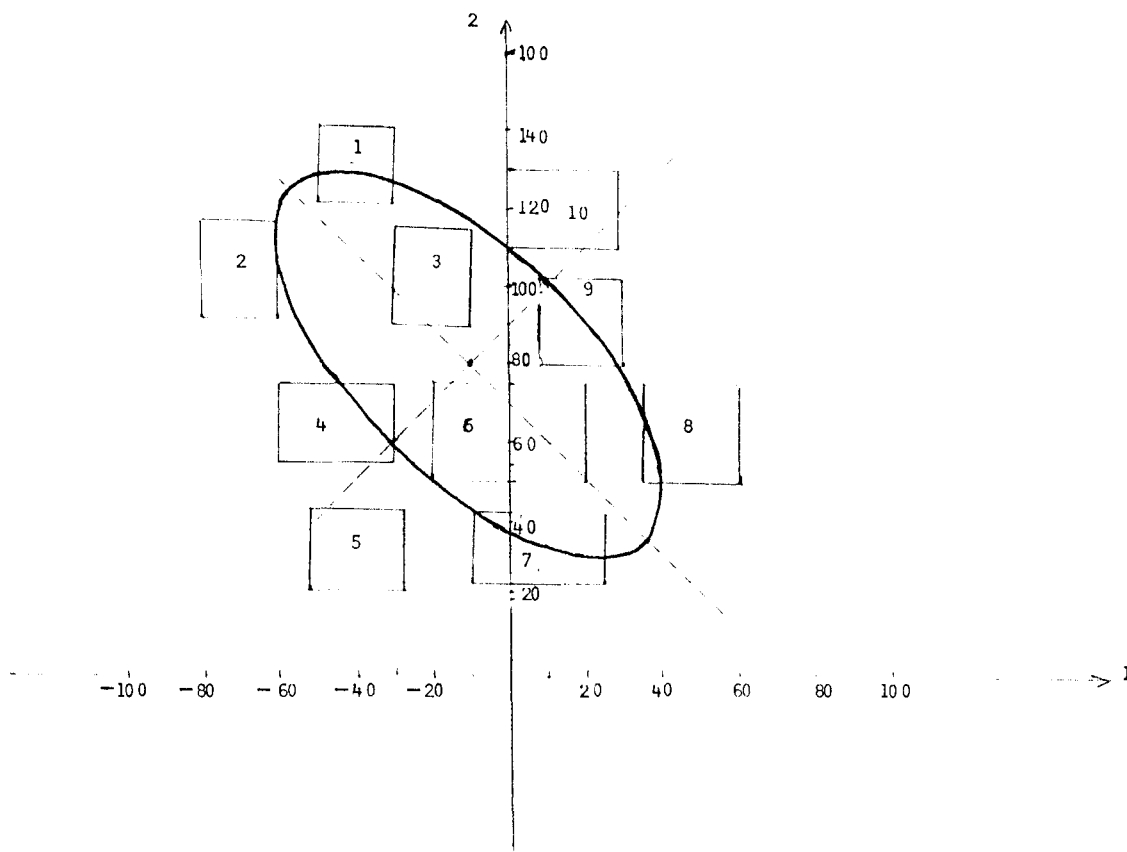


Figura 4.1

Quando aumentamos as nossas regiões, ou seja, a elipse onde estão localizados os pontos normais bivariados fica mais coberta, da forma a seguir:

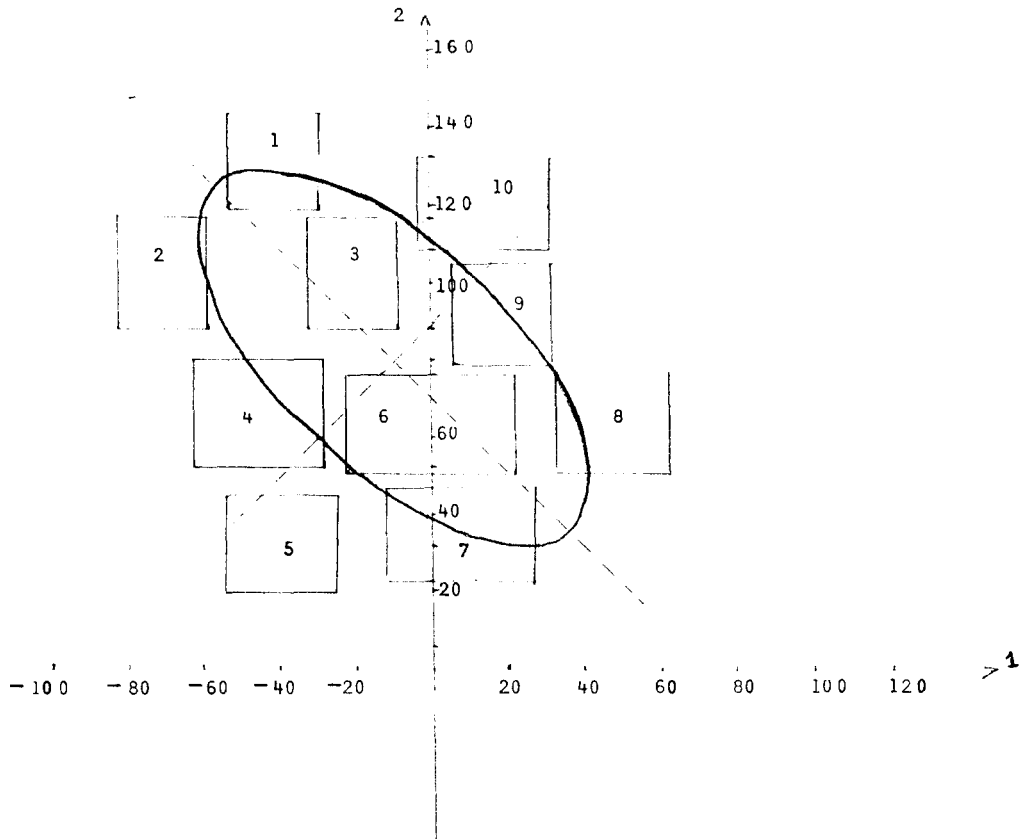


Figura 4.2

A situação melhora e as classes são:

CLASSE	EIXO 1	EIXO 2
1	(-52, -28)	(118, 142)
2	(-82, -58)	(88, 117)
3	(-32, -8)	(88, 117)
4	(-62, -28)	(53, 77)
5	(-54, -25)	(20, 44)
6	(-22, 22)	(48, 77)
7	(-12, 27)	(22, 44)
8	(33, 62)	(48, 77)
9	(6, 32)	(78, 104)
10	(-2, 32)	(108, 132).

Os valores estimados foram $\hat{p} = 3.01$ e $\hat{q} = 20.88$ e obtivemos as seguintes frequências

INTERVALO	FREQUÊNCIA ABSOLUTA	FREQUÊNCIA ESTIMADA
.0000 — .0184	28	16.9
.0184 — .0369	82	84.3
.0369 — .0553	133	158.0
.0553 — .0738	183	211.0
.0738 — .0922	221	235.0
.0922 — .1106	253	235.0
.1106 — .1291	235	218.0
.1291 — .1475	211	191.0

.1475	→	.1660	180	160.0
.1660	→	.1844	137	129.0
.1844	→	.2028	97	101.0
.2028	→	.2213	76	76.5
.2213	→	.2397	42	56.5
.2397	→	.2581	48	40.8
.2581	→	.2766	18	28.8
.2766	→	.2950	17	19.8
.2950	→	.3135	14	13.4
.3135	→	.3319	9	8.8
.3319	→	.3503	8	5.7
.3503	→	.3688	1	3.6
.3688	→	.3872	4	2.2
.3872	→	.4057	0	1.3
.4057	→	.4241	1	.8
.4241	→	.4425	1	.5
.4425	→	.4610	1	.3

TABELA 4.5.

E χ^2 observado foi 33.77 com 17 graus de liberdade

$$\chi_{5\%}^2 = 27.587.$$

E finalmente tomamos as nossas regiões independentes cobrindo praticamente toda a elipse, como podemos ver no gráfico seguinte:

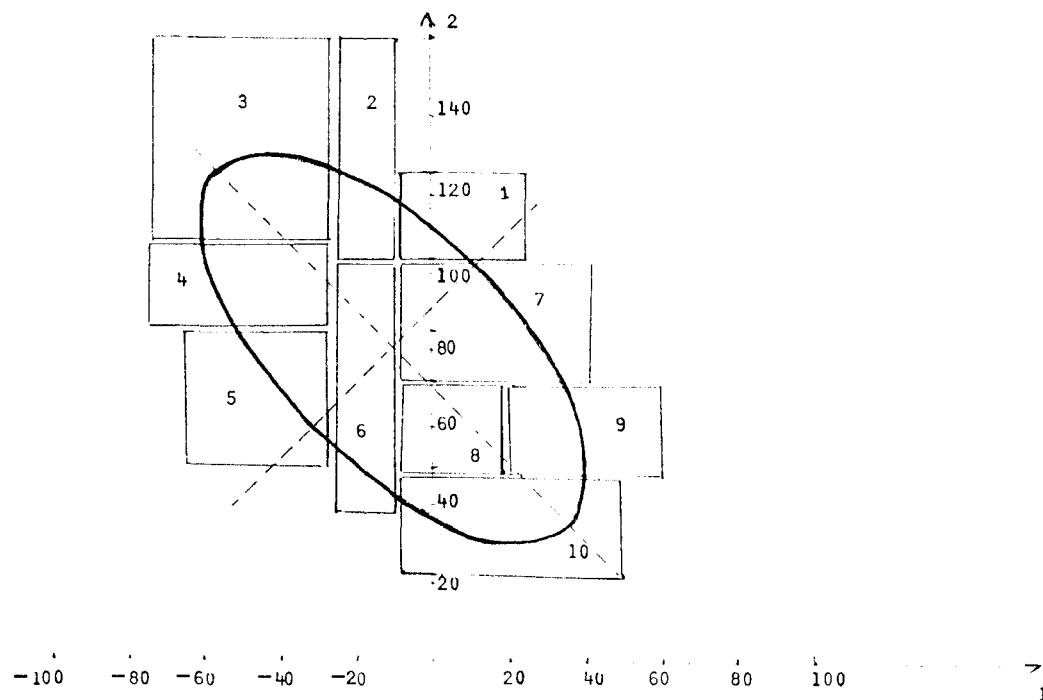


Figura 4.3

As 10 classes são:

CLASS IS	EIXO 1	EIXO 2
1	(-8, 25)	(104, 126)
2	(-25, -9)	(104, 160)
3	(-72, -26)	(108, 160)
4	(-72, -26)	(86, 107)
5	(-65, -26)	(50, 85)

6	(-25, -9)	(38, 103)
7	(-8, 42)	(71, 103)
8	(-8, 18)	(48, 70)
9	(19, 60)	(48, 70)
10	(-8, 50)	(22, 47)

Obtivemos os seguintes resultados:

$$\hat{p} = 2.55 \quad , \quad \hat{q} = 12.64$$

e χ^2 observado = 27.46 com 18 graus de liberdade.

