EQUAÇÕES FUNCIONAIS E FUNDAMENTOS AXIO-
MÁTICOS DAS MEDIDAS DE INFORMAÇÃO

LILIAN TORNG SHENG

ORIENTADOR

PUSHPA NARAYAN RATHIE

Tese apresentado no Instituto de Mat-
emática, Estatística e Ciência da Co-
mputação como requisito parcial para
a obtenção do título de Doutor em Ci-
ências

SUMÁRIO

Neste trabalho iremos discutir sobre várias medidas de informação generalizada, (tais como a entropia não-adtiva de ordem α, a entropia de ordem (α,β), a divergência dirigida não-aditiva de ordem α, a imprecisão de ordem α, a divergência dirigida não-aditiva de ordem (α,β), a divergência-J de ordem α, etc.), suas propriedades e equações funcionais as quais tem aplicações na caracterização de algumas das medidas de informação.

Capítulo 1 dá uma idea de medidas de informação e suas definições. Algumas equações funcionais que serão usadas nos capítulos posteriores e alguns teoremas relevantes que caracterizam axiomaticamente medidas de informação são também introduzidos. As aplicações das várias medidas de informação são também mencionadas na ultimo secção.

No Capítulo 2 discute-se duas equações funcionais as quais são versões generalizadas da equaão funcional de Caundy e McLeod (1960). A equação funcional em uma variável é usada na caracterização axiomática da entropia não-aditiva de ordem β enquando que a equação funcional em duas variáveis é usada para caracterizar a medida de informação teórica $I_{\gamma,\delta}^{\alpha,\beta}$ de um evento . As propriedades destas medidas de informação são listadas.

No Capítulo 3 discute-se duas equações funcionais as quais surgem naturalmente da não-adtividade das medidas de informação envolvendo uma e duas distribuições de probabilidade respectivamente. A equação funcional em uma variável é aplicada na caracterização da entropia não-adtiva de ordem β enquando a equação funcional em duas variáveis é aplicada na caracterização axiomática de uma medida de informação generalizada de ordem (α,β), a divergência dirigida de ordem α, e a imprecisão de ordem $1+\beta$ respectivamente. Propriedades das medidas de informação são também listadas.

Capítulo 4 é dedicado à divergência-J de ordem α e suas propriedades. Um teorema caracterizando a divergência-J de ordem α é provado assumindo-se um conjunto de cinco postulados.

Parte desta tese foi submetida para publicação nos três artigos seguintes:

(a) On The Measurable Solutions of Certain Functional Equations.

(b) On The Measurable Solutions of Non-Additive Functional Equations.

(c) The J-Divergence of Order α.

O artigo listado em (b) esta em fase de revisão para sua publicação em Annales Polonici Mathematici.

SUMMARY

In this work we shall deal with various generalized information measures, (such as non-additive entropy of order α, the entropy of order (α,β), the non-additive directed divergence of order α, the inaccuracy of order α, the non-additive directed divergence of order (α,β), the J-divergence of order α, etc.) their properties and functional equations which have applications in characterizing some of the information measures.

Chapter 1 gives an idea of information measures and their definitions. Some functional equations which will be used in the later chapters and some relevant axiomatic characterization theorems of information measures are also introduced. The applications of various information measures are also mentioned in the last section.

Chapter 2 deals with two functional equations which are generalized versions of the functional equation of Chaundy and McLeod (1960). The functional equation in one variable is used in an axiomatic characterization of the non-additive entropy of order β of an event while th functional equation in two variables is used to characterize the information theoretic measure $I^{\alpha,\beta}_{\gamma,\delta}$ of an event. The properties of both of these information measures are listed.

Chapter 3 deals with two generalized functional equations which arise naturally from the non-additivity of the information measures involving one and two probability distributions respectively. The functional equation in one variable is applied in axiomatic characterization of the

non-additive entropy of order β while the functional equation in two variables is applied in axiomatic characterization of a generalized information measure of order (α,β), The directed divergence of order α and the inaccuracy of order $1+\beta$ respectively. Properties of the information measures are also listed.

Chapter 4 is devoted to the J-divergence of order α and its properties. A theorem characterizing the J-divergence of order α is proved by assuming a set of five postulates.

Part of this thesis has been submitted for publication in the form of following three papers:

(a) On The Measurable Solutions of Certain Functional Equations.

(b) On The Measurable Solutions of Non-Additive Functional Equations.

(c) The J-Divergence of Order α.

The paper listed in (b) is under revision for its publication in Annales Polonici Mathematici.

Agradeço ao professor Pushpa N. Rathie, orientador desse trabalho, pela sua valiosa orientação, inspiração e encorajamento na preparação desta tese.

A autora é imensamente grata a muitos colegas Brasileiros, em particular ao professor Ademir J. Petenate, pelo auxílio prestado na tradução do manuscrito para o português.

ÍNDICE

1. MEDIDAS DE INFORMAÇÃO E EQUAÇÕES FUNCIONAIS: UMA INTRODUÇÃO

1.1 Introdução 1

1.2 Definições 2

1.3 Equações Funcionais 7

1.4 Teoremas de Caracterização 11

1.5 Aplicações 15

2. EQUAÇÕES FUNCIONAIS E CARACTERIZAÇÕES DE MEDIDAS DE INFORMAÇÃO

2.1 Introdução 16

2.2 Propriedades 22

2.3 A Solução de (2.1.1) 27

2.4 A Solução de (2.1.8) 35

3. EQUAÇÕES FUNCIONAIS E MEDIDAS DE INFORMAÇÃO NÃO-ADITIVAS

3.1 Introdução 44

3.2 Propriedades 49

3.3 Equação Funcional (3.1.1) Em Uma Variável 53

3.4 Equação Funcional (3.1.7) Em Duas Variáveis 65

4. A DIVERGÊNCIA-J DE ORDEM α

4.1 Introdução 90

4.2 Propriedades 93

4.3 O Teorema de Caracterização 95

REFERENCIA BIBLIOGRAFICA 112

CAPÍTULO 1

MEDIDAS DE INFORMAÇÃO E EQUAÇÕES FUNCIONAIS : UMA INTRODUÇÃO

1.1 INTRODUÇÃO

Shannon (1948) introduziu o trabalho pioneiro de uma medida de informação ou entropia de distribuição geral de probabilidade finita e completa. Ele deu, originalmente, um teorema de caracterização da entropia introduzida por ele, desde então, vários artigos de pesquisa tem sido publicados os quais simplificam e estendem o trabalho original de Shannon.

Entropia pode ser interpretada como uma medida da incerteza ou uma medida de informação. A entropia de Shannon é generalizada para a entropia de ordem α, e isto será discutido em detalhes nos Capítulos 2 e 3. Generalizações adicionais de entropia (chamadas imprecisão, divergência dirigida, divergência-J, e suas generalizações de ordem α, etc.) são também discutidas nos próximos capítulos. Estas generalizações envolvem duas distribuições de probabilidade, falando vagamente, a maioria delas são para medir a discriminação de uma distribuição de probabilidade contra a outra, ou a discriminação entre duas distribuições de probabilidade.

1.2 DEFINIÇÕES

A fim de permitir uma introdução global sobre este trabalho, começamos por definir as várias medidas de informação e estatísticas, embora elas serão definidas novamente em capítulos respectivos a fim de que os capítulos sejam auto-suficientes.

Seja

$$\Delta_n = \{(p_1,\ldots,p_n): p_i \geqq 0, \sum_{i=1}^{n} p_i = 1\} \ , \quad e$$

$$\Delta'_n = \{(p_1,\ldots,p_n): p_i \geqq 0, \sum_{i=1}^{n} p_i \leqq 1\}$$

para n = 1, 2, ...

Usaremos a convenção $0 \cdot \log 0 = 0$ e $0^\alpha = 0$ $(\alpha \neq 0)$ nas expressões ao longo do texto e os logarítmos serão tomados na base 2.

Definição 1. A entropia de Shannon de uma distribuição $P = (p_1,\ldots,p_n)$ em Δ_n é definida por

$$(1.2.1) \qquad H_n(P) = -\sum_{i=1}^{n} p_i \log p_i$$

Definição 2. A entropia aditiva de ordem α [Rényi (1960)] de um distribuição $P = (p_1,\ldots,p_n) \in \Delta_n$ é definida por

$$(1.2.2) \qquad \mathcal{H}_{n,\alpha}(P) = (1-\alpha)^{-1} \log \left(\sum_{i=1}^{n} p_i^\alpha \right), \qquad \alpha \neq 1$$

Definição 3. A entropia não-aditiva de ordem α [Havrda e Charvát (1967)] de uma distribuição $P = (p_1,\ldots,p_n) \in \Delta_n$ é definida por

$$(1.2.3) \qquad H_{n,\alpha}(P) = \frac{\sum_{i=1}^{n} p_i^\alpha - 1}{2^{1-\alpha} - 1} \ , \qquad \alpha \neq 1$$

Definição 4. A entropia de ordem (α,β) [Sharma e Taneja (1975)] de uma distribuição $P = (p_1,\ldots,p_n) \in \Delta_n$ é definida por

$$(1.2.4) \qquad H_{n,\alpha,\beta}(P) = \frac{\sum_{i=1}^{n} (p_i^{\alpha} - p_i^{\beta})}{2^{1-\alpha} - 2^{1-\beta}}, \qquad \alpha \neq \beta$$

Definição 5. A divergência dirigida [Kullback (1959)] ou ganho de informação [Rényi (1960)] entre duas distribuições $P = (p_1,\ldots,p_n)$ e $Q = (q_1,\ldots,q_n)$ em Δ_n é definida por

$$(1.2.5) \qquad I_n(P:Q) = \sum_{i=1}^{n} p_i \log(p_i/q_i)$$

onde sempre que um $q_i = 0$, o correspondente $p_i = 0$.

Definição 6. A divergência dirigida aditiva de ordem α [Rényi (1960)] entre duas distribuições $P = (p_1,\ldots,p_n)$ e $Q = (q_1,\ldots,q_n)$ em Δ_n é definida por

$$(1.2.6) \qquad \hat{I}_{n,\alpha}(P:Q) = \frac{1}{\alpha - 1} \log\left(\sum_{i=1}^{n} p_i^{\alpha} q_i^{1-\alpha}\right), \qquad \alpha \neq 1$$

Definição 7. A divergência dirigida não-aditiva de ordem α [Rathie e Kannappan (1972)] entre duas distribuições $P = (p_1,\ldots,p_n)$ e $Q = (q_1,\ldots,q_n)$ em Δ_n é definida por

$$(1.2.7) \qquad I_{n,\alpha}(P:Q) = \frac{\sum_{i=1}^{n} p_i^{\alpha} q_i^{1-\alpha} - 1}{2^{\alpha-1} - 1}, \qquad \alpha \neq 1$$

Definição 8. A divergência dirigida não-aditiva de ordem (α,β) [Kannappan e Tathie (1975 a)] [Sharma e Taneja (1975)] entre duas distribuições

$P = (p_1,\dots,p_n) \in \Delta_n$ e $Q = (q_1,\dots,q_n) \in \Delta'_n$ é definida por

(1.2.8) $$I_{n,\alpha,\beta}(P:Q) = \frac{\sum_{i=1}^{n} p_i^{\alpha} q_i^{\beta} - 1}{2^{-\beta} - 1}, \qquad \beta \neq 0$$

Definição 9. A medida de informação teórica de ordem $\binom{\alpha,\beta}{\gamma,\delta}$ [Sharma e Taneja (1975)] entre duas distribuições $P = (p_1,\dots,p_n)$ e $Q = (q_1,\dots,q_n)$ em Δ_n é definida por

(1.2.9) $$I^{\alpha,\beta}_{\gamma,\delta}(P:Q) = \frac{\sum_{i=1}^{n} (p_i^{\alpha} q_i^{\beta} - p_i^{\gamma} q_i^{\delta})}{2^{-\beta} - 2^{-\delta}}, \qquad (\alpha-\gamma)(\delta-\beta) < 0$$

Definição 10. A divergência-J ou divergência simétrica [Kullback (1959)] entre duas distribuições $P = (p_1,\dots,p_n)$ e $Q = (q_1,\dots,q_n)$ é definida por

(1.2.10) $$J_n(P:Q) = I_n(P:Q) + I_n(Q:P)$$

$$= \sum_{i=1}^{n} (p_i - q_i) \log (p_i/q_i)$$

onde $q_i = 0$ se e sòmente se o correspondente $p_i = 0$.

Definição 11. A divergência-J ou divergênci simetrica de ordem α [neste trabalho, Capítulo 4] entre duas distribuições $P = (p_1,\dots,p_n)$ e $Q = (q_1,\dots,q_n)$ em Δ_n é definida por

(1.2.11) $$J_{n,\alpha}(P:Q) = I_{n,\alpha}(P:Q) + I_{n,\alpha}(Q:P)$$

$$= (2^{\alpha-1} - 1)^{-1}\{\sum_{i=1}^{n} p_i^{\alpha} q_i^{1-\alpha} + \sum_{i=1}^{n} q_i^{\alpha} p_i^{1-\alpha} - 2\}, \qquad \alpha \neq 1$$

Definição 12. A imprecisão [Kerridge (1961)] entre duas distribuições $P = (p_1,\dots,p_n)$ e $Q = (q_1,\dots,q_n)$ em Δ_n é definida por

$$(1.2.12)\quad K(P:Q) = -\sum_{i=1}^{n} p_i \log q_i$$

onde sempre que um $q_i = 0$, o correspondente $p_i = 0$.

Definição 13. A imprecisão de ordem β [Rathie e Kannappan (1973)] entre duas distribuições é definida por

$$(1.2.13)\quad K^{\beta}(P:Q) = \frac{\sum_{i=1}^{n} p_i q_i^{\beta} - 1}{2^{-\beta} - 1}, \qquad \beta \neq 0$$

Definição 14. A afinidade [Matusita (1961)] entre duas distribuições $P = (p_1,\ldots,p_n)$ e $Q = (q_1,\ldots,q_n)$ em Δ_n é definida por

$$(1.2.14)\quad A_n(P:Q) = \sum_{i=1}^{n} (p_i q_i)^{1/2}$$

Definição 15. A medida de distância quadratica [Matusita (1966)] entre duas distribuições $P = (p_1,\ldots,p_n)$ e $Q = (q_1,\ldots,q_n) \in \Delta_n$ é definida por

$$(1.2.15)\quad D_n(P:Q) = 2\sum_{i=1}^{n} (p_i^{1/2} - q_i^{1/2})^2$$

$$= 2\{1 - \sum_{i=1}^{n} (p_i q_i)^{1/2}\}$$

Definição 16. A estatistica χ^2 de Pearson é definida por

$$(1.2.16)\quad \chi^2 = \sum_{i=1}^{k} \{(n_i - np_i)^2/(np_i)\} = n\{\sum_{i=1}^{k} q_i^2 p_i^{-1} - 1\}$$

onde $n = \sum_{i=1}^{k} n_i$, n_i e np_i são as frequências observadas e esperadas correspondentes ao i-ésimo grupo respectivamente e p_i é a probabilidade de obtermos uma observação no i-ésimo grupo, e $q_i = n_i/n$.

É claro que, quando $\alpha \to 1$, os dois tipos de entropia de ordem α em (1.2.2) e (1.2.3) reduzem-se à entropia de Shannon em (1.2.1), portanto são entropias generalizadas. $\hat{H}_{n,\alpha}$ e $H_{n,\alpha}$ são ligadas pela relação

$$(1.2.17) \quad H_{n,\alpha} = \frac{1}{2^{1-\alpha}-1}\{2^{(1-\alpha)\hat{H}_{n,\alpha}}-1\}, \qquad \alpha \neq 1$$

A entropia de ordem (α,β) em (1.2.4) envolve dois parametros , quando $\beta = 1$, é óbvio que (1.2.3) e (1.2.4) coincidem.

Temos relaçoes similares entre divergência dirigida em (1.2.5) e dois tipos de divergêcia dirigida de ordem α em (1.2.6) e (1.2.7).

Quando $\alpha \to 1$, ambos os tipos de divergência dirigida de ordem α reduzem-se à divergência dirigida em (1.2.5), é claramente, de (1.2.6) e (1.2.7), $\hat{I}_{n,\alpha}$ e $I_{n,\alpha}$ tem a seguinte relação

$$(1.2.18) \quad I_{n,\alpha} = \frac{1}{2^{\alpha-1}-1}\{2^{(\alpha-1)\hat{I}_{n,\alpha}}-1\} \qquad \alpha \neq 1$$

$I_{n,\alpha,\beta}$ definida em (1.2.8) envolve dois parametros, é óbvio que quando $\beta = 1 - \alpha$, (1.2.8) coincide com (1.2.7).

Também é evidente que quando $\alpha \to 1$, a divergência-J de ordem α em (1.2.11) reduz-se à divergência-J em (1.2.10).

A medida $I^{\alpha,\beta}_{\gamma,\delta}$ definida em (1.2.9) envolve mais parametros. Quando $\gamma = 1$, $\delta = 0$ e $\alpha + \beta = 1$, ela dá a divergência dirigida de ordem α. Quando $\delta=0$, $\gamma = 1$ e $\alpha = 1$, ela dá a imprecisão de ordem β.

As relação entre divergência de ordem α e medidas estatisticas tais como afinidade, estatistica χ^2 de Pearson, e medida de distancia quadratica são explanadas no Capítulo 4.

1.3 EQUAÇÕES FUNCIONAIS

A equação funcional desempenha um papel importante na caracterização de medidas de informação, desde que quase todas as medidas de informação tem "representação de soma $\Sigma\, f(\;)$", é fácil obter uma equação funcional envolvendo $f(\;)$ através do uso de poucas propriedades das medidas de informação.

Começando com a equação funcional

$$(1.3.1)\qquad \sum_{i=1}^{n}\sum_{j=1}^{m} f(x_i y_j) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i^{\alpha} f(y_j) + \sum_{i=1}^{n}\sum_{j=1}^{m} y_j^{\beta} f(x_i)$$

para $X = (x_1,\ldots,x_n) \in \Delta_n$, $Y = (y_1,\ldots,y_m) \in \Delta_m$, $\alpha \neq \beta$.

Veremos no Capítulo 2 que com a condição de normalização $f(1/2) = 1/2$, a solução mensurável é da forma

$$(1.3.2)\qquad f(x) = \frac{x^{\alpha} - x^{\beta}}{2^{1-\alpha} - 2^{1-\beta}}$$

a qual junto com

$$(1.3.3)\qquad H_{n,\alpha,\beta}(X) = \sum_{i=1}^{n} f(x_i), \qquad X = (x_1,\ldots,x_n) \in \Delta_n$$

produz

$$H_{n,\alpha,\beta}(X) = \frac{\sum_{i=1}^{n} (p_i^{\alpha} - p_i^{\beta})}{2^{1-\alpha} - 2^{1-\beta}}$$

Similarmente, começando a equação funcional

$$(1.3.4)\qquad \sum_{i=1}^{n}\sum_{j=1}^{m} F(x_i y_j, u_i v_j)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} x_i^{\alpha} u_i^{\beta} F(y_j, v_j) + \sum_{i=1}^{n}\sum_{j=1}^{m} y_j^{\gamma} v_j^{\delta} F(x_i, u_i)$$

para $\alpha,\beta,\gamma,\delta$ positivos e $(\alpha-\gamma)(\delta-\beta)<0$.

onde $X=(x_1,\ldots,x_n)$, $U=(u_1,\ldots,u_n)\ \varepsilon\ \Delta_n$, $Y=(y_1,\ldots,y_m)$, $V=(v_1,\ldots,v_m)\ \varepsilon\ \Delta_m$.

Veremos no Capítulo 2 que com a condição de normalização $F(1,1/2)=1$, a solução mensurável é da forma

$$(1.3.5)\qquad F(x,y)=\frac{x^{\alpha}y^{\beta}-x^{\gamma}y^{\delta}}{2^{-\beta}-2^{-\delta}}\qquad (\alpha-\gamma)(\delta-\beta)<0$$

a qual conjuntamente com

$$(1.3.6)\qquad I_{\gamma,\delta}^{\alpha,\beta}(X:Y)=\sum_{i=1}^{n}F(x_i,y_j)$$

para $X=(x_1,\ldots,x_n)$, $Y=(y_1,\ldots,y_n)\ \varepsilon\ \Delta_n$.

produz

$$I_{\gamma,\delta}^{\alpha,\beta}(X:Y)=\frac{\sum_{i=1}^{n}(x_i^{\alpha}y_i^{\beta}-x_i^{\gamma}y_i^{\delta})}{2^{-\beta}-2^{-\delta}}$$

As propriedades não-adtiva dos tipos

$$(1.3.7)\qquad F_{mn}(X*Y)=F_n(X)+F_m(Y)+\lambda\cdot F_n(X)\,F_m(Y)$$

$$(1.3.8)\qquad F_{mn}(X*Y:U*V)=F_n(X:U)+F_m(Y:V)+\lambda\cdot F_n(X:U)\,F_m(Y:V)$$

para $X,U\ \varepsilon\ \Delta_n$, $Y,V\ \varepsilon\ \Delta_m$.

são possuidas por diversas medidas de informação.

Assim, no Capítulo 3, definimos

$$(1.3.9)\qquad f(x)=\frac{x^{\beta}-x}{2^{1-\beta}-1}\ ;\quad F(x,y)=\frac{x^{\alpha}y^{\beta}-x}{2^{-\beta}-1}$$

Entao, pela representação de soma

(1.3.10) $H_{n,\beta}(X) = \sum_{i=1}^{n} f(x_i)$; $I_{n,\alpha,\beta}(X:Y) = \sum_{i=1}^{n} F(x_i,y_i)$

para $X = (x_1,\ldots,x_n)$, $Y = (y_1,\ldots,y_n) \in \Delta_n$

a não-aditividade produz as seguintes equções funcionais

(1.3.12)
$$\sum_{i=1}^{n}\sum_{j=1}^{m} f(x_iu_j) = \sum_{i=1}^{n} f(x_i) + \sum_{j=1}^{m} f(u_j) + (2^{1-\beta}-1)\{\sum_{i=1}^{n} f(x_i)\}\{\sum_{j=1}^{m} f(u_j)\}$$

e

(1.3.13)
$$\sum_{i=1}^{n}\sum_{j=1}^{m} F(x_iu_j,y_iv_j) = \sum_{i=1}^{n} F(x_i,y_i) + \sum_{j=1}^{m} F(u_j,v_j) + (2^{-\beta}-1)\{\sum_{i=1}^{n} F(x_i,y_i)\}\{\sum_{j=1}^{m} F(u_j,v_j)\}$$

para $X = (x_1,\ldots,x_n) \in \Delta_n$, $U = (u_1,\ldots,u_m) \in \Delta_m$, $Y = (y_1,\ldots,y_n) \in \Delta'_n$ e $V = (v_1,\ldots,v_m) \in \Delta'_m$.

Recursividade é outra propriedade importante partilhada pela maioria das medidas de informação. Recursividade em uma variável significa que a diferença de $F_n(p_1,\ldots,p_n)$ e $F_{n-1}(p_1+p_2,p_3,\ldots,p_n)$ é uma função de p_1 e p_2 sómente. Analogamente, para o caso de duas variáveis, significa que a diferença de $F_n(p_1,\ldots,p_n:q_1,\ldots,q_n)$ e $F_{n-1}(p_1+p_2,p_3,\ldots,p_n:q_1+q_2,q_3,\ldots,q_n)$ é uma função de $(p_1,p_2:q_1,q_2)$ sómente.

Assim, no Capítulo 4, temos

(1.3.14) $f(x,y) = I_{2,\alpha}(x,1-x:y,1-y)$, $\quad (x,y) \in [0,1]\times[0,1]$

e

(1.3.15) $g(x,y) = J_{2,\alpha}(x,1-x:y,1-y)$, $\quad (x,y) \in [0,1]\times[0,1]$

Então a recursividade e simetria de $I_{3,\alpha}$ da a seguinte equação funcional

$$g(x,y) + (1-x)^{\alpha}(1-y)^{1-\alpha}f\left(\frac{u}{1-x}, \frac{v}{1-y}\right) + (1-y)^{\alpha}(1-x)^{1-\alpha}f\left(\frac{v}{1-y}, \frac{u}{1-x}\right)$$
$$= g(u,v) + (1-u)^{\alpha}(1-v)^{1-\alpha}f\left(\frac{x}{1-u}, \frac{y}{1-v}\right) + (1-v)^{\alpha}(1-u)^{1-\alpha}f\left(\frac{y}{1-v}, \frac{x}{1-u}\right) \tag{1.3.16}$$

1.4 TEOREMAS DE CARACTERIZAÇÃO

Existem diversos teoremas de caracterização e resultados na estrutura de equações funcionais relacionadas a este trabalho os quais serão listados abaixo sem prova.

<u>Teorema 1.4.1</u> A única solução continua da equação funcional (1.3.1) sob a condição de normalização f(1/2) é dada por

$$f(p) = \frac{p^{\alpha} - p^{\beta}}{2^{1-\alpha} - 2^{1-\beta}}$$

A correspondente representação de soma de H em termos de f é

$$H(P) = \frac{\sum_{i=1}^{n} (p_i^{\alpha} - p_i^{\beta})}{2^{1-\alpha} - 2^{1-\beta}}$$

a qual é a entropia de ordem (α,β).

[Sharma e Taneja (1975)]

<u>Teorema 1.4.2</u> A única solução contína da equação funcional

$$(1.4.1) \qquad \sum_{i=1}^{n} \sum_{j=1}^{m} f(x_i y_j) = \sum_{i=1}^{n} f(x_i) + \sum_{j=1}^{m} f(y_j) + C\cdot \sum_{i=1}^{n} \sum_{j=1}^{m} f(x_i) f(y_j)$$

para $X = (x_1,\ldots,x_n) \in \Delta_n$, $Y = (y_1,\ldots,y_m) \in \Delta_m$.

satisfazendo a condição $f(1/2) = 1/2$ é dada por

$$f(p) = \frac{p - p^{\alpha}}{1 - 2^{1-\alpha}}, \quad p \in [0,1] \qquad \alpha \geqq 0, \quad \alpha \neq 1$$

[Behara e Nath (1973)]

<u>Teorema 1.4.3</u> As funções f, g, h $\in C[0,1]$ satisfazendo a equação funcional

$$(1.4.2) \qquad \sum_{i=1}^{m} \sum_{j=1}^{n} f(x_i y_j) = \sum_{i=1}^{m} g(x_i) + \sum_{j=1}^{m} h(y_j) + \lambda \sum_{i=1}^{m} \sum_{j=1}^{n} g(x_i) h(y_j)$$

para $X = (x_1,\dots,x_m) \in \Delta_m$, $Y = (y_1,\dots,y_n) \in \Delta_n$

são dadas por

(a) $f(x) = \frac{x}{\lambda}\,(abx^{\beta} - 1)$, $g(x) = \frac{x}{\lambda}\,(ax^{\beta-1} - 1)$, $h(x) = \frac{x}{\lambda}\,(ax^{\beta-1} - 1)$

para todo $x \in [0,1[$, onde $ab \neq 0$, $\beta \geqq 0$ são constantes arbitrárias.

(b) $f(x) = g(x) = -\,x/\lambda$, $h(x)$: arbitrária.

(c) $f(x) = h(x) = -\,x/\lambda$, $g(x)$: arbitrária.

<u>Corolário do Teorema 1.4.3</u> A entropia H de uma distribuição de probabilidade completa $P = (p_1,\dots,p_m) \in \Delta_m$ com não-aditividade e a "representação de soma" $H(P) = \sum_{i=1}^{n} f(p_i)$ onde f continua, sob a condição $f(1/2) = 1/2$ e $f(1) = 0$, é

$$H(P) = \frac{\sum_{i=1}^{n} p_i^{\beta} - 1}{2^{-\beta} - 1}, \qquad \beta > 0, \beta \neq 1$$

[Mittal (1976)]

<u>Teorema 1.4.4</u> A solução contínua da equação funcional (1.3.4) sob a condiçaõ $F(1,1/2)$ é dada por

$$F(p,q) = \frac{p^{\alpha} q^{\beta} - p^{\gamma} q^{\delta}}{2^{-\beta} - 2^{-\delta}} \qquad \alpha \neq \gamma\ ,\ (\beta \neq \delta) \text{ quando } \beta = \delta, (\alpha = \gamma)$$

A correspondente "representação de soma" de I associadada com as distribuições P e Q é

$$I(P:Q) = \frac{\sum_{i=1}^{n} (p_i^{\alpha} q_i^{\beta} - p_i^{\gamma} q_i^{\delta})}{2^{-\beta} - 2^{-\delta}}$$

onde os parametros são os mesmos como acima.

[Sharma e Taneja (1975)]

<u>Teorema 1.4.5</u> Suponha que $J(P,Q)$ satisfaz as hipotéses 1 e 2 seguintes:

Hipotése 1. $J(P,Q)$ assume a forma

$$J(P,Q) = \int F(p_\lambda(x),\ q_\lambda(x))p_\lambda(x)\ \lambda(dx)$$

onde F é função mensurável fixada. λ é uma medida tal que ambos P e Q absolutamente continuas relativas a λ enquando $p_\lambda(x)$ e $q_\lambda(x)$ são as densidades correspondentes.

