METRIZATION OF SETS OF SUB-σ-ALGEBRAS AND THEIR CONDITIONAL ENTROPIES

METRIZATION OF SETS OF SUB-σ-ALGEBRAS AND THEIR CONDITIONAL ENTROPIES

By

J. M. SINGH, M.A.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

September 1971

MASTER OF SCIENCE (1971)
(Mathematics)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Metrization of Sets of Sub-σ-Algebras and their Conditional Entropies

AUTHOR: J. M. Singh, M.A. (Delhi University)

SUPERVISOR: Dr. M. Behara

NUMBER OF PAGES: 32

SCOPE AND CONTENTS: This thesis deals with metrizations of sets of conditional entropies and sets of sub-σ-algebras. C. Rajski's Theorem ([9]) on the metric space of discrete probability distributions can be deduced as a particular case of a theorem on the metric space of sub-σ-algebras given in Chapter III, the proof of which is comparatively very concise. The completeness of this metric space and some other properties are also proved.

PREFACE

The first chapter of this thesis deals with known results on probability theory, metric spaces, conditional entropies and generalized conditional entropies.

The second chapter deals with the metrization of sets of conditional entropies and sets of finite sub-σ-algebras.

In the third chapter it is shown that a metric can be defined on the set of all sub-σ-algebras of a given algebra. It is observed that C. Rajski's ([9]) theorem on the metric space of discrete probability distribution turns out to be particular case of this theorem. The completeness and other properties of this metric space are also established.

ACKNOWLEDGEMENTS

The author wishes to express his sincere appreciation to his supervisor Dr. M. Behara for his willing assistance and guidance and for the generosity with which he gave his valuable time during the course of this research.

The author also wishes to express his gratitude to Dr. C. E. Billigheimer for a critical review of the manuscript and to Dr. P. Nath for helpful discussions.

Thanks also go to Miss Irene Bojkiwskyj for her prompt and efficient typing of the manuscript and to McMaster University for financial support.

TABLE OF CONTENTS

CHAPTER I

PRELIMINARIES

1.1 Introduction

This chapter presents known results on probability, metric spaces and conditional entropies. Throughout the discussion of this chapter and the ensuing chapters, unless otherwise stated, the probability space under consideration is denoted by $(\Omega, \mathcal{R}, P)$ where Ω is the abstract space, $\mathcal{R}$ is the σ-algebra of all subsets of Ω and P is the probability measure over $\mathcal{R}$. By a finite σ-algebra we shall mean a σ-algebra consisting of a finite number of subsets of Ω. It is easy to verify that there is a one-to-one correspondence between finite measurable partitions of Ω and finite sub-σ-algebras of $\mathcal{R}$.

1.2 Metric Spaces

Definition 1.2.1: A metric space (x,d) is a non-empty set X of elements (which we call points) together with a real-valued function d defined on X×X such that for all x,y and z in X:

i) $d(x,y) \geq 0$

ii) $x = y \Rightarrow d(x,y) = 0$

iii) $d(x,y) = 0 \Rightarrow x = y$

iv) $d(x,y) \leq d(x,z) + d(y,z)$

The function d is called a metric.

Note 1.2.1: Putting $z = x$ in (iv), we get

$$d(x,y) \leq d(x,x) + d(y,x) = d(y,x). \text{ ----(a)}$$

But as x and y are arbitrary points, therefore a similar argument gives

$$d(y,x) \leq d(x,y). \text{ ----(b)}$$

From (a) and (b) we have $d(x,y) = d(y,x)$. Hence $d(x,y)$ is symmetric function.

Definition 1.2.2: A pair (X,d) is called a pseudometric space if d satisfies all the conditions of a metric except that $d(x,y) = 0$ need not imply $x = y$.

Definition 1.2.3: Let (X,d) be the given metric space. If there exists a positive number k such that $d(x,y) \leq k$ for every pair of points x and y of X, we say that the metric space (X,d) is a bounded metric space. A metric space which is not bounded is said to be unbounded; in that case $d(x,y)$ takes values as large as we please.

Definition 1.2.4: The sequence of points $\{x_n\}$ in a metric space (X,d) is said to converge to a point $x \in X$ if the distance $d(x_n,x)$ tends to zero as $n \to \infty$, that is, if for every positive value of ε there exists an integer n_0, depending on ε, such that

$$d(x_n, x) < \varepsilon, \text{ whenever } n \geq n_0$$

The point x is called the limit point of the sequence.

<u>Definition 1.2.5</u>: The sequence of points $\{x_n\}$ in a metric space (X,d) is called a Cauchy sequence (or a fundamental sequence), if, for every positive value of ε, there exists an integer $n_0(\varepsilon)$ such that

$$d(x_n, x_{n+p}) < \varepsilon \text{ whenever } n \geq n_0(\varepsilon) \text{ and } p > 0.$$

It is to be noted that every convergent sequence is a Cauchy sequence in a metric space but the converse is not necessarily true.

<u>Definition 1.2.6</u>: A metric space (X,d) is said to be complete if every Cauchy sequence $\{x_n\}$ of points of X converges to a point of X.

1.3 Probability

1.3.1 <u>Conditional Probability</u>: For any two sets $A, B \in \mathcal{R}$, such that $P(B) > 0$, the conditional probability of A given that B has occured is defined as:

$$\frac{P(A \cap B)}{P(B)} \text{ and is denoted by } P(A/B)$$

Now, we define the conditional probability of $A \in \mathcal{R}$

given the sub-σ-algebra $\mathcal{R}_0$ of $\mathcal{R}$. We denote this conditional probability by $P_{\mathcal{R}_0}(A)$ and define it as an integrable random variable which is

1. measurable w.r.t. the sub-σ-algebra $\mathcal{R}_0$, and
2. satisfies the functional equation

$$\int_G P_{\mathcal{R}_0}(A)dP = P(A \cap G);\ G \varepsilon \mathcal{R}_0.$$

Applying the Radon-Nikodym theorem the existence of the random variables $P_{\mathcal{R}_0}(A)$ can be proved (For details, cf. (1), Chapter III, p. 95).

If $\mathcal{R}_0$ is generated by the finite or countable measurable partition $\{B_j\}$ of Ω, then $P_{\mathcal{R}_0}(A)$ may also be defined as follows:

$$P_{\mathcal{R}_0}(A) = \sum_j \frac{P(A \cap B_j)}{P(B_j)} \chi(B_j);\ A \varepsilon \mathcal{R}$$

up to equivalence ($\chi(B_j)$ is the characteristic function of the set B_j).

1.4 Entropy

Notation 1.4.1: If $\{\mathcal{R}_\alpha : \alpha \varepsilon \Lambda\}$ is a family of sub-σ-algebras of $\mathcal{R}$, then, $\bigvee_{\alpha \varepsilon \Lambda} \mathcal{R}_\alpha$ denotes the sub-σ-algebra of $\mathcal{R}$ generated by $\bigcup_{\alpha \varepsilon \Lambda} \mathcal{R}_\alpha$ and $\bigwedge_{\alpha \varepsilon \Lambda} \mathcal{R}_\alpha$ denotes the largest sub-σ-algebra of $\mathcal{R}$ contained in each of the sub-σ-algebra

$\mathcal{R}_\alpha$ (In the finite case we also write $\bigvee_{i=1}^{n} \mathcal{R}_i = \mathcal{R}_1 \vee \mathcal{R}_2 \vee \ldots \vee \mathcal{R}_n$ and $\bigwedge_{i=1}^{n} \mathcal{R}_i = \mathcal{R}_1 \wedge \mathcal{R}_2 \wedge \ldots \wedge \mathcal{R}_n$).

<u>Definition 1.4.1</u>: Let $\mathcal{R}_0 \subseteq \mathcal{R}$ be the finite σ-algebra whose atoms are $A_1, A_2, \ldots, A_n$. The entropy $H(\mathcal{R}_0)$ of the finite σ-algebra $\mathcal{R}_0$ is defined as

$$H(\mathcal{R}_0) = - \sum_{k=1}^{n} p_k \log_2 p_k, \text{ where } p_k = P(A_k);$$

$$k = 1, 2, \ldots, n.$$

<u>Definition 1.4.2</u>: Let $\mathcal{R}_0, \mathcal{R}_0'$ be two finite sub-σ-algebras of $\mathcal{R}$ whose atoms are $A_i (1 \leq i \leq \mathfrak{n})$ and $A_k' (1 \leq k \leq \mathfrak{n}')$ respectively. We define

$$p_{ik} = \begin{cases} \dfrac{P(A_i \cap A_k')}{P(A_i)} & \text{if } P(A_i) > 0 \\ \dfrac{1}{\mathfrak{n}'} \text{ (say)} & \text{if } P(A_i) = 0 \end{cases}$$

and $P^{(i)} = (p_{i1}, p_{i2}, \ldots, p_{i\mathfrak{n}'})$.

