24. *On 3-dimensional Interaction Information*

By Seiji TAKANO
Department of Mathematics, Tokyo Institute of Technology

(Comm. by Kôsaku YOSIDA, M. J. A., Feb. 12, 1974)

1. Introduction. A notion of interaction information was introduced by McGill [2] and it was used in multivariate information analysis. Lately, it is shown that it plays an important role especially in the variables connected with Markov dependence [5]. However, it seems that the essential problem suggested by McGill: *under what conditions does* 3-*dimensional interaction information take positive, zero and negative values*? is not yet solved. The purpose of the present note is to show that it deeply relates to the trace of the product matrix of the three transition matrices, each of which represents the relationship between the variables. It is also shown that if the trace is equal to 1, the variables having zero interaction information constitute some intermediate dependence lying between Markovian dependence and independence. In addition, we show some processes realizing positive and negative interactions.

2. Definitions and basic properties. Consider the random variables X, Y and Z, taking only finite number of states $\{a_1, a_2, \cdots, a_L\}$, $\{b_1, b_2, \cdots, b_M\}$ and $\{c_1, c_2, \cdots, c_N\}$, respectively, where L, M and N are positive integers. $P(XYZ)$ denotes the function taking the joint probability value $p(ijk)=P(X=a_i, Y=b_j, Z=c_k)$ when $X=a_i, Y=b_j$ and $Z=c_k$. Then, 3-dimensional interaction information of the variables X, Y and Z is defined by

$$J=J(XYZ)=E\{\log [P(XYZ)P(X)P(Y)P(Z) \\ /(P(XY)P(YZ)P(ZX))]\} \tag{2.1}$$

or, equivalently, using the conditional probabilities,

$$J=E\{\log [P(XYZ)/P(Y\,|\,X)P(Z\,|\,Y)P(X\,|\,Z))]\} \tag{2.2}$$

where E means the expectation over all $P(XYZ)$. This may be more instructive if we rewrite it as

$$\begin{aligned} J&=E\{\log [P(XZ\,|\,Y)/(P(X\,|\,Y)P(Z\,|\,Y))]\} \\ &\quad -E\{\log [P(XZ)/(P(X)P(Z))]\} \\ &=EI(X, Z\,|\,Y)-I(X, Z). \end{aligned} \tag{2.3}$$

The first term of (2.3) is the conditional information between X and Z, given Y, and the second, the mutual information between X and Z. Thus, J may be considered as a measure which suggests the effect of Y with respect to X and Z. Since J is symmetric with respect to each variable (cf. [5]), these representations do not lose the generality.

Now, consider the form (2.2) and rewrite it as

$$J=E\{\log [P(XYZ)/((1/C)P(Y\mid X)P(Z\mid Y)P(X\mid Z))]\}-\log C \tag{2.4}$$

where C is given by

$$C=\sum P(Y\mid X)P(Z\mid Y)P(X\mid Z). \tag{2.5}$$

Here the summation is taken over all values of X, Y and Z, and we assume that even if one of the conditional probabilities reduces to zero, it affects nothing to J. Clearly the first term of (2.4) is a discrimination between the two 3-dimensional distributions (cf. [3]), so that we have

Lemma 1. *$J\geq -\log C$. The equality holds only when $C=1$ and then*

$$P(XYZ)=P(Y\mid X)P(Z\mid Y)P(X\mid Z) \tag{2.6}$$

holds good.

Proof. The first assertion is almost clear. We shall assume that the equality holds. Clearly, it is required that

$$P(XYZ)=(1/C)P(Y\mid X)P(Z\mid Y)P(X\mid Z).$$

From this we have

$$P(XY)=(1/C)P(Y\mid X)\sum_Z P(Z\mid Y)P(X\mid Z), \qquad \text{i.e.,}$$

$$P(X)=(1/C)\sum_Z P(Z\mid Y)P(X\mid Z),$$

provided $P(Y\mid X)\neq 0$. Since $P(Y\mid X)=0$ implies only the trivial case, we shall have $1=(1/C)\sum_X \sum_Z P(Z\mid Y)P(X\mid Z)$ by summation. This proves that $C=1$ and (2.6) holds. Q.E.D.

