About the Mutual (Conditional) Information

Renato Renner and Ueli Maurer[1]
Institute of Theoretical Computer Science, ETH Zürich
CH-8092 Zürich, Switzerland
{renner,maurer}@inf.ethz.ch

Abstract — **We give a necessary, sufficient, and easily verifiable criterion for the conditional probability distribution $P_{Z|XY}$ (where X, Y and Z are arbitrary random variables), such that $I(X;Y) \geq I(X;Y|Z)$ holds for any distribution P_{XY}. Furthermore, the result is generalized to the case where Z is specified by a conditional probability distribution depending on more than two random variables.**

I. Introduction

The mutual information $I(X;Y)$ between two random variables X and Y is one of the basic measures in information theory [1]. It can be interpreted as the amount of information that X gives on Y (or vice versa). In general, additional information, i.e., conditioning on an additional random variable Z, can either increase or decrease this mutual information. Without loss of generality, Z can be seen as the output of a channel C with input (X,Y), being fully specified by the conditional probability distribution $P_{Z|XY}$. It is thus a natural question whether for a fixed $P_{Z|XY}$ (i.e., a fixed channel C with input (X,Y) and output Z) conditioning on Z can possibly (i.e., for some distribution P_{XY}) increase the mutual information between X and Y.

In the following, we give an answer to this question as well as to a generalization thereof. Our result has different applications, e.g., in the context of information theoretical secret key agreement.

II. Definitions and Results

Let the function $p : (z,x,y) \mapsto p(z|x,y)$ be a conditional probability distribution (i.e., $p(\cdot|x,y)$ is a probability distribution for each pair (x,y)).

Definition 1. The conditional probability distribution p is called *correlation free* if for all random variables X, Y and Z with $P_{Z|XY} = p$ (and arbitrary distribution P_{XY})

$$I(X;Y) \geq I(X;Y|Z).$$

In view of this definition, our main goal can be formulated as finding a simple criterion for p to be correlation free.

Definition 2. The conditional probability distribution p is called *multiplicative* if it is the product of two functions r and s depending only on (z,x) and (z,y), respectively, i.e.,

$$p(z|x,y) = r(z,x) \cdot s(z,y)$$

for all x, y and z.

Our main theorem states that the multiplicative property is exactly the criterion needed.

Theorem 3. *The conditional probability distribution p is correlation free if and only if it is multiplicative.*

In fact, it is easy to decide whether a given conditional probability distribution p is multiplicative. The following lemma shows that one only has to check the conditional independence of a certain pair of random variables.

Lemma 4. *The conditional probability distribution p is multiplicative if and only if $I(X;Y|Z) = 0$ for two independent and uniformly distributed random variables X, Y (i.e., P_{XY} is constant), and Z with $P_{Z|XY} = p$.*

Combining Theorem 3 and Lemma 4, our result can be formulated as follows: Let C be a fixed channel, taking as input a pair of random variables. If and only if the mutual information of a uniformly distributed input pair $(\bar{X},\bar{Y})$ does not increase when conditioning on the channel output $\bar{Z}$ (i.e., it equals 0), then the mutual information of any arbitrary input pair (X,Y) does not increase when conditioning on the output Z:

$$0 = I(\bar{X};\bar{Y}) \geq I(\bar{X};\bar{Y}|\bar{Z}) \iff \forall\, P_{XY} : I(X;Y) \geq I(X;Y|Z).$$

III. Generalizations

The conditional probability distribution p considered in the previous section corresponds to a channel taking two random variables as input. However, our result can be extended to conditional probability distributions of the form $p : (z,x_1,\ldots,x_n) \mapsto p(z|x_1,\ldots,x_n)$ for $n \in \mathbf{N}$. The generalization of Definition 1 and Definition 2 is straightforward.

Definition 5. The conditional probability distribution p is called *correlation free* if

$$\sum_{i=1}^{n} I(X_i;Z) \geq I(X_1 \cdots X_n;Z)$$

for any choice of random variables $X_1,\ldots,X_n$ and Z with $P_{Z|X_1\cdots X_n} = p$.

Definition 6. The conditional probability distribution p is called *multiplicative* if it can be written as a product

$$p(z|x_1,\ldots,x_n) = \prod_{i=1}^{n} r_i(z,x_i)$$

for appropriate functions $r_1,\ldots,r_n$.

It turns out that Theorem 3 still holds for these extended definitions.

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pt. I, pp. 379–423, 1948; pt. II, pp. 623–656, 1948.

[1]This work was supported by the Swiss National Science Foundation (SNF).