UNIVERSIDADE FEDERAL DE SANTA CATARINA

UM ESTUDO DE GENERALIZAÇÃO DE IMPRECISÃO PARA VARIÁVEIS ALEATÓRIAS GENERALIZADAS

Joanete Maria Costa Valente

julho - 1984

UM ESTUDO DE GENERALIZAÇÕES DA IMPRECISÃO
PARA VARIAVEIS ALEATÓRIAS GENERALIZADAS

por

Joanete Maria Costa Valente

Esta dissertação foi julgada adequada para a obtenção do título de

"MESTRE EM CIÊNCIAS"

Especialidade em MATEMÁTICA, e aprovada em sua forma final pelo Curso de Pós-Graduação em Matemática da UNIVERSIDADE FEDERAL DE SANTA CATARINA

William G Whitley

Prof. William Glenn Whitley, PhD.
Coordenador

Banca Examinadora:

Gurdial

Prof. Gur Dial, PhD.
Orientador

Donald Morison Silberger

Prof. Donald Morison Silberger, PhD.

P Stange

Prof. Plinio Stange, PhD.

À Cláudia Regina,

à Izabela Carla

ao Mílton Luíz Jr. e

ao Mílton Luíz

A G R A D E C I M E N T O S

Ao Professor GUR DIAL pela orientação e dedicação.

A meus filhos pela compreensão nos momentos difíceis.

A meu esposo pelo apoio, incentivo e compreensão.

A meus pais e irmãos pela atenção e incentivo.

A colega e amiga Rosemari Pereira por sua colaboração.

Ao Curso de Pós-Graduação, ao Departamento de Matemática e a seus funcionários e professores que tornaram possível este trabalho.

A Direção e colegas do Colégio Estadual Professor Simão José Hess e da Faculdade de Educação da UDESC pelo apoio e incentivo.

R E S U M O.

O conceito de imprecisão introduzido por Kerridge em 1961 é uma generalização da entropia de Shannon para variáveis aleatórias completas. Este conceito tem sido aplicado para problemas de inferência em estatística.

No capítulo I, a imprecisão de Kerridge e suas generalizações são definidas e as suas propriedades são citadas.

No capítulo II, o conceito das variáveis aleatórias generalizadas é introduzido e imprecisões são definidas para estas variáveis. A desigualdade de Shannon é generalizada - para estas variáveis e condições são determinadas sob as quais a desigualdade é verdadeira ou não, para imprecisão de Kerridge e imprecisão de ordem α. A característica monotônica da imprecisão de ordem α é estudada com relação ao parâmetro α.

A medida de imprecisão é definida na forma geral e algumas generalizações são estudadas. Para distribuições potência de probabilidade uma nova medida, dita beta -imprecisão, é introduzida.

No capítulo III, que é o desenvolvimento original deste trabalho, as imprecisões de ordem 1 e grau β e ordem α e grau β são introduzidas. A desigualdade de Shannon é generalizada para imprecisão de ordem α e grau β e condições são determinadas sob as quais a desigualdade é verdadeira ou não. Finalmente a monotonicidade da imprecisão em relação a α e β é estudada e gama-imprecisão de ordem α e tipo β é introduzida.

ABSTRACT

The idea of inaccuracy, which is an generalization of Shannon's entropy, was introduced in 1961 by Kerridge for complete random variables. This concept has found applications for problems of statistical inference.

In chapter I, Kerridge's inaccuracy and its generalizations are defined and the properties of these measures are given.

In chapter II, the concept of the generalized random variables is introduced and various inaccuracies are defined for these variables. Shannon's inequality is generalized for these variables, and conditions are determined, under which the inequality is true or not, in the case of Kerridge's inaccuracy; and the inaccuracy of order α is studied with relation to the parameter α. The measure of inaccuracy is defined in the general form and some generalizations are studied. For power distributions, gama-inaccuracy is introduced.

In chapter III, which is the original work of this dissertation, the inaccuracies of order 1 and type β and order α and type β are introduced. Shannon's inequality is generalized for inaccuracy of order α and type β, and conditions are determined under which this inequality is true or not. The monotonicity of inaccuracy with relation to α and β is studied, and gama-inaccuracy of order α and type β is introduced.

Í N D I C E

CAPÍTULO I - RESULTADOS PRELIMINARES

1.1. - Introdução 2

1.2. - Entropia de Shannon 3

1.3. - Medidas de imprecisão 6

1.4. - Generalizações da imprecisão de Kerridge 10

CAPÍTULO II - ALGUMAS GENERALIZAÇÕES DE IMPRECISÃO PARA VARIÁVEIS ALEATÓRIAS GENERALIZADAS

2.1. - Introdução 14

2.2. - Entropia e imprecisão 17

2.3. - Características monotônicas da imprecisão de ordem α 21

2.4. - Imprecisão como valor médio 27

2.5. - β-imprecisão 35

CAPÍTULO III- IMPRECISÃO DE ORDEM 1 E GRAU β E DE ORDEM α E GRAU β

3.1. - Definição 37

3.2. - Generalização da desigualdade de Shannon .. 37

3.3. - Características monotônicas da imprecisão de ordem α e grau β 40

3.4. - Gama-imprecisão 42

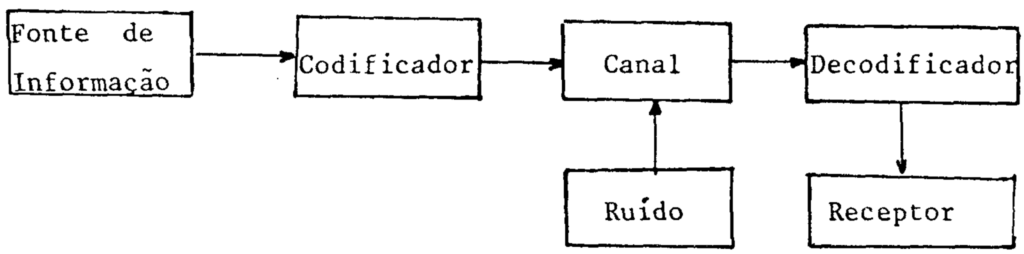
BIBLIOGRAFIA 44

C A P Í T U L O I - RESULTADOS PRELIMINARES

1.1. - Introdução:

A Teoria da Informação, localizada como uma das mais recentes áreas de estudo da matemática, por suas aplicações, tem sido objeto de vários trabalhos nos mais conceituados centros de pesquisa em todo o mundo. Consiste esta teoria, basicamente , no estudo da transmissão de informação entre uma fonte e um receptor estabelecendo parâmetros ou medidas para que esta transmissão seja possível.

Shannon [16] apresenta o esquema abaixo que mostra claramente o problema:



(figura 1)

Fonte de Informação: É onde se origina ou se produzem as mensagens a serem transmitidas. Pode ser uma pessoa ou uma máquina. Na comunicação telefônica, por exemplo, a fonte é uma pessoa.

As informações geradas por uma fonte são geralmente representadas por uma variável aleatória X, discreta ou contínua.

Codificador: Transforma a linguagem da fonte para a linguagem do canal sem alterar o conteúdo da mensagem.

Decodificador: Decifra a mensagem transmitida em linguagem inteligível ao receptor.

Receptor: Destino da mensagem. Representado usualmente por uma variável aleatória Y, discreta ou contínua.

Ruído: Dificuldades ou perturbações que podem acontecer no decorrer da transmissão da mensagem pelo canal.

Ao discutir sobre a natureza estatística da comunicação de mensagens Shannon [16] considerou um modelo de sistema de comunicação como o que aparece na figura 1 anterior. É interessante observar que a Teoria da Informação procura construir um modelo matemático para cada bloco da citada figura. Desta forma, em tais modelos, fontes, codificador, canal, decodificador e receptor devem ser estatísticamente definidos. Isto nos leva a afirmar que os sistemas de comunicação que estudamos são de natureza estatística e que, portanto, seus comportamentos não podem ser descritos de forma determinística. Por exemplo, ao estudarmos uma fonte não estamos interessados em uma fonte particular que emite informações perfeitamente determinadas. Nos interessa é que a mesma emita uma sequência de símbolos ao acaso, baseados em alguma estatística.

Um dos resultados principais da Teoria de Informação é que é possível transmitir informações por um canal com ruídos, em qualquer velocidade menor que a capacidade do canal, com uma pequena e arbitrária "probabilidade de erro". Simbolicamente: Se R é a velocidade de transmissão da informação e C é a capacidade do canal, então:

Se $R<C$, pode-se transmitir a informação com pequena probabilidade de erro, em qualquer canal ruidoso. Este resultado é conhecido como teorema fundamental da Teoria da Informação. A primeira demonstração deste teorema devemos a Shannon [16]. Para tanto, ele introduziu uma medida de incerteza (denominada entropia de Shannon) que é definida a seguir.

1.2. - Entropia de Shannon

Seja $\Gamma_n = \{P=(p_1,p_2,\ldots p_n);\ p_k\geq 0,\ \sum_{k=1}^{n} p_k=1\}$, $n\geq 2$ o conjunto das distribuições de probabilidades completas.

Seja $X=(x_1,x_2,\ldots,x_n)$ uma variável aleatória discreta com distribuição de probabilidade P pertencente a Γ_n.

A medida de incerteza introduzida por Shannon [16], denominada entropia de Shannon, é dada por:

$$H(X) = H_n(P) = -\sum_{k=1}^{n} p_k \log p_k;\quad P\epsilon\Gamma_n.$$

Observações: 1. Usaremos por convenção que $0\log 0=0$.

2. Caso não seja especificado o logarítmo será usado com base dois.

3. No decorrer deste trabalho Σ denotará $\sum_{k=1}^{n}$ quando n e k são bem entendidos.

Esta medida tem muitas aplicações e foi generalizada por vários pesquisadores [1] e [7].

Na literatura da Teoria de Informação existem várias caracterizações da entropia usando equações funcionais e aplicando algumas propriedades que são essenciais do ponto de vista desta teoria. Citamos abaixo algumas propriedades importantes:

P_1- Simetria:

$H_n(p_1,p_2,\ldots,p_n) = H_n(p_{k(1)},p_{k(2)},\ldots,p_{k(n)})$, para todo $(p_1,p_2,\ldots,p_n)\varepsilon\Gamma_n$, onde k é uma permutação arbitrária de $\{1,2,\ldots,n\}$.

O sentido intuitivo de simetria é que a quantidade de informação é invariante sob a troca de ordem dos eventos.

P_2 - Normalidade:

$H_2(1/2;1/2) = 1.$

Esta propriedade estabelece que em um experimento com duas consequências equiprováveis, temos uma unidade de informação (1 bit).

P_3 - Decisividade:

$H_2(1;0) = H_2(0;1) = 0.$

Esta propriedade estabelece que não há incerteza em um experimento com duas consequências, uma de probabilidade 1 e outra de probabilidade zero.

P_4 - Expansividade

$H_n(p_1,p_2,\ldots,p_n) = H_{n+1}(p_1,p_2,\ldots,p_n,0)$

A interpretação dessa propriedade é que a adição de conse - quências de probabilidade zero não muda a incerteza do experimento.

P_5 - Aditividade forte:

$H_{mn}(p_1q_{11},\ldots,p_1q_{1m};\ p_2q_{21},\ldots,p_2q_{2n};\ldots;p_2q_{m1},\ldots p_mq_{mn})$ =

$H_m(p_1,\ldots,p_m) + \Sigma p_k H_n(q_{k1},\ldots q_{kn})$, para todo

$(p_1,\ldots,p_m)\varepsilon\Gamma_m$, $(q_{k1},\ldots q_{kn})\varepsilon\Gamma_n$.

P_6 - Aditividade:

$H_{mn}(p_1q_1,\ldots,p_1q_n;\ldots;p_mq_1,\ldots,p_mq_n)$ =

$H_m(p_1,\ldots,p_m) + H_n(q_1,\ldots,q_n)$, para todo

$(p_1,\ldots,p_m)\varepsilon\Gamma_m$ e $(q_1,\ldots,q_n)\varepsilon\Gamma_n$.

P_7 - Recursividade:

$H_n(p_1,\ldots,p_n) = H_{n-1}(p_1+p_2,\ p_3,\ p_4,\ldots,p_n)$ +

$(p_1+p_2)\ H_2(p_1/p_1+p_2;p_2/p_1+p_2)$, para todo

$(p_1,\ldots,p_n)\varepsilon\Gamma_n$ e para $p_1+p_2>0$.

P_8 - Maximalidade:

$H_n(p_1,\ldots,p_n) \leqslant H_n(1/n,\ldots,1/n)$ para todo $(p_1,\ldots,p_n)\varepsilon\Gamma_n$.