Hipotése 2. Suponha que P_1,Q_1 são medidas de probabilidade em Ω_1 e P_2,Q_2 são medidas de probabilidade em Ω_2. Seja $P_1 \times P_2$ e $Q_1 \times Q_2$ as correspondentes medidas produto em $\Omega_1 \times \Omega_2$. Então

$$J(P_1 \times P_2, Q_1 \times Q_2) = J(P_1,Q_1) + J(P_2,Q_2)$$

adicionalmente com

$$J(P,Q) = J(Q,P)$$

Então $J(P,Q)$ é (a menos de uma constante multiplicativa) a divergência simétrica

$$J(P,Q) = a \int (p_\lambda - q_\lambda)\ \log\left(\frac{p_\lambda(x)}{q_\lambda(x)}\right)\ \lambda(dx)$$

[Haaland, Brockett e Levine (1979)]

<u>Teorema 1.4.6</u> Os cinco postulados seguintes determinam a J-divergência $J(P{:}Q) = \sum_{i=1}^{n} p_i \log(p_i/q_i)$ únicamente para $P = (p_1,\ldots,p_n) \in \Delta_n$, $Q = (q_1,\ldots,q_n) \in \Delta_n$.

Postulado 1. (recursividade) Para $p_1+p_2 > 0$, $q_1+q_2 > 0$ e para todo

$$J_n(p_1,\ldots,p_n:q_1,\ldots,q_n)$$

$$= J_{n-1}(p_1+p_2,p_3,\ldots,p_n:q_1+q_2,q_3,\ldots,q_n) +$$

$$(p_1+p_2)\ I_2\left(\frac{p_1}{p_1+p_2},\frac{p_2}{p_1+p_2}:\frac{q_1}{q_1+q_2},\frac{q_2}{q_1+q_2}\right) +$$

$$(q_1+q_2)\ I_2\left(\frac{q_1}{q_1+q_2},\frac{q_2}{q_1+q_2}:\frac{p_1}{p_1+p_2},\frac{p_2}{p_1+p_2}\right)$$

Postulado 2. (simetria) Para qualquer permutação arbitrária (a_1,a_2,a_3) de $(1,2,3)$,

$$J_3(p_1,p_2,p_3:q_1,q_2,q_3) = J_3(p_{a_1},p_{a_2},p_{a_3})$$

Postulado 3. (derivatividade) Seja $f(x,y) = I_2(x,1-x:y,1-y)$, f tem derivadas parciais com respeito a $x,y \in (0,1)$ até segunda ordem.

Postulado 4. (nulidade) Para $x \in [0,1]$

$$J_2(x,1-x:x,1-x) = 0$$

Postulado 5. (normalização)

$$J_2(2/3,1/3 : 1/3,2/3) = 2/3$$

[Kannappan e Rathie (1979)]

Algumas outras caracterizações dessas medidas de informação generalizadas são citadas por Daróczy (1970), Sharma e Autar (1974), Kannappan e Rathie (1973), Rathie e Kannappan (1973), Forte e Ng (1973 e 1975)

Para detalhes sobre teoremas de caracterização, estudas sobre equações funcionais relacionadas com os teoremas e referências a respeito de várias aplicações, etc., veja Mathai e Rathie (1975) e Aczél e Daróczy (1975).

1.5 APLICAÇÕES

A quantidade - log p_i é interpretada como a informação contida no evento E_i [Shannon (1948)] com probabilidade p_i. Assim $\log(p_i/q_i) = \log p_i - \log q_i$ pode ser tomada como ganho de informação em predizer o evento E_i. Para interpretações de ganho de informação em previsão de tempo e problemas economicos, veja Theil (1967). O ganho médio de informação, isto é $\sum_{i=1}^{n} p_i \log(p_i/q_i)$, é interpretado como a informação média para discriminação em favor da hipotése H_1 contra H_2 em problemas de inferência estatística [Kullback (1959)]. Kerridge (1961) interpreta $\sum_{i=1}^{n} p_i \log(p_i/q_i)$ como uma medida do erro feita por um observador ao estimar um distribuição de probabilidade discreta como $Q = (q_1,\ldots,q_n)$ a qual, de fato, é $P = (p_1,\ldots,p_n)$.

Em análise de informação obtida de questionários, a divergência -J $J(P:Q)$ é interpretada como a quantidade de informação esperada em uma questão para distinguir entre individuos do grupo um e do grupo dois. [Haaland, Brockett e Levine (1979)]:

Desde que o conceito de afinidade pode ser interpretada como a distância angular entre duas distribuições, este conceito, em algun sentido, mede a proximidade das distribuições. Para aplicações da afinidade em estimação pontual e por intervalo, veja [Matusita (1954),(1955), (1961)]. É mostrado em George e Mathai (1974) que afinidade é muito útil em classificação de população de acordo com algumas características socio-antropológica.

CAPÍTULO 2

EQUAÇÕES FUNCIONAIS E CARACTERIZAÇÕES DE MEDIDAS DE INFORMAÇÃO

RESUMO

Generalizaçoes da equação funcional de Chaundy e McLeod (1960) em uma variával e duas variáveis são discutidas sob o ponto de vista de mensurábilidade. Uma das equações é util na caracterização axiomática da entropia não aditiva de ordem β .

2.1 INTRODUÇÃO

Seja $\Delta_n = \{(p_1,\ldots,p_n): p_i \geqq 0$ para $i=1,\ldots,n,\ \sum_{i=1}^{n} p_i = 1\}$ para $n \geqq 1$ denote o conjunto de todas n-uplas de distribuições de probabilidade.

Neste capítulo, nós consideramos duas equações funcionais.

(i) Equação Funcional em Uma Variável

Seja $f: [0,1] \to R$ (Reais) mensuravel (ou continua em um ponto, ou limitada em um intervalo pequeno) e satisfazendo a equação funcional

$$(2.1.1) \qquad \sum_{i=1}^{m} \sum_{j=1}^{n} f(x_i y_j) = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i^{\alpha} f(y_j) + \sum_{i=1}^{m} \sum_{j=1}^{n} y_j^{\beta} f(x_i)$$

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

onde $X = (x_1,\ldots,x_m) \in \Delta_m$, $Y = (y_1,\ldots,y_n) \in \Delta_n$ e α, β são parametros.

Quando $\alpha = \beta = 1$, (2.1.1) reduz-se a equação funcional de Chaundy e McLeod (1960)

$$(2.1.2) \qquad \sum_{i=1}^{m} \sum_{j=1}^{n} f(x_i y_j) = \sum_{i=1}^{m} f(x_i) + \sum_{j=1}^{n} f(y_j)$$

As soluções da equação funcional generalizada em uma variável (2.1.1) conduz a uma caracterização da entropia de evento. A derivação da equação funcional (2.1.1) através da construção da informação é discutida abaixo.

Considere uma variável aleatória discreta X assumindo valores finitos nos eventos $E_1, E_2, \ldots, E_n$ com distribuições de probabilidade $P = (p_1, \ldots, p_n)$, $\sum_{i=1}^{n} p_i = 1$.

A entropia de ordem (α,β) de uma distribuição de probabilidade $P = (p_1, \ldots, p_n)$ é definida como

$$(2.1.3) \qquad H_{n,\alpha,\beta}(P) = \frac{\sum_{i=1}^{n} (p_i^{\alpha} - p_i^{\beta})}{2^{1-\alpha} - 2^{1-\beta}}$$

Quando $\alpha = 1$, (2.1.3) reduz-se a entropia de ordem β [Havrda e Charvát (1967)] ou [Daróczy (1970)].

$$(2.1.4) \qquad H_{n,\beta}(P) = \frac{\sum_{i=1}^{n} p_i^{\beta} - 1}{2^{1-\beta} - 1}$$

Quando $\beta \to 1$. (2.1.4) reduz-se a entropia de Shannon $- \sum_{i=1}^{n} p_i \log p_i$.

Assim (2.1.3) é uma generalização a qual inclui não sòmente a entropia de Shannon mas também a entropia de ordem β como um caso partícular.

Considere

$$(2.1.5) \qquad f(p) = \frac{p^{\alpha} - p^{\beta}}{2^{1-\alpha} - 2^{1-\beta}}$$

Então a entropia de ordem (α,β) de uma distribuição de probabilidade $P=(p_1, \ldots, p_n)$ é

(2.1.6) $H_{n,\alpha,\beta}(p_1,\ldots,p_n) = \sum_{i=1}^{n} f(p_i)$

Seja $f(p_i)$ a entropia de ordem (α,β) do eventos E_i, $i=1,\ldots,n$, em outras palavras, $f(p_i)$ ẽ um tipo de informação contido no evento E_i, a incerteza do evento E_i. Ê claro que $f(p)$ tem a condição de nomalização que $f(1/2) = 1/2$.

Para duas distribuições de probabilidade $P = (p_1,\ldots,p_n) \in \Delta_n$, $Q = (q_1,\ldots,q_m) \in \Delta_m$, usando (2.1.5), a entropia de ordem (α,β) do produto de P e Q ẽ da seguinte forma

(2.1.7) $H_{mn,\alpha\ \beta}(P\ Q) = \sum_{i=1}^{n}\sum_{j=1}^{m} f(p_i q_j)$

$$= \frac{1}{2^{1-\alpha}-2^{1-\beta}} \sum_{i=1}^{n}\sum_{j=1}^{m} (p_i^\alpha q_j^\alpha - p_i^\beta q_j^\beta)$$

$$= \frac{1}{2^{1-\alpha}-2^{1-\beta}} \left\{ \sum_{i=1}^{n}\sum_{j=1}^{m} p_i^\alpha (q_j^\alpha - q_j^\beta) - \sum_{i=1}^{n}\sum_{j=1}^{m} q_j^\beta (p_i^\alpha - p_i^\beta) \right\}$$

Isto da a equação funcional (2.1.1).

(ii) Equação Funcional em Duas Variãvels

Seja F: $[0,1]\times[0,1] \to R$ (Reais) mensurãval em cada variãvel e satisfazendo a equação funcional

(2.1.8) $\sum_{i=1}^{n}\sum_{j=1}^{m} F(x_i y_j, u_i v_j)$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} x_i^\alpha u_i^\beta F(y_j,v_j) + \sum_{i=1}^{n}\sum_{j=1}^{m} y_j^\gamma v_j^\delta F(x_i,u_i)$$

para $\alpha,\beta,\gamma,\delta > 0, (\alpha-\gamma)(\delta-\beta) < 0$.

onde $X=(x_1,\ldots,x_n)$, $U = (u_1,\ldots,u_n) \in \Delta_m$ e $Y = (y_1,\ldots,y_m)$, $V=(v_1,\ldots,v_m)$

$\varepsilon\ \Delta_m$, $n = 2,3$; $n = 2,3$

e $F(1,0) = 0$.

Esta é uma generalização da equação funcional de Chaundy e McLeod em duas variáveis.

As soluçõea da equação funcional generalizada em duas variáveis (2.1.8) da uma medida de informação teórica para um evento associado com duas distribuições de probabilidade. A derivação do equação funcional (2.1.8) através da construção da informação teórica é discutida abaixo.

Considere uma variável aleatória discreta X assumindo um número finito de valores nos eventos $E_1, E_2, \ldots, E_n$, sejam duas distribuições de probabilidade $P = (p_1, \ldots, p_n)$ e $Q = (q_1, \ldots, q_n)$ em Δ_n associadas a X.

Uma medida de informação teórica associada com as distribuições de probabilidade P e Q é

$$(2.1.9) \qquad I^{\alpha,\beta}_{\gamma,\delta}\ (P:Q) = \frac{\sum_{i=1}^{n} (p_i^{\alpha} q_i^{\beta} - p_i^{\gamma} q_i^{\delta})}{2^{-\beta} - 2^{-\delta}}$$

Se $\gamma = 1$, $\delta = 0$ e $\alpha+\beta = 1$, (2.1.9) reduz-se à divergência dirigida de ordem α [Rathie e Kannappan (1972)].

$$(2.1.10) \qquad I_{\alpha}\ (P:Q) = \frac{\sum_{i=1}^{n} p_i^{\alpha} q_i^{1-\alpha} - 1}{2^{\alpha-1} - 1}$$

Quando $\alpha \to 1$, (2.1.10) reduz-se a

$$(2.1.11) \qquad I\ (P:Q) = \sum_{i=1}^{n} p_i \log(p_i/q_i)$$

a qual é a divergência de Kullback (1959).

Em (2.1.9), se $\delta = 0$, $\gamma = 1$ e $\alpha = 1$, obtemos a imprecisão de ordem β [Rathie e Kannappan (1973)].

$$(2.1.12) \quad K^{\beta}(P:Q) = \frac{\sum_{i=1}^{n} p_i q_i^{\beta} - 1}{2^{-\beta} - 1}$$

Quando $\beta \to 0$, (2.1.12) reduz-se a imprecisão de Kerridge (1961)

$$(2.1.13) \quad K(P:Q) = -\sum_{i=1}^{n} p_i \log q_i$$

Assim (2.1.9) é uma generalização que inclui a divergência dirigida não-aditiva de ordem α, a divergência de Kullback, imprecisão de ordem β, e imprecisão de Kerridge como casos particulares.

Considere

$$(2.1.14) \quad F(p,q) = \frac{p^{\alpha} q^{\beta} - p^{\gamma} q^{\delta}}{2^{-\beta} - 2^{-\delta}}$$

Então a medida de informação teórica em (2.1.9) pode ser escrita como

$$(2.1.15) \quad I_{\gamma,\delta}^{\alpha,\beta}(P:Q) = \sum_{i=1}^{n} F(p_i,q_i)$$

Aqui $F(p_i,q_i)$ denota a medida de informação teórica contido no eventos E_i, $i=1,\ldots,n$, é claro que $F(p,q)$ tem a condição de normalização que $F(1.1/2) = 1$.

Para distribuições de probabilidade $X = (x_1,\ldots,x_n)$, $U = (u_1,\ldots,u_n)$ em Δ_n e $Y = (y_1,\ldots,y_m)$, $V = (v_1,\ldots,v_m)$ em Δ_m, $I_{\gamma,\delta}^{\alpha,\beta}$ tem a seguinte relação

$$(2.1.16) \quad I_{\gamma,\delta}^{\alpha,\beta}(X*Y: U*V)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} F(x_i y_j, u_i v_j)$$

$$= (2^{-\beta} - 2^{-\delta})^{-1} \{ \sum_{i=1}^{n} \sum_{j=1}^{m} [(x_i y_j)^{\alpha} (u_i v_j)^{\beta} - (x_i y_j)^{\gamma} (u_i v_j)^{\delta}] \}$$

$$= (2^{-\beta} - 2^{-\delta})^{-1} \{ \sum_{i=1}^{n} \sum_{j=1}^{m} [x_i^{\alpha} u_i^{\beta} (y_j^{\alpha} v_j^{\beta} - y_j^{\gamma} v_j^{\delta}) + y_j^{\gamma} v_j^{\delta} (x_i^{\alpha} u_i^{\beta} - x_i^{\gamma} u_i^{\delta})] \}$$

Isto da a equação funcional (2.1.8).

O objectivo deste capítulo é encontra a solução mensurável (ou continua em um ponto, ou limitada em um intervalo pequeno) de (2.1.1) e (2.1.8).

Assim a equação funcional (2.1.1) foi resolvida por Sharma e Taneja (1975) sob a condição de continuidade. No mesmo artígo (2.1.8) foi resolvida son condição de continuidade para

$$\sum_{i=1}^{n} x_i = \sum_{j=1}^{m} y_j = 1, \quad \sum_{i=1}^{n} u_i \leqq 1 \text{ e } \sum_{j=1}^{m} v_j \leqq 1.$$

o tratamento feito por eles não é claro em vários lugares porque os domínios dos parametros não são propriamente definidos.

2.2 PROPRIEDADES

I. Para A Entropia de Ordem (α,β)

$$H_{n,\alpha,\beta}(P) = \frac{\sum_{i=1}^{n}(p_i^{\alpha} - p_i^{\beta})}{2^{1-\alpha} - 2^{1-\beta}} \qquad \alpha > 0,\ \beta > 0,\ \alpha \neq \beta$$

(i) Não-Negatividade

$H_{n,\alpha,\beta}(P)$ é não-negativa, e é zero se e sòmente se $p_i = 1$ para algun $i = 1,2,\ldots,n$. Isto é claro da definição de $H_{n,\alpha,\beta}(P)$.

Isto significa que a entropia de ordem (α,β) para a distribuição de probabilidade é generalmente positiva e será zero se e sòmente se um dos enevtos acontece com probabilidade 1.

(ii) Simetria

$$H_{n,\alpha,\beta}(p_1,\ldots,p_n) = H_{n,\alpha,\beta}(p_{a(1)},\ldots,p_{a(n)})$$

onde $\{a(1),\ldots,a(n)\}$ é uma permutação arbitrária de $\{1,\ldots,n\}$.

A entropia não depende da ordem em que os eventos são fixados.

(iii) Expandibilidade

$$H_{n,\alpha,\beta}(p_1,\ldots,p_n) = H_{n+1,\alpha,\beta}(p_1,\ldots,p_n,o)$$

A adição de um evento de probabilidade zero não altera o valor da entropia.

(iv) Recursividade

$$H_{n,\alpha,\beta}(p_1,\ldots,p_n)$$

$$= H_{n-1,\alpha,\beta}(p_1+p_2,p_3,\ldots,p_n) +$$

$$\frac{1}{2^{1-\alpha} - 2^{1-\beta}}\{(2^{1-\alpha} - 1)(p_1+p_2)^{\alpha} H_{2,\alpha}\left(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2}\right) -$$

$$(2^{1-\beta}-1)(p_1+p_2)^\beta H_{2,\beta}\left(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2}\right)\}$$

para $p_1+p_2 > 0$, e onde $H_{2,\alpha}$ ($H_{2,\beta}$) é a entropia não-aditiva de ordem α (β).

(v) Não-Aditividade

Para $P \in \Delta_n$, $Q \in \Delta_m$

$$H_{mn,\alpha,\beta}(P\,Q) = H_{n,\alpha,\beta}(P) + H_{m,\alpha,\beta}(Q) + \frac{1}{2^{1-\alpha}-2^{1-\beta}} \times$$

$$\{(2^{1-\alpha}-1)^2 H_{n,\alpha}(P) H_{m,\alpha}(Q) - (2^{1-\beta}-1)^2 H_{n,\beta}(P) H_{m,\beta}(Q)\}$$

onde $H_{n,\alpha}$ ($H_{n,\beta}$) é a entropia não-aditiva de ordem α (β).

(vi) Não-Aditividade Forte

Para $P = (p_1,\ldots,p_m) \in \Delta_m$; $Q_j = (q_{j1},\ldots,q_{jn}) \in \Delta_n$, $j = 1,\ldots,m$.

$$H_{mn,\alpha,\beta}(p_1q_{11},\ldots,p_1q_{1n},\ldots,p_mq_{m1},\ldots,p_mq_{mn})$$

$$= H_{m,\alpha,\beta}(P) + \frac{1}{2^{1-\alpha}-2^{1-\beta}} \{2^{1-\alpha}-1) \sum_{j=1}^{m} p_j^\alpha H_{n,}\ (Q_j) +$$

$$(2^{1-\beta}-1) \sum_{j=1}^{m} p_j^\beta H_{m,\beta}(Q_j)\}$$

onde $H_{n,\alpha}$ ($H_{n,\beta}$) é a entropia não-aditiva de ordem α (β).

(vii) Continuidade

$H_{n,\alpha,\beta}(P)$ é uma função continua de n variáveis.

(viii) Normalização

$H_{2,\alpha}(1/2,1/2) = 1$

II. Para A Medida de Informação Teórica $I_{\gamma,\delta}^{\alpha,\beta}$

$$I_{\gamma,\delta}^{\alpha,\beta}(P:Q) = \frac{\sum_{i=1}^{n} (p_i^\alpha q_i^\beta - p_i^\gamma q_i^\delta)}{2^{-\beta} - 2^{-\delta}} , \qquad (\alpha - \gamma)(\delta - \beta) < 0$$

(i) Não-Negatividade

$I_{\gamma,\delta}^{\alpha,\beta}(P:Q)$ é não-negativa, e é zero se e sómente se $p_i = q_i = 1$ para algun $i = 1,\ldots,n$. Isto é claro da definição de $I_{\gamma,\delta}^{\alpha,\beta}(P:Q)$.

Isto significa que a medida de informação entre duas distribuições de probabilidade é geralmente positiva e será zero se e sómente se as duas distribuições coincidem e um dos eventos acontece com probabilidade 1.

(ii) Simetria

$$I_{\gamma,\delta}^{\alpha,\beta}(p_1,\ldots,p_n:q_1,\ldots,q_n) = I_{\gamma,\delta}^{\alpha,\beta}(p_{a(1)},\ldots,p_{a(n)}:q_{a(1)},\ldots,q_{a(n)})$$

onde $\{a(1),\ldots,a(n)\}$ é uma permutação arbitrária de $\{1,\ldots,n\}$.

A permutação conjunta de indices correspondentes de eventos não muda o valor da medida de informação teórica $I_{\gamma,\delta}^{\alpha,\beta}(P:Q)$.

(iii) Expandibilidade

$$I_{\gamma,\delta}^{\alpha,\beta}(p_1,\ldots,p_n:q_1,\ldots,q_n) = I_{\gamma,\delta}^{\alpha,\beta}(p_1,\ldots,p_n,0:q_1,\ldots,q_n,0)$$

A adição de um evento de probabilidade zero não altera o valor da medida de informação teórica $I_{\gamma,\delta}^{\alpha,\beta}(P:Q)$.

(iv) Recursividade

$$I_{\gamma,\delta}^{\alpha,\beta}(p_1,\ldots,p_n:q_1,\ldots,q_n)$$

$$= I_{\gamma,\delta}^{\alpha,\beta}(p_1+p_2,p_3,\ldots,p_n:q_1+q_2,q_3,\ldots,q_n) +$$

$$\frac{1}{2^{-\beta}-2^{-\delta}}\{(2^{-\beta}-1)(p_1+p_2)^{\alpha}(q_1+q_2)^{\beta}I_{\alpha,\beta}\left(\frac{p_1}{p_1+p_2},\frac{p_2}{p_1+p_2}:\frac{q_1}{q_1+q_2},\frac{q_2}{q_1+q_2}\right)$$

$$-(2^{-\delta}-1)(p_1+p_2)^{\gamma}(q_1+q_2)^{\delta}\ I_{\gamma,\delta}\left(\frac{p_1}{p_1+p_2},\frac{p_2}{p_1+p_2}:\frac{q_1}{q_1+q_2},\frac{q_2}{q_1+q_2}\right)\}$$

para $p_1+p_2>0$, $q_1+q_2>0$.

onde $I_{\alpha,\beta}$ $(I_{\gamma,\delta})$ é a divergência dirigida não-aditiva de ordem (α,β) $((\gamma,\delta))$.

(v) Não-Aditividade

Para P, Q ε Δ_n e R, S ε Δ_m

$$I^{\alpha,\beta}_{\gamma,\delta}\ (P*R : Q*S)$$

$$= I^{\alpha,\beta}_{\gamma,\delta}\ (P:Q) + I^{\alpha,\beta}_{\gamma,\delta}\ (R:S) + \frac{1}{2^{-\beta}-2^{-\delta}}\{(2^{-\beta}-1)^2 I_{\alpha,\beta}(P:Q)I_{\alpha,\beta}(R:S) -$$

$$(2^{-\delta}-1)^2 I_{\gamma,\delta}(P:R)I_{\gamma,\delta}(Q:S)\}$$

Isto é claro das seguintes relações:

$$I^{\alpha,\beta}_{\gamma,\delta}\ (P:Q) = \frac{1}{2^{-\beta}-2^{-\delta}}\{(2^{-\beta}-1)I_{\alpha,\beta}(P:Q) - (2^{-\delta}-1)I_{\gamma,\delta}(P:Q)\}$$

e

$$I_{\alpha,\beta}(P*R : Q*S) = I_{\alpha,\beta}(P:Q) + I_{\alpha,\beta}(R:S) + (2^{-\beta}-1)I_{\alpha,\beta}(P:Q)I_{\alpha,\beta}(R:S)$$

onde $I_{\alpha,\beta}$ é a divergência dirigida não-adtiva de ordem (α,β).

(vi) Não-Aditividade Forte

Para $(p_1,\ldots,p_n)$, $(q_1,\ldots,q_n)$ ε Δ_n, e $(P_{1i},\ldots,P_{mi})$, $(Q_{1i},\ldots,Q_{mi})$ ε Δ_m para $i = 1,\ldots,n$.

$$I^{\alpha,\beta}_{\gamma,\delta}\ (p_1P_{11},\ldots,p_1P_{m1},\ldots,p_nP_{1n},\ldots,p_nP_{mn}: q_1Q_{11},\ldots,q_1Q_{m1},\ldots,$$

$$q_nQ_{1n},\ldots,q_nQ_{mn})$$

$$= I_{\gamma,\delta}^{\alpha,\beta}(p_1,\ldots,p_n:q_1,\ldots,q_n) + \frac{1}{2^{-\beta}-2^{-\delta}} \times$$

$$\sum_{i=1}^{n} \{(2^{-\beta}-1)p_i^{\alpha} q_i^{\beta} I_{\alpha,\beta}(P_{1i},\ldots,P_{mi}:Q_{1i},\ldots,Q_{mi}) -$$

$$(2^{-\delta}-1)p_i^{\gamma} q_i^{\delta} I_{\gamma,\delta}(P_{1i},\ldots,P_{mi}:Q_{1i},\ldots,Q_{mi})\}$$

Isto ẽ claro da seguinte relação

$$I_{\gamma,\delta}^{\alpha,\beta}(p_1,\ldots,p_n:q_1,\ldots,q_n)$$

$$= I_{\gamma,\delta}^{\alpha,\beta}(p_1+\ldots+p_k,p_{k+1},\ldots,p_n:q_1+\ldots+q_k,q_{k+1},\ldots,q_n) +$$

$$\frac{1}{2^{-\beta}-2^{-\delta}}\{(2^{-\beta}-1)r_k^{\alpha} s_k^{\beta} I_{\alpha,\beta}(\frac{p_1}{r_k},\ldots,\frac{p_k}{r_k}:\frac{q_1}{s_k},\ldots,\frac{q_k}{s_k}) -$$

$$(2^{-\delta}-1)r_k^{\gamma} s_k^{\delta} I_{\gamma,\delta}(\frac{p_1}{r_k},\ldots,\frac{p_k}{r_k}:\frac{q_1}{s_k},\ldots,\frac{q_k}{s_k})$$

para $r_k = p_1+\ldots+p_k$, $s_k = q_1+\ldots+q_k$, $1 \leqq k \leqq n$.

onde $I_{\alpha,\beta}$ $(I_{\gamma,\delta})$ ẽ a divergência dirigida não-adtiva de ordem (α,β) $((\gamma,\delta))$.

(vii) Continuidade

$I_{\gamma,\delta}^{\alpha,\beta}(P:Q)$ ẽ uma funçaõ continua de 2n variãveis, para $P, Q \in \Delta_n$.

(viii) Normalização

$$I_{\gamma,\delta}^{\alpha,\beta}(1,0:1/2,1/2) = 1$$

2.3 A SOLUÇÃO DE (2.1.1)

A fim de derivar a solução mensurável da equação funcional (2.1.1), nós provamos alguns lemas que seguem

Lema 2.3.1 $f(0) = f(1) = 0$

Prova:

Substituindo $(x_1,x_2) = (0,1)$, $(y_1,y_2,y_3) = (0,1,0)$ em (2.1.1). então

$$2f(0) = f(1) \tag{2.3.1}$$

Substituindo $(x_1,x_2) = (0,1)$, $(y_1,y_2,y_3) = (y,1-y,0)$ para $y \in (0,1)$ em (2.1.1), produz

$$3f(0) = \{y^{\beta} + (1-y)^{\beta}\}\{f(1) + f(0)\}, \qquad y \in (0,1) \tag{2.3.2}$$

Usando (2.3.1), (2.3.2) produz $f(0) = 0$, e protanto de (2.3.1), temos

$$f(1) = f(0) = 0 \tag{2.3.3}$$

Lema 2.3.2 Para $X = (x_1,\dots,x_n) \in \Delta_n$, $n = 2,3$

$$\sum_{i=1}^{n} f(x_i) = C\cdot\sum_{i=1}^{n} [x_i^{\alpha} - x_i^{\beta}] \qquad \alpha,\beta > 0\ , \qquad \alpha \neq \beta$$

onde C é uma constante arbitraria.

Prova:

Para $n = 2,3$, desde que

$$\sum_{i=1}^{n}\sum_{j=1}^{n} f(x_i y_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} f(y_j x_i)$$

Usando (2.1.1), Temos

$$(\sum_{i=1}^{n} x_i^{\alpha})\{\sum_{j=1}^{n} f(y_j)\} + (\sum_{j=1}^{n} y_j^{\beta})\{\sum_{i=1}^{n} f(x_i)\}$$

$$= \Big(\sum_{j=1}^{n} y_j^{\alpha}\Big)\Big\{\sum_{i=1}^{n} f(x_i)\Big\} + \Big(\sum_{i=1}^{n} x_i^{\beta}\Big)\Big\{\sum_{j=1}^{n} f(y_j)\Big\}$$

Assim ,

$$\frac{\sum_{j=1}^{n} f(y_j)}{\sum_{j=1}^{n} (y_j^{\alpha} - y_j^{\beta})} = \frac{\sum_{i=1}^{n} f(x_i)}{\sum_{i=1}^{n} (x_i^{\alpha} - x_i^{\beta})} = C$$

para x_i, $y_j \in (0,1)$

Note que desde que $\alpha \neq \beta$, o denominador não se anulará. Assim

(2.3.4) $$\sum_{i=1}^{n} f(x_i) = C\cdot\sum_{i=1}^{n} \left[x_i^{\alpha} - x_i^{\beta}\right], \qquad x_i \in (0,1)$$

O qual junto com o fato que $f(1) = f(0) = 0$, temos (2.3.3) verdadeiro para todo $x \in [0,1)$, com $\sum_{i=1}^{n} x_i = 1$.