The conditional entropy $H(\mathcal{R}_0'/\mathcal{R}_0)$ of the finite sub-σ-algebra $\mathcal{R}_0'$ w.r.t. the finite sub-σ-algebra $\mathcal{R}_0$ is defined as:

$$H(\mathcal{R}_0'/\mathcal{R}_0) = \sum_{i=1}^{\mathfrak{n}} p_i H(P^{(i)}), \text{where } H(P^{(i)}) = - \sum_{k=1}^{\mathfrak{n}'} p_{ik} \log_2 p_{ik}.$$

Remark 1.4.1: $H(\mathcal{R}_0'/\mathcal{R}_0)$ is not changed if we replace $\mathcal{R}_0, \mathcal{R}_0'$ by equivalent finite sub-σ-algebras of $\mathcal{R}$.

Definition 1.4.3: Let S be the system of all finite sub-σ-algebras of any given σ-algebra $\mathcal{R}_0 \subseteq \mathcal{R}$. The entropy $H(\mathcal{R}_0)$ of the σ-algebra $\mathcal{R}_0$ is defined as

$$H(\mathcal{R}_0) = \sup_{C \in S} H(C).$$

Definition 1.4.4: Let S and S' be the systems of all finite sub-σ-algebras of $\mathcal{R}_0 \subseteq \mathcal{R}$ and $\mathcal{R}_0' \subseteq \mathcal{R}$ respectively; then the conditional entropy $H(\mathcal{R}_0'/\mathcal{R}_0)$ of $\mathcal{R}_0'$ w.r.t. $\mathcal{R}_0$ is defined as

$$H(\mathcal{R}_0'/\mathcal{R}_0) = \sup_{C' \in S'} \inf_{C \in S} H(C'/C).$$

Note 1.4.1: Let $\mathcal{R}_0, \mathcal{R}_1, \mathcal{R}_0', \mathcal{R}_1'$ be arbitrary σ-algebras $\subseteq \mathcal{R}$; then, we have the following (cf. [5], p. 260, Theorem 4).

A. Equalities

i) $H(\mathcal{R}_0 \vee \mathcal{R}_0') = H(\mathcal{R}_0) + H(\mathcal{R}_0'/\mathcal{R}_0)$.

ii) $H(\mathcal{R}_0' \vee \mathcal{R}_1'/\mathcal{R}_0) = H(\mathcal{R}_0'/\mathcal{R}_0) + H(\mathcal{R}_1'/\mathcal{R}_0' \vee \mathcal{R}_0)$.

iii) $H(\mathcal{R}_0'/\mathcal{R}_0) = H(\mathcal{R}_0' \vee \mathcal{R}_1'/\mathcal{R}_0)$ if $\mathcal{R}_1' \subseteq \mathcal{R}_0$.

iv) $H(\mathcal{R}_0'/\mathcal{R}_0) = 0$ if $\mathcal{R}_0' \subseteq \mathcal{R}_0$ up to equivalence.

B. Inequalities

v) $H(\mathcal{R}_0) \leq H(\mathcal{R}_1)$ if $\mathcal{R}_0 \subseteq \mathcal{R}_1$.

vi) $H(\mathcal{R}_0'/\mathcal{R}_0) \leq H(\mathcal{R}_1'/\mathcal{R}_0)$ if $\mathcal{R}_0' \subseteq \mathcal{R}_1'$.

vii) $H(\mathcal{R}_0'/\mathcal{R}_0) \geq H(\mathcal{R}_0'/\mathcal{R}_1)$ if $\mathcal{R}_0 \subseteq \mathcal{R}_1$.

viii) $0 \geq H(\mathcal{R}_0'/\mathcal{R}_0) - H(\mathcal{R}_0') \geq H(\mathcal{R}_1'/\mathcal{R}_0) - H(\mathcal{R}_1')$

if $\mathcal{R}_0' \subseteq \mathcal{R}_1'$ and $H(\mathcal{R}_1') < \infty$.

ix) $H(\mathcal{R}_1') - H(\mathcal{R}_0') \leq H(\mathcal{R}_1'/\mathcal{R}_0) - H(\mathcal{R}_0'/\mathcal{R}_0)$

if $H(\mathcal{R}_1') < \infty$ and $\mathcal{R}_0' \subseteq \mathcal{R}_1'$.

$H(\mathcal{R}_1'/\mathcal{R}_0) - H(\mathcal{R}_0'/\mathcal{R}_0) \geq H(\mathcal{R}_1'/\mathcal{R}_1) - H(\mathcal{R}_0'/\mathcal{R}_1)$

if $H(\mathcal{R}_1'/\mathcal{R}_0) < \infty$ and $\mathcal{R}_0' \subseteq \mathcal{R}_1'$, $\mathcal{R}_0 \subseteq \mathcal{R}_1$.

x) $H(\mathcal{R}_0 V \mathcal{R}_1) \leq H(\mathcal{R}_0) + H(\mathcal{R}_1)$ with equality if $\mathcal{R}_0$ and $\mathcal{R}_1$ are independent, i.e., if

$P(E \cap F) = P(E).P(F) \qquad (E \epsilon \mathcal{R}_0,\ F \epsilon \mathcal{R}_1)$.

If $H(\mathcal{R}_o V \mathcal{R}_1) < \infty$, then equality implies independence.

xi) $H(\mathcal{R}_0' V \mathcal{R}_1'/\mathcal{R}_0) \leq H(\mathcal{R}_0'/\mathcal{R}_0) + H(\mathcal{R}_1'/\mathcal{R}_0)$.

1.4.5 Generalized Conditional Entropies

A few generalized conditional entropies are given by M. Behara and P. Nath for finite measurable partitions (see [1]) from which Renyi's conditional entropy and Shannon's conditional entropy can be deduced as a particular case. Here, we define some of these generalized entropies.

If $\mathcal{R}_0'$ and $\mathcal{R}_0$ are two finite sub-σ-algebras of $\mathcal{R}$

whose atoms are $A'_k (1 \le k \le n')$ and $A_i (1 \le i \le n)$ respectively, then we know that the conditional probability of A given $\mathcal{R}_0$ is defined up to equivalence by

$$P_{\mathcal{R}_0}(A) = P(A/\mathcal{R}_0) = \sum_{i=1}^{n} \frac{P(A \cap A_i)}{P(A_i)} \chi(A_i); \quad A \varepsilon \mathcal{R}'_0.$$

where $\chi(A_i)$ is the characteristic function of A_i and the conditional entropy of $\mathcal{R}'_0$ w.r.t. $\mathcal{R}_0$ is defined as

$$\hat{I}_\alpha(\mathcal{R}'_0/\mathcal{R}_0) = \sum_{k=1}^{n'} \int_\Omega Z_\alpha(P(A'_k/\mathcal{R}_0))dP$$

$$\text{where} \quad Z_\alpha(t) = \begin{cases} \dfrac{t-t^\alpha}{1-2^{1-\alpha}} & ; \quad t \varepsilon (0,1], \ \alpha \varepsilon [0,\infty) \\ 1 & ; \quad t = 0, \ \alpha = 0 \\ 0 & ; \quad t = 0, \ \alpha \varepsilon (0,\infty). \end{cases}$$

Another conditional entropy is defined as

$$I_\alpha(\mathcal{R}'_0/\mathcal{R}_0) = \sum_{k=1}^{n'} \int_\Omega P^{\alpha-1}(\mathcal{R}_0) \ Z_\alpha(P(A'_k/\mathcal{R}_0))dP.$$

An equivalent form of this definition is

$$I_\alpha(\mathcal{R}'_0/\mathcal{R}_0) = \sum_{i=1}^{n} P^\alpha(A_i) \ I_\alpha(\mathcal{R}'_0/A_i)$$

where

$$I_\alpha(\mathcal{R}'_0/A_i) = \sum_{k=1}^{n'} Z_\alpha(P(A'_k/A_i)).$$

If $\mathcal{R}_0$, $\mathcal{R}'_0$ and $\mathcal{R}''_0$ be the finite sub-σ-algebras of $\mathcal{R}$, then the following properties hold; (cf. [1], p. 34-37).

i) $\mathcal{R}_0' \subseteq \mathcal{R}_0$ up to equivalence $\Leftrightarrow I_\alpha(\mathcal{R}_0'/\mathcal{R}_0) = 0$; $\alpha > 0$.