Esta propriedade estabelece que a entropia, como uma medida de incerteza, é máxima quando todos os eventos admissíveis tem probabilidades iguais.

P_9 - Sub-aditividade:

$H_{mn}(p_{11},p_{12},\ldots,p_{1n};\ p_{21},p_{22},\ldots,p_{2n};\ldots;p_{m1},p_{m2},\ldots,p_{mn}) \leqslant$

$H_m(\Sigma p_{1k},\Sigma p_{2k},\ldots,\Sigma p_{mk})+H_n(\Sigma p_{j1},\Sigma p_{j2},\ldots,\Sigma p_{jn})$, para todo

$(p_{11},p_{12},\ldots,p_{mn})\varepsilon\Gamma_{mn}$.

O significado intuitivo desta propriedade é que a informação esperada de dois experimentos (não necessariamente independentes) não é maior que a soma das informações esperadas dos experimentos individuais.

P_{10} - Não-negatividade:

$H_n(p_1,\ldots,p_n)\geqslant 0$, para todo $(p_1,\ldots,p_n)\varepsilon\Gamma_n$. Com a igualdade se $p_i=1$ para algum i e $p_j=0$, $j=1,2,\ldots,n$, $j\neq i$,

P_{11} - Continuidade:

$H_n(p_1,\ldots,p_n)$ é uma função contínua de suas variáveis.

P_{12} - Monotonicidade:

$H_n(1/n,\ldots,1/n)$ é uma função monótona crescente de n.

P_{13} - Integrabilidade a Lebesgue:

$H_2(p;1-p)$ é Lebesgue integrável em $[0,1]$

P_{14} - Mensurabilidade:

$H_2(p,1-p)$ é uma função mensurável de $p\varepsilon[0,1]$.

1.3. - Medidas de Imprecisão.

Vamos supor que um certo experimento assuma probabilidades dos possíveis resultados de um certo experimento dado. As falhas ou imprecisões a que este experimento pode estar sujeito devem considerar as possibilidades de não se ter informações suficientes, sendo portanto de enunciado vago, ou ter alguma informação incorreta. Devemos lembrar que isto é o que ocorre também na prática, pois as pesquisas estatísticas são raramente conhecidas com perfeita exatidão.

Todo cálculo estatístico e problemas inferidos estão relacionados com a elaboração de resultados que podem ser incorretos em qualquer um ou em ambos os aspectos acima citados.

A teoria de informação em [16] fornece uma teoria geral de incerteza que nos possibilita lidar com o aspecto vago de imprecisão. A utilidade da imprecisão nos problemas estatísticos tem sido mostrada por vários autores [6]. Contudo, a teoria de informação desenvolvida por Shannon [16], não trabalha com imprecisão no sentido mais geral. Com o intuito de remover esta limitação e mostrar sua utilidade em um sentido mais amplo, Kerridge [5] introduziu a idéia de imprecisão que é definida a seguir:

Definição 1.1. - Sejam $P,Q \in \Gamma_n$. Consideremos que a distribuição P de um certo experimento está condicionada a ocorrência da distribuição Q. Então, a imprecisão de Kerridge, é definida como:

$$H_n(P//Q) = - \Sigma p_k \log q_k \tag{1}$$

Observações: 1. - Usaremos por convenção que, se $q_k=0$, o correspondente $p_k=0$.

2. - É fácil ver que quando $P = Q$, (1) reduz-se para $H_n(P) = - \Sigma p_k \log p_k$ que é a entropia de Shannon.

Considerando que a medida de imprecisão é uma generalização da entropia de Shannon, várias propriedades são comuns para as duas medidas. Abaixo listamos as propriedades da imprecisão de Kerridge.

P_1 - Não-negatividade:

$H_n(P//Q) \geq 0$, para $P,Q \in \Gamma_n$.

Com a igualdade se e somente se, $p_k = q_k = 1$ para algum k e $p_k=q_k=0$, para todos os outros distintos k. Isto significa que a imprecisão zero implica em um resultado correto, obtido com certeza absoluta.

P_2 - Expansibilidade:

$H_{n+1}(p_1,\ldots,p_n,0//q_1,\ldots,q_n,0)=H_n(P//Q)$, para $P,Q \in \Gamma_n$.

P_3 - Simetria:

$$H_n(p_{k(1)},\dots,p_{k(n)}//q_{k(1)},\dots,q_{k(n)}) =$$

$H(p_1,\dots,p_n//q_1,\dots,q_n)$, para $P,Q\varepsilon\Gamma_n$, com $\{k(1),\dots,k(n)\}$ permutação arbitrária de $\{1,2,\dots,n\}$.

P_4 - Recursividade:

$$H_n(p_1,\dots,p_n//q_1,\dots,q_n) =$$

$$H_{n-1}(p_1+p_2,p_3,\dots,p_n//q_1+q_2,q_3,\dots,q_n) \quad +$$

$(p_1+p_2)H_2(p_1/p_1+p_2,p_2/p_1+p_2//q_1/q_1+q_2,q_2/q_1+q_2)$, para $P,Q\varepsilon\Gamma_n$,

com $n=3,4,\dots$ e $p_1+p_2>0$ e $q_1+q_2>0$.

P_5 - Aditividade:

$$H_{mn}(p_1q_1,\dots,p_1q_n;\dots;p_mq_1,\dots,p_mq_n//r_1s_1,\dots,r_1s_n;\dots; r_ms_1,\dots,r_ms_n) =$$

$H_m(p_1,\dots,p_m//r_1,\dots,r_m) + H_n(q_1,\dots,q_n//s_1,\dots,s_n)$, para

p_i, q_j, r_i, $s_j \geqslant 0$, $i=1,2,\dots,n$; $j=1,2,\dots,m$ e $\Sigma p_i=1$, $\Sigma q_j=1$, $\Sigma r_i=1$ e $\Sigma s_j=1$.

P_6 - Aditividade forte:

$$H_{mn}(p_{11},\dots,p_{nm}//q_{11},\dots,q_{nm}) =$$

$$H_m(p_1,\dots,p_m//q_1,\dots,q_m) + \Sigma p_j H_n(p_{1j}/p_j,\dots,p_{nj}/p_j// q_{1j}/q_j,\dots,q_{nj}/q_j),$$

para $\Sigma\Sigma p_{ij}=1$, $p_j=\Sigma p_{ij}>0$, $q_j=\Sigma q_{ij}>0$, $\Sigma\Sigma q_{ij}=1$ e $i=1,\dots n$ e $j=1,\dots,m$.

P_7 - Continuidade:

$H_n(p_1,...,p_n//q_1,...,q_n)$ é uma função contínua para 2n variáveis.

P_8 - Monotonicidade:

Quando n-alternativas igualmente prováveis são estabelecidas, a imprecisão é uma função monotonicamente crescente de n, isto é;

$H_n(1/n,...,1/n//1/n,...,1/n)$ é uma função monotonicamente - crescente de n.

P_9 -

A imprecisão de uma proposição é inalterada se duas alternativas, sobre as quais a mesma afirmação é feita, são combinadas. Por exemplo:

$$H_3(p_1,p_2,p_3//q_1,q_2,q_2) = H_2(p_1,p_2+p_3//q_1,q_2)$$

A função que satisfaz somente as propriedades P_4, P_7, P_8 e P_9 é dada por:

$$H(P//Q) = -k\sum_k p_k \log_D q_k,$$ onde k é uma constante arbitrária e D>1.

P_{10} -

Uma imprecisão assume um valor infinito se $q_k=0$ e $p_k \neq 0$, para algum k. O significado disto é claro se considerarmos a aplicação quando q_k é uma estimativa de probabilidade de que a hipótese é verdadeira. Se $q_k=0$ está enunciado incorretamente, isto pode conduzir a uma teoria sempre verdadeira a qual, por não ser de interesse, será permanentemente descartada. O enunciado implica em que a verdade é considerada como infinitamente avaliável.

Observações: 1. As afirmações dadas por Kerridge são intuitivamente corretas.

2. Em conformidade com a teoria da informação tomamos $k=1$. $(D=2)$.
3. Escrevemos $H_n(P//Q)$ como:

$$H_n(P//Q) = -\Sigma p_k \log p_k - \Sigma p_k \log q_k/p_k.$$

Nesta equação, o primeiro termo do lado direito representa a incerteza e o segundo termo a imprecisão de erro. Esta interpretação é baseada no fato de que o termo erro desaparece quando $p_k=q_k$ para todo k e todo q_k positivo.

Existe uma interessante dualidade entre a informação e a entropia (ou incerteza) na teoria da informação. Isto está relacionado ao fato de que uma medida de incerteza de uma situação é a quantidade de informação a qual deve ser obtida anteriormente.

De modo semelhante, a imprecisão pode ser relacionada para uma quantidade de informações ausentes.

1.4. - Generalizações da imprecisão de Kerridge.

Definição 1.2. - Sejam $P,Q \epsilon \Gamma_n$. Então a imprecisão não aditiva de grau β $(0<\beta\neq 1)$ é definida como:

$$H_n^\beta(P//Q) = A_\beta(\Sigma p_k q_k^{\beta-1}-1);\ 0<\beta\neq 1, \qquad (2)$$

onde $A_\beta = (2^{1-\beta}-1)^{-1}$, sendo β um parâmetro.

Observações: 1. - Quando $P=Q$ temos que $H_n^\beta(P//P) = H_n^\beta(P) = A_\beta(\Sigma p_k^\beta-1)$ que é a entropia de grau β.

2. - $\lim\limits_{\beta\to 1} H_n^\beta(P//Q) = H_n(P//Q)$

Definição 1.3. - Sejam $P,Q \epsilon \Gamma_n$. Então a medida aditiva de imprecisão de ordem α $(\alpha\neq 1)$ é definida como:

$$H_{n,\alpha}(P//Q) = (1-\alpha)^{-1}\log \Sigma p_k q_k^{\alpha-1},$$ para $0<\alpha\neq 1$, onde α é um parâmetro. (3)

Observações: 1. - Quando P=Q temos que $H_{n,\alpha}(P//Q)=(1-\alpha)^{-1}\log \Sigma p_k^{\alpha}$, a qual é a entropia de ordem α.

2. - $\lim_{\alpha\to 1} H_{n,\alpha}(P//Q) = H_n(P//Q)$.

3. - Existe a seguinte relação entre (2) e (3):

$$H_n^{\alpha}(P//Q) = (2^{(1-\alpha)H_{n,\alpha}(P//Q)}-1).\ A_{\alpha};\quad 0<\alpha\neq 1.$$

As seguintes propriedades são válidas para $H_{\beta}(P//Q)$ e $H_{n,\alpha}(P//Q)$:

P_1 - Não-negatividade:

$H_n^{\beta}(P//Q)$ e $H_{n,\alpha}(P//Q)$ ambas são não negativas. A igualdade ocorre se e somente se $p_k=1$ e $q_k=1$ e os demais são zeros.

P_2 - Expansibilidade (ou zero-indiferente):

$$H_{n+1,\alpha}(p_1,p_2,\ldots,p_n,0//q_1,q_2,\ldots,q_n,0) =$$

$$H_{n,\alpha}(p_1,\ldots,p_n//q_1,\ldots,q_n)$$ e

$$H_{n+1}^{\beta}(p_1,p_2,\ldots,p_n,0//q_1,q_2,\ldots,q_n,0) =$$

$$H_n^{\beta}(p_1,\ldots,p_n//q_1,\ldots q_n)$$

P_3 - Simetria:

$H_{n,\alpha}(P//Q)$ e $H_n^{\beta}(P//Q)$ ficam inalterados se os elementos de P e Q são rearranjos na mesma forma tal que a correspondência um a um entre eles não é alterada.

P_4 - Recursividade:

$$H_n^{\beta}(P//Q) = H_{n-1}^{\beta}(p_1+p_2,p_3,\ldots,p_n//q_1+q_2,q_3,\ldots,q_n) +$$

$$(p_1+p_2)(q_1+q_2)H_2^{\beta}(p_1/p_1+p_2,p_2/p_1+p_2//$$

$$q_1/q_1+q_2,q_2/q_1+q_2),$$ para

$P,Q\varepsilon\Gamma_n$ e $p_1+p_2>0$, $q_1+q_2>0$.

P_5 - Aditividade:

$$H_{m,n,\alpha}(p_1r_1,\dots p_1r_m;\dots;p_nr_1,\dots,p_nr_m;q_1s_1,\dots,q_1s_m,\dots q_ns_m)=$$

$$H_{n,\alpha}(p_1,\dots,p_n//q_1,\dots,q_n)+H_{m,\alpha}(r_1,\dots,r_m//s_1,\dots,s_m);$$

para $P=(p_1,\dots,p_n)$, $Q=(q_1,\dots,q_n)\varepsilon\Gamma_n$ e $S=(s_1,\dots,s_m)$, $R=(r_i,\dots,r_m)\varepsilon\Gamma_m$.

