複合事象の情報エントロピー

—相互情報量と相補的分枝構造—

6Ｌ－3

古閑　政
九州東海大学工学部経営管理学科

1.はじめに

いわゆる熱力学的現象がエントロピー増大の法則で説明されるように、「社会現象の情報エントロピーは増大する」という仮定の下で説明できる事例は幾つかみられる[1]。また秤量問題（別名 Odd Ball Problem）におけるように、情報エントロピーがヒューリスティックスとして役立つこともある[2],[3]。

しかし工学的応用として方式設計に活用した例は通信分野以外には見当たらないようである。その原因は、情報エントロピーと情報量との関係や情報量を与えるとされる相互情報量について、具体的問題を取り扱った研究報告が少ない点にあると思われる。

かかる視点から、本研究では情報エントロピーの分枝分解性に基ずく分枝構造を中心に、その具体的応用を検討する。

2.複合事象と相互情報量

有機的な統合体としてのシステムを考察の対象とし、これが幾つかのサブシステムに分割できる構造をもつとする。そして、このサブシステムがとる状態を事象といい、その生起が確率的現象とみなせる場合を対象とする。したがって、サブシステム間の相互関係とは確率値の集合間になんらかの関係があることを言う。

このように、ひとつのシステム内に相互関係のある事象群が存在するとき、それを複合事象ということにする。その最も単純な例として、ＸとＹの２事象を含むシステムについて考える。

Ｘ，Ｙの情報エントロピーをＨ（Ｘ），Ｈ（Ｙ）で表現すれば、相互情報量Ｉ（Ｘ；Ｙ）はＸを先行事象とすると次式の上側、Ｙを先行事象とすると下側で与えられる。

$$I(X;Y)=H(X)-H(X\mid Y) \quad (1)$$

$$=H(Y)-H(Y\mid X) \quad (2)$$

或いはＸの要素事象の集合を $[x_1,x_2,\cdots x_m]$ 、Ｙの要素事象の集合を $[y_1,y_2,\cdots y_n]$ とすれば、両者の複合事象は、両集合の直積によって与えられるので、Ｘを先行事象とするとき次式が得られる。

$$X\cdot Y=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}\cdot[y_1\ y_2\ \cdots\ y_n]$$

$$=\begin{bmatrix} x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ x_2y_1 & x_2y_2 & \cdots & x_2y_n \\ \vdots & & & \vdots \\ x_my_1 & x_my_2 & \cdots & x_my_n \end{bmatrix} \quad (3)$$

Ｙを先行事象としても同様の結果が得られるので、相互情報量は次式の形でも表される。

$$I(X;Y)=H(X)+H(Y)-H(X\cdot Y) \quad (4)$$

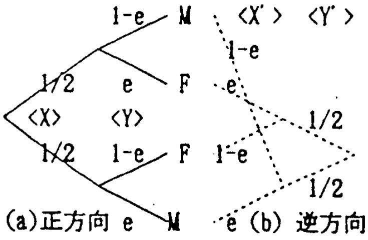
上式でＸ，Ｙの結合が強ければ、複合事象のエントロピーＨ（Ｘ・Ｙ）は小さくなるので、Ｉ（Ｘ；Ｙ）は両事象の多様性を表しながらその結合の強さも含んだ値であることがわかる。

3.相補的分枝構造

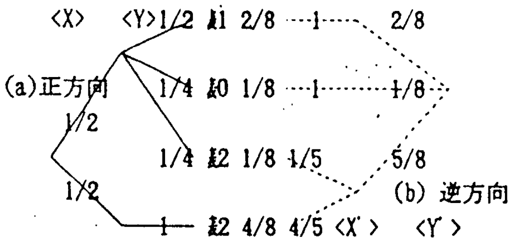
情報エントロピーはグラフ化したとき分枝分解性を有するので、次頁図１，２の実線に示す通り、生起確率値の分布はツリー構造を成す。ここで図１は、1 bit の情報量をもつ情報の例として引用されることの多い「赤ちゃん誕生」の場合である。同図のｅは赤子の性が誤って伝えられる確率である。また第一

Information Entropy of Complex Events System
-Mutual Information and Conplementary Branch Structure