ii) $\mathcal{R}_0' \subseteq \mathcal{R}_0'' \Rightarrow I_\alpha(\mathcal{R}_0/\mathcal{R}_0'') \leq I_\alpha(\mathcal{R}_0/\mathcal{R}_0')$; $\alpha \geq 1$.

iii) $I_\alpha(\mathcal{R}_0' V \mathcal{R}_0''/\mathcal{R}_0) = I_\alpha(\mathcal{R}_0'/\mathcal{R}_0) + I_\alpha(\mathcal{R}_0''/\mathcal{R}_0 V \mathcal{R}_0')$; $\alpha \geq 0$.

iv) $I_\alpha(\mathcal{R}_0 V \mathcal{R}_0') = I_\alpha(\mathcal{R}_0) + I_\alpha(\mathcal{R}_0'/\mathcal{R}_0)$; $\alpha \geq 0$.

v) If $\mathcal{R}_0 \subseteq \mathcal{R}_0'$, then $I_\alpha(\mathcal{R}_0'/\mathcal{R}_0) = I_\alpha(\mathcal{R}_0') - I_\alpha(\mathcal{R}_0)$; $\alpha \geq 0$.

vi) If $\mathcal{R}_0' \subseteq \mathcal{R}_0''$, then $I_\alpha(\mathcal{R}_0'/\mathcal{R}_0) \leq I_\alpha(\mathcal{R}_0''/\mathcal{R}_0)$; $\alpha \geq 0$.

A few more generalized conditional entropies are also discussed in [1].

CHAPTER II

METRIC SPACES OF CONDITIONAL ENTROPIES AND FINITE SUB-σ-ALGEBRAS

2.1 Metric Space of Conditional Entropies

Definition 2.1.1: Let $\mathcal{B}$ and $\mathcal{C}$ be finite sub-σ-algebras of $\mathcal{R}$ generated by the measurable partitions $\{B_i\}_{i=1}^{n}$ and $\{C_j\}_{j=1}^{m}$ respectively. Let $\mathcal{R}_0$ be a sub-σ-algebra of $\mathcal{R}$. Then the function $H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C})$ is almost everywhere defined by

$$H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C})(x) = \sum_{j=1}^{m} P_{\mathcal{R}_0}(C_j)(x) \sum_{i=1}^{n} Z\left(\frac{P_{\mathcal{R}_0}(B_i \cap C_j)(x)}{P_{\mathcal{R}_0}(C_j)(x)}\right)$$

where $Z(x) = -x \log_2 x,\ 0 < x \leq 1$

and $Z(0) = 0$.

In particular, for $\mathcal{C} = \{\phi, \Omega\}$ the function $H_{\mathcal{R}_0}(\mathcal{B})$ is almost everywhere defined by

$$H_{\mathcal{R}_0}(\mathcal{B})(x) = -\sum_{i=1}^{n} P_{\mathcal{R}_0}(B_i)(x) \log_2 P_{\mathcal{R}_0}(B_i)(x).$$

Remark 2.1.1: The following properties hold; (cf. [8], Proposition 2.1).

a) $H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C})$ is a $\mathcal{R}_0$-measurable function on Ω.

b) For almost every $x \in \Omega$ we conclude by the definition of $P_{\mathcal{R}_0}$ that the mapping $P_{\mathcal{R}_0} : A \longrightarrow P_{\mathcal{R}_0}(A)(x)$ is a measure on $\mathcal{B}_1 V \mathcal{C}_1$ and it follows that for any pair of finite sub-σ-algebras $\mathcal{B}_1$ and $\mathcal{C}_1$ of $\mathcal{R}$; $\mathcal{B}_1 \supset \mathcal{B}$ and $\mathcal{C}_1 \supset \mathcal{C}$, $H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C})$, $H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1)$ and $H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C})$ are conditional entropies w.r.t. the measure space $(\Omega, \mathcal{B}_1 V \mathcal{C}_1, P_{\mathcal{R}_0})$ and the following result is true:

$$H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}_1) \leq H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) \leq H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}) \quad \text{a.e.}$$

c) $$H(\mathcal{B}/\mathcal{R}_0 V \mathcal{C}) = \int_\Omega H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C})\, dP.$$

<u>Remark 2.1.2</u>: From (a) of Remark 2.1.1, we know that

$H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C})(x)$ is a $\mathcal{R}_0$-measurable function

$\Longrightarrow H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C})(x)$ is a $\mathcal{R}$-measurable function

$\Longrightarrow H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C})(x)$ is a random variable.

<u>Theorem 2.1.1</u>: If Z is the space of all equivalence classes of functions $H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C})(x)$; where $\mathcal{R}_0$ is a fixed sub-σ-algebra of $\mathcal{R}$ and $\mathcal{B}$ and $\mathcal{C}$ are any two finite sub-σ-algebras of $\mathcal{R}$ then $(Z, d_{\mathcal{R}_0})$ is a metric space, where

$$d_{\mathcal{R}_0}(H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1), H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)) = \int \frac{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}\, dP$$

and $\mathcal{B}_1, \mathcal{C}_1, \mathcal{B}_2, \mathcal{C}_2$ are finite sub-σ-algebras of $\mathcal{R}$.

Proof:

a) $d_{\mathcal{R}_0}(H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1), H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)) \geq 0.$

b) $H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) \overset{P}{=} H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2) \Longrightarrow H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2) = 0$ a.e.

$$\Longrightarrow \frac{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|} = 0 \quad \text{a.e.}$$

$$\Longrightarrow \int \frac{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}\, dP = 0$$

$$\Longrightarrow d_{\mathcal{R}_0}(H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1), H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)) = 0.$$

c) $d_{\mathcal{R}_0}(H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1), H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)) = 0$

$$\Longrightarrow \int \frac{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}\, dP = 0$$

$$\Longrightarrow \frac{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|} = 0 \quad \text{a.e.}$$

$$\Longrightarrow |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)| = 0 \quad \text{a.e.}$$

$$\Longrightarrow H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) \overset{P}{=} H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2).$$

d) For almost every $x\varepsilon\Omega$, we have

$$|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)| \leq |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3)|$$

$$+ |H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)| \quad \text{———} \quad (1)$$

and

$$\frac{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3)|} + \frac{|H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}$$

$$\geq \frac{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3)| + |H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3)| + |H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}$$

$$= 1 \Big/ \left(1 + \frac{1}{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3)| + |H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}\right)$$

$$\geq 1 \Big/ \left(1 + \frac{1}{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}\right) \quad \text{(using (1))}$$

$$= \frac{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}$$

Thus for a.e. $x\varepsilon\Omega$, we have

$$\frac{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|} \leq \frac{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3)|}$$

$$+ \frac{|H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}$$

$$\Longrightarrow \int \frac{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}\, dP$$

$$\leq \int \frac{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3)|}\, dP$$

$$+ \int \frac{|H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}\, dP$$

$$\Longrightarrow d_{\mathcal{R}_0}(H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1), H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)) \leq d_{\mathcal{R}_0}(H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1), H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3))$$

$$+ d_{\mathcal{R}_0}(H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3), H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)).$$

Hence the triangular inequality is also satisfied.

<u>Corollary 2.1.1</u>: Let the metric $d_{\mathcal{R}_0}$ on the set Z of Theorem 2.1.1 be defined as:

$$d_{\mathcal{R}_0}(H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1), H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)) = \int |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|\, dP$$

then $(Z, d_{\mathcal{R}_0})$ is a metric space.

2.2 Metric Spaces of Finite Sub-σ-Algebras

Theorem 2.2.1: Let $\mathcal{R}_0$ be any given sub-σ-algebra of $\mathcal{R}$. If Z is the set of all equivalence classes of finite sub-σ-algebras of $\mathcal{R}$ and if $(\mathcal{B}_1,\mathcal{C}_1)\varepsilon Z\times Z$ and $(\mathcal{B}_2,\mathcal{C}_2)\varepsilon Z\times Z$, then $(Z\times Z,\ d_{\mathcal{R}_0})$ is a pseudo-metric space, where, for almost every $x\varepsilon\Omega$, $d_{\mathcal{R}_0}$ is defined as follows:

$$d_{\mathcal{R}_0}\{(\mathcal{B}_1,\mathcal{C}_1),\ (\mathcal{B}_2,\mathcal{C}_2)\} = |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|.$$

Proof: For almost every $x\varepsilon\Omega$, we have

a) $d_{\mathcal{R}_0}\{(\mathcal{B}_1,\mathcal{C}_1),\ (\mathcal{B}_2,\mathcal{C}_2)\} \geq 0.$

b) $(\mathcal{B}_1,\mathcal{C}_1) = (\mathcal{B}_2,\mathcal{C}_2) \Longrightarrow \mathcal{B}_1 \overset{P}{=} \mathcal{B}_2$ and $\mathcal{C}_1 \overset{P}{=} \mathcal{C}_2$

$\Longrightarrow H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) = H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)$ a.e $\Longrightarrow |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)| = 0$ a.