Observação: Lema 2.3.2 é provado acima sem condição de regularidade. Quando $\alpha = 1$ e $f(1/2)= 1/2$, o lema 2.2.2 da

$$\sum_{i=1}^{n} f(x_i) = \frac{\sum_{i=1}^{n} x_i - 1}{2^{1-\beta} - 1}$$

o qual é a entropia não-aditiva de ordem β. Veja outra equação funcional em Rathie e Kannappan (1971). O mesmo resultato com um teorema na p.213 foi provado por Sharma e Taneja (1975) assumindo continuidade.

<u>Lema 2.3.3</u> Para x fixo, $x \in [0,1]$, se

(2.3.5) $$A_x(t) = f(xt) + f((1-x)t) - [x^{\alpha} + (1-x)^{\alpha}]f(t) - t^{\beta}[f(x)+f(1+x)]$$

para $t \in [0,1]$

Então

(2.3.6) $$A_x(u+v) = A_x(u) + A_x(v), \qquad u,v,u+v \in [0.1]$$

Prova:

Seja $(x_1,x_2) = (x,1-x)$, $(y_1,y_2,y_3) = (u,v,1-u-v)$ em (2.1.1)

para $x,u,v,u+v \in [0,1]$. Então

(2.3.7) $f(xu) + f((1-x)u) + f(xv) + f((1-x)v) + f(x(1-u-v)) +$

$f((1-x)(1-u-v))$

$= \{x^\alpha+(1-x)^\alpha\}\{f(u) + f(v) + f(1-u-v)\} + \{u^\beta+v^\beta+(1-u-v)^\beta\}.$

$= \{f(x) + f(1-x)\}$

Seja $(x_1,x_2) = (x,1-x)$, $(y_1,y_2,y_3) = (u+v,\ 1-u-v,0)$ em (2.1.1)

para $x,u,v,u+v$ como acima, então usando Lema 2.3.1

(2.3.8) $f(x(u+v)) + f((1-x)(1-u-v)) + f(x(1-u-v)) + f((1-x)(1-u-v))$

$= \{x^\alpha+(1-x)^\alpha\}\{f(u+v) + f(1-u-v)\} + \{(u+v)^\beta+(1-u-v)^\beta\}.$

$\{f(x) + f(1-x)\}$

Substraindo (2.3.8) de (2.3.7)

(2.3.9) $f(x(u+v)) + f((1-x)(u+v)) - f(xu) - f((1-x)u)-f(xv)-f((1-x)v)$

$= \{x^\alpha+(1-x)^\alpha\}\{f(u+v) - f(u) - f(v)\} + \{(u+v)^\beta-u^\beta-v^\beta\}\{f(x) + f(1-x)\}$

Para x fixo, $x \in [0,1]$, seja

$A_x(t) = f(xt) + f((1-x)t) - \{x^\alpha+(1-x)^\alpha\}f(t) - t^\beta\{f(x) + f((1-x)\}$

Então (2.3.9) fornece-se

$A_x(u+v) = A_x(u) + A_x(v)$, $\quad u,v,u+v \in [0,1]$

Isto significa que para x fixo, $x \in [0,1]$, $A_x(t)$ é aditiva em t.

O seguinte lema foi provado por Daróczy e Losonczi (1967),

portanto a prova é omitida.

<u>Lema 2.3.4</u> Qualquer solução de equação funcional de Cauchy

(2.3.10) $g(x+y) = g(x) + g(y)$

Pode ser extendida de

(2.3.11) $\{(x,y): x,y,x+y \in [0,1]\}$

para R^2. Isto é, se g é definida em $[0,1]$ e satisfaz (2.3.10) em (2.3.11), então existe uma equação funcional de Cauchy na reta real

$$h(x+y) = h(x) + h(y), \qquad x,y \in (-\infty,\infty)$$

e tal que

$$h(x) = g(x) \quad , \qquad x \in [0,1]$$

A prova do seguinte lema é também omitida [ver Aczél (1975)].

<u>Lema 2.3.5</u> Se a equação funcional de Cauchy

$$h(x+y) = h(x) + h(y)$$

é satisfeita por todos os reais x,y, e se a função $h(x)$ é mensurável (ou continua em um ponto, ou limitada em um intervalo pequeno), então

$$h(x) = cx$$

para todo real x, e onde c é uma constante arbitraria.

Agora, nós provamos o seguinte teorema.

<u>Teorema 2.3.1</u> Se f: $[0,1] \rightarrow R$ (Reais) satisfaz a equação funcional (2.1.1) e f tem qualquer das seguintes propriedades:

(a) f é continua em um ponto.

(b) f é limitada em um intervalo pequeno.

(c) f é mensurável.

Então

$$f(x) = C\cdot\{x^{\alpha} - x^{\beta}\} \qquad \alpha,\beta > 0,\ \alpha \neq \beta$$

onde C é uma constante arbitraria.

Prova:

Desde que $A_x(t)$ definida em (2.3.5) é mensurável (ou continua em um ponto, ou limitada em um intervalo pequeno), entretanto, como mostrado no Lema 2.3.3, para x fixo, $x \in [0,1]$, $A_x(t)$ é aditiva em t na região definida em (2.3.11), concluimos pelos Lemas 2.3.4 e 2.3.5 que

(2.3.12) $A_x(t) = t \cdot A_x(1)$, $\qquad t \in [0,1]$

Desde que $f(1) = 0$, de (2.3.5)

$$A_x(1) = f(x) + f(1-x) - \{x^\alpha + (1-x)^\alpha\} f(1) - f(1-x) - f(x) = 0$$

Este fato conduz a

$$A_x(t) = 0 , \qquad t \in [0,1]$$

ou

(2.3.13) $f(xu) + f((1-x)u)$

$$= \{x^\alpha + (1-x)^\alpha\} f(u) + u^\beta \{f(x) + f(1-x)\}$$

Usando Lema 2.3.2 para $f(x) + f(1-x)$, temos

(2.3.14) $f(xu) + f((1-x)u)$

$$= \{x^\alpha + (1-x)^\alpha\} f(u) + C \cdot u^\beta \{x^\alpha - x^\beta + (1-x)^\alpha - (1-x)^\beta\}$$

Novamente do Lema 2.3.2, temos que

(2.3.15) $f(1-u) + f(xu) + f((1-x)u)$

$$= C\{(1-u)^\alpha - (1-u)^\beta + (xu)^\alpha - (xu)^\beta + [(1-x)u]^\alpha - [(1-x)u]^\beta\}$$

Substraindo (2.3.14) de (2.3.15)

(2.3.16) $f(1-u) = C\{(1-u)^\alpha - (1-u)^\beta + (xu)^\alpha - (xu)^\beta + [(1-x)u]^\alpha - [(1-x)u]^\beta\} -$

$$\{x^\alpha + (1-x)^\alpha\} f(u) - u^\beta \{f(x) + f(1-x)\}$$

Substituindo $x = 1/2$ em (2.3.16),

(2.3.17) $f(1-u) = C \cdot \{(1-u)^\alpha - (1-u)^\beta + 2(u/2)^\alpha - 2(u/2)^\beta\} - 2^{1-\alpha} f(u) -$

$$u^\beta \{2^{1-\alpha} - 2^{1-\beta}\} \cdot C$$

Mas de (2.3.4)

(2.3.18) $f(u) + f(1-u) = C\{u^\alpha - u^\beta + (1-u)^\alpha - (1-u)^\beta\}$

Combinando (2.3.17) e (2.3.18), temos

(2.3.19) $f(u) = C\{u^{\alpha}-u^{\beta}\}, \qquad \alpha \neq 1, \qquad u \in |0,1|$

Para $\alpha = 1$, o lema 2.3.2 da

(2.3.20) $\sum_{i=1}^{n} f(x_i) = C\{1 - \sum_{i=1}^{n} x_i^{\beta}\}$

Portanto (2.1.1) e (2.3.20) produz

(2.3.21) $\sum_{i=1}^{m} \sum_{j=1}^{m} f(x_i y_j)$

$$= \sum_{j=1}^{m} f(y_j) + C(\sum_{j=1}^{m} y_j^{\beta})(1 - \sum_{i=1}^{m} x_i^{\beta})$$

$$= \sum_{j=1}^{m} \{f(y_j)+Cy_j^{\beta}\} - C\sum_{i=1}^{m} \sum_{j=1}^{m} x_i^{\beta} y_j^{\beta}$$

Seja

(2.3.22) $h(t) = f(t) + Ct^{\beta}$

Portanto (2.3.21) torna-se

(2.3.23) $\sum_{i=1}^{m} \sum_{j=1}^{m} h(x_i y_j) = \sum_{j=1}^{m} h(y_j)$

O fato $f(0) = f(1) = 0$ conduz a

(2.3.24) $h(0) = 0, \ h(1) = C$

Mas desde que

$$\sum_{i=1}^{m} \sum_{j=1}^{m} h(x_i y_j) = \sum_{i=1}^{m} \sum_{j=1}^{m} h(y_j x_i)$$

É claro que

(2.3.25) $\sum_{j=1}^{m} h(y_j) = \sum_{i=1}^{m} h(x_i)$

para $X = (x_1,\ldots,x_m)$, $Y = (y_1,\ldots,y_m)$ em Δ_m e $m = 2,3$.

De (2.3.25), é claro que

(2.3.26) $\sum_{i=1}^{m} h(x_i) = C,$ para $m = 2,3$

onde C é uma constante.

Para x fixo, $x \in [0,1]$, se tomamos

(2.3.27) $A_x(t) = h(xt) + h((1-x)t),$ $t \in [0,1]$

E usando o método emprecado no lema 2.3.3, com a ajuda de (2.3.26), obtemos

(2.3.28) $A_x(u+v) = A_x(u) + A_x(v)$

para $u,v,u+v \in [0,1]$.

Novamente usando o lema 2.3.4, o qual é um resultado de Daróczy e Losonczi (1967), concluimos que

(2.3.29) $A_x(u) = u\cdot A_x(1)$

Desde que $A_x(t)$ é definida como em (2.3.27), junto com (2.3.26) temos

(2.3.30) $A_x(1) = h(x) + h(1-x) = C$

Portanto (2.3.29) torna-se

(2.3.31) $A_x(u) = Cu$

ou

(2.3.32) $h(xu) + h((1-x)u) = Cu,$ $u,x \in [0,1]$

Para $x = 1$, (2.3.32) produz

(2.3.33) $h(u) = Cu$

Com a definição de $h(t)$ em (2.3.22), nós vemos que

(2.3.34) $f(u) = C\{u-u^{\beta}\},$ $u \in [0,1]$

Assim (2.3.19) (quando $\alpha \neq 1$) e (2.3.34) (quand0 $\alpha = 1$)

provamos o Teorema 2.3.1.

Como uma aplicação do Teorema 2.3.1, nós temos o seguinte Lema.

<u>Lema 2.3.6</u> A função mensurável $f\colon [0,1] \to R$ (Reais) satisfaendo a equação funcional (2.1.1) e com a condição de normalização $f(1/2) = 1/2$, então $f(x)$ é dado por

$$f(x) = \frac{1}{2^{1-\alpha} - 2^{1-\beta}} (x^{\alpha} - x^{\beta}) \tag{2.3.35}$$

a qual é entropia de um evento com probabilidade x.

Prova:

As solução mensurável da equação funcional (2.1.1) é dada na seguinte forma

$$f(x) = C\{x^{\alpha} - x^{\beta}\}$$

Usando a condição de normalização $f(1/2) = 1/2$, é fácil ver que

$$C = (2^{1-\alpha} - 2^{1-\beta})$$

Isto prova o lema.

2.4 A SOLUÇÃO DE (2.1.8)

Seja $(x_1,x_2) = (u_1,u_2) = (0,1)$, $(y_1,y_2,y_3) = (v_1,v_2,v_3) =$ $(0,1,0)$ em (2.1.8), então temos

$$2F(0,0) = F(1,1) \tag{2.4.1}$$

Seja $(x_1,x_2) = (u_1,u_2) = (0,1)$, $(y_1,y_2,y_3) = (y,1-y,0)$ e $(v_1,v_2,v_3) = (v,1-v,0)$ para $y,v \in [0,1]$ em (2.1.8), então

$$\begin{aligned} 4F(0,0) &+ F(y,v) + F(1-y,1-v) \\ &= \{F(y,v) + F(1-y,1-v) + F(0,0)\} + \{y^\gamma v^\delta + (1-y)^\gamma (1-v)^\delta\}\cdot \\ &\quad \{F(0,0) + F(1,1)\} \end{aligned} \tag{2.4.2}$$

Ou

$$3F(0,0) = \{y^\gamma v^\delta + (1-y)^\gamma (1-v)^\delta\}\{F(0,0) + F(1,1)\} \tag{2.4.3}$$

Combinando (2.4.3) e (2.4.1), temos

$$F(0,0) = F(1,1) = 0 \tag{2.4.4}$$

A fim de derivar a solução para a equação funcional (2.1.8), nós precisamos provar alguns lemas.

<u>Lema 2.4.1</u> Para $X = (x_1,\ldots,x_m)$, $U = (u_1,\ldots,u_m)$ em Δ_m, $m = 2,3$

$$\sum_{i=1}^{m} F(x_i,u_i) = C \sum_{i=1}^{m} (x_i^\alpha u_i^\beta - x_i^\gamma u_i^\delta) \tag{2.4.5}$$

onde $\alpha,\beta,\gamma,\delta$ são parametros satisfazendo $(\alpha-\gamma)(\delta-\beta) < 0$.

Prova:

Quando $n = m$ (2.1.8) ē o que se segue

$$\sum_{i=1}^{m} \sum_{j=1}^{m} F(x_i y_j, u_i v_j) = \sum_{i=1}^{m} \sum_{j=1}^{m} x_i^\alpha u_i^\beta F(y_j,v_j) + \sum_{i=1}^{m} \sum_{j=1}^{m} y_j^\gamma v_j^\delta F(x_i,u_i) \tag{2.4.6}$$

e

$$\sum_{i=1}^{m} \sum_{j=1}^{m} F(y_j x_i, v_j u_i) = \sum_{i=1}^{m} \sum_{j=1}^{m} y_j^\alpha v_j^\beta F(x_i,u_i) + \sum_{i=1}^{m} \sum_{j=1}^{m} x_i^\gamma u_i^\delta F(y_j,v_j) \tag{2.4.7}$$

Desde que

$$\sum_{i=1}^{m}\sum_{j=1}^{m} F(x_i y_j, u_i v_j) = \sum_{i=1}^{m}\sum_{j=1}^{m} F(y_j x_i, v_j u_i)$$

De (2.4.6) e (2.4.7), segue que

$$\sum_{i=1}^{m}\sum_{j=1}^{m} x_i^{\alpha}u_i^{\beta}F(y_j,v_j) + \sum_{i=1}^{m}\sum_{j=1}^{m} y_j^{\gamma}v_j^{\delta}F(x_i,u_i) \tag{2.4.8}$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{m} y_j^{\alpha}v_j^{\beta}F(x_i,u_i) + \sum_{i=1}^{m}\sum_{j=1}^{m} x_i^{\gamma}u_i^{\delta}F(y_j,v_j)$$

Ou equivalentemente

$$\frac{\sum_{j=1}^{m} F(y_j,v_j)}{\sum_{j=1}^{m} (y_j^{\alpha}v_j^{\beta} - y_j^{\gamma}v_j^{\delta})} = \frac{\sum_{i=1}^{m} F(x_i,u_i)}{\sum_{i=1}^{m} (x_i^{\alpha}u_i^{\beta} - x_i^{\gamma}u_i^{\delta})}, \quad \text{para } (\alpha-\gamma)(\delta-\beta) < 0 \tag{2.4.9}$$

para $X = (x_1,\ldots,x_m)$, $Y = (y_1,\ldots,y_m)$, $U = (u_1,\ldots,u_m)$ e $V = (v_1,\ldots,v_m)$ em Δ_m.

Deste que os parametros $\alpha,\beta,\gamma,\delta$ são positivos e satisfazem a relação $(\alpha-\gamma)(\delta-\beta) < 0$, o denominador de 2.3.9 não se anulará nunca para $x_i,u_i,y_i,v_i \in (0,1)$.

Seja $(x_1,x_2) = (y_1,y_2) = (1,0)$, (2.1.8) torna-se

$F(1,0) + F(0,1) = 0$

O qual junto com o fato que $F(0,0) = f(1,1) = 0$, temos (2.4.9) verdadeiro para todos X,Y,U,V em Δ_m.

Portanto (2.4.9) significa que

$$\sum_{i=1}^{m} F(x_i,u_i) = C \sum_{i=1}^{m} (x_i^{\alpha}u_i^{\beta} - x_i^{\gamma}u_i^{\delta}) \tag{2.4.10}$$

para todo $X = (x_1,\dots,x_m)$, $U = (u_1,\dots,u_m)$ em Δ_m, onde C é uma constante.

<u>Lema 2.4.2</u> Para x fixo, $x \in [0,1]$, se

$$(2.4.11)\quad A_{xu}(p,q) = F(xp,uq) + F((1-x)p,(1-u)q) - F(p,q)\{x^\alpha u^\beta + (1-x)^\alpha(1-u)^\beta\} - p^\gamma q^\delta\{F(x,u) + F(1-x,1-u)\}$$

para $p,q \in [0,1]$.

Então $A_{xu}(p,q)$ é aditiva em duas variáveis, ou equivalentemente

$$(2.4.12)\quad A_{xu}(y+v,w+t) = A_{xu}(y,w) + A_{xu}(v,t)$$

para $y,w,v,t,y+w,v+t \in [0,1]$.

Prova:

Substituindo $(x_1,x_2) = (x,1-x)$, $(u_1,u_2) = (u,1-u)$, $(y_1,y_2,y_3) = (y,v,1-y-v)$, $(v_1,v_2,v_3) = (w,t,1-w-t)$ em (2.1.8).

$$(2.4.13)\quad F(xy,uw) + F(xv,ut) + F(x(1-y-v),u(1-w-t)) + F((1-x)y,(1-u)w) + F((1-x)v,(1-u)t) + F((1-x)(1-y-v),(1-u)(1-w-t))$$
$$= \{x^\alpha u^\beta + (1-x)^\alpha(1-u)^\beta\}\{F(y,w) + F(v,t) + F(1-y-v,1-w-t)\} + \{y^\gamma w^\delta + v^\gamma t^\delta + (1-y-v)^\gamma(1-w-t)^\delta\}\{F(x,u) + F(1-x,1-u)\}$$

Novamente substituindo $(x_1,x_2) = (x,1-x)$, $(u_1,u_2) = (u,1-u)$ $(y_1,y_2,y_3) = (y+v,1-y-v,0)$, $(v_1,v_2,v_3) = (w+t,1-w-t,0)$ em (2.1.8) com o fato que $F(0,0) = 0$, temos

$$(2.4.14)\quad F(x(y+v),u(w+t)) + F(x(1-y-v),u(1-w-t)) + F((1-x)(y+v),(1-u)(w+t)) + F((1-x)(1-y-v),(1-v)(1-w-t))$$
$$= \{x^\alpha u^\beta + (1-x)^\alpha(1-u)^\beta\}\{F(y+v,w+t) + F(1-y-v,1-w-t)\} + \{(y+v)^\gamma(w+t)^\delta + (1-y-v)^\gamma(1-w-t)^\delta\}\{F(x,u) + F(1-x,1-u)\}$$

Agora, substraindo (2.4.13) de (2.4.14)

(2.4.15) $F(x(y+v),u(w+t)) + F((1-x)(y+v),(1-u)(w+t)) - F(xy,uw) -$

$F((1-x)y,(1-u)w) - F(xv,ut) - F((1-x)v,(1-u)t)$

$= \{x^{\alpha}u^{\beta}+(1-x)^{\alpha}(1-u)^{\beta}\}\{F(y+v,w+t) - F(y,w) - F(v,t)\} +$

$\{(y+v)^{\gamma}(w+t)^{\delta} - y^{\gamma}w^{\delta} - v^{\gamma}t^{\delta}\}\{F(x,u) + F(1-x,1-u)\}$

Seja

(2.4.16) $A_{xu}p,q) = F(xp,uq) + F((1-x)p,(1-u)q) - \{x^{\alpha}u^{\beta}+(1-x)^{\alpha}(1-u)^{\beta}\}F(p,q)$

$-p^{\gamma}q^{\delta}\{F(x,u) + F(1-x,1-u)\}$

Então (2.4.9) é simplificada para

$A_{xu}(y+v,w+t) = A_{xu}(y,w) + A_{xu}(v,t)$

onde $y,w,v,t,y+v,w+t \in [0,1]$.

<u>Lema 2.4.3</u> Para x,u fixos, $x,u \in [0,1]$, se $A_{xu}(p,q)$ é mensurável (ou continua em um ponto, ou limitada em um intervalo pequeno), em cada uma de suas variáveis e satisfaz (2.4.12), então

(2.4.17) $A_{xu}(p,q) = A_{xu}(1,0)\cdot p + A_{xu}(0,1)\cdot q$

Prova:

Fazendo $w = t = 0$ em (2.4.12), temos

(2.4.18) $A_{xu}(y+v,0) = A_{xu}(y,0) + A_{xu}(v,0)$

para $y,v,y+v \in [0,1]$.

Esta é uma função aditiva em uma variável, como no lema 2.3.4, vemos que

(2.4.19) $A_{xu}(y,0) = A_{xu}(1,0)\cdot y$

Similarmente, se fazemos $y = v = 0$ em (2.4.12), isto conduz a

(2.4.20) $A_{xu}0,t) = A_{xu}(0,1)\cdot t$

Portanto

(2.4.21) $A_{xu}(p,q) = A_{xu}(p,0) + A_{xu}(0,q)$

$=A_{xu}(1,0)\cdot p + A_{xu}(0,1)\cdot q,$ para $p,q \in [0,1]$

Lema 2.4.4 Para todo $p \in [0,1]$, temos

(2.4.22) (a) $F(p,0) = pF(1,0)$, $F(0,p) = pF(0,1)$

(b) $F(1,0) = F(0,1) = F(p,0) = F(0,p) = 0$

Prova:

A equação (2.4.17) para $q = 0$ produz

(2.4.23) $A_{xu}(p,0) = A_{xu}(1,0)\cdot p$

A qual usando a definição de $A_{xu}(p,q)$ em (2.4.11) da

(2.4.24) $F(xp,0) + F((1-x)p,0) - F(p,0)\{x^\alpha u^\beta + (1-x)^\alpha (1-u)^\beta\}$

$= p\{F(x,0) + F(1-x,0) - F(1,0)[x^\alpha u^\beta + (1-x)^\alpha (1-u)^\beta]\}$

Em (2.4.24), seja $x = 1$, junto com o fato $F(0,0) = 0$. Então

(2.4.25) $F(p,0) - F(p,0)u^\beta = p\{F(1,0) - F(1,0)u^\beta\}$, $u \in [0,1]$

ou

(2.4.26) $F(p,0) = pF(1,0)$, $p \in [0,1]$

Similarmente, a equação (2.4.17) para $p = 0$ produz

(2.4.27) $A_{xu}(0,q) = A_{xu}(0,1)\cdot q$

Da qual, usando a definição de $A_{xu}(p,q)$ em (2.4.11) quando $u = 1$, podemos obter

(2.4.28) $F(0,p) = pF(0,1)$, $p \in [0,1]$

Seja $(x_1,x_2) = (1,0)$, $(u_1,u_2) = (0,1)$ em (2.4.5) de Lema 2.4.1, temos

(2.4.29) $F(0,1) + F(1,0) = 0$

Desde que $F(1,0) = 0$, isto implica

$F(0,1) = 0$

e

$F(p,0) = F(0,p) = 0 \ , \qquad p \in [0,1]$

Teorema 2.4.1 Se $F: [0,1]\times[0,1] \to R$ (Reais) satisfaz a equação funcional (2.1.8) com $F(1,0) = 0$, e F é mensurável (ou continua em um ponto, ou limitada em um intervalo pequeno) em cada uma de suas variáveis. Então

$$F(p,q) = C\{p^{\alpha}q^{\beta} - p^{\gamma}q^{\delta}\} \tag{2.4.30}$$

onde $\alpha,\beta,\gamma,\delta$ são parâmetros positivos satisfazendo a relaão $(\alpha-\gamma)(\delta-\beta) < 0$ e C é uma constante.

Prova:

Do Lema 2.4.4, sabemos que

$F(1,0) = (0,1) = F(p,0) = F(0,p) = 0, \qquad p \in [0,1]$

Portanto, pela definição de $A_{xu}(p,q)$ em (2.4.16)

$$A_{xu}(1,0) = F(x,0) + F(1-x,0) - \{x^{\alpha}u^{\beta} + (1-x)^{\alpha}(1-u)^{\beta}\}F(1,0) = 0 \tag{2.4.31}$$

e

$$A_{xu}(0,1) = F(0,u) + F(0,1-u) - \{x^{\alpha}u^{\beta} + (1-x)^{\gamma}(1-u)^{\beta}\}F(0,1) = 0 \tag{2.4.32}$$

A função $A_{xu}(p,q)$ como definida em (2.4.11) é aditiva em duas variáveis, portanto, pelo Lema 2.4.3, ela pode ser representada da seguinte maneira

$A_{xu}(p,q) = A_{xu}(1,0)\cdot p + A_{xu}(0,1)\cdot q$

De (2.4.31) e (2.4.32), concluimos que

$$A_{xu}p,q) = 0 \tag{2.4.33}$$

A qual quando combinando com (2.4.11), produz

$$F(xp,uq) + F((1-x)p,(1-u)q) \tag{2.4.34}$$

$$= F(p,q)\{x^{\alpha}u^{\beta}+(1-x)^{\alpha}(1-u)^{\beta}\} + p^{\gamma}q^{\delta}\{F(x,u) + F(1-x,1-u)\}$$

Usando o Lema 2.4.1

(2.3.35) O lado esquerdo de (2.3.34)

$$= F(p,q)\{x^{\alpha}u^{\beta}+(1-x)^{\alpha}(1-u)^{\beta}\} + Cp^{\gamma}q^{\delta}\{x^{\alpha}u^{\beta}-x^{\gamma}u^{\delta}+(1-x)^{\alpha}(1-u)^{\delta}\}$$

Mas sabemos do Lema 2.4.1 que

(2.4.36) $F(1-p,1-q) + F(xp,uq) + F((1-x)p,(1-u)q)$

$$= C\{ (1-p)^{\alpha}(1-q)^{\beta}-(1-p)^{\gamma}(1-q)^{\delta}+(xp)^{\alpha}(uq)^{\beta}- (xp)^{\gamma}(uq)^{\delta}+$$
$$[(1-x)p]^{\alpha}[(1-u)q]^{\beta}-[(1-x)p]^{\gamma}[(1-u)q]^{\delta}\}$$

Agora, substraindo (2.4.35) de (2.4.36)

(2.4.37) $F(1-p,1-q) = C\{ (1-p)^{\alpha}(1-q)^{\beta}-(1-p)^{\gamma}(1-q)^{\delta}+(xp)^{\alpha}(uq)^{\beta}+$
$$[(1-x)p]^{\alpha}[(1-u)q]^{\beta}-x^{\alpha}u^{\beta}p^{\gamma}q^{\delta}-(1-x)^{\alpha}(1-u)^{\beta}p^{\gamma}q^{\delta}\} -$$
$$F(p,q)\{x^{\alpha}u^{\beta}+(1-x)^{\alpha}(1-u)^{\beta}\}$$

Substituindo $x = u = 1/2$ em (2.4.37), temos

(2.4.38) $F(1-p,1-q) = 2^{1-\alpha-\beta}\{Cp^{\alpha}q^{\beta}-Cp^{\gamma}q^{\delta}-F(p,q)\} + C\{ (1-p)^{\alpha}(1-q)^{\beta}-$
$$(1-p)^{\gamma}(1-q)^{\delta}\}$$

Mas sabemos do,Lema 2.4.1 que

(2.4.39) $F(p,q) + F(1-p,1-q)$

$$= C\{p^{\alpha}q^{\beta}-p^{\gamma}q^{\delta}+(1-p)^{\alpha}(1-q)^{\beta}-(1-p)^{\gamma}(1-q)^{\delta}\}$$

Substraindo (2.4.38) de (2.4.39)

(2.4.40 $F(p,q) = C\{1-2^{1-\alpha-\beta}\}\{p^{\alpha}q^{\beta}-p^{\gamma}q^{\delta}\} + 2^{1-\alpha-\beta}F(p,q)$

ou

(2.4.41) $F(p,q) = C\{p^{\alpha}q^{\beta}-p^{\gamma}q^{\delta}\}$

para $\alpha + \beta \neq 1$, $p,q \in [0,1]$.