P_6 - Não-aditividade:

$$H^{\beta}_{m,n}(p_1,r_{11},\dots,p_1r_{m1},\dots,p_nr_{1n},\dots,p_nr_{mn}//$$
$$q_1s_{11},\dots,q_1s_{m1},\dots,q_ns_{1n},\dots,q_ns_{mn})=$$

$$H^{\beta}_n(P//Q)+\Sigma p_iq_i^{\beta-1}H^{\beta}_m(r_{1i},\dots,r_{mi}//s_{1i},\dots,s_{mi});\quad \text{para}$$

$P=(p_1,\dots,p_n)$, $Q=(q_1,\dots,q_n)\varepsilon\Gamma_n$ e $R=(r_{1i},\dots,r_{mi})$, $S=(s_{1i},\dots,s_{mi})\varepsilon\Gamma_m$ e para algum $i=1,2,\dots,n$.

P_7 - Continuidade:

H^{β}_n e $H_{n,\alpha}$ são funções contínuas em suas variáveis.

As seguintes generalizações de (2) e (3) foram dadas por Rathie [14].

Definição 1.4. - $H_n(p_1,\dots p_n;\beta_1,\dots,\beta_n//q_1,\dots,q_n)=$ (4)

$$-\Sigma p_k^{\beta_k}\log q_k/\Sigma p_k;$$ $P,Q\varepsilon\Gamma_n$ e $\beta_1,\dots,\beta_n$ são n-parâmetros.

Caso particular:

Para $\beta_1=\beta_2=\dots=\beta_n=1$, (4) reduz-se para (1).

Definição 1.5. - Para $P,Q\varepsilon\Gamma_n$, $0<\alpha\neq1$,

$$H_{n,\alpha}(P;\beta_1,\dots,\beta_n//Q)=(\alpha-1)^{-1}\log\Sigma\, p_k^{\beta_k}q_k^{\alpha-1}/\Sigma p_k^{\beta_k}. \quad (5)$$

Caso particular:

Para $\beta_1=\beta_2=\ldots=\beta_n=1$, (5) reduz-se para (3).

Sharma e Gupta [15] introduziram a imprecisão de ordem 1 e grau β e de ordem α e grau β respectivamente, da seguinte maneira:

$$H_n^*(P//Q;1,\beta) = A_\beta(2^{(\beta-1)\Sigma p_k \log q_k} - 1), \quad 0<\beta\neq 1 \quad \text{e}$$

$$H_n^*(P//Q;\alpha,\beta) = A_\beta\left[(\Sigma p_k q_k^{\alpha-1})^{\beta-1/\alpha-1} - 1\right], \quad 0<\alpha\neq 1 \text{ e } 0<\beta\neq 1;$$

onde α,β são parâmetros e $P,Q\varepsilon\Gamma_n$.

Observações: 1. - $\lim\limits_{\beta\to 1} H_n^*(P//Q;\alpha,\beta) = H_{n,\alpha}(P//Q)$.

2. - Quando $\alpha=\beta\neq 1$, $H_n^*(P//Q;\alpha,\beta) = H_n^\beta(P//Q)$.

3. - $\lim\limits_{\alpha\to 1}\lim\limits_{\beta\to 1} H_n^*(p//Q;\alpha,\beta) = H_n(P//Q)$.

No próximo capítulo estudaremos estas medidas de imprecisão para variáveis aleatórias generalizadas.

CAPÍTULO II - ALGUMAS GENERALIZAÇÕES DE IMPRECISÃO PARA VARIÁVEIS ALEATÓRIAS GENERALIZADAS.

2.1. - Introdução:

Neste capítulo apresentaremos algumas medidas de imprecisão para variáveis aleatórias generalizadas.

Existem situações onde alguns eventos não são observáveis ou são perdidos. Usando este conceito Renyi [13] introduziu o conceito das variáveis aleatórias incompletas e generalizadas. Se alguns eventos, em certas situações, não são observáveis, então a variável aleatória resultante, se finita e discreta, será incompleta no sentido que terá menos valores que anteriormente (tal variável aleatória poderia ser assumida teoricamente), e portanto, a soma das probabilidades associadas será estritamente menor que 1. Uma variável aleatória que é completa ou incompleta é chamada uma variável aleatória generalizada.

Seja $X = (x_1,\dots,x_n)$ uma variável aleatória generalizada com distribuição de probabilidade $P = (p_1,\dots,p_n)$; $p_k \geq 0$ e $\Sigma p_k \leq 1$.
A equação

$$W_1(P) = \Sigma p_k \tag{2.1}$$

é denominada peso de P ou distribuição de X. Claramente, para uma dada distribuição generalizada P se tem que $W_1(P) \leq 1$.

Se $W_1(P)=1$ a variável aleatória X é dita variável aleatória completa e sua correspondente distribuição denominada distribuição de probabilidade completa.

$$\text{Sejam: } \Delta_n = \{P: P=(p_1,\dots,p_n);\ p_k > 0 \text{ e } \Sigma p_k \leq 1\},\quad n \geq 1, \quad \text{e}$$

$$\Delta_n^* = \{P: P=(p_1,\dots,p_n);\ p_k > 0 \text{ e } \Sigma p_k = 1\},\quad n \geq 1.$$

Obviamente $\Delta_n^* \subset \Delta_n$, $n=1,2,\dots$. Denotando por (X) a distribuição generalizada com elementos x, $0 < x \leq 1$, claramente Δ_1^* consiste em apenas um elemento. Mais precisamente temos a distribuição singular completa (1).

A seguir apresentaremos algumas definições que precisaremos para o desenvolvimento deste capítulo.

Sejam, para $n \geq 2$, $P=(p_1,p_2,\ldots,p_n)$ e $Q=(q_1,q_2,\ldots q_n) \varepsilon \Delta_n$.

Definição 2.1. - P e Q são ditas equivalentes, anotamos "$P \sim Q$", se os elementos de P são permutações dos elementos de Q.

Observemos que $P \sim P$; se $P \sim Q$ então $Q \sim P$ e que se $P \sim Q$ e $Q \sim R$ então $P \sim R$, onde $R=(r_1,r_2,\ldots,r_n) \varepsilon \Delta_n$, $n \geq 2$. Desta forma "$\sim$" é uma relação de equivalência sobre Δ_n e assim Δ_n pode ser decomposto em classes de equivalência disjuntas.

Se $E(P)=\{Q:P \sim Q\}$ e $E^*(P)=\{Q:W_1(P)=W_1(Q)\}$, então $E(P) \subset E^*(P)$ porque se $P \sim Q$ então $W_1(P)=W_1(Q)$, mas a recíproca não é verdadeira.

Por exemplo: Se $P=(1/3,1/2,1/6)$ e $Q=(1/6,1/3,1/2)$, então $W_1(P)=1=W_1(Q)$, com $P \sim Q$. Agora, se $P=(1/3,1/3,1/6)$ e $Q=(1/2,1/6,1/6)$ então $W_1(P)=5/6=W_1(Q)$, embora $P \not\sim Q$.

Definição 2.2 - P é dita dominada por Q, anotemos "$P<<Q$", se existe uma permutação $k(1), k(2),\ldots,k(n)$ de $1,2,\ldots,n$ tal que $p_j \leq q_{k(j)}$, $j=1,\ldots,n$.

É óbvio que se $P<<Q$ então $W_1(P) \leq W_1(Q)$, mas a recíproca não é válida. Observemos também que $P<<P$, $P<<Q$ não implica em $Q<<P$, e que se $P<<Q$ e $Q<<R$ então $P<<R$, sendo portanto Δ_n parcialmente ordenado por "$<<$" para $n=2,3,\ldots$.

Analisemos os exemplos:

(a) Se $P=(1/6,1/4,1/8)$ e $Q=(1/5,1/3,1/7)$, como $1/6<1/5$, $1/4<1/3$ e $1/8<1/7$, segue que $P<<Q$ e também que $W_1(P)=13/24<71/105=W_1(Q)$.

(b) Se $P=(1/4,1/2)$ e $Q=(5/12,5/12)$; como $1/4<5/12$ mas $1/2>5/12$ então $P \not<< Q$ (nem $Q<<P$), porém $W_1(P)=3/4<10/12=W_1(Q)$; isto é, existem distribuições que não são dominadas uma pela outra.

Sejam $G(P)=\{Q:P<<Q\}$ e $G^*(P)=\{Q:W_1(P) \leq W_1(Q)\}$, claramente:

(I) $G(P) \subset G^*(P)$, pois se $P<<Q$ então $W_1(P) \leq W_1(Q)$;

(II) $E^*(P) \subset G^*(P)$, pois se, $Q \varepsilon E^*(P)$ então $W_1(Q)=W_1(P)$ e daí $W_1(Q) \leq W_1(P)$ de onde $Q \varepsilon G^*(P)$;

(III) $E^*(P) \not\subset G(P)$; por exemplo: Se $P=(1/2,1/3)$ e $Q=(1/3,1/2)$, então $Q \in E^*(P)$ (também a $E(P)$), mas $Q \notin G(P)$, pois $P \not\prec\prec Q$ (também $Q \not\prec\prec P$).

Definição 2.3. - P é dita r-dominada por Q, anotemos "$P \overset{r}{\propto} Q$", se para qualquer real $r>0$, $M_r(P) \leq M_r(Q)$, onde $M_r(P)=\Sigma p_k^r$.

A medida $M_r(P)$ é chamada r-ésima soma ordenada dos elementos de P.

Claramente se $P \sim Q$ então $M_r(P)=M_r(Q)$, mas a recíproca não é válida. Também se $P \prec\prec Q$ então $P \overset{1}{\propto} Q$ e a recíproca não é válida.

Observações: (a) Se $P \sim Q$, então os elementos de P e Q são os mesmos, porém, possivelmente permutados. Desta forma $\Sigma p_k^r = \Sigma q_k^r$, ou seja, $M_r(P)=M_r(Q)$.

Observemos porém que, se $P=(1/3,1/4)$, $Q=(\sqrt{5}/12,\sqrt{5}/6)$ e $r=2$, então: $M_r(P)=M_2(P)=(1/3)^2+(1/4)^2=25/144=(\sqrt{5}/12)^2+(\sqrt{5}/6)^2=M_2(Q)$, mas $P \not\sim Q$.

(b) Se $P \prec\prec Q$, então $p_k \leq q_k$ para cada $k=1,2,\ldots,n$, assim $\Sigma p_k \leq \Sigma q_k$, ou seja, $P \overset{1}{\propto} Q$. Agora se, por exemplo: $P=(1/3,1/4)$ e $Q=(1/2,1/3)$ então $M_1(P)=7/12<5/6=M_1(Q)$ e $P \overset{1}{\propto} Q$ embora $P \not\prec\prec Q$.

Consideremos que $P \overset{r}{\sim} Q$ designe $M_r(P)=M_r(Q)$. Vemos então que qualquer que seja r, $P \overset{r}{\sim} P$, se $P \overset{r}{\sim} Q$ então $Q \overset{r}{\sim} P$ e se $P \overset{r}{\sim} Q$ e $Q \overset{r}{\sim} R$ então $P \overset{r}{\sim} R$, ou seja, "$\overset{r}{\sim}$" é uma relação de equivalência sobre Δ_n e daí Δ_n pode ser decomposto em classes de equivalência disjuntas.

Se $E_r(P)=\{Q: P \overset{r}{\sim} Q\}$, $E_r^*(P)=\{Q: P \overset{r}{\propto} Q\}$, então $E_r(P) \subset E_r^*(P)$, pois se $Q \in E_r(P)$ então $\Sigma p_k^r = \Sigma q_k^r$ e daí $M_r(P) \leq M_r(Q)$ de onde $Q \in E_r^*(P)$.

Por outro lado notemos que se $P \sim Q$ então $P \overset{1}{\sim} Q$ mas a recíproca não é válida.

Por exemplo: Se $P=(1/6,1/3,1/2)$ e $Q=(1/3,1/3,1/3)$, temos que $M_1(P)=1/6+1/3+1/2=1=1/3+1/3+1/3=M_1(Q)$, ou seja, $P \overset{1}{\sim} Q$, mas $P \not\sim Q$.

Vale a pena observar ainda que para dois distintos valores de r, r_1 e r_2, não é necessário que $E_{r_1}(P)$ e $E_{r_2}(P)$ sejam distintos.