Obtivemos as seguintes frequências:

INTERVALOS	FREQUÊNCIA ABSOLUTA	FREQUÊNCIA ESTIMADA
.0000 — .0230	30	24.5
.0230 — .0459	87	90.6
.0459 — .0689	128	150.0
.0689 — .0919	176	189.0
.0919 — .1148	216	207.0
.1148 — .1378	212	209.0
.1378 — .1608	204	199.0
.1608 — .1837	193	180.0
.1837 — .2067	184	158.0
.2067 — .2297	118	134.0
.2297 — .2527	120	110.0

.2527	→	.2726	85	88.9
.2726	→	.2986	67	70.0
.2986	→	.3216	40	53.8
.3216	→	.3445	29	40.6
.3445	→	.3675	36	29.9
.3675	→	.3905	19	21.6
.3905	→	.4134	21	15.3
.4134	→	.4364	15	10.3
.4364	→	.4594	8	7.1
.4594	→	.4823	3	4.7
.4823	→	.5053	6	3.0
.5053	→	.5283	2	1.9
.5283	→	.5512	0	1.1
.5512	→	.5742	1	.7

TABELA 4.6.

Não temos nenhuma evidência em rejeitar a hipótese que as variáveis β_i 's tem distribuição $B(p,q)$ ao nível de confiança de 5%. Pois a esse nível de confiança e com 18 graus de liberdade só rejeitaremos a hipótese para valores do qui-quadrado maiores que 28.869.

Este raciocínio, isto é, deixar cobrir as classes o espaço mais provável criado pela Σ da multinormal, foi seguido para outros valores de médias e matrizes de covariância e também para diferentes dimensões n e obtivemos resultados semelhantes.

Como já citamos, na etapa 11 da nossa simulação, calculamos o χ^2 para diferentes valores de p e q . Variamos p e q em torno de até 2% de \hat{p} e \hat{q} , a fim de verificar o comportamento das estimativas obtidas pelo método de máxima verossimilhança.

Vejamos em duas situações diferentes.

Primeiro, quando tivemos uma situação em que as classes não cobriam a elipsóide que é região de concentração dos pontos padrões, as estimativas obtidas pela máxima verossimilhança não justificavam a explicação de β por uma Beta. A estimativa do parâmetro q apresentou um valor demasiadamente grande. Fomos variando os valores dos parâmetros p e q e notamos que quando diminuimos ambos p e q , o qui-quadrado também decresceu até certo valor e voltando a crescer novamente. E quando houve aumento de valores de p e q , χ^2 cresceu sempre.

Podemos observar no seguinte exemplo. Para variáveis aleatórias normas bidimensionais com média $\mu = (0,0)$ e matriz de covariância $\Sigma = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$. Os pontos normais estão concentrados na elipse seguinte (fig. 4.4) e tomamos as 10 classes da seguinte forma

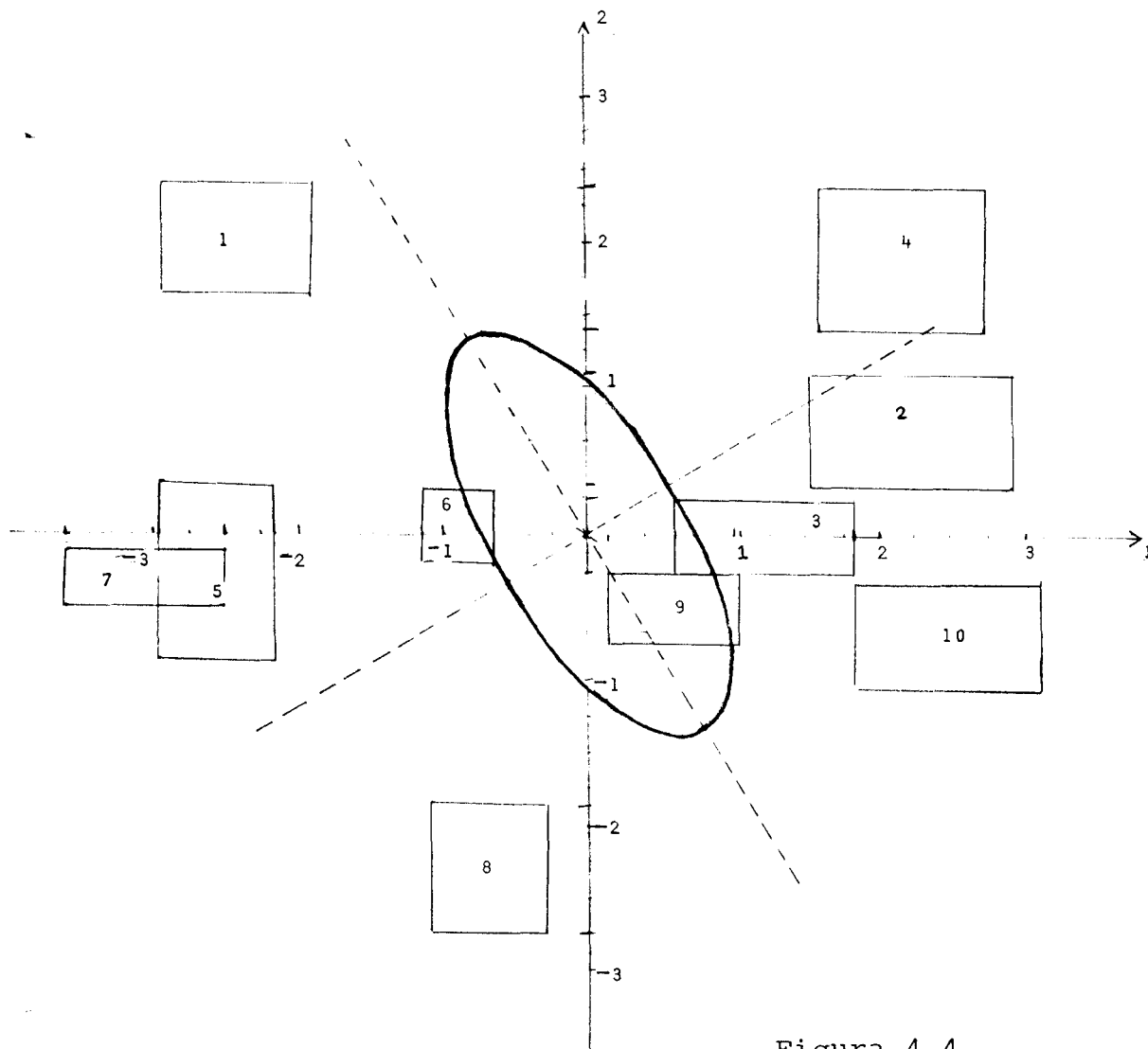


Figura 4.4

Obtivemos $\hat{p} = 3.41$ e $\hat{q} = 50.27$ e $\chi^2 = 105.7$ com 18 graus de liberdade

$$\chi_{5\%}^2 = 28.869.$$

O gráfico da variação de p e q em torno de até 2% de \hat{p} e \hat{q} é apresentado em seguida.

Os valores de χ^2 para

$$\hat{p}' = \hat{p} \pm .0005i\hat{p}$$

$$\hat{q}' = \hat{q} \pm .0005i\hat{q}$$

$$i = 0, 1, \dots, 4.$$

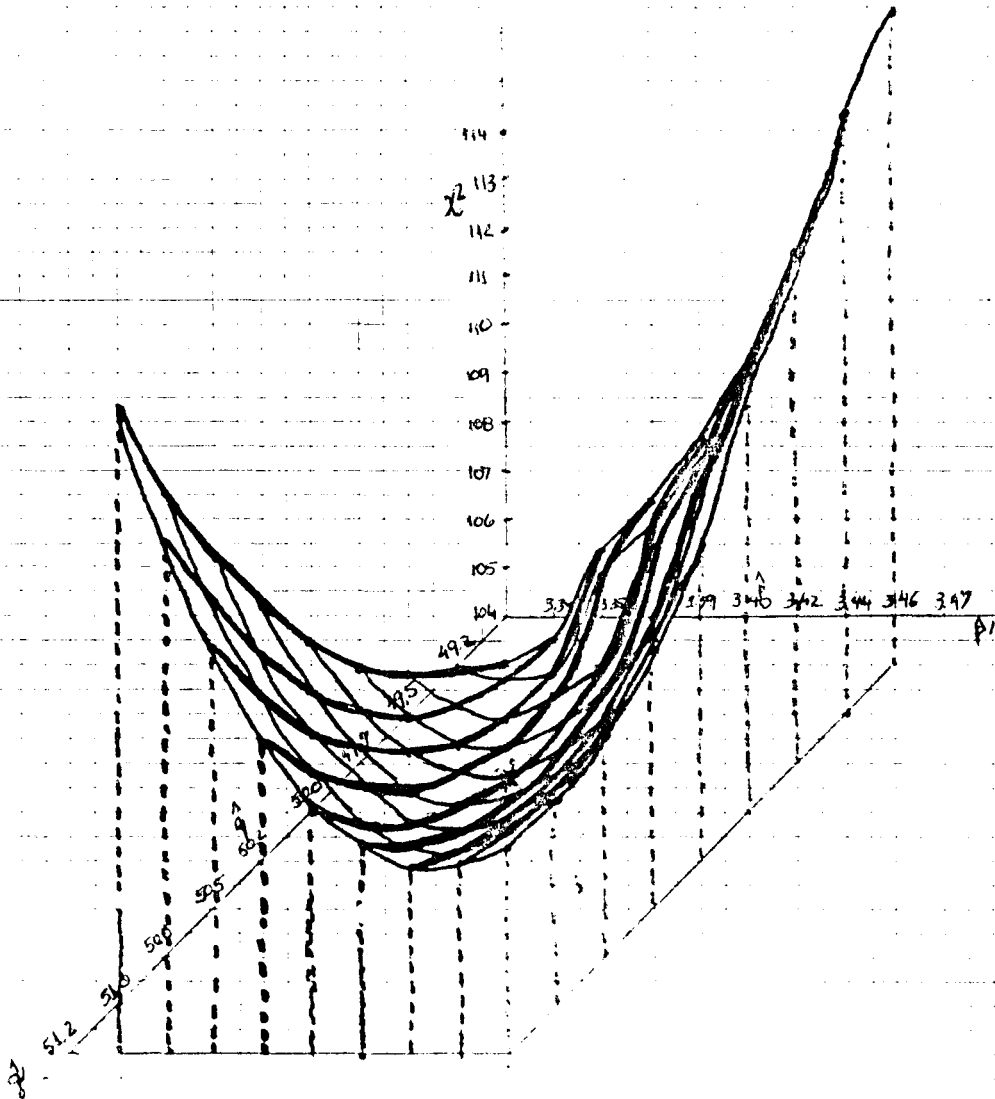


Figura 4.5

Na vizinhança de até 2% que tomamos, sempre rejeitamos a hipótese que a variável β tem distribuição $Be+a(p,q)$.

Temos a segunda situação, quando as classes cobrem a elipse. Nesse caso, calculamos o χ^2 para p e q na vizinhança de até 1% em torno das estimativas e obtivemos valores próximos a aqueles obtidos anteriormente para \hat{p} e \hat{q} . As vezes chega até a "melhorar" o valor de χ^2 obtido para as estimativas, mas essa "melhora" é pouco significativa. Já para p e q próximos a 2% de \hat{p} e \hat{q} , obtivemos valores de χ^2 grandes, isto é, a medida que fomos afastando dos valores estimados \hat{p} e \hat{q} , o valor de χ^2 foi aumentando de tal forma que chegamos em alguns casos, até a rejeitar a hipótese que a variável aleatória β_i tem distribuição $B(p,q)$. Então o uso do estimador de máxima verossimilhança, mais uma vez, se justificou plenamente, mesmo em um caso de simulação.

