

UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE MATEMÁTICA

f -ENTROPIAS GENERALIZADAS E A PROBABILIDADE DE ERRO

Aquiles Leite Nascimento

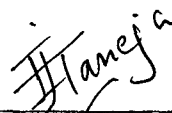
Setembro

1983

ESTA TESE FOI JULGADA ADEQUADA PARA A OBTENÇÃO DO TÍTULO DE

"MESTRE EM CIÊNCIAS"

ESPECIALIDADE EM MATEMÁTICA E APROVADA EM SUA FORMA FINAL
PELO CURSO DE PÓS-GRADUAÇÃO.



Prof. Inder Jeet Taneja, Ph.D.
Coordenador

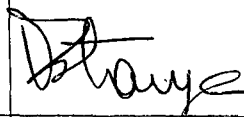
BANCA EXAMINADORA:



Prof. Inder Jeet Taneja, Ph.D.
Orientador



Prof. Annibal Parracho Santanna, Dr.



Prof. Plínio Stange, Dr.

DEDICATÓRIA

Dedico este trabalho para meus pais João Pedro e
Mariana

para meus irmãos

Ana Maria

Diomedes

Sebastião

Neuza

Alexandrino

Joamil

Sebastiana

Joacemil

Joary

Mariene

Deomildes

Juramil

pelo esforço e pela dedicação que sempre consagraram à minha
formação.

AGRADECIMENTOS

Ao professor Inder Jeet Taneja pela escolha do assunto de pesquisa, orientação capaz, eficiente e incansável, sem a qual este trabalho não seria realizado.

Aos professores Plínio Slange e Annibal Parracho Santanna, que compuseram a banca examinadora, pelas críticas e sugestões.

A todos os meus professores, principais responsáveis pela formação que tenho.

Aos colegas de turma: Zildo, Sanches e Marcelo, pelo incentivo e apoio durante o curso.

Aos demais amigos, colegas e funcionários do Departamento de Matemática da U.F.S.C..

Ao Curso de Pós-Graduação em Matemática da Universidade Federal de Santa Catarina.

À Universidade Federal de Mato Grosso pela oportunidade que me concedeu.

RESUMO

Neste trabalho apresentaremos as cotas superior e inferior da probabilidade de erro através de uma família de funções satisfazendo a propriedade da soma, sendo que essas funções são contínuas e existem as suas segundas derivadas.

Os casos particulares desta família são entropia de Shannon, entropia de grau β , α -log entropia e entropia de grau (α, β) . Também, apresentaremos as cotas superiores da probabilidade de erro em função de algumas destas entropias, levando em consideração a propriedade recursiva.

ABSTRACT

In this work we present upper and lower bounds to Probability of Error by a family of functions satisfying the sum property and the functions considered are continuous and exist second derivative.

The particular cases are: Shannon's entropy, Quadratic entropy, entropy of degree β , α -log entropy and entropy of degree (α, β) . Also, we present the upper bounds to probability of error in terms of some of these entropies taking into consideration the recursive property.

INTRODUÇÃO, OBJETIVO E IMPORTÂNCIA

Introdução

Kovalevski (1968) apresentou uma cota superior da probabilidade de erro em função do parâmetro t ,

$$H(C/X) \geq \log t + t(t+1) \cdot [\log(\frac{t+1}{t})] \cdot [P(e) - \frac{t-1}{t}] , \quad (1)$$

onde t é um inteiro satisfazendo a inequação

$$\frac{t-1}{t} \leq P(e) < \frac{t}{t+1} \quad (2)$$

e $H(C/X)$ é a entropia condicional (Shannon (1948) dada por

$$H(C/X) = \text{Ex} \left\{ - \sum_{i=1}^m P(C_i/X) \log P(C_i/X) \right\}. \quad (3)$$

Observação: Denotaremos $\log_2 b$ por $\log b$.

Já a cota inferior dada por Kovalevski (1968) é igual à cota de Fano, ou seja

$$H(C/X) \leq -P(e) \log P(e) - (1 - P(e)) \log (1 - P(e)) + P(e) \log (m-1). \quad (4)$$

Baseado na idéia de Kovalevski, Bem-Bassat (1978) estendeu os resultados para funções gerais satisfazendo a propriedade da soma e considerou

$$T(f) = \{H(P) / H: \Delta_m \rightarrow \mathbb{R}, m \geq 2, H(P) = \sum_{i=1}^m f(p_i), \text{ onde } f \text{ é estritamente côncava, } f'' \text{ existe e } f(0) = \lim_{p \rightarrow 0} f(p) = 0\}. \quad (5)$$

com Δ_m sendo o conjunto de todas as distribuições de probabilidades completas.

As cotas obtidas por Bem-Bassat (1978) são as seguintes:

$$H(P) \geq t \cdot f\left(\frac{1}{t}\right) + t(t+1) \left\{ (t+1) \cdot f\left(\frac{1}{t+1}\right) - t \cdot f\left(\frac{1}{t}\right) \right\} \cdot \left(p_e - \frac{t-1}{t}\right). \quad (6)$$

e

$$H(P) \leq f(1-p_e) + (m-1) \cdot f\left(\frac{p_e}{m-1}\right), \quad (7)$$

onde t é dado em (2) e

$$p_e = 1 - \max \{p_1, p_2, \dots, p_m\}.$$

Os casos particulares considerados por Bem-Bassat são entropia de Shannon, entropia quadrática (Vajda (1968)) e entropia de grau β (Daróczy (1970)) dadas por

$$H(P) = - \sum_{i=1}^m p_i \log p_i. \quad (9)$$

$$h_2(P) = \sum_{i=1}^m p_i (1 - p_i). \quad (10)$$

$$e \quad H^\beta(P) = (2^{1-\beta} - 1)^{-1} \cdot \left\{ \sum_{i=1}^m p_i^\beta - 1 \right\}, \quad \beta \neq 1, \quad \beta > 0, \quad (11)$$

respectivamente.

Sharma e Taneja (1977) estudaram as seguintes medi-

das generalizadas:

$$H_{\ell}^{\alpha}(P) = - 2^{\alpha-1} \cdot \sum_{i=1}^m p_i^{\alpha} \log p_i, \alpha > 0 \quad (12)$$

$$H^{(\alpha, \beta)}(P) = (2^{1-\alpha} - 2^{1-\beta})^{-1} \cdot \sum_{i=1}^m (p_i^{\alpha} - p_i^{\beta}), \alpha \neq \beta, \alpha, \beta > 0 \quad (13)$$

$$e \quad H_S^{(\alpha, \beta)}(P) = - \frac{2^{\alpha-1}}{\text{sen}\beta} \cdot \sum_{i=1}^m p_i^{\alpha} \text{sen}(\beta \log p_i), \alpha, \beta > 0, \quad (14)$$

$$\beta \neq n \cdot \pi, \quad n = 0, 1, 2, \dots$$

para todo $P = (p_1, p_2, \dots, p_m) \in \Delta_m$.

As medidas dadas em (12), (13) e (14) pertencem à família $T(f)$, desde que as funções abaixo

$$f(p) = - 2^{\alpha-1} p^{\alpha} \log p, \quad \alpha > 0 \quad (15)$$

$$f(p) = (2^{1-\alpha} - 2^{1-\beta})^{-1} (p^{\alpha} - p^{\beta}), \quad \alpha \neq \beta, \alpha, \beta > 0 \quad (16)$$

$$e \quad f(p) = - \frac{2^{\alpha-1}}{\text{sen}\beta} p^{\alpha} \text{sen}(\beta \log p), \quad \alpha, \beta > 0, \beta \neq n\pi, \quad (17)$$

$$n = 0, 1, 2, \dots$$

sejam côncavas no intervalo $[0, 1]$.

Observação: Usamos as seguintes convenções:

$$0 \cdot \log 0 = 0, \quad 0^{\alpha} = 0, \quad \alpha > 0 \quad e \quad 0 \cdot \text{sen}(\beta \log 0) = 0.$$

No capítulo II, apresentaremos cotas superior e inferior da probabilidade de erro em função das medidas dadas em (15) e (16) aplicando as técnicas de Ben-Bassat (1978).

No capítulo III, apresentaremos as cotas superiores da probabilidade de erro usando a bem conhecida propriedade re

cursiva das medidas (11), (12) e (13), enquanto, os casos (9) e (10) foram estudados por Hellman e RaviV (1970) e Vajda (1968), respectivamente.

Objetivo

O objetivo principal deste trabalho é estender os resultados obtidos por Ben-Bassat (1978) usando a propriedade da soma para outras medidas existentes, tais como α -log entropia, entropia de grau (α, β) , etc. e também obter, alternativamente, as cotas superiores da probabilidade de erro aplicando a propriedade recursiva de algumas dessas medidas.

Importância

As medidas estudadas por Sharma e Taneja em 1977 tais como α -log entropia, entropia de grau (α, β) e seno-entropia não tiveram nenhuma aplicação até agora e pela primeira vez foram aplicadas para obter cotas superior e inferior da probabilidade de erro e deixando em aberto os resultados ligados com seno-entropia. O objetivo de vários pesquisadores é obter a melhor aproximação da probabilidade de erro, descobrindo várias medidas de informação ou distância. E mais tarde pode ser provado que os resultados estudados têm sua aproximação melhor do que outros existentes, dependendo dos valores dos parâmetros.

Observaremos também que as propriedades da soma e da recursividade que serão aplicadas neste trabalho têm sua impor

tância maior na existência e caracterização das medidas de informação.

CAPÍTULO I

MEDIDAS DE INFORMAÇÃO E A PROBABILIDADE DE ERRO

Neste capítulo, apresentaremos várias medidas de informação, iniciando com a entropia de Shannon e dando suas generalizações. Definiremos o conceito de probabilidade de erro e também apresentaremos cotas superior e inferior da probabilidade de erro em função de algumas medidas de informação.

