Conditional (p, q) entropy

By R. W. ALLEN, Jr. (Florence, South Carolina)

1. Introduction

Let $(\Omega, \mathscr{A}, P)$ be a probability space. The purpose of this paper is to investigate properties of a conditional entropy of a finite σ-subfield of the σ-field $\mathscr{A}$. SHANNON [9] has defined an entropy based on information theoretic principles, while DARÓCZY [3] has defined a more general family of entropies that is based primarily on properties of mean values. A conditional entropy has been formulated based on the Shannon entropy (see, for example, [1]), by replacing probabilities with conditional probabilities and then taking the expected value.

Several interesting and useful results have been obtained using this conditional entropy. In this paper we shall define a family of conditional entropies based on the family of Daróczy (p, q)-entropies and consider the properties of this family of (p, q)-entropies. In particular we shall consider which properties are common to both conditional Shannon and conditional (p, q)-entropy and which properties are unique to one or the other.

Throughout the entire paper, a basic working knowledge of conditional expectations will be assumed. Further, the following notation will be used throughout: $\Pi(\mathscr{F})$ will denote the atoms of a finite σ-field $\mathscr{F}$, $\sum\limits_{\mathscr{F}}$ will denote the summation over the atoms of $\mathscr{F}$, $P^{\mathscr{B}}(F)$ will denote the conditional probability of F with respect to the σ-field $\mathscr{B}$, and $E(\cdot)$ will denote the expected value.

2. Definition of conditional (p, q)-entropy and properties of (p, q)-entropy

We consider the following

Definition: Let $\mathscr{F}$ be a finite σ-subfield of $\mathscr{A}$ and $\mathscr{B}$ any σ-subfield of $\mathscr{A}$. Then the $\mathscr{B}$-conditional (p, q)-entropy of $\mathscr{F}$, denoted by $H^{\mathscr{B}}_{p,q}(\mathscr{F})$, is defined by

$$H^{\mathscr{B}}_{p,q}(\mathscr{F}) = E\Big[-\log\Big\{\sum_{\mathscr{F}} P^{\mathscr{B}}(F)]^{p+q}\Big/\sum_{\mathscr{F}}[P^{\mathscr{B}}(F)]^{q}\Big\}^{1/p}\Big], \qquad p\neq 0. \tag{1}$$

$$H^{\mathscr{B}}_{0,q}(\mathscr{F}) = E\Big[-\sum_{\mathscr{F}}[P^{\mathscr{B}}(F)]^{q}\log P^{\mathscr{B}}(F)\Big/\sum_{\mathscr{F}}[P^{\mathscr{B}}(F)]^{q}\Big], \tag{2}$$

respectively, where p and q are arbitrary real numbers and (p, q) is the order of the function.

One can see that the $\mathscr{B}$-conditional Shannon entropy of $\mathscr{F}$ is such an entropy of order (0, 1). Further, we shall define $\mathscr{B}$-conditional Rényi entropy of $\mathscr{F}$ of order α (see [8]) to be the above entropy of order $(p, 1)$, with $p+1 = \alpha$, and it will be denoted $H_\alpha^{\mathscr{B}}(\mathscr{F})$.

As immediate consequences of the definition of (p, q)-entropy of $\mathscr{F}$, we have the following

Lemma 1. *Let $\mathscr{F}$ be a finite σ- subfield of $\mathscr{A}$. Then for all real α,*

$$H_\alpha(\mathscr{F}) = H_{(\alpha-1,1)}(\mathscr{F}) = H_{(1-\alpha,\alpha)}(\mathscr{F}) \tag{3}$$

and for all $\alpha \neq 0$,

$$H_{(-\alpha,\alpha)}(\mathscr{F}) = 1/\alpha \log N - \frac{1-\alpha}{\alpha} H_\alpha(\mathscr{F}), \tag{4}$$

where N is the cardinality of $\Pi(\mathscr{F})$.

Now using a Lagrange multiplier technique to maximize or minimize a function $f(x_1, x_2, \ldots, x_N)$ under the condition $x_1+x_2+\cdots+x_N = 1$ plus the definition of Rényi entropy of order α we have

Lemma 2. *Let $\mathscr{F}$ be a finite σ-subfield of $\mathscr{A}$ and let N=cardinality of $\Pi(\mathscr{F})$. Then*

$$H_\alpha(\mathscr{F}) \leqq \log N \quad \text{if} \quad \alpha \geqq 0, \qquad H_\alpha(\mathscr{F}) \geqq \log N \quad \textit{if} \quad \alpha \leqq 0.$$

Daróczy in his work ([2], Theorem 5) has proved the following result that is central to the boundedness properties that we shall obtain.