Quando $\alpha + \beta = 1$, Lema 2.4.1 fornece

(2.4.42) $$\sum_{i=1}^{n} F(x_i,u_i) = C \sum_{i=1}^{n} (x_i^{\alpha}u_i^{1-\alpha}-x_i^{\gamma}u_i^{\delta}) \quad , \qquad n = 2,3$$

Para $x_i=u_i$, $i=1,\ldots,m$; $y_j= v_j$, $j=1,\ldots,n$ e $\alpha+\beta=1$, a equação

(2.1.8) reduz-se a

(2.4.43) $$\sum_{i=1}^{m} \sum_{j=1}^{n} F(x_iy_j,x_iy_j) = \sum_{i=1}^{m} \sum_{j=1}^{n} x_iF(y_j,y_j) + \sum_{i=1}^{m} \sum_{j=1}^{n} y_j^{\gamma+\delta}F(x_i,x_i)$$

Esta é uma função de uma variável, portanto usando o Teorema2.3.1

temos

(2.4.44) $$F(x,x) = C\{x-x^{\gamma+\delta}\} \quad , \qquad \gamma + \delta \neq 1$$

Sem perda de generalidade, para qualquer p,q em $[0.1]$,podemos

supor $p < q$, então (2.4.43) produz

(2.4.45) $$F(p,q) + F(1-q,1-q) + F(q-p,0)$$

$$= C\{p^{\alpha}q^{1-\alpha}-p^{\gamma}q^{\delta}+(1-q)-(1-q)^{\gamma+\delta}\}$$

Desde que sabemos de (2.4.44) que

(2.4.46) $$F(1-p,1-p) = C\{(1-p)-(1-p)^{\gamma+\delta}\}$$

e também que

(2.4.47) $$F(q-p,0) = 0$$

Substraindo (2.4.46) e (2.4.47) de (2.4.45), temos

(2.4.48) $$F(p,q) = C\{p^{\alpha}q^{1-\alpha}-p^{\gamma}q^{\delta}\} \quad , \qquad \gamma + \delta \neq 1, \quad p,q \in [0,1]$$

Assim (2.4.41) e (2.4.48) provam o Teorema 2.4.1

Como uma aplicação do Teorema 2.4.1, nós temos o seguinte Lema.

<u>Lema 2.4.5</u> A função mensurável $F: [0,1]\times[0,1] \to R$ (reais) satisfazendo a equação funcional (2.1.8) e com a condição de normalização $F(1,1/2) = 1$, então $F(x,y)$ é dado por

(2.4.49) $F(x,y) = \dfrac{1}{2^{-\beta}-2^{-\delta}} (x^{\alpha}y^{\beta}-x^{\gamma}y^{\delta})$

a qual é uma medida de informação teórica de um evento com probabilidade x ou y sob a distribuição de probabilidade X ou Y respectivamente.

Prova:

Aplicando o Teorema 2.4.1, a solução mensurável da equação funcional (2.1.8) é dado por

(2.4.50) $F(p,q) = C\{p^{\alpha}q^{\beta}-p^{\gamma}q^{\delta}\}$

Então usando a condição de normalização $F(1,1/2) = 1$, é fácil ver que

(2.4.51) $C = \dfrac{1}{2^{-\beta}- 2^{-\delta}}$

Isto prova o lema.

CAPÍTULO 3

EQUAÇÕES FUNCIONAIS E MEDIDAS DE INFORMAÇÃO NÃO-ADITIVAS

RESUMO

Neste capítulo, as soluções mensuráveis da equação funcional em uma e duas variáveis são obtidas. Estas equações funcionais surgem naturalmente de algumas propriedades de não-aditividade de medidas de informação. As soluções mensuráveis, em casos especiais, são utilizadas para a caracterização axiomatica de medidas de informação não-aditivas.

3.1 INTRODUÇÃO

Vamos denotar

$$\Delta_n = \{ (p_1,\dots,p_n): p_i \geqq 0,\ i=1,\dots,n,\ \sum_{i=1}^{n} p_i = 1 \}, \qquad n \geqq 1$$

$$\Delta'_n = \{ (p_1,\dots,p_n): p_i \geqq 0,\ i=1,\dots,n,\ \sum_{i=1}^{n} p_i \leqq 1 \}, \qquad n \geqq 1$$

(i) Equação Funcional em Uma Variável

Considerando funções mensuráveis $f, g, h : [0,1] \to R$ (Reais) satisfazendo a equação funcional

$$\sum_{i=1}^{n} \sum_{j=1}^{m} f(x_i y_j) = \sum_{i=1}^{n} g(x_i) + \sum_{j=1}^{m} h(y_j) + \lambda \sum_{i=1}^{n} \sum_{j=1}^{m} g(x_i) h(y_j) \tag{3.1.1}$$

onde $X = (x_1,\dots,x_n) \in \Delta_n$, $Y = (y_1,\dots,y_m) \in \Delta_m$ para $n=2,3$; $m=2,3$ e λ é uma constante não nula.

As soluçoes do caso especial da equação funcional (3.1.1) conduz a uma caracterização da entropia não-aditiva de ordem β. A derivação da equação funcional (3.1.1) quando $f = g = h$ atrvés da não-aditividade da entropia de ordem β é discutida abaixo.

A entropia não-aditiva de ordem β [Havrda e Charvát (1967)] é definida como

$$(3.1.2)\qquad H_{n,\beta}(P) = \frac{\sum_{i=1}^{n} p_i^{\beta} - 1}{2^{1-\beta}-1}$$

para todo $P = (p_1,\ldots,p_n) \in \Delta_n$, $n \geq 1$, $\beta \neq 1$, $\beta > 0$.

Aqui usando a convenção $0^{\beta}= 0$.

Seja

$$(3.1.3)\qquad f_{\beta}(x) = \frac{x^{\beta}- 1}{2^{1-\beta}-1}, \qquad x \in [0,1]$$

a correspondente entropia do evento o qual é de probabilidade x, Então podemos escrever $H_{n,\beta}(P)$ como uma soma de $f_{\beta}(p_i)$ para todo $i=1,\ldots,n$.

$$(3.1.4)\qquad H_{n,\beta}(P) = \sum_{i=1}^{n} f_{\beta}(p_i)$$

para todo $P = (p_1,\ldots,p_n) \in \Delta_n$, $\beta \neq 1$, $\beta > 0$.

A propriedade não-aditividade de $H_{n,\beta}$ é dada abaixo:

Para $P = (p_1,\ldots,p_n) \in \Delta_n$, $Q = (q_1,\ldots,q_m) \in \Delta_m$

$$(3.1.5)\qquad H_{mn,\beta}(P*Q) = H_{n,\beta}(P) + H_{m,\beta}(Q) + (2^{-\beta}-1)H_{n,\beta}(P)H_{m,\beta}(Q)$$

Ou, se usando a representação de soma para $H_{n,\beta}$, então

$$(3.1.6)\qquad \sum_{i=1}^{n}\sum_{j=1}^{m} f_{\beta}(x_i y_j) = \sum_{i=1}^{n} f_{\beta}(x_i) + \sum_{j=1}^{m} f_{\beta}(y_j) +$$

$$(2^{1-\beta}-1)\{\sum_{i=1}^{n} f_\beta(x_i)\}\{\sum_{j=1}^{m} f_\beta(y_j)\}$$

Este é um caso especial da equação funcional (3.1.1), quando $f = g = h$.

(ii) Equação Funcional em Duas Variáveis

Um generalização de (3.1.1) para duas variáveis segue-se.

Para $f, g, h : [0,1] \times [0,1] \to R$ (Reais) mensurável em cada uma de suas variáveis satisfazendo a equação funcional

$$(3.1.7) \quad \sum_{i=1}^{n} \sum_{j=1}^{m} f(x_i y_j, u_i v_j)$$

$$= \sum_{i=1}^{n} g(x_i, u_i) + \sum_{j=1}^{m} h(y_j, v_j) + \lambda \sum_{i=1}^{n} \sum_{j=1}^{m} g(x_i, u_i) h(y_j, v_j)$$

onde $X = (x_1, \ldots, x_n) \in \Delta_n$, $U = (u_1, \ldots, u_n) \in \Delta'_n$, $Y = (y_1, \ldots, y_m) \in \Delta_m$, $V = (v_1, \ldots, v_m) \in \Delta'_m$ para $n = 2,3$; $m = 2,3$ e λ é uma constante não nula.

As soluções do caso especial da equação funcional (3.1.7) conduz à caracterização de várias medidas de informação generalizada não-aditiva da equação funcional (3.1.7) quando $f = g = h$ através da não-aditividade da medida de informação generalizada é discutida abaixo.

A Divergência dirigida não-aditiva de ordem (α,β) é dada na seguinte forma:

$$(3.1.8) \quad I_{n,\alpha,\beta}(P:Q) = \frac{\sum_{i=1}^{n} p_i^\alpha q_i^\beta - 1}{2^{-\beta} - 1}, \qquad \beta \neq 0, \qquad \alpha, \beta > 0.$$

para todo $P = (p_1, \ldots, p_n) \in \Delta_n$, $Q = (q_1, \ldots, q_n) \in \Delta'_n$.

Similarmente,

$$(3.1.9) \quad f_{\alpha,\beta}(p,q) = \frac{p^{\alpha}q^{\beta} - p}{2^{-\beta} - 1}, \qquad \beta \neq 0 \ , \ \alpha,\beta > 0$$

denota a medida de informação contida no evento o qual assume probabilidade p ou q acordo com a distribuição de probabilidade P ou Q respectivamente associadas do espaço de probabilidade.

Então, é claro que $I_{n,\alpha,\beta}(P:Q)$ é uma soma de $f_{\alpha,\beta}(p_i,q_i)$ para $i = 1,\ldots,n$, enquando $P = (p_1,\ldots,p_n) \in \Delta_n$, $Q = (q_1,\ldots,q_n) \in \Delta'_n$.

$$(3.1.10) \quad I_{n,\alpha,\beta}(P:Q) = \sum_{i=1}^{n} f_{\alpha,\beta}(p_i,q_i)$$

De (3.1.9), podemos ver que $f_{\alpha,\beta}(1,1/2) = 1$

Se sob a condição $f_{\alpha,\beta}(1/2,1/2) = 0$, (3.1.9) dará a relação $\alpha + \beta = 1$, e nesse caso

$$(3.1.11) \quad I_{n,\alpha,1-\alpha}(P:Q) = \frac{\sum_{i=1}^{n} p_i^{\alpha}q_i^{1-\alpha} - 1}{2^{\alpha-1} - 1}, \qquad \alpha \neq 1 \ , \ \alpha > 0$$

a qual é uma divergência dirigida de ordem α $(\alpha \neq 1)$

Quando $f(1/2,1/2) = 1/2$, (3.1.9) dará $\alpha = 1$, e portanto

$$(3.1.12) \quad I_{n,1,\beta}(P:Q) = \frac{\sum_{i=1}^{n} p_i q_i^{\beta} - 1}{2^{-\beta} - 1}, \qquad \beta \neq 0 \ , \ \beta > 0$$

a qual é uma imprecisão de ordem $1+\beta$.

A propriedade não-aditividade de $I_{n,\alpha,\beta}$ é dada abaixo:

Para $P = (p_1,\ldots,p_n) \in \Delta_n$, $Q = (q_1,\ldots,q_m) \in \Delta_m$, $R = (r_1,\ldots,r_n) \in \Delta'_n$, $S = (s_1,\ldots,s_m) \in \Delta'_m$.

$$(3.1.13) \quad I_{mn,\alpha,\beta}(P*Q : R*S) = I_{n,\alpha,\beta}(P:R) + I_{m,\alpha,\beta}(Q:S) + (2^{-\beta}-1)\, I_{n,\alpha,\beta}(P:Q)\, I_{m,\alpha,\beta}(R:S)$$

Ou, se usando a representação de soma para $I_{n,\alpha,\beta}$, então

$$(3.1.14) \quad \sum_{i=1}^{n} \sum_{j=1}^{m} f_{\alpha,\beta}(p_i q_j, r_i s_j)$$

$$= \sum_{i=1}^{n} f_{\alpha,\beta}(p_i, r_i) + \sum_{j=1}^{m} f_{\alpha,\beta}(q_j, s_j) +$$

$$(2^{-\beta}-1)\{\sum_{i=1}^{n} f_{\alpha,\beta}(p_i, r_i)\}\{\sum_{j=1}^{m} f_{\alpha,\beta}(q_j, s_j)\}$$

Este é uma caso especial da equação funcional (3.1.7), quando $f = g = h$.

A solução continua de (3.1.1) quando $f = g = h$ foi dada anteriormente por Behara e Nath (1973) e Sharma e Taneja (1975). Mittal (1976) fornece a solução continua de (3.1.1). A solução continua de (3.1.7) quando $f = g = h$ foi dada por Sharma e Taneja (1975).

O objectivo deste trabalho é encontra a solução mensurável das equações funcionais (3.1.1) e (3.1.7). Uma caracterização de entropia não-aditiva é indicada como uma aplicação de (3.1.1), enquando outra medida de informação generalizada não-aditiva de ordem (α,β) é caracterizada como uma aplicação de (3.1.7).

3.2 PROPIEDADES

I. Para A Entropia Não-Aditiva de Ordem β

$$H_{n,\beta}(P) = \frac{\sum_{i=1}^{n} p_i^{\beta} - 1}{2^{1-\beta} - 1}, \qquad \beta > 0,\ \beta \neq 1$$

(i) Não-Negatividade

$H_{n,\beta}(P)$ é não-negativa, e é zero se e sómente se $p_i = 1$ para algun $i = 1,\ldots,n$, e os restos são zeros.

A entropia é geralmente positiva para a distribuição de probabilidade e será zero se e sómente se um dos eventos acontece com probabilidade 1.

(ii) Simetria

$$H_{n,\beta}(p_1,\ldots,p_n) = H_{n,\beta}(p_{a(1)},\ldots,p_{a(n)})$$

onde $\{a(1),\ldots,a(n)\}$ é uma permutaçaõ arbitrária de $\{1,\ldots,n\}$.

A entropia não depende da ordem em que os eventos são fixados.

(iii) Expandibilidade

$$H_{n,\beta}(p_1,\ldots,p_n) = H_{n+1,\beta}(p_1,\ldots,p_n,0)$$

A adição de um evento de probabilidade zero não altera o valor da entropia.

(iv) Recursividade

$$H_{n,\beta}(p_1,\ldots,p_n)$$

$$= H_{n-1,\beta}(p_1+p_2,p_3,\ldots,p_n) + (p_1+p_2)^{\beta} H_{2,\beta}\left(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2}\right)$$

para $p_1+p_2 > 0$.

(v) Não-Aditividade

Para $P \in \Delta_n$, $Q \in \Delta_m$

$$H_{mn,\beta}(P*Q)$$

$$= H_{n,\beta}(P) + H_{m,\beta}(Q) + (2^{1-\beta} - 1) H_{n,\beta}(P) H_{m,\beta}(Q)$$

(vi) Não-Aditividade Forte

Para $P = (p_1,\ldots,p_m) \in \Delta_m$; $Q_j = (q_{j1},\ldots,q_{jn}) \in \Delta_n, j = 1,2,\ldots,m$

$$H_{mn,\beta}(p_1q_{11},\ldots,p_1q_{1n},\ldots,p_mq_{m1},\ldots,p_mq_{mn})$$

$$= H_{m,\beta}(p_1,\ldots,p_m) + \sum_{j=1}^{m} p_j^{\beta} H_{n,\beta}(q_{j1},\ldots,q_{jn})$$

(vii) Continuidade

$H_{n,\beta}(P)$ é uma função contínua de n variáveis para $P \in \Delta_n$.

(viii) Normalização

$$H_{2,\beta}(1/2,1/2) = 1$$

II. Para A Divergência Dirigida Não Aditiva de Ordem (α,β)

$$I_{n,\alpha,\beta}(P:Q) = \frac{\sum_{i=1}^{n} p_i^{\alpha} q_i^{\beta} - 1}{2^{-\beta} - 1} \quad , \qquad \alpha,\beta > 0$$

(i) Simetria

$$I_{n,\alpha,\beta}(p_1,\ldots,p_n:q_1,\ldots,q_n)$$

$$= I_{n,\alpha,\beta}(p_{a(1)},\ldots,p_{a(n)}: q_{a(1)},\ldots,q_{a(n)})$$

onde $\{a(1),\ldots,a(n)\}$ é uma permutação arbitrária de $\{1,\ldots,n\}$.

A permutação conjuta de indices correspondentes de eventos não muda a valor da medida de informação $I_{n,\alpha,\beta}$.

(ii) Expandibilidade

$$I_{n,\alpha,\beta}(p_1,\dots,p_n:q_1,\dots,q_n) = I_{n+1,\alpha,\beta}(p_1,\dots,p_n,0:q_1,\dots,q_n,0)$$

A adição de um evento de probabilidade zero não altera o valor da $I_{n,\alpha,\beta}$.

(iii) Recursividade

$$I_{n,\alpha,\beta}(p_1,\dots,p_n:q_1,\dots,q_n)$$

$$= I_{n-1,\alpha,\beta}(p_1+p_2,p_3,\dots,p_n:q_1+q_2,q_3,\dots,q_n) +$$

$$(p_1+p_2)^{\alpha}(q_1+q_2)^{\beta} I_{2,\alpha,\beta}\left(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2} : \frac{q_1}{q_1+q_2}, \frac{q_2}{q_1+q_2}\right)$$

para $p_1+p_2 > 0$, $q_1+q_2 > 0$.

(iv) Não-Aditividade

Para P, $Q \in \Delta_n$, R, $S \in \Delta_m$

$$I_{mn,\alpha,\beta}(P*R : Q*S)$$

$$= I_{n,\alpha,\beta}(P:Q) + I_{m,\alpha,\beta}(R:S) + (2^{-\beta}-1) I_{n,\alpha,\beta}(P:Q) I_{m,\alpha,\beta}(R:S)$$

(v) Não-Aditividade Forte

Para $(p_1,\dots,p_n)$, $(q_1,\dots,q_n) \in \Delta_n$; $(P_{1i},\dots,P_{mi})$, $(Q_{1i},\dots,Q_{mi})$ $\in \Delta_m$ para $i = 1,2,\dots,n$.

$$I_{mn,\alpha,\beta}(p_1P_{11},\dots,p_1P_{m1},\dots,p_nP_{1n},\dots,p_nP_{mn}: q_1Q_{11},\dots,q_1Q_{m1},\dots,$$
$$q_nQ_{1n},\dots,q_nQ_{mn})$$

$$= I_{n,\alpha,\beta}(p_1,\dots,p_n:q_1,\dots,q_n) +$$

$$\sum_{i=1}^{m} p_i^{\alpha} q_i^{\beta} I_{m,\alpha,\beta}(P_{1i},\dots,P_{mi}: Q_{1i},\dots,Q_{mi})$$

Isto é claro da seguinte relação

$$I_{n,\alpha,\beta}\,(p_1,\ldots,p_n:q_1,\ldots,q_n)$$

$$= I_{n-k,\alpha,\beta}\,(p_1+\ldots+p_k,p_{k+1},\ldots,p_n:q_1+\ldots+q_k,q_{k+1},\ldots,q_n) +$$

$$(2^{-\beta}-1)\,I_{k,\alpha,\beta}\,\left(\frac{p_1}{r_k},\ldots,\frac{p_k}{r_k}:\frac{q_1}{s_k},\ldots,\frac{q_k}{s_k}\right)$$

onde $r_k = p_1+\ldots+p_k$, $s_k = q_1+\ldots+q_k$ para $1 \leqq k \leqq n$.

(vi) Continuidade

$I_{n,\alpha,\beta}$ (P:Q) é uma função de 2n variáveis para P, Q ε Δ_n.

(vii) Normalização

$$I_{2,\alpha,\beta}(1,1/2) = 1$$

3.3 EQUAÇÃO FUNCIONAL (3.1.1) EM UMA VARIÁVEL

Nesta secção, daremos a solução mensurável da equação funcional (3.1.1), e aplicamos o caso especial de (3.1.1) para $f = g = h$ para derivar uma caracterização para a entropia não-aditiva de ordem β.

A fim de provar o teorema nesta secção, nos necessitamos desenvolver vários lemas que segue-se, o lema 3.3.1 pode ser encontrado no livro de Aczél (1966), e então a prova será omitida.

<u>Lema 3.3.1</u> A função mensurável f satisfazendo a equação de Cauchy

(3.3.1) $\quad f(x+y) = f(x) \cdot f(y)$

para totos os números reais positivos x,y, tem a solução não-trivial

(3.3.2) $\quad f(x) = e^{cx}$

onde c é uma constante arbitrária.

<u>Lema 3.3.2</u> A função mensurável k katisfazendo a equação de Cauchy

(3.3.3) $\quad k(xy) = k(x) \cdot k(y), \qquad x,y \in (0,1]$

tem a solução não-trivial

(3.3.4) $\quad k(x) = x^{\alpha}, \qquad x \in (0,1]$

onde α é uma constante arbitrária.

Prova:

Seja

(3.3.5) $\quad g(u) = k(e^{u})$

Para $x,y \in (0,1]$, podemos escrever

(3.3.6) $\quad x = e^{u},\ y = e^{v}$ para algun $u,v \in (-\infty,0]$

Portanto, (3.3.3) torna-se

(3.3.7) $g(u+v) = g(u) \cdot g(v)$, $\qquad u,v \in (-\infty,0]$

Novamente, seja

(3.3.8) $h(t) = g(-t)$

Então, (3.3.7) pode ser escrevida como

(3.3.9) $h(a+b) = h(a) \cdot h(b)$, $\qquad a,b \in [0,\infty)$

Pelo Lema 3.3.1, h tem seguinte solução não-trivial

(3.3.10) $h(x) = e^{cx}$, $\qquad x \in [0,\infty)$

Da transformação (3.3.5) e (3.3.8), então, k tem a solução não-trivial como em (3.3.4).

Lema 3.3.3 A solução geral, com k,m,n mensuráveis, da equação de Pexider

(3.3.11) $k(xy) = m(x) \cdot n(y)$, $\qquad x,y \in (0,1]$

é

Caso I. Se $m(1) \neq 0$ e $n(1) \neq 0$

(3.3.12) $k(t) = abt^{\alpha}$, $m(t) = at^{\alpha}$, $n(t) = bt^{\alpha}$

onde α é um parametro.

suplementada com as soluções triviais

Caso II. Se $m(1) = 0$ ou $n(1) = 0$

(3.3.13) $\begin{cases} k(t) = 0 \\ m(t) = 0 \\ n(t): \text{arbitrária} \end{cases}$ ou $\begin{cases} k(t) = 0 \\ m(t): \text{arbitrária} \\ n(t) = 0 \end{cases}$

Prova:

Seja $y = 1$ em (3.3.11), nós temos

(3.3.14) $k(x) = b\,m(x)$

onde $b = n(1)$.

Seja $x = 1$ em (3.3.11), então

(3.3.15) $k(y) = a \cdot n(1)$

onde $a = m(1)$

Caso I. Quando $m(1) \neq 0$ e $n(1) \neq 0$

Substituindo (3.3.14) em (3.3.11), temos

(3.3.16) $abk(xy) = k(x) \cdot k(y), \qquad x,y \in (0,1]$

Usando a transformação

(3.3.17) $\phi(t) = \frac{k(t)}{ab}$

Então (3.3.16) torna-se

(3.3.18) $\phi(xy) = \phi(x) \cdot \phi(y), \qquad x,y \in (0,1]$

Pelo Lema 3.3.2, nós temos a solução não-trivial como

(3.3.19) $\phi(x) = x^{\alpha}$

onde α é uma constante arbitrária.

Correspondentemente, nós temos a solução não-trivial para k,m e n como em (3.3.12).

Caso II. Quando $m(1) = 0$ ou $n(1) = 0$

De (3.3.14) e (3.3.15), é fácil ver que

(3.3.20) $k(x) = 0, \qquad x \in (0,1]$

e correspondentemente, nós temos a solução trivial como (3.3.13).

Observação: No livre de Aczél (1966), a equação de Pexider é resolvida na região tal que x,y são números reais positivos, mas aqui nós nos restringiremos ao caso em que x e y variam no intervalo $(0,1]$.

<u>Teorema 3.3.1</u> Se f,g e h são soluções mensuráveis da equação funcional (3.1.1) para $X = (x_1,\ldots,x_n) \in \Delta_n$, $Y = (y_1,\ldots,y_m) \in \Delta_m$, onde $n = 2,3; m = 2,3$ e λ é uma constante não nula, então elas são dadas por

$$(3.3.21)\quad \begin{cases} f(x) = \frac{x}{\lambda}\,(abx^{\beta-1} - 1) \\ g(x) = \frac{x}{\lambda}\,(ax^{\beta-1} - 1) \\ h(x) = \frac{x}{\lambda}\,(bx^{\beta-1} - 1) \end{cases}$$

para todo $x \in [0,1]$, onde $ab \neq 0$ e β ẽ um parametro .

ou

$$(3.3.22)\quad f(x) = g(x) = -\frac{x}{\lambda}\,,\ h(x)\text{: arbitrãria}$$

ou

$$(3.3.23)\quad f(x) = h(x) = -\frac{x}{\lambda}\,,\ g(x)\text{: arbitrãria}$$

Prova:

Substituindo $Y = (y,u,1-y-u) \in \Delta_3$ em (3.1.1), temos

$$(3.3.24)\quad \sum_{i=1}^{n} \{f(x_i y) + f(x_i u) + f(x_i(1-y-u))\}$$

$$= \sum_{i=1}^{n} g(x_i) + \{h(y) + h(u) + h(1-y-u)\} +$$

$$\lambda\cdot\{\sum_{i=1}^{n} g(x_i)\}\{h(y) + h(u) + h(1-y-u)\}$$

Novamente, substituindo $Y = (y+u,1-y-u) \in \Delta_2$ em (3.1.1), temos

$$(3.3.25)\quad \sum_{i=1}^{n} \{f(x_i(y+u)) + f(x_i(1-y-u))\}$$

$$= \{\sum_{i=1}^{n} g(x_i)\} + \{\ h(y+u) + h(1-y-u)\ \} +$$

$$\lambda\cdot\{\sum_{i=1}^{n} g(x_i)\}\{\ h(y+u) + h(1-y-u)\}$$

Substraindo (3.3.24) de (3.3.25), então

(3.3.26) $$\sum_{i=1}^{n} f(x_i(y+u)) - h(y+u) - \lambda\cdot\{\sum_{i=1}^{n} g(x_i)\}h(y+u)$$

$$= \sum_{i=1}^{n} f(x_i y) - h(y) - \lambda\cdot\{\sum_{i=1}^{n} g(x_i)\}\cdot h(y) + \sum_{i=1}^{n} f(x_i u) - h(u) -$$

$$\lambda\cdot\{\sum_{i=1}^{n} g(x_i)\}\cdot h(u)$$

Para $X = (x_1,\ldots,x_n) \in \Delta_n$, defina

(3.3.27) $$A_X(t) = \sum_{i=1}^{n} f(x_i t) - h(t) - \lambda\cdot\{\sum_{i=1}^{n} g(x_i)\}\cdot h(t)$$

para $t \in [0,1]$.

Então, é fácil ver de (3.3.26) que $A_X(\cdot)$ é aditiva em J, onde

(3.3.28) $$J = \{(x,y): x,y,x+y \in [0,1]\}$$

Ou equivalentemente, dizamos

(3.3.29) $$A_X(y+u) = A_X(y) + A_X(u) \qquad \text{em } J$$

Concluimos de Lema 2.3.4, o qual é um resultado de Daróczy e Losonczi (1967) que a solução mensurável de (3.3.29) é

(3.3.30) $$A_X(t) = t\cdot A_X(1), \qquad t \in [0,1]$$

Para ver a expressão de $A_X(1)$, nós precisamos encontrar a expressão para

$$\sum_{i=1}^{n} f(x_i) - h(1) - \lambda\cdot\{\sum_{i=1}^{n} g(x_i)\}\cdot h(1)$$

Portanto, substituindo $Y = (1,0)$ em (3.1.1), temos

(3.3.31) $$\sum_{i=1}^{n} f(x_i) + n\cdot f(0)$$

$$= \sum_{i=1}^{n} g(x_i) + h(1) + h(0) + \lambda\cdot\{\sum_{i=1}^{n} g(x_i)\}\{h(1) + h(0)\}$$

Novamente, substituindo Y = (1,0,0) em (3.1.1), temos

$$(3.3.32)\quad \sum_{i=1}^{n} f(x_i) + 2n\cdot f(0)$$

$$= \sum_{i=1}^{n} g(x_i) + h(1) + 2\, h(0) + \lambda\cdot\{\sum_{i=1}^{n} g(x_i)\}\{h(1) + 2\, h(0)\}$$

Substraindo (3.3.31) de (3.3.32), então

$$(3.3.33)\quad n\, f(0) = h(0) + \lambda\, h(0) + \sum_{i=1}^{n} g(x_i)$$

Conbinando (3.3.33) com (3.3.31), temos

$$(3.3.34)\quad \sum_{i=1}^{n} f(x_i) - h(1) - \lambda\cdot h(1)\cdot\{\sum_{i=1}^{n} g(x_i)\} = \sum_{i=1}^{n} g(x_i)$$

Assim,

$$(3.3.35)\quad A_X(1) = \sum_{i=1}^{n} g(x_i)$$

De (3.3.30) e (3.3.35), e da definição de $A_X(t)$ em (3.3.27),

temos a seguinte relação

$$(3.3.36)\quad \sum_{i=1}^{n} f(x_i t) - h(t) - \lambda\cdot h(t) \sum_{i=1}^{n} g(x_i) = t\cdot \sum_{i=1}^{n} g(x_i)$$

para todo $X = (x_1,\dots,x_n) \in \Delta_n$, $n = 2,3$ e $t \in [0,1]$.

Seja X = (x,v,1-x-v) em (3.3.36), temos

$$(3.3.37)\quad f(xt) + f(vt) + f((1-x-v)t) - h(t) -$$
$$\lambda\cdot h(t)\cdot\{g(x) + g(v) + g(1-x-v)\}$$
$$= t\cdot\{g(x) + g(v) + g(1-x-v)\}$$

Novamente, seja X = (x+v, 1-x-v) em (3.3.36), temos

$$(3.3.38)\quad f((x+v)t) + f((1-x-v)t) - h(t) - \lambda\cdot h(t)\{g(x+v) + g(1-x-v)\}$$
$$= t\cdot\{g(x+v) + g(1-x-v)\}$$

Substraindo (3.3.37) de (3.3.38), conseguimos

$$(3.3.39)\quad f((x+v)t) - \lambda\cdot h(t)g(x+v) - t\cdot g(x+v)$$
$$= f(xt) - \lambda\cdot h(t)g(x) - t\cdot g(x) + f(vt) - \lambda\cdot h(t)g(v) - t\cdot g(v)$$

Para $t \in [0,1]$, definimos

$$(3.3.40)\quad B_t(w) = f(wt) - \lambda\cdot h(t)g(w) - t\cdot g(w), \qquad w \in [0,1]$$

Então, (3.3.39) pode ser reescrita como

$$(3.3.41)\quad B_t(x+v) = B_t(x) + B_t(v)$$

para $x,v,x+v \in [0,1]$.