3. A condition for interaction information zero. Lemma 1 suggests the classification of situations by introducing the trace of

$$\boldsymbol{PQR}=\text{the product of the transition matrices;}$$

$$\begin{aligned}&\boldsymbol{P}=(p_{ij}),\ \boldsymbol{Q}=(q_{jk}) \text{ and } \boldsymbol{R}=(r_{ki}), \text{ where } p_{ij}=P(Y=b_j\mid X=a_i),\\ &q_{jk}=P(Z=c_k\mid Y=b_j) \text{ and } r_{ki}=P(X=a_i\mid Z=c_k),\\ &i=1,\cdots,L;\ j=1,\cdots,M;\ k=1,\cdots,N.\end{aligned} \tag{3.1}$$

Theorem 1. *A 3-dimensional interaction information becomes zero if and only if (2.6) holds, provided the trace of $\boldsymbol{PQR}$ is equal to 1. Moreover, such a distribution is constructed from any transition matrices $\boldsymbol{P}, \boldsymbol{Q}$ and $\boldsymbol{R}$ satisfying*

$$\tilde{\boldsymbol{P}}\tilde{\boldsymbol{Q}}=0,\ \tilde{\boldsymbol{Q}}\tilde{\boldsymbol{R}}=0 \quad \textit{and} \quad \tilde{\boldsymbol{R}}\tilde{\boldsymbol{P}}=0, \tag{3.2}$$

where $\tilde{\boldsymbol{P}}=(\tilde{p}_{ij})$, $\tilde{\boldsymbol{Q}}=(\tilde{q}_{jk})$ and $\tilde{\boldsymbol{R}}=(\tilde{r}_{ki})$ are defined by

$$\begin{aligned}&\tilde{p}_{ij}=p_{ij}-p_{1j},\ \tilde{q}_{jk}=q_{jk}-q_{1k} \quad \textit{and} \quad \tilde{r}_{ki}=r_{ki}-r_{1i}\\ &\qquad (i=2,\cdots,L;\ j=2,\cdots,M;\ k=2,\cdots,N)\end{aligned} \tag{3.3}$$

Proof. Let $\operatorname{tr}(\boldsymbol{PQR})=C=1$. Then, by Lemma 1, J is zero if and only if (2.6) holds. Now assume that (2.6) holds. Then

$$P(XY)=P(Y\mid X)\sum_Z P(Z\mid Y)P(X\mid Z).$$

Thus, excepting the trivial case of $P(Y\mid X)=0$, we have

$$P(X)=\sum_Z P(Z\mid Y)P(X\mid Z). \tag{3.4.1}$$

Similarly, we obtain

$$P(Y)=\sum_{X} P(X \mid Z)P(Y \mid X) \tag{3.4.2}$$

$$P(Z)=\sum_{Y} P(Y \mid X)P(Z \mid Y) \tag{3.4.3}$$

provided $P(XYZ)\neq 0$. We shall then show that (3.4.1)–(3.4.3) imply (3.2). Let (3.4.1) be satisfied. Since the right hand side of (3.4.1) again forms a stochastic matrix and does not depend on Y, we must have $\boldsymbol{QR}=\boldsymbol{P}_1^*$, where $\boldsymbol{P}_1^*$ is defined by the distribution of X such that

$$\boldsymbol{P}_1^* = \begin{bmatrix} p_1, p_2, \cdots, p_L \\ p_1, p_2, \cdots, p_L \\ \vdots \quad \vdots \qquad \vdots \\ p_1, p_2, \cdots, p_L \end{bmatrix}$$

and $p_i=P(X=a_i)$, $i=1,\cdots,L$. This implies that for any j and i, $\sum_k q_{jk}r_{ki}=\sum_k q_{1k}r_{ki}$, or, $\sum_k (q_{jk}-q_{1k})(r_{ki}-r_{1i})=0$, $i=2,\cdots,L$; $j=2,\cdots,M$. Thus, we have $\tilde{\boldsymbol{Q}}\tilde{\boldsymbol{R}}=0$. Analogously, we see that the other two of (3.2) follow from (3.4.2) and (3.4.3). On the other hand, if we construct a 3-dimensional distribution from (2.6), using these $\boldsymbol{P}, \boldsymbol{Q}$ and $\boldsymbol{R}$ which satisfy (3.2), then clearly it satisfies our requirement. Q.E.D.

Corollary. *Any transition matrix* $\boldsymbol{P}$ *satisfying* $\tilde{\boldsymbol{P}}^2=0$ *defines the distribution for which* $J=0$.

Theorem 1 implies the previous results in [5] for the special case. In fact, if the variables form a Markov chain, then the interaction zero occurs only when the first and the last variables are mutually independent. Further, if the chain is homogeneous, it will generate a 1-*dependent Markov chain*, i.e., the corresponding variables $(X_i, X_{i+1}, \cdots, X_j)$ and $(X_s, X_{s+1}, \cdots, X_t)$ are mutually independent whenever $s-j>1$ $(1\leq i\leq j<s\leq t;\ i, j, s, t$: integers).