Lemma 3. *Let $\mathscr{F}$ be a finite σ-subfield of $\mathscr{A}$. Then if $p_1 \leqq p_2$ $H_{p_1, q}(\mathscr{F}) \geqq H_{p_2, q}(\mathscr{F})$ and if $q_1 \leqq q_2$, $H_{p, q_1}(\mathscr{F}) \geqq H_{p, q_2}(\mathscr{F})$; that is, $H_{p,q}(\mathscr{F})$ is a decreasing function of both p and q.*

3. Boundedness properties of (p, q) entropy and conditional (p,q) entropy

We now seek boundedness conditions on $H_{p,q}(\mathscr{F})$ and $H_{p,q}^B(\mathscr{F})$. Using the above lemmas we can prove the following

Theorem 1. *Let $\mathscr{F}$ be a finite σ-subfield of $\mathscr{A}$ and let N=cardinality of $\Pi(\mathscr{F})$. If p and q are real numbers such that either $q \geqq 1$ and $p+q \geqq 0$ or $0 \leqq q < 1$ and $p+q \geqq 1$, then $H_{p,q}(\mathscr{F}) \leqq \log N$. If p and q are real numbers such that either $q \leqq 0$ and $p+q \leqq 1$ or $0 < q \leqq 1$ and $p+q \leqq 0$, then $H_{p,q}(\mathscr{F}) \geqq \log N$.*

Proof. We prove the theorem in the cases $q \geqq 1$ and $p+q \geqq 0$ or $0 \leqq q < 1$ and $p+q \geqq 1$. The other cases can be proved by a similar method with the obvious changes. First suppose $q \geqq 1$ and $p+q \geqq 0$. By lemma 1, equation (4), we have $H_{(-q,q)}(\mathscr{F}) = \frac{1}{q} \log N - \frac{1-q}{q} H_q(\mathscr{F})$, and by lemma 2 and the fact that $1-q \leqq 0$, we have $H_{(-q,q)}(\mathscr{F}) \leqq \frac{1}{q} \log N + \frac{q-1}{q} \log N = \log N$. Since $p \geqq -q$, by Lemma 3. we have $H_{p,q}(F) \leqq H_{(-q,q)}(\mathscr{F}) \leqq \log N$.

Next suppose $0 \leqq q < 1$ and $p+q \geqq 1$. Since $p \geqq 1-q$, Lemma 3., Lemma 2. and Lemma 1. give $H_{p,q}(\mathscr{F}) \leqq H_{1-q,q}(\mathscr{F}) = H_q(\mathscr{F}) \leqq \log N$ as desired.

The obvious question now is what can one say about $H_{p,q}(\mathscr{F})$ when p and q are not in the ranges covered by theorem 1. The answer is that $\log N$ may be neither an upper bound nor a lower bound for $H_{p,q}(\mathscr{F})$ for arbitrary $\mathscr{F}$ such that N=cardinality of $\Pi(\mathscr{F})$ if p and q are not in the ranges of Theorem 1. To be more specific, the complement of the sets of (p, q) considered in Theorem 1. consists of the union of the sets S_1, S_2, and S_3, where

$$S_1 = \{(p, q): q > 1,\ p+q < 0\}$$

$$S_2 = \{(p, q): 0 < q < 1,\ 0 < p+q < 1\}$$

$$S_3 = \{(p, q): q < 0,\ p+q > 1\}.$$

It is possible to find a point (p_0, q_0) in each of these three sets such that for different σ-subfields $\mathscr{F}_1$ and $\mathscr{F}_2$ with cardinality of $\Pi(\mathscr{F}_1)$=cardinality of $\Pi(\mathscr{F}_2)=N$, and such that $H_{p_0,q_0}(\mathscr{F}_1) < \log N$ and $H_{p_0,q_0}(\mathscr{F}_2) > \log N$. Consider the "points" $(-9/2, 3) \in S_1$, $(1/6, 1/3) \in S_2$, and $(5/2, -1/2) \in S_3$. With each of these points there exist σ-subfields $\mathscr{F}_1$ and $\mathscr{F}_2$ with cardinality $\Pi(\mathscr{F}_1), \Pi(\mathscr{F}_2)=2$ that have the property $H_{p_0,q_0}(\mathscr{F}_1) < \log 2 < H_{p_0,q_0}(\mathscr{F}_2)$. Two such σ-subfields for $(5/2, -1/2) \in S_3$ are the σ-subfields $\mathscr{F}_1$ (with $P(F_1')=1/4$, $P(F_1'')=3/4$) and $\mathscr{F}_2$ (with $P(F_2')=1/25$, $P(F_2'')= =24/25$). These particular $\mathscr{F}_1$ and $\mathscr{F}_2$ do not necessarily give the desired results for the "points" in S_1 and S_2, but appropriate σ-fields can be found.