Aplicando o resultado de Daróczy e Losonczi (1967), (ou Lema 2.3.4). Concluimos que

$$(3.3.42)\quad B_t(t) = x\cdot B_t(1), \qquad x \in [0,1]$$

Para calcular para $B_t(1)$, nós precisamos calcular para

$$(3.3.43)\quad f(t) - \lambda\cdot h(t)g(1) - t\cdot g(1), \qquad t \in [0,1]$$

Fazendo $X = (1,0)$ em (3.3.36), temos

$$(3.3.44)\quad f(t) + f(0) - h(t) - \lambda\cdot h(t)\{g(1) + g(0)\}$$
$$= T\cdot\{g(1) + g(0)\}$$

Novamente, tomando $X = (1,0,0)$ em (3.3.36), temos

$$(3.3.45)\quad f(t) + 2\, f(0) - h(t) - \lambda\cdot h(t)\{g(1) + 2\, g(0)\}$$
$$= t\cdot\{g(1) + 2\, g(0)\}$$

Substraindo (3.3.44) de (3.3.45)

$$(3.3.46)\quad f(0) - \lambda\cdot h(t)g(0) = t\cdot g(0)$$

Então conbinando (3.3.46) junto com (3.3.44), temos

$$(3.3.47)\quad B_t(1) = f(t) - \lambda\cdot h(t)g(1) - t\cdot g(1) = h(t)$$

Para $t \in [0,1]$.

Assim de (3.3.42) e da definição de $B_t(x)$ em (3.3.40), temos a seguinte relação

(3.3.48) $f(xt) = \lambda \cdot h(t)g(x) + t \cdot g(x) + x \cdot h(t)$

para todo $x,t \in [0,1]$.

Em particular, (3.3.48) é verdade para $x,t \in (0,1]$.

Seja

$$(3.3.49) \quad \begin{cases} F(x) = 1/x\ f(x) \\ G(x) = 1/x\ g(x) \\ H(x) = 1/x\ h(x) \end{cases}$$

Então, (3.3.48) torna-se

(3.3.50) $xt \cdot F(xt) = \lambda xt \cdot H(t)G(x)$

para todo $x,y \in (0,1)$

Podemos reescrever (3.3.50) na seguinte equação funcional

(3.3.51) $F(xt) = \lambda \cdot H(t)G(x) + G(x) + H(t)$

para $x,t \in (0,1]$.

Agora usando outra transformação.

Seja

$$(3.3.52) \quad \begin{cases} l(x) = 1 + \lambda \cdot F(x) \\ m(x) = 1 + \lambda \cdot G(x) \\ n(x) = 1 + \lambda \cdot H(x) \end{cases}$$

Se multiplicamos (3.3.51) por λ ($\lambda \neq 0$) e depois adicionamos 1 em ambos lados, usando a transformação em (3.3.52), é fácil ver que (3.3.51) torna-se

(3.3.53) $k(xt) = m(x)n(t)$, $\qquad x,t \in (0,1]$

Isto é a equação de Pexider a qual nós mencionamos no Lema

3.3.3, então tem a seguinte solução

Caso I. Se $m(1) \neq 0$ e $n(1) \neq 0$

A solução mensurável de (3.3.51) é

$$(3.3.54)\quad \begin{cases} k(x) = abx^{1-\beta} \\ m(x) = ax^{1-\beta} \\ n(x) = bx^{1-\beta} \end{cases}$$

para $x \in (0,1]$

onde a,b são constantes não nulas.

De (3.3.52), a solução (3.3.54) significa

$$(3.3.55)\quad \begin{cases} F(x) = \lambda^{-1}(abx^{1-\beta} - 1) \\ G(x) = \lambda^{-1}(ax^{1-\beta} - 1) \\ H(x) = \lambda^{-1}(bx^{1-\beta} - 1) \end{cases}$$

para $x \in (0,1]$.

Então pela transformação (3.3.49), (3.3.55) é o mesmo que

$$(3.3.56)\quad \begin{cases} f(x) = \frac{x}{\lambda}\,(abx^{\beta-1} - 1) \\ g(x) = \frac{x}{\lambda}\,(ax^{\beta-1} - 1) \\ h(x) = \frac{x}{\lambda}\,(bx^{\beta-1} - 1) \end{cases}$$

para $x \in (0,1]$.

E isto é a solução (3.3.21) do Teorema 3.3.1 enquando $x \neq 0$.

Caso II. Se $m(1) = 0$ ou $n(1) = 0$

Então

(3.3.57) $k(x) = m(x) = 0$, $n(x)$:arbitrária

ou

(3.3.58) $k(x) = n(x) = 0$, $m(x)$:arbitrária

Correspondendo à solução suplementado (3.3.57) ou (3.3.58), a transformação (3.3.52) dará

(3.3.59) $F(x) = G(x) = -1/\lambda$, $H(x)$: arbitrária

ou

(3.3.60) $F(x) = H(x) = -1/\lambda$, $G(x)$: arbitrária

Então da transformação (3.3.49)

(3.3.61) $f(x) = g(x) = -x/\lambda$, $h(x)$: arbitrária, $x \varepsilon (0,1]$

ou

(3.3.62) $f(x) = h(x) = -x/\lambda$, $g(x)$: arbitrária, $x \varepsilon (0,1]$

Assim as soluções (3.3.56),(3.3.61) e (3.3.62) implicam respectivamente as soluções (3.3.21),(3.3.22) e (3.3.23) no teorema enquendo $x \neq 0$. Para provar que (3.3.21),(3.3.22) e (3.3.23) também valem para $x = 0$, procedemos como segue.

Quando $x = 0$

Seja $x = 0$ em (3.3.48), temos

(3.3.63) $f(0) = g(0)\{\lambda \cdot h(t) + t\}$, $t \varepsilon [0,1]$

Analogamente, seja $t = 0$ em (3.3.48), temos

(3.3.64) $f(0) = h(0)\{\lambda g(x) + x\}$, $x \varepsilon [0,1]$

Caso I. Quando $m(1) \neq 0$ e $n(1) \neq 0$

A solução (3.3.56) implica que

(3.3.65) $f(0) = g(0) = h(0) = 0$

Então (3.3.56) é verdadeireo para todo $x \varepsilon [0,1]$.

Caso II. Quando $m(1) = 0$ ou $n(1) = 0$

Tanto a solução em (3.3.61) ou em (3.3.62) implica que

$f(0) = 0$

Quando (3.3.61) for a solução para $x \varepsilon (0,1]$, teremos por

(3.3.63) que $g(0) = 0$, e portanto $h(0)$ é arbitrário.

Portanto (3.3.61) é verdadeiro para todo $x \in [0,1]$.

Similarmente, (3.3.62) é verdadeiro para todo $x \in [0,1]$.

Isto prova o teorema.

Observação: Quando $f = g = h$, a solução em (3.3.21) reduz-se a

$$f(x) = \frac{x}{\lambda}(x^{\beta-1}-1)$$

enquando (3.3.22) e (3.3.23) reduzem-se a

$$f(x) - \frac{x}{\lambda}.$$

Agora aplicando o Teorema 3.3.1, provamos um lema de caracterização para a entropia não-adtiva de ordem β

<u>Lema 3.3.4</u> Se $H(P)$ em Δ_n, $n \geq 1$, satisfaz os seguintes postulados:

(a) Postulado da soma

(3.3.64) $$H(P) = \sum_{i=1}^{n} f(p_i), \qquad P = (p_1,\dots,p_n) \in \Delta_n,\ n = 2,3$$

onde f é mensurável.

(b) Postulado de não-aditividade

(3.3.65) $$H(P*Q) = H(P) + H(Q) + (2^{1-\beta}-1)H(P)\cdot H(Q)$$

onde $P = (p_1,\dots,p_n) \in \Delta_n$, $Q = (q_1,\dots,q_m) \in \Delta_m$, para $n, m = 2,3$, e $\beta \neq 1$.

(c) Postulado de normalização

(3.3.66) $f(1/2) = 1/2$

então $H(P)$ é a entropia não-adtiva de ordem β dada por (3.1.2).

Prova:

De (3.3.64) e (3.3.65), como vimos antes, tem-se a equação funcional

$$(3.3.67)\quad \sum_{i=1}^{n} \sum_{j=1}^{m} f(p_i q_j)$$

$$= \sum_{i=1}^{n} f(p_i) + \sum_{j=1}^{m} f(q_j) + (2^{1-\beta} - 1)\{\sum_{i=1}^{n} f(p_i)\}\{\sum_{j=1}^{m} f(q_j)\}$$

para $P = (p_1,\ldots,p_n) \in \Delta_n$, $Q = (q_1,\ldots,q_m) \in \Delta_m$, $n = 2,3;\ m = 2,3,\ \beta \neq 1.$

onde f é mensurável.

Este é o caso especial da equação funcional (3.1.1) quando $f = g = h$, e $\lambda = (2^{1-\beta}-1)$ onde $\beta \neq 1$. Portanto pelo Teorema 3.3.1, sua solução mensurável é dada por

$$(3.3.68)\quad f(p) = \frac{p}{2^{1-\beta}-1}(p^{\beta-1} - 1)$$

A outra solução $f(p) = -\dfrac{p}{2^{1-\beta} - 1}$ é desprezada pelo postulado da normalização. Do postulado da soma vemos

$$(3.3.69)\quad H(P) = \sum_{i=1}^{n} f(p_i) = \frac{\sum_{i=1}^{n} p_i^{\beta} - 1}{2^{1-\beta} - 1}, \qquad \beta \neq 1$$

Portanto H(P) é a entropia não-aditiva de ordem β.

Isto prova o lema.

3.4 EQUAÇÃO FUNCIONAL (3.1.7) EM DUAS VARIÁVEIS

Neste secção, a solução mensurável da equação funcional (3.1.7) é derivada. A fim de realizar isto, primeiro resolveremos o caso especial da equação funcional (3.1.7) quando $f = g = h$, ou a seguinte equação funcional

$$(3.4.1) \quad \sum_{i=1}^{n} \sum_{j=1}^{m} f(x_i y_j, u_i v_j)$$

$$= \sum_{i=1}^{n} f(x_i, u_i) + \sum_{j=1}^{m} f(y_j, v_j) + \lambda \sum_{i=1}^{n} \sum_{j=1}^{m} f(x_i, u_i) f(y_j, v_j)$$

para $X = (x_1, \ldots, x_n) \in \Delta_n$, $Y = (y_1, \ldots, y_m) \in \Delta_m$, $U = (u_1, \ldots, u_n) \in \Delta'_n$, $V = (v_1, \ldots, v_m) \in \Delta'_m$, $n = 2,3$; $m = 2,3$ e λ é uma constante não nula.

A fim de provar os teoremas neste secção, precisamos provar primeira varios lemas.

O lema seguinte é uma extensão do Lema 3.3.2.da equação de Cauchy em uma variável em $(0,1]$ para a função de duas variáveis na região $(0,1] \times (0,1]$.

<u>Lema 3.4.1</u> A função mensurável $L(\cdot,\cdot)$ satisfazendo a equação funcional

$$(3.4.2) \quad L(xy, uv) = L(x,u) \cdot L(y,v)$$

para (x,u), $(y,v) \in (0,1] \times (0,1]$.

tem a solução não- trivial

$$(3.4.3) \quad L(x,v) = x^{\alpha} v^{\beta}, \qquad x,v \in (0,1]$$

onde α,β são constantes arbitrarias.

Prova:

Seja $u = v = 1$, (3.2) reduz-se à equação funcional em uma variável, pelo Lema 3.3.2, tem a solução não-trivial

(3.4.4) $L(x,1) = x^{\alpha}$, $x \in (0,1]$

Similarmente, seja $x = y = 1$ em (3.4.2), nós temos

(3.4.5) $L(1,u) = u^{\beta}$, $u \in (0,1]$

Agora, seja $y = u = 1$ em (3.4.2), produz

(3.4.6) $L(x,v) = L(x,1)\cdot L(1,v)$

e usando (3.4.4) e (3.4.5), temos

(3.4.7) $L(x,v) = x^{\alpha}v^{\beta}$, $(x,v) \in (0,1] \times (0,1]$

onde α,β são constantes arbitrárias.

<u>Lema 3.4.2</u> Se f é a solução mensurável de (3.4.1), então para $z,w \in [0,1]$,

(3.4.8)
$$\sum_{i=1}^{n} G(x_i z, x_i w) - \left(\sum_{i=1}^{n} x_i^{\gamma}\right) G(z,w)$$
$$= z\left\{\sum_{i=1}^{n} G(x_i,0) - \left(\sum_{i=1}^{n} x_i^{\gamma}\right) G(1,0)\right\} + w\left\{\sum_{i=1}^{n} G(0,x_i) - \left(\sum_{i=1}^{n} x_i^{\gamma}\right) G(0,1)\right\}$$

onde

(3.4.9) $G(x,y) = \lambda\cdot f(x,y) + x$

Prova:

Para $u_i = x_i$, $1 \leqq i \leqq n$, $v_j = y_j$, $1 \leqq j \leqq m$, (3.4.1) tem a seguinte forma

(3.4.10)
$$\sum_{i=1}^{n} \sum_{j=1}^{m} f(x_i y_j, x_i y_j)$$
$$= \sum_{i=1}^{n} f(x_i,x_i) + \sum_{j=1}^{m} f(y_j,y_j) + \lambda \sum_{i=1}^{n} \sum_{j=1}^{m} f(x_i,x_i)\, f(y_j,y_j)$$

Esta equação funcional envolve somente uma variável, e de fato, este é um caso especial da equação funcional (3.1.1) quando $f = g = h$. Portanto, do Teorema 3.3.1, temos as seguintes soluções.

(3.4.11) $$f(x,x) = \frac{x}{\lambda}\,(x^{\gamma-1} - 1),$$

ou

(3.4.12) $$f(x,x) = -\frac{x}{\lambda}$$

Se multiplicamos λ em ambos lados da equação funcional (3.4.1) e depois adicionamos 1 a ambos lados também, (3.4.1) se reduzira a

(3.4.13) $$\sum_{i=1}^{n}\sum_{j=1}^{m} \lambda\, f(x_iy_j, u_iv_j) + 1$$

$$= \{\sum_{i=1}^{n} \lambda\, f(x_i,u_i) + 1\}\{\sum_{j=1}^{m} \lambda\, f(y_j,v_j) + 1\}$$

para $X = (x_1,\ldots,x_n) \in \Delta_n$, $Y = (y_1,\ldots,y_m) \in \Delta_m$, $U = (u_1,\ldots,u_n) \in \Delta'_n$, $V = (v_1,\ldots,v_m) \in \Delta'_m$, $n = 2,3$; $m = 2,3$ e λ é uma constante não nula.

Assim, seja

(3.4.14) $$G(x,y) = \lambda\, f(x,y) + x$$

(3.4.1) torna-se

(3.4.15) $$\sum_{i=1}^{n}\sum_{j=1}^{m} G(x_iy_j, u_iv_j) = \{\sum_{i=1}^{n} G(x_i,u_i)\}\{\sum_{j=1}^{m} G(y_j,v_j)\}$$

para $X = (x_1,\ldots,x_n) \in \Delta_n$, $Y = (y_1,\ldots y_m) \in \Delta_n$, $U = (u_1,\ldots,u_n) \in \Delta'_n$, $V = (v_1,\ldots,v_m) \in \Delta'_m$ e $n = 2,3$; $m = 2,3$.

Quando $X = U$, $Y = V$, correspondentmente a (3.4.11) e (3.4.12), a solução não-trivial de (3.4.15) é

(3.4.16) $$G(x,x) = x^{\gamma}, \qquad x \in \left]0,1\right]$$

Especialmente,

(3.4.17) $G(0,0) = 0;\ G(1,1) = 1$

Para X, U fixos, $X \in \Delta_n$, $U \in \Delta'_n$, seja $Y = (y_1, y_2, 1-y_1-y_2)$, $V = (v_1, v_2, 1-v_1-v_2)$ em (3.4.5), então

(3.4.18) $$\sum_{i=1}^{n} \{G(x_iy_1, u_iv_1) + G(x_iy_2, u_iv_2) + G(x_i(1-y_1-y_2), u_i(1-v_1-v_2))\}$$

$$= \{\sum_{i=1}^{n} G(x_i, u_i)\}\{G(y_1, v_1) + G(y_2, v_2) + G(1-y_1-y_2, 1-v_1-v_2)\}$$

Novamente, seja $Y = (y_1+y_2, 1-y_1-y_2)$ e $V = (v_1+v_2, 1-v_1-v_2)$ em (3.4.15), com a ajuda de (3.4.17), temos

(3.4.19) $$\sum_{i=1}^{n} \{G(x_i(y_1+y_2), u_i(v_1+v_2)) + G(x_i(1-y_1-y_2), u_i(1-v_1-v_2))\}$$

$$= \{\sum_{i=1}^{n} G(x_i, u_i)\}\{G(y_1+y_2, v_1+v_2) + G(1-y_1-y_2, 1-v_1-v_2)\}$$

Substraindo (3.4.18) de (3.4.19), então

(3.4.20) $$\sum_{i=1}^{n} G(x_i(y_1+y_2), u_i(v_1+v_2)) - \{\sum_{i=1}^{n} G(x_i, u_i)\}G(y_1+y_2, v_1+v_2)$$

$$= \{\sum_{i=1}^{n} G(x_iy_1, u_iv_1) - [\sum_{i=1}^{n} G(x_i, u_i)]G(y_1, v_1)\} +$$

$$\{\sum_{i=1}^{n} G(x_iy_2, u_iv_2) - [\sum_{i=1}^{n} G(x_i, u_i)]\ G(y_2, v_2)\}$$

Para $X \in \Delta_n$, $U \in \Delta'_n$, X e U fixos, definimos

(3.4.21) $$A_{XU}(z,w) = \sum_{i=1}^{n} G(x_iz, u_iw) - \{\sum_{i=1}^{n} G(x_i, u_i)\}\ G(z,w)$$

É fécil ver que $A_{XU}(\cdot,\cdot)$ é adtiva, ou

(3.4.22) $$A_{XU}(y+w, v+t) = A_{XU}(y,v) + A_{XU}(w,t)$$

para $(y,v),(w,t),(y+w,v+t) \in [0,1] \times [0,1]$

Pelo Lema 2.3.4 o qual é uma extensão simples do resultado de Daróczy e Losonczi (1967) em uma variável, sabemos que

$$(3.4.23) \quad A_{XU}(z,w) = z\, A_{XU}(1,0) + w\, A_{XU}(0,1), \qquad (z,w) \in [0,1] \times [0,1]$$

Quando $X = U$, desde que $A_{XU}(z,w)$ é definido como em (3.4.21) com a ajuda de (3.4.16), temos

$$(3.4.24) \quad A_{XX}(z,w) = \sum_{i=1}^{n} G(x_i z, x_i w) - \left(\sum_{i=1}^{n} x_i^{\gamma} \right) G(z,w)$$

Portanto, quando $X = U$, (3.4.23) produz (3.4.8).

Isto prova o Lema 3.4.2.

Agora estamos em condições de provar nosso primeiro teorema.

As soluções de (3.4.1) são dadas no teorema seguinte.

<u>Teorema 3.4.1</u> Se $f: [0,1] \times [0,1] \to R$ (Reais) é mensurável em cada de suas variáveis, e satisfaz (3.4.1) para $X \in \Delta_n$, $Y \in \Delta_m$, $U \in \Delta_n'$, $V \in \Delta_m'$, $n = 2,3$; $m = 2,3$ e λ é uma constante não nula, então as soluções não-triviais de f são dadas por

$$(3.4.25) \quad f(x,y) = \frac{1}{\lambda}(x^{\alpha}y^{\beta} - x), \qquad \alpha \neq 0,\ \beta \neq 0$$

ou

$$(3.4.26) \quad f(x,y) = \frac{1}{\lambda}(x^{\gamma} - x), \qquad \gamma \neq 0$$

ou

$$(3.4.27) \quad f(x,y) = \frac{1}{\lambda}(y^{\gamma} - x), \qquad \gamma \neq 0$$

Prova:

Seja $X = Y = (1,0)$, $U = V = (0,1)$ em (3.4.15), então, com a ajuda de (3.4.17), temos

(3.4.28) $G(1,0) + G(0,1) = \{G(1,0) + G(0,1)\}^2$

Portanto

(3.4.29) $G(1,0) + G(0,1) = 0$

Ou

(3.4.30) $G(1,0) + G(0,1) = 1$

Caso I. $G(1,0) + G(0,1) = 0$

Seja $Y = (1,0)$, $V = (0,1)$ em (3.4.15), então

(3.4.31) $$\sum_{i=1}^{n} G(x_i,0) + \sum_{i=1}^{n} G(0,u_i) = 0$$

Substituindo $X = (x_1,x_2,1-x_1-x_2)$, $U = (u_1,u_2,1-u_1-u_2)$ em (3.4.31), produz

(3.4.32) $$G(x_1,0) + G(x_2,0) + G(1-x_1-x_2,0) + G(0,u_1) + G(0,u_2) + G(0,1-u_1-u_2) = 0$$

Novamente, substituindo $X = (x_1+x_2,1-x_1-x_2)$, $U = (u_1,u_2,1-u_1-u_2)$ em (3.4.31), produz

(3.4.33) $$G(x_1+x_2,0) + G(1-x_1-x_2,0) + G(0,u_1+u_2) + G(0,1-u_1-u_2) = 0$$

Substraindo (3.4.32) de (3.4.33), temos

(3.4.34) $$G(x_1+x_2,0) + G(0,u_1+u_2) = G(x_1,0) + G(0,u_1) + G(x_2,0) + G(0,u_2)$$

Então, definimos

(3.4.35) $A(a,b) = G(a,0) + G(0,b)$

É claro que (3.4.34) pode ser reduzida a uma função aditiva de A, ou

(3.4.36) $A(x_1+x_2,u_1+u_2) = A(x_1,u_1) + A(x_2,u_2)$

Portanto

(3.4.37) $A(a,b) = a\cdot(1,0) + b\cdot A(0,1)$

para $(a,b) \in [0,1] \times [0,1]$.

A relação (3.4.37) com a definição de $A(a,b)$ em (3.3.35), produz

(3.4.38) $G(a,0) + G(0,b) = a\cdot G(1,0) + b\cdot G(0,1)$

para todo $(a,b) \in [0,1] \times [0,1]$.

Em particular,

(3.4.39) $G(a,0) = a\cdot G(1,0),\qquad a \in [0,1]$

e

(3.4.40) $G(0,b) = b\cdot G(0,1),\qquad b \in [0,1]$

Agora, seja $x_i = u_i$ para $1 \leqq i \leqq n$; $Y = (v,1-v)$, $V = (v,0)$ para $v \in (0,1)$ em (3.4.15), temos

(3.4.41) $$\sum_{i=1}^{n} G(x_i v, x_i v) + \sum_{i=1}^{n} G(x_i(1-v),0)$$

$$= \{\sum_{i=1}^{n} G(x_i,x_i)\}\{G(v,v) + G(1-v,0)\}$$

Desde que conhecemos de (3.4.16) que $G(t,t) = t^{\gamma}$, para todo $t \in [0,1]$ e de (3.4.39), temos

(3.4.42) $$\sum_{i=1}^{n} (x_i v)^{\gamma} + G(1,0)\ \{\sum_{i=1}^{n} x_i(1-v)\}$$

$$= (\sum_{i=1}^{n} x_i^{\gamma})\{v^{\gamma} + (1-v)\ G(1,0)\}$$

Desde que $\sum_{i=1}^{n} x_i = 1$ e $v \neq 0$, (3.4.42) é equivalente a

(3.4.43) $$G(1,0) = (\sum_{i=1}^{n} x_i^{\gamma})\ G(1,0)$$

para todo $X = (x_1,\dots,x_n) \in \Delta_n$

Conclui-se que

(3.4.44) $G(1,0) = 0$

De (3.4.29), é claro que

(3.4.45) $G(0,1) = 0$

De (3.4.39) e (3.4.40), vemos que

(3.4.46) $G(a,0) = 0, \qquad a \in [0,1]$

e

(3.4.47) $G(0,b) = 0, \qquad b \in [0,1]$

Desde que (3.4.23) com a definição de $A_{XU}(z,w)$ dada em (3.4.21) produz

(3.4.48) $$\sum_{i=1}^{n} G(x_i z, u_i w)$$

$$= \{\sum_{i=1}^{n} G(x_i,u_i)\}\, G(z,w) + z\{\sum_{i=1}^{n} G(x_i,0) - [\sum_{i=1}^{n} G(x_i,u_i)]\cdot G(1,0)\} +$$

$$a\{\sum_{i=1}^{n} G(0,u_i) - [\sum_{i=1}^{n} G(x_i,u_i)]\cdot G(0,1)\}$$

Com a ajuda de (3.4.44)–(3.4.47), (3.4.48) reduz-se a

(3.4.49) $$\{\sum_{i=1}^{n} G(x_i,u_i)\}\cdot G(z,w) = \sum_{i=1}^{n} G(x_i z, u_i w)$$

para $X = (x_1,\dots,x_n) \in \Delta_n$, $U = (u_1,\dots,u_n) \in \Delta'_n$, $\qquad z,w \in [0,1]$

Agora, para z,w fixos, $z,w \in [0,1]$, nós definimos

(3.4.50) $B_{zw}(t,s) = G(tz,sw) - G(t,s)\cdot G(z,w)$

Pela relação em (3.4.49), é fácil ver que para z.w fixos, $z,w \in [0,1]$, $B_{zw}(t,s)$ é aditiva em (t,s) para $(t,s) \in [0,1] \times [0,1]$, ou

(3.4.51) $B_{zw}(x_1+x_2,u_1+u_2) = B_{zw}(x_1,u_1) + B_{zw}(x_2,u_2)$

onde $(x_1,u_1),(x_2,u_2),(x_1+x_2,u_1+u_2) \in [0,1] \times [0,1]$.

Similarmente, sabemos que

(3.4.52) $B_{zw}(t,s) = t \cdot B_{zw}(1,0) + s \cdot B_{zw}(0,1)$

Da definição de $B_{zw}(t,s)$ em (3.4.50), vemos que

(3.4.53) $B_{zw}(1,0) = G(z,0) - G(1,0) \cdot G(z,w) = 0$

e

(3.4.54) $B_{zw}(0,1) = G(0,w) - G(0,1) \cdot G(z,w) = 0$

Combinando (3.4.52), (3.4.53) e (3.4.54), nós temos que

(3.4.55) $B_{zw}(t,s) = 0$

ou

(3.4.56) $G(tz,sw) = G(t,s) \cdot G(z,w)$

onde $(t,s),(z,w) \in [0,1] \times [0,1]$.

Especialmente, (3.4.56) é verdadeiro em $(0,1] \times (0,1]$.

Pelo Lema 3.4.1, (3.4.56) tem a solução mensurável como

(3.4.57) $G(x,y) = x^{\alpha}y^{\beta}, \qquad (x,y) \in (0,1] \times (0,1]$

onde α,β são parametros abtrários.

A fim de termos (3.4.16) consistente com (3.4.57), vemos que

$\alpha + \beta = \gamma$

O fato que de (3.4.46) e (3.4.47)

$G(a,0) = G(0,a) = 0, \qquad a \in [0,1]$

tornara a solução em (3.4.57) verdadeiro para todo $(x,y) \in [0,1] \times [0,1]$.

Agora, pela definição de $G(x,y)$ em (3.4.14), temos

(3.4.58) $f(x,y) = \frac{1}{\lambda}(x^{\alpha}y^{\beta} - x), \qquad \alpha,\beta \neq 0, \ (x,y) \in [0,1] \times [0,1]$

Caso II. $G(1,0) + G(0,1) = 1$

Seja $Y = (0,1)$, $V = (1,0)$ em (3.4.15), então

$$\sum_{i=1}^{n} G(x_i,0) + \sum_{i=1}^{n} G(0,u_i) = \sum_{i=1}^{n} G(x_i,u_i) \tag{3.4.59}$$

Portanto, nós definimos

$$B(a,b) = G(a,b) - G(a,0) - G(0,b) \tag{3.4.60}$$

para $(a,b) \varepsilon [0,1] \times [0,1]$.

Então, como antes, usando a relação em (3.4.59), depois de uma pequena manipulação, concluimos que $B(a,b)$ como definida acima em (3.4.60) é aditiva em duas variáveis na região $[0,1] \times [0,1]$.