Podemos observar no mesmo exemplo anterior, onde as classes foram tomadas conforme a figura a seguir.

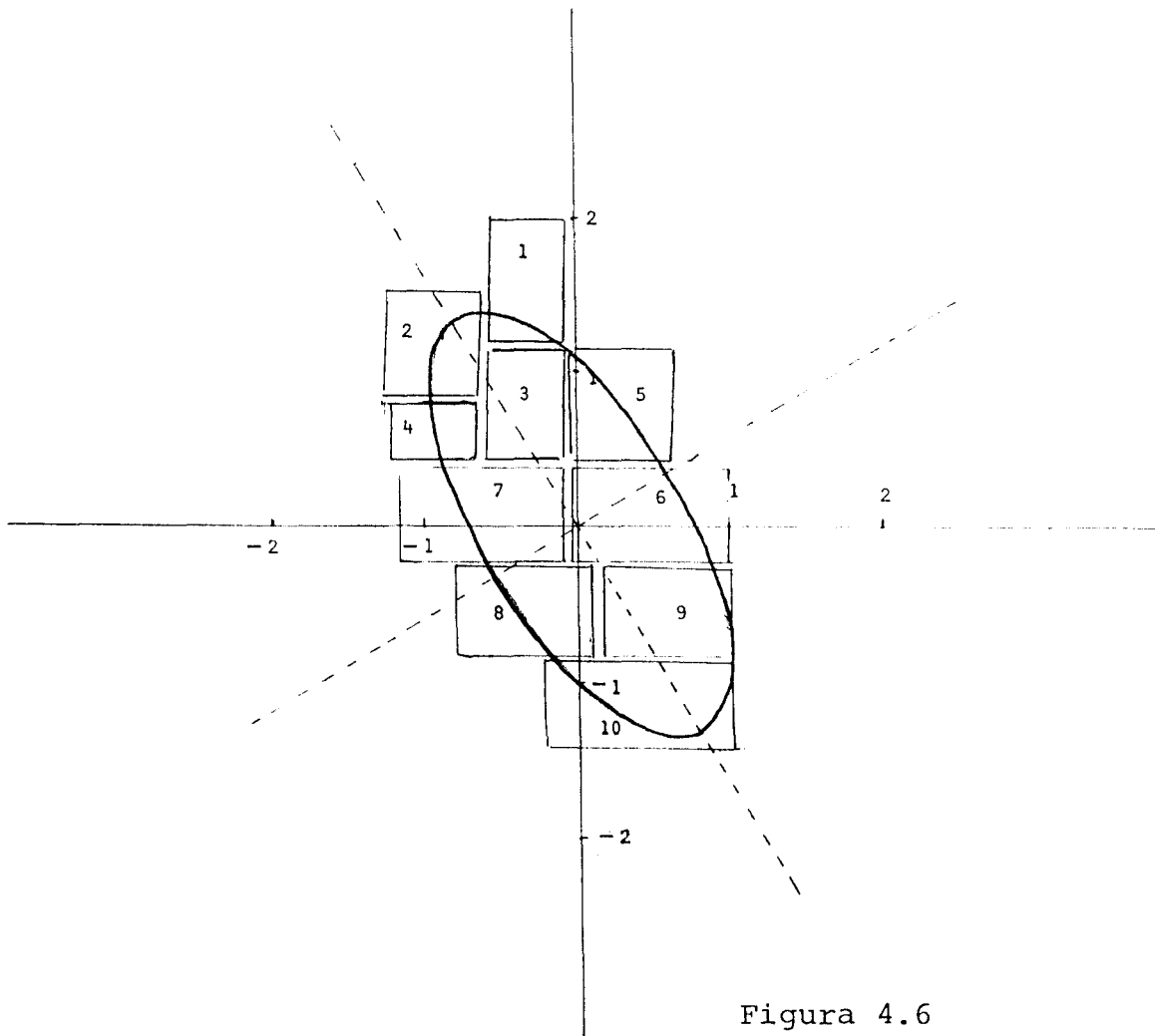


Figura 4.6

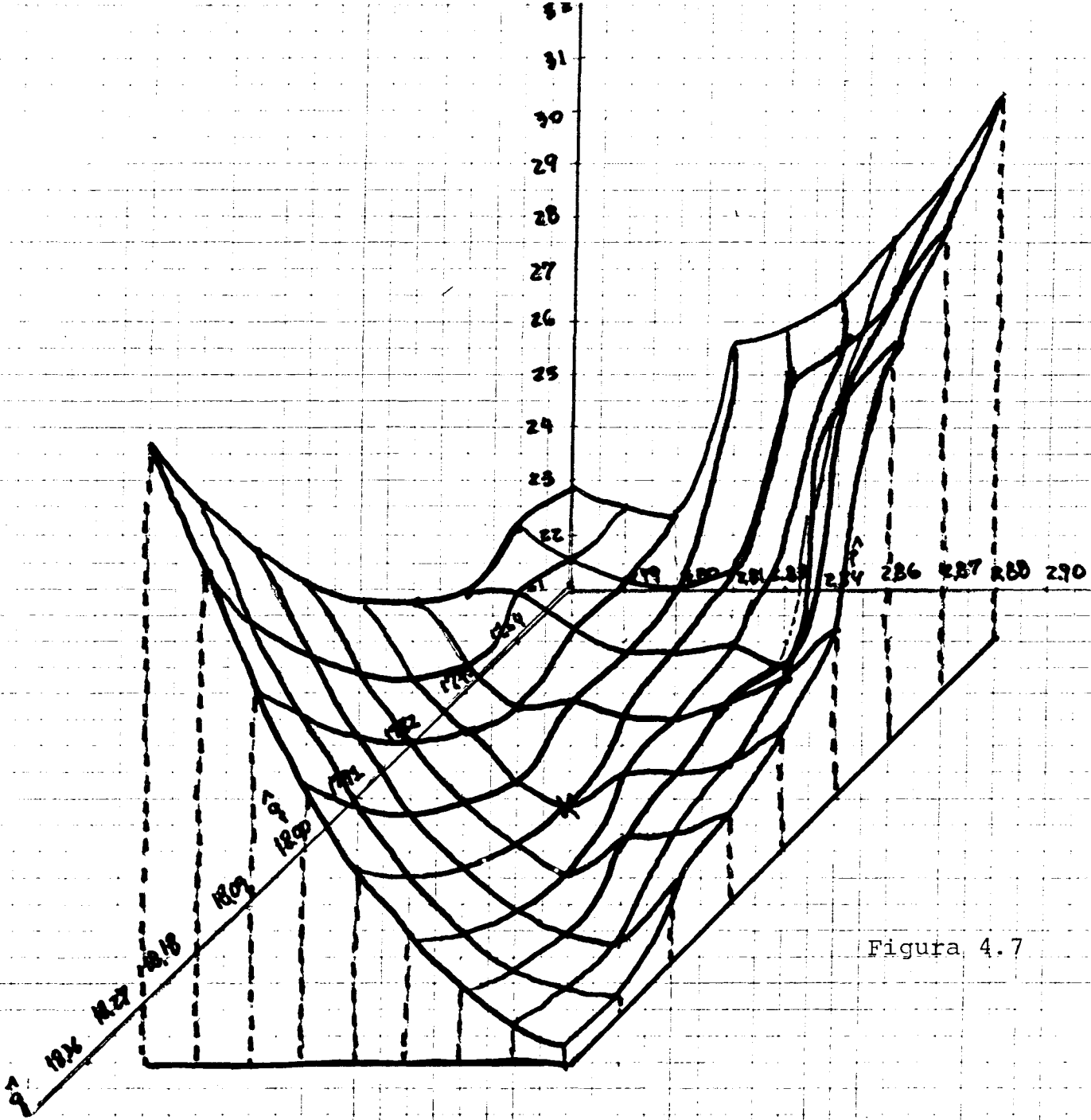
E obtivemos o seguinte comportamento do χ^2 para as varia-
ções de p e q .

$$\hat{p} = 2.85 \quad , \quad \hat{q} = 18.01$$

χ^2 observado = 21.81 com 17 graus de liberdade.

$$\chi_{5\%}^2 = 27.587.$$

A figura 4.7 representa as variações de p e q em torno
de até 2% dos valores estimados \hat{p} e \hat{q} .



Os valores de χ^2 para
 $\hat{p}' = \hat{p} \pm .005 i \hat{p}$, $i=0,1,\dots,4$
 $\hat{q}' = \hat{q} \pm .005 i \hat{q}$

Figura 4.7

Apresentamos alguns resultados obtidos, ilustrando casos nos quais a distribuição da variável pode ser explicado por uma Beta (p,q)

$$n = 2 \quad , \quad k = 10$$

$$\mu = (1 \quad 1) \quad \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

CLASSES	EIXO 1	EIXO 2
1	(1.251, 2.000)	(1.660, 2.430)
2	(.350, 1.250)	(2.001, 2.500)
3	(.350, 1.250)	(1.252, 2.000)
4	(-.200, .349)	(.980, 2.000)
5	(1.251, 2.120)	(.900, 1.649)
6	(.350, 1.250)	(.252, 1.251)
7	(-.200, .349)	(.000, .979)
8	(1.251, 2.120)	(.252, .899)
9	(.350, 1.250)	(-.500, .251)
10	(1.251, 1.850)	(-.450, .251)

Distribuição beta com $\hat{p} = 3.14$ e $\hat{q} = 15.38$.

INTERVALOS	FREQUÊNCIA ABSOLUTA	FREQUÊNCIA ESTIMADA
.0000 — .0206	11	7.9
.0206 — .0412	54	45.8
.0412 — .0618	87	97.7
.0618 — .0824	131	145.0
.0824 — .1030	181	180.0
.1030 — .1237	174	198.0
.1237 — .1443	197	202.0
.1443 — .1649	220	194.0
.1649 — .1855	190	178.0
.1855 — .2061	156	157.0
.2061 — .2267	140	134.0
.2267 — .2473	122	111.0
.2473 — .2679	87	89.4
.2679 — .2885	66	70.3
.2885 — .3091	56	54.0
.3091 — .3298	32	40.6
.3298 — .3504	33	29.8
.3504 — .3710	16	21.4
.3710 — .3916	14	15.1
.3916 — .4122	8	10.4
.4122 — .4328	11	7.0
.4328 — .4534	8	4.6

.4534 — .4740	3	2.9
.4740 — .4946	1	1.8
.4946 — .5152	1	1.1

Qui-quadrado observado = 23.617 com 19 graus de liberdade.