$\Longrightarrow d_{\mathcal{R}_0}\{(\mathcal{B}_1,\mathcal{C}_1),\ (\mathcal{B}_2,\mathcal{C}_2)\} = 0$ a.e.

c) $d_{\mathcal{R}_0}\{(\mathcal{B}_1,\mathcal{C}_1),\ (\mathcal{B}_2,\mathcal{C}_2)\} = 0$ a.e. $\Longrightarrow H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) = H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)$ a.e

$\not\Longrightarrow (\mathcal{B}_1,\mathcal{C}_1) = (\mathcal{B}_2,\mathcal{C}_2).$

d) Let $(\mathcal{B}_3,\mathcal{C}_3)\varepsilon Z\times Z$. For almost every $x\varepsilon\Omega$, we have

$$|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)| \leq |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3)|$$

$$+ |H_{\mathcal{R}_0}(\mathcal{B}_3/\mathcal{C}_3) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|$$

$$\Longrightarrow \quad d_{\mathcal{R}_0}\{(\mathcal{B}_1,\mathcal{C}_1), (\mathcal{B}_2,\mathcal{C}_2)\} \leq d_{\mathcal{R}_0}\{(\mathcal{B}_1,\mathcal{C}_1), (\mathcal{B}_3,\mathcal{C}_3)\}$$

$$+ d_{\mathcal{R}_0}\{(\mathcal{B}_3,\mathcal{C}_3), (\mathcal{B}_2,\mathcal{C}_2)\}.$$

Thus it is proved that $(Z\times Z, d_{\mathcal{R}_0})$ is pseudometric space.

Corollary 2.2.1: $(Z\times Z, d_{\mathcal{R}_0})$ is a bounded pseudometric space, if $d_{\mathcal{R}_0}$ is defined for almost every $x\varepsilon\Omega$, as:

$$d_{\mathcal{R}_0}\{(\mathcal{B}_1,\mathcal{C}_1), (\mathcal{B}_2,\mathcal{C}_2)\} = \frac{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|} .$$

Corollary 2.2.2: Putting $\mathcal{C}_1 = \mathcal{C}_2 = \{\Omega, \phi\}$ in the metric of Theorem 2.2.1, we find that $(Z, d_{\mathcal{R}_0})$ is a pseudometric space where d is defined for almost every $x\varepsilon\Omega$ as:

$$d_{\mathcal{R}_0}(\mathcal{B}_1,\mathcal{B}_2) = |H_{\mathcal{R}_0}(\mathcal{B}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2)| \text{ or}$$

$$d_{\mathcal{R}_0}(\mathcal{B}_1,\mathcal{B}_2) = \frac{|H_{\mathcal{R}_0}(\mathcal{B}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2)|} .$$

Other Examples of Pseudometric Spaces

Let $\mathcal{R}_0$ and Z be as defined in Theorem 2.2.1.

Consider the following functions on Z×Z.

1. $d_{\mathcal{R}_0}\{(\mathcal{B}_1, \mathcal{C}_1), (\mathcal{B}_2, \mathcal{C}_2)\} = \int |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)| dP.$

2. $d_{\mathcal{R}_0}\{(\mathcal{B}_1, \mathcal{C}_1), (\mathcal{B}_2, \mathcal{C}_2)\} = \int \frac{|H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C}_1) - H_{\mathcal{R}_0}(\mathcal{B}_2/\mathcal{C}_2)|} dP.$

For each of the functions $d_{\mathcal{R}_0}$ given above $(Z\times Z, d_{\mathcal{R}_0})$ is a pseudometric space.

Now consider the following functions on Z.

3. $d_{\mathcal{R}_0}(\mathcal{B}, \mathcal{C}) = \int |H_{\mathcal{R}_0}(\mathcal{B}) - H_{\mathcal{R}_0}(\mathcal{C})| dP.$

4. $d_{\mathcal{R}_0}(\mathcal{B}, \mathcal{C}) = \int \frac{|H_{\mathcal{R}_0}(\mathcal{B}) - H_{\mathcal{R}_0}(\mathcal{C})|}{1 + |H_{\mathcal{R}_0}(\mathcal{B}) - H_{\mathcal{R}_0}(\mathcal{C})|} dP.$

For each $d_{\mathcal{R}_0}$ given by 3 and 4 $(Z, d_{\mathcal{R}_0})$ is a pseudometric space.

Theorem 2.2.2: If $\mathcal{B}$ and $\mathcal{C}$ are finite sub-σ-algebras of $\mathcal{R}$, then we have the following results:

1. $H_{\mathcal{R}_0}(\mathcal{B} \vee \mathcal{C}) = H_{\mathcal{R}_0}(\mathcal{B}) + H_0(\mathcal{C}/\mathcal{B})$ a.e.

2. $H_{\mathcal{R}_0}(\mathcal{B} \vee \mathcal{C}/\mathcal{B}_1) = H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{B}_1) + H_{\mathcal{R}_0}(\mathcal{C}/\mathcal{B} \vee \mathcal{B}_1)$ a.e.

where $\mathcal{B}_1$ is also a finite sub-σ-algebra of $\mathcal{R}$.

<u>Proof</u>: For almost every $x \varepsilon \Omega$, we know by the definition of $P_{\mathcal{R}_0}$ that the mapping $P_{\mathcal{R}_0} : A \longrightarrow P_{\mathcal{R}_0}(A)(x)$; $A \varepsilon \mathcal{B} V \mathcal{C}$ is a measure on $\mathcal{B} V \mathcal{C}$ (cf. [8], Proposition 2.1). Therefore $H_{\mathcal{R}_0}(\mathcal{B} V \mathcal{C})$ is the entropy w.r.t. the measure space $(\Omega, \mathcal{B} V \mathcal{C}, P_{\mathcal{R}_0})$.

Hence $H_{\mathcal{R}_0}(\mathcal{B} V \mathcal{C}) = H_{\mathcal{R}_0}(\mathcal{B}) + H_{0}(\mathcal{C}/\mathcal{B})$ a.e.

Similarly $P_{\mathcal{R}_0} : A \longrightarrow P_{\mathcal{R}_0}(A)(x)$; $A \varepsilon \mathcal{B} V \mathcal{C} V \mathcal{B}_1$, is a measure on $\mathcal{B} V \mathcal{C} V \mathcal{B}_1$. Therefore $H_{\mathcal{R}_0}(\mathcal{B} V \mathcal{C}/\mathcal{B}_1)$ is the conditional entropy w.r.t. the measure space $(\Omega, \mathcal{B} V \mathcal{C} V \mathcal{B}_1, P_{\mathcal{R}_0})$.

Hence $H_{\mathcal{R}_0}(\mathcal{B} V \mathcal{C}/\mathcal{B}_1) = H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{B}_1) + H_{\mathcal{R}_0}(\mathcal{C}/\mathcal{B} V \mathcal{B}_1)$ a.e.

<u>Note 2.2.1</u>: If $\mathcal{B}, \mathcal{C}, \mathcal{B}_1, \mathcal{C}_1$ are finite sub-σ-algebras of $\mathcal{R}$, then the following results can be similarly proved.

3. $H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}) = H_{\mathcal{R}_0}(\mathcal{B} V \mathcal{B}_1/\mathcal{C})$ a.e if $\mathcal{B}_1 \subseteq \mathcal{C}$.

4. $H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}) = 0$ a.e if $\mathcal{B} \subseteq \mathcal{C}$ up to equivalence .

5. $H_{\mathcal{R}_0}(\mathcal{B}) \leq H_{\mathcal{R}_0}(\mathcal{C})$ a.e if $\mathcal{B} \subseteq \mathcal{C}$.

6. $H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}) \leq H_{\mathcal{R}_0}(\mathcal{B}_1/\mathcal{C})$ a.e if $\mathcal{B} \subseteq \mathcal{B}_1$.

7. $H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}) \geq H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}_1)$ a.e if $\mathcal{C} \subseteq \mathcal{C}_1$.

Note 2.2.2: If Z is the set of equivalence classes of sub-σ-algebras of $\mathcal{R}$, then (Z, d) is a metric space where d is given by

$$d(\mathcal{B},\mathcal{C}) = H(\mathcal{B}/\mathcal{C}) + H(\mathcal{C}/\mathcal{B}); \quad \mathcal{B},\mathcal{C}\varepsilon Z$$

(cf. [5], Theorem 7, P. 265).