Because of this erratic behaviour of the (p, q)-entropy for some p and q not in the ranges indicated by Theorem 1, we shall henceforth exclude values of p and q in the sets S_1, S_2, and S_3 from our further consideration of (p, q)-entropy. From information theoretic principles, we note that a measure of entropy should be bounded above by the entropy of a finite σ-field in which each of the atoms has equal probability. For this reason, we also exclude values of p and q which lie in the ranges indicated by the second half of theorem 1 from our further consideration. Therefore we restrict our attention to the study of $H_{p,q}(\mathscr{F})$ for values of p and q such that either $q \geqq 1$ and $p+q \geqq 0$ or $0 \leqq q < 1$ and $p+q \geqq 1$. It is also easy to show by examining the definition that $H_{p,q}(\mathscr{F}) \geqq 0$. We are now in a position to state and prove

Theorem 2. *Let $\mathscr{F}$ be a finite σ-subfield of $\mathscr{A}$, such that $\Pi(\mathscr{F})$ has cardinality N, and $\mathscr{B}$ be any σ-subfield of $\mathscr{A}$. Then $0 \leqq H^{\mathscr{B}}_{p,q}(\mathscr{F}) \leqq \log N$, provided either $q \geqq 1$ and $p+q \geqq 0$ or $0 \leqq q < 1$ and $p+q \geqq 1$.*

Proof. For any $\omega \in \Omega$, almost everywhere we have, according to Theorem 1. that

$$0 \leqq -\frac{1}{p} \log \left[\frac{\sum\limits_{\mathscr{F}} [P^{\mathscr{B}}(F)(\omega)]^{p+q}}{\sum\limits_{\mathscr{F}} [P^{\mathscr{B}}(F)(\omega)]^q} \right] \leqq \log N \quad \text{if} \quad p \neq 0,$$

$$0 \leqq \frac{-\sum\limits_{\mathscr{F}} [P^{\mathscr{B}}(F)(\omega)]^q \log [P^{\mathscr{B}}(F)(\omega)]}{\sum\limits_{\mathscr{F}} [P^{\mathscr{B}}(F)(\omega)]^q} \leqq \log N \quad \text{if} \quad p = 0.$$

Thus the expected value of each of these functions must be bounded between 0 and $\log N$. But the expected value of the above functions is exactly $H^{\mathscr{B}}_{p,q}(\mathscr{F})$ and so $0 \leqq H^{\mathscr{B}}_{p,q}(\mathscr{F}) \leqq \log N$ if p and q are in the ranges indicated.

We next prove a theorem which gives us an interpretation of the statement that the $\mathscr{B}$-conditional (p, q)-entropy of $\mathscr{F}$ is 0.

Theorem 3. *Let p and q be real numbers such that either $q \geqq 1$ and $p+q \geqq 0$ or $0 \leqq q < 1$ and $p+q \geqq 1$. Then $H^{\mathscr{B}}_{p,q}(\mathscr{F})=0$ if and only if $\mathscr{F} \subset \mathscr{B}$ a.e.*

Proof. First suppose $\mathscr{F} \subset \mathscr{B}$. Then if $F \in \mathscr{F}$, $P^{\mathscr{B}}(F)=1_F$ a.e. and if m is any real number and ω is a point such that $P^{\mathscr{B}}(F)(\omega)=1_F(\omega)$ for all $F \in \Pi(\mathscr{F})$, then $\sum_{\mathscr{F}} [P^{\mathscr{B}}(F)(\omega)]^m = 1$. Thus given $\omega \in \Omega$, if $p \neq 0$,

$$-\frac{1}{p} \log \left[\frac{\sum_{\mathscr{F}} [P^{\mathscr{B}}(F)(\omega)]^{p+q}}{\sum_{\mathscr{F}} [P^{\mathscr{B}}(F)(\omega)]^{q}} \right] \overset{\text{a.e.}}{=} \frac{1}{p} \log \frac{1}{1} = 0$$

and hence $H^{\mathscr{B}}_{p,q}(\mathscr{F})=0$. If $p=0$, then

$$\frac{-\sum_{\mathscr{F}} [P^{\mathscr{B}}(F)(\omega)]^q \log [P^{\mathscr{B}}(F)(\omega)]}{\sum_{\mathscr{F}} [P^{\mathscr{B}}(F)(\omega)]^q} \overset{\text{a.e.}}{=} -\sum_{\mathscr{F}} [P^{\mathscr{B}}(F)(\omega)]^q \log [P^{\mathscr{B}}(F)(\omega)] \overset{\text{a.e.}}{=} 0$$

since $-0 \log 0$ is defined to be zero and $\log 1$ is zero. Hence $H^{\mathscr{B}}_{0,q}(\mathscr{F})=0$.