Qui-quadrado crítico = 30.144 com $\alpha = 5\%$

$$n = 3, \quad k = 10$$

$$\mu = (60, -45, 75.9) \quad \Sigma = \begin{pmatrix} 1087.02 & 728.30 & 837.76 \\ 728.30 & 2209.00 & 325.71 \\ 837.76 & 325.71 & 1089.00 \end{pmatrix}$$

CLASSES	EIXO 1	EIXO 2	EIXO 3
1	(20, 60)	(-80, -10)	(100, 150)
2	(60, 100)	(-80, -10)	(100, 150)
3	(-10, 60)	(-150, -80)	(50, 100)
4	(60, 130)	(-150, -80)	(50, 100)
5	(-10, 60)	(-80, -10)	(50, 100)
6	(60, 130)	(-80, -10)	(50, 100)
7	(-10, 60)	(-10, 60)	(50, 100)
8	(60, 130)	(-10, 60)	(50, 100)
9	(20, 60)	(-80, -10)	(0, 50)
10	(60, 100)	(-80, -10)	(0, 50)

Distribuição beta com $\hat{p} = 3.88$ e $\hat{q} = 11.12$

INTERVALOS	FREQUÊNCIA ABSOLUTA	FREQUÊNCIA ESTIMADA
.0000 — .0252	3	1.0
.0252 — .0504	15	10.5
.0504 — .0756	33	32.9
.0756 — .1008	61	64.6
.1008 — .1260	85	99.6
.1260 — .1512	138	132.0
.1512 — .1764	132	158.0
.1764 — .2016	177	175.0
.2016 — .2268	193	182.0
.2268 — .2520	201	180.0
.2520 — .2772	173	171.0
.2772 — .3024	150	156.0
.3024 — .3276	133	137.0
.3276 — .3528	122	117.0
.3528 — .3780	94	97.2
.3780 — .4032	79	78.2
.4032 — .4284	56	61.1
.4284 — .4536	44	46.3
.4536 — .4788	39	34.1
.4788 — .5040	22	24.3
.5040 — .5292	20	16.8
.5292 — .5544	16	11.2

.5544 → .5796	7	7.2
.5796 → .6048	2	4.5
.6048 → .6300	4	2.6

Qui-quadrado observado = 18.414 com 20 graus de liberdade.

Qui-quadrado crítico = 31.410 com $\alpha = 5\%$

$$n = 5, \quad k = 10$$

$$\mu = (20, \quad -10, \quad 32, \quad 14, \quad -30)$$

$$\Sigma = \begin{pmatrix} 51.69 & 8.81 & 12.31 & 15.39 & 11.12 \\ 8.81 & 76.56 & 12.26 & 15.32 & 22.08 \\ 12.31 & 12.26 & 26.94 & 6.06 & 10.56 \\ 15.39 & 15.32 & 6.06 & 94.67 & 9.50 \\ 11.12 & 22.08 & 10.56 & 9.50 & 66.26 \end{pmatrix}$$

CLASSES	EIXO 1	EIXO 2	EIXO 3	EIXO 4	EIXO 5
1	(10, 20)	(-4, 12)	(28, 35)	(20, 34)	(-36, -25)
2	(20, 30)	(-4, 12)	(28, 35)	(20, 34)	(-36, -25)
3	(5, 20)	(-16, -4)	(21, 28)	(7, 20)	(-47, -36)
4	(20, 35)	(-16, -4)	(21, 28)	(7, 20)	(-47, -36)
5	(5, 20)	(-16, -4)	(28, 35)	(7, 20)	(-36, -25)
6	(20, 35)	(-16, -4)	(28, 35)	(7, 20)	(-36, -25)

7	(5, 20)	(-16, -4)	(35, 43)	(7, 20)	(-25, -13)
8	(20, 35)	(-16, -4)	(35, 43)	(7, 20)	(-25, -13)
9	(10, 20)	(-28, -16)	(28, 35)	(-6, 7)	(-36, -25)
10	(20, 30)	(-28, -16)	(28, 35)	(-6, 7)	(-36, -25)

Distribuição beta com $\hat{p} = 6.10$ e $\hat{q} = 14.76$.

INTERVALOS	FREQUÊNCIA ABSOLUTA	FREQUÊNCIA ESTIMADA
.0000 → .0249	0	.1
.0249 → .0497	1	.5
.0497 → .0746	5	3.9
.0746 → .0995	15	14.2
.0995 → .1243	35	34.3
.1243 → .1492	63	63.7
.1492 → .1740	96	99.2
.1740 → .1989	139	135.0
.1989 → .2238	154	167.0
.2238 → .2486	188	189.0
.2486 → .2735	201	200.0
.2735 → .2984	188	198.0
.2984 → .3232	174	186.0
.3232 → .3481	187	166.0
.3481 → .3729	139	141.0
.3729 → .3978	120	115.0

.3978 — .4227	100	89.7
.4227 — .4475	66	66.9
.4475 — .4724	54	47.9
.4724 — .4973	28	32.7
.4973 — .5221	21	21.4
.5221 — .5470	17	13.3
.5470 — .5718	4	7.9
.5718 — .5967	3	4.4
.5967 — .6216	2	2.4

Qui-quadrado observado = 11.794 com 18 graus de liberdade

Qui-quadrado crítico = 28.869 com $\alpha = 5\%$.

CAPÍTULO V

CONCLUSÃO

O nosso sistema "capaz de aprender" (Ω, S, g, f, v) , aqui como sistema de aprendizagem reduzido, sabe que existem várias classes C_i sobre Ω e qualquer padrão a ser classificação pertence a uma das classes C_i .

As classes do professor são hipercubos distribuídos uniformemente inicialmente, as classes do sistema são hiperbolas (a regra de decisão e a da mínima distância), e o ambiente é gerado por variáveis aleatórias multinormais.

Após uma análise sobre o comportamento da variável aleatória β , a função correção dada em (1.3.3), chegamos que uma distribuição Beta com parâmetros p e q dá um ajuste satisfatório quando as classes aprendidas estão sobre a região onde os pontos "eventos" estão concentrados. Há então, indicação pelos resultados das simulações que podemos por como indicador de uma aprendizagem correta o ajuste da distribuição amostral da variável aleatória β a uma distribuição teórica Beta.

Lembramos que da maneira como foi construída, as variáveis β 's não são independentes. Pois os pontos referentes podem ser alterados em qualquer das etapas iterativas e estas alterações dependem do valor de β numa certa etapa anterior. Mesmo assim, obtivemos que

a variável β pode ser explicada por uma distribuição Beta (p,q) nas condições descritas anteriormente. O fato de uma classificação ser correta ou incorreta também não influí no ajuste de uma distribuição Beta (p,q) .

ESTE PROGRAMA TESTA O AJUSTE DOS VALORES

DISTANCIA (X(I),Y(L))
REF(A(I))= ----- , L=1,....,NG
MAXIMA DISTANCIA (X(I),Y(L))

ONDE:

X : DISTRIBUICAO NORMAL N(U,A)

Y : DISTRIBUICAO UNIFORME U(A,B), INICIALMENTE ;

COM UMA DISTRIBUICAO OPTA(P,Q).

PARA CALCULAR AUTOVALORES E AUTOVETORES
DA MATRIZ DE COVARIANCIA UTILIZA AS SEGUINTE
SUBROTINAS SAS:

- 1 - F01AHE.TEX
- 2 - F02AHE.TEX
- 3 - P01AHE.TEX
- 4 - P02AHE.TEX
- 5 - X02AAE.TEX
- 6 - X02APE.TEX
- 7 - X04AAE.TEX

OS INFERIOS ABATXO SIGNIFICAM:

MIN - UNIDADE DE ENTRADA

NDUT - UNIDADE DE SAIDA

N - DIMENSAO DO ESPACO QUE SE TRABALHA

DN - TAMANHO DA AMOSTRA

K0 - NUMERO DE CLASSES

NCL - NUMERO DE INTERVALOS DO HISTOGRAMA

DIMENSION H(2), TEL(10), XL(2), Y(10), A(10,10), D(12), VH(10), E(12)
DIMENSION H(10), TER(28), REF(20,10), S(10,10), P1(10,10), ELE(10,10)
DIMENSION P2(10,10), PGV(30), Z(10,10), PT(10,20), DIST(20)
DIMENSION AS(28), PE(12), GB(3000), R(10,10), T(10,10), AL(28), PR(28)
DIMENSION MC(28), QUL(15), XIS(110), YPS(110), GE1(2200), GE2(2200)
DIMENSION YPA(60), PRO(30), RT(10,20), TT(30), YIMP(60), RES(30)
DOUBLE PRECISION G1, G2, BETA(3000)
COMMON/12/P, C2, C1, XX
EXTERNAL F16
DATA G1/30/
DATA ICI/4/
DATA LOUF/30/
DATA N/5/
DATA DN/2000/
DATA K0/10/
DATA NCL/25/

LE OS DADOS DE ENTRADA QUE SAO:

VM=VEIOR DE MEDIA, DIMENSAO N

H = MATRIZ DE COEFICIENTES DE CORRELACAO

S = MATRIZ DIAGONAL DOS DESVIOS SIGMA(I,I)

E CALCULA A MATRIZ DE COVARIANCIA A(I,J).

READ(NIN,999)(VM(I), I=1, N)

```

9       DO 9 J=1,4
10      READ(RIN,995) (CR(I,J),J=1,4)
11      DO 10 I=1,4
12      READ(RIN,995) (S(I,I),J=1,4)
13      CALL PROC(V,S,R,P1)
14      CALL PROC(V,P1,S,A)
15      WRITE (GOUT,991)
16      WRITE (GOUT,993) (V(I),I=1,4)
17      WRITE (GOUT,994)
18      DO 16 I=1,4
19      WRITE (GOUT,993) (A(I,J),J=1,4)
20      IA=10
21      IB=10
22
23      FORMA A PLACADA DO SUBDADO
24
25      CALL PROC2AF(I,AF2AF(IT),A,IA,0,3,Z,IT)
26      IFAIL=1
27
28      AN OVALORES E AUTOVECTORES DA MATRIZ ORIGINAL
29
30      CALL PROC2AF(I,AF2AF(IT),0,3,Z,IT,IFAIL)
31      IF (IFAIL.EQ.0) GO TO 26
32      WRITE (GOUT,998) IFAIL
33      WRITE (GOUT,997) (D(I),I=1,4)
34      CONTINUE
35
36      CIA FOR = MATRIZ DIAGONAL DOS AUT. VALORES DE A
37      S = MATRIZ DE AUTOVECTORES DE A.
38
39      DO 13 I=1,4
40      DO 13 J=1,4
41      IF (I.NE.J) GO TO 12
42      C(I,J)=C(I)
43      GO TO 13
44
45      C(I,J)=0.0
46      CONTINUE
47      WRITE (GOUT,938)
48      DO 19 I=1,4
49      WRITE (GOUT,993) (C(I,J),J=1,4)
50      CONTINUE
51      DO 15 I=1,4
52      DO 15 J=1,4
53      GA=AF(I,J)
54      C(I,J)=GA**C.B
55      CALL PROC(V,Z,C,C,P2)
56
57      LE 36 EM NIBRECURSOS QUE REPRESENTAM AS CLASSES
58      E ESCOLHE ALFABETICAMENTE UM PONTO CONTIGUO MELES
59
60      FORMAT(//,2X,' OS PONTOS INICIAIS ALFABETICOS',//)
61      DO 1 I=1,40
62      DO 2 K=1,4
63      READ(RIN,975) RFI(K,I),RFI(K,I)
64      WRITE (GOUT,995) I
65      DO 4 K=1,4
66      WRITE (GOUT,982) K,RFI(K,I),RFI(K,I)
67      CONTINUE
68      WRITE (GOUT,760)
69      G=2.31567

```

DO 3 I=1,N

DO 3 K=1,P

G=KAB(G)

3 RHP(K,I)=(RHP(I,K)-RPI(I,K))*G+RPI(I,K)

DO 4 I=1,N

4 WRITE(UNIT,994)I,(RHP(I,K),K=1,P)

NCC=0

NCF=0

H=0

30 IF(.NOT.H) GO TO 31

C
C
C
C
C

GENA ALBERT EINSTEIN A UN VECTOR COLUMN DE DIMENSION
* SSE VECTOR DE DISTRIBUCION NORMAL (μ, Σ)

L=0

7 IF(L.GE.1) GO TO 6

5 DO 17 I=1,2

G=KAB(G)

H(I)=0

AL(I)=ZC(I)-1

17 CONTINUE

SS=AL(1)*P+AL(2)*P

IF(SS.GT.1.E) GO TO 6

DO 8 I=1,2

8 $\lambda(L+I)=AL(I)*SQRT(-P*LOG(SS)/SS)$

L=L+2

G) TO 7

6
C
C
C
C
C
C
C

CRIM 3 VECTOR K, DE DIMENSION N, DE DISTRIBUCION
* DISTRIBUCION NORMAL MULTIVARIADA (μ, Σ)

DO 22 I=1,N

S22A=0.