Note 2.2.3: For almost every $x\varepsilon\Omega$, a metric on the set Z of equivalence classes of finite sub-σ-algebras of $\mathcal{R}$ may be defined as follows:

$$d^1_{\mathcal{R}_0} = d^1_{\mathcal{R}_0}(\mathcal{B},\mathcal{C}) = H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}) + H_{\mathcal{R}_0}(\mathcal{C}/\mathcal{B}); \quad \mathcal{B},\mathcal{C}\varepsilon Z.$$

We know that if d is a metric defined on some set, then, $\frac{d}{1+d}$ is also a metric on the same set. Thus another metric on the set Z for almost every $x\varepsilon\Omega$ is given by

$$d^2_{\mathcal{R}_0}(\mathcal{B},\mathcal{C}) = \frac{H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}) + H_{\mathcal{R}_0}(\mathcal{C}/\mathcal{B})}{1 + H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}) + H_{\mathcal{R}_0}(\mathcal{C}/\mathcal{B})}; \quad \mathcal{B},\mathcal{C}\varepsilon Z.$$

Now each of these metrics generates one more metric given by

$$d^3_{\mathcal{R}_0}(\mathcal{B},\mathcal{C}) = \int\left[H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}) + H_{\mathcal{R}_0}(\mathcal{C}/\mathcal{B})\right] dP; \quad \mathcal{B},\mathcal{C}\varepsilon Z.$$

and $$d^4_{\mathcal{R}_0}(\mathcal{B},\mathcal{C}) = \int \frac{H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}) + H_{\mathcal{R}_0}(\mathcal{C}/\mathcal{B})}{1 + H_{0}(\mathcal{B}/\mathcal{C}) + H_{\mathcal{R}_0}(\mathcal{C}/\mathcal{B})} dP; \quad \mathcal{B},\mathcal{C}\varepsilon Z.$$

CHAPTER III

A COMPLETE METRIC SPACE OF SUB-σ-ALGEBRAS

3.1 Metric Space of Sub-σ-Algebras

C. Rajski proved that the functional

(1) $d(x,y) = 1 - \frac{I(x,y)}{H(x,y)}$ where $H(x,y) \neq 0$ and $I(x,y) = H(x) + H(y) - H(x,y)$

is a distance in the set X of all discrete probability distributions (cf. [9], Theorem p. 372). It is a consequence of this theorem, that in Information Theory the dependence between the transmitted and the received discrete signals may be expressed as a distance.

Replacing x by $\mathcal{R}_0 \subseteq \mathcal{R}$ and y be $\mathcal{R}_0' \subseteq \mathcal{R}$ in (1), we prove in Theorem 3.1.1 that,

$$d(\mathcal{R}_0,\mathcal{R}_0') = 1 - \frac{I(\mathcal{R}_0,\mathcal{R}_0')}{H(\mathcal{R}_0 \vee \mathcal{R}_0')}, \quad H(\mathcal{R}_0 \vee \mathcal{R}_0') \neq 0$$

is a metric in the set of all equivalence classes of sub-σ-algebras of $\mathcal{R}$.

It is observed that the theorem given by C. Rajski ([9]) is a particular case of Theorem 3.1.1 and the proof is by comparison concise. In order to show this, we need the following lemma:

Lemma 3.1.1: If $(\Omega,\mathfrak{F},\mu)$ is a probability space where $\Omega = \{w: 0 \leq w \leq 1\}$, $\mathfrak{F}$ is the σ-algebra consisting of all Borel subsets of Ω and μ is Lebesgue measure, then corresponding to every discrete probability distribution there exists a sub-σ-algebra of $\mathcal{F}$.

<u>Proof</u>: A discrete probability distribution is the collection of various values of a random variable which correspond to the atoms of a finite or countable measurable partition as the case may be together with the probability measure of these atoms.

To prove the lemma, we consider the following discrete probability distribution:

$$\text{Prob}(X = x_i) = p_i;\ i = 1,2,\ldots;\ \sum_i p_i = 1.$$

Let the atoms corresponding to the values $x_1, x_2, x_3, \ldots$ of random variables be as follows:

$\{x:\ 0 \leq x < p_1\}$, $\{x:\ p_1 \leq x < p_1 + p_2\}$, $\{x:\ p_1 + p_2 \leq x < p_1 + p_2 + p_3\}, \ldots$

The above subsets of Ω, obviously forms a finite or countable measurable partition of $\Omega = [0,1]$ and hence there exists a sub-σ-algebra of $\mathcal{F}$ corresponding to this countable measurable partition.

<u>Theorem 3.1.1</u>: If Z is the set of equivalence classes of sub-σ-algebras of $\mathcal{R}$ with finite entropy then the functional

$$d(\mathcal{R}_0, \mathcal{R}_0') = 1 - \frac{I(\mathcal{R}_0, \mathcal{R}_0')}{H(\mathcal{R}_0 \vee \mathcal{R}_0')};\quad H(\mathcal{R}_0 \vee \mathcal{R}_0') \neq 0$$

is a distance in the set Z.

<u>Proof:</u>

a) $$d(\mathcal{R}_0,\mathcal{R}_0') = 1 - \frac{I(\mathcal{R}_0,\mathcal{R}_0')}{H(\mathcal{R}_0 \vee \mathcal{R}_0')} = 1 - \frac{H(\mathcal{R}_0) + H(\mathcal{R}_0') - H(\mathcal{R}_0 \vee \mathcal{R}_0')}{H(\mathcal{R}_0 \vee \mathcal{R}_0')}$$

$$= \frac{\{H(\mathcal{R}_0 \vee \mathcal{R}_0') - H(\mathcal{R}_0)\} + \{H(\mathcal{R}_0 \vee \mathcal{R}_0') - H(\mathcal{R}_0')\}}{H(\mathcal{R}_0 \vee \mathcal{R}_0')}$$

$$= \frac{H(\mathcal{R}_0/\mathcal{R}_0') + H(\mathcal{R}_0'/\mathcal{R}_0)}{H(\mathcal{R}_0 \vee \mathcal{R}_0')} \geq 0.$$

b) $d(\mathcal{R}_0,\mathcal{R}_0') = 0 \Longrightarrow H(\mathcal{R}_0'/\mathcal{R}_0) + H(\mathcal{R}_0/\mathcal{R}_0') = 0$

$\Longrightarrow \mathcal{R}_0' \subseteq \mathcal{R}_0$ up to equivalence and $\mathcal{R}_0 \subseteq \mathcal{R}_0'$ up to equivalence

$\Longrightarrow \mathcal{R}_0 = \mathcal{R}_0'$ up to equivalence.

c) Let $\mathcal{R}_0 = \mathcal{R}_0'$ up to equivalence.

therefore $H(\mathcal{R}_0/\mathcal{R}_0') = 0$ and $H(\mathcal{R}_0'/\mathcal{R}_0) = 0$

$$\Longrightarrow \frac{H(\mathcal{R}_0/\mathcal{R}_0') + H(\mathcal{R}_0'/\mathcal{R}_0)}{H(\mathcal{R}_0 \vee \mathcal{R}_0')} = 0 \Longrightarrow d(\mathcal{R}_0,\mathcal{R}_0') = 0.$$

d) Now we establish the triangle inequality. We have

$H(\mathcal{R}_0 \vee \mathcal{R}_0') \leq H(\mathcal{R}_0 \vee \mathcal{R}_0' \vee \mathcal{R}_0'')$; $\mathcal{R}_0,\mathcal{R}_0',\mathcal{R}_0'' \varepsilon Z$

$$\Longrightarrow \frac{H(\mathcal{R}_0 \vee \mathcal{R}_0') - H(\mathcal{R}_0')}{H(\mathcal{R}_0 \vee \mathcal{R}_0')} \leq \frac{H(\mathcal{R}_0 \vee \mathcal{R}_0' \vee \mathcal{R}_0'') - H(\mathcal{R}_0')}{H(\mathcal{R}_0 \vee \mathcal{R}_0' \vee \mathcal{R}_0'')}$$

$$\Rightarrow \frac{H(\mathcal{R}_0/\mathcal{R}_0')}{H(\mathcal{R}_0 V \mathcal{R}_0')} \leq \frac{H(\mathcal{R}_0 V \mathcal{R}_0' V \mathcal{R}_0'') - H(\mathcal{R}_0')}{H(\mathcal{R}_0 V \mathcal{R}_0' V \mathcal{R}_0'')}$$

$$= \frac{H(\mathcal{R}_0 V \mathcal{R}_0' V \mathcal{R}_0'') - H(\mathcal{R}_0' V \mathcal{R}_0'') + H(\mathcal{R}_0' V \mathcal{R}_0'') - H(\mathcal{R}_0')}{H(\mathcal{R}_0 V \mathcal{R}_0' V \mathcal{R}_0'')}$$

$$= \frac{H(\mathcal{R}_0/\mathcal{R}_0' V \mathcal{R}_0'')}{H(\mathcal{R}_0 V \mathcal{R}_0' V \mathcal{R}_0'')} + \frac{H(\mathcal{R}_0''/\mathcal{R}_0')}{H(\mathcal{R}_0 V \mathcal{R}_0' V \mathcal{R}_0'')} \leq \frac{H(\mathcal{R}_0/\mathcal{R}_0' V \mathcal{R}_0'')}{H(\mathcal{R}_0 V \mathcal{R}_0' V \mathcal{R}_0'')} + \frac{H(\mathcal{R}_0''/\mathcal{R}_0')}{H(\mathcal{R}_0' V \mathcal{R}_0'')} .$$