Seja

$$\Delta_m = \{P = (p_1, p_2, \dots, p_m) / p_i \geq 0, \sum_{i=1}^m p_i = 1\},$$

o conjunto de todas as distribuições de probabilidades.

1.1 - Medidas de Informação

i) Entropia de Shannon (Shannon (1948)).

É definida por

$$H(P) = - \sum_{i=1}^m p_i \log p_i, \quad (1.1)$$

para todo $P = (p_1, p_2, \dots, p_m) \in \Delta_m$.

ii) Entropia de ordem α (Rényi (1961))

É definida por

$$H_{\alpha}(P) = \frac{1}{1-\alpha} \log \left(\sum_{i=1}^m p_i^{\alpha} \right), \alpha \neq 1, \alpha > 0 \quad (1.2)$$

para todo $P = (p_1, p_2, \dots, p_m) \in \Delta_m$.

Esta entropia tende para a entropia de Shannon quando $\alpha \rightarrow 1$.

iii) Entropia quadrática (Vajda (1968))

É definida por

$$h_2(P) = \sum_{i=1}^m (p_i - p_i^2) = \sum_{i=1}^m p_i(1 - p_i), \quad (1.3)$$

para todo $P = (p_1, p_2, \dots, p_m) \in \Delta_m$.

iv) Entropia de grau β (Havrda e Charvát (1967) e Daróczy (1970)).

É definida por

$$H^{\beta}(P) = (2^{1-\beta} - 1)^{-1} \cdot \left(\sum_{i=1}^m p_i^{\beta} - 1 \right), \beta \neq 1, \beta > 0. \quad (1.4)$$

para todo $P = (p_1, p_2, \dots, p_m) \in \Delta_m$.

Esta entropia também tende para a entropia de Shannon quando $\beta \rightarrow 1$.

A entropia quadrática é caso particular desta entropia.

v) δ -Entropia (Arimoto (1971))

É definida por

$${}_{\delta}H(P) = (\delta - 1)^{-1} \cdot \left[\left(\sum_{i=1}^m p_i^{1/\delta} \right)^{\delta} - 1 \right], \quad \delta \neq 1, \delta > 0, \quad (1.5)$$

para todo $P = (p_1, p_2, \dots, p_m) \in \Delta_m$

É fácil ver que

$$\lim_{\delta \rightarrow 1} {}_{\delta}H(P) = H(P).$$

vi) Entropia de grau (α, β) (Sharma e Taneja (1977))

É definida por

$$H^{(\alpha, \beta)}(P) = (2^{1-\alpha} - 2^{1-\beta})^{-1} \cdot \sum_{i=1}^m (p_i^{\alpha} - p_i^{\beta}), \quad (1.6)$$

$$\alpha \neq \beta, \alpha, \beta > 0.$$

para todo $P = (p_1, p_2, \dots, p_m) \in \Delta_m$.

É fácil de verificar que:

$$H^{(\alpha, \beta)}(P) = \frac{A_{\alpha}}{A_{\alpha} - A_{\beta}} \cdot H^{\alpha}(P) + \frac{A_{\beta}}{A_{\beta} - A_{\alpha}} \cdot H^{\beta}(P), \quad \alpha \neq \beta, \quad (1.7)$$

onde $H^{\alpha}(P)$ é a entropia de grau α dada em (1.4) e

$$A_{\alpha} = 2^{1-\alpha} - 1, \quad \alpha \neq 1, \alpha > 0.$$

Também,

$$\lim_{\beta \rightarrow \alpha} H^{(\alpha, \beta)}(P) = H_{\alpha}^{\alpha}(P) = \left[2^{\alpha-1} \cdot \sum_{i=1}^m p_i^{\alpha} \log p_i \right], \quad \alpha > 0, \quad (1.8)$$

a qual se reduz para a entropia de Shannon quando $\alpha = 1$.

(1.8) é chamada de α -log entropia.

✓vii) Seno - Entropia (Sharma e Taneja (1977))

É definida por

$$H_S^{(\alpha, \beta)}(P) = -\frac{2^{\alpha-1}}{\text{sen } \beta} \cdot \sum_{i=1}^m p_i^\alpha \text{sen } (\beta \log p_i), \quad \alpha > 0, \quad \beta > 0, \\ \beta \neq n\pi, \quad n = 0, 1, 2, \dots \quad (1.9)$$

para todo $P = (p_1, p_2, \dots, p_m) \in \Delta_m$.

Neste caso também,

$$\lim_{\beta \rightarrow 0} H_S^{(\alpha, \beta)}(P) = H_\alpha^\alpha(P) = -2^{\alpha-1} \sum_{i=1}^m p_i^\alpha \log p_i, \quad \alpha > 0$$

Observação: Analogamente a (1.7) e (1.9), Aczél e Daróczy (1963) apresentaram as seguintes generalizações:

$$H_{(\alpha, \beta)}(P) = \frac{1}{\beta - \alpha} \log \left[\frac{\sum_{i=1}^m p_i^\alpha}{\sum_{i=1}^m p_i^\beta} \right], \quad \alpha \neq \beta, \quad \alpha, \beta > 0 \quad (1.10)$$

e

$$H_{(\alpha, \beta)}(P) = -\frac{1}{\alpha} \arctg \left[\frac{\sum_{i=1}^m p_i^\beta \text{sen } (\alpha \log p_i)}{\sum_{i=1}^m p_i^\beta \cos(\alpha \log p_i)} \right], \\ \alpha \neq 0. \quad (1.11)$$

1.2 - Probabilidade de Erro

Consideremos o problema da teoria da decisão de classificar uma observação X como vinda de m possíveis classes (h_i

hipóteses) $C = (C_1, C_2, \dots, C_m)$. Denotamos $p_i = \Pr \{C = C_i\}$, $i = 1, 2, \dots, m$ a probabilidade "a priori" das classes e $f_1(x), f_2(x), \dots, f_m(x)$, a função densidade condicional dada a verdadeira classe ou hipótese, isto é, $f_i(x) = \Pr \{X=x/C = C_i\}$, $i = 1, 2, \dots, m$. Suponhamos que $f_i(x)$ e p_i sejam completamente conhecidas. Dada qualquer observação $X = x$, podemos calcular "a posteriori" a probabilidade condicional de C pela regra de Bayes

$$P(C_i/x) = \frac{p_i f_i(x)}{\sum_{j=1}^m p_j f_j(x)}, \quad j = 1, 2, \dots, m.$$

É bem conhecido que a regra de decisão que minimiza a probabilidade de erro é a regra de decisão de Bayes, a qual associa x à classe que tem a maior probabilidade "a posteriori". Usando esta regra, a probabilidade de erro para um dado $X = x$ é expressa por

$$P(e/x) = 1 - \max \{P(C_1/x), P(C_2/x), \dots, P(C_m/x)\}.$$

Antes de observar X , a probabilidade de erro $P(e)$ associada a X é definida como a probabilidade de erro esperada após observá-lo, isto é,

$$P(e) = E_x \{1 - \max [P(C_1/x), P(C_2/x), \dots, P(C_m/x)]\}.$$

o que equivale a dizer

$$P(e) = 1 - \int_X \max [P(C_i/x)] p(x) dx.$$

1.3 - Equivocação e suas generalizações

Na secção (1.1) apresentamos várias medidas de informação. Nesta secção apresentaremos as medidas na sua forma generalizada em termos de probabilidades condicionais e as cotas superior e inferior da probabilidade de erro com suas medidas.

i) Equivocação de Shannon

É dada por:

$$H(C/X) = E_X \left[- \sum_{i=1}^m P(C_i/x) \log P(C_i/x) \right], \quad (1.12)$$

onde $P(C_i/x)$ representa a probabilidade de $C = C_i$ dado $X = x$.

Hellman e Raviv (1970) obtiveram a cota superior da probabilidade de erro dada por

$$P(e) \leq \frac{1}{2} H(C/X), \quad (1.13)$$

onde $H(C/X)$ é dado em (1.12). Por outro lado Chu e Chueh (1966), Kovalevski (1968) apresentaram a seguinte cota inferior para a probabilidade de erro

$$\begin{aligned} H(C/X) \leq & -P(e) \log P(e) - (1 - P(e)) \log (1 - P(e)) + \\ & + P(e) \log (m - 1) \end{aligned} \quad (1.14)$$

ii) Entropia Condicional de ordem α

É dada por

$$H^\alpha(C/X) = E_X [H^\alpha(C/x)], \quad (1.15)$$

onde

$$H^\alpha(C/X) = \frac{1}{1-\alpha} \log \sum_{i=1}^m P(C_i/x)^\alpha, \quad \alpha \neq 1, \alpha > 0 \quad (1.16)$$

Ben-Bassat e Raviv (1978) apresentaram as seguintes cotas superiores para a probabilidade de erro

$$P(e) \leq \frac{1}{2} H^\alpha(C/X), \quad 0 < \alpha \leq 2 \quad (1.17)$$

e

$$P(e) \leq \frac{1}{2} H^\alpha(C/X), \quad \alpha > 0, \text{ se } P(e/x) \geq \frac{1}{2} \quad (1.18)$$

Usando o resultado que

$H_\alpha(P) \leq H(P)$, $\alpha > 1$, podemos afirmar que as cotas (1.17) e (1.18) são melhores que (1.13).

Por outro lado, Toussaint (1978) apresentou a seguinte cota inferior:

$$H^\alpha(C/X) \leq (1-\alpha)^{-1} \log \{ (m-1)^{1-\alpha} \cdot P(e)^\alpha + (1-P(e)^\alpha) \}, \quad 0 < \alpha < 1, \quad (1.19)$$

a qual foi estendida por Taneja (1983b), para $\alpha \geq 1$.