DO 19 K=1,P

19 S22A=S22A+S22(I,K)*K(K)

22 S(I)=(S22A+M(I))

C
C
C
C
C

SE D VECTOR NORMAL UNIDICO AND PERTENECE A UNA DAS CLASSES
* GENA MAIC DE MATHE, DE PERTENECER CALCULA LAS DISTANCIAS DE
* TODOS LOS REPRESENTES DE PUNTO A ESTE VECTOR

IT=0

K=0

21 IF(.NOT.G) GO TO 171

259 K=K+1

I=0

22 I=I+1

IF(S(I).GT.P2(I)) .AND. (I).LE.NP(I,K)) GO TO 21

IF(I.LT.A) GO TO 22

IT=IT+1

C
C
C
C
C

VECTOR NORMAL A DISTANCIA DE LAS CLASSES

IF(0) SIGNIFICAN LAS CLASSES = 1, 0, 0, ...

P=0. PERTENECER; P=1, 2, ...

IT(I)=K

IF(K.LE.P) GO TO 259

```

171 IF(IT.EQ.0)GO TO 30
    IF(IT.EQ.1) GO TO 172
    IU=0
272 IU=IU+1
    ANT=(IU-1.)/FLOAT(IF)
    POS=FLOAT(IU)/FLOAT(IF)
    IF(G.GE.ANT)GO TO 274
    GO TO 272
274 IF(G.LT.POS)GO TO 276
    GO TO 272
276 K=IAL(IU)
    GO TO 271
172 K=IAL(IF)
271 M=M+1
    DO 24 I1=1,K0
    SOMA=0.0
    DO 25 J=1,M
25 SOMA=(REF(I1,J)-B(J))*2+SOMA
24 DIST(I1)=SORT(SOMA)
    AMAX=DIST(1)
    DO 26 I1=2,K0
    IF(DIST(I1).LT.AMAX) GO TO 26
    AMAX=DIST(I1)
26 CONTINUE
    AMIN=DIST(1)
    KC=1
    DO 27 I1=2,K0
    IF(DIST(I1).GT.AMIN) GO TO 27
    AMIN=DIST(I1)
    KC=I1
27 CONTINUE
    BETA(M)=DIST(K)/AMAX
    IF(KC.EQ.K) GO TO 50
    NCF=NCF+1
    GO TO 51
50 NCC=NCC+1
    *
    *
    * CORREGO O REFERENTE.
    *
51 DO 28 I=1,M
28 REF(K,I)=BETA(M)*B(I)+(1.0-BETA(M))*REF(K,I)
    GO TO 30
    *
    *
    * FAZ O HISTOGRAMA DE BETA, DIVIDINDO EM NCC CLASSES
    *
31 CONTINUE
    WRITE(NOUT,501)MM
501 FORMAT(//.2X,'NUMERO TOTAL DE PONTOS CLASSIFICADOS =',I6)
    *WRITE(NOUT,500)NCC,NCE
500 FORMAT(//.2X,'CEPTOS =',I6,10X,'ERRAMOS =',I4,////)
    NFN=NA
    DO 35 I=1,NH
55 GS(I)=BETA(I)
444 BMAX=BFIA(1)
    NVA=1
    DO 35 I=2,NH
    IF(BFIA(I).GT.BMAX)GO TO 35
    BMAX=BFIA(I)
    NVA=I
35 CONTINUE

```

```

BMAX2=BETA(1)
IF(NVF.DG.1)BMAX2=BETA(2)
DO 445 I=1,N
IF(I.EQ.NVA)GO TO 445
I=BETA(1)+I*(BMAX2)GO TO 445
BMAX2=BETA(1)
445 CONTINUE
VA=BMAX-BMAX2
IF(VA.GT..000)GO TO 442
GO TO 443
442 BETA(BVA)=A.0
N=N-1
GO TO 444
443 DO 33 I=1,NCL
33 VC(I)=0

I=0
447 I=I+1
IF(BETA(I).EQ.0.)GO TO 447
IF(BETA(I).EQ.BMAX)GO TO 36
K=NCL
GO TO 37
36 K=INDEX(GO(I)+NCL/BMAX)+1
37 NC(K)=NC(K)+1
IF(I.EQ.N)GO TO 447
32 CONTINUE
DO 34 J=0,NCL-1
AL(J+1)=J*(BMAX/NCL)
AS(J+1)=(J+1)*(BMAX/NCL)
FR(I+1)=AL(J+1)/FLOI*(N)
34 CONTINUE

CABCOULA:
G1=(Y1*Y2*...*YN)**(1./N)
G2=((1-Y1)*(1-Y2)*...*(1-YN))**((1./N))
PARK=POISSON DE "KAMA" FERUSSIMIL(UN)
461 G1=0.0+00
G2=0.0+00
DO 53 I=1,N
IF(BETA(1).EQ.0.)GO TO 53
IF(BETA(I).EQ.1.)BETA(I)=.999999
G1=G1+(LOG(BETA(I)))
G2=G2+DLOG(1.-BETA(I))
53 CONTINUE
G1=G1/(FLOI*(N))
G2=G2/(FLOI*(N))
G1=EXP(G1)
G2=EXP(G2)
G1=SQRT(G1)
G2=SQRT(G2)
998 FORMAT(25,'LOGO DE FORTALE PALE =',I?)
997 FORMAT(//,12X,'MOTIVADO'S',//,(1X,10E12.6))
996 FORMAT(//,12X,'CLASSE ',I?)
995 FORMAT(10(F8.3))
994 FORMAT(' CLASSE ',I1,2X,'PDMT',10(E8.2))
993 FORMAT(10(B18.6))
991 FORMAT(//,12X,'VETOR DE MEDIAS,')
990 FORMAT(//,12X,'MTRIE DE COVARIANCIA')

```



```

IF(F1.GE.1.0.AND.F3.GE.0.0)GO TO 114
A1=F3-F1
A2=F2+F3
GO TO 115
114 A1=F3-F1
A2=F2-F3
115 YP=F2-F1
IF(YF.NE.0.0)GO TO 117
A2=(A1+Z3)/2.
GO TO 118
117 A2=(A1+A2+A2+A2)/YP
A4=X3-F3/A3
F4=F1*(A4)
IF(X4.NE.0.0)GO TO 121
IF(X5.NE.0.0)GO TO 119
GO TO 122
121 IF(ABS((X4-X3)/YA).LE.EP)GO TO 116
122 F1=F2
F2=F3
F3=F4
A1=X2
A2=X3
A3=X4
IF(Y1.GE.0)GO TO 111
116 XDOF=X4
NS=F4
ZF=RCDF
CONTINUE
C
C
C RESOLVE A INTEGRAL, PELO METODO DE SIMPSON,
C DA FUNCAO DISTRIBUCIONAL BETA COM OS PARAMETROS ESTIMADOS
C NOS VARIOS INTERVALOS
C
PE(0)=P
XPE(0)=0
DO 201 I=1,ICI
PE(I)=P-I*.005*P
XPE(I)=XPE(I)+.005*XE
DO 202 J=1,2*ICI
XPE(ICI+J)=XPE(I)+.005*XE
n=(XPE(1)-XPE(0))/h1.
DO 203 I=1,ICL
DO 204 J=1,ICL
XIS(LL-1)=XIS(LL-1)+LL*n
YPS(LL-1)=YPS(LL-1)**(PE(I)-1.)*(1.-YPS(LL-1))**(XPE(J)-1.)
204 CONTINUE
DO 205 J=1,29
YPAR(0)=0.0
YIMP(0)=0.0
YPAR(L1)=YPS(5*L1)+YPAR(LL-1)
YIMP(L1)=YPS(2*L1-1)+YIMP(LL-1)
CONTINUE
YIMP(30)=YPS(50)+YIMP(29)
YPS(L)=4*(YPS(0)+YPS(10)+2*YPS(20)+4*YIMP(30))/3.
LUPA=LUPA+YPS(L)

```



```

213      CONTINUE
      *WRITE(NOUT,974)PR(1),QUT(J)
974      FORMAT(///,2X,'DISTRIBUICAO BRITA COM P =',F9.1,5X,'Q =',F9.1)
C
C      FAZ O TESTE QUI-QUADRADO DE FUNCAO DISTRIBUIDA COM
C      A FUNCAO DISTRIBUICAO BRITA(P,Q)
C
      IF(I.NE.0)GO TO 216
      IF(J.NE.0)GO TO 210
      *WRITE(NOUT,972)
972      FOR=AT(4X,'INTEGRAVAVELS',7X,'FREQ',2X,'FREQ',8X,'PREO',5X,
1'PROBAB',/,3X,'INF.',5X,'SUP.',5X,'A-S.',3X,'RST.',9X,'RGL.1)
216      Q=700 L=1,NCL
      IF(L)=RES(L)/TOTL
      PRO(L)=VEN*TT(L)
      NOV(L)=C(L)
      IF(I.NE.0)GO TO 206
      IF(J.NE.0)GO TO 206
      IPK(L)=IFIX(PRO(L)+.5)
      *WRITE(NOUT,973)AL(L),ES(L),NCL(L),IPK(L),PRO(L),TT(L)
973      FORMAT(/,3X,F5.1,3X,F5.1,7X,I3,4X,I3,9X,F7.5,3X,F7.5)
206      CONTINUE
      QUIT=.0
      L=NCL+1
207      U=U-1
210      IF(PRO(L).GT.5.0)GO TO 208
      QUITO=((NOV(L)-PRO(L))*2)/PRO(L)+QUIT
      IF(L.NE.12)GO TO 207
      L=0
844      L=L+1
843      IF(PRO(L).GT.5.0)GO TO 842
      QUITO=((NOV(L)-PRO(L))*2)/PRO(L)+QUITO
      IF(L.NE.11)GO TO 843
      GO TO 209
842      L=L+1
      NGL=NGL-1
      PRO(L)=PRO(L-1)+PRO(L)
      NOV(L)=NOV(L-1)+NOV(L)
      GO TO 843
208      L=L-1
      NGL=NGL-1
      PRO(L)=PRO(L+1)+PRO(L)
      NOV(L)=NOV(L+1)+NOV(L)
      GO TO 210
209      NGL=NGL-1
      *WRITE(NOUT,969)QUITO,NCL
969      FORMAT(/,2X,'QUT-QUADRADO OBS =',F9.3,3X,'COM',I6,2X,'G. L.1)
202      CONTINUE
      END

```

```

C
C      A SEGUIR AS SUBROTINAS UTILIZADAS
C      PARA ESTIMAR OS PARAMETROS P E Q
C      DA DISTRIBUICAO BRITA.