Hence $\frac{H(\mathcal{R}_0/\mathcal{R}_0')}{H(\mathcal{R}_0 V \mathcal{R}_0')} \leq \frac{H(\mathcal{R}_0/\mathcal{R}_0' V \mathcal{R}_0'')}{H(\mathcal{R}_0 V \mathcal{R}_0' V \mathcal{R}_0'')} + \frac{H(\mathcal{R}_0''/\mathcal{R}_0')}{H(\mathcal{R}_0' V \mathcal{R}_0'')}$. —— (1)

Now, we prove that $\frac{H(\mathcal{R}_0/\mathcal{R}_0' V \mathcal{R}_0'')}{H(\mathcal{R}_0 V \mathcal{R}_0' V \mathcal{R}_0'')} \leq \frac{H(\mathcal{R}_0/\mathcal{R}_0'')}{H(\mathcal{R}_0 V \mathcal{R}_0'')}$. —— (2)

We have $H(\mathcal{R}_0 V \mathcal{R}_0' V \mathcal{R}_0'') = H(\mathcal{R}_0 V \mathcal{R}_0'') + H(\mathcal{R}_0'/\mathcal{R}_0 V \mathcal{R}_0'')$

$$\Rightarrow H(\mathcal{R}_0 V \mathcal{R}_0' V \mathcal{R}_0'') \geq H(\mathcal{R}_0 V \mathcal{R}_0'') \Rightarrow \frac{H(\mathcal{R}_0/\mathcal{R}_0' V \mathcal{R}_0'')}{H(\mathcal{R}_0 V \mathcal{R}_0' V \mathcal{R}_0'')}$$

$$\leq \frac{H(\mathcal{R}_0/\mathcal{R}_0' V \mathcal{R}_0'')}{H(\mathcal{R}_0 V \mathcal{R}_0'')}$$

$$\Rightarrow \frac{H(\mathcal{R}_0/\mathcal{R}_0' V \mathcal{R}_0'')}{H(\mathcal{R}_0 V \mathcal{R}_0' V \mathcal{R}_0'')} \leq \frac{H(\mathcal{R}_0/\mathcal{R}_0'')}{H(\mathcal{R}_0 V \mathcal{R}_0'')} .$$

Thus (2) is proved.

Now from (1) and (2) we have

$$\frac{H(\mathcal{R}_0/\mathcal{R}_0')}{H(\mathcal{R}_0 V \mathcal{R}_0')} \leq \frac{H(\mathcal{R}_0/\mathcal{R}_0'')}{H(\mathcal{R}_0 V \mathcal{R}_0'')} + \frac{H(\mathcal{R}_0''/\mathcal{R}_0')}{H(\mathcal{R}_0' V \mathcal{R}_0'')}. \quad \text{—— (3)}$$

Interchanging the roles of $\mathcal{R}_0$ and $\mathcal{R}_0'$, we obtain

$$\frac{H(\mathcal{R}_0'/\mathcal{R}_0)}{H(\mathcal{R}_0 V \mathcal{R}_0')} \leq \frac{H(\mathcal{R}_0'/\mathcal{R}_0'')}{H(\mathcal{R}_0' V \mathcal{R}_0'')} + \frac{H(\mathcal{R}_0''/\mathcal{R}_0)}{H(\mathcal{R}_0 V \mathcal{R}_0'')}. \text{ —— } (4)$$

Adding (3) and (4) we obtain

$$\frac{H(\mathcal{R}_0/\mathcal{R}_0') + H(\mathcal{R}_0'/\mathcal{R}_0)}{H(\mathcal{R}_0 V \mathcal{R}_0')} \leq \frac{H(\mathcal{R}_0/\mathcal{R}_0'') + H(\mathcal{R}_0''/\mathcal{R}_0)}{H(\mathcal{R}_0 V \mathcal{R}_0'')} + \frac{H(\mathcal{R}_0'/\mathcal{R}_0'') + H(\mathcal{R}_0''/\mathcal{R}_0')}{H(\mathcal{R}_0' V \mathcal{R}_0'')}$$

$$\Longrightarrow d(\mathcal{R}_0,\mathcal{R}_0') \leq d(\mathcal{R}_0,\mathcal{R}_0'') + d(\mathcal{R}_0'',\mathcal{R}_0'). \text{ —— } (5)$$

<u>Note 3.1.1</u>: Lemma 3.1.1 shows that the set of sub-σ-algebras, each corresponding to a discrete probability distribution, is a subset of all possible sub-σ-algebras of $\mathfrak{F}$.

Thus if the probability space under consideration is $(\Omega,\mathfrak{F},\mu)$ as defined in Lemma 3.1.1, then the proof of C. Rajski's Theorem ([9]) on a metric space of discrete probability distributions follows immediately.

<u>Note 3.1.2</u>: Replacing $\mathcal{R}_0''$ by $\mathcal{R}_0 V \mathcal{R}_0'$ we obtain

$$d(\mathcal{R}_0,\mathcal{R}_0'') = \frac{H(\mathcal{R}_0/\mathcal{R}_0'') + H(\mathcal{R}_0''/\mathcal{R}_0)}{H(\mathcal{R}_0 V \mathcal{R}_0'')} = \frac{H(\mathcal{R}_0/\mathcal{R}_0 V \mathcal{R}_0') + H(\mathcal{R}_0 V \mathcal{R}_0'/\mathcal{R}_0)}{H(\mathcal{R}_0 V \mathcal{R}_0')}$$

$$= \frac{H(\mathcal{R}_0'/\mathcal{R}_0)}{H(\mathcal{R}_0 V \mathcal{R}_0')}. \text{ —— } (6)$$

Similarly $d(\mathcal{R}_0'',\mathcal{R}_0') = \dfrac{H(\mathcal{R}_0/\mathcal{R}_0')}{H(\mathcal{R}_0 V \mathcal{R}_0')}$. —— (7)

Adding (6) and (7), we obtain

$$d(\mathcal{R}_0,\mathcal{R}_0') = \frac{H(\mathcal{R}_0/\mathcal{R}_0') + H(\mathcal{R}_0'/\mathcal{R}_0)}{H(\mathcal{R}_0 V \mathcal{R}_0')} = d(\mathcal{R}_0,\mathcal{R}_0'') + d(\mathcal{R}_0'',\mathcal{R}_0').$$

Thus the inequality in (5) becomes equality if $\mathcal{R}_0''$ is replaced by $\mathcal{R}_0 V \mathcal{R}_0'$.

Corollary 3.1.1:

$$d = \frac{H(\mathcal{R}_0/\mathcal{R}_0') + H(\mathcal{R}_0'/\mathcal{R}_0)}{H(\mathcal{R}_0 V \mathcal{R}_0')} \Bigg/ \left(1 + \frac{H(\mathcal{R}_0/\mathcal{R}_0') + H(\mathcal{R}_0'/\mathcal{R}_0)}{H(\mathcal{R}_0 V \mathcal{R}_0')}\right)$$

$$= \frac{H(\mathcal{R}_0/\mathcal{R}_0') + H(\mathcal{R}_0'/\mathcal{R}_0)}{H(\mathcal{R}_0 V \mathcal{R}_0') + H(\mathcal{R}_0/\mathcal{R}_0') + H(\mathcal{R}_0'/\mathcal{R}_0)}$$

then (Z,d) is also a bounded metric space.

Definition 3.1.1: A metric space (X,d) is convex if for any two distinct elements $x,y \in X$, there exists an element z different from both x and y and such that

$$d(x,y) = d(x,z) + d(y,z)$$

Theorem 3.1.2: If Z_1 is the sub-space of sub-σ-algebras of $\mathcal{R}$ which are such that for any two sub-σ-algebras one is not contained in the other, then (Z_1,d) is a convex metric space.

Proof: The proof immediately follows from Theorem 3.1.1, Note 3.1.2 and the Definition 3.1.1.