Quando $\alpha = 2$, obtemos:

$$\frac{m-1}{m} \cdot [1 - \sqrt{\frac{mB(C/X) - 1}{m-1}}] \leq P(e), \quad (1.20)$$

iii) Entropia Condicional Quadrática

É dada por:

$$h(C/X) = E_X \left[\sum_{i=1}^m (P(C_i/x) - P(C_i/x)^2) \right] \quad (1.21)$$

Devijver (1974) definiu a seguinte medida, chamada

distância Bayesiana:

$$\begin{aligned} B(C/X) &= \int p(x) P(C_i/x)^2 dx \\ &= 1 - h(C/X), \end{aligned} \quad (1.22)$$

e também apresentou as seguintes desigualdades:

$$\begin{aligned} \frac{1}{2}(1 - B(C/X)) &\leq 1 - \sqrt{B(C/X)} \leq \frac{m-1}{m} \cdot \left[1 - \sqrt{\frac{mB(C/X) - 1}{m-1}}\right] \\ &\leq \frac{h(C/X)}{1 + \sqrt{1 - 2h(C/X)}} \leq Pe \leq h(C/X) \\ &\leq \frac{1}{2} H(C/X), \end{aligned} \quad (1.23)$$

onde um dos termos das desigualdades (1.23) coincide com (1.20)

iv) Entropia Condicional de grau β

É dada por

$$\begin{aligned} H^\beta(C/X) &= E_X \left[(2^{1-\beta} - 1)^{-1} \cdot \left(\sum_{i=1}^m P(C_i/x)^\beta - 1 \right) \right], \\ &\beta \neq 1, \beta > 0. \end{aligned} \quad (1.24)$$

Ben-Bassat (1978) apresentou a seguinte cota superior:

$$Pe \leq \frac{1}{2} H^\beta(C/X), \quad 0 < P(e/x) \leq \frac{1}{2}, \quad \beta > 0. \quad (1.25)$$

Devijver (1977) apresentou a seguinte cota inferior:

$$H^\beta(C/X) \leq (2^{1-\beta} - 1)^{-1} \cdot \left[(1 - P(e))^\beta + (m-1) \cdot \left(\frac{P(e)}{m-1} \right)^\beta - 1 \right], \quad \beta > 1, \quad (1.26)$$

a qual foi estendido por Taneja (1983a), para $0 < \beta < 1$.

v) Entropia -Condicional

É dada por

$$\delta H(C/X) = \text{Ex} \left\{ \frac{1}{\delta-1} \cdot \left[\left(\sum_{i=1}^m P(C_i/X)^{1/\delta} \right)^\delta - 1 \right] \right\},$$

$$\delta \neq 1, \delta > 0. \quad (1.27)$$

Taneja (1982) apresentou a seguinte cota superior:

$$P(e) \leq \frac{1}{2} \delta H(C/X), \quad \delta > 0 \quad (1.28)$$

Usando a desigualdade:

$$\delta H(P) \leq H_\alpha(P) \leq H(P), \quad 0 < \delta = \alpha^{-1} < 1, \quad (1.29)$$

podemos concluir que a cota dada em (1.28) é mais exata do que (1.13), (1.17) e (1.18).

Por outro lado, Boekee e Van der Lubbe (1980) apresentaram a seguinte cota inferior:

$$\delta H(C/X) \leq (\delta-1)^{-1} \cdot \left\{ 1 - \left[(1-P(e))^{1/\delta} + (m-1) \cdot \left(\frac{P(e)}{m-1} \right)^{1/\delta} \right]^\delta \right\}. \quad (1.30)$$

Boekee e Van der Lubbe (1980) também apresentaram a seguinte desigualdade

$$(1 - \delta) \cdot \delta H(C/X) \leq P(e) \leq 1 - \left[1 - (1 - \delta) \cdot \delta H(C/X) \right]^{1/\delta-1} \quad (1.31)$$

Podemos ver que as medidas apresentadas em (1.1), (1.3), (1.4), (1.6), (1.8) e (1.9) têm uma propriedade em comum, isto é,

$$H(P) = \sum_{i=1}^m f(p_i) \quad (1.32)$$

onde

$$f(p) = -p \log p,$$

$$f(p) = p(1 - p),$$

$$f(p) = (2^{1-\beta} - 1)^{-1} \cdot (p^\beta - p), \quad \beta \neq 1, \beta > 0,$$

$$f(p) = -2^{\alpha-1} p^\alpha \log p,$$

$$f(p) = (2^{1-\alpha} - 2^{1-\beta})^{-1} \cdot [p^\alpha - p^\beta], \quad \alpha \neq \beta, \alpha, \beta > 0,$$

$$f(p) = -\frac{2^{\alpha-1}}{\operatorname{sen} \beta} \cdot p^\alpha \operatorname{sen}(\beta \log p), \quad \alpha, \beta > 0,$$

$$\beta \neq n \cdot \pi, \quad n = 0, 1, 2, \dots,$$

respectivamente.

A propriedade (1.32) é chamada de propriedade da soma e que tem sua maior importância na caracterização destas medidas.

No capítulo II, apresentaremos as cotas superior e inferior da probabilidade de erro, que têm por base o trabalho de Ben-Bassat (1978), em termos da função f dada em (1.32), juntamente com outras condições.

Em casos particulares, Ben-Bassat (1978) considera somente as medidas (1.1), (1.3) e (1.4). Estenderemos o raciocínio para as medidas (1.6) e (1.8).

No capítulo III, apresentaremos as cotas superiores

aplicando principalmente a propriedade da recursividade para algumas dessas medidas.

A medida (1.3) é caso particular simples de (1.4) e o caso (1.1) foi estudado extensivamente por Hellman e Raviv (1970). Entretanto, não incluiremos o estudo destas medidas no capítulo III.

CAPÍTULO II

f-FAMÍLIA DE MEDIDAS E A PROBABILIDADE DE ERRO

Neste capítulo apresentaremos os resultados de Ben-Bassat (1978) o qual se baseia no trabalho de Kovalevski (1968) e aplicaremos os casos particulares para obter cotas superior e inferior da probabilidade de erro.

Na secção 2.1, introduziremos a família de medidas de informação às quais se aplicam os resultados de Ben-Bassat. Na secção 2.2, enunciaremos esses resultados. Na secção 2.3, aplicaremos os resultados para obter cotas superior e inferior para a probabilidade de erro, usando as entropias de Shannon, entropia quadrática, entropia de grau β , α -log entropia e entropia de grau (α, β) .

2.1 - f-Família de Medidas

Define-se uma família de medidas de informação $T(f)$ por

$$T(f) = \{H(P)/H: \Delta_m \rightarrow \mathbb{R}, m \geq 2, H(P) = \sum_{i=1}^m f(p_i),$$

onde f é côncava, f'' existe e

$$f(0) = \lim_{p \rightarrow 0} f(p) = 0 \}, \quad (2.1)$$

Δ_m definido como no capítulo I.

Por simplicidade, usaremos as seguintes notações:

$$\Delta_m(p_e) = \{P = (p_1, p_2, \dots, p_m) \in \Delta_m / 1 - \max \{p_1, p_2, \dots, p_m\} = p_e\} \quad (2.2)$$

$$\bar{h}(p_e) = \sup_P \{H(P) / P \in \Delta_m(p_e)\} \quad (2.3)$$

$$\underline{h}(p_e) = \inf_P \{H(P) / P \in \Delta_m(p_e)\} \quad (2.4)$$

$$\bar{H}(P(e)) = \sup_X \{H(X) / X \in F, E_X \{P(e/x)\} = P(e)\} \quad (2.5)$$

$$\underline{H}(P(e)) = \inf_X \{H(X) / X \in F, E_X \{P(e/x)\} = P(e)\}, \quad (2.6)$$

onde F é o conjunto de todas as características observáveis e cada membro de F é representado por variáveis aleatórias.

2.2 - Resultados Principais

Citaremos aqui os teoremas e os lemas que foram considerados por Ben-Bassat (1978). Suas demonstrações podem ser vistas no apêndice ao final deste trabalho, as quais foram feitas com mais detalhes.

Teorema 2.1:

Para todo $H(P) \in T(f)$ e para um dado p_e , $0 \leq p_e \leq 1 - \frac{1}{m}$, $m \geq 2$, $\bar{h}(p_e)$ é atingido quando toda componente de P , exceto talvez uma, é igual a $\frac{p_e}{m-1}$ e, então,

$$\bar{h}(p_e) = f(1 - p_e) + (m - 1) \cdot f\left(\frac{p_e}{m-1}\right). \quad (2.7)$$

Teorema 2.2:

Para todo $H(P) \in T(f)$ e para todo p_e , $0 \leq p_e \leq 1 - \frac{1}{m}$, $m \geq 2$, $\underline{h}(p_e)$ é atingido quando toda componente de P , exceto talvez uma, é igual a $1 - p_e$ ou 0 (zero) e, então

$$\underline{h}(p_e) = t \cdot f(1 - p_e) + f[1 - t(1 - p_e)], \quad (2.8)$$

onde t é um número inteiro determinado pelas inequações:

$$\frac{t-1}{t} \leq p_e < \frac{t}{t+1}. \quad (2.9)$$

Lema 2.1:

Para todo $H(P) \in T(f)$, $\underline{h}(p_e)$ é uma função convexa sobre o intervalo $[0, 1 - \frac{1}{m}]$, $m \geq 2$.

Lema 2.2:

Para todo $H(P) \in T(f)$ e para qualquer número inteiro t tal que $t \leq m - 1$, $m \geq 2$, $\underline{h}(p_e)$ é estritamente côncava sobre o intervalo $[\frac{t-1}{t}, \frac{t}{t+1}]$.