4. Positive and negative interactions. We show some conditions for positive and negative interactions. Firstly, Lemma 1 suggests the following theorem.

Theorem 2. *A* 3-*dimensional interaction information takes positive values if* $\mathrm{tr}\,(\boldsymbol{PQR})\leq 1$, *provided that we do not take into account of the system derived from* (2.6) *when* $\mathrm{tr}\,(\boldsymbol{PQR})=1$.

On the other hand, for the remaining case, we have

Theorem 3. *For the variables satisfying* $\mathrm{tr}\,(\boldsymbol{PQR})>1$, *the corresponding* 3-*dimensional interaction information can take positive, zero and negative values.*

Proof. It suffices to give an example of the distribution satisfying the condition of the theorem such that the corresponding J takes positive, zero and negative values. For this purpose consider the next distribution:

i	j	k	$p(ijk)$	i	j	k	$p(ijk)$
1	1	1	$1/18+\varepsilon$	2	1	1	$1/9-\varepsilon$
1	1	2	$1/9-\varepsilon$	2	1	2	$2/9+\varepsilon$
1	2	1	$2/9+\delta$	2	2	1	$1/9-\delta$
1	2	2	$1/9-\delta$	2	2	2	$1/18+\delta$

where we assume $|\varepsilon|, |\delta| \leq 1/18$. From this we have

$$\mathrm{tr}\,(\boldsymbol{PQR}) = 82/81 + 4(\varepsilon+\delta)/9.$$

Let $\varepsilon+\delta=0$ and $\varepsilon=2x/9$ ($|x|\leq 1/4$). Then, we obtain easily

$$J = (5/9)\log(9/10) + (4/9)\log(9/8) + H(x),$$

where

$$\begin{aligned} H(x) = {} & (1/18)[(1+4x)\log(1+4x)+(1-4x)\log(1-4x)] \\ & +(2/9)[(1+2x)\log(1+2x)+(1-2x)\log(1-2x)] \\ & +(2/9)[(1+x)\log(1+x)+(1-x)\log(1-x)]. \end{aligned}$$

Since $H(x)$ is a continuous, symmetric and convex function of x and J takes its minimum value $J_m = (5/9)\log(9/10) + (4/9)\log(9/8)$ when the variables form a Markov chain, J takes positive, zero and negative values. In fact, J_m is negative and $|J_m| < H(1/4)$. Q.E.D.

Finally, consider a process which realizes positive or negative interaction. Generally, we can write

$$p(ijk) = p_i p_{ij} q_{jk} + \varepsilon_{ijk}, \qquad |\varepsilon_{ijk}| \leq p_i p_{ij} q_{jk} \tag{4.1}$$

with

$$\sum_i \varepsilon_{ijk} = 0 \quad \text{and} \quad \sum_k \varepsilon_{ijk} = 0. \tag{4.2}$$

Lemma 2. *Assume that*

$$\sum_j \varepsilon_{ijk} = 0. \tag{4.3}$$

Then, for any distribution of the form (4.1)–(4.3), *the trace of* $\boldsymbol{PQR}$ *is not less than* 1. *It equals to* 1 *only when* X *and* Z *are mutually independent.*

The proof follows from the direct computation and Schwartz inequality.

Now let us introduce random transformations such that

$$p_{ij}^{(n+1)} = \sum_u p_{iu}^{(n)} p_{uj} \quad \text{and} \quad q_{jk}^{(n+1)} = \sum_v q_{jv}^{(n)} q_{vk}. \tag{4.4}$$

This means that some number of channels are inserted between X and Y, and Y and Z, respectively. Then the distribution becomes

$$p_n(ijk) = p_i p_{ij}^{(n)} q_{jk}^{(n)} + \varepsilon_{ijk}^{(n)}, \qquad |\varepsilon_{ijk}^{(n)}| \leq p_i p_{ij}^{(n)} q_{jk}^{(n)} \tag{4.5}$$

$$\sum_i \varepsilon_{ijk}^{(n)} = \sum_j \varepsilon_{ijk}^{(n)} = \sum_k \varepsilon_{ijk}^{(n)} = 0. \tag{4.6}$$

Lemma 3 (cf. [1]). *Let* $m_n = \mathrm{Min}_{i,j}\,[p_{ij}^{(n)}]\,\mathrm{Min}_{j,k}\,[q_{jk}^{(n)}]$. *Then the sequence* $\{m_n : n=1,2,\cdots\}$ *is nondecreasing in* n.