3.2 Completeness of the Metric Space of Theorem 3.1.1

Theorem 3.2.1: The metric space (Z,d) of Theorem 3.1.1 is a complete metric space.

Proof: We are to show that any fundamental sequence $\mathcal{R}_1, \mathcal{R}_2, \ldots$ converges in Z. It is sufficient to consider the case $d(\mathcal{R}_n, \mathcal{R}_{n+p}) < \frac{1}{2^n}$ $(p > 0)$; for from any fundamental sequence we can select a subsequence satisfying this condition and a fundamental sequence that contains a convergent subsequence is convergent.

We put

$$\bar{\mathcal{R}} = \bigwedge_{\ell=1}^{\infty} \bigvee_{k=\ell}^{\infty} \mathcal{R}_k$$

and show that $\bar{\mathcal{R}} \varepsilon Z$ and $d(\bar{\mathcal{R}}, \mathcal{R}_n) \longrightarrow 0$.

We prove this theorem under the assumption that $H(\mathcal{R}_0) \leq k$ (a fixed non-zero and positive constant) $\forall \mathcal{R}_0 \varepsilon Z$.

We have

$$d(\mathcal{R}_n, \bar{\mathcal{R}}) \leq d(\mathcal{R}_n, \bigvee_{k=n}^{\infty} \mathcal{R}_k) + d(\bigvee_{k=n}^{\infty} \mathcal{R}_k, \bar{\mathcal{R}}). \quad \text{—— (1)}$$

$$d(\mathcal{R}_n, \bigvee_{k=n}^{\infty} \mathcal{R}_k) = \frac{H(\mathcal{R}_n / \bigvee_{k=n}^{\infty} \mathcal{R}_k) + H(\bigvee_{k=n}^{\infty} \mathcal{R}_k / \mathcal{R}_n)}{H(\bigvee_{k=n}^{\infty} \mathcal{R}_k)}$$

$$= \frac{H(\bigvee_{k=n}^{\infty} \mathcal{R}_k / \mathcal{R}_n)}{H(\bigvee_{k=n}^{\infty} \mathcal{R}_k)} = \frac{H(\bigvee_{k=n+1}^{\infty} \mathcal{R}_k / \mathcal{R}_n)}{H(\bigvee_{k=n}^{\infty} \mathcal{R}_k)}. \quad \text{—— (2)}$$

Now,

$$d\Big(\bigvee_{k=n}^{\infty}\mathcal{R}_k,\bar{\mathcal{R}}\Big) = \frac{H\big(\bar{\mathcal{R}}/\bigvee_{k=n}^{\infty}\mathcal{R}_k\big) + \big(\bigvee_{k=n}^{\infty}\mathcal{R}_k/\bar{\mathcal{R}}\big)}{H\big(\bigvee_{k=n}^{\infty}\mathcal{R}_k \vee \bar{\mathcal{R}}\big)}$$

but $\bar{\mathcal{R}} \subseteq \bigvee_{k=n}^{\infty}\mathcal{R}_k$.

Therefore

$$d\Big(\bigvee_{k=n}^{\infty}\mathcal{R}_k,\bar{\mathcal{R}}\Big) = \frac{H\big(\bigvee_{k=n}^{\infty}\mathcal{R}_k/\bar{\mathcal{R}}\big)}{H\big(\bigvee_{k=n}^{\infty}\mathcal{R}_k\big)} = \frac{H\big(\bigvee_{k=n}^{\infty}\mathcal{R}_k\big) - H(\bar{\mathcal{R}})}{H\big(\bigvee_{k=n}^{\infty}\mathcal{R}_k\big)}$$

$$= 1 - \frac{H(\bar{\mathcal{R}})}{H\big(\bigvee_{k=n}^{\infty}\mathcal{R}_k\big)} . \qquad (3)$$

From (1), (2) and (3) we have

$$d(\mathcal{R}_n,\bar{\mathcal{R}}) \leq \frac{H\big(\bigvee_{k=n+1}^{\infty}\mathcal{R}_k/\mathcal{R}_n\big)}{H\big(\bigvee_{k=n}^{\infty}\mathcal{R}_k\big)} + 1 - \frac{H(\bar{\mathcal{R}})}{H\big(\bigvee_{k=n}^{\infty}\mathcal{R}_k\big)} . \qquad (4)$$

Now for $\ell > n$

$$H\Big(\bigvee_{k=\ell}^{\infty}\mathcal{R}_k/\bigvee_{k=n}^{\ell-1}\mathcal{R}_k\Big) = H\Big(\mathcal{R}_\ell/\bigvee_{k=n}^{\ell-1}\mathcal{R}_k\Big) + H\Big(\bigvee_{k=\ell+1}^{\infty}\mathcal{R}_k/\bigvee_{k=n}^{\ell}\mathcal{R}_k\Big)$$

$$\Rightarrow \sum_{\ell=n+1}^{\infty} H\Big(\bigvee_{k=\ell}^{\infty}\mathcal{R}_k/\bigvee_{k=n}^{\ell-1}\mathcal{R}_k\Big) = \sum_{\ell=n+1}^{\infty} H\Big(\mathcal{R}_\ell/\bigvee_{k=n}^{\ell-1}\mathcal{R}_k\Big) + \sum_{\ell=n+1}^{\infty} H\Big(\bigvee_{k=\ell+1}^{\infty}\mathcal{R}_k/\bigvee_{k=n}^{\ell}\mathcal{R}_k\Big)$$

$$\Rightarrow H\Big(\bigvee_{k=n+1}^{\infty}\mathcal{R}_k/\mathcal{R}_n\Big) = \sum_{\ell=n+1}^{\infty} H\Big(\mathcal{R}_\ell/\bigvee_{k=n}^{\ell-1}\mathcal{R}_k\Big) \leq \sum_{\ell=n+1}^{\infty} H(\mathcal{R}_\ell/\mathcal{R}_{\ell-1})$$

$$\Longrightarrow \frac{H(\bigvee_{k=n+1}^{\infty} \mathcal{R}_k/\mathcal{R}_n)}{H(\bigvee_{k=n}^{\infty} \mathcal{R}_k)} \leq \frac{\sum_{\ell=n+1}^{\infty} H(\mathcal{R}_\ell/\mathcal{R}_{\ell-1})}{H(\bigvee_{k=n}^{\infty} \mathcal{R}_k)}$$

$$\Longrightarrow \frac{H(\bigvee_{k=n+1}^{\infty} \mathcal{R}_k/\mathcal{R}_n)}{H(\bigvee_{k=n}^{\infty} \mathcal{R}_k)} \leq \frac{H(\mathcal{R}_{n+1}/\mathcal{R}_n)}{H(\bigvee_{k=n}^{\infty} \mathcal{R}_k)} + \frac{H(\mathcal{R}_{n+2}/\mathcal{R}_{n+1})}{H(\bigvee_{k=n}^{\infty} \mathcal{R}_k)} + \cdots$$

$$\Longrightarrow \frac{H(\bigvee_{k=n+1}^{\infty} \mathcal{R}_k/\mathcal{R}_n)}{H(\bigvee_{k=n}^{\infty} \mathcal{R}_k)} \leq \frac{H(\mathcal{R}_{n+1}/\mathcal{R}_n)}{H(\mathcal{R}_n \vee \mathcal{R}_{n+1})} + \frac{H(\mathcal{R}_{n+2}/\mathcal{R}_{n+1})}{H(\mathcal{R}_{n+1} \vee \mathcal{R}_{n+2})} + \cdots$$

$$\Longrightarrow \frac{H(\bigvee_{k=n+1}^{\infty} \mathcal{R}_k/\mathcal{R}_n)}{H(\bigvee_{k=n}^{\infty} \mathcal{R}_k)} \leq d(\mathcal{R}_n, \mathcal{R}_{n+1}) + d(\mathcal{R}_{n+1}, \mathcal{R}_{n+2}) + \cdots$$

$$< \frac{1}{2^n} + \frac{1}{2^{n+1}} + \frac{1}{2^{n+2}} + \cdots$$

$$= \frac{\frac{1}{2^n}}{1-\frac{1}{2}} = \frac{1}{2^{n-1}}. \qquad (5)$$

Now from (4) and (5) we obtain

$$d(\mathcal{R}_n, \bar{\mathcal{R}}) \leq \frac{1}{2^{n-1}} + 1 - \frac{H(\bar{\mathcal{R}})}{H(\bigvee_{k=n}^{\infty} \mathcal{R}_k)}.$$

Hence

$$\lim_{n\to\infty} d(\mathcal{R}_n,\bar{\mathcal{R}}) \leq \lim_{n\to\infty} \frac{1}{2^{n-1}} + 1 - \lim_{n\to\infty} \frac{H(\bar{\mathcal{R}})}{H(\bigvee_{k=n}^{\infty}\mathcal{R}_k)}$$

$$\Rightarrow \lim_{n\to\infty} d(\mathcal{R}_n,\bar{\mathcal{R}}) \leq 0 + 1 - \frac{H(\bar{\mathcal{R}})}{H(\bar{\mathcal{R}})} \quad \text{(see [10], 5.8 - P. 16)}$$

$$\Rightarrow \lim_{n\to\infty} d(\mathcal{R}_n,\bar{\mathcal{R}}) \leq 0 \Rightarrow \lim_{n\to\infty} d(\mathcal{R}_n,\bar{\mathcal{R}}) = 0.$$

Now we prove that $H(\bar{\mathcal{R}})$ is finite.