```

```

C
C      ESTIMA O PARAMETRO P

```

C

```

FUNCTION FID(YP)
  CD=MOD(Y2/P,C2,C1,XX
  EXTERRGRD F=1
  IF(YF.LD.0.)YP=.30000F1
  AX=PG(YP)
  XX=C1-C2+PG(YP)
  EPS=1.E-06
  RR=10.0
  AA=0.
  IF(AX.DF.1.1)AA=1./XX-.1
  EPS=1.E-06
  NZ=100
  XX1=AA
  IF(A.PD.0)DF=RR+1.
  XX3=FF
  XX2=(RR+AA)/2.
  FF1=PHI(XX1)
  FF3=PHI(XX3)
  FF2=PHI(XX2)
  IF(ABS(FF2).LE.XPS(FF1))GO TO 131
  YY=FF2
  FF2=FF1
  FF1=YY
  YY=XX2
  XX2=XX1
  XX1=YY
  IF(ABS(FF3).LE.XPS(FF2))GO TO 133
  IF(ABS(FF3).LE.ABS(FF1))GO TO 132
  YY=FF3
  FF3=FF1
  FF1=YY
  YY=XX3
  XX3=XX1
  XX1=YY
  IF(ABS(FF3).LE.ABS(FF2))GO TO 133
  YY=FF2
  FF2=FF3
  FF3=YY
  YY=XX2
  XX2=XX3
  XX3=YY
  NZ=NZ-1
  YY=XX3-XX1
  IF(YY.EQ.0.)GO TO 130
  ZZ=XX3-XX1
  IF(ZZ.EQ.0.)GO TO 136
  AA1=(FF3-FF2)/YY
  AA2=(FF3-FF1)/ZZ
  IF(FF1.LD.0.AND.FF3.LE.0.)GO TO 134
  IF(FF1.GD.0.AND.FF3.GD.0.)GO TO 134
  AA1=-FF3-FF1
  AA2=FF2+FF3
  GO TO 135
  AA1=FF3-FF1
  AA2=FF2-FF3
  YY=FF2-FF1
  IF(YY.EQ.0.)GO TO 137
  AS2=(XX1+XX3)/2.
  GO TO 132

```



```
SUBROUTINE PROD(Q,R,FILE,D)
DIMENSION Q(10,10),ELL(1,10),T(10,10)
DO 1 I=1,10
DO 1 J=1,10
SUMA=0.0
DO 2 K=1,10
SUMA=SUMA+Q(I,K)*ELL(K,J)
1 I(I,J)=SUMA
END
```

SUBROUTINE F01A(F, ATOL, A, IA, O, S, Z, IZ)

88

MARK 2 RELEASED. MAG COPYRIGHT 1979

EDITED BY JOYCE CLARK OXFORD OPG NUCLEAR PHYSICS 05TH NOV 1970

FORTRAN MACRO VERSION F01A26.TEC

MARK 3 REVISED

MARK 4.5 REVISED

MARK 5C REVISED

TRIDZ

THIS SUBROUTINE REDUCES THE GIVEN LOWER TRIANGLE OF A SYMMETRIC MATRIX, A, STORED IN THE ARRAY A(N,N), TO TRIANGULAR FORM USING HOUSEHOLDERS REDUCTION. THE DIAGONAL OF THE RESULT IS STORED IN THE ARRAY D(N) AND THE SUB-DIAGONAL IN THE LAST N-1 STORES OF THE ARRAY Z(N) (WITH THE ADDITIONAL ELEMENT Z(1) = 0). THE TRANSFORMATION MATRICES ARE ACCUMULATED IN THE ARRAY Z(N,N). THE ARRAY A IS LEFT UNALTERED UNLESS THE ACTUAL PARAMETERS CORRESPONDING TO A AND Z ARE IDENTICAL.

1ST AUGUST 1971

INTEGER I, IA, II, IZ, J1, J, K, L, N

REAL ATOL, F, G, H, HA, A(IA,N), D(N), Z(IZ,N)

DO 40 I=1,N

DO 20 J=1,I

Z(J,J) = A(I,J)

20 CONTINUE

40 CONTINUE

IF (N.EQ.1) GO TO 280

DO 200 II=2,N

I = N - II + 2

L = I - 2

F = Z(I,I-1)

G = 0.0

IF (L.EQ.0) GO TO 80

DO 60 K=1,L

G = G + Z(I,K)*Z(I,K)

60 CONTINUE

80 H = G + F*F

IF G IS TOO SMALL FOR ORTHOGONALITY TO BE GUARANTEED THE TRANSFORMATION IS SKIPPED

IF (G.GT.ATOL) GO TO 100

F(I) = F

H = 0.0

GO TO 240

100 L = L + 1

G = SQRT(H)

IF (F.GT.0.0) G = -G

Z(I) = G

H = H - F*G

Z(I,I-1) = F - G

F = 0.0

DO 180 J=1,L

Z(I,J) = Z(I,J)/H

H = 0.0

FOR EACH ELEMENT OF A*G

DO 120 K=1,J

G = G + Z(J,K)*Z(I,K)

120 CONTINUE

J1 = J + 1

IF (J1.GT.L) GO TO 100

```

      DO 110 K=1,L
      G = G + Z(K,J)*Z(I,K)
140    CONTINUE
C     FORM ELEMENT OF P
160    H(J) = G/H
      F = F + G*Z(J,I)
180    CONTINUE
C     FORM R
      HH = F/(H+L)
C     FORM REDUCED A
      DO 220 J=1,L
      F = Z(I,J)
      G = H(J) - HH*F
      H(J) = G
      DO 200 K=1,J
      Z(J,K) = Z(J,K) - F*H(K) - G*Z(I,K)
200    CONTINUE
220    CONTINUE
240    R(I) = H
260    CONTINUE
280    R(I) = 0.0
      D(I) = 0.0
C     ACCUMULATION OF TRANSFORMATION MATRICES
      DO 400 I=1,L
      E = I - 1
      IF (E(1).EQ.0.0) GO TO 360
      DO 340 J=1,E
      G = 0.0
      DO 300 K=1,L
      G = G + Z(I,K)*Z(K,J)
300    CONTINUE
      DO 320 K=1,E
      Z(I,J) = Z(K,J) - G*Z(K,I)
320    CONTINUE
340    CONTINUE
360    H(I) = Z(I,I)
      Z(I,I) = 1.0
      IF (L.EQ.1) GO TO 400
      DO 380 J=1,L
      Z(I,J) = 0.0
      Z(J,I) = 0.0
380    CONTINUE
400  CONTINUE
      RETURN
      END
SUBROUTINE T0ZAMP(D, ACORP, D, U, Z, I7, IFAIL)
MARK 2 RELEASE AND COPYRIGHT 1975
EDITED BY JOYCE CLARKE OXFORD ENG NUCLEAR PHYSICS 11TH SEP 1976
      PORTMAN HALL VERSION F01A24.TEC
MARK 3 REVISED
MARK 4 REVISED
MARK 4.5 REVISED
T0Z2
THIS SUBROUTINE FINDS THE EIGENVALUES AND EIGENVECTORS OF A
TRIANGULAR MATRIX T, GIVEN WITH ITS DIAGONAL ELEMENTS IN
THE ARRAY D(K) AND ITS SUBDIAGONAL ELEMENTS IN THE LAST N
- 1 SCANS OF THE ARRAY Z(K), USING QR TRANSFORMATIONS. THE
EIGENVALUES ARE OBTAINED BY THE DIAGONAL ELEMENTS IN THE
ARRAY D IN ASCENDING ORDER. THE EIGENVECTORS ARE FORMED IN

```

0000000000

FILE ARRAY Z(N,N), OVERWRITING THE ACCUMULATED
 TRANSFORMATIONS AS SUPPLIED BY THE SUBROUTINE FOIAJE. THE
 SUBROUTINE WILL FAIL IF ANY ONE EIGENVALUE TAKES MORE THAN 30
 ITERATIONS.
 151 APRIL 1972

INTEGER FOIAJE, ISAVE, IPABL, N, I, L, J, M, I1, N1, IJ, K, L2
 DOUBLE PRECISION SREAME

REAL R, F, H, ACHEPS, G, P, B, C, S, D(N), E(N), Z(L2,N)
 DIMENSION Z(999,999) /

ISAVE = IPABL
 IF (G.EQ.1) GO TO 30
 DO 20 I=2,N
 E(I-1) = E(I)
 20 CONTINUE

40 E(N) = 1.0
 B = 0.5
 F = 0.0

DO 300 L=1,N
 J = 0
 H = ZCHEPS*(ABS(D(L))+ABS(E(L)))
 IF (R.LT.H) H = H

LOOK FOR SMALL SUB-DIAG ELEMENT
 DO 60 M=L,N
 IF (ABS(B(M)).LE.F) GO TO 80
 60 CONTINUE

80 IF (M.EQ.L) GO TO 290
 100 IF (J.EQ.30) GO TO 400
 J = J + 1

FORM SHIFT
 G = B(L)
 H = B(L+1) - G
 IF (ABS(H).GE.ABS(E(L))) GO TO 120
 P = H*0.5/E(L)
 R = SQRT(P*P+1.0)
 U = P + R
 IF (F.LT.0.0) W = P - R
 D(L) = E(L)/W
 GO TO 140

120 D = 2.0*E(L)/H
 R = SQRT(P*P+1.0)
 D(L) = E(L)*P/(1.0+U)

140 H = G - B(L)
 I1 = L + 1
 IF (I1.GT.N) GO TO 190
 DO 160 I=I1,N
 D(I) = D(I) - H

160 CONTINUE
 180 F = F + H
 200 TRANSFORMATION

P = D(I)
 C = 1.0
 S = 0.0
 N1 = N - 1
 DO 280 IJ=L,N1
 I = I1 - IJ + L
 G = C*D(I)
 H = S*P
 IF (ABS(G).LT.ABS(P(I))) GO TO 200
 C = E(L)/P

```

      H = SQRT(C*C+1.0)
      E(I+1) = S*R*E
      S = C/P
      C = 1.0/P
      GO TO 220
200   C = P/E(I)
      R = SQRT(C*C+1.0)
      E(I+1) = S*R(I)*E
      S = 1.0/R
      C = C/R
220   P = C*D(I) - S*G
      D(I+1) = H + E*(C*G+D*D(I))
C     FORM VECTOR
      DO 240 K=1,N
          E = Z(K,I+1)
          Z(F,I+1) = S*Z(K,I) + C*H
          Z(F,I) = C*Z(K,I) - S*H
240   CONTINUE
260   CONTINUE
      H(L) = S*P
      D(L) = C*P
      IF (ABS(E(L)).GT.B) GO TO 100
280   D(L) = D(L) + P
300   CONTINUE
C     ORDER EIGENVALUES AND EIGENVECTORS
      DO 380 I=1,N
          K = I
          P = D(I)
          II = I + 1
          IF (II.GT.N) GO TO 340
          DO 320 J=II,N
              IF (D(J).GE.P) GO TO 320
              K = J
              P = D(J)
320   CONTINUE
340   IF (F.EQ.I) GO TO 380
          D(K) = D(I)
          D(I) = P
          DO 360 J=1,N
              E = Z(J,I)
              Z(J,I) = Z(I,E)
              Z(J,K) = P
360   CONTINUE
380   CONTINUE
      IFAIL = 0
      RETURN
400   IFAIL = PDZARF(ISAVAL,1,ORNAME)
      RETURN
      END
      INTEGER FUNCTION PDZARF(IFAIL, ERROR, SQUARE)
C     MARK 1 REVISED.  NAG COPYRIGHT 1971
C     MARK 3 REVISED
C     MARK 4A REVISED.  JUL-75
C     MARK 4.5 REVISED
C     MARK 7 REVISED (DEC 1978) (APR 1979)
C     RETURNS THE VALUE OF ERROR.  IF TERMINALS THE PROGRAM.
C     IF A HARD FAILURE OCCURS, THIS ROUTINE CALLS A FURTHER AUXILIARY
C     ROUTINE PDZAR2 WHICH GIVES A TRACE, A FAILURE MESSAGE AND HALTS
C     THE PROGRAM.
      INTEGER ERROR, IFAIL, BDD