Lema 2.3:

Para todo $H(P) \in T(f)$, $L(p_e)$ é uma função convexa ponto a ponto que liga os pontos

$$\underline{h}\left(\frac{t-1}{t}\right) \text{ e } \underline{h}\left(\frac{t}{t+1}\right), \quad t = 1, 2, \dots, m-1,$$

onde $L(p_e)$ é dada por

$$\begin{aligned}
L(p_e) &= \underline{h}\left(\frac{t-1}{t}\right) + \frac{\underline{h}\left(\frac{t}{t+1}\right) - \underline{h}\left(\frac{t-1}{t}\right)}{\frac{t}{t+1} - \frac{t-1}{t}} \cdot \left(p_e - \frac{t-1}{t}\right) \\
&= t \cdot f\left(\frac{1}{t}\right) + t(t+1) \cdot \left[(t+1) \cdot f\left(\frac{1}{t+1}\right) - \right. \\
&\quad \left. - t f\left(\frac{1}{t}\right) \right] \left(p_e - \frac{t-1}{t}\right). \tag{2.10}
\end{aligned}$$

Teorema 2.3:

Para todo $H(P) \in T(f)$ e para qualquer valor dado $P(e)$, $0 \leq P(e) \leq 1 - \frac{1}{m}$, $m \geq 2$, $\bar{H}(P(e)) = \bar{h}(P(e))$, $\bar{h}(P(e))$ é atingido quando toda a probabilidade de erro é igual a $P(e)$.

Teorema 2.4:

Para todo $H(P) \in T(f)$ e para qualquer valor dado $P(e)$, $0 < P(e) \leq 1 - \frac{1}{m}$, $m \geq 2$,

$$\underline{H}(P(e)) = L(P(e)), \tag{2.11}$$

onde $L(P(e))$ é dado em (2.10).

2.3 - Aplicações

A seguir, daremos alguns dos membros da família definida em (2.1) com suas respectivas concavidades e com aplicações em probabilidade de erro.

2.3.1 - Entropia de Shannon

$$\text{Seja } f(p) = -p \log p, p \in [0,1], \text{ com } 0 \cdot \log 0 = 0 \quad (2.12)$$

Então,

$$H(P) = \sum_{i=1}^m f(p_i) = - \sum_{i=1}^m p_i \log p_i, \quad (2.13)$$

a qual é a entropia de Shannon.

Concavidade

$$\text{Temos } f'(p) = -\log e - \log p$$

$$f''(p) = -\frac{1}{p} \log e$$

Portanto, $-\frac{1}{p} \cdot \log e \leq 0$, para $p \in (0,1]$. Isto implica que $f(p)$ é uma função côncava em $[0,1]$ pois $0 \cdot \log 0 = 0$ ou seja $H(P)$ é uma função côncava em Δ_m .

Portanto, $H(P) \in T(f)$.

Cotas superior e inferior

Aplicando os teoremas 2.1 e 2.3, temos

$$\bar{H}(P(e)) = -P(e) \log P(e) - (1 - P(e)) \log (1 - P(e)) + P(e) \log (m-1)$$

isto é,

$$\begin{aligned} H(C/X) \leq & -P(e) \log P(e) - (1 - P(e)) \log (1 - P(e)) + \\ & + P(e) \log (m - 1). \end{aligned} \quad (2.14)$$

Pelo teorema 2.4, temos

$$\underline{H}(P(e)) = \log t + t(t+1) \log \frac{t+1}{t} \left(P(e) - \frac{t-1}{t} \right), \quad (2.15)$$

onde t é definido em (2.9).

Kovalevski (1968) observou que no caso $0 < P(e) \leq \frac{1}{2}$, o parâmetro t atinge o valor 1 e então $\underline{H}(P(e))$ se reduz para $\underline{H}(P(e)) = 2P(e)$.

Isto implica na seguinte cota superior

$$P(e) \leq \frac{1}{2} H(C/X), \quad (2.16)$$

a qual foi estudada por Hellman e Raviv em 1970, para qualquer valor de $P(e)$.

2.3.2 - Entropia quadrática

Seja $f(p) = p(1-p)$, $p \in [0,1]$, então

$$H(P) = \sum_{i=1}^m f(p_i) = \sum_{i=1}^m p_i(1-p_i), \quad (2.17)$$

Concavidade:

$$\text{Temos } f(p) = -p^2 + p$$

$$f'(p) = -2p + 1$$

$$\text{e } f''(p) = -2$$

Portanto, $f''(p) \leq 0$. Isto implica que $f(p)$ é côncava para todo p tal que $p \in [0,1]$.

Logo, $f(p)$ é côncava ou seja $H(P)$ é uma função côncava em Δ_m .

Portanto, $H(P) \in T(f)$.

Cotas inferior e superior

Aplicando os teoremas 2.1 e 2.3, temos

$$\begin{aligned}
 \bar{H}(P(e)) &= (1 - P(e)) \cdot P(e) + (m - 1) \cdot \left(\frac{P(e)}{m - 1}\right) \cdot \left(1 - \frac{P(e)}{m - 1}\right) \\
 &= (1 - P(e)) \cdot P(e) + P(e) \cdot \left(1 - \frac{P(e)}{m - 1}\right) = \\
 &= P(e) + P(e) - P(e)^2 - \frac{P(e)}{m - 1} \\
 &= 2 \cdot P(e) - \left[P(e)^2 \cdot \left(1 + \frac{1}{m - 1}\right)\right] \\
 &= 2 \cdot P(e) - P(e)^2 \cdot \left(\frac{m - 1 + 1}{m - 1}\right) \\
 &= \left(-\frac{m}{m - 1}\right) \cdot P(e)^2 + 2P(e).
 \end{aligned}$$

$$\text{Portanto, } \bar{H}(P(e)) = \left(-\frac{m}{m - 1}\right) P(e)^2 + 2P(e),$$

$$\text{isto é, } h(C/X) \leq -\frac{m}{m - 1} P(e)^2 + 2P(e) \quad (2.18)$$

Pelo teorema 2.4, temos:

$$\begin{aligned}
 \underline{H}(P(e)) &= t \cdot \frac{1}{t} \left(1 - \frac{1}{t}\right) + t \cdot (t + 1) \cdot \left[\left(t + 1\right) \cdot \frac{1}{t + 1} \cdot \left(1 - \frac{1}{t + 1}\right) - \right. \\
 &\quad \left. - t \cdot \frac{1}{t} \left(1 - \frac{1}{t}\right)\right] \cdot \left(P(e) - \frac{t - 1}{t}\right) \\
 &= \frac{t - 1}{t} + t \cdot (t + 1) \cdot \left[\frac{t + 1 - 1}{t + 1} - \frac{t - 1}{t}\right] \cdot \left(P(e) - \frac{t - 1}{t}\right) \\
 &= \frac{t - 1}{t} + t \cdot (t + 1) \cdot \left[\frac{t^2 - t^2 + 1}{t(t + 1)}\right] \cdot \left(P(e) - \frac{t - 1}{t}\right)
 \end{aligned}$$

$$= \frac{t-1}{t} + \frac{t \cdot (P(e)) - t + 1}{t} = P(e). \quad (2.19)$$

Portanto, $H(P(e)) = P(e)$.

Para uma dada variável X , (2.18) e (2.19), obtemos:

$$P(e) \leq h(C/X) \leq -\frac{m}{m-1} P(e)^2 + 2P(e).$$

Resolvendo a inequação do lado direito, isto é,

$$h(C/X) \leq -\frac{m}{m-1} P(e)^2 + 2P(e), \text{ temos:}$$

$$\frac{h(C/X)}{2} \leq -\frac{m}{2(m-1)} \cdot P(e)^2 + \frac{2P(e)}{2} \rightarrow$$

$$\rightarrow P(e) - \frac{m}{2(m-1)} P(e)^2 \geq \frac{h(C/X)}{2}$$

$$- \frac{2mP(e)}{m-1} + \frac{m^2 P(e)^2}{(m-1)^2} \leq \frac{mh(C/X)}{m-1}$$

$$1 - \frac{2mP(e)}{m-1} + \frac{m^2 P(e)^2}{(m-1)^2} \leq 1 - \frac{mh(C/X)}{m-1}$$

$$\left(1 - \frac{mP(e)}{m-1}\right)^2 \leq 1 - \frac{mh(C/X)}{m-1}$$

$$1 - \frac{mP(e)}{m-1} \leq \sqrt{1 - \frac{mh(C/X)}{m-1}}$$

$$- \frac{mP(e)}{m-1} \leq -1 + \sqrt{1 - \frac{mh(C/X)}{m-1}}$$

$$m \cdot P(e) \geq (m-1) \cdot \left[1 - \sqrt{1 - \frac{mh(C/X)}{m-1}}\right]$$

$$P(e) \geq \frac{m-1}{m} \cdot \left[1 - \sqrt{1 - \frac{mh(C/X)}{m-1}}\right],$$

a qual é dada em (1.20)

Mas, $h(C/X) = 1 - B(C/X)$, (veja (1.22))

$$\text{logo, } P(e) \geq \frac{m-1}{m} \left[1 - \sqrt{\frac{mB(C/X) - 1}{m-1}} \right].$$

Esta cota foi estudada pela primeira vez por Devijver (1974).

2.3.3 - Entropia de grau β

$$\text{Seja } f(p) = \frac{p^\beta - p}{2^{1-\beta} - 1}, \quad \beta \neq 1, \beta > 0,$$

$$\text{então } H(P) = \sum_{i=1}^m f(p_i) = (2^{1-\beta} - 1)^{-1} \cdot \left[\sum_{i=1}^m p_i^\beta - 1 \right], \quad (2.21)$$

$$\beta \neq 1, \beta > 0.$$

Concavidade:

$$\text{Temos } f'(p) = \frac{\beta p^{\beta-1} - 1}{2^{1-\beta} - 1}$$

$$\text{e } f''(p) = \frac{\beta \cdot (\beta-1) \cdot p^{\beta-2}}{2^{1-\beta} - 1}$$

$$\text{Portanto, } f''(p) \leq 0, \text{ se } \frac{\beta(\beta-1)p^{\beta-2}}{2^{1-\beta} - 1} \leq 0$$

1º caso: Quando $0 < \beta < 1$. Neste caso, $2^{1-\beta} - 1 > 0$
e $\beta - 1 < 0$, portanto, $f''(p) \leq 0$.