Lemma 4 (cf. [5, 6]). *Define* $\delta_{ik}^{(n)}$ *by*

$$p_n(ik) = p_i r_k^{(n)} + \delta_{ik}^{(n)}, \qquad r_k^{(n)} = \sum_i p_i \sum_j p_{ij}^{(n)} q_{jk}^{(n)}. \tag{4.7}$$

Then we have $\delta_{ik}^{(n)} \to 0$ *uniformly for each* (i, k) *as* $n \to \infty$ *if and only if the relation*

$$\lim_{n\to\infty} \tilde{\boldsymbol{P}}^n \tilde{\boldsymbol{Q}}^n = 0 \tag{4.8}$$

holds. Here $\tilde{\boldsymbol{P}}$ *and* $\tilde{\boldsymbol{Q}}$ *are given by* (3.3).

Theorem 4. *Let some* $\{p_i : i=1,2,\cdots,L\}$, $\boldsymbol{P}$ *and* $\boldsymbol{Q}$ *be given.*

(1) *If the variables* X_n, Y_n *and* Z_n $(n\geq 1)$ *define the distribution*

$$p_n(ijk) = p_i p_{ij}^{(n)} q_{jk}^{(n)} + \varepsilon_{ijk} \tag{4.9}$$

where ε_{ijk} *satisfies* (4.2), (4.3) *and* $|\varepsilon_{ijk}| \leq p_i m_1$, *and* $\boldsymbol{P}, \boldsymbol{Q}$ *are so selected as to satisfy* (4.8), *then the interaction information becomes positive for sufficiently large* n, *provided not all* ε_{ijk} *are zeros.*

(2) *If the variables* X_n, Y_n *and* Z_n $(n\geq 1)$ *define the distribution*

$$p_n(ijk) = p_i p_{ij} q_{jk} + \varepsilon_{ijk}^{(n)}, \tag{4.10}$$

where $\varepsilon_{ijk}^{(n)}$ *satisfies* (4.6) *and* $|\varepsilon_{ijk}^{(n)}| \leq p_i p_{ij} q_{jk}$, *then the interaction information becomes negative for sufficiently large* n *when* $\varepsilon_{ijk}^{(n)} \to 0$ $(n\to\infty)$ *uniformly for each state, provided* $\boldsymbol{P}, \boldsymbol{Q}$ *are selected such that* $\tilde{\boldsymbol{P}}\tilde{\boldsymbol{Q}} \neq 0$.

The proof follows from the Lemmas 2–4 and the inequality

$$\begin{aligned}&(1/2)\{\textstyle\sum^+ [\varepsilon_{ijk}^2/(p_i p_{ij} q_{jk})] - \sum^+ [\varepsilon_{ijk}^3/(p_i p_{ij} q_{jk})^2]\}\\ &\quad -\log\Big(1+\sum_{i,k} [\delta_{ik}^2/(p_i r_k)]\Big) \leq J \leq \log\Big(1+\sum_{i,j,k} [\varepsilon_{ijk}^2/(p_i p_{ij} q_{jk})]\Big)\\ &\quad -(1/2)\{\textstyle\sum^+ [\delta_{ik}^2/(p_i r_k)] - \sum^+ [\delta_{ik}^3/(p_i r_k)^2]\},\end{aligned}$$

where $\sum^+$ means the sum over all values of $\varepsilon_{ijk}>0$ or $\delta_{ik}>0$, and (4.1), (4.2) and (4.7) with $p(ijk)$ substituted by $p_1(ijk)$ are assumed.

Acknowledgement. The author is grateful to Professor K. Kunisawa for kind advices and suggestions in the course of the preparation of the paper.

References

[1] Doob, J. L.: Stochastic Processes. John-Wiley & Sons, New York (1953).

[2] McGill, W. J.: Multivariate Information Transmission. I. R. E. Trans. I. T., pp. 93–111 (1954).

[3] Kullback, S.: Information Theory and Statistics. John-Wiley & Sons, New York (1959).

[4] Pinsker, M. S.: Information and Information Stability of Random Variables and Processes. Holden-Day, San Francisco/London/Amsterdom (1964).

[5] Takano, S.: Interaction information and Markov chain. J. Multivariate Analysis, 3, 93–101 (1973).

[6] ——: Characterization of regular Markov chains for asymptotically independent variables. J. Multivariate Analysis, 3, December (1973).