```



```

C      DJOB,PRECISION,SAVE
C      TEST IF NO ERROR DETECTED
C      IF (ERROR.EQ.0) GO TO 20
C      DELIVER OUTPUT UNIT FOR MESSAGE
C      CALL XERRM (0,000)
C      TEST FOR SOFT FAILURE
C      IF (<0(1PMIL,10).EQ.1) GO TO 10
C      HARD FAILURE
C      WRITE (000,99999) ORGANE, SPGRP
C      STOPPING MECHANISM HAS BEEN TRIPPED
C      CALL PA112Z (A)
C      STOP
C      SOFT FAIL
C      TEST IF ERROR MESSAGE SUPPRESSED
10  IF (<0(1P/10,10).EQ.0) GO TO 20
C      WRITE (000,99999) ORGANE, SPGRP
20  P112Z = ERROR
C      RETURN
99999  FORMAT (15X, 'ERROR DETECTED BY MAG LIBRARY ROUTINE', A6,
C      * 114 - 1PMIL = , 15//)
C      EQU
C      SUBROUTINE P112Z
C      MARK 2 RELEASED BY TOM BRADNER AND JOYCE CLARKE OXFORD 00700000
C      MARK 6 REVISED.                                     00800000
C      CALL TRACE                                         00810000
C      STOP IF FAILURE DETECTED BY MAG LIBRARY ROUTINE' 02100000
C      EQU                                               02200000
C      AUTO EDIT 20 SEP 76
C      REAL FUNCTION X02AF(A)
C      MAG COPYRIGHT 1975
C      EDITED BY JOYCE CLARKE OXFORD OEG NUCLEAR PHYSICS 03RD OCT 1976
C      FORTRAN MACRO VERSION F01'26.TEC
C      MARK 4.5 RELEASED
C
C      * EPS *
C      RETURNS THE VALUE EPS WHEN EPS IS THE SMALLEST
C      POSITIVE
C      NUMBER SUCH THAT 1.0 + EPS > 1.0
C      THE X PARAMETER IS NOT USED
C      FOR ICB 1966
C      X02AF = 2.0**(-37.0)
C      REAL X
C      X02AF = *140400000000
C      RETURN
C      EQU
C      AUTO EDIT 17 OCT 76
C      REAL FUNCTION X02AF(X)
C      MAG COPYRIGHT 1975
C      EDITED BY JOYCE CLARKE OXFORD OEG NUCLEAR PHYSICS 03RD OCT 1976
C      FORTRAN MACRO VERSION F01'26.TEC
C      MARK 4.5 RELEASED
C
C      * I00 *
C      RETURNS THE RATIO OF THE SMALLEST POSITIVE REAL FLOATING-
C      POINT NUMBER REPRESENTABLE ON THE COMPUTER TO EPS
C      FOR ICB 1966
C      X02AF = 2.0**(-220.0)
C      REAL X
C      X02AF = *033400000000
C      RETURN

```

```

END
SUBROUTINE X02AAE(I, NERR)
MARK 7 RELEASED NAG COPYRIGHT 1978
EDITED BY JOYCE CLARKE OXFORD OEG NUCLEAR PHYSICS 05TH NOV 1976
FORTRAN MACRO VERSION F01A26.TEC
IF I = 0, SETS NERR TO CURRENT ERROR MESSAGE UNIT NUMBER
(STORED IN NERR1)
IF I = 1, CHANGES CURRENT ERROR MESSAGE UNIT NUMBER TO
VALUE SPECIFIED BY NERR.

```

```

*** NOTE ***

```

```

THIS ROUTINE ASSUMES THAT THE VALUE OF NERR1 IS SAVED
BETWEEN CALLS. IN SOME IMPLEMENTATIONS IT MAY BE
NECESSARY TO STORE NERR1 IN A LABELLED COMMON
BLOCK /X02AA/ TO ACHIEVE THIS.

```

```

.. SCALAR ARGUMENTS ..
INTEGER I, NERR

```

```

.. LOCAL SCALARS ..
INTEGER NERR1

```

```

DATA NERR1 /-1/
IF (I.EQ.0) NERR = NERR1
IF (I.EQ.1) NERR = NERR
RETURN
END

```

```

SUBROUTINE F02ABE(A, IA, N, R, V, IV, E, IFAIL)
MARK 2 RELEASED NAG COPYRIGHT 1972
EDITED BY JOYCE CLARKE OXFORD OEG NUCLEAR PHYSICS 11TH SEP 1976
FORTRAN MACRO VERSION F01A24.TEC
MARK 3 REVISED
MARK 4.5 REVISED

```

```

EIGENVALUES AND EIGENVECTORS OF A REAL SYMMETRIC MATRIX
1ST AUGUST 1971

```

```

INTEGER POINTE, ISAVE, IFAIL, N, IA, IV
DOUBLE PRECISION SRNAME
REAL TOL, XXXX, A(IA,N), R(N), V(IV,N), F(N), X02ADE, X02ABE
DATA SRNAME /OH F02ABE /
ISAVE = IFAIL
IFAIL = 1
TOL = X02ADE(XXXX)
CALL F01AJF(N, TOL, A, IA, R, E, V, IV)
TOL = X02ABE(XXXX)
CALL F02ABE(N, TOL, R, E, V, IV, IFAIL)
IF (IFAIL.NE.0) IFAIL = F01AAF(ISAVE, IFAIL, SRNAME)
RETURN
END

```

VECTOR DE MEDIAS

0.200000E+02 -0.100000E+02 0.320000E+02 0.140000E+02 -0.300000E+02

MATRIZ DE COVARIANCIAS

0.516961E+02 0.680775E+01 0.123143E+02 0.153909E+02 0.111201E+02
 0.800775E+01 0.765625E+02 0.122614E+02 0.153248E+02 0.220797E+02
 0.123143E+02 0.122614E+02 0.269361E+02 0.605984E+01 0.105616E+02
 0.153909E+02 0.153248E+02 0.605984E+01 0.940729E+02 0.959426E+01
 0.111201E+02 0.220797E+02 0.105616E+02 0.959426E+01 0.662596E+02

AUTOVALORES

.263379E+02 .443009E+02 .515866E+02 .774783E+02 .122343E+03

MATRIZ DE AUTOVECTORES

-0.314935E+00 0.688059E+00 0.577738E+00 0.871580E-02 -0.305834E+00
 -0.129718E+00 0.245557E+00 -0.634929E+00 -0.491519E+00 -0.527398E+00
 0.956251E+00 0.216751E+00 0.151300E+00 -0.125824E+00 -0.194243E+00
 0.261229E-01 -0.194807E+00 -0.107310E+00 0.728759E+00 -0.647119E+00
 -0.821059E-01 -0.617533E+00 0.478202E+00 -0.459795E+00 -0.414516E+00

CLASSE 1

RT 1(10.000, 20.000)
 RT 2(-4.000, 12.000)
 RT 3(28.000, 35.000)
 RT 4(20.000, 34.000)
 RT 5(-38.000, -25.000)

CLASSE 2

RT 1(20.000, 30.000)
 RT 2(-4.000, 12.000)
 RT 3(28.000, 35.000)
 RT 4(20.000, 34.000)
 RT 5(-38.000, -25.000)

CLASSE 3

RT 1(5.000, 20.000)
 RT 2(-16.000, -4.000)
 RT 3(21.000, 28.000)
 RT 4(7.000, 20.000)
 RT 5(-47.000, -36.000)

CLASSE 4

RT 1(20.000, 35.000)
 RT 2(-16.000, -4.000)
 RT 3(21.000, 28.000)
 RT 4(7.000, 20.000)
 RT 5(-47.000, -36.000)

CLASSE 5

RT 1(5.000, 20.000)
 RT 2(-16.000, -4.000)
 RT 3(28.000, 35.000)
 RT 4(7.000, 20.000)
 RT 5(-36.000, -25.000)

CLASSE 6
 RT 10 20.000, 35.000)
 RT 20 -16.000, -4.000)
 RT 30 28.000, 35.000)
 RT 40 7.000, 20.000)
 RT 50 -35.000, -25.000)

CLASSE 7
 RT 10 5.000, 20.000)
 RT 20 -16.000, -4.000)
 RT 30 35.000, 43.000)
 RT 40 7.000, 20.000)
 RT 50 -25.000, -13.000)

CLASSE 8
 RT 10 20.000, 35.000)
 RT 20 -16.000, -4.000)
 RT 30 35.000, 43.000)
 RT 40 7.000, 20.000)
 RT 50 -25.000, -13.000)

CLASSE 9
 RT 10 10.000, 20.000)
 RT 20 -28.000, -16.000)
 RT 30 28.000, 35.000)
 RT 40 -6.000, 7.000)
 RT 50 -36.000, -25.000)

CLASSE 10
 RT 10 20.000, 30.000)
 RT 20 -28.000, -16.000)
 RT 30 28.000, 35.000)
 RT 40 -6.000, 7.000)
 RT 50 -36.000, -25.000)

OS PONTOS INICIAIS ALFABETICOS

CLASSE 1	PONTO	11.948	0.731	32.753	33.280	-35.580
CLASSE 2	PONTO	27.325	2.275	30.739	21.638	-34.800
CLASSE 3	PONTO	14.131	-11.450	23.086	12.614	-40.683
CLASSE 4	PONTO	24.839	-9.540	22.273	10.475	-37.394
CLASSE 5	PONTO	0.627	-5.394	30.559	7.485	-20.290
CLASSE 6	PONTO	22.828	-11.087	28.949	13.108	-31.548
CLASSE 7	PONTO	5.926	-0.402	40.191	11.701	-17.020
CLASSE 8	PONTO	30.699	-10.769	37.447	12.703	-21.343
CLASSE 9	PONTO	10.379	-27.763	31.198	1.211	-32.828
CLASSE 10	PONTO	20.488	-20.396	30.583	3.711	-32.126

NUMERO TOTAL DE PONTOS CLASSIFICADOS = 2000

CELTOS = 1709

ERRADOS = 291

DISTRIBUCION NETA COM P = 6.1046

Q = 14.7630

INTERVALOS	FREQ.	FREQ.	FREQ.	PROBAB.	
INF. SUP.	ABS.	EST.	REL.		
.5000	.6249	0	0	.00000	.000001
.6249	.6497	1	1	.00050	.000025
.6497	.6746	5	4	.00250	.00197
.6746	.6995	15	14	.00750	.00712
.6995	.1243	35	34	.01750	.01713
.1243	.1492	61	64	.03150	.03185
.1492	.1740	96	99	.04900	.04961
.1740	.1989	139	135	.06950	.06774
.1989	.2238	154	167	.07700	.08346
.2238	.2486	198	189	.09400	.09455
.2486	.2735	261	206	.10250	.09978
.2735	.2984	188	198	.09400	.09901
.2984	.3232	174	196	.08700	.09297
.3232	.3481	187	166	.09250	.08300
.3481	.3729	139	141	.06950	.07060
.3729	.3978	120	115	.06000	.05754
.3978	.4227	100	90	.05000	.04481
.4227	.4475	66	67	.03300	.03347
.4475	.4724	54	46	.02700	.02393
.4724	.4973	28	35	.01400	.01637
.4973	.5221	21	21	.01050	.01070
.5221	.5470	17	13	.00950	.00667
.5470	.5718	4	8	.00200	.00395
.5718	.5967	3	4	.00150	.00222
.5967	.6216	2	2	.00100	.00110

QUI-QUADRADO OBS = 11.794 COM 10 G. T.