2º caso: Quando $\beta > 1$. Neste caso, $2^{1-\beta} - 1 < 0$
e $\beta - 1 > 0$, portanto, $f''(p) \leq 0$.

Quando $\beta = 1$, $f(p)$ se reduz para o caso 2.3.1, a qual é côncava em $[0,1]$.

Assim, $f(p)$ é uma função côncava em $[0,1]$, para todo $\beta > 0$, isto é, $H^\beta(P)$ é uma função côncava em Δ_m , $\beta > 0$.

Cotas superior e inferior

mos: Aplicando os teoremas 2.1 e 2.3, respectivamente, te

$$\begin{aligned}\bar{H}(P(e)) &= (2^{1-\beta} - 1)^{-1} \cdot [(1 - P(e))^\beta - (1 - P(e))] + \\ &+ (2^{1-\beta} - 1)^{-1} \cdot (m - 1) \cdot \left[\left(\frac{P(e)}{m-1} \right)^\beta - \frac{P(e)}{m-1} \right] \\ &= (2^{1-\beta} - 1)^{-1} \cdot \left\{ (1 - P(e))^\beta + (m-1) \cdot \left(\frac{P(e)}{m-1} \right)^\beta - 1 \right\}, \\ &\beta \neq 1, \beta > 0,\end{aligned}$$

i.é.,

$$\begin{aligned}H^\beta(\mathbf{c}/X) &\leq (2^{1-\beta} - 1)^{-1} \cdot \left\{ (1 - P(e))^\beta + (m-1) \left(\frac{P(e)}{m-1} \right)^\beta - 1 \right\}, \\ &\beta \neq 1, \beta > 0. \quad (2.22)\end{aligned}$$

Pelo teorema 2.3, temos:

$$\begin{aligned}\underline{H}(P(e)) &= (2^{1-\beta} - 1)^{-1} \cdot [t^{1-\beta} - 1] + (2^{1-\beta} - 1) \cdot [t(t+1)^{2-\beta} \\ &- (t+1) \cdot t^{2-\beta}] \cdot \left(P(e) - \frac{t-1}{t} \right).\end{aligned}$$

Para $0 < P(e) \leq \frac{1}{2}$, t atinge o valor 1 e, então

$$\begin{aligned}\underline{H}(P(e)) &= (2^{1-\beta} - 1)^{-1} \cdot [1^{1-\beta} - 1 + 1 \cdot (1+1)^{2-\beta} - \\ &- (1+1) \cdot 1^{2-\beta}] \cdot \left(P(e) - \frac{1-1}{1} \right) \\ &= (2^{1-\beta} - 1)^{-1} \cdot (2^{2-\beta} - 2) \cdot P(e) = 2P(e),\end{aligned}$$

que aplica na seguinte cota superior:

$$P(e) \leq \frac{1}{2} H(X), \quad 0 < P(e) \leq \frac{1}{2}, \quad \text{isto é,}$$

$$P(e) \leq \frac{1}{2} H^\beta(C/X), \quad 0 < P(e) \leq \frac{1}{2}. \quad (2.23)$$

2.3.4 - α -log Entropia.

$$\text{Seja } f(p) = - 2^{\alpha-1} \cdot p^\alpha \log p, \quad \alpha > 0$$

então,
$$H(P) = \sum_{i=1}^m f(p_i)$$

$$= - 2^{\alpha-1} \cdot \sum_{i=1}^m p_i^\alpha \log p_i, \quad \alpha > 0 \quad (2.24)$$

Concavidade:

$$\text{Temos } f(p) = - 2^{\alpha-1} p^\alpha \log p$$

$$= - 2^{\alpha-1} p^\alpha \log_e p \cdot \log e$$

$$f'(p) = - 2^{\alpha-1} \cdot \log e (\alpha p^{\alpha-1} \log_e p + p^\alpha \cdot \frac{1}{p})$$

$$= - 2^{\alpha-1} \log e (\alpha p^{\alpha-1} \cdot \log_e p + p^{\alpha-1})$$

$$f''(p) = - 2^{\alpha-1} \log e (\alpha(\alpha-1)p^{\alpha-2} \cdot \log_e p +$$

$$+ \alpha p^{\alpha-1} \cdot \frac{1}{p} + (\alpha-1)p^{\alpha-2}).$$

$$= - 2^{\alpha-1} \log e \{ \alpha(\alpha-1)p^{\alpha-2} \log_e p + \alpha p^{\alpha-2} +$$

$$+ (\alpha-1)p^{\alpha-2} \}$$

$$= - 2^{\alpha-1} \log_e \{ \alpha(\alpha-1)p^{\alpha-2} \log_e p + (2\alpha-1)p^{\alpha-2} \} \leq 0,$$

se $1 - \alpha \geq 0$ e $2\alpha - 1 \geq 0$, isto é,

$$1 \geq \alpha \quad \text{e} \quad \alpha \geq \frac{1}{2}, \text{ isto é, } \frac{1}{2} \leq \alpha \leq 1.$$

Logo, $f(p)$ é uma função côncava em $[0,1]$, para todo $\frac{1}{2} \leq \alpha \leq 1$, isto é, $H_\ell^\alpha(P)$ é uma função côncava em Δ_m , para todo $\frac{1}{2} \leq \alpha \leq 1$.

Cotas superior e inferior

Aplicando os teoremas 2.1 e 2.3, respectivamente, temos:

$$\begin{aligned} \bar{H}(P(e)) &= - 2^{\alpha-1} (1 - P(e))^\alpha \log (1 - P(e)) - \\ &\quad - 2^{\alpha-1} (m-1) \cdot \left(\frac{P(e)}{m-1}\right)^\alpha \log \left(\frac{P(e)}{m-1}\right) \\ &= - 2^{\alpha-1} \cdot \{ (1 - P(e))^\alpha \log (1 - P(e)) + \\ &\quad + (m-1)^{1-\alpha} P(e)^\alpha \cdot \log \left(\frac{P(e)}{m-1}\right) \}, \quad \frac{1}{2} \leq \alpha \leq 1, \end{aligned}$$

isto é,

$$\begin{aligned} H_\ell^\alpha(\mathbb{C}/X) &\leq - 2^{\alpha-1} \{ (1 - P(e))^\alpha \log (1 - P(e)) + \\ &\quad + (m-1)^{1-\alpha} P(e)^\alpha \log \left(\frac{P(e)}{m-1}\right) \}, \quad \frac{1}{2} \leq \alpha \leq 1, \end{aligned} \tag{2.25}$$

Aplicando o teorema 2.4, temos:

$$\begin{aligned} \underline{H}(P(e)) &= - 2^{\alpha-1} \cdot \left\{ t \left(\frac{1}{t}\right)^\alpha \log \left(\frac{1}{t}\right) + t \cdot (t+1) \cdot \right. \\ &\quad \left. \cdot \left[(t+1) \left(\frac{1}{t+1}\right)^\alpha \log \left(\frac{1}{t+1}\right) - t \cdot \left(\frac{1}{t}\right)^\alpha \log \left(\frac{1}{t}\right) \right] \right\}. \end{aligned}$$

$$\begin{aligned}
& \cdot [P(e) - \frac{t-1}{t}] \} \\
& = - 2^{\alpha-1} \cdot \{ t^{1-\alpha} \log(\frac{1}{t}) + t(t+1) \cdot [(t+1)^{1-\alpha} \log(\frac{1}{t+1}) - \\
& - t^{1-\alpha} \log(\frac{1}{t})] \cdot [P(e) - \frac{t-1}{t}] \}. \quad (2.26)
\end{aligned}$$

Para $0 < P(e) \leq \frac{1}{2}$, t atinge o valor 1 e, então,

$$\begin{aligned}
\underline{H}(P(e)) & = - 2^{\alpha-1} \cdot \{ 1^{1-\alpha} \cdot \log(\frac{1}{1}) + 1 \cdot (1+1) \cdot \\
& \cdot [2^{1-\alpha} \log(\frac{1}{2}) - 1^{1-\alpha} \log(\frac{1}{1})] \cdot [P(e) - \frac{1-1}{1}] \}. \\
& = - 2^{\alpha-1} \cdot \{ 2 \cdot [2^{1-\alpha} \cdot (-1)] \cdot [P(e)] \\
& = 2^{\alpha-1} \cdot 2 \cdot 2^{1-\alpha} \cdot P(e) \\
& = 2^{\alpha-1+1+1-\alpha} \cdot P(e) \\
& = 2 \cdot P(e),
\end{aligned}$$

que implica na seguinte cota superior

$$\begin{aligned}
P(e) & \leq \frac{1}{2} H(X), \quad 0 < P(e) \leq \frac{1}{2}, \text{ isto é,} \\
P(e) & \leq \frac{1}{2} H_{\ell}^{\alpha}(C/X), \quad 0 < P(e) \leq \frac{1}{2}, \quad \frac{1}{2} \leq \alpha \leq 1. \quad (2.27)
\end{aligned}$$

2.3.5 - Entropia de grau (α, β) .