Por exemplo: Se $P=(1/6,1/6,2/3)$ e $Q=(1/14,2/7,9/14)$, então $M_1(P)=1/6+1/6+2/3=1=1/14+2/7+9/14=M_1(Q)$ e daí $Q\varepsilon E_1(P)$. Mas, por outro lado, $M_2(P)=(1/6)^2+(1/6)^2+(2/3)^2=1/2=(1/14)^2+(2/7)^2+(9/14)^2=$ $M_2(Q)$ e também que $Q\varepsilon E_2(P)$. Assim $E_1(P)$ e $E_2(P)$ não são disjuntos.

Este exemplo mostra também que se definirmos $d_r(P,Q)=$ $|M_r(P)-M_r(Q)|$ então (Δ_n,d_n) $n\geqslant 2$ são pseudo-métricas.

Observemos que:

(i) $d_r(P,Q)\geqslant 0$;

(ii) $d_r(P,Q)=|M_r(P)-M_r(Q)|=|M_r(Q)-M_r(P)|=$ $d_r(Q,P)$; e

(iii) $d_r(P,Q)+d_r(Q,R)=|M_r(P)-M_r(Q)|+$ $|M_r(Q)-M_r(R)| \geqslant |[M_r(P)-M_r(Q)]+$ $[M_r(Q)-M_r(R)]| = |M_r(P)-M_r(R)|=$ $d_r(P,R)$.

Porém $d_r(P,Q)=0$ não implica necessariamente em que $P=Q$, pois $d_r(P,Q)=0$ implica em $M_r(P)=M_r(Q)$ e daí que $P\neq Q$. (Observemos P e Q do exemplo anterior).

2.2. - Entropia e Imprecisão:

Apresentaremos agora a extensão do conceito de imprecisão (definida no capítulo I) para distribuições de probabilidade generalizadas.

Definição 2.4. - $H_\alpha(p//q)=\log 1/q$, $p\varepsilon\Delta_1$, $q\varepsilon\Delta_1$ e $\alpha>0$. (2.2)

Definição 2.5. - Imprecisão de ordem α.

$$H_\alpha(P//Q)=(1-\alpha)^{-1}\log(\Sigma p_k q_k^{\alpha-1}/\Sigma p_k),\quad 0<\alpha\neq 1,\ P\varepsilon\Delta_n,$$

$Q\varepsilon\Delta_n$ e $n=2,3,\ldots$. (2.3)

Definição 2.6. $H_\alpha(P//Q)=-\Sigma p_k\log(q_k/\Sigma p_k)$, $\alpha=1$, $P\varepsilon\Delta_n$, $Q\varepsilon\Delta_n$ e $n=2,3,\ldots$. (2.4)

Observação: Nota-se claramente que $H_\alpha(p//q)$ independe do parâmetro α.

Definição 2.7. - $H_\alpha(P) = H_\alpha(p//p) = \log 1/p$, onde $p\varepsilon\Delta_1$ e $\alpha>0$. (2.5)

Definição 2.8. - $H_\alpha(P//P) = H_\alpha(P)$, onde:

$$H_\alpha(P)=(1-\alpha)^{-1}\log(\Sigma p_k^\alpha/\Sigma p_k),\ 1\neq\alpha>0,\ P\varepsilon\Delta_n,\ n=2,3,\ldots \quad (2.6)$$

Observação: A medida $H_\alpha(P)$ é chamada entropia de Renyi [13].

Definição 2.9. - $H_1(P) = -\Sigma p_k \log p_k/\Sigma p_k$, $P\varepsilon\Delta_n$ e $n=2,3,\ldots$. (2.7)

Observação: $H_1(P)$ é chamada entropia de Shannon.

Definição 2.10. - Desigualdade de Jensen.

Seja $f(x)$ uma função côncava e contínua definida em $[0,1]$. Então para quaisquer $x_1,x_2,\ldots,x_n\varepsilon[0,1]$ e qualquer que seja o conjunto dos números $\lambda_1,\lambda_2,\ldots,\lambda_n$, reais e não negativos tais que $\Sigma\lambda_k=1$, temos:

$$\Sigma\lambda_k f(x_k) \leqslant f(\Sigma\lambda_k x_k).$$

Para funções convexas, temos:

$$\Sigma\lambda_k f(x_k) \geqslant f(\Sigma\lambda_k x_k).$$

Obviamente:

Se $P\sim Q$ então $H_\alpha(P) = H_\alpha(Q)$, $\alpha>0$, $P,Q\varepsilon\Delta_n$.

Se $P\overset{\alpha}{\sim}Q$ então $H_\alpha(P) = H_\alpha(Q)$, $0<\alpha\neq 1$ e $P,Q\ \varepsilon\Delta_n^*$, isto é;

Se $P,Q\varepsilon\Delta_n^*$ então $\Sigma p_k=1=\Sigma q_k$.

Se $P\overset{\alpha}{\sim}Q$ então $M_\alpha(P)=M_\alpha(Q)$ e daí $\Sigma p_k^\alpha=\Sigma q_k^\alpha$. Portanto, $\Sigma p_k^\alpha/\Sigma p_k = \Sigma q_k^\alpha/\Sigma q_k$ assim $\log(\Sigma p_k^\alpha/\Sigma p_k)= \log(\Sigma q_k^\alpha/\Sigma q_k)$, de onde $(1-\alpha)^{-1}\log(\Sigma p_k^\alpha/\Sigma p_k) = (1-\alpha)^{-1}\log(\Sigma q_k^\alpha/\Sigma q_k)$, e finalmente:

$$H_\alpha(P) = H_\alpha(Q).$$

Um fato de conhecimento geral em Teoria da Informação é que se $P\varepsilon\Delta_n^*$, $Q\varepsilon\Delta_n^*$, $n=1,2,\ldots$, então:

$$H_1(P) \leqslant H_1(P//Q) \tag{2.8}$$

Vemos que se $p_k<q_k$, $k=1,2,\ldots,n$, então para $P\varepsilon\Delta_n$ e $Q\varepsilon\Delta_n$, $n=1,2,\ldots$, temos:

$$H_1(P) = -\ \Sigma p_k \log(p_k)/\Sigma p_k > -\Sigma p_k \log(q_k)/\Sigma p_q = H_1(P//Q), \quad \text{e}$$

portanto (2.8) não vale. Contudo se $P\varepsilon G^*(Q)$ então:

$$W_1(Q) \leqslant W_1(P) \Rightarrow \Sigma q_k \leqslant \Sigma p_k \Rightarrow \Sigma p_k/\Sigma q_k \geqslant 1.$$

Mas,

$$\Sigma(p_k/q_k) \geqslant \Sigma p_k/\Sigma q_k \geqslant 1.$$

Desta forma:

$$\log\Sigma(p_k/q_k) \geqslant 0$$

Agora, sendo $p_k>0$,

$$p_k \log\Sigma(p_k/q_k) \geqslant 0 \Rightarrow \log\left[\Sigma(p_k/q_k)\right]^{p_k} \geqslant 0$$

e daí, pela definição 2.10,

$$\Sigma p_k \log(p_k/q_k) \geqslant \log\left[\Sigma(p_k/q_k)\right]^{p_k} \geqslant 0$$

assim:

$$\Sigma p_k \log p_k/q_k \geqslant 0 \Rightarrow \Sigma p_k \log\, p_k - \Sigma p_k \log\, q_k \geqslant 0$$

portanto:

$$\Sigma p_k \log\, p_k \geqslant \Sigma p_k \log\, q_k \Rightarrow -\ \Sigma p_k \log\, p_k \leqslant -\ \Sigma p_k \log\, q_k$$

então:

$$H_1(P) \leqslant H_1(P//Q).$$

Por outro lado a desigualdade:

$$H_\alpha(P) \leqslant H_\alpha(P//Q), \quad P\varepsilon\Delta_n,\ Q\varepsilon\Delta_n,\ n\geqslant 2 \tag{2.9}$$

nem sempre é válida. Arora [2] provou que (2.9) vale quando $P\sim Q$.

Exemplo: Se $P=(1/2;1/3)$ e $Q=(1/2;1/2)$ para $\alpha=2$, temos que:

$H_\alpha(P) = \log 30/13 > \log 2 = H_\alpha(P//Q)$, isto mostra que ainda que $Q\varepsilon\Delta_n^*$, como $P\varepsilon\Delta_n$ (2.9) não é necessariamente válida.

Consideremos o problema de determinar condições sob as quais (2.9) é válida. Para determinar estas condições definimos o seguinte:

Definição 2.11.- Desigualdade de Hölder [3].

$$\Sigma a_k b_k \geqslant (\Sigma a_k^{1/\lambda})^\lambda (\Sigma b_k^{1/(1-\lambda)})^{1-\lambda}, \quad \text{onde } a_k,\ b_k>0 \text{ e } \lambda>1,$$

com a igualdade se, e somente se, para algum c, $a_k^{1-\lambda} = b_k^\lambda c$, para qualquer $k=1,2,\ldots,n$. Para $\lambda<1$ a desigualdade se inverte.

Na definição (2.11) substituindo $a_k=p_k$, $b_k=q_k^{\alpha-1}$ e $\lambda=1/\alpha$ temos que:

$$\Sigma p_k q_k^{\alpha-1} \leqslant \left[M_\alpha(P)\right]^{1/\alpha}\left[M_\alpha(Q)\right]^{(\alpha-1)/\alpha}, \quad \alpha>1. \tag{2.10}$$

para $\alpha<1$, (2.10) se inverte.

Assim, se $P\varepsilon E^*(Q)$ $\Sigma q_k^\alpha \leqslant \Sigma p_k^\alpha$, para $1\neq\alpha>0$, usando-se (2.10), temos que:

(a) para $\alpha>1$.

$\Sigma p_k q_k^{\alpha-1} \leqslant (\Sigma p_k^\alpha)^{1/\alpha}(\Sigma q_k^\alpha)^{(\alpha-1)/\alpha} \leqslant \Sigma p_k^\alpha$, tomando-se log em ambos os membros da desigualdade, vem:

$\log(\Sigma p_k q_k^{\alpha-1}) \leqslant \log(\Sigma p_k^\alpha)$; multiplicando-se por $1/1-\alpha<0$

vem que:

$1/(1-\alpha)\log(\Sigma p_k q_k^{\alpha-1}) \geqslant 1/(1-\alpha)\log(\Sigma p_k^\alpha)$, isto é, que

$$H_\alpha(P) \leqslant H_\alpha(P//Q)$$

(b) Para $\alpha<1$:

$\Sigma p_k q_k^{\alpha-1} \geqslant (\Sigma p_k^\alpha)^{1/\alpha}(\Sigma q_k^\alpha)^{(\alpha-1)/\alpha}$, como $(\alpha-1)/\alpha<0$, vem que:

$$\Sigma p_k q_k^{\alpha-1} \geqslant (\Sigma p_k)^{1/\alpha}(\Sigma p_k^{\alpha})^{(\alpha-1)/\alpha} = \Sigma p_k^{\alpha},$$ daí a demonstração

segue análoga a anterior.

Observação: Quando $P\varepsilon\Delta_n^*, Q\varepsilon\Delta_n^*, \alpha=1$ (2.8) é a própria desigualdade de Shannon, cuja importância é bastante reconhecida.

A desigualdade (2.9) é uma generalização da desigualdade de Shannon para distribuições generalizadas. A condição $P\varepsilon E_\alpha^*(Q)$, onde $1\neq\alpha>0$ é necessária mas não é suficiente. Observemos o exemplo:

Se $P=(1/2;1/2)$ e $Q=(1/4;3/4)$ então $M_2(P)=8/16$ e $M_2(Q)=10/16$ e assim $Q\varepsilon E_2^*(P)$, mas $H_\alpha(P)=1$ bit e também $H_2(P)=1$ bit. Desta forma temos que, para um α fixo em (2.9)a igualdade pode valer mesmo quando $P\neq Q$. Se, para os mesmos P e Q usarmos $\alpha=3/2$, temos:

$M_{3/2}(P) = 0{,}707$ e $M_{3/2}(Q) = 0{,}7745$, desta forma $Q\varepsilon E_\alpha^*(P)$ pois:

$$M_{3/2}(P) < M_{3/2}(Q), \quad \text{mas} \quad H_{3/2}(P) < H_{3/2}(P//Q).$$

2.3. - Característica monotônica da Imprecisão de Ordem α.

Se $p\varepsilon\Delta_1$ e $q\varepsilon\Delta_1$, então é óbvio que $H_\alpha(p//q)$ é independente de α, isto é;

$$H_\alpha(p//q) = (1-\alpha)^{-1}\log(pq^{\alpha-1}/p) = (1-\alpha)^{-1}(1-\alpha)\log(1/q)$$
$$= \log(1/q).$$

Desta forma, a questão da discussão de característica monotônica da imprecisão de ordem α com relação a α não pode acontecer quando as distribuições em consideração são distribuições generalizadas simples.