DISTRIBUICAO BETA COM P =	6.1046	Q =	14.6692
QU1-QUADRADO OBS =	11.452	COM	1º G. r.
DISTRIBUICAO BETA COM P =	6.1046	Q =	14.6153
QU1-QUADRADO OBS =	11.555	COM	1º G. r.
DISTRIBUICAO BETA COM P =	6.1046	Q =	14.5415
QU1-QUADRADO OBS =	12.105	COM	1º G. r.
DISTRIBUICAO BETA COM P =	6.1046	Q =	14.4677
QU1-QUADRADO OBS =	13.102	COM	1º G. r.
DISTRIBUICAO BETA COM P =	6.1046	Q =	14.3368
QU1-QUADRADO OBS =	12.582	COM	1º G. r.
DISTRIBUICAO BETA COM P =	6.1046	Q =	14.9106
QU1-QUADRADO OBS =	13.815	COM	1º G. r.
DISTRIBUICAO BETA COM P =	6.1046	Q =	14.9844
QU1-QUADRADO OBS =	15.495	COM	1º G. r.
DISTRIBUICAO BETA COM P =	6.1046	Q =	15.0582
QU1-QUADRADO OBS =	17.021	COM	1º G. r.
DISTRIBUICAO BETA COM P =	6.6741	Q =	14.7630
QU1-QUADRADO OBS =	12.537	COM	1º G. r.
DISTRIBUICAO BETA COM P =	6.6741	Q =	14.6692
QU1-QUADRADO OBS =	11.754	COM	1º G. r.

DISTRIBUICAO BETA COM P =	6.0741	Q =	14.6153
QU1-QUADRADO OBS =	11.418	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.0741	Q =	14.5415
QU1-QUADRADO OBS =	11.527	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.0741	Q =	14.4677
QU1-QUADRADO OBS =	12.083	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.0741	Q =	14.8368
QU1-QUADRADO OBS =	13.765	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.0741	Q =	14.9106
QU1-QUADRADO OBS =	15.439	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.0741	Q =	14.9844
QU1-QUADRADO OBS =	17.561	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.0741	Q =	15.0582
QU1-QUADRADO OBS =	20.353	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.0436	Q =	14.7630
QU1-QUADRADO OBS =	13.764	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.0436	Q =	14.6892
QU1-QUADRADO OBS =	12.541	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.0436	Q =	14.6153
QU1-QUADRADO OBS =	11.764	COM	18 G. T.

DISTRIBUICAO BETA COM P =	0.0436	Q =	14.5415
QU1-QUADRADO OBS =	11.433	COM	1º G. T.
DISTRIBUICAO BETA COM P =	0.0436	Q =	14.4677
QU1-QUADRADO OBS =	11.548	COM	1º G. T.
DISTRIBUICAO BETA COM P =	0.0436	Q =	14.8368
QU1-QUADRADO OBS =	15.433	COM	1º G. T.
DISTRIBUICAO BETA COM P =	0.0436	Q =	14.9106
QU1-QUADRADO OBS =	17.760	COM	1º G. T.
DISTRIBUICAO BETA COM P =	0.0436	Q =	14.9844
QU1-QUADRADO OBS =	20.316	COM	1º G. T.
DISTRIBUICAO BETA COM P =	0.0436	Q =	15.0582
QU1-QUADRADO OBS =	23.522	COM	1º G. T.
DISTRIBUICAO BETA COM P =	0.0130	Q =	14.7630
QU1-QUADRADO OBS =	15.674	COM	1º G. T.
DISTRIBUICAO BETA COM P =	0.0130	Q =	14.6892
QU1-QUADRADO OBS =	13.612	COM	1º G. T.
DISTRIBUICAO BETA COM P =	0.0130	Q =	14.6153
QU1-QUADRADO OBS =	12.596	COM	1º G. T.
DISTRIBUICAO BETA COM P =	0.0130	Q =	14.5415
QU1-QUADRADO OBS =	11.823	COM	1º G. T.

DISTRIBUICAO BETA COM P = 0.0130 Q = 14.4677
 QU1-QUADRADO OBS = 11.497 COM 18 G. T.

DISTRIBUICAO BETA COM P = 0.0130 Q = 14.8368
 QU1-QUADRADO OBS = 17.778 COM 19 G. T.

DISTRIBUICAO BETA COM P = 0.0130 Q = 14.9106
 QU1-QUADRADO OBS = 20.329 COM 19 G. T.

DISTRIBUICAO BETA COM P = 0.0130 Q = 14.9844
 QU1-QUADRADO OBS = 23.331 COM 19 G. T.

DISTRIBUICAO BETA COM P = 0.0130 Q = 15.0582
 QU1-QUADRADO OBS = 26.783 COM 19 G. T.

DISTRIBUICAO BETA COM P = 5.9825 Q = 14.7630
 QU1-QUADRADO OBS = 17.846 COM 19 G. T.

DISTRIBUICAO BETA COM P = 5.9825 Q = 14.6892
 QU1-QUADRADO OBS = 15.747 COM 19 G. T.

DISTRIBUICAO BETA COM P = 5.9825 Q = 14.6153
 QU1-QUADRADO OBS = 14.696 COM 19 G. T.

DISTRIBUICAO BETA COM P = 5.9825 Q = 14.5415
 QU1-QUADRADO OBS = 12.891 COM 19 G. T.

DISTRIBUICAO BETA COM P = 5.9825 Q = 14.4677
 QU1-QUADRADO OBS = 11.931 COM 18 G. T.

DISTRIBUICAO BETA COM P =	5.9825	Q =	14.8368
QUI-QUADRADO OBS =	20.394	COM	19 G. T.
DISTRIBUICAO BETA COM P =	5.9825	Q =	14.9106
QUI-QUADRADO OBS =	23.391	COM	19 G. T.
DISTRIBUICAO BETA COM P =	5.9825	Q =	14.9844
QUI-QUADRADO OBS =	26.839	COM	19 G. T.
DISTRIBUICAO BETA COM P =	5.9825	Q =	15.0582
QUI-QUADRADO OBS =	30.740	COM	19 G. T.
DISTRIBUICAO BETA COM P =	6.1351	Q =	14.7630
QUI-QUADRADO OBS =	11.535	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.1351	Q =	14.6692
QUI-QUADRADO OBS =	11.030	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.1351	Q =	14.6153
QUI-QUADRADO OBS =	12.177	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.1351	Q =	14.5415
QUI-QUADRADO OBS =	13.166	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.1351	Q =	14.4677
QUI-QUADRADO OBS =	14.008	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.1351	Q =	14.8368
QUI-QUADRADO OBS =	11.583	COM	18 G. T.

DISTRIBUICAO BETA COM P =	0.1351	Q =	14.9106
QU1-QUADRADO OBS =	12.076	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1351	Q =	14.9844
QU1-QUADRADO OBS =	13.914	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1351	Q =	15.0582
QU1-QUADRADO OBS =	15.599	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1656	Q =	14.7630
QU1-QUADRADO OBS =	11.760	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1656	Q =	14.6892
QU1-QUADRADO OBS =	12.299	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1656	Q =	14.6153
QU1-QUADRADO OBS =	13.284	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1656	Q =	14.5415
QU1-QUADRADO OBS =	14.717	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1656	Q =	14.4677
QU1-QUADRADO OBS =	16.000	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1656	Q =	14.3938
QU1-QUADRADO OBS =	11.067	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1656	Q =	14.9106
QU1-QUADRADO OBS =	12.020	COM	18 G. r.

DISTRIBUICAO BETA COM P =	0.1056	Q =	14.9844
QU1-QUADRADO OBS =	12.819	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1056	Q =	15.0582
QU1-QUADRADO OBS =	14.063	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1962	Q =	14.7630
QU1-QUADRADO OBS =	12.469	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1962	Q =	14.6892
QU1-QUADRADO OBS =	13.449	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1962	Q =	14.6153
QU1-QUADRADO OBS =	14.877	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1962	Q =	14.5415
QU1-QUADRADO OBS =	16.754	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1962	Q =	14.4677
QU1-QUADRADO OBS =	19.081	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1962	Q =	14.3368
QU1-QUADRADO OBS =	11.936	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1962	Q =	14.9106
QU1-QUADRADO OBS =	11.849	COM	18 G. r.
DISTRIBUICAO BETA COM P =	0.1962	Q =	14.9844
QU1-QUADRADO OBS =	12.207	COM	18 G. r.

DISTRIBUICAO BETA COM P =	6.1962	Q =	15.0582
QUI-QUADRADO OBS =	13.011	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.2267	Q =	14.7630
QUI-QUADRADO OBS =	13.664	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.2267	Q =	14.6892
QUI-QUADRADO OBS =	15.088	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.2267	Q =	14.6153
QUI-QUADRADO OBS =	16.957	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.2267	Q =	14.5415
QUI-QUADRADO OBS =	19.279	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.2267	Q =	14.4677
QUI-QUADRADO OBS =	22.653	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.2267	Q =	14.8366
QUI-QUADRADO OBS =	12.689	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.2267	Q =	14.9106
QUI-QUADRADO OBS =	12.161	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.2267	Q =	14.9844
QUI-QUADRADO OBS =	12.079	COM	18 G. T.
DISTRIBUICAO BETA COM P =	6.2267	Q =	15.0582
QUI-QUADRADO OBS =	12.443	COM	18 G. T.

BIBLIOGRAFIA

- BAKHVALOV, N.S. (1977). Numerical Methods. MIR Publishers.
- BECKMAN, R.J. e TIETJEN, G.L. (1978). Maximum Likelihood Estimation for the Beta Distribution. J. Statist. Comput. Simul., 7, 253-258.
- DAVIS, P.J. e RABINOWITZ, P. (1975). Methods of Numerical Integration. Academic Press.
- FISZ, M. (1963). Probability Theory and Mathematical Statistics. 3^a edition. John Wiley & Son.
- FUKUNAGA, K. (1972). Introduction to Statistical Pattern Recognition. Academic Press.
- HUMPHREYS, L.G. (1939). Acquisition and Extinction of Verbal Expectations in a Situation Analogous to Conditioning. Journal of Experimental Psychology, 25(3), 294-301.
- IOSIFESCU, M. e THEODORECU, R. (1969). Random Processes and Learning. Springer-Verlag.

LUKE, Y.L. (1969). The Special Functions and their Approximations.
Academic Press.

RATHIE, P.N. e ROHRER, H.G. (1983). The Exact Distribution of Products of Random Variables (submetido a publicação).

ROHRER, H.G. (1975). A Supervised Network of Adaptive Automata for Pattern Recognition - Arbeitsberichte des IMMD, Band 8, nº 8, 1-41.

ROHRER, H.G. (1978). Processos de Aprendizagem. Atas do 3º SINAPE, 153-159.

ROHRER, H.G. (1980). Sobre a Construção de Operadores em Processos de Aprendizagem para Classificação Automática. Atas do 4º SINAPE, 291-299.