Masashi Koga
Kyushutokai University
9-1-1 Toroku,Kumamoto City,862 Japan



図－1 誕生問題の分枝構造



図－2 やや複雑な事象の分枝構造

事象Xは「どちらの性の赤子がうまれたか」を問題とし、第二事象Yは「どちらの性が伝えられたか」を問題とする事象である。よってXの確率分布は {1/2,1/2} であり、Yの確率分布は同図の破線で示す相補的分枝構造から得られ {1/2,1/2} となる。

要素事象による行列表現をすると

x_1:男の赤子が生まれる事象

x_2:女の赤子が生まれる事象

y_1:性が正しく伝えられる事象

y_2:性が逆に伝えられる事象

$$X\cdot Y = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \cdot [y_1\ y_2] = \begin{bmatrix} x_1y_1 & x_1y_2 \\ x_2y_1 & x_2y_2 \end{bmatrix} \qquad (5)$$

逆方向の相補的分枝構造については

$$Y\cdot X = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \cdot [x_1\ x_2] = \begin{bmatrix} y_1x_1 & y_2x_2 \\ y_2x_1 & y_2x_2 \end{bmatrix} \qquad (6)$$

この時の相互情報量は

$$I(X;Y) = H(1/2,1/2) - H(1-e,e) \qquad (7)$$

図2は「硬貨が2枚あり、ひとつは両面とも表で他方は表裏があるとする。どちらの硬貨を選ぶかという試行をXとし、選んだ硬貨を2回なげて表裏がどうでるかという試行をYとする」問題に対応した分枝構造である。これを要素事象の行列で表現すると

$$X\cdot Y = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \cdot [y_1\ y_2\ y_3] = \begin{bmatrix} x_1y_1 & x_1y_2 & x_1y_3 \\ x_2y_1 & x_2y_2 & x_2y_3 \end{bmatrix} \qquad (8)$$

$$Y\cdot X = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \cdot [x_1\ x_2] = \begin{bmatrix} y_1x_1 & y_1x_2 \\ y_2x_1 & y_2x_2 \\ y_3x_1 & y_3x_2 \end{bmatrix} \qquad (9)$$

要素事象の内容を考慮すると、X・YとY・Xとは異なるとみるべきだが、H(X·Y)=H(Y·X)であるから相互情報量は（４）式で与えられる。

4．相互情報量の意義

4.1原因推定の例　通信の場合と同じように、条件付平均エントロピーH（X｜Y）が小さいほど情報源Xで何が起こったかを現象側Yにおける生起事象から推定するのが可能である。

具体例として「２種類の壺から紅白の球を取り出す」問題を考える。事象Xは、A，Bのどちらの壺を選ぶかであり、事象Yは、紅白の球のどちらを取り出すかである。そこで壺の選択確率はいずれも1/2で、紅白の球の数が以下の３通りの場合の相互情報量を求める。

（１）Aに紅２個と白３個、Bに紅３個と白２個

（２）Aに紅１個と白２個、Bに紅２個と白１個

（３）Aに紅１個と白３個、Bに紅３個と白１個

いずれの場合も H(X)= H(Y)= H(1/2,1/2) であり H(X·Y)の相違によって相互情報量が (1)0.028 (2)0.082 (3)0.185 bit となっている。

4.2データ処理の例

小売業のＰＯＳデータの処理について考える。この場合 $[x_1,x_2,\cdots x_m]$ は客毎の購買額を示すトランザクション·データで、$[y_1,y_2,\cdots y_n]$ は商品分類毎の売上額を表している。

そこで、現実の販売動向に対応したYの比率分布を $A=(a_1\ a_2\ \cdots a_n)$ とし、粗利益額を望ましい構成とする比率分布を $B=(b_1\ b_2\ \cdots b_n)$ と置く。また商品分類毎の設定粗利益率を $P=(p_1\ p_2\ \cdots p_n)$ とすれば、販売の立場から望ましいのは

$$H(Y) = -\sum_{i=1}^{n} b_i \operatorname{Log} b_i \quad (10) \qquad P(Y) = \sum_{i=1}^{n} p_i b_i \quad (11)$$

の両者をできるだけ大きくすることである。よって H(Y)·P(Y)を最大にする分布Bをラグランジュの未定乗数法により求め、分布Aをその値に近付けるのが販売戦略の目標となる。

5．おわりに

システムの内部事象について、相互情報量を検討すれば、設計及び運用の指針が得られることを述べた。

（文献 (1)(2)(3)は、発表時に紹介する）