Além disso, se $P\varepsilon\Delta_n$ e $Q\varepsilon\Delta_n$, $n\geqslant 2$, existe a possibilidade de $H_\alpha(P//Q)$ ser independente de α. Por exemplo:

Para qualquer P, se $Q=(W_1(Q)/n,\ldots,W_1(Q)/n)$ temos que:

$$H_\alpha(P//Q) = (1-\alpha)^{-1}\log(\Sigma p_k(W_1(Q)/n)^{\alpha-1}/\Sigma p_k)$$
$$= (1-\alpha)^{-1}\log\left[(W_1(Q)^{\alpha-1}/n)(\Sigma p_k/\Sigma p_k)\right]$$

$$= (1-\alpha)^{-1}\log(W_1(Q)/n)^{\alpha-1}$$

$$= (1-\alpha)^{-1}(\alpha-1)\log(W_1(Q)/n)$$

$$= -\log(W_1(Q)/n)$$

$$= \log(n/W_1(Q)),$$ que é independente de α.

Agora, se usarmos a definição 2.10, poderemos observar que:

$$\frac{d}{d\alpha}(H_\alpha(P//Q)) \leqslant 0 \qquad (2.11)$$

De fato:

$$\frac{d}{d\alpha}(H_\alpha(P//Q)) = \frac{1}{(1-\alpha)^2}\Big[\log(\Sigma p_k q_k^{\alpha-1}/\Sigma p_k) + (\Sigma p_k q_k^{\alpha-1}\log(q_k^{1-\alpha}))/(\Sigma p_k q_k^{\alpha-1})\Big],$$

desde que, pela definição 2.10,

$$\begin{aligned} A &= (\Sigma p_k q_k^{\alpha-1}\log(q_k^{1-\alpha}))/(\Sigma p_k q_k^{\alpha-1}) \\ &\leqslant \log((\Sigma p_k q_k^{1-\alpha} q_k^{\alpha-1})/(\Sigma p_k q_k^{\alpha-1})) \\ &\leqslant \log((\Sigma p_k)/\Sigma p_k q_k^{\alpha-1})) = B, \end{aligned}$$

temos que:

$$\frac{d}{d\alpha}(H_\alpha(P//Q) = 1/(1-\alpha)^2(-B+A) \leqslant 0.$$

Observemos que:

$\frac{d}{d\alpha}(H_\alpha(P//Q) = 0$ se $Q = (W_1(Q)/n,\dots,W_1(Q)/n)$ já que, neste caso, $H_\alpha(P//Q)$ independe de α.

Pelo que acima mostramos, temos que $H_\alpha(P//Q)$ é uma função monotonicamente decrescente de α.

Por conveniência, admitamos que Q não é uma distribuição retangular, isto é;

$$Q \neq (W_1(Q)/n,\dots,W_1(Q)/n).$$

Então:

$$\alpha_1 < \alpha_2 \Rightarrow H_{\alpha_2}(P//Q) < H_{\alpha_1}(P//Q).$$

A seguir, analisemos o seguinte problema:

É possível encontrar um número real $x>1$, tal que:

$$1 < \frac{H_{\alpha_1}(P//Q)}{H_{\alpha_2}(P//Q)} < x, \quad \text{sendo} \quad 0<\alpha_1<\alpha_2<\infty \tag{2.12}$$

O teorema que segue responde a questão para $1<\alpha_1<\alpha_2<\infty$. A resposta ainda não é conhecida para o caso $0<\alpha_1<\alpha_2<1$.

Teorema 1.4. - Se $P\varepsilon\Delta_n$, $Q\varepsilon\Delta_n$, $n=2,3,\ldots$ e Q não é uma distribuição retangular, então:

$$1 < \frac{H_{\alpha_1}(P//Q)}{H_{\alpha_2}(P//Q)} < \frac{\alpha_2-1}{\alpha_1-1}, \qquad 1<\alpha_1<\alpha_2<\infty. \tag{2.13}$$

Demonstração:

Dados $P=(p_1,p_2,\ldots,p_n)$ e $Q=(q_1,q_2,\ldots,q_n)$ é possível construir um número infinito não enumerável de distribuições:

$$R_m=(r_1(m),r_2(m),\ldots,r_n(m)) \quad \text{onde} \quad r_k(m)=(p_kq_k^m)/(\Sigma p_kq_k^m),\ m>0.$$

Notemos que:

$$W_1(R_m)=\Sigma r_k(m)=(\Sigma p_kq_k^m)/(\Sigma p_kq_k^m)=1 \text{ e portanto que } R_m\varepsilon\Delta_n^*.$$

Observemos que:

$$H_\alpha(R_m//Q) = (1-\alpha)^{-1}\log \frac{(\Sigma p_kq_k^m/\Sigma p_kq_k^m)q_k^{\alpha-1}}{\Sigma p_kq_k^m/\Sigma p_kq_k^m}$$

$$= (1-\alpha)^{-1}\log(\Sigma p_kq_k^{\alpha+m-1}/\Sigma p_kq_k^m)$$

$$= (1-\alpha)^{-1}\log \frac{\Sigma(p_kq_k^{\alpha+m-1}/\Sigma p_k)}{\Sigma(p_kq_k^m/\Sigma p_k)}$$

$$= (1-\alpha)^{-1}\,\frac{(1-\alpha-m)}{(1-\alpha-m)}\,\log(\Sigma p_kq_k^{\alpha+m-1}/\Sigma p_k) -$$

$$(1-\alpha)^{-1}(m/m)\log\,(\Sigma p_kq_k^m/\Sigma p_k)$$

$$= \frac{(1-\alpha-m)}{(1-\alpha)} H_{\alpha+m}(P//Q) + \frac{m}{(1-\alpha)} H_{m+1}(P//Q) \quad (2.14)$$

Desde que $H_\alpha(R_m//Q) > 0$, segue, para $1<\alpha<\infty$ e $m>0$ que:

$$\frac{1-\alpha-m}{1-\alpha} H_{\alpha+m}(P//Q) + \frac{m}{(1-\alpha)} H_{m+1}(P//Q) > 0$$

$$\frac{m}{1-\alpha} H_{m+1}(P//Q) > -\frac{1-\alpha-m}{1-\alpha} H_{\alpha+m}(P//Q)$$

$$\frac{m}{\alpha-1} H_{m+1}(P//Q) < \frac{m+\alpha-1}{\alpha-1} H_{\alpha+m}(P//Q)$$

$$\frac{H_{m+1}(P//Q)}{H_{\alpha+m}(P//Q)} < \frac{m+\alpha-1}{m} \quad (2.15)$$

Colocando $\alpha_1=m+1$ e $\alpha_2=\alpha+m$, o resultado segue.

Demonstração alternativa: O teorema acima também poderia ser provado de outra forma, como mostramos a seguir.

Desde que: $1<\alpha_1<\alpha_2$ $q^{\alpha_1}>q^{\alpha_2}$, $0<q<1$, segue que:

$$p_k q_k^{\alpha_1-1} > p_k q_k^{\alpha_2-1} \qquad \text{daí:}$$

$$\Sigma p_k q_k^{\alpha_1-1} > \Sigma p_k q_k^{\alpha_2-1}$$

$$\frac{\Sigma p_k q_k^{\alpha_1-1}}{\Sigma p_k} > \frac{\Sigma p_k q_k^{\alpha_2-1}}{\Sigma p_k}$$

$$\log \frac{\Sigma p_k q_k^{\alpha_1-1}}{\Sigma p_k} > \log \frac{\Sigma p_k q_k^{\alpha_2-1}}{\Sigma p_k}$$

$$\frac{\alpha_1-1}{\alpha_1-1}\log\frac{\Sigma p_k q_k^{\alpha_1-1}}{\Sigma p_k} > \frac{\alpha_2-1}{\alpha_2-1}\log\frac{\Sigma p_k q_k^{\alpha_2-1}}{\Sigma p_k}$$

$$(1-\alpha_1)H_{\alpha_1}(P//Q) > (1-\alpha_2)H_{\alpha_2}(P//Q),$$

como $\alpha_1,\alpha_2>1$, vem que:

$$\frac{H_{\alpha_1}(P//Q)}{H_{\alpha_2}(P//Q)} < \frac{1-\alpha_2}{1-\alpha_1}$$

ou

$$\frac{H_{\alpha_1}(P//Q)}{H_{\alpha_2}(P//Q)} < \frac{\alpha_2-1}{\alpha_1-1}.$$

Corolário: Dado $x>1$, é também possível ter α_1 e α_2 ambos maiores que 1 dependendo de x, tais que:

$$1 < \frac{H_{\alpha_1}(P//Q)}{H_{\alpha_2}(P//Q)} < x.$$

Demonstração:

Desde que $x>1$ é sempre possível obter $y\neq 0$ tal que $x=1+(1/y)$.

Seja $c>0$, constante arbitrária. Definimos:

$\alpha_1= cy+1$ e $\alpha_2= c(y+1)+1$.

Então:

$$\frac{H_{cy+1}(P//Q)}{H_{c(y+1)+1}(P//Q)} = \frac{(-cy)^{-1}\log(\Sigma p_k q_k^{cy}/\Sigma p_k)}{\left[-c(y+1)\right]^{-1}\log(\Sigma p_k q_k^{c(y+1)}/\Sigma p_k)}$$

$$= \frac{c(y+1)}{cy}\log\frac{\Sigma p_k q_k^{cy}/\Sigma p_k}{\Sigma p_k q_k^{c(y+1)}/\Sigma p_k}$$

$$= \frac{y+1}{y} \log \frac{\Sigma p_k q_k^{cy}}{\Sigma p_k q_k^{c(y+1)}}$$

mas, desde que $0<q_k<1$, para cada $k=2,3,\ldots,n$, $(q_k^{cy}/q_k^{c(y+1)})>1$, daí

$$1 < \frac{H_{\alpha_1}(P//Q)}{H_{\alpha_2}(P//Q)} < \frac{y+1}{y} = x, \quad \text{onde} \quad 1<\alpha_1<\alpha_2.$$

É óbvio que as escolhas de α_1 e α_2 não são únicas. Ao contrário, podem-se escolher infinitos pares não enumeráveis (α_1,α_2). Porém é também claro que os valores escolhidos de α_1 e α_2 diferem para cada $c>0$, cuja escolha é arbitrária.

Observações:

I) Na demonstração do Teorema 2.4. vimos que $\alpha_2-\alpha_1=(\alpha+m-(m+1))=\alpha-1$. Assim especificados α_1 e α_2 a escolha de α em (2.11) é unicamente determinada. Um vez conhecido α, o correspondente m pode ser encontrado porque α_2 é dado e $m=\alpha_2-\alpha$. Desta forma as escolhas de α e m não são arbitrárias mas dependem dos valores atribuidos a α_1 e α_2.

II) O limite superior em (2.13) é independente de P e Q. Assim a identificação de Q com P resolve o problema correspondente para a Entropia de Renyi. Além disso, no caso da Entropia de Renyi, o limite superior em (2.13) pode ser refinado e um melhor limite superior foi estabelecido em [8].

III) Conforme já observamos é difícil conseguir um limite superior conveniente quando $0<\alpha_1<\alpha_2<1$. Contudo, é possível conhecer a forma com que imprecisões de ordem maiores que a unidade se relacionam com as outras, por aplicação da desigualdade de Cauchy.

Definição 2.12. - Desigualdade de Cauchy:

$$(\Sigma a_k b_k)^2 \leqslant (\Sigma a_k^2)(\Sigma b_k^2)$$

onde $a_k,b_k>0$. Com a igualdade se, e somente se, para algum c, $a_k^{1/2}=b_k^{1/2}c_k$ para qualquer $k=1,2,\ldots,n$.

Substituindo: $a_k=p_k^{1/2}q_k^{(\alpha-1)/2}$, $b_k=p_k^{1/2}q_k^{(m-1)/2}$ na definição 2.12, temos que:

$$(\Sigma p_k q_k^{(\alpha+m-1)/2})^2 \leqslant (\Sigma p_k q_k^{\alpha-1})(\Sigma p_k q_k^{m-1}) \qquad \text{(A)}$$

Por definição:

$$H_{(\alpha+m)/2}(P//Q) = (1-\frac{\alpha+m}{2})^{-1}\log\frac{\Sigma p_k q_k^{(\alpha+m)/2-1}}{\Sigma p_k}$$

$$= \frac{2}{2-\alpha-m}\log\frac{\Sigma p_k q_k^{(\alpha+m)/2-1}}{\Sigma p_k}$$

$$= \frac{1}{2-\alpha-m}\log\left(\frac{\Sigma p_k q_k^{(\alpha+m)/2-1}}{\Sigma p_k}\right)^2$$

usando (A)

$$\leqslant \frac{\log\left[(\Sigma p_k q_k^{\alpha-1}/\Sigma p_k)(\Sigma p_k q_k^{m-1}/\Sigma p_k)\right]}{2-\alpha-m}$$

$$= \frac{\frac{1-\alpha}{1-\alpha}\log(\Sigma p_k q_k^{\alpha-1}/\Sigma p_k) + \frac{1-m}{1-m}\log(\Sigma p_k q_k^{m-1}/\Sigma p_k)}{2-\alpha-m}$$

com $0<\alpha<1$ e $0<m<1$.