Since $\lim_{n\to\infty} d(\bar{\mathcal{R}},\mathcal{R}_n) = 0$

we have

$$d(\bar{\mathcal{R}},\mathcal{R}_n) < \tfrac{1}{2} \text{ (say) for some } n$$

$$\Rightarrow \frac{H(\bar{\mathcal{R}}/\mathcal{R}_n) + H(\mathcal{R}_n/\bar{\mathcal{R}})}{H(\bar{\mathcal{R}}\vee\mathcal{R}_n)} < \frac{1}{2}$$

$$\Rightarrow \frac{H(\bar{\mathcal{R}}/\mathcal{R}_n)}{H(\bar{\mathcal{R}}\vee\mathcal{R}_n)} < \frac{1}{2} \Rightarrow \frac{H(\bar{\mathcal{R}}\vee\mathcal{R}_n) - H(\mathcal{R}_n)}{H(\bar{\mathcal{R}}\vee\mathcal{R}_n)} < \frac{1}{2}$$

$$\Rightarrow 1 - \frac{H(\mathcal{R}_n)}{H(\bar{\mathcal{R}}\vee\mathcal{R}_n)} < \frac{1}{2} \Rightarrow \frac{1}{2} < \frac{H(\mathcal{R}_n)}{H(\bar{\mathcal{R}}\vee\mathcal{R}_n)}$$

$$\Rightarrow H(\bar{\mathcal{R}}\vee\mathcal{R}_n) < 2H(\mathcal{R}_n) \Rightarrow H(\bar{\mathcal{R}}) \leq H(\bar{\mathcal{R}}\vee\mathcal{R}_n) < 2H(\mathcal{R}_n)$$

$\Rightarrow H(\bar{\mathcal{R}})$ is a finite constant.

<u>Note 3.2.1</u>: Let Z be the set of equivalence classes of finite sub-σ-algebras of $\mathcal{A}$. Let $\mathcal{B},\mathcal{C}\in Z$

and $H_{\mathcal{R}_0}(\mathcal{B} \vee \mathcal{C}) \neq 0$, then, as in Note 2.2.3,

for almost every $x \in \Omega$ metrics on the set Z may be defined as follows:

$$d^5_{\mathcal{R}_0} = d^5_{\mathcal{R}_0}(\mathcal{B},\mathcal{C}) = \frac{H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}) + H_{\mathcal{R}_0}(\mathcal{C}/\mathcal{B})}{H_{\mathcal{R}_0}(\mathcal{B} \vee \mathcal{C})}.$$

$$d^6_{\mathcal{R}_0} = d^6_{\mathcal{R}_0}(\mathcal{B},\mathcal{C}) = \frac{H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}) + H_{\mathcal{R}_0}(\mathcal{C}/\mathcal{B})}{H_{\mathcal{R}_0}(\mathcal{B} \vee \mathcal{C}) + H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}) + H_{\mathcal{R}_0}(\mathcal{C}/\mathcal{B})}.$$

The other two metrics on the set Z are:

$$d^7_{\mathcal{R}_0} = d^7_{\mathcal{R}_0}(\mathcal{B},\mathcal{C}) = \int \frac{H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}) + H_{\mathcal{R}_0}(\mathcal{C}/\mathcal{B})}{H_{\mathcal{R}_0}(\mathcal{B} \vee \mathcal{C})}\, dP.$$

$$d^8_{\mathcal{R}_0} = d^8_{\mathcal{R}_0}(\mathcal{B},\mathcal{C}) = \int \frac{H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}) + H_{\mathcal{R}_0}(\mathcal{C}/\mathcal{B})}{H_{\mathcal{R}_0}(\mathcal{B} \vee \mathcal{C}) + H_{\mathcal{R}_0}(\mathcal{B}/\mathcal{C}) + H_{\mathcal{R}_0}(\mathcal{C}/\mathcal{B})}\, dP.$$

3.3 Families of Metrics on the Set of Finite Sub-σ-Algebras

From Section 1.4.5, we know that the generalized conditional entropy $I_\alpha(\mathcal{R}_0'/\mathcal{R}_0)$ for the finite sub-σ-algebras $\mathcal{R}_0$ and $\mathcal{R}_0'$ of $\mathcal{R}$ satisfies the properties (i) to (vi). Properties (iii) to (vi) are true for $\alpha \geq 0$ and (i) and (ii) are valid for $\alpha > 0$ and $\alpha \geq 1$ respectively.

Now we define two new entropies on the set Z of all finite sub-σ-algebras of $\mathcal{R}$ as follows:

$$d_\alpha^1 = d_\alpha^1(\mathcal{R}_0,\mathcal{R}_0') = I_\alpha(\mathcal{R}_0/\mathcal{R}_0') + I_\alpha(\mathcal{R}_0'/\mathcal{R}_0);\ \mathcal{R}_0,\mathcal{R}_0'\varepsilon Z,$$

$$d_\alpha^2 = d_\alpha^2(\mathcal{R}_0,\mathcal{R}_0') = \frac{I_\alpha(\mathcal{R}_0/\mathcal{R}_0') + I_\alpha(\mathcal{R}_0'/\mathcal{R}_0)}{I_\alpha(\mathcal{R}_0 V \mathcal{R}_0')};\ \mathcal{R}_0,\mathcal{R}_0'\varepsilon Z$$

$$I_\alpha(\mathcal{R}_0 V \mathcal{R}_0') \neq 0.$$

It can be easily verified that (Z,d_α^1) is a complete metric space for all $\alpha > 0$ and (Z,d_α^2) is a metric space for all $\alpha \geq 1$.

<u>Note 3.3.1</u>: For $\alpha = 1$, $I_\alpha(\mathcal{R}_0'/\mathcal{R}_0)$ reduces to Shannon's conditional entropy. Thus if Z is the set of equivalence classes of finite sub-σ-algebras of $\mathcal{R}$, then

$D_1 = \{d_\alpha^1;\ \alpha\varepsilon(0,\infty)\}$ and $D_2 = \{d_\alpha^2;\ \alpha\varepsilon[1,\infty)\}$ represent

two families of metrics on Z such that the metric given $d = H(\mathcal{R}_0/\mathcal{R}_0') + H(\mathcal{R}_0'/\mathcal{R}_0)$ belongs to D_1 and the metric

given by $d = \dfrac{H(\mathcal{R}_0/\mathcal{R}_0') + H(\mathcal{R}_0'/\mathcal{R}_0)}{H(\mathcal{R}_0 V \mathcal{R}_0')}$ belongs to D_2.

References

1. Behara, M. and Nath, P. (1971), Additive and Non-additive Entropies of Finite Measurable Partitions, Lecture notes in Mathematics, Probability and Information Theory II, Springer-Verlag.

2. Billingsley, P., Ergodic Theory and Information, New York, Wiley 1965.

3. Brown, T. A. (1963), Entropy and Conjugacy, Ann. Math. Stat., 34, 226-232.

4. Copson, E. T., Metric Spaces, Cambridge University Press - 1968.

5. Jacobs, K., Lecture Notes on Ergodic Theory, Part II, Matematisk Institut, Aarhus Universitet.

6. Loeve, M., Probability Theory, New York, Van Nostrand 1955.

7. Neveu, J., Mathematical Foundations of the Calculas of Probability, Holden-Day, Inc., San Francisco, London, Amsterdam.

8. Nijst, A. G. P. M., Some Remarks on Conditional Entropy, 22 Z. Wahrscheinlichkeitstheorie Verw. Geb; Bd. 12, P. 307-319.

9. Rajski, C., A Metric Space of Discrete Probability Distributions, Information and Control 4, 371-377 (1961),

10. Rokhlin, V. A., Lectures on the Entropy Theory of Measure Preserving Transformations, Russ. Math. Surveys 22, 1-52.(1967).

11. Royden, H. L., Real Analysis, Second Edition, The Macmillan Company, New York, Collier-MacMillan Limited, London.