Pelo Lema 2.4.3 no capítulo anterior, então

$$B(a,b) = a\cdot B(1,0) + b\cdot B(0,1) \tag{3.4.61}$$

De (3.4.60) e (3.4.17), temos

$$B(1,0) = 0 \tag{3.4.62}$$

e

$$B(0,1) = 0 \tag{3.4.63}$$

Portanto da definição de $B(a,b)$ em (3.4.60), sabemos que

$$G(a,b) = G(a,0) + G(0,b), \qquad (a,b) \varepsilon [0,1] \times [0,1] \tag{3.4.64}$$

Substituindo (3.4.64) em (3.4.15), temos

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \{G(x_iy_j,0) + G(0,u_iv_j)\} \tag{3.4.65}$$

$$= \{ \sum_{i=1}^{n} [G(x_i,0) + G(0,u_i)]\}\{ \sum_{i=1}^{n} [G(y_j,0) + G(0,v_j)]\}$$

Seja $U = (1,0)$, $V = (1,0)$ em (3.4.65), usando o fato $G(0,0) = 0$, temos

$$\sum_{i=1}^{n} \sum_{j=1}^{m} G(x_iy_j,0) + G(0,1) \tag{3.4.66}$$

$$= \{ \sum_{i=1}^{n} G(x_i,0) + G(0,1)\}\{ \sum_{j=1}^{m} G(y_j,0) + G(0,1)\}$$

Seja

(3.4.67) $k(t) = G(t,0) + t\cdot G(0,1)$

Então, é claro que

(3.4.68) $k(0) = 0$

e

(3.4.69) $k(1) = 1$

E portanto, (3.4.66) torna-se

(3.4.70) $\sum_{i=1}^{n} \sum_{j=1}^{m} k(x_i y_j) = \{ \sum_{i=1}^{n} k(x_i)\}\{ \sum_{j=1}^{m} k(y_j)\}$

Usando outra transformação, seja

(3.4.71) $k(t) = \lambda m(t) + t$

Então, (3.4.70) reduz-se à forma seguinte

(3.4.72) $\sum_{i=1}^{n} \sum_{j=1}^{m} \lambda\cdot m(x_i y_j) + 1 = \{ \sum_{i=1}^{n} \lambda\cdot m(x_i) + 1\}\{ \sum_{j=1}^{m} \lambda\cdot m(y_j) + 1\}$

Ou

(3.4.73) $\sum_{i=1}^{n} \sum_{j=1}^{m} m(x_i y_j) = \sum_{i=1}^{n} m(x_i) + \sum_{j=1}^{m} m(y_j) + \lambda \sum_{i=1}^{n} \sum_{j=1}^{m} m(x_i)m(y_j)$

para $X = (x_1,\dots,x_n) \in \Delta_n$, $Y = (y_1,\dots,y_m) \in \Delta_m$, $n= 2,3$; $m = 2,3$ e λ é uma constante não nula.

Este é o caso especial da equação funcional (3.1.1) quando $f = g = h$, portanto podemos aplicar o resultado do Teorema 3.3.1 e obter as soluções não-triviais como segue:

(3.4.74) $m(t) = \frac{t}{\lambda} (t^{\alpha-1} - 1)$, $\qquad t \in [0,1]$

onde α é um parametro arbitrário com a convenção $0^{\beta} = 1$.

Suplementado com a solução trivial como

(3.4.75) $m(t) = -\frac{t}{\lambda}$

Correspondendo a (3.4.74) e (3.4.75), temos o respectivamente o seguinte.

(3.4.76) $k(t) = t^{\alpha}, \qquad t \in [0,1]$

e

(3.4.77) $k(t) = 0, \qquad t \in [0,1]$

Da definição de k em (3.4.67), (3.4.76) e (3.4.77) significa as duas relações seguintes respectivamente.

(3.4.78) $G(t,0) + t\,G(0,1) = t^{\alpha}, \qquad t \in [0,1]$

e

(3.4.79) $G(t,0) + t\,G(0,1) = 0, \qquad t \in [0,1]$

Similarmente, começamos fazendo X = (1,0), Y = (1,0) em (3.4.65), temos

(3.4.80) $\sum_{i=1}^{n} \sum_{j=1}^{m} G(0,u_i v_j) + G(1,0)$

$$= \{ \sum_{i=1}^{n} G(0,u_i) + G(1,0)\}\{ \sum_{j=1}^{m} G(0,v_j) + G(1,0)\}$$

Seja

(3.4.81) $n(t) = G(0,t) + t\,G(1,0)$

Novamente, nós temos

(3.4.82) $\sum_{i=1}^{n} \sum_{j=1}^{m} n(u_i v_j) = \{ \sum_{i=1}^{n} n(u_i)\}\{ \sum_{j=1}^{m} n(v_j)\}$

Usando a transformação

(3.4.83) $n(t) = \lambda\cdot k(t) + t$

Então, (3.4.82) reduz-se do caso especial de (3.1.1) quando $f = g = h$, aplicando o resultado do Teorema 3.3.1, obtemos as soluções correspondentes para n no que se segue.

(3.4.84) $n(t) = t^{\beta}$, $t \in [0,1]$

e

(3.4.85) $n(t) = 0$, $t \in [0,1]$

Desde que $n(t)$ é definido como em (3.4.81), (3.4.84) e (3.4.85) significam respectivamente as seguintes rela ções.

(3.4.86) $G(0,t) + t\, G(1,0) = t^{\beta}$, $t \in [0,1]$

e

(3.4.87) $G(0,t) + t\, G(1,0) = 0$, $t \in [0,1]$.

Desde que conhecemos que

$G(t,t) = t^{\gamma}$, $t \in [0,1]$

e

$G(0,1) + G(1,0) = 1$

De (3.4.78), (3.4.86) e (3.4.64), é claro que

(3.4.88) $t^{\gamma} + t = t^{\alpha} + t^{\beta}$, $t \in [0,1]$

ou

(3.4.88)' $t^{\gamma} + t = 0$, $t \in [0,1]$

(3.4.88)' é impossível acontecer, portanto nós sòmente consideramos (3.4.88).

Os parametros envolvendo a equação em (3.4.88) para todo $t \in [0,1]$ tem que satisfazer uma dos seguintes relações.

(3.4.89) $\alpha = 1,\ \beta = \gamma \neq 1$

ou

(3.4.90) $\beta = 1, \ \alpha = \gamma \neq 1$

Para ver isto, de (3.4.88), temos

(3.4.91) $(t^{\gamma} + 1)^2 = (t^{\alpha} + t^{\beta})^2, \qquad t \in [0,1]$

ou

(3.4.92) $t^{2\gamma} + 2t^{\gamma+1} + t^{2} = t^{2\alpha} + t^{2\beta} + 2t^{\alpha+\beta}$

Usando a relação em (3.4.88), substituindo t por t^2, pode ser visto que

(3.4.93) $t^{2\alpha} + t^{2\beta} = (t^2)^{\alpha} + (t^2)^{\beta} = (t^2)^{\gamma} + (t^2) = t^{2\gamma} + t^2$

Portanto (3.4.92) é equivalente a

(3.4.94) $2t^{\gamma+1} = 2t^{\alpha+\beta}, \qquad t \in [0,1]$

Isto implica que

(3.4.95) $\gamma + 1 = \alpha + \beta$

Substituindo esta relação (3.4.95) na relação

(3.4.96) $t^{\gamma} - t^{\alpha} = t^{\beta} - t, \qquad t \in [0,1]$

ou

(3.4.97) $t^{\gamma} - t^{\beta} = t^{\alpha} - t, \qquad t \in [0,1]$

Nós temos respectivamente

(3.4.98) $t^{\alpha+\beta-1} - t^{\alpha} = t^{\beta} - t, \qquad t \in [0,1]$

ou

(3.4.99) $t^{\alpha+\beta-1} - t^{\beta} = t^{\alpha} - t, \qquad t \in [0,1]$

(3.4.98) é o mesmo como

(3.4.100) $t^{\alpha-1}(t^{\beta} - 1) = t^{\beta} - t, \qquad t \in [0,1]$

Isto implica que

(3.4.101) $\alpha = 1$

E por outro lado de (3.4.95), temos

(3.4.102) $\gamma = \beta \neq 1$

Similarmente, (3.4.99) dará que, se $\alpha \neq 1$, então

(3.4.103) $\beta = 1,\ \gamma = \beta \neq 1$

Portanto, (3.4.101), (3.4.102) e (3.4.103) provam (3.4.89) e (3.4.90).

Quando (3.4.89) é verdade, então de (3.4.78), temos

(3.4.104) $G(t,0) + t\,G(0,1) = t$

Seja $Y = (v,1-v)$, $V = (v,0)$ para $v \in (0,1]$ e $U = X = (x_1,\ldots,x_n) \in \Delta_n$ em (3.4.15), temos

(3.4.105) $$\sum_{i=1}^{n} G(x_i v,\ x_i v) + \sum_{i=1}^{n} G(x_i(1-v),\ 0)$$

$$= \{\sum_{i=1}^{n} G(x_i,x_i)\}\{G(v,0) + G(1-v,0)\}$$

De (3.4.104) e (3.4.16), sabemos o seguinte

(3.4.106) $G(t,0) = t - t\,G(0,1);\ G(t,t) = t^{\gamma}, \qquad t \in [0,1]$

Substituindo (3.4.106) em (3.4.105), então

(3.4.107) $$v^{\gamma} \sum_{i=1}^{n} x_i^{\gamma} + (1-v) - (1-v)\,G(0,1)$$

$$= (\sum_{i=1}^{n} x_i^{\gamma})\{v^{\gamma} + (1-v) - (1-v)\,G(0,1)\}$$

ou

(3.4.108) $$1 - G(0,1) = (\sum_{i=1}^{n} x_i^{\gamma})\{1 - G(0,1)\}$$

para todo $(x_1,\ldots,x_n) \in \Delta_n$.

Isto implica que

(3.4.109) $G(0,1) = 1$

(3.4.109) e (3.4.30) dá

(3.4.110) $G(1,0) = 0$

Portanto de (3.4.104), sabemos que

(3.4.111) $G(t,0) = 0, \qquad t \in [0.1]$

Por outro lado, combinando (3.4.16), (3.4.111) e (3.4.64), temos

(3.4.112) $G(0,t) = t^{\gamma}. \qquad t \in [0,1]$

e

(3.4.113) $G(a,b) = G(a,0) + G(0,b) = b^{\gamma}$

para $(a,b) \in [0,1] \times [0,1]$.

Correpondentemente, da definição de G em (3.4.9), temos

(3.4.114) $f(a,b) = \frac{1}{\lambda}(b^{\gamma} - a), \qquad (a,b) \in [0,1] \times [0,1]$

Similarmente, se (3.4.90) é o caso, então

(3.4.115) $f(a,b) = \frac{1}{\lambda}(a^{\gamma} - a), \qquad (a,b) \in [0,1] \times [0,1]$

Isto prova o Teorema 3.4.1.

Agora estamos para derivar a solução da equação funcional (3.1.7) a qual é uma versão generalizado da equação funcional (3.4.1) com três funções mensuráveis diferentes. O seguinte teorema é provado utilizando-se o Teorema 3.3.1 e a discussão a qual é usado na prova do Teorema 3.4.1.

<u>Teorema 3.4.2</u> Se $f,g,h : [0,1] \times [0,1] \to R$ (Reais) são mensuráveis em cada uma de suas variáveis e satisfazem (3.1.7), então as soluções não-triviais de f, g e h são dadas por

$$(3.4.116)\quad \begin{cases} f(x,y) = 1/\lambda\cdot(abx^{\alpha}y^{\beta} - x), & \alpha,\beta \neq 0 \\ g(x,y) = 1/\lambda\cdot(ax^{\alpha}y^{\beta} - x), & \alpha,\beta \neq 0 \\ h(x,y) = 1/\lambda\cdot(bx^{\alpha}y^{\beta} - x), & \alpha,\beta \neq 0 \end{cases}$$

$$(3.4.117)\quad \begin{cases} f(x,y) = 1/\lambda\cdot(abx^{\gamma} - x), & \gamma \neq 0 \\ g(x,y) = 1/\lambda\cdot(ax^{\gamma} - x), & \gamma \neq 0 \\ h(x,y) = 1/\lambda\cdot(bx^{\gamma} - x), & \gamma \neq 0 \end{cases}$$

$$(3.4.118)\quad \begin{cases} f(x,y) = 1/\lambda\cdot(ay^{\gamma} - x), & \gamma \neq 0 \\ g(x,y) = 1/\lambda\cdot(ay^{\gamma} - x), & \gamma \neq 0 \\ h(x,y) = 1/\lambda\cdot(by^{\gamma} - x), & \gamma \neq 0 \end{cases}$$

suplementada com as soluções triviais

(3.4.119) $f(x,y) = g(x,y) = -x/\lambda$, $h(x,y)$: arbitrária.

(3.4.120) $f(x,y) = h(x,y) = -x/\lambda$, $g(x,y)$: arbitrária.

Prova:

Na equação funcional (3.1.7), fazendo $X = U = (x_1,\ldots,x_n) \in \Delta_n$, $Y = V = (y_1,\ldots,y_m) \in \Delta_m$, então

$$(3.4.121)\quad \sum_{i=1}^{n}\sum_{j=1}^{m} f(x_iy_j, x_iy_j)$$

$$= \sum_{i=1}^{n} g(x_i,x_i) + \sum_{j=1}^{m} h(y_j,y_j) + \lambda \sum_{i=1}^{n}\sum_{j=1}^{m} g(x_i,x_i)h(y_j,y_j)$$

Esta é uma equação funcional de uma variável, portanto, pelo Teorema 3.3.1, temos

$$f(x,x) = 1/\lambda\cdot(abx^{\gamma-1} - 1)$$

$$g(x,x) = 1/\lambda\cdot(ax^{\gamma-1} - 1)$$

$$h(x,x) = 1/\lambda\cdot(bx^{\gamma-1} - 1)$$

para $ab \neq 0$.

(3.4.123) $f(x,x) = g(x,x) = -x/\lambda$, $h(x,x)$: arbitrária

ou

(3.4.124) $f(x,x) = h(x,x) = -x/\lambda$, $g(x,x)$: arbitrária

Similar à transformação nós temos para resolver (3.1.1), seja

$$(3.4.125)\quad \begin{cases} F(x,y) = \lambda\, f(x,y) + x \\ G(x,y) = \lambda\, g(x,y) + x \\ H(x,y) = \lambda\, h(x,y) + x \end{cases}$$

em (3.1.7), então torna-se

$$(3.4.126)\quad \sum_{i=1}^{n} \sum_{j=1}^{m} F(x_i y_j, u_i v_j) = \{\sum_{i=1}^{n} G(x_i, u_i)\}\{\sum_{j=1}^{m} H(y_j, v_j)\}$$

para $X = (x_1,\dots,x_n) \in \Delta_n$, $U = (u_1,\dots,u_n) \in \Delta'_n$, $Y = (y_1,\dots,y_m) \in \Delta_m$, $V = (v_1,\dots,v_m) \in \Delta'_m$ e $n = 2,3$; $m = 2,3$.

Correspondendo as soluções em (3.4.122), (3.4.123) e (3.4.124), temos o seguinte para F, G e H respectivamente,

$$(3.4.127)\quad \begin{cases} F(x,x) = abx^{\gamma} \\ G(x,x) = ax^{\gamma} \\ H(x,x) = bx^{\gamma} \end{cases}$$

onde γ é um parametro e $ab \neq 0$.

(3.4.128) $F(x,x) = G(x,x) = 0$, $H(x,x)$: arbitrária

ou

(3.4.129) $F(x,x) = H(x,x) = 0$, $G(x,x)$: arbitrária

Caso I. Quando $F(x,x) = abx^{\gamma}$, $G(x,x) = ax^{\gamma}$, $H(x,x) = bx^{\gamma}$

Em particular, (3.4.123) da

(3.4.130) $F(0,0) = G(0,0) = H(0,0) = 0$

e

(3.4.131) $F(1,1) = ab$, $G(1,1) = a$, $H(1,1) = b$

para $ab \neq 0$.

Seja $Y = (1,0)$, $V = (1,0)$ em (3.4.126), com (3.4.130) e (3.4.131), então temos

(3.4.132) $$\sum_{i=1}^{n} F(x_i,u_i) = b\cdot \sum_{i=1}^{n} G(x_i,u_i)$$

para todo $X = (x_1,\ldots,x_n) \in \Delta_n$, $U = (u_1,\ldots,u_n) \in \Delta'_n$, $n = 2.3$.

Similarmente, seja $X = (1,0)$, $U = (1,0)$ em (3.4.126), temos

(3.4.133) $$\sum_{j=1}^{m} F(y_j,v_j) = a\cdot \sum_{j=1}^{m} H(y_j,v_j)$$

para todo $Y = (y_1,\ldots,y_m) \in \Delta_m$, $V = (v_1,\ldots,v_m) \in \Delta'_m$, $m = 2,3$.

Substituindo (3.4.132), (3.4.133) em (3.4.126), então

(3.4.134) $$\sum_{i=1}^{n} \sum_{j=1}^{m} F(x_iy_j,u_iv_j) = \frac{1}{ab}\{\sum_{i=1}^{n} F(x_i,u_i)\}\{\sum_{j=1}^{m} F(y_j,v_j)\}$$

Esta é uma equação funcional envolvendo sòmente uma função F, seja

(3.4.135) $$R(x,y) = \frac{1}{ab} F(x,y)$$

(3.4.134) torna-se

(3.4.136) $$\sum_{i=1}^{n} \sum_{j=1}^{m} R(x_iy_j,u_iv_j) = \{\sum_{i=1}^{n} R(x_i,u_i)\}\{\sum_{j=1}^{m} R(y_j,v_j)\}$$

Esta é a equação funcional do tipo (3.4.15) a qual nós discutímos na prova do Teorema 3.4.1 e tem as soluções não-triviais como segue.

(3.4.137) $R(x,y) = x^{\alpha}y^{\beta}$, $\qquad \alpha,\beta \neq 0$

(3.4.138) $R(x,y) = x^{\gamma}$, $\qquad \gamma \neq 0$

(3.4.139) $R(x,y) = y^{\gamma}$, $\qquad \gamma \neq 0$

De (3.4.135), assim as soluções correspondentes para F são

(3.4.140) $F(x,y) = abx^{\alpha}y^{\beta}$, $\alpha,\beta \neq 0$

(3.4.141) $F(x,y) = abx^{\gamma}$, $\gamma \neq 0$

(3.4.142) $F(x,y) = aby^{\gamma}$, $\gamma \neq 0$

Combinando cada uma de (3.4.140), (3.4.141) e (3.4.142) com (3.4.132), podemos resolver $G(x,y)$ como segue:

I(i) Quando $F(x,y) = abx^{\alpha}y^{\beta}$, $\alpha,\beta \neq 0$

De (3.4.132), usando (3.4.130), temos

(3.4.143) $$\sum_{i=1}^{n} G(x_i,u_i) = \frac{1}{b} \cdot \sum_{i=1}^{n} F(x_i,u_i) = a\cdot \sum_{i=1}^{n} x_i^{\alpha}u_i^{\beta}$$

para $X = (x_1,\ldots,x_n) \in \Delta_n$, $U = (u_1,\ldots,u_n) \in \Delta_n'$; $\alpha,\beta \neq 0$.

Seja

(3.4.144) $A(t,s) = G(t,s) - a\cdot t^{\alpha}s^{\beta}$

para $(t,s) \in [0,1] \times [0,1]$; $\alpha,\beta \neq 0$.

Usando a relação em (3.4.143), depois de uma pequena manipulação, então é fácil ver que $A(t,s)$ é aditiva em $[0,1] \times [0,1]$, pelo Lema 2.4.3, temos

(3.4.145) $A(x,y) = x\,A(1,0) + y\,A(0,1)$, $(x,y) \in [0,1] \times [0,1]$

Com a definição de $A(t,s)$ em (3.4.144), (3.4.145) torna-se

(3.4.146) $G(x,y) - ax^{\alpha}y^{\beta} = x\,G(1,0) + y\,G(0,1)$

Para encontrar os valores exactos para $G(1,0)$ e $G(0,1)$, seja $X = (x,1-x)$, $U = (u,0)$ para $x,u \in (0,1)$ em (3.4.132), temos

(3.4.147) $F(x,u) + F(1-x,0) = b\,\{G(x,u) + G(1-x,0)\}$

Substituindo (3.4.140) para F e (3.4.146) para G em (3.4.147), produz

(3.4.148) $b\,G(1,0) + u\,G(0,1) = 0, \qquad u \in (0,1)$

Isto implica que

(3.4.149) $G(1,0) = G(0,1) = 0$

Portanto de (3.4.146), concluimos que

(3.4.150) $G(x,y) = ax^{\alpha}y^{\beta}$

para $(x,y) \in [0,1] \times [0,1] \qquad \alpha,\beta \neq 0$

Similarmente, de (3.4.133) e (3.4.140), temos

(3.4.151) $\sum_{j=1}^{m} H(y_j,v_j) = b \sum_{j=1}^{m} y_j^{\alpha}v_j^{\beta}$

Pelo mesmo método de antes, concluimos que

(3.4.152) $H(x,y) = bx^{\alpha}y^{\beta}$

para $(x,y) \in [0,1] \times [0,1] \quad \alpha,\beta \neq 0$

Por um método paralelo como em I(i), podemos obter os seguintes resultados.

I(ii) Quando $F(x,y) = abx^{\gamma}, \qquad \gamma \neq 0$

Então

(3.4.153) $G(x,y) = ax^{\gamma}, \quad H(x,y) = bx^{\gamma}$

para $(x,y) \in [0,1] \times [0,1]$, e $\gamma \neq 0$.

I(iii) Quando $F(x,y) = aby^{\gamma}, \qquad \gamma \neq 0$

Então

(3.4.154) $G(x,y) = ay^{\gamma}, \; H(x,y) = by^{\gamma}$

para $(x,y) \in [0,1] \times [0,1]$, e $\gamma \neq 0$.

Atraves da transformação em (3.4.125), as soluções que obtivemos em I(i), I(ii) e I(iii) são (3.4.116), (3.4.117) e (3.4.118) respectivamente.

Caso II. Quando $F(x,x) = G(x,x) = 0$, $H(x,x)$: arbitrária, $x \in [0,1]$

Seja $X = U = (1,0)$ em (3.4.126), com a ajuda de (3.4.128), temos

(3.4.155) $\sum_{j=1}^{m} F(y_j,v_j) = 0$

para $Y = (y_1,\ldots,y_m) \in \Delta_m$, $V = (v_1,\ldots v_m) \in \Delta_m$.

Seja $Y = (y_1,y_2,1-y_1-y_2)$, $V = (v_1,v_2,1-v_1-v_2) \in \Delta_3$ em (3.4.155), produz

(3.4.156) $F(y_1,v_1) + F(y_2,v_2) + F(1-y_1-y_2,1-v_1-v_2) = 0$

Novamente, seja $Y = (y_1+y_2,1-y_1-y_2)$, $V = (v_1+v_2,1-v_1-v_2)$ em (3.4.155), produz

(3.4.157) $F(y_1+y_2,v_1+v_2) + F(1-y_1-y_2,1-v_1-v_2) = 0$

Combinando (3.4.156) e (3.4.157), é fácil ver que F é aditiva em $[0,1] \times [0,1]$, portanto como antes, temos a relação

(3.4.158) $F(y,v) = y\, F(1,0) + v\, F(0,1)$

Com a condição que

$F(0,0) = F(1,1) = 0$

É trivial ver de (3.4.155) que

(3.4.159) $F(1,0) = F(0,1) = 0$

Portanto

(3.4.160) $F(y,v) = 0, \qquad (y,v) \in [0,1] \times [0,1]$

Seja $Y = V = (y_1,\ldots,y_m)$ em (3.4.126), com (3.4.160), é fácil ver que a função G também anula-se em todo parte, e portanto a função H poderia ser arbitrãia em $[0,1] \times [0,1]$.

Portanto, correspondente as soluções desenvolvidas a qual é dado em (3.4.119).

Caso III. Quando $F(x,x) = H(x,x) = 0$, $G(x,x)$: arbitrária $x \in [0,1]$

A prova é paralela ao Caso II feito acima, e portanto tem as soluções como dado em (3.4.120).

Isto prova o Teorema 3.4.2.

Agora, usando o Teorema 3.4.1, provamos um lema de caracterização para algumas medidas de informação generalizada.

Lema 3.4.3 Se $I(P:Q)$ em $\Delta_n \times \Delta'_n$, $n \geqq 1$, satisfazendo os seguintes postulados:

(a) Postulado da soma

(3.4.161) $I(P:Q) = \sum_{i=1}^{n} f(p_i,q_i)$

para todo $P = (p_1,\ldots,p_n) \in \Delta_n$, $Q = (q_1,\ldots,q_n) \in \Delta'_n$, $n = 2,3$; e f é mensurável em cada uma de suas variáveis.

(b) Postulado de não-aditividade

(3.4.162) $I(P*R:Q*S)$

$= I(P:Q) + I(R:S) + (2^{-\beta}-1)\, I(P:Q)I(R:S)$

para todo $P = (p_1,\ldots,p_n) \in \Delta_n$, $Q = (q_1,\ldots,q_n) \in \Delta'_n$, $R = (r_1,\ldots,r_m) \in \Delta_m$, $S = (s_1,\ldots,s_m) \in \Delta'_m$, $n = 2,3$; $m = 2,3$.

(c) Postulado de normalização

(3.4.163) (i) $f(1,1/2) = 1$, $f(0,1) = 0$

(3.4.164) (ii) $f(1,1/2) = 1$, $f(0,1) = 0$, $f(1/2,1/2) = 0$

(3.4.165) (iii) $f(1,1/2) = 1$, $f(0,1) = 0$, $f(1/2,1/2) = 1/2$

então $I(P:Q)$ é

(i) A divergência dirigida não-aditiva de ordem (α,β)

(ii) a divergência dirigida de ordem α $(\alpha \neq 1)$

(iii) a imprecisão de ordem $1+\beta$

respectivamente.

Prova:

Dos postulados (a) e (b), temos a seguinte equação funcional

$$(3.4.166) \quad \sum_{i=1}^{n} \sum_{j=1}^{m} f(p_i r_j, q_i s_j)$$

$$= \sum_{i=1}^{n} f(p_i, q_i) + \sum_{j=1}^{m} f(r_j, s_j) + (2^{-\beta} - 1) \sum_{i=1}^{n} \sum_{j=1}^{m} f(p_i, q_i) f(r_j, s_j)$$

para $P = (p_1, \ldots, p_n) \in \Delta_n$, $Q = (q_1, \ldots, q_n) \in \Delta'_n$, $R = (r_1, \ldots, r_m) \in \Delta_m$, $S = (s_1, \ldots, s_m) \in \Delta'_m$, $n = 2,3$; $m = 2,3$ e $\beta \neq 0$.

Esta é a equação funcional (3.4.1) quando $\lambda = 2^{-\beta} - 1$. Portanto pelo Teorema 3.4.1, as soluções são

$$(3.4.167) \quad f(x,y) = \frac{x^{\alpha} y^{\delta} - x}{2^{-\beta} - 1}$$

$$(3.4.168) \quad f(x,y) = \frac{x^{\gamma} - x}{2^{-\beta} - 1}$$

$$(3.4.169) \quad f(x,y) = \frac{y^{\gamma} - x}{2^{-\beta} - 1}$$

A condição de normalização $f(1,1/2) = 1$ elemina a possibilidade da solução em (3.4.168), também da a relação

$$(3.4.170) \quad \delta = \beta$$

em (3.4.167) e

$$(3.4.171) \quad \gamma = \beta$$

em (3.4.169)

Mas com a condição de normalização $f(0,1) = 0$, a solução em (3.4.169) é também eleminada, e a única solução que existe é

$$(3.4.172)\quad f(x,y) = \frac{x^{\alpha}y^{\beta} - x}{2^{-\beta} - 1}, \qquad \beta \neq 0$$

Novamente, a condição de normalização $f(1/2,1/2) = 0$ dá

$$(3.4.173)\quad \alpha + \beta = 1$$

Neste caso, (3.4.172) torna-se

$$(3.4.174)\quad f(x,y) = \frac{x^{\alpha}y^{1-\alpha} - x}{2^{\alpha-1} - 1}$$

A condição de normalização $f(1/2,1/2) = 1/2$ em (3.4.172) produzira

$$(3.4.175)\quad \alpha = 1$$

E portanto, (3.4.172) dá

$$(3.4.176)\quad f(x,y) = \frac{xy^{\beta} - x}{2^{-\beta} - 1}, \qquad \beta \neq 0$$

Pelo postulado da soma em (a), vemos que (3.4.172), (3.4.174) e (3.4.176) produzem várias medidas de informação generalizadas (a), (b) e (c) no teorema respectivamente.

Isto prova o lema.

CAPÍTULO 4

A DIVERGÊNCIA-J DE ORDEM α

RESUMO

A divergência-J de ordem α é definida e suas propriedades são listadas. Um teorema de caracterização é também provado assumindo-se um conjunto de cinco postulados.

4.1 INTRODUÇÃO

Denotemos por $\Delta_n = \{(p_1,\dots,p_n): p_i \geq 0,\ 1 \leq i \leq n,\ \sum_{i=1}^{n} p_i = 1\}$ o conjunto de distribuiçõens de probabilidade discretos. Para duas dessas distribuições $P=(p_1,\dots,p_n)$ e $Q=(q_1,\dots,q_n)$ em Δ_n, a divergência dirigida de ordem α [Rathie e Kannappan (1972)] é definida como

$$(4.1.1)\qquad I_{n,\alpha}(P:Q) = \frac{\sum_{i=1}^{n} p_i^{\alpha} q_i^{1-\alpha} - 1}{2^{\alpha-1}-1}, \qquad \alpha \neq 1$$

Entre duas distribuições de probabilidade $P=(p_1,\dots,p_n)$ e $Q=(q_1,\dots,q_n)$ em Δ_n, existem diversas medidas estatísticas as quais estão relacionadas com divergência dirigida.