Fazendo $Q=P$, o correspondente resultado para $H_\alpha(P)$ é:

$$H_{(\alpha+m)/2}(P) \leqslant \frac{(1-\alpha)\,H_\alpha(P) + (1-m)H_m(P)}{2-\alpha-m}.$$

2.4. - Imprecisão como um valor médio.

Seja ϕ uma função estritamente crescente monotônica e contínua para todo $x\varepsilon(-\infty,+\infty)$. Nath em [11] definiu ϕ-imprecisão como:

$$H_\phi(P//Q) = \phi^{-1}(\Sigma p_k \phi(\log(1/q_k))/\Sigma p_k)$$

$$= M_\phi(P;Q) \qquad (2.16)$$

e provou que as únicas formas de ϕ-imprecisão que são aditivas são aquelas que correspondem a:

$\phi(x) = ax+b$, $a\neq 0$ e b constantes arbitrárias

ou

$\phi(x) = \phi_\alpha(x)$ onde $\phi_\alpha(x) = a.2^{(1-\alpha)x}+b$, $a\neq 0$ e b constantes arbitrárias e $\alpha\neq 1$.

Usaremos ϕ escrita por ϕ_l ou ϕ_e conforme seja linear ou exponencial, respectivamente. Então:

$$H_{\alpha_l}(P//Q) = \alpha_l^{-1}(\Sigma p_k \alpha_l(\log(1/q_k))/\Sigma p_k),$$

sendo: $\alpha_l=\phi_{\alpha_l}(x)= ax+b$ e portanto:

$$\alpha_l^{-1} =\phi_{\alpha_l}^{-1}(x) = (x-b)/a, \text{ de onde,}$$

$$H_{\alpha_l}(P//Q) = \frac{(\Sigma p_k(a\log(1/q_k)+b)/\Sigma p_k)-b}{a}$$

$$= \frac{a\Sigma p_k \log(1/q_k)+b\Sigma p_k-b\Sigma p_k}{a\Sigma p_k}$$

$$= -\frac{\Sigma p_k \log q_k}{\Sigma p_k}$$

$$= H_1(P//Q).$$

e

$$H_{\alpha_e}(P//Q) = \alpha_e^{-1}(\Sigma p_k \alpha_e(\log(1/q_k))/\Sigma p_k), \ \alpha\neq 1, \text{ onde}$$

$\alpha_e=\phi_{\alpha_e}(x)=a.2^{(1-\alpha)x}+b$ e $\alpha_e^{-1}=\phi_{\alpha_e}^{-1}(x)=(1-\alpha)^{-1}\log((x-b)/a)$, daí:

$$H_{\alpha_e}(P//Q) = (1-\alpha)^{-1}\log \frac{(\Sigma p_k(a.2^{(1-\alpha)\log(1/q_k)}+b)/\Sigma p_k)-b}{a}$$

$$= (1-\alpha)^{-1}\log \frac{a\Sigma p_k(1/q_k)^{(1-\alpha)}+b\Sigma p_k-b\Sigma p_k}{a\Sigma p_k}$$

$$= (1-\alpha)^{-1}\log(\Sigma p_k q_k^{\alpha-1}/\Sigma p_k)$$

$$= H_\alpha(P//Q).$$

São de nosso interesse as ϕ-imprecisões que são homogêneas, isto é;

$$\phi^{-1}(\Sigma\rho_k\phi(m\log(1/q_k))) = m\phi^{-1}(\Sigma\rho_k\phi\log(1/q_k))),$$

com $\rho_k = p_k/\Sigma p_k$, $m>0$ (2.17)

Suponhamos que $\phi(1)=0$. Então de [1] segue que as únicas formas que ϕ pode assumir são:

$$\phi(x) = c\log x \qquad \text{ou} \qquad \phi(x) = ax^{\alpha-1}+b, \text{ onde } \alpha\neq 1, \text{e}$$

$a\neq 0$, b constantes arbitrárias. Esta observação mostra que a ϕ-imprecisão não pode ser simultaneamente aditiva e homogênea, exceto quando ϕ é linear.

Também:

$$M_\phi(P;Q) = 2^{\Sigma p_k\log(1/q_k)/\Sigma p_k}, \quad \text{quando } \phi(x) = c\log x. \tag{2.18}$$

$$K_\alpha(P;Q) = M_\phi(P;Q) = \left[\Sigma p_k(\log(1/q_k))^{\alpha-1}/\Sigma p_k\right]^{1/(\alpha-1)},$$

quando $\phi(x) = ax^{\alpha-1}+b$, $\alpha\neq 1$, $a\neq 0$ e b constantes arbitrárias.

(2.19)

Portanto $H_1(P//Q)$ é a única imprecisão que é ao mesmo tempo aditiva e homogênea. Notemos que:

$$K_2(P;Q) = \left[\Sigma p_k(\log(1/q_k))^{2-1}/\Sigma p_k\right]^{1/(2-1)}$$

$$= -\Sigma p_k \log q_k/\Sigma p_k$$

$$= H_1(P//Q).$$

Sejam:

$$H_{\phi_1}(P//Q;m) = \phi_1^{-1}(\Sigma\rho_k\phi_1(m\log(1/q_k))) \qquad (2.20)$$

e

$$H_{\phi_e}(P//Q;m) = \phi_e^{-1}(\Sigma\rho_k\phi_e(m\log(1/q_k))) \qquad (2.21)$$

onde $\rho_k=p_k/\Sigma p_k$. Temos que:

$$H_{\phi_1}(P//Q;m) = \left[\Sigma\rho_k(am\log(1/q_k)+b)-b\right]/a$$

$$= (ma\Sigma\rho_k\log(1/q_k)+b\Sigma\rho_k-b)/a$$

$$= (ma\Sigma p_k\log(1/q_k)/\Sigma p_k+b\Sigma p_k/\Sigma p_k-b)/a$$

$$= m.\ -(\Sigma p_k\log q_k)/\Sigma p_k$$

$$= m.\ H_1(P//Q)$$

Mas,

$$H_{\phi_e}(P//Q;m) \neq mH_{\phi_e}(P//Q) \qquad \text{(logo não é homogênea).}$$

Define-se [10]:

$$H_\alpha(P//Q;m) = (1-\alpha)^{-1}\log(\Sigma p_k q_k^{m\alpha-m}/\Sigma p_k), \quad 1\neq\alpha>0 \text{ e } m>0. \qquad (2.22)$$

Então temos que:

$$H_{\phi_1}((P//Q;m) = H_\alpha(P//Q;m) = mH_{m\alpha-m+1}(P//Q).$$

Usando a monotonicidade de $H_\alpha(P//Q)$ segue que

$$H_\alpha(P//Q;m) \leqq m\, H_{\alpha-m+1}(P//Q), \quad \text{se } m>1 \quad \text{e}$$

$$H_\alpha(P//Q;m) \geqq m\, H_{\alpha-m+1}(P//Q) \quad \text{se } 0<m<1.$$

A imprecisão $H_{\alpha-m+1}(P//Q)$ é definida para todo $\alpha>0$, contanto que $0<m<1$. Mas, se $m>1$, é definida para intervalos da forma $m-1<\alpha<m$ e $m<\alpha<\infty$ tais que $\alpha-m+1$ torna-se positivo.

A quantidade $H_\alpha(P//Q;m)$ pode ser considerada como uma medida aditiva de múltiplos positivos de imprecisão sendo que

$$\lim_{\alpha\to 1} H_\alpha(P//Q;m) = m\, H_1(P//Q)$$

e é um valor médio.

Em geral, as expressões para múltiplos positivos de imprecisões não são únicas. Por exemplo, consideremos:

$$H^*_\alpha(P//Q;m) = (1-\alpha)^{-1}\log\,(\Sigma p_k q_k^{\alpha^m-1}/\Sigma p_k), \quad 1\neq\alpha>0 \text{ e } m>0.$$

É claro que $H^*_\alpha(P//Q;m)$ é aditiva e, além disso,

$$\lim_{\alpha\to 1} H^*_\alpha(P//Q;m) = \lim_{\alpha\to 1}\,(-m(\alpha^m-1)\Sigma p_k q_k^{\alpha^m-1}\log q_k/\Sigma p_k q_k^{\alpha^m-1})$$

$$= -\,m\,\Sigma p_k \log q_k/\Sigma p_k$$

$$= m\, H_1(P//Q)$$

e também,

$$H^*_\alpha(P//Q;1) = (1-\alpha)^{-1}\log(\Sigma p_k q_k^{\alpha-1}/\Sigma p_k)$$

$$= H_\alpha(P//Q).$$

Mas $H^*_\alpha(P//Q;m)$ não pode ser visto como um valor médio. Observemos que, se $1\neq\alpha>0$ e $m>1$,

$$m(\alpha-1) < \alpha^m-1 \quad \text{e, para cada } k=1,2,\dots,n, \text{ vem que}$$

$$q_k^{m(\alpha-1)} < q_k^{\alpha^m-1}$$

$$p_k q_k^{m\alpha-m} < p_k q_k^{\alpha^m-1}$$

$$\frac{\Sigma p_k q_k^{m\alpha-m}}{\Sigma p_k} < \frac{\Sigma p_k q_k^{\alpha^m-1}}{\Sigma p_k}$$

$$\log\left(\frac{\Sigma p_k q_k^{m\alpha-m}}{\Sigma p_k}\right) < \log\left(\frac{\Sigma p_k q_k^{\alpha^m-1}}{\Sigma p_k}\right) \qquad \text{e ainda que,}$$

$$m(\alpha-1)\log\left(\frac{\Sigma p_k q_k^{m\alpha-m}}{\Sigma p_k}\right) < (\alpha^m-1)\log\left(\frac{\Sigma p_k q_k^{\alpha^m-1}}{\Sigma p_k}\right)$$

isto é, que:

$$\frac{\log(\Sigma p_k q_k^{m\alpha-m}/\Sigma p_k)}{\log(\Sigma p_k q_k^{\alpha^m-1}/\Sigma p_k)} < \frac{\alpha^m-1}{m(\alpha-1)}$$

$$\frac{(1-\alpha)^{-1}\log(\Sigma p_k q_k^{m\alpha-m}/\Sigma p_k)}{(1-\alpha)^{-1}\log(\Sigma p_k q_k^{\alpha^m-1}/\Sigma p_k)} < \frac{1-\alpha^m}{m(1-\alpha)}$$

e finalmente temos que:

$$\frac{H_\alpha(P//Q;m)}{H^*_\alpha(P//Q;m)} < \frac{1}{m}\cdot\frac{1-\alpha^m}{1-\alpha}\ , \quad 0<\alpha\neq1,\ m>1. \qquad (2.23)$$

Analogamente, se $1\neq\alpha>0$ e $0<m<1$, vem que

$\alpha^m-1 < m(\alpha-1)$, e com procedimento semelhante ao acima desenvolvido, que:

$$\frac{H_\alpha(P//Q;m)}{H^*_\alpha(P//Q;m)} < \frac{m(1-\alpha)}{1-\alpha^m}, \quad 0<\alpha\neq 1 \text{ e } 0<m<1. \qquad (2.24)$$

Também, se $W_1(P) = W_1(Q)$ e aplicando a definição 2.11, vem que:

$$H_\alpha(P) \leqq \alpha H_\alpha(P//Q) + (1-\alpha m)H_{\alpha m}(Q), \quad \alpha>0,\ \alpha\neq 1, \quad m>0.$$

Para obtermos este resultado substituimos, na definição 2.11, $a_k=p_k$, $b_k=q_k^{m(\alpha-1)}$ e $\lambda=1/\alpha$ e temos,

$$\Sigma p_k q_k^{m(\alpha-1)} \leqq (\Sigma p_k^\alpha)^{1/\alpha}(\Sigma q_k^{m(\alpha-1)\cdot\frac{\alpha}{\alpha-1}})^{\frac{\alpha-1}{\alpha}}$$

$$\log \Sigma p_k q_k^{m(\alpha-1)} \leqq \frac{1}{\alpha}\log(\Sigma p_k^\alpha)+\frac{(\alpha-1)}{\alpha}\log(\Sigma q_k^{m\alpha})$$

multiplicando ambos os lados da desigualdade por α, vem que:

$$\alpha\log \Sigma p_k q_k^{m(\alpha-1)} \leqq \log(\Sigma p_k^\alpha)+(\alpha-1)\log(\Sigma q_k^{m\alpha})$$