Next suppose $H^{\mathscr{B}}_{p,q}(\mathscr{F})=0$. If $p \neq 0$, it follows that

$$\frac{\sum_{\mathscr{F}} [P^{\mathscr{B}}(F)(\omega)]^{p+q}}{\sum_{\mathscr{F}} [P^{\mathscr{B}}(F)(\omega)]^{q}} = 1 \quad \text{a.e.} \tag{5}$$

Let ω be any point such that (5) holds and let $\Pi(\mathscr{F}) = \{F_1, F_2, \ldots, F_n\}$. For each $F_i \in \Pi(\mathscr{F})$, $0 \leqq P^{\mathscr{B}}(F_i)(\omega) \leqq 1$ a.e. and hence if $p<0$, $[P^{\mathscr{B}}(F_i)(\omega)]^{p+q} \geqq [P^{\mathscr{B}}(F_i)]^q$ a.e. and if $p>0$, $[P^{\mathscr{B}}(F_i)(\omega)]^{p+q} \leqq [P^{\mathscr{B}}(F_i)(\omega)]^q$ a.e. Further we have strict inequality unless $P^{\mathscr{B}}(F_i)(\omega)=0$ or 1. Thus since (5) holds we have equality of $[P^{\mathscr{B}}(F_i)(\omega)]^{p+q}$ and $[P^{\mathscr{B}}(F_i)(\omega)]^q$ for all i and hence we have for one j, $1 \leqq j \leqq n$, $P^{\mathscr{B}}(F_j)(\omega)=1$ and for $1 \leqq i \leqq n$, $i \neq j$ we have $P^{\mathscr{B}}(F_i)(\omega)=0$. Since ω was arbitrary, we have for every $i=1, 2, \ldots, n$, $P^{\mathscr{B}}(F_i)(\omega)=0$ or 1 a.e.

If $p=0$, it follows that $\sum_{\mathscr{F}} [P^{\mathscr{B}}(F)(\omega)]^q \log \dfrac{1}{P^{\mathscr{B}}(F)(\omega)} = 0$ a.e. Since each of these terms is non-negative, we must have for each i, $1 \leqq i \leqq n$, $[P^B(F_i)(\omega)]^q \log [P^B(F_i)(\omega)]=0$ a.e. which implies that either $P^B(F_i)(\omega)=0$ or 1 a.e.

Hence in either case we have that $P^B(F_i)=0$ or 1 a.e. for each i, $1 \leqq i \leqq n$, and since $P^{\mathscr{B}}(F_i)(\omega)$ is a $\mathscr{B}$-measurable function, we have that it is the indicator function of some set B in $\mathscr{B}$. Now $B = \{\omega : P^{\mathscr{B}}(F_i)(\omega)=1\} = \{\omega : P^{\mathscr{B}}(F_i)(\omega)>0\} \supset F_i$ by [4]. Further

$$P(B) = \int_B 1_B(\omega)\, dP = \int_B P^{\mathscr{B}}(F_i)(\omega)\, dP - \int_B 1_{F_i}(\omega)\, dP = P(B \cap F_i).$$

Now $P(B-F_i) = P(B - B \cap F_i) = 0$ and hence $B \subset F_i$. Therefore $B=F_i$ a.e. and hence $\mathscr{F} \subset \mathscr{B}$.

4. Further properties of conditional (p, q) entropy

One of the most useful and important properties of conditional Shannon entropy is the additivity property:

$$H_{0,1}^{\mathscr{B}}(\mathscr{F} \vee \mathscr{F}') = H_{0,1}^{\mathscr{B}}(\mathscr{F}) + H_{0,1}^{\mathscr{B} \vee \mathscr{F}}(\mathscr{F}')$$

where $\mathscr{F}$ and $\mathscr{F}'$ are finite σ-subfields of $\mathscr{A}$ and $\mathscr{B}$ is an arbitrary σ-subfield of $\mathscr{A}$. It is this additivity property that enables one to prove the useful Sinai theorem (see [10]) which facilitates the actual computation of the entropy of a measure-preserving, invertible transformation τ. However, the additivity property is one of the properties that characterizes the Shannon entropy and hence one expects that for arbitrary p and q, the conditional (p, q) entropy of $\mathscr{F}$ given $\mathscr{B}$ will not satisfy this property and indeed this is generally true. One can define an invariant of a transformation τ in the same manner as the Kolmogoroff—Sinai invariant ([1], [6], [10]) using conditional (p, q) entropy, but the absence of this additivity property prevents us from proving an analogous Sinai theorem and hence the computation of this invariant will be difficult.