A afinidade [Matusita (1961)] é dada por

$$A_n(p_1,\dots,p_n:q_1,\dots,q_n) = \sum_{i=1}^{n} (p_i q_i)^{1/2}$$

A medida de distância quadratica [Matusita (1966)] é dada por

$$D_n(p_1,\dots,p_n:q_1,\dots,q_n) = \sum_{i=1}^{n} (p_i^{1/2} - q_i^{1/2})^2 = 2\{1 - \sum_{i=1}^{n} (p_i q_i)^{1/2}\}$$

E a estatística χ^2 de Pearson é definida por

$$\chi^2 = \sum_{i=1}^{k} \{(n_i - np_i)^2/(np_i)\} = n\,\{\sum_{i=1}^{k} q_i^2 p_i^{-1} - 1\}$$

onde $n = \sum_{i=1}^{k} n_i$, n_i e np_i são as frequências observadas e esperadas correspondentes as i-ésimo grupo respectivamente e p_i é a probabilidade de obtermos uma observações no i-ésimo grupo, e $q_i = n_i/n$.

É claro que

$$I_{n,1/2}(P:Q) = \{A_n(P:Q) - 1\}/(2^{-1/2} - 1)$$

$$I_{n,1/2}(P:Q) = D_n/(2-2^{1/2})$$

e

$$I_{k,2}(q_1,\ldots,q_k:p_1,\ldots,p_k) = \frac{\chi^2}{n}$$

Aqui usamos a convenção $0^0=1$, $0^\alpha=0$ $(\alpha\neq 0)$

Quando n=2, (4.1.1) produz

(4.1.2) $$I_{2,\alpha}(P:Q) = \frac{p^\alpha q^{1-\alpha} + (1-p)^\alpha (1-q)^{1-\alpha} - 1}{2^{\alpha-1} - 1}, \qquad \alpha\neq 1$$

Para $\alpha\to 1$, (4.1.1) reduz-se a

(4.1.3) $$I_n(P:Q) = \sum_{i=1}^{n} p_i \log (p_i/q_i)$$

onde $0 \cdot \log 0 = 0$ e sepre que um $q_i=0$, o correspondente $p_i=0$.

$I_n(P:Q)$ é a medida de divergência dirigida [Kullback (1959)] ou ganho de informação [Renyi (1960)]

A divergência-J (ou divergência simetrica) entre duas distribuições P e Q é definida por [Kullback (1959)]

(4.1.4) $$J_n(P:Q) = \sum_{i=1}^{n} (p_i - q_i) \log (p_i/q_i) = J_n(P:Q) + J_n(Q:P)$$

Nós definimos a divergência-J (ou divergência simetrica) de ordem α entre duas distribuições $P = (p_1,\ldots,p_n)$ e $Q = (q_1,\ldots,q_n)$ em Δ_n

$$(4.1.5) \quad J_{n,\alpha}(P:Q) = I_{n,\alpha}(P:Q) + I_{n,\alpha}(Q:P)$$

$$= \frac{\sum_{i=1}^{n} p_i^{\alpha} q_i^{1-\alpha} + \sum_{i=1}^{n} q_i^{\alpha} p_i^{1-\alpha} - 2}{2^{\alpha-1} - 1} \qquad \alpha \neq 1$$

Para $\alpha \to 1$, (4.1.5) fornece (4.1.4). Portanto (4.1.5) é uma generalização a um parametro de (4.1.4)

Para n=2, (4.1.5) fornece a forma sequente

$$(4.1.6) \quad J_{2,\alpha}(P:Q) = \frac{p^{\alpha}q^{1-\alpha} + (1-p)^{\alpha}(1-q)^{1-\alpha} + q^{\alpha}p^{1-\alpha} + (1-q)^{\alpha}(1-p)^{1-\alpha} - 2}{2^{\alpha-1} - 1}$$

$$\alpha \neq 1.$$

Se H_1, H_2 são as hipoteses que a variavel aleatória X é de uma população estatística com distribuição de probabilidade P e Q repectivamente, para o j-ésimo evento, nós definimos $\frac{(p_j/q_j)^{\alpha-1} - 1}{2^{\alpha-1} - 1}$ como a informação do evento para discriminação em favor de H_1 contra H_2, então $I_{n,\alpha}(P:Q)$ é a informação média para discriminação em favor de H_1 contra H_2 e $J_{n,\alpha}(P:Q)$ é a informação média para discriminação entre H_1 e H_2.

Um teorema de caracterização para (4.1.4) foi dado recentemente por Kannappan e Rathie (1979)

O objetivo desse trabalho é dar algumas propriedades e um teorema de caracterização para (4.1.5) para $\alpha \neq 1$, usando um conjunto de cinco postulados. esta é a primeira tentativa para dar uma fundamentação axiomática para a divergência-J de ordem α.

4.2 PROPRIEDADES

A divergência-J de ordem α tem as seguintes propriedades:

(i) Não-Negatividade

$J_{n,\alpha}(P:Q)$ é não-negativa, e é zero se e somente se $p_i = q_i$ para totos i. Isto é claro da desigualidade

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

Isto significa que a informação média para discriminação entre duas distribuições diferentes é geralmente positivo e será zero se e sómente se as duas distribuições coincidirem.

(ii) Simetria

$$J_{n,\alpha}(p_1,\dots,p_n:q_1,\dots,q_n) = J_{n,\alpha}(p_{a(1)},\dots,p_{a(n)}:q_{a(1)},\dots,q_{a(n)})$$

onde $\{ a(1),\dots,a(n) \}$ é uma permutação arbitraria de $\{1,\dots,n\}$.

A permutação conjunta de indices correspondentes de eventos não muda o valor da divergência-J de ordem α.

(iii) Expandibilidade

$$J_{n+1,\alpha}(p_1,\dots,p_n,0:q_1,\dots,q_n,0) = J_{n,\alpha}(p_1,\dots,p_n:q_1,\dots,q_n)$$

A adição de um evento de probabilidade zero não altera o valor da divergência-J de ordem α.

(iv) Recursividade

$$J_{n,\alpha}(p_1,\dots,p_n:q_1,\dots,q_n) = J_{n-1,\alpha}(p_1+p_2,p_3,\dots,p_n:q_1+q_2,q_3,\dots,q_n)$$

$$+ (p_1+p_2)^{\alpha}(q_1+q_2)^{1-\alpha}I_{2,\alpha}\left(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2} : \frac{q_1}{q_1+q_2}, \frac{q_2}{q_1+q_2} \right)$$

$$+ (q_1+q_2)^{\alpha}(p_1+p_2)^{1-\alpha}I_{2,\alpha}\left(\frac{q_1}{q_1+q_2}, \frac{q_2}{q_1+q_2} : \frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2} \right)$$

para $p_1+p_2 > 0$, $q_1+q_2 > 0$.

(v) Não-Aditividade

Para P, Q ε Δ_n e R, S ε Δ_m

$$J_{mn,\alpha}(P*R : Q*S)$$

$$= J_{n,\alpha}(P:Q) + J_{m,\alpha}(R:S) + \frac{1}{2^{\alpha-1}-1} \{ I_{n,\alpha}(P:Q)\cdot I_{m,\alpha}(R:S) +$$

$$I_{n,\alpha}(Q:P)\cdot I_{m,\alpha}(S:R) \}$$

(vi) Não-Aditividade Forte

Para $(p_1,\ldots,p_n)$, $(q_1,\ldots,q_n)$ em Δ_n e $(P_{1i},\ldots,P_{mi})$, $(Q_{1i},\ldots,Q_{mi})$ em Δ_m para i=1,...,n.

$$J_{mn,\alpha}(p_1P_{11},p_1P_{21},\ldots,p_1P_{m1},\ldots,p_nP_{1n},\ldots,p_nP_{mn}: q_1Q_{11},q_1Q_{21},\ldots,$$

$$q_1Q_{m1},\ldots,q_nQ_{1n},\ldots,q_nQ_{mn})$$

$$= J_{n,\alpha}(p_1,\ldots,p_n: q_1,\ldots,q_n) + \sum_{i=1}^{n} q_i^{\alpha}p_i^{1-\alpha} I_{m,\alpha}(Q_{1i},\ldots,Q_{mi}:P_{1i},\ldots,P_{mi})$$

$$\sum_{i=1}^{n} p_i^{\alpha}q_i^{1-\alpha} I_{m,\alpha}(P_{1i},\ldots,P_{mi}: Q_{1i},\ldots,Q_{mi})$$

(vii) Continuidade

$J_{n,\alpha}(P:Q)$ é uma função continua de 2n variaveis.

(viii) Intercâmbiobilidade de distribuições

Para $P = (p_1,\ldots,p_n)$, $Q = (q_1,\ldots,q_n)$ em Δ_n

$$J_{n,\alpha}(P:Q) = J_{n,\alpha}(Q:P)$$

(ix) Nulidade

Seja $g(x,y) = J_{2,\alpha}(x,1-x:y,1-y)$, onde x,y ε $[0,1]$.

$g(x,x) = 0$ para x ε $[0,1]$

A divergência-J de um evento e seu complemento entre duas distribuições de probabilidade identicas é zero.

(x) Normalização

$$g(0,1) = 2(1-2^{\alpha-1})^{-1}$$

4.3 O TEOREMA DE CARACTERIZAÇÃO

Seja $f(x,y) = I_{2,\alpha}(x,1-x:y,1-y)$ para $x,y \in K$, onde

$K = (0,1)\times(0,1)\cup\{(0,z)\}\cup\{(1,z_1)\}$ com $z \in [0,1)$ e $z_1 \in (0,1]$

Nesta secção, usamos quatro das propriedades acima e a existencia de derivadas parciais atē segunda ordem de $f(x,y)$ com respeito a $x,y \in (0,1)$ como postulados para provar um teorema de caracterização para a divergência-J de ordem α.

<u>Teorema 4.3.1</u> As seguintes propriedades

a. Simetria (ii) quando $n = 3$.

b. Recursividade (iv).

c. Nulidade (ix)

d. Normalização (x)

e. A existencia de derivadas parciais atē segunda ordem de $f(x,y)$ com respeito a $x,y \in (0,1)$

determina de maneira unica a divergência-J de ordem α para $\alpha \neq 1$.

Prova:

Quando $n = 3$, seja $X = (x,u,1-x-u)$, $Y = (y,v,1-y-v)$.

Da recursividade nós podemos obter o seguinte

(4.3.1) $$J_{3,\alpha}(x,u,1-x-u:y,v,1-y-v) = J_{2,\alpha}(x+u,1-x-u:y+v,1-y-v) +$$

$$(x+u)^{\alpha}(y+v)^{1-\alpha}I_{2,\alpha}\left(\frac{x}{x+u}, \frac{u}{x+u} : \frac{y}{y+v}, \frac{v}{y+v}\right) +$$

$$(y+v)^{\alpha}(x+u)^{1-\alpha}I_{2,\alpha}\left(\frac{y}{y+v}, \frac{v}{y+v} : \frac{x}{x+u}, \frac{u}{x+u}\right)$$

(4.3.2) $$J_{3,\alpha}(u,1-x-u,x:v,1-y-v,y) = J_{2,\alpha}(1-x,x:1-y,y) +$$

$$(1-x)^{\alpha}(1-y)^{1-\alpha}I_{2,\alpha}\left(\frac{u}{1-x}, \frac{1-x-u}{1-x} : \frac{v}{1-y}, \frac{1-y-v}{1-y}\right) +$$

$$(1-y)^{\alpha}(1-x)^{1-\alpha}I_{2,\alpha}\left(\frac{v}{1-y}, \frac{1-y-v}{1-y} : \frac{u}{1-x}, \frac{1-x-u}{1-x}\right)$$

(4.3.3) $\quad J_{3,\alpha}(x,1-x-u,u:y,1-y-v,v) = J_{2,\alpha}(1-u,u:1-v,v) +$

$$(1-u)^{\alpha}(1-v)^{1-\alpha}I_{2,\alpha}\left(\frac{x}{1-u},\frac{1-x-u}{1-u} : \frac{y}{1-v},\frac{1-y-v}{1-v}\right) +$$

$$(1-v)^{\alpha}(1-u)^{1-\alpha}I_{2,\alpha}\left(\frac{y}{1-v},\frac{1-y-v}{1-v} : \frac{x}{1-u},\frac{1-x-u}{1-u}\right)$$

De simitria quando $n=3$, nós sabemos que (4.3.1), (4.3.2) e (4.3.3) são equivalentes, usando a notação de f e g como definida antes, temos a equação funcional

(4.3.4) $\quad g(x,y) + (1-x)^{\alpha}(1-y)^{1-\alpha}f\left(\frac{u}{1-x},\frac{v}{1-y}\right) + (1-y)^{\alpha}(1-x)^{1-\alpha}f\left(\frac{v}{1-y},\frac{u}{1-x}\right)$

$$= g(u,v) + (1-u)^{\alpha}(1-v)^{1-\alpha}f\left(\frac{x}{1-u},\frac{y}{1-v}\right) + (1-v)^{\alpha}(1-u)^{1-\alpha}f\left(\frac{y}{1-v},\frac{x}{1-u}\right)$$

para $x,y,u,v \in [0,1)$, $x+u$, $y+v \in [0,1]$.

e

(4.3.5) $\quad$ O lado direito de (4.3.4)

$$= g(x+u,y+v) + (x+u)^{\alpha}(y+v)^{1-\alpha}f\left(\frac{x}{x+u},\frac{y}{y+v}\right) + f\left(\frac{y}{y+v},\frac{x}{x+u}\right)$$

para $x,y,u,v \in [0,1)$ e $x+u$, $y+v \in (0,1]$.

Para $u = v = 0$, (4.3.4) produz

(4.3.6) $\quad g(x,y) = f(x,y) + f(y,x) - \{(1-x)^{\alpha}(1-y)^{1-\alpha}+(1-y)^{\alpha}(1-x)^{1-\alpha}\}f(0,0)$

$$+ g(0,0)$$

Desde que f tem derivadas parciais até segunda ordem, o mesmo vale para g.

A prova do teorema é basiada no seguinte lema.

<u>Lema 4.3.1</u> A solução geral do equação funcional (4.3.4) $(\alpha\neq 1)$ quando $f(x,y)$ tem derivadas parciais até segunda ordem com respeito a $x,y \in K$, é dada por

Caso I. Quando $\alpha \neq 1/2$

(a) Quando $\alpha \neq 0$

(4.3.7) $$g(x,y) = \{A-f(0,0)\}\{(1-x)^{\alpha}(1-y)^{1-\alpha}+(1-y)^{\alpha}(1-x)^{1-\alpha}\} + B\{x^{\alpha}y^{1-\alpha}+y^{\alpha}x^{1-\alpha}\}+C+g(0,0)$$

(b) Quando $\alpha = 0$

(4.3.8) $$g(x,y) = A\{(y-1)\log(1-x)+y\cdot\log x+(x-1)\log(1-y)+x\cdot\log y\}+B\{y\cdot\log y +(1-y)\log(1-y)+x\cdot\log x+(1-x)\log(1-x)\}+(x+y)\{C+f(0,0)\}+g(0,0)$$

Caso II. Guando $\alpha = 1/2$

(4.3.9) $$g(x,y) = Ax^{1/2}y^{1/2}-B\{x^{1/2}y^{1/2}+(1-x)^{1/2}(1-y)^{1/2}-1\}+ +2\{1-(1-x)^{1/2}(1-y)^{1/2}\}f(0,0)+g(0,0)$$

Prova:

Para y e v fixados, a equação funcional (4.3.4) produz

(4.3.10) $$A(x)+(1-x)^{\alpha}F\left(\frac{u}{1-x}\right)+(1-x)^{1-\alpha}H\left(\frac{u}{1-x}\right)$$

$$=B(u)+(1-u)^{\alpha}L\left(\frac{x}{1-u}\right)+(1-u)^{1-\alpha}M\left(\frac{x}{1-u}\right)$$

onde

(4.3.11) $$\begin{cases} A(x)=g(x,y) \\ F(x)=(1-y)^{1-\alpha}f\left(x,\frac{v}{1-y}\right) \\ H(x)=(1-y)^{\alpha}f\left(\frac{v}{1-y},x\right) \end{cases} \qquad \begin{cases} B(x)=g(x,v) \\ L(x)=(1-v)^{1-\alpha}f\left(x,\frac{y}{1-v}\right) \\ M(x)=(1-v)^{\alpha}f\left(\frac{y}{1-v},x\right) \end{cases}$$

Diferenciamos (4.3.10) primeiro com respeito a x, nós temos

(4.3.12) $$A'(x)+u(1-x)^{\alpha-2}F'\left(\frac{u}{1-x}\right)-\alpha(1-x)^{\alpha-1}F\left(\frac{u}{1-x}\right)+u(1-x)^{-\alpha-1}H'\left(\frac{u}{1-x}\right)$$

$$-(1-\alpha)(1-x)^{-\alpha}H\left(\frac{u}{1-x}\right)$$

$$=(1-u)^{\alpha-1}L'\left(\frac{x}{1-u}\right)+(1-u)^{-\alpha}M'\left(\frac{x}{1-u}\right)$$

Então diferenciamos (4.3.12) novamente com respeito a u, temos

(4.3.13) $$(1-\alpha)(1-x)^{\alpha-2}F'\left(\frac{u}{1-x}\right)+u(1-x)^{\alpha-3}F''\left(\frac{u}{1-x}\right)+\alpha(1-x)^{-\alpha-1}H'\left(\frac{u}{1-x}\right)$$

$$+u(1-x)^{-\alpha-1}H''\left(\frac{u}{1-x}\right)$$

$$=(1-\alpha)(1-u)^{\alpha-2}L'\left(\frac{x}{1-u}\right)+x(1-u)^{\alpha-3}L''\left(\frac{x}{1-u}\right)+\alpha(1-u)^{-\alpha-1}M'\left(\frac{x}{1-u}\right)$$

$$+x(1-u)^{-\alpha-2}M''\left(\frac{x}{1-u}\right)$$

Agora, seja $t=\frac{u}{1-x}$, $s=\frac{x}{1-u}$, então

$$x=\frac{s(t-1)}{st-1}\quad,\qquad 1-x=\frac{s-1}{st-1}$$

$$u=\frac{t(s-1)}{st-1}\quad,\qquad 1-u=\frac{t-1}{st-1}$$

e (4.3.13) torna-se

(4.3.14) $$(1-\alpha)\left(\frac{1-s}{1-st}\right)^{\alpha-2}F'(t)+t\left(\frac{1-s}{1-st}\right)^{\alpha-2}F''(t)+\alpha\left(\frac{1-s}{1-st}\right)^{-\alpha-1}H'(t)$$

$$+t\left(\frac{1-s}{1-st}\right)^{-\alpha-1}H''(t)$$

$$=(1-\alpha)\left(\frac{1-t}{1-st}\right)^{\alpha-2}L'(s)+s\left(\frac{1-t}{1-st}\right)^{\alpha-2}L''(s)+\alpha\left(\frac{1-t}{1-st}\right)^{-\alpha-1}M'(s)$$

$$+s\left(\frac{1-t}{1-st}\right)^{-\alpha-1}M''(s)$$

Multiplicando (4.3.14) por $\left\{\frac{1-st}{(1-s)(1-t)}\right\}^{\alpha-2}$

(4.3.15) $$(1-t)^{2-\alpha}\{(1-\alpha)F'(t)+tF''(t)\}+\left(\frac{1-s}{1-st}\right)^{1-2\alpha}(1-t)^{2-\alpha}\{\alpha H'(t)+tH''(t)\}$$

$$=(1-s)^{2-\alpha}\{(1-\alpha)L'(s)+sL''(s)\}+\left(\frac{1-t}{1-st}\right)^{1-2\alpha}(1-s)^{2-\alpha}\{\alpha M'(s)+sM''(s)\}$$

Podemos reescrever (4.3.15) na seguinte forma

(4.3.16) $$\{\,p(t)-q(s)\,\}\left\{\frac{1-st}{(1-s)(1-t)}\right\}^{1-2\alpha}=m(s)-l(t)$$

onde

$$(4.3.17)\quad \begin{cases} p(t) = (1-t)^{2-\alpha}\{(1-\alpha)F'(t)+tF''(t)\} \\ q(s) = (1-s)^{2-\alpha}\{(1-\alpha)L'(s)+sL''(s)\} \\ l(t) = (1-t)^{\alpha+1}\{\alpha H'(t)+tH''(t)\} \\ m(s) = (1-s)^{\alpha+1}\{\alpha M'(s)+sM''(s)\} \end{cases}$$

Caso I. Quando $\alpha \neq 1/2$

(4.3.16) e (4.3.17) implicam que

$p(t)-q(s) = m(s)-l(t) = 0$

Isto é

$$(4.3.18)\quad (1-t)^{2-\alpha}\{(1-\alpha)F'(t)+tF''(t)\} = (1-s)^{2-\alpha}\{(1-\alpha)L'(s)+sL''(s)\} = a_1$$

onde a_1 é uma constante

e

$$(4.3.19)\quad (1-s)^{\alpha+1}\{\alpha M'(s)+sM''(s)\} = (1-t)^{\alpha+1}\{\alpha H'(t)+tH''(t)\}= b_1$$

onde b_1 é uma constante

Resolvendo a equação funcional (4.3.18), temos

$$(4.3.20)\quad F(t) = (1-y)^{1-\alpha}f(t, \frac{v}{1-y})$$

$$= -\frac{a_1}{\alpha(1-\alpha)}(1-t)^{\alpha}+\frac{a_2}{\alpha}t^{\alpha}+a_3\ , \qquad \text{se } \alpha \neq 0$$

$$= a_1\log(1-t) + a_2'\log t + a_3'\ , \qquad \text{se } \alpha = 0$$

onde a_i, a_i', $1\leq i\leq 3$, são funções de y e v.

similarmente, a equação diferencial (4.3.19) produz

$$(4.3.21)\quad M(s) = (1-v)^{\alpha}f(\frac{y}{1-v}, s)$$

$$= -\frac{b_1}{\alpha(1-\alpha)}(1-s)^{1-\alpha}+\frac{b_2}{1-\alpha}s^{1-\alpha}+b_3\ , \qquad \text{se } \alpha \neq 0$$

$$= b_1\{s\log s + (1-s)\log(1-s)\}+b_2' s + b_3'\ , \qquad \text{se } \alpha = 0$$

onde b_i, b_i', $1\leq i\leq 3$, são funções de y e v.

Caso I(a). Quando $\alpha \neq 0$

(4.3.20) implica que

$$(4.3.22) \quad f(t, \frac{v}{1-y}) = -\frac{a_1}{\alpha(1-\alpha)}(1-y)^{\alpha-1}(1-t)^{\alpha} + \frac{a_2}{\alpha}(1-y)^{\alpha-1}t^{\alpha} + a_3(1-y)^{\alpha-1}$$

Seja $s = \frac{v}{1-y}$, a parte a esquerda é uma função de t e s, portanto, podemos escrever

$$(4.3.23) \quad f(t,s) = \alpha_1(s)(1-t)^{\alpha} + \alpha_2(s)\, t^{\alpha} + \alpha_3(s), \qquad s,t \in (0,1)$$

E similarmente, (4.3.21) implica que

$$(4.3.24) \quad f(t,s) = \beta_1(t)(1-s)^{1-\alpha} + \beta_2(t)\, s^{1-\alpha} + \beta_3(t), \qquad s,t \in (0,1)$$

onde α_i e β_i, $1 \leq i \leq 3$, são funções arbitrarias de s e t respectivamente.

Combinando (4.3.23) e (4.3.24) obtemos

$$(4.3.25) \quad \alpha_1(s)(1-t)^{\alpha} + \alpha_2(s)\, t^{\alpha} + \alpha_3(s) \equiv \beta_1(t)(1-s)^{1-\alpha} + \beta_2(t)\, s^{1-\alpha} + \beta_3(t)$$

Diferenciando (4.3.25) com respeito a t, temos

$$(4.3.26) \quad \alpha\{-\alpha_1(s)(1-t)^{\alpha-1} + \alpha_2(s)\, t^{\alpha-1}\} \equiv \beta_1'(t)(1-s)^{1-\alpha} + \beta_2'(t)\, s^{1-\alpha} + \beta_3'(t)$$

Desde que (4.3.26) é verdade para todo $s,t \in (0,1)$, seja $t = 1/2$, (4.3.26) torna-se

$$(4.3.27) \quad \alpha\{-\alpha_1(s)\cdot 2^{1-\alpha} + \alpha_2(s)\cdot 2^{1-\alpha}\} \equiv \beta_1'(\tfrac{1}{2})(1-s)^{1-\alpha} + \beta_2'(\tfrac{1}{2})\, s^{1-\alpha} + \beta_3'(\tfrac{1}{2})$$

Seja $t = 1/3$, então (4.3.26) torna-se

$$(4.3.28) \quad \alpha\{-\alpha_1(s)\cdot 2^{\alpha-1}\cdot 3^{1-\alpha} + \alpha_2(s)\cdot 3^{1-\alpha}\} \equiv \beta_1'(\tfrac{1}{3})(1-s)^{1-\alpha} + \beta_2'(\tfrac{1}{3})\, s^{1-\alpha} + \beta_3'(\tfrac{1}{3})$$

Combinando (4.3.27) e (4.3.28), e resolvendo para $\alpha_1(s)$ e $\alpha_2(s)$ respectivamente, temos

$$(4.3.29) \quad \alpha_1(s) = \delta_1(1-s)^{1-\alpha} + \delta_2 s^{1-\alpha} + \delta_3$$

e

$$(4.3.30) \quad \alpha_2(s) = \gamma_1(1-s)^{1-\alpha} + \gamma_2 s^{1-\alpha} + \gamma_3$$

onde δ_i e γ_i, $1 \leq i \leq 3$, são constantes arbitrarias.

Substituindo (4.3.29) e (4.3.30) para $\alpha_1(s)$ e $\alpha_2(s)$ respectivamente em (4.3.26), então

(4.3.31) $\beta_1'(t)(1-s)^{1-\alpha}+ \beta_2'(t)\ s^{1-\alpha}+ \beta_3'(t)$

$$= \alpha\{-[\alpha_1(1-s)^{1-\alpha}+ \delta_2 s^{1-\alpha}+ \delta_3](1-t)^{\alpha-1}+[\gamma_1(1-s)^{1-\alpha}+ \gamma_2 s^{1-\alpha}+ \gamma_3]\ t^{\alpha-1}\}$$

Comparando os coificientes de $(1-s)^{1-\alpha}$, $s^{1-\alpha}$ e constantes respectivamente, é claro que

(4.3.32) $\beta_i'(t) = \alpha\{\gamma_i t^{\alpha-1}- \delta_i(1-t)^{\alpha-1}\}$, $\qquad i = 1,2,3$

Isto da

(4.3.33) $\beta_i(t) = \gamma_i t^{\alpha}+ \delta_i(1-t)^{\alpha}+ \xi_i$, $\qquad t \in (0,1)$, $\quad i = 1,2,3$

onde γ_i, δ_i e ξ_i para $1\leq i\leq 3$ são constantes arbitrarias. Portanto (4.3.24) produz

(4.3.34) $f(t,s) = \{\gamma_1 t^{\alpha}+ \delta_1(1-t)^{\alpha}+ \xi_1\}(1-s)^{1-\alpha}+ \{\gamma_2 t^{\alpha}+ \delta_2(1-t)^{\alpha}+ \xi_2\}\cdot s^{1-\alpha}$

$$+\{\gamma_3 t^{\alpha}+ \delta_3(1-t)^{\alpha}+ \xi_i\}\ , \qquad s,t \in (0.1)$$

Substituindo a expressão (4.3.34) para f em (4.3.6), portanto

(4.3.35) $g(t,s) = \{\gamma_1 t^{\alpha}+[\delta_1-f(0,0)](1-t)^{\alpha}+ \xi_1\}(1-s)^{1-\alpha}+ \{\gamma_2 t^{\alpha}+ \delta_2(1-t)^{\alpha}+$

$$\xi_2\}\cdot s^{1-\alpha}+\{\gamma_3 t^{\alpha}+ \delta_3(1-t)^{\alpha}+ \xi_3\}+\{\gamma_1 s^{\alpha}+[\delta_1-f(0,0)](1-s)^{\alpha}\ \xi_i\}\cdot$$

$$(1-t)^{1-\alpha}+\{\gamma_2 s^{\alpha}+ \delta_2(1-s)^{\alpha}+ \xi_2\}\cdot t^{1-\alpha}+\{\gamma_3 s^{\alpha}+ \delta_3(1-s)^{\alpha}+\xi_3\}+g(0,0)$$

Agora substituindo as expressões (4.3.34) para f e (4.3.35) para g respectivamente em (4.3.4), torna-se

(4.3.36) $\{\gamma_1 x^{\alpha}+[\delta_1-f(0,0)](1-x)^{\alpha}+ \xi_1\}(1-y)^{1-\alpha}+\{\gamma_2 x^{\alpha}+ \delta_2(1-x)^{\alpha}+ \xi_2\}\cdot y^{1-\alpha}+$

$$\{\gamma_3 x^{\alpha}+ \delta_3(1-x)^{\alpha}+ \xi_3\}+\{\gamma_1 y^{\alpha}+[\delta_1-f(0,0)](1-y)^{\alpha}+ \xi_1\}(1-x)^{1-\alpha}+$$

$$\{\gamma_2 y^{\alpha}+ \delta_2(1-y)^{\alpha}+ \xi_2\}\cdot x^{1-\alpha} +\{\gamma_3 y^{\alpha}+ \delta_3(1-y)^{\alpha}+ \xi_3\} + g(0,0) +$$

$$(1-x)^{\alpha}(1-y)^{1-\alpha}\{[\gamma_1(\frac{u}{1-x})^{\alpha}+\delta_1(1-\frac{u}{1-x})^{\alpha}+\xi_1](1-\frac{v}{1-y})^{1-\alpha}+$$

$$[\gamma_2(\frac{u}{1-x})^{\alpha}+\delta_2(1-\frac{u}{1-x})^{\alpha}+\xi_2](\frac{v}{1-y})^{1-\alpha}+[\gamma_3(\frac{u}{1-x})^{\alpha}+$$

$$\delta_3(1-\frac{u}{1-x})^{\alpha}+\xi_3]\}+(1-y)^{\alpha}(1-x)^{1-\alpha}\{[\gamma_1(\frac{v}{1-y})^{\alpha}+\delta_1(1-\frac{v}{1-y})^{\alpha}+$$

$$\xi_1](1-\frac{u}{1-x})^{1-\alpha}+[\gamma_2(\frac{v}{1-y})^{\alpha}+\delta_2(1-\frac{v}{1-y})^{\alpha}+\xi_2](\frac{u}{1-x})^{1-\alpha}+$$

$$[\gamma_3(\frac{v}{1-y})^{\alpha}+\delta_3(1-\frac{v}{1-y})^{\alpha}+\xi_3]\}$$

$$=\{\gamma_1u^{\alpha}+[\delta_1-f(0,0)](1-u)^{\alpha}+\xi_1\}(1-v)^{1-\alpha}+\{\gamma_2u^{\alpha}+\delta_2(1-u)^{\alpha}+\xi_2\}\cdot v^{1-\alpha}+$$

$$\{\gamma_3u^{\alpha}+\delta_3(1-u)^{\alpha}+\xi_3\}+\{\gamma_1-[\delta_1-f(0,0)](1-v)^{\alpha}+\xi_1\}(1-u)^{1-\alpha}+$$

$$\{\gamma_2v^{\alpha}+\delta_2(1-v)^{\alpha}+\xi_2\}\cdot u^{1-\alpha}+\{\gamma_3v^{\alpha}+\delta_3(1-v)^{\alpha}+\xi_3\}+g(0,0)+$$

$$(1-u)^{\alpha}(1-v)^{1-\alpha}\{[\gamma_1(\frac{x}{1-u})^{\alpha}+\delta_1(1-\frac{x}{1-u})^{\alpha}+\xi_1](1-\frac{y}{1-v})^{1-\alpha}+$$

$$[\gamma_2(\frac{x}{1-u})^{\alpha}+\delta_2(1-\frac{x}{1-u})^{\alpha}+\xi_2](\frac{y}{1-v})^{1-\alpha}+$$

$$[\gamma_3(\frac{x}{1-u})^{\alpha}+\delta_3(1-\frac{x}{1-u})^{\alpha}+\xi_3]\}+(1-v)^{\alpha}(1-u)^{1-\alpha}\{[\gamma_1(\frac{y}{1-v})^{\alpha}+$$

$$\delta_1(1-\frac{y}{1-v})^{\alpha}+\xi_1](1-\frac{x}{1-u})^{1-\alpha}+[\gamma_2(\frac{y}{1-v})^{\alpha}+\delta_2(1-\frac{y}{1-v})^{\alpha}+\xi_2]\cdot$$

$$(\frac{x}{1-u})^{1-\alpha}+[\gamma_3(\frac{y}{1-v})^{\alpha}+\delta_3(1-\frac{y}{1-v})^{\alpha}+\xi_3]\}$$

para $x,y,u,v \in [0,1)$ e $x+u,\ y+v \in [0,1]$.