Para $\alpha>1$, temos que,

$$\alpha/(\alpha-1)\log \Sigma p_k q_k^{m(\alpha-1)} \leqq 1/(\alpha-1)\log(\Sigma p_k^\alpha)+\frac{1-\alpha m}{1-\alpha m}\log(\Sigma q_k^{m\alpha})$$

ou

$$1/(1-\alpha)\log(\Sigma p_k^\alpha) \leqq \alpha\,\frac{1}{1-\alpha}\log(\Sigma p_k q_k^{m(\alpha-1)})+\frac{1-\alpha m}{1-\alpha m}\log(\Sigma q_k^{m\alpha})$$

Como para distribuições generalizadas completas temos que $\Sigma p_k=\Sigma q_k=1$, então:

$$1/(1-\alpha)\log(\Sigma p_k^{\alpha}/\Sigma p_k) \leq \alpha.\frac{1}{1-\alpha}\log(\Sigma p_k q_k^{m(\alpha-1)}/\Sigma p_k)+$$

$$\frac{1-\alpha m}{1-\alpha m}\log(\Sigma q_k^{m\alpha}/\Sigma q_k)$$

isto é;

$$H_\alpha(P) \leq \alpha H_\alpha(P//Q;m) + (1-\alpha m)H_{\alpha m}(Q)$$

Agora, para distribuições incompletas, isto é, $\Sigma p_k = \Sigma q_k \leq 1$, temos que:

$$\frac{1}{1-\alpha}\log(\Sigma p_k^{\alpha}) \leq \frac{\alpha}{1-\alpha}\log(\Sigma p_k q_k^{m(\alpha-1)}) + \frac{1-\alpha m}{1-\alpha m}\log(\Sigma q_k^{\alpha m}) + A,$$

onde

$$A = \frac{1}{1-\alpha}\log(\Sigma p_k) - \frac{\alpha}{1-\alpha}\log(\Sigma p_k) - \log(\Sigma p_k) = 0,$$

logo:

$$\frac{1}{1-\alpha}\log(\Sigma p_k^{\alpha}) \leq \frac{\alpha}{1-\alpha}\log(\Sigma p_k q_k^{m(\alpha-1)}) + \frac{1-\alpha m}{1-\alpha m}\log(\Sigma q_k^{\alpha m}) +$$

$$\frac{1}{1-\alpha}\log(\Sigma p_k) - \frac{\alpha}{1-\alpha}\log(\Sigma p_k) - \log(\Sigma p_k)$$

daí:

$$\frac{1}{1-\alpha}\log(\Sigma p_k^{\alpha}) - \frac{1}{1-\alpha}\log(\Sigma p_k) \leq \frac{\alpha}{1-\alpha}\log(\Sigma p_k q_k^{m(\alpha-1)}) - \frac{\alpha}{1-\alpha}\log(\Sigma p_k)$$

$$+ \frac{1-\alpha m}{1-\alpha m}\log(\Sigma q_k^{\alpha m}) - \log(\Sigma p_k)$$

ou:

$$\frac{1}{1-\alpha}\log(\Sigma p_k^{\alpha}/\Sigma p_k) \leq \frac{\alpha}{1-\alpha}\log(\Sigma p_k q_k^{m(\alpha-1)}) + \frac{1-\alpha m}{1-\alpha m}\log(\Sigma q_k^{\alpha m}/\Sigma q_k)$$

ou ainda:

$$H_\alpha(P) \leq \alpha H_\alpha(P//Q;m) + (1-\alpha m)H_{\alpha m}(Q).$$

Para $0<\alpha<1$, a demonstração segue analogamente.

2.5. - β-Imprecisões.

Dados $P=(p_1,p_2,\ldots,p_n)\varepsilon\Delta_n$, $n\geq 2$, é possível construir um número infinito não enumerável de distribuições de probabilidades;

$$P^{(\beta)}=(p_1^{(\beta)},p_2^{(\beta)},\ldots,p_n^{(\beta)}),$$

ditas β-distribuições potências derivadas de P, onde:

$$p_k^{(\beta)} = \frac{p_k^\beta}{\Sigma p_k^\beta}\ , \quad \beta\varepsilon(-\infty;+\infty) \text{ e } k=1,2,\ldots,n.$$

Observemos que:

$$H_1(P^{(\beta)}//P) = -\frac{\Sigma p_k^\beta \log(p_k)/\Sigma p_k^\beta}{\Sigma p_k^\beta/\Sigma p_k^\beta}$$

$$= -\frac{\Sigma p_k^\beta \log(p_k)}{\Sigma p_k^\beta}\ , \qquad (2.25)$$

e também,

$$H_\alpha(P^{(\beta)}//P) = (1-\alpha)^{-1}\log\left(\frac{\Sigma(p_k^\beta/\Sigma p_k^\beta)\cdot p_k^{\alpha-1}}{\Sigma p_k^\beta/\Sigma p_k^\beta}\right)$$

$$= (1-\alpha)^{-1}\log\left(\frac{p_k^{\alpha+\beta-1}}{\Sigma p_k^\beta}\right);\ 1\neq\alpha>0. \qquad (2.26)$$

Uma vez que o lado direito das equações (2.25) e (2.26) dependem dos elementos de P, denotemos:

$$H_1(P^{(\beta)}//P) = I_1^\beta(P)$$

e

$$H_\alpha(P^{(\beta)}//P) = I_\alpha^\beta(P).$$

É fácil ver que, para β=1, vem que:

$$I_1^1(P) = H_1(P)$$

e

$$I_\alpha^1(P) = H_\alpha(P).$$

Denominaremos $I_1^\beta(P)$ e $I_\alpha^\beta(P)$ por β-imprecisões de ordem 1 e ordem α, respectivamente.

Se $\beta>1$, então $I_1^\beta(P)$ e $I_\alpha^\beta(P)$ representam as entropias de Kapur [4] de grau β e ordem 1 e grau β e ordem α respectivamente.

É fácil ver que $I_1^\beta(P)$ é função monotônica de β, assim:

$$I_1^\beta(P) \leqq H_1(P), \quad \beta>1$$

e (*)

$$I_1^\beta(P) \geqq H_1(P), \quad 0<\beta<1.$$

Além disso, pela inequação de Shannon,

$$H_1(P^{(\beta)}) \leqq H_1(P^{(\beta)}//P), \quad \text{quando} \quad W_1(P) \leqq 1 = W_1(P^{(\beta)}).$$

Por outro lado notemos que as equações (*) quando substituimos $H_1(P)$ por $H_\beta(P)$ nem sempre são válidas. Por exemplo, se $\beta=2$ com $P=(1/2;1/3)$ temos que:

$$I_1^\beta(P) = - \frac{1/4.\log(1/2)+(1/9)\log(1/3)}{(1/4) + (1/9)} \cong 5,07$$

$$H_\beta(P) = - \left(\frac{(1/2)\log(1/4)}{(1/2)+(1/3)} + \frac{(1/3)\log(1/3)}{(1/2)+(1/3)}\right) \cong 3,25$$

logo: $I_1^\beta(P) > H_\beta(P)$.

CAPÍTULO III - IMPRECISÃO DE ORDEM 1 E GRAU β E DE ORDEM α E GRAU β.

Neste capítulo estudaremos as imprecisões de ordem 1 e grau β e de ordem α e grau β para distribuições de probabilidades generalizadas imcompletas definidas em 3.1.

3.1. - Definição: Sejam $P,Q\varepsilon\Delta_n$, as imprecisões de ordem 1 e grau β e de ordem α e grau β definidas em [17] são dadas respectivamente por:

$$H_1^\beta(P//Q) = A_\beta[\Sigma 2^{(1-\beta)\Sigma p_k \log(q_k)/\Sigma p_k} -1], \quad 0<\beta\neq 1. \tag{3.1}$$

$$H_\alpha^\beta(P//Q) = A_\beta[(\Sigma p_k q_k^{\alpha-1}/\Sigma p_k)^{(\beta-1)/(\alpha-1)} -1] \tag{3.2}$$

com $\alpha,\beta>0,\alpha,\beta\neq 1$ e $\alpha\neq\beta$, onde $A_\beta=(2^{1-\beta}-1)^{-1}$.

Observações:

I. - Quando P=Q;

$$H_\alpha^\beta(P//Q) = H_\alpha^\beta(P) = A_\beta[(\Sigma p_k^\alpha/\Sigma p_k)^{(\beta-1)/(\alpha-1)} -1], \quad \alpha,\beta>0,$$

$\alpha,\beta\neq 1$ e $\alpha\neq\beta$, que é a entropia de Sharma Mital.

II. - Quando $\alpha=\beta\neq 1$,

$$H_\alpha^\beta(P//Q) = H_\beta(P//Q).$$

III. - $\lim\limits_{\beta\to 1} H_\alpha^\beta(P//Q) = H_\alpha(P//Q).$

IV. - $\lim\limits_{\alpha\to 1} \lim\limits_{\beta\to 1} H_\alpha^\beta(P//Q) = H_1(P//Q).$

3.2. - Generalização da desigualdade de Shannon.

Sabemos que se $P,Q\varepsilon\Delta_n^*$, $n\geq 1$, então:

$$H_1(P) \leq H_1(P//Q) \tag{3.3}$$

Se $p_k<q_k$, k=1,...,n então para $P,Q\varepsilon\Delta_n$, temos

$$H_1(P) > H_1(P//Q);$$

isto é, (3.3) não é válida para distribuições generalizadas.

Vimos no capítulo II, que

$$H_\alpha(P) \leqq H_\alpha(P//Q), \quad P,Q\varepsilon\Delta_n, \tag{3.4}$$

também nem sempre é válida. Ou seja, se $\alpha<1$ e $p_k>q_k$ ou $\alpha>1$ e $p_k<q_k$ então:

$$H_\alpha(P) > H_\alpha(P//Q).$$

A desigualdade:

$$H_\alpha^\beta(P) \leqq H_\alpha^\beta(P//Q) \tag{3.5}$$

$P,Q\varepsilon\Delta_n$, $n\geq 2$, $\alpha,\beta>0$ e $\alpha,\beta\neq 1$, também nem sempre é válida.

Procuraremos estabelecer condições para a validade ou não de (3.5). Tais condições são estabelecidas nos teoremas que seguem. O teorema abaixo mostra sob quais condições (3.5) não se verifica.

Teorema 3.2. - Sejam $P,Q\varepsilon\Delta_n$ tais que $p_k<q_k$, $k=1,\ldots,n$ então (3.5) não é válida.

Demonstração: (1) Se $\alpha>1$ e $0<\beta<1$ então $\alpha-1>0$ e para $p_k<q_k$, temos que:

$$p_k^{\alpha-1} < q_k^{\alpha-1}$$

e assim,

$$\Sigma p_k^\alpha < \Sigma p_k q_k^{\alpha-1} \tag{i}$$

Portanto, com $\Sigma p_k>0$, vem que,

$$\frac{\Sigma p_k^\alpha}{\Sigma p_k} < \frac{\Sigma p_k q_k^{\alpha-1}}{\Sigma p_k}.$$

Agora, como $(\beta-1)/(\alpha-1)<0$, segue que:

$$\left(\frac{\Sigma p_k^\alpha}{\Sigma p_k}\right)^{(\beta-1)/(\alpha-1)} > \left(\frac{\Sigma p_k q_k^{\alpha-1}}{\Sigma p_k}\right)^{(\beta-1)/(\alpha-1)}. \tag{ii}$$

Sendo $0<\beta<1$, temos que $A_\beta>0$ e daí que:

$$A_\beta[(\frac{\Sigma p_k^\alpha}{\Sigma p_k})^{(\beta-1)/(\alpha-1)}-1] > A_\beta[(\frac{\Sigma p_k q_k^{\alpha-1}}{\Sigma p_k})^{(\beta-1)/(\alpha-1)}-1] \qquad \text{(iii)}$$

isto é;

$$H_\alpha^\beta(P) > H_\alpha^\beta(P//Q).$$

(2) Se $\alpha>1$ e $\beta>1$ então $(\beta-1)/(\alpha-1)>0$ e (ii) tem seu sinal invertido. Mas como $A_\beta<0$ o sinal de (iii) permanece como no item (1) desta demonstração.