Seja $f(p) = (2^{1-\alpha} - 2^{1-\beta})^{-1} \cdot (p^{\alpha} - p^{\beta})$, $\alpha \neq \beta$, $\alpha, \beta > 0$,

então,
$$H(P) = \sum_{i=1}^m f(p_i)$$

$$= (2^{1-\alpha} - 2^{1-\beta})^{-1} \cdot \sum_{i=1}^m (p_i^\alpha - p_i^\beta), \quad (2.28)$$

$$\alpha \neq \beta, \quad \alpha, \beta > 0.$$

Concavidade:

$$\text{Temos, } f(p) = (2^{1-\alpha} - 2^{1-\beta})^{-1} (p^\alpha - p^\beta),$$

$$f'(p) = (2^{1-\alpha} - 2^{1-\beta})^{-1} \cdot (\alpha p^{\alpha-1} - \beta p^{\beta-1}),$$

$$\text{e } f''(p) = (2^{1-\alpha} - 2^{1-\beta}) \cdot (\alpha(\alpha-1)p^{\alpha-2} - \beta(\beta-1)p^{\beta-2})$$

Agora, $f''(p) \leq 0$, se

$$(2^{1-\alpha} - 2^{1-\beta})^{-1} \cdot (\alpha(\alpha-1)p^{\alpha-2} - \beta(\beta-1)p^{\beta-2}) \leq 0$$

1º caso: $2^{1-\alpha} - 2^{1-\beta} > 0$, se $\alpha < \beta$

$$\text{e } \alpha(\alpha-1)p^{\alpha-2} - \beta(\beta-1)p^{\beta-2} \leq 0$$

se $\alpha(\alpha-1) \leq 0$, se $0 < \alpha \leq 1$

e $\beta(\beta-1) \geq 0$, se $\beta \geq 1$ ou $1 \leq \beta$

i.é., $f''(p) \leq 0$, se $0 < \alpha \leq 1 < \beta$.

2º caso: $2^{1-\alpha} - 2^{1-\beta} < 0$, se $\beta < \alpha$

$$\alpha(\alpha-1)p^{\alpha-2} - \beta(\beta-1)p^{\beta-2} \geq 0$$

se $\alpha(\alpha-1) \geq 0 \rightarrow \alpha \geq 1$

e se $\beta(\beta-1) \leq 0 \rightarrow \beta \leq 1$

isto é, $f''(p) \leq 0$, se $0 < \beta \leq 1 < \alpha$

Portanto,

$f'(p)$ é côncava em $[0,1]$ quando $0 < \alpha \leq 1 < \beta$

isto é, $H^{(\alpha, \beta)}(P)$ é côncava em Δ_m quando $0 < \alpha \leq 1 < \beta$ ou

$0 < \beta \leq 1 < \alpha$.

Cotas superior e inferior

Aplicando os teoremas 2.1 e 2.3, respectivamente, temos

$$\begin{aligned} \bar{H}(P(e)) &= (2^{1-\alpha} - 2^{1-\beta})^{-1} \cdot \{[(1 - P(e))^\alpha - (1 - P(e))^\beta] + \\ &+ (m-1) \cdot [(\frac{P(e)}{m-1})^\alpha - (\frac{P(e)}{m-1})^\beta]\}, \end{aligned}$$

$$0 < \alpha \leq 1 < \beta \quad \text{ou} \quad 0 < \beta \leq 1 < \alpha$$

$$\begin{aligned} \text{isto é, } H^{(\alpha, \beta)}(C/X) &\leq (2^{1-\alpha} - 2^{1-\beta})^{-1} \cdot \{[(1 - P(e))^\alpha - (1 - P(e))^\beta] + \\ &+ (m-1) \cdot [(\frac{P(e)}{m-1})^\alpha - (\frac{P(e)}{m-1})^\beta]\}, \end{aligned} \quad (2.29)$$

$$0 < \alpha \leq 1 < \beta \quad \text{ou} \quad 0 < \beta \leq 1 < \alpha.$$

Aplicando o teorema 2.4, temos

$$\begin{aligned} \underline{H}(P(e)) &= (2^{1-\alpha} - 2^{1-\beta})^{-1} \cdot \{t[(\frac{1}{t})^\alpha - (\frac{1}{t})^\beta] + \\ &+ t(t+1) \cdot [(t+1) \cdot ((\frac{1}{t+1})^\alpha - (\frac{1}{t+1})^\beta) - \\ &- t((\frac{1}{t})^\alpha - (\frac{1}{t})^\beta)] \cdot (P(e) - \frac{t-1}{t})\} \\ &= (2^{1-\alpha} - 2^{1-\beta})^{-1} \cdot \{(t^{1-\alpha} - t^{1-\beta}) + \\ &+ t(t+1) \cdot [(t+1)^{1-\alpha} - (t+1)^{1-\beta}] - \end{aligned}$$

$$- (t^{1-\alpha} - t^{1-\beta})] \cdot (P(e) - \frac{t-1}{t}) \}. \quad (2.30)$$

$$0 < \alpha \leq 1 < \beta \text{ ou } 0 < \beta \leq 1 < \alpha.$$

Para $0 < P(e) \leq \frac{1}{2}$, t atinge o valor 1 e, então,

$$\begin{aligned} \underline{H}(P(e)) &= (2^{1-\alpha} - 2^{1-\beta})^{-1} \{ 2 \cdot (2^{1-\alpha} - 2^{1-\beta}) \cdot P(e) \} \\ &= 2P(e), \end{aligned}$$

isto é, $P(e) \leq \frac{1}{2} H(X), \quad 0 < P(e) \leq \frac{1}{2}$

ou $P(e) \leq \frac{1}{2} H^{(\alpha, \beta)}(C/X), \quad 0 < P(e) \leq \frac{1}{2}, \quad (2.31)$

$$0 < \alpha \leq 1 < \beta \text{ ou } 0 < \beta \leq 1 < \alpha.$$

CAPÍTULO III

COTAS SUPERIORES DA PROBABILIDADE DE ERRO

Neste capítulo apresentaremos as cotas superiores em função das entropias de grau β , entropia de grau (α, β) , α -log entropia, dadas respectivamente em (1.4), (1.6), (1.8) e aplicando principalmente a propriedade da recursividade destas medidas. Os casos das medidas (1.1) e (1.3) não serão incluídos neste capítulo, pois (1.1) foi estudada por Hellman e Raviv (1970) e (1.3) é caso particular simples de (1.4).

No presente capítulo demonstraremos que podemos obter os resultados do capítulo II, aplicando diretamente as medidas dadas, para obter cotas superiores para a probabilidade de erro, aplicando a propriedade da recursividade dessas medidas.

3.1 - Aplicações da Recursividade

Sem perda de generalidade, usamos:

$$M = \max \{p_1, p_2, \dots, p_m\} = p_m.$$

Lema 3.1:

Para $\beta \neq 1$, $\beta > 0$, $P \in \Delta_m$, $m \geq 2$, temos

$$H^\beta(1 - M, M) \leq H^\beta(P), \quad (3.1)$$

onde $H^\beta(P)$ é a entropia de grau β dada em (1.4).

Prova:

Podemos escrever:

$$\begin{aligned}
 & H^\beta(p_1, p_2, \dots, p_m) - H^\beta(p_1+p_2, p_3, \dots, p_m) = \\
 & = (2^{1-\beta} - 1)^{-1} \cdot \left\{ \left(\sum_{i=1}^m p_i^\beta - 1 \right) - (p_1+p_2)^\beta - \sum_{i=3}^m p_i^\beta + 1 \right\} \\
 & = (2^{1-\beta} - 1)^{-1} \cdot \{ p_1^\beta + p_2^\beta - (p_1+p_2)^\beta \} \\
 & = (2^{1-\beta} - 1)^{-1} (p_1+p_2)^\beta \cdot \left\{ \left(\frac{p_1}{p_1+p_2} \right)^\beta + \right. \\
 & \left. + \left(\frac{p_2}{p_1+p_2} \right)^\beta - 1 \right\}, \quad p_1 + p_2 > 0 \\
 & = (p_1+p_2)^\beta \cdot H_2^\beta \left(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2} \right), \quad \beta \neq 1, \quad \beta > 0.
 \end{aligned}$$

$$\text{Mas } H_2^\beta \left(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2} \right) \geq 0.$$

Portanto,

$$H^\beta(p_1, p_2, \dots, p_m) \geq H^\beta(p_1+p_2, p_3, \dots, p_m)$$

Por indução podemos concluir que:

$$H^\beta(p_1, p_2, \dots, p_m) \geq H^\beta(p_1+p_2 + \dots + p_{m-1}, p_m), \text{ isto é,}$$

$$H^\beta(P) \geq H^\beta(1-M, M), \quad \beta \neq 1, \quad \beta > 0.$$

Lema 3.2:

Para todo $0 < \alpha \leq 1$, $P \in \Delta_m$, $m \geq 2$, temos,

$$H_{\ell}^{\alpha}(1 - M, M) \leq H_{\ell}^{\alpha}(P), \quad (3.2)$$

onde $H_{\ell}^{\alpha}(P)$ é a α -log entropia dada em (1.8).

Prova:

$$\begin{aligned} & H_{\ell}^{\alpha}(p_1, p_2, \dots, p_m) - H_{\ell}^{\alpha}(p_1 + p_2, p_3, \dots, p_m) \\ &= - 2^{\alpha-1} \sum_{i=1}^m p_i^{\alpha} \log p_i + 2^{\alpha-1} (p_1 + p_2)^{\alpha} \log (p_1 + p_2) + \\ &+ 2^{\alpha-1} \cdot \sum_{i=3}^m p_i^{\alpha} \log p_i. \\ &= - 2^{\alpha-1} \{ p_1^{\alpha} \log p_1 + p_2^{\alpha} \log p_2 - (p_1 + p_2)^{\alpha} \log (p_1 + p_2) \} \\ &= - 2^{\alpha-1} \cdot (p_1 + p_2)^{\alpha} \left\{ \left(\frac{p_1}{p_1 + p_2} \right)^{\alpha} \log p_1 + \left(\frac{p_2}{p_1 + p_2} \right)^{\alpha} \log p_2 - \right. \\ &\quad \left. - \log (p_1 + p_2) \right\}. \\ &= (p_1 + p_2)^{\alpha} \cdot H_{\ell}^{\alpha} \left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2} \right) - 2^{\alpha-1} \log (p_1 + p_2) \cdot \\ &\quad \cdot \left\{ \left(\frac{p_1}{p_1 + p_2} \right)^{\alpha} + \left(\frac{p_2}{p_1 + p_2} \right)^{\alpha} - 1 \right\} \\ &= (p_1 + p_2)^{\alpha} \cdot H_{\ell}^{\alpha} \left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2} \right) + 2^{\alpha-1} \log \frac{1}{p_1 + p_2} \cdot \\ &\quad \cdot \left\{ \left(\frac{p_1}{p_1 + p_2} \right)^{\alpha} + \left(\frac{p_2}{p_1 + p_2} \right)^{\alpha} - 1 \right\}. \end{aligned}$$

≥ 0 , se

$$\left(\frac{p_1}{p_1 + p_2} \right)^{\alpha} + \left(\frac{p_2}{p_1 + p_2} \right)^{\alpha} \geq 1 \quad (\text{pois todos os fatores são})$$

positivos), o que é válido para $0 < \alpha \leq 1$ pela conhecida desigualdade:

$$\sum_{i=1}^m p_i^\alpha \geq \left(\sum_{i=1}^m p_i \right)^\alpha, \quad 0 < \alpha \leq 1.$$

Portanto,

$$H_\lambda^\alpha(p_1, p_2, \dots, p_m) \geq H_\lambda^\alpha(p_1 + p_2, p_3, \dots, p_m).$$

Por indução temos o resultado desejado.