Even with the absence of this additivity property, for certain p and q we can prove a monotone property that is an obvious corollary of the additivity property when considering conditional Shannon entropy.

Theorem 4. *Let $\mathscr{F}$ and $\mathscr{F}'$ be finite σ-fields with $\mathscr{F} \subset \mathscr{F}'$ and let $\mathscr{B}$ be an arbitrary σ-field. Then if p and q are real numbers such that either $0 < q \leqq 1$ and $p+q \geqq 1$ or $q \geqq 1$ and $0 < p+q \leqq 1$, then*

$$H_{p,q}^{\mathscr{B}}(\mathscr{F}) \leqq H_{p,q}^{\mathscr{B}}(\mathscr{F}').$$

Proof. First we note that if $f(t)$ is convex (concave) in the interval $[0, 1]$ and $f(0)=0$, then for arbitrary x, y, and $x+y \in [0, 1]$ we have the inequality (see [5], p. 132),

$$f(x+y) \geqq (\leqq) f(x) + f(y). \tag{6}$$

We also note that since $\mathscr{F} \subset \mathscr{F}'$, then for every set $A_i \in \Pi(\mathscr{F})$, $i = 1, 2, \dots, n$, there exist finitely many sets $C_{ij} \in \Pi(\mathscr{F}')$ (j runs through a finite index set) so that $A_i = \bigcup_j C_{ij}$, from which we have $P(A_i) = \sum_j P(C_{ij})$. Moreover, for any σ-subfield $\mathscr{B}$ of $\mathscr{A}$, we have $P^{\mathscr{B}}(A_i) = P^{\mathscr{B}}(\bigcup_j C_{ij}) = \sum_j P^{\mathscr{B}}(C_{ij})$ a.e. We select representatives of $P^{\mathscr{B}}(A_i)$ for all $1 \leqq i \leqq n$ and representatives of $P^{\mathscr{B}}(C_{ij})$ for all $1 \leqq i \leqq n$ and all j in a finite index set. Let ω be a point in the intersection of the sets of definition of all $P^{\mathscr{B}}(A_i)$ and $P^{\mathscr{B}}(C_{ij})$. Define x_i and y_{ij} associated with this point ω as follows: $x_i = P^{\mathscr{B}}(A_i)(\omega)$, $y_{ij} = P^{\mathscr{B}}(C_{ij})(\omega)$. We observe that $x_i \geqq 0$ and $\sum_{i=1}^{n} x_i = 1$ and also $y_{ij} \geqq 0$ and $\sum_{i=1}^{n} \sum_j y_{ij} = 1$. Further we have $x_i = \sum_j y_{ij}$ for almost all $\omega \in \Omega$.

First assume $p=0$ in either of the cases. Then $q=1$ in either case and we are considering conditional Shannon entropy which is known to satisfy this monotone property.

Assume now that $p \neq 0$ and let $F(t) = -\log(t^{1/p})$. First assume $0 < q \leqq 1$ and $p+q \geqq 1$. We note that since $p \neq 0$, we must have $p > 0$ and hence $F(t)$ is a decreasing function of t. Since t^{p+q} is a convex function in the interval [0, 1] and t^q is a concave function in the interval [0, 1], using (6) we have that the following inequality is true:

$$H^{\mathscr{B}}_{p,q}(\mathscr{F}) = E\left[-\log\left\{\frac{\sum_{i=1}^{n}[P^{\mathscr{B}}(A_i)]^{p+q}}{\sum_{i=1}^{n}[P^{\mathscr{B}}(A_i)]^{q}}\right\}^{\frac{1}{p}}\right] = E\left[F\left\{\frac{\sum_{i=1}^{n}\left(\sum_j y_{ij}\right)^{p+q}}{\sum_{i=1}^{n}\left(\sum_j y_{ij}\right)^{q}}\right\}\right] \leqq$$

$$\leqq E\left[F\left\{\frac{\sum_{i=1}^{n}\sum_j y_{ij}^{p+q}}{\sum_{i=1}^{n}\sum_j y_{ij}^{q}}\right\}\right] = E\left[-\log\left\{\frac{\sum_{i=1}^{n}\sum_j [P^{\mathscr{B}}(C_{ij})]^{p+q}}{\sum_{i=1}^{n}\sum_j [P^{\mathscr{B}}(C_{ij})]^{q}}\right\}^{\frac{1}{p}}\right] = H_{p,q}(F').$$

If we now assume $q \geqq 1$ and $0 < p+q \leqq 1$, we note $p<0$ and hence $F(t)$ is an increasing function of t. In this case t^{p+q} is a concave function and t^q is a convex function in the interval [0, 1]. Again, using (6) in a similar fashion we have the desired inequality.