(3) Se $0<\alpha<1$ e $0<\beta<1$ então (i) troca de sinal. Sendo $(\beta-1)/(\alpha-1)>0$ (ii) continua com mesmo sinal do item (1) e o mesmo acontece com (iii), pois $A_\beta>0$.

(4) Se $0<\alpha<1$ e $\beta>1$, novamente (i) troca de sinal com relação a demonstração do item (1). Agora $(\beta-1)/(\alpha-1)<0$ e $A_\beta<0$ e o sinal da desigualdade se inverte duas vezes tornando-se ao final ainda como em (iii).

Portanto, se $0<\alpha\neq 1$ e $0<\beta\neq 1$ e $p_k<q_k$ então (3.5) não é válida.

Nos teoremas a seguir estabeleceremos condições sob as quais (3.5) se verifica.

Teorema 3.3. - Sejam $P,Q\in\Delta_n$ tais que $P\in E_\alpha^*(Q)$, com $0<\alpha\neq 1$ e $0<\beta\neq 1$. Então (3.5) se verifica.

Demonstração: Vimos no capítulo II que se $P\in E_\alpha^*(Q)$, então $H_\alpha(P)\leqq H_\alpha(P//Q)$. Por outro lado, sabemos que:

$$H_\alpha^\beta(P) = A_\beta[2^{(1-\beta)H_\alpha(P)}-1]$$

e

$$H_\alpha^\beta(P//Q) = A_\beta[2^{(1-\beta)H_\alpha(P//Q)}-1].$$

(a) Para $\beta>1$, temos que $(1-\beta)<0$ e assim que

$$(1-\beta)H_\alpha(P) \geqq (1-\beta)H_\alpha(P//Q), \qquad \text{(I)}$$

daí,

$$2^{(1-\beta)H_\alpha(P)}-1 \geqq 2^{(1-\beta)H_\alpha(P//Q)}-1$$

mas $A_\beta<0$, portanto:

$$A_\beta[2^{(1-\beta)H_\alpha(P)}-1] \leqq A_\beta[2^{(1-\beta)H_\alpha(P//Q)}-1] \qquad \text{(II)}$$

(b) Para $0<\beta<1$ segue que $(1-\beta)>0$ e daí que (I) tem seu sinal invertido. Porém agora $A_\beta>0$, portanto (II) segue da mesma forma. Logo, em ambos os casos (3.5) se verifica.

Teorema 3.4. - Sejam $P,Q\varepsilon\Delta_n$ tais que $p_k>q_k$, $k=1,2,\ldots,n$. Então (3.5) é válida.

Demonstração: Se $0<\alpha<1$ vem que $\Sigma p_k^\alpha<\Sigma p_k q_k^{\alpha-1}$. Agora, se $0<\beta<1$, então $(\beta-1)/(\alpha-1)>0$ e também $A_\beta>0$ e o sinal da desigualdade (iii) do teorema 3.1 se inverte. Além disso, se $\beta>1$, $(\beta-1)/(\alpha-1)<0$ e também $A_\beta<0$ e de novo o sinal de (iii) do teorema 3.1 se inverte.

Se $\alpha>1$, $\Sigma p_k^\alpha>\Sigma p_k q_k^{\alpha-1}$. Assim, se $\beta>1$ então $(\beta-1)/(\alpha-1)>0$ enquanto que $A_\beta<0$, daí o sinal de (iii) se inverte. Finalmente, se $0<\beta<1$ então $(\beta-1)/(\alpha-1)<0$ enquanto que $A_\beta>0$ e mais uma vez o sinal de (iii) se inverte. Concluímos então que, se $p_k>q_k$ então (3.5) é válida.

3.3. - Características monotônicas da Imprecisão de ordem α e grau β.

Se $p\varepsilon\Delta_1$ e $q\varepsilon\Delta_1$ então:

$$H_\alpha^\beta(p//q) = A_\beta[(pq^{\alpha-1}/p)^{(\beta-1)/(\alpha-1)}-1]$$

$$= A_\beta[q^{\beta-1}-1],$$

isto é, $H_\alpha^\beta(p//q)$ é independente de α. Então para distribuições generalizadas simples a questão da característica monotônica da imprecisão de ordem α e grau β não aparece. Existem alguns tipos de distribuições para as quais $H_\alpha^\beta(P//Q)$ é independente de α. No seguinte teorema apresentamos um tipo de distribuição para a qual $H_\alpha^\beta(P//Q)$ não depende de α.

Teorema 3.5. - Para qualquer $P\varepsilon\Delta_n$ seja $Q=(W_1(Q)/n,\ldots,W_1(Q)/n)$ então $H_\alpha^\beta(P//Q)$ é independente de α.

Demonstração: Para $P\varepsilon\Delta_n$ e $Q=(W_1(Q)/n,\ldots,W_1(Q)/n)$, temos que

$$H_\alpha^\beta(P//Q) = A_\beta[(\Sigma p_k(W_1(Q)/n)^{\alpha-1}/\Sigma p_k)^{(\beta-1)/(\alpha-1)}-1]$$

$$= A_\beta[(W_1(Q)/n)^{\beta-1}-1],$$

que é independente de α.

No seguinte teorema é provada a monotonicidade da função $H_\alpha^\beta(P//Q)$ com relação a α.

Teorema 3.6. - Para $\beta>0$, $\beta\neq 1$, $H_\alpha^\beta(P//Q)$ é monotonicamente decrescente com relação a α.

Demonstração: $H_\alpha^\beta(P//Q)$ pode ser escrita na forma:

$$H_\alpha^\beta(P//Q) = A_\beta[2^{(1-\beta)H_\alpha(P//Q)}-1], \text{ onde}$$

$$H_\alpha(P//Q) = 1/(1-\alpha).\log(\Sigma p_k q_k^{\alpha-1}/\Sigma p_k), \quad \alpha>0 \text{ e } \alpha\neq 1.$$

Nath [12] mostrou que para $\alpha_1<\alpha_2$, temos,

$$H_{\alpha_2}(P//Q) \leqq H_{\alpha_1}(P//Q).$$

Agora consideremos dois casos:

(a) Para $0<\beta<1$, vem que,

$$(1-\beta)H_{\alpha_2}(P//Q) \leqq (1-\beta)H_{\alpha_1}(P//Q)$$

$$[2^{(1-\beta)H_{\alpha_2}(P//Q)}-1] \leqq [2^{(1-\beta)H_{\alpha_2}(P//Q)}-1]$$

$$A_\beta[2^{(1-\beta)H_{\alpha_2}(P//Q)}-1] \leqq A_\beta[2^{(1-\beta)H_{\alpha_1}(P//Q)}-1]$$

$$H_{\alpha_2}^\beta(P//Q) \leqq H_{\alpha_1}^\beta(P//Q)$$

(b) Para $\beta>1$ a prova segue da mesma forma e, portanto, qualquer que seja $0<\beta\neq 1$ temos que $H_\alpha^\beta(P//Q)$ é monotonicamente decrescente com respeito a α.

3.4. - γ-Imprecisões.

Para $P\varepsilon\Delta_n$ é possível obter um número infinito não enumerável de distribuições de probabilidades

$$P^{(\gamma)} = (p_1^{(\gamma)}, p_2^{(\gamma)}, \ldots, p_n^{(\gamma)})$$

ditas γ-potência de distribuições derivadas de P, onde

$$p_k^{(\gamma)} = \frac{p_k^\gamma}{\Sigma p_k^\gamma}, \quad \gamma\varepsilon(-\infty;+\infty).$$

Então,

$$H_1^\beta(P^{(\gamma)}//P) = A_\beta[2^{(1-\beta)(\Sigma p_k^\gamma \log(p_k)/\Sigma p_k^\gamma)}-1] \qquad (A)$$

e

$$H_\alpha^\beta(P^{(\gamma)}//P) = A_\beta[2^{(1-\beta)(1-\alpha)^{-1}\log(\Sigma p_k^\gamma p_k^{\alpha-1}/\Sigma p_k^\gamma)}-1], \quad (B)$$

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

Notemos que

$$\lim_{\beta\to 1} H_1^\beta(P^{(\gamma)}//P) = H_1(P^{(\gamma)}//P)$$

e

$$\lim_{\beta\to 1} H_\alpha^\beta(P^{(\gamma)}//P) = H_\alpha(P^{(\gamma)}//P).$$

Onde $H_1(P^{(\gamma)}//P)$ e $H_\alpha(P^{(\gamma)}//P)$ são, respectivamente, as gama-distribuições potências de probabilidade de ordem 1 e de ordem α definidas por Nath em [11].

Como (A) e (B) dependem apenas dos elementos de P, denotemos

$$H_1^\beta(P^{(\gamma)}//P) = I_1^{\beta,\gamma}(P) \qquad (C)$$

e

$$H_\alpha^\beta(P^{(\gamma)}//P) = I_\alpha^{\beta,\gamma}(P) \qquad (D)$$

Para $\gamma=1$, temos,

$$I_1^{\beta,1}(P) = H_1^{\beta}(P)$$

e

$$I_\alpha^{\beta,1}(P) = H_\alpha^{\beta}(P)$$

$I_1^{\beta,\gamma}(P)$ e $I_\alpha^{\beta,\gamma}(P)$ são γ-imprecisões de ordem 1 e grau β e ordem α e grau β, respectivamente.

No seguinte teorema provaremos a relação entre $I_1^{\beta,\gamma}(P)$ e $H_1^{\beta}(P)$.

Teorema 3.7. - (i) Para $\gamma>1$, temos que $I_1^{\beta,\gamma}(P) \leqq H_1^{\beta}(P)$ e

(ii) Para $0<\gamma<1$, temos que $I_1^{\beta,\gamma}(P) \geqq H_1^{\beta}(P)$.

Demonstração: (i) Seja $\gamma>1$, então temos que

$$I_1^{\gamma}(P) \leqq H_1(P) \qquad \text{(E)}$$

$I_1^{\beta,\gamma}(P)$ pode ser escrita da seguinte forma:

$$I_1^{\beta,\gamma}(P) = A_\beta[2^{(\beta-1)I_1^{\gamma}(P)}-1] \qquad \text{(F)}$$

Agora temos dois casos:

(a) Seja $\beta>1$. Então usando (E) e (F) podemos ver que,

$$I_1^{\beta,\gamma}(P) \leqq H_1^{\beta}(P).$$

(b) Seja $0<\beta<1$. A demonstração segue analogamente.

(ii) Seja $0<\gamma<1$. A demonstração é similar porque agora (E) tem seu sinal invertido.

B I B L I O G R A F I A

1. - Aczel, J. and Daroczi, A.; On measures of information and their characterizations; Academic Press, London, 1975.

2. - Arora, P. N.; Studies on entropy, inaccuracy and relative information of order α; Ph.D. thesis, Delhi University, 1970.

3. - Gallager, R. G.; Information theory and reliable communication; John Wiley and Sons, New York, 1968.

4. - Kapur, J. N.; Generalized entropy of order α and type β; The mathematics Seminar, 4, 78-94, 1968.

5. - Kerridge, D. F.; Inaccuracy and inference; J. Roy. Stat. Soc.; Série B, 23, 184-199, 1961.

6. - Lindley, D. V.; Binomial sampling and the concept of information; Biometrika, 44, 179-186, 1957.

7. - Mathai, A. M. and Rathie, P. N.; Basic concept of information theory and statistics; Wiley Eastern Limited, New Delhi, 1975.

8. - Nath, P.; Entropy and inaccuracy;Cahiers due Centre D'etudes de Recherche Operationnelle, 4, 192-205, 1967.

9. - Nath, P.; Inaccuracy and coding theory; Metrika 13, 123-135, 1968.

10.- Nath, P.; On the measures of errors in information; J. Math. Sciences 3(1), 1-16, 1968.

11.- Nath, P.; An axiomatic characterization of inaccuracy for discrete generalized probability distributions; Opsearch 7(2), 115-133, 1970.

12.- Nath, P.;Remarks on some measures of inaccuracy of finite discrete generalized probability distributions; Entropy and Ergodic Theory, Vol II, 77-100, 1974.

13.- Renyi, A.; On measures of entropy and information; Proc. 4th Berk. Symp. Math. Stat. and Prob. Bol 1, 547-561. 1961.

14.- Rathie, P. N.; On generalized measures of inaccuracies, information and errors in information; Statistics, 30, 340-349. 1970.

15.- Sharma, B. D. and Gupta, H. C.; Sub-additive measures of relative information and inaccuracy; Metrika, 23, 155-165. 1976.

16.- Shannon, C. E.; A mathematical theory of communication; BSTJ, 27, 379-423. 1948.

17.- Van der Pyl, T.; Information d'order α et type β; Axiomatique et propriétés, Thèse de 3º Cycle, Paris, 1977.