Lema 3.3:

Para todo $P \in \Delta_m$, $m \geq 2$, $0 < \alpha \leq 1 < \beta$ ou $0 < \beta \leq 1 < \alpha$, temos

$$H^{(\alpha, \beta)}(1 - M, M) \leq H^{\alpha, \beta}(P), \quad (3.3)$$

onde $H^{(\alpha, \beta)}(P)$ é a entropia de grau (α, β) dada em (1.6).

Prova:

Podemos escrever

$$H^{(\alpha, \beta)}(P) = \frac{A_\alpha}{A_\alpha - A_\beta} H^\alpha(P) + \frac{A_\beta}{A_\beta - A_\alpha} H^\beta(P), \quad (3.4)$$

onde $H^\eta(P)$ é a entropia de grau η e $A_\eta = 2^{1-\eta} - 1$, $\eta = \alpha$ ou β

De (3.1) e (3.4), obtemos:

$$H^{(\alpha, \beta)}(P) \geq \frac{A_\alpha}{A_\alpha - A_\beta} H^\alpha(1 - M, M) + \frac{A_\beta}{A_\beta - A_\alpha} H^\beta(1 - M, M), \quad (3.5)$$

desde que $\frac{A_\alpha}{A_\alpha - A_\beta} \geq 0$ e $\frac{A_\beta}{A_\beta - A_\alpha} \geq 0$, isto é,

$$A_\alpha \geq 0, A_\alpha > A_\beta \quad e \quad A_\beta \leq 0$$

$$A_\alpha = 2^{\alpha-1} - 1 \geq 0 \quad \rightarrow \quad \alpha \leq 1$$

$$e \quad A_\alpha > A_\beta \rightarrow 2^{1-\alpha} - 1 > 2^{1-\beta} - 1 \rightarrow \\ \rightarrow 1 - \alpha > 1 - \beta \rightarrow \alpha < \beta$$

$$e \quad A_\beta \leq 0 \rightarrow \beta > 1$$

Portanto, $0 < \alpha \leq 1 < \beta$

Similarmente, (3.5) é válida para $0 < \beta \leq 1 \leq \alpha$.

Mas,

$$\frac{A_\alpha}{A_\alpha - A_\beta} \cdot H^\alpha(1-M, M) + \frac{A_\beta}{A_\beta - A_\alpha} H^\beta(1-M, M) = H^{(\alpha, \beta)}(1-M, M). \quad (3.6)$$

De (3.5) e (3.6) temos o resultado (3.3).

3.2 - Cotas Superiores

Nas três entropias citadas em (1.4), (1.8) e (1.6) temos as seguintes propriedades em comum:

$$\frac{1}{2} H(0, 1) = \frac{1}{2} H(1, 0) = 0$$

$$e \quad \frac{1}{2} H\left(\frac{1}{2}, \frac{1}{2}\right) = \frac{1}{2}. \quad (3.7)$$

Por outro lado como o gráfico de $1 - \max\{p, 1 - p\}$, $0 \leq p \leq 1$ consiste de duas retas entre $(0, 0)$ e $(\frac{1}{2}, \frac{1}{2})$ e entre $(\frac{1}{2}, \frac{1}{2})$ e $(1, 0)$, obtemos o seguinte resultado usando a concavidade de $H(1-p, p)$ (com suas respectivas condições de parâme-

tros), isto é,

$$1 - \max \{p, 1 - p\} \leq \frac{1}{2} H(1 - p, p) \quad (3.8)$$

ou

$$1 - M \leq \frac{1}{2} H(1 - M, M) \quad (3.9)$$

Agora, usando (3.1), (3.2) e (3.3), temos

$$p_e \leq \frac{1}{2} H^\beta(P), \quad (\beta > 0) \quad (3.10)$$

$$p_e \leq \frac{1}{2} H_\ell^\alpha(P), \quad \left(\frac{1}{2} \leq \alpha \leq 1\right) \quad (3.11)$$

$$p_e \leq \frac{1}{2} H^{(\alpha, \beta)}(P), \quad (0 < \alpha \leq 1 < \beta) \quad (3.12)$$

$$\text{ou } 0 < \beta \leq 1 < \alpha$$

respectivamente.

Substituindo p_i por $P(C/x)$, para cada $X = x$ em (3.10), (3.11), (3.12) e tomando a expectativa com respeito a X , temos

$$P(e) \leq \frac{1}{2} H^\beta(C/X), \quad \beta > 0 \quad (3.13)$$

$$P(e) \leq \frac{1}{2} H_\ell^\alpha(C/X), \quad \frac{1}{2} \leq \alpha \leq 1 \quad (3.14)$$

$$P(e) \leq \frac{1}{2} H^{(\alpha, \beta)}(C/X), \quad (3.15)$$

$$0 < \alpha \leq 1 < \beta \quad \text{ou} \quad 0 < \beta \leq 1 < \alpha$$

respectivamente.

Observação: De (3.8) temos $\max \{p, 1 - p\} \geq \frac{1}{2}$, isto é, de (3.9) temos $M \geq \frac{1}{2}$, o que implica que $0 < p(e/x) \leq \frac{1}{2}$ em (3.13), (3.14) e (3.15).

APÊNDICE

DEMONSTRAÇÃO DO TEOREMA 2.1

Como $p_e = 1 - \max \{p_1, p_2, \dots, p_m\}$ e a soma dos p_k é igual a 1 (um), o valor $\max \{p_k\}$ não pode ser menor que $\frac{1}{m}$ e m é o número de classes. Assim, segue-se que $0 \leq p_e \leq 1 - \frac{1}{m}$.

Sem perda de generalidade, vamos supor que $\max \{p_1, p_2, \dots, p_m\} = p_m$. Então,

$$\begin{aligned} H(P) &= \sum_{i=1}^m f(p_i) \\ &= \sum_{i=1}^m f(p_i) + f(p_m) \\ &= (m-1) \cdot \sum_{i=1}^m \frac{1}{(m-1)} \cdot f(p_i) + f(1 - p_e) \end{aligned}$$

Agora, como f é uma função côncava e

$$\sum_{i=1}^{m-1} \frac{1}{(m-1)} \cdot f(p_i) \leq f\left(\frac{1}{m-1} \cdot \sum_{i=1}^{m-1} p_i\right), \text{ temos}$$

$$\begin{aligned} H(P) &\leq (m-1) \cdot f\left(\frac{1}{m-1} \cdot \sum_{i=1}^{m-1} p_i\right) + f(1 - p_e) \\ &= (m-1) \cdot f\left(\frac{1}{m-1} \cdot (1 - p_m)\right) + f(1 - p_e) \end{aligned}$$

$$H(P) = (m-1) \cdot f\left(\frac{p_e}{m-1}\right) + f(1 - p_e)$$

Portanto,

$$\text{Sup } H(P) = \bar{h}(P(e)) = f(1 - p_e) + (m-1) \cdot f\left(\frac{p_e}{m-1}\right),$$

$p \in \Delta_m(p_e)$, o que completa a demonstração do teorema.

Demonstração do teorema 2.2

Como no caso anterior, logo que a probabilidade de erro é dada, $\max_k p_k = p_1 = 1 - p_e$, é também dado. Consequentemente, para todos os valores de p_k , a inequação

$$0 \leq p_k \leq 1 - p_e \text{ é válida.}$$

A idéia da demonstração é que separando os valores do par de argumentos p_k e p_l , com sua soma permanecendo constante e, conseqüentemente, com a probabilidade de erro permanecendo constante, diminuimos o valor médio da função convexa destes argumentos, isto é, diminuimos a entropia. Alguma separação a mais, torna-se impossível quando todos os valores de p_k exceto, talvez um, estiverem no limite do intervalo dos valores possíveis.

De fato, suponhamos que alguns pares quaisquer de valores de p_k sejam diferentes de 0 (zero) e de $1 - p_e$. Sejam p_2 e p_3 esses valores. Então, separando-os, isto é, substituindo seus valores por p_2^* e p_3^* que estão mais próximos de 0 (zero) e de $1 - p_e$, de modo que as inequações

$$p_2^* < p_2 < p_3^*$$

$$p_2^* < p_3 < p_3^* ,$$

estejam satisfeitas.

Esta mudança sempre pode ser feita, conservando a soma:

$$p_2^* + p_3^* = p_2 + p_3.$$

A₁

e conseqüentemente, não transgredindo a condição $\sum_{i=1}^m p_k = p$.

Os valores p_2 e p_3 que se encontram entre p_2^* e p_3^* , podem ser representados por médias ponderadas dos últimos valores:

$$p_2 = \alpha p_2^* + (1 - \alpha) p_3^*$$

$$p_3 = \beta p_2^* + (1 - \beta) p_3^*$$

Substituindo estas relações na fórmula A_1 e na base da relação $p_3^* - p_2^* \neq 0$, encontramos $\alpha + \beta = 1$.