In particular, one should note that among the values of p and q for which the inequality is valid are those p and q which represent conditional Shannon entropy and conditional Rényi entropy of order $\alpha > 0$.

We now seek values of p and q for which the convexity property of $H^{\mathscr{B}}_{p,q}(\mathscr{F})$ holds. That is, if $\mathscr{B}$ and $\mathscr{B}'$ are σ-fields such that $\mathscr{B} \subset \mathscr{B}'$, then do there exist values of p and q such that $H^{\mathscr{B}}_{p,q}(\mathscr{F}) \geqq H^{\mathscr{B}'}_{p,q}(\mathscr{F})$ for ever yfinite σ-field $\mathscr{F}$? This problem is very closely related to the problem of finding values of p and q for which the function

$$F_{p,q}(x_1, \ldots, x_n) = -\frac{1}{p}\log\left[\frac{\sum_{i=1}^{n} x_i^{p+q}}{\sum_{i=1}^{n} x_i^{q}}\right],$$

under the condition $\sum x_i = 1$, is a concave function of the n variables. We state the following

Definition: Let $F(x_1, \ldots, x_n)$ be a real-valued function of n variables defined on a convex set G in n-space. Then F is a *concave function* of the n-variables if and only if

$$(7) \qquad F\big(a(x_1, \ldots, x_n) + b(y_1, \ldots, y_n)\big) \geqq aF(x_1, \ldots, x_n) + bF(y_1, \ldots, y_n)$$

for all non-negative real numbers a and b such that $a+b=1$.

Restricting our attention to conditional Rényi entropy of order α, we can prove that for $0<\alpha<1$, the convexity property does hold. Under these restrictions we can prove the following

Lemma. *Let* $F_\alpha(x_1, \ldots, x_n) = \dfrac{1}{1-\alpha}\log[x_1^\alpha + \cdots + x_n^\alpha]$ *where* $0<\alpha<1$, $x_i > 0$, $\sum_{i=1}^{n} x_i = 1$. *Then* $F_\alpha(x_1, \ldots, x_n)$ *is a concave function.*

PROOF. First note that $x_1^\alpha + \cdots + x_n^\alpha$ is a concave function since each x_1^α is a concave function for $0<\alpha<1$ and $0<x_i<1$. The added restriction that $\sum_{i=1}^{n} x_i = 1$ is a linear restriction of a convex set and this will not affect the concavity of the function. It is easily shown that an increasing concave function of a concave function is a concave function. Now $h(y) = \frac{1}{1-\alpha} \log y$ is an increasing concave function since $0<\alpha<1$ and hence $F_\alpha(x_1, \ldots, x_n) = \frac{1}{1-\alpha} \log [x_1^\alpha + \cdots + x_n^\alpha]$ is a concave function.

We note that the restrictions must be made to Rényi entropy of order α, $0<\alpha<1$, since the functions used to obtain other (p, q)-entropies are not in general concave. We can now state and prove the following convexity theorem.

Theorem 5. *Let $\mathscr{B}$ and $\mathscr{B}'$ be a σ-subfields of $\mathscr{A}$ such that $\mathscr{B} \subset \mathscr{B}'$, $\mathscr{F}$ be a finite σ-subfield of $\mathscr{A}$, and $0<\alpha<1$. Then $H_\alpha^{\mathscr{B}}(\mathscr{F}) \geqq H_\alpha^{\mathscr{B}'}(\mathscr{F})$.*

PROOF. Since $\mathscr{B} \subset \mathscr{B}'$, we have that for any summable function f, $E^{\mathscr{B}}(f) = E^{\mathscr{B}}(E^{\mathscr{B}'}(f))$. Hence if $F \in \Pi(\mathscr{F})$, we have that $P^{\mathscr{B}}(F) = E^{\mathscr{B}}(P^{\mathscr{B}'}(F))$ and from this we obtain $\sum_{\mathscr{F}} [P^{\mathscr{B}}(F)]^\alpha = \sum_{\mathscr{F}} [E^{\mathscr{B}}(P^{\mathscr{B}'}(F))]^\alpha$. Further

$$\frac{1}{1-\alpha} \log \sum_{\mathscr{F}} [P^{\mathscr{B}}(F)]^\alpha = \frac{1}{1-\alpha} \log \sum_{\mathscr{F}} [E^{\mathscr{B}}(P^{\mathscr{B}'}(F))]^\alpha,$$

and finally

$$H_\alpha^{\mathscr{B}}(\mathscr{F}) = E\left(\frac{1}{1-\alpha} \log \sum_{\mathscr{F}} [E^{\mathscr{B}}(P^{\mathscr{B}'}(F))]^\alpha\right). \tag{8}$$