Comparando os coificientes de $(1-y)^{1-\alpha}$ na equação acima, então

(4.3.37) $\{\gamma_1x^{\alpha}+[\delta_1-f(0,0)](1-x)^{\alpha}+\xi_1\}+\gamma_3u^{\alpha}+\delta_3(1-x-u)^{\alpha}+\xi_3(1-x)^{\alpha}=0$

Isto implica que

(4.3.38) $\gamma_1=\xi_1=\gamma_3=\delta_3=0$

Comparando os coificientes de $y^{1-\alpha}$ em (4.3.36), então

(4.3.39) $\gamma_2 x^\alpha + \delta_2(1-x)^\alpha + \xi_2 = \gamma_2 x^\alpha + \delta_2(1-x-u)^\alpha + \xi_2(1-u)^\alpha$

Isto implica que

(4.3.40) $\delta_2 = \xi_2 = 0$

Usando (4.3.38) e (4.3.40), assim f é reduzida na seguinte forma.

(4.3.41) $f(t,s) = \delta_1(1-t)^\alpha(1-s)^{1-\alpha} + \gamma_2 t^\alpha s^{1-\alpha} + \xi_3$

Substituindo (4.3.41) em (4.3.6), então g é de seguinte forma

(4.3.42) $g(x,y) = [\delta_1 - f(0,0)]\cdot\{(1-x)^\alpha(1-y)^{1-\alpha} + (1-y)^\alpha(1-x)^{1-\alpha}\}$

$$\gamma_2\{x^\alpha y^{1-\alpha} + y^\alpha x^{1-\alpha}\} + 2\xi_3 + g(0,0)$$

Tomando $\delta_1 = A$, $\gamma_2 = B$ e $2\xi_3 = C$, (4.3.42) é o mesmo que (4.3.7).

Caso I(b). Quando $\alpha = 0$

(4.3.20) implica que

$$(1-y)\cdot f(t, \frac{v}{1-y}) = a_1\log(1-t) + a_2'\log t + a_3'$$

Ou

(4.3.43) $f(t, \frac{v}{1-y}) = a_1(1-y)^{-1}\log(1-t) + a_2'(1-y)^{-1}\log t + a_3'(1-y)^{-1}$

Seja $s = \frac{v}{1-y}$, desde que o lado esquerdo de (4.3.43) é uma função de t e s somente, portanto podemos reescrever (4.3.43) com

(4.3.44) $f(t,s) = \alpha_1(s)\log(1-t) + \alpha_2(s)\log t + \alpha_3(s)$

E similarmente (4.3.21) implica que

(4.3.45) $f(t,s) = \beta_1(t)\{s\cdot\log s + (1-s)\log(1-s)\} + \beta_2(t)s + \beta_3(t)$

De (4.3.44) e (4.3.45), temos

(4.3.46) $\alpha_1(s)\log(1-t) + \alpha_2(s)\log t + \alpha_3(s) \equiv \beta_1(t)\cdot\{s\log s +$

$$(1-s)\log(1-s)\} + \beta_2(t)s + \beta_3(t)$$

Diferenciando (4.3.46) com respeito a t, temos

$$(4.3.47)\quad \frac{\alpha_1(s)}{t-1} + \frac{\alpha_2(s)}{t-1} \equiv \beta_1'(t)\{\ s\cdot\log s + (1-s)\log(1-s)\ \} + \beta_2'(t)s + \beta_3'(t)$$

$$s,t \in (0,1)$$

Novamente seja $t = 1/2$, então

$$(4.3.48)\quad -2\alpha_1(s) + 2\alpha_2(s)$$

$$\equiv \beta_1'(\tfrac{1}{2})\ \{\ s\cdot\log s + (1-s)\log(1-s)\} + \beta_2'(\tfrac{1}{2})s + \beta_3'(\tfrac{1}{2})$$

Substituindo $t = 1/3$ em (4.3.47)

$$(4.3.49)\quad -\frac{3}{2}\alpha_1(s) + 3\alpha_2(s)$$

$$\equiv \beta_1'(\tfrac{1}{3})\ \{\ s\log s + (1-s)\log(1-s)\} + \beta_2'(\tfrac{1}{3})s + \beta_3'(\tfrac{1}{3})$$

Resolvendo $\alpha_1(s)$ e $\alpha_2(s)$ pela combinação de (4.3.48) e (4.3.49), temos

$$(4.3.50)\quad \alpha_1(s) = \delta_1\{\ s\log s + (1-s)\log(1-s)\} + \delta_2 s + \delta_3$$

e

$$(4.3.51)\quad \alpha_2(s) = \gamma_1\{\ s\log s + (1-s)\log(1-s)\} + \gamma_2 s + \gamma_3$$

onde δ_i e γ_i, $1\leq i\leq 3$, são constantes arbitrarias.

Substituindo $\alpha_1(s)$ e $\alpha_2(s)$ de (4.3.50) e (4.3.51) respectivamente em (4.3.47) produz

$$(4.3.52)\quad \beta_1'(t)\ \{\ s\log s + (1-s)\log(1-s)\} + \beta_2'(t) + \beta_3'(t)$$

$$\equiv \{\frac{\delta_1}{t-1} + \frac{\gamma_1}{t}\}\{\ s\log s + (1-s)\log(1-s)\} + \{\frac{\delta_2}{t-1} + \frac{\gamma_2}{t}\}s + \frac{\delta_3}{t-1} + \frac{\gamma_3}{t}$$

Comparando os coificientes em (4.3.52), obtemos

$$(4.3.53)\quad \beta_i'(t) = \frac{\delta_i}{t-1} + \frac{\gamma_i}{t}\,, \qquad t \in (0,1), \quad 1<i<3$$

Ou

(4.3.54) $\beta_i(t) = \delta_i \log(1-t) + \gamma_i \log t + \xi_i$, $t \in (0,1)$, $1 \leq i \leq 3$

onde δ_i, γ_i e ξ_i, $1 \leq i \leq 3$, são constantes arbitrarias.

Substituindo a expressão (4.3.54) para βi em (4.3.45), então

(4.3.55) $$f(t,s) = \{ \delta_1 \log(1-t) + \gamma_1 \log t + \xi_1\}\{ s \log s + (1-s)\log(1-s)\} + \{ \delta_2 \log(1-t) + \gamma_2 \log t + \xi_2\}s + \{ \delta_3 \log(1-t) + \gamma_3 \log t + \xi_3\}$$

Agora substituindo (4.3.55) em (4.3.6), g é da seguinte forma

(4.3.56) $$g(t,s) = \{\delta_1 \log(1-t) + \gamma_1 \log t + \xi_1\}A(s) + \{\delta_2 \log(1-t) + \gamma_2 \log t + \xi_2\}s + \{\delta_3 \log(1-t) + \gamma_3 \log t + \xi_3\} + \{\delta_1 \log(1-s) + \gamma_1 \log s + \xi_1\}A(t) + \{\delta_2 \log(1-s) + \gamma_2 \log s + \xi_2\}t + \{\delta_3 \log(1-s) + \gamma_3 \log s + \xi_3\} + (t+s-2)f(0,0) + g(0,0)$$

onde $A(t) = t \log t + (1-t)\log(1-t)$

Substituindo a expressão (4.3.55) para f e (4.3.56) para g respect-ivamente em (4.3.4) quando $\alpha=0$.

(4.3.57) $$\{\delta_1 \log(1-x) + \gamma_1 \log x + \xi_1\}A(y) + \{\delta_2 \log(1-x) + \gamma_2 \log x + \xi_2\}y + \{\delta_3 \log(1-x) + \gamma_3 \log x + \xi_3\} + \{\delta_1 \log(1-y) + \gamma_1 \log y + \xi_1\}A(x) + \{\delta_2 \log(1-y) + \gamma_2 \log y + \xi_2\}x + \{\delta_3 \log(1-y) + \gamma_3 \log y + \xi_3\} + (x+y-2)f(0,0) + g(0,0) + (1-y)\{[\delta_1 \log(1-\frac{u}{1-x}) + \gamma_1 \log(\frac{u}{1-x}) + \xi_1] \cdot [\frac{v}{1-y}\log\frac{v}{1-y} + (1-\frac{v}{1-y})\log(1-\frac{v}{1-y})] + [\delta_2 \log(1-\frac{u}{1-x}) + \gamma_2 \log\frac{u}{1-x} + \xi_2](\frac{v}{1-y}) + [\delta_3 \log(1-\frac{u}{1-x}) + \gamma_3 \log\frac{u}{1-x} + \xi_3]\} + (1-x)\{[\delta_1 \log(1-\frac{v}{1-y}) + \gamma_1 \log\frac{v}{1-y} + \xi_1] \cdot [\frac{u}{1-x}\log\frac{u}{1-x} + (1-\frac{u}{1-x})\log(1-\frac{u}{1-x})] + [\delta_2 \log(1-\frac{v}{1-y}) + \gamma_2 \log\frac{v}{1-y} + \xi_2](\frac{u}{1-x})$$

$$|\delta_3\log(1-\frac{v}{1-y}) + \gamma_3\log\frac{v}{1-y} + \xi_3|\}$$

$$= \{\delta_1\log(1-u) + \gamma_1\log u + \xi_1\}A(v) + \{\delta_2\log(1-u) + \gamma_2\log u + \xi_2\}v +$$

$$\{\delta_3\log(1-u) + \gamma_3\log u + \xi_3\} + \{\delta_1\log(1-v) + \gamma_1\log v + \xi_1\}A(u) +$$

$$\{\delta_2\log(1-v) + \gamma_2\log v + \xi_2\}u + \{\delta_3\log(1-v) + \gamma_3\log v + \xi_3\} +$$

$$(u+v-2)f(0,0) + g(0,0) + (1-v)\{|\delta_1\log(1-\frac{x}{1-u}) + \gamma_1\log\frac{x}{1-u} + \xi_1|\cdot$$

$$A(\frac{y}{1-v}) + [\delta_2\log(1-\frac{x}{1-u}) + \gamma_2\log\frac{x}{1-u} + \xi_2|(\frac{y}{1-v}) + |\delta_3\log(1-\frac{x}{1-u})$$

$$+ \gamma_3\log\frac{x}{1-u} + \xi_3|\} + (1-u)\{|\delta_1\log(1-\frac{y}{1-v}) + \gamma_1\log\frac{y}{1-v} + \xi_1|A(\frac{x}{1-u})$$

$$+ |\delta_2\log(1-\frac{y}{1-v}) + \gamma_2\log\frac{y}{1-v} + \xi_2|(\frac{x}{1-u}) + |\delta_3\log(1-\frac{y}{1-v}) +$$

$$\gamma_3\log\frac{y}{1-v} + \xi_3|\}$$

Agora, comparando os coificientes de y log y

(4.3.58) $\delta_1\log(1-x) + \gamma_1\log x + \xi_1 = \delta_1\log(1-\frac{x}{1-u}) + \gamma_1\log\frac{x}{1-u} + \xi_1$

Isto implica que

(4.3.59) $\delta_1 = \gamma_1 = 0$

Novamente comparando os coificientes de y em (4.3.57)

(4.3.60) $\delta_2\log(1-x) + \gamma_2\log x + \xi_2 + f(0,0) - |\delta_3\log(1-\frac{u}{1-x}) + \gamma_3\log\frac{u}{1-x} + \xi_3|$

$$= \delta_2\log(1-\frac{x}{1-u}) + \gamma_2\log\frac{x}{1-u} + \xi_2$$

Isto implica que

(4.3.61) $\delta_2 = \gamma_2 = -\delta_3$, $\xi_3 = f(0,0)$, $\gamma_3 = 0$

Substituimos (4.3.59) e (4.3.61) em (4.3.55) e (4.3.56) respectivamente. portanto

(4.3.62) $f(t,s) = \xi_1 A(s) + \delta_2\{(s-1)\log(1-t) + s\log t\} + \xi_2 s + f(0,0)$

e

$$(4.3.63)\quad g(t,s) = \xi_1\{A(s) + A(t)\} + \delta_2\{(s-1)\log(1-t) + s\log t + (t-1)\log(1-s)$$
$$+ t\log s\} + (s+t)\{\xi_2 + f(0,0)\} + g(0,0)$$

Se temos $A = \delta_2$, $B = \xi_1$, $C = \xi_2$, então (4.3.63) e da forma da (4.3.8),

Caso II. Quando $\alpha = 1/2$

Neste caso, a equação funcional (4.3.4) e (4.3.5) torna-se

$$(4.3.64)\quad g(x,y) + (1-x)^{1/2}(1-y)^{1/2}\{ f(\tfrac{u}{1-x}, \tfrac{v}{1-y}) + f(\tfrac{v}{1-y}, \tfrac{u}{1-x}) \}$$
$$= g(u,v) + (1-u)^{1/2}(1-v)^{1/2}\{ f(\tfrac{x}{1-u}, \tfrac{y}{1-v}) + f(\tfrac{y}{1-v}, \tfrac{x}{1-u}) \}$$

(4.3.65) O lado direito do (4.3.64)

$$= g(x+u,y+v) + (x+u)^{1/2}(y+v)^{1/2}\{ f(\tfrac{x}{x+u}, \tfrac{y}{y+v}) + f(\tfrac{y}{y+v}, \tfrac{x}{x+u})\}$$

para $x,y,u,v \in [0,1)$ e $x+u$, $y+v \in (0,1]$.

Seja $u = v = 0$ em (4.3.64), então temos

$$(4.3.66)\quad g(x,y) = f(x,y) + f(y,x) - 2(1-x)^{1/2}(1-y)^{1/2}f(0,0) + g(0,0)$$

Substituindo (4.3.66) em (4.3.64), então temos envolvida somente uma função desconhecida f.

$$(4.3.67)\quad f(x,y) + f(y,x) + (1-x)^{1/2}(1-y)^{1/2}\{ f(\tfrac{u}{1-x}, \tfrac{v}{1-y}) + f(\tfrac{v}{1-y}, \tfrac{u}{1-x}) - 2f(0,0)\}$$
$$= f(u,v) + f(v,u) + (1-u)^{1/2}(1-v)^{1/2}\{ f(\tfrac{x}{1-u}, \tfrac{y}{1-v}) + f(\tfrac{y}{1-v}, \tfrac{x}{1-u}) - 2f(0,0)\}$$

Seja

$$(4.3.68)\quad h(x,y) = f(x,y) + f(y,x) - 2f(0,0)$$

(4.3.67) ē simplificada para a seguinte forma

$$(4.3.69)\quad h(x,y) + (1-x)^{1/2}(1-y)^{1/2}h(\tfrac{u}{1-x}, \tfrac{v}{1-y}) = h(u,v) + (1-u)^{1/2}(1-v)^{1/2}h(\tfrac{x}{1-u}, \tfrac{y}{1-v})$$

Esta equação funcional (4.3.69) foi resolvida por Kannappan e Rathie (1975), a solução de (4.3.69) é da seguinte forma.

(4.3.70) $h(x,y) = Ax^{1/2}y^{1/2} - B\{x^{1/2}y^{1/2} + (1-x)^{1/2}(1-y)^{1/2} - 1\}, \quad x,y \in [0,1]$

De (4.3.66) e (4.3.68), é claro que

(4.3.71) $g(x,y) = h(x,y) - 2\{(1-x)^{1/2}(1-y)^{1/2} - 1\}f(0,0) + g(0,0)$

Portanto de (4.3.70)

$$g(x,y) = Ax^{1/2}y^{1/2} - B\{x^{1/2}y^{1/2} + (1-x)^{1/2}(1-y)^{1/2} - 1\} + 2\{1 - (1-x)^{1/2}(1-y)^{1/2}\}f(0,0) + g(0,0)$$

E isto é (4.3.9)

Isto prova o lema.

<u>Continuação da prova do teorema.</u>

Caso I. Quando $\alpha \neq 1/2$

(a) Quando $\alpha \neq 0$

Da propriedade (iv), $g(x,x) = 0$ para $x \in [0,1]$, temos

(4.3.72) $g(0,0) = 0$

E de (4.3.7)

(4.3.73) $0 = g(x,x) = \{A - f(0,0)\}\{2 - 2x\} + 2xB + C$

Isto implica que

(4.3.74) $f(0,0) = A = -B, \quad C = -2B$

Usando (4.3.72) e (4.3.74), g torna-se

(4.3.75) $g(x,y) = \lambda_1\{(1-x)^{\alpha}(1-y)^{1-\alpha} + (1-y)^{\alpha}(1-x)^{1-\alpha} + x^{\alpha}y^{1-\alpha} + y^{\alpha}x^{1-\alpha} - 2\}$

onde $\lambda_1 = B$.

(b) Quand $\alpha = 0$

A propriedade (ix) produz

(4.3.76) $g(0,0) = 0$

E de (4.3.8)

(4.3.77) $0 = g(x,x) = 2A\{(x+1)\log(1-x) + x\log x\} + 2B\{x\log x + (1-x)\log(1-x)\} + 2\{C + f(0,0)\}x$

Isto implica que

(4.3.78) $A = B = 0\ ,\qquad f(0,0) = -C$

Usando (4.3.76) e (4.3.78), temos que (4.3.8) torna-se

(4.3.79) $g(x,y) = 0$

Isto é uma solução trivial.

Caso II. Quando $\alpha = 1/2$

A propriedade (ix) produz

(4.3.80) $g(0,0) = 0$

E de (4.3.9)

(4.3.81) $0 = g(x,x) = Ax + 2xf(0,0)$

Isto implica que

(4.3.82) $f(0,0) = -A/2$

Usando (4.3.80) e (4.3.82), portanto (4.3.9) torna-se

(4.3.83) $g(x,y) = \lambda_2\{x^{1/2}y^{1/2} + (1-x)^{1/2}(1-y)^{1/2} - 1\}$

onde $\lambda_2 = A - B$.

Da propriedade (x), é claro que

$$\lambda_i = (2^{\alpha-1} - 1)^{-1}.2^{i-1}$$

Assim podemos escrever g na seguinte forma

(4.3.84) $g(x,y) = \dfrac{1}{2^{\alpha-1}-1}\{x^\alpha y^{1-\alpha} + y^\alpha x^{1-\alpha} + (1-x)^\alpha(1-y)^{1-\alpha} + (1-y)^\alpha(1-x)^{1-\alpha} - 2\}$

Quando $\alpha \neq 1/2$, seja $u = v = 0$ em (4.3.5), então

(4.3.85) $\{(1-x)^\alpha(1-y)^{1-\alpha} + (1-y)^\alpha(1-x)^{1-\alpha}\}f(0,0)$

$= \{x^\alpha y^{1-\alpha} + y^\alpha x^{1-\alpha}\}f(1,1)\ ,\qquad x,y \in (0,1)$

Isto fornece

(4.3.86) $f(0,0) = f(1,1) = 0$

Substituindo (4.3.86) em (4.3.41), temos

(4.3.87) $\delta_1 = \gamma_2 = -\xi_3$

Isto é

(4.3.88) $f(t,s) = \delta_1\{(1-t)^{\alpha}(1-s)^{1-\alpha} + t^{\alpha}s^{1-\alpha} - 1\}$

Usando o fato $f(0,0) = g(0,0) = 0$, vemos de (4.3.6) que

(4.3.89) $g(x,y) = f(x,y) + f(y,x)$

Desde que $g(x,y)$ é explicitamente dado em (4.3.84), isto levá a

que

(4.3.90) $\delta_1 = \dfrac{1}{2^{\alpha-1}-1}$

E

(4.3.91) $f(t,s) = \dfrac{1}{2^{\alpha-1}-1}\{(1-t)^{\alpha}(1-s)^{1-\alpha} + t^{\alpha}s^{1-\alpha} - 1\}$

Quando $\alpha = 1/2$, seja $u = v = 0$ em (4.3.65), nós novamente temos

(4.3.92) $f(0,0) = f(1,1) = 0$

Usando o fato $f(0,0) = g(0,0) = 0$, vemos de (4.3.66) que

(4.3.93) $g(x,y) = f(x,y) + f(y,x)$

Desde que $g(x,y)$ é explicitamente dado em (4.3.84) (tome $\alpha = 1/2$),

portanto

(4.3.94) $f(x,y) + f(y,x) = \dfrac{2}{2^{-1/2}-1}\{x^{1/2}y^{1/2} + (1-x)^{1/2}(1-y)^{1/2} - 1\}$

O uso sucessivo da propriedade (iv) da

(4.3.95) $J_{n,\alpha}(p_1,\ldots,p_n:q_1,\ldots,q_n)$

$$= J_{2+\alpha}(p_1+\ldots+p_{n-1},p_n:q_1+\ldots+q_{n-1},q_n) +$$

$$= \sum_{i=1}^{n-1} r_i^{\alpha}s_i^{1-\alpha}f\left(\frac{p_i}{r_i}, \frac{q_i}{s_i}\right) + \sum_{i=1}^{n-1} s_i^{\alpha}r_i^{1-\alpha}f\left(\frac{q_i}{s_i}, \frac{p_i}{r_i}\right)$$

onde $r_i = p_1+\dots+p_i$; $s_i = q_1+\dots+q_i$.

Quando $\alpha \neq 1/2$

(4.3.84) e (4.3.91) produz

(4.3.96) $J_{n,\alpha}(p_1,\dots,p_n:q_1,\dots,q_n)$

$$= \frac{1}{2^{\alpha-1}-1}\{[p_n^{\alpha}q_n^{1-\alpha}+(1-p_n)^{\alpha}(1-q_n)^{1-\alpha}+q_n^{\alpha}p_n^{1-\alpha}+(1-q_n)^{\alpha}(1-p_n)^{1-\alpha}-2] +$$

$$= \sum_{i=2}^{n-1}[p_i^{\alpha}q_i^{1-\alpha}+r_{i-1}^{\alpha}s_{i-1}^{1-\alpha}-r_i^{\alpha}s_i^{1-\alpha}] + \sum_{i=2}^{n-1}[q_i^{\alpha}p_i^{1-\alpha}+s_{i-1}^{\alpha}r_{i-1}^{1-\alpha}-s_i^{\alpha}r_i^{1-\alpha}]\}$$

$$= \frac{1}{2^{\alpha-1}-1}\{\sum_{i=1}^{n}p_i^{\alpha}q_i^{1-\alpha}+\sum_{i=1}^{n}q_i^{\alpha}p_i^{1-\alpha}-2\}$$

Quando $\alpha = 1/2$

(4.3.84) e (4.3.94) produz

(4.3.97) $J_{n,1/2}(p_1,\dots,p_n:q_1,\dots,q_n)$

$$= g(p_n,q_n) + \sum_{i=1}^{n-1} r_i^{1/2}s_i^{1/2}\{f(\frac{p_i}{r_i}, \frac{q_i}{s_i}) + f(\frac{q_i}{s_i}, \frac{p_i}{r_i})\}$$

$$= \frac{2}{2^{-1/2}-1}\{\sum_{i=1}^{n}p_i^{1/2}q_i^{1/2}-1\}$$

Isto prova o teorema.

REFERENCIA BIBLIOGRAFICA

Aczél, J. (1966). Lectures on Functional Equations and Their Applications. Academic Press, New York.

Aczél, J. e Daróczy, Z. (1975). On Measures of Information and Their Characterizations. Academic Press, New York.

Behara, M. e Nath, P. (1973). Additive and Non-Additive Entropies of Finite Measurable Partitions. Lecture Notes in Mathematics, No. 296, Springer-Verlag, 102-138.

Chaundy, T. W. e McLeod, J. B. (1960). On A Functional Equation. Proc. Edin. Math. Notes, 43, 7-8.

Daróczy, Z. e Losonczi, L. (1967). Uber die Erweiterung der auf einer Punkmenge additive Funcktionen. Publ. Math. Debrecen 14, 239-245.

Forte, B. e Ng, C. T. (1975). Derivation of A Class of Entropies Including Those of Degree β. Information and Control, 28, 335-351.

George, A. e Mathai, A. M. (1974). Applications of The Concepts of Affinity and Distance to Population Problems. J. Biosocial Sciences, 6, 347-356.

Haaland, P., Brockett, P. L. e Levine, A. (1979) A Characterization of

Divergence With Applications to Questionnaire Information. A ser publicado.

Havrda, J. e Charvát, F. (1967). Quantification Method of Classification processes. The Concept of Structural a-entropy, Kybernetika,3, 30-35.

Kannappan, PL. e Rathie, P. N. (1975 a). On The Solution of A Functional Equation Connected With Inaccuracy. J. Indian Math. Soc., 39,131-147.

Kannappan, PL. e Rathie, P. N. (1975 b). On Generalized Information Function. Tôhoku Math. J.,27, 207-212.

Kannappan, PL. e Rathie, P. N. (1979). An Axiomatic Characterization of J-Divergence. A ser publicado.

Kerridge, D. F. (1961). Inaccuracy and Inference. J. Roy. Statist. Soc. , ser. B., 23, 184-194.

Kullback, S. (1959). Information Theory and Statistics. Wiley, New York.

Mathai, A. M. e Rathie, P. N. (1975). Basic Concepts in Information Theory and Statistics: Axiomatic Foundations and Applications. Wiley, New York.

Matusita, K. (1954). On The Estimation by The Minimum Distance Method. Ann. Inst. Statist. Math., 5, 59-65.

Matusita, K. (1955). Decision Rules Based on The Distance, for Problem of Fit, Two Samples and Estimations. Ann. Math. Statist., 26, 631-640.

Matusita, K. (1961). Interval Estimation Based on The Notion of Affinity. Bull. International Statist. Inst., 38, 241-244.

Mittal, D. P. (1975). On Continuous Solution of A Functional Equation. Metrika, 22, 31-40.

Rathie, P. N. (1971). On The Solution of Functional Inequality and Its Applications. Tôhoku Math. J., 23, 681-690.

Rathie e Kannappan, PL. (1971). On A Functional Equation Connected With Shannon's Entropy. Funkcialaj Ekvacioj, 14, 153-159.

Rathie, P. N. e Kannappan, PL. (1972). A Directed Divergence Function of Type β. Information and Control, 20, 38-45.

Rathie, P. N. e Kannappan, PL. (1973). An Inaccuracy Function of Type β. Ann. Inst. Statist. Math., 25, 205-214.

Rényi, A. (1960). On Measure of Entropy and Information. Proc. Fourth Berkeley Symp. Math. Statist. Probability, 1, 547-561.

Shannon, C. E. (1948). A Mathematical Theory of Communication. Bell System Tech. J., 27, 379-423, 623-656.

Sharma, B. D. e Autar, R. (1974). Relative Information Functions and Their Type (α,β) Generalizations. Metrika, 21, 41-50.

Sharma, B. D. e Taneja, I. J. (1975). Entropy of Type (α,β) and Other Generalized Measures in Information Theory. Metrika, 22, 205-215.

Theil, H. (1967). Economic and Information Theory. North-Holland, Amsterdam.