Conforme a condição:

$$f\left(\sum_k \alpha_k z_k\right) \geq \sum_k \alpha_k f(z_k),$$

levando em conta o fato de que p_2^* e p_3^* não são iguais às suas médias, obtemos

$$f(p_2) > \alpha f(p_2^*) + (1 - \alpha) f(p_3^*),$$

$$f(p_3) > \beta f(p_2^*) + (1 - \beta) f(p_3^*).$$

Levando-se em conta a inequação A_2 , encontramos:

$$f(p_2) + f(p_3) > f(p_2^*) + f(p_3^*).$$

Chegamos à conclusão que, quando um par de valores de p_k é "separado" o par correspondente de termos na expressão:

$$H(P) = f(1 - p_e) + \sum_{k=2}^m f(p_k),$$

diminui, e ao mesmo tempo a entropia $H(P)$ também diminui.

Uma diminuição adicional é impossível se todos os valores de p_k são iguais a 0 (zero) ou a $1 - p_e$, porque com uma

probabilidade de erro dada p_e , os valores p_k não podem se estender para além dos limites do segmento $[0, 1 - p_e]$.

O decréscimo é também impossível se somente um dos valores p_k não ficar sobre o limite do intervalo, porque este valor isolado não pode ser mudado sem se mudar o valor dado da probabilidade de erro.

A fim de encontrar o valor de $\underline{h}(p_e)$, que é o limite inferior exato da entropia, dividimos 1 (um) pelo máximo valor de p_k igual $1 - p_e$.

A parte inteira m do quociente obtido mostrará quantos valores de p_k deve ser igual a $1 - m \cdot (1 - p_e)$ e o resto deve ser igual a 0 (zero). Conseqüentemente,

$$\inf H(P) = \underline{h}(p_e) = t \cdot f(1 - p_e) + f[1 - t(1 - p_e)].$$

Demonstração do lema 2.1

Sejam p_e e p'_e pontos em $[0, 1 - \frac{1}{m}]$ e seja λ um número real tal que $0 \leq \lambda \leq 1$. Pelo teorema 2.1 e a concavidade de f , segue-se para um dado $H(P) = T(f)$,

$$\begin{aligned} \bar{h}(\lambda p_e + (1 - \lambda)p'_e) &= f(1 - \lambda p_e - (1 - \lambda)p'_e) + (m-1) f\left(\frac{\lambda p_e + (1 - \lambda)p'_e}{m-1}\right) \\ &= f(\lambda(1 - p_e) + (1 - \lambda)(1 - p'_e)) + \\ &+ (m-1) \cdot f\left(\frac{\lambda p_e}{m-1} + (1 - \lambda) \cdot \frac{p'_e}{m-1}\right) \\ &> \lambda f(1 - p_e) + (1 - \lambda) \cdot f(1 - p'_e) + \\ &+ (m-1) \cdot \lambda f\left(\frac{p_e}{m-1}\right) + (m-1)(1 - \lambda) \cdot f\left(\frac{p'_e}{m-1}\right) \\ &= \lambda \bar{h}(p_e) + (1 - \lambda) \bar{h}(p'_e). \end{aligned}$$

Demonstração do lema 2.2

Sejam p_e e p'_e pertencentes ao intervalo $[\frac{t-1}{t}, \frac{t}{t+1}]$, com t inteiro e seja λ um número real tal que $0 \leq \lambda \leq 1$. Pelo teorema 2.2 e pela convexidade de f , segue-se que, para um dado $H(P) \in T(f)$, temos:

$$\begin{aligned} \underline{h}(\lambda p_e + (1 - \lambda)p'_e) &= t \cdot f(\lambda(1 - p_e) + (1 - \lambda) \cdot (1 - p'_e)) + \\ &+ f[\lambda(1 - t(1 - p_e)) + (1 - \lambda) \cdot (1 - t(1 - p'_e))] \\ &+ \lambda \cdot t f(1 - p_e) + (1 - \lambda)t \cdot f(1 - p'_e) + \\ &+ \lambda f[1 - t(1 - p_e)] + (1 - \lambda) \cdot f[1 - t(1 - p'_e)] \\ &= \lambda \underline{h}(p_e) + (1 - \lambda) \underline{h}(p'_e). \end{aligned}$$

Demonstração do lema 2.3

Uma função linear por partes é convexa se sua inclinação nunca decrescer. Considere a inclinação de $L(p_e)$ sem o fator positivo $t \cdot (1+t)$,

$$a(t) = (t+1) \cdot f\left(\frac{1}{t+1}\right) - t \cdot f\left(\frac{1}{t}\right),$$

$$a'(t) = \left[f\left(\frac{1}{t+1}\right) - \frac{1}{t+1} \cdot f'\left(\frac{1}{t+1}\right) \right] - \left[f\left(\frac{1}{t}\right) - \frac{1}{t} \cdot f'\left(\frac{1}{t}\right) \right]$$

Tomando a derivada de:

$$b(t) = f\left(\frac{1}{t}\right) - \left(\frac{1}{t}\right) \cdot f'\left(\frac{1}{t}\right), \text{ obtemos}$$

$$b'(t) = \left(\frac{1}{t^3}\right) \cdot f''\left(\frac{1}{t}\right), \text{ que é não positivo, para } t \text{ posi-}$$

tivo e f côncava, portanto, $b(t)$ é uma função decrescente de

t e, portanto, $a'(t)$ é não-negativa, para t positivo, o que implica que essa inclinação $a(t)$ nunca decresce.

Demonstração do teorema 2.3

Trocando p por $P(e/x)$ em (2.7) e tomando o valor esperado de ambos os lados, obtemos o resultado desejado, pois

$$\begin{aligned} H(X) &= E_X \bar{h}(P(e/x)) = E_X \left\{ f(1 - P(e/x)) + \right. \\ &\quad \left. + (m-1) \cdot f\left(\frac{P(e/x)}{m-1}\right) \right\} \\ &\leq f(1 - E_X(P(e/x))) + (m-1) \cdot f\left(\frac{E_X(P(e/x))}{m-1}\right), \text{ pois } f \text{ é c\u00f4ncava.} \end{aligned}$$

$$= f(1 - P(e)) + (m-1) f\left(\frac{P(e)}{m-1}\right)$$

$$= \text{Sup } H(X) = \bar{h}(P(e))$$

$$= \bar{H}(P(e)) = \bar{h}(P(e)).$$

Demonstração do teorema 2.4

Sua demonstração segue-se imediatamente usando os lemas 2.2 e 2.3 e tomando o valor esperado da informação e da probabilidade de erro.

BIBLIOGRAFIA

- ACZÉL, J. & Z. Daróczy (1963). Über verallgemeinerte quasi-lineare Mittelwerte, die mit Gewichts-funktionen gebildet sind. Publicationes Mathematicae Debrecen, vol. 10, pág. 171-190.
- ARIMOTO, S. (1971). Information theoretical considerations on estimation problems. Information and Control, vol. 19, pág. 181-194.
- BEN-BASSAT, M. (1978). f-entropies, probability of error and feature selection. Information and Control, vol. 39, pág. 227-242.
- BEN-BASSAT, M. and J. RAVIV, (1978). Rényi's entropy and the probability of error. IEEE Trans. Inform. Theory, vol. IT-24, pág. 324-331.
- BOEKEE, D.E. & J.C.A. Van der Lubbe (1980). The r-norm information measure. Information and Control, vol. 45, pág. 136-155.
- CHU, T.J. & J.C. CHUEH (1966). Inequalities Between Information Measures and Error Probability. J. Franklin Inst., vol. 282, pág. 121-125.
- DARÓCZY, Z. (1970). Generalized Information Functions. Information and Control, vol. 16, pág. 36-51.
- DEVIJVER, P.A. (1974). "On a new class of bounds on Bayes riski in multihypotesis pattern recognition". IEEE Trans. Computgrs, vol. C-23, pág. 70-80.
- DEVIJVER, P.A. (1977). Entropies of degree β and lower bounds for the avarage error rate. Information and Control, vol. 34, n^o 3, pág. 222-226.
- HAVRDA, J. & F. Charvát (1967). Quantification Method of Classification Process, the concept of Structural α -Entropy. Kibernetika, vol. 3, pág. 30-35.
- HELLMAN, M.E. & J. Raviv (1970). Probability of error, equivocation and chernoff - bound. IEEE Trans. on Information Theory, vol. II-16, pág. 368-372.

- KOVALEVSKI, V.A. (1968). The problem of character recognition from the point of view of mathematical statistics, in: "Character Readers and Pattern Recognition", pág. 3-30.
- RÉNYI, A. (1961). On measures of entropy and information, proc. 4th Berkeley Symp. Math. Statistics and Probability, vol. 1, pág. 547-561.
- SHANNON, C.E. (1948). A mathematical theory of communication. Bell System Tech. J., vol. 27, pág. 623-658.
- SHARMA, B.D. & I.J. Taneja (1977). Three generalized-additive measures of entropy. Elektronische Informationsverarbeitung und Kybernetik, vol. 13, pág. 419-433.
- TANEJA, I.J. (1982). Generalized γ -entropy and probability of erro, proc. International Conference on Cybernetics and Society Scattle, Washington, pág. 463-466.
- TANEJA, I.J. (1983a). Comentário sobre o artigo "Entropies of degree β and lower bounds for the average error rate". IEEE Trans. Syst. Man and Cybernetics, SMC 13, nº 2, pág. 241-242.
- TANEJA, I.J. (1983b). Comentário sobre o paper "A generalization of Shannon's equivocation and the Fano bound". IEEE Trans. Syst., Man and Cybernetics, SMC 13, nº 3.
- TOUSSAINT, G.T. (1977). A generalization of Shannon's Equivocation and the Fano bound. IEEE Trans. Syst., Man Cybernetics, SMC 7, pág. 300-302.
- VAJDA, I. (1968). Bounds on the minimal error probability and checking a finite or countable number of hypotheses. Information Transm. Problems, vol. 4, pág. 9-17.