Since, by the lemma, $\frac{1}{1-\alpha} \log \sum x_i^\alpha$, under the conditions that $x_i > 0$ and $\sum x_i = 1$, and $0<\alpha<1$, is a concave function, we have by Jensen's inequality that

$$\frac{1}{1-\alpha} \log \sum_{\mathscr{F}} [E^{\mathscr{B}}(P^{\mathscr{B}'}(F))]^\alpha \geqq E^{\mathscr{B}}\left[\frac{1}{1-\alpha} \log \sum_{\mathscr{F}} [P^{\mathscr{B}'}(F)]^\alpha\right].$$

Then from (8) we have

$$H_\alpha^{\mathscr{B}}(\mathscr{F}) \geqq E\left[E^{\mathscr{B}}\left\{\frac{1}{1-\alpha} \log \sum_{\mathscr{F}} [P^{\mathscr{B}'}(F)]^\alpha\right\}\right] = E\left(\frac{1}{1-\alpha} \log \sum_{\mathscr{F}} [P^{\mathscr{B}'}(F)]^\alpha\right) = H_\alpha^{\mathscr{B}'}(\mathscr{F}),$$

and hence $H_\alpha^{\mathscr{B}}(F) \geqq H_\alpha^{\mathscr{B}'}(\mathscr{F})$.

We now state a special form of the martingale convergence theorem (see [7]).

L_2-Martingale Convergence Theorem: Let $\{\mathscr{B}_n\}$ be an increasing (decreasing) sequence of σ-subfields of $\mathscr{A}$ and let $\mathscr{B} = \lim_{n\to\infty} \mathscr{B}_n$. Then for every $f \in L_2(\Omega, \mathscr{A}, P)$, $E^{\mathscr{B}_n}(f) \xrightarrow{L_2} E^{\mathscr{B}}(f)$.

We now use the above convergence theorem to prove a continuity theorem for conditional (p, q) entropy if p and q are as in the first part of Theorem 1.

Theorem 6. (Continuity Theorem.) *Suppose $\{\mathcal{B}_n\}$ is an increasing (decreasing) sequence of σ-subfields of $\mathcal{A}$ and $\lim_{n\to\infty} \mathcal{B}_n = \mathcal{B}$ and that either $q \geqq 1$ and $p+q \geqq 0$ or $0 \leqq q < 1$ and $p+q \geqq 1$. If $\mathcal{F}$ is finite σ-field, then $H^{\mathcal{B}_n}_{p,q}(\mathcal{F}) \to H^{\mathcal{B}}_{p,q}(\mathcal{F})$. For conditional Rényi entropy of order α, $0<\alpha<1$, we have $H^{\mathcal{B}_n}_\alpha(\mathcal{F}) \uparrow (\downarrow) H^{\mathcal{B}}_\alpha(\mathcal{F})$.*

PROOF. Suppose p and q are real numbers in either of the above ranges. We know by the martingale convergence theorem that for $F \in \Pi(\mathcal{F})$ $P^{\mathcal{B}_n}(F) \xrightarrow{L_2} P^{\mathcal{B}}(F)$ and hence $P^{\mathcal{B}_n}(F) \xrightarrow{P} P^{\mathcal{B}}(F)$.

We first assume $p \neq 0$. Since t^m is uniformly continuous on $[0, 1]$ for $m \geqq 0$, then

$$\sum_{\mathcal{F}} [P^{\mathcal{B}_n}(F)]^{p+q} \xrightarrow{P} \sum_{\mathcal{F}} [P^{\mathcal{B}}(F)]^{p+q} \quad \text{and} \quad \sum_{\mathcal{F}} [P^{\mathcal{B}_n}(F)]^{q} \xrightarrow{P} \sum_{\mathcal{F}} [P^{\mathcal{B}}(F)]^{q}.$$

Now $\log t$ is uniformly continuous on $[a, \infty)$, $a>0$ and so

$$-\frac{1}{p}\log\Big\{\sum_{\mathcal{F}} [P^{\mathcal{B}_n}(F)]^{p+q}\Big\} \xrightarrow{P} -\frac{1}{p}\log\Big\{\sum_{\mathcal{F}} [P^{\mathcal{B}}(F)]^{p+q}\Big\}$$

and similarly when the power $(p+q)$ is replaced by q. From this we obtain

$$-\frac{1}{p}\log\Big\{\sum_{\mathcal{F}} [P^{\mathcal{B}_n}(F)]^{p+q}\Big\} + \frac{1}{p}\log\Big\{\sum_{\mathcal{F}} [P^{\mathcal{B}_n}(F)]^{q}\Big\} \xrightarrow{P} -\frac{1}{p}\log\Big\{\sum_{\mathcal{F}} [P^{\mathcal{B}}(F)]^{p+q}\Big\} +$$

$$+\frac{1}{p}\log\Big\{\sum_{\mathcal{F}} [P^{\mathcal{B}}(F)]^{q}\Big\}.$$

That is,

$$-\frac{1}{p}\log\left\{\frac{\sum_{\mathcal{F}} [P^{\mathcal{B}_n}(F)]^{p+q}}{\sum_{\mathcal{F}} [P^{\mathcal{B}_n}(F)]^{q}}\right\} \xrightarrow{P} -\frac{1}{p}\log\left\{\frac{\sum_{\mathcal{F}} [P^{\mathcal{B}}(F)]^{p+q}}{\sum_{\mathcal{F}} [P^{\mathcal{B}}(F)]^{q}}\right\}.$$

Since, by theorem 1, all of the functions are bounded by $\log N$, where N is the cardinality of $\Pi(\mathcal{F})$, we have L_1-convergence. Thus

$$E\left[-\frac{1}{p}\log\left\{\frac{\sum_{\mathcal{F}} [P^{\mathcal{B}_n}(F)]^{p+q}}{\sum_{\mathcal{F}} [P^{\mathcal{B}_n}(F)]^{q}}\right\}\right] \to E\left[-\frac{1}{p}\log\left\{\frac{\sum_{\mathcal{F}} [P^{\mathcal{B}}(F)]^{p+q}}{\sum_{\mathcal{F}} [P^{\mathcal{B}}(F)]^{q}}\right\}\right]$$

and hence $\lim_{n\to\infty} H^{\mathcal{B}_n}_{p,q}(F) = H^{\mathcal{B}}_{p,q}(\mathcal{F})$ if $p \neq 0$. If $p=0$, the proof is analogous, if we note that $-t\log t$ is uniformly continuous on $[0, 1]$ provided we define $0 \log 0$ to be 0.

If we consider conditional Rényi entropy of order α, $0<\alpha<1$, then by T heorem 5. we have $H^{\mathcal{B}_n}_\alpha(F) \uparrow (\downarrow) H^{\mathcal{B}}_\alpha(\mathcal{F})$.

5. Open questions

Several open questions can now be posed. First, as mentioned in the beginn ing of section 4, one may define a new entropy, $\hat{h}'_{p,q}(\tau)$, of a measure-preservin- transformation in a completely analogous way as when discussing conditiolng Shannon entropy. This $\hat{h}'_{p,q}(\tau)$ is a conjugacy invariant. An interesting questioa would be to determine if $\hat{h}_{p,q}(\tau)$ as formulated by Daróczy [3] is the same as $\hat{h}'_{p,q}(\tau)$. That is, is the entropy of τ as evaluated using two different methods the same, or are $\hat{h}'_{p,q}(\tau)$ and $\hat{h}_{p,q}(\tau)$ two different measures of τ. Another question is whether there is an axiomatic characterization of the (p, q) entropy.

References

[1] P. Billingsley, Ergodic Theory and Information, *New York*, 1965.
[2] Z. Daróczy, Einige Ungleichungen über die mit Gewichtsfunktionen gebildeten Mittlewerte, *Monatsh. Math.*, **68**, (1964), 102—112.
[3] Z. Daróczy, Über eine Verallgemeinerung des Entropiebegriffs der maßtreuen Abbildungen, *Publ. Math. (Debrecen)*, **12**, (1965), 117—125.
[4] A. Hanen and J. Neveu, Atomes conditionnels d'un espace de probabilité, *Acta Math. Acad. Sci. Hungar.*, **17**, 1966, 443—449.
[5] E. Hille, Functional Analysis and Semi-Groups, *A. M. S. Colloquium Publ., New York*, 1948.
[6] A. N. Kolmogoroff, A new invariant for transitive dynamical systems, *Dokl. Akad. Nauk. SSSR*, **119**, (1958), 861—864.
[7] M. Loeve, Probability Theory, 3rd edition, *Princeton*, 1963.
[8] A. Rényi, On measures of entropy and information, *Proc. Fourth Berkeley Sympos. Statist. and Probility*, 1960, 547—561.
[9] C. Shannon, A mathematical theory of communication, *Bell System Tech. J.* **27**, (1948), 379—423 and 623—656.
[10] Ya. G. Sinai, On the notion of the entropy of a dynamical system, *Dokl. Akad. Nauk. SSSR* **124**, (1959 a), 768—771.

(Received November 23, 1970.)