The Society for Economic Studies
The University of Kitakyushu
Working Paper Series No. 2015-8
(accepted in March 31, 2016)

# 情報の確率論的形式化

朱　乙文*

人々は、日常生活の中で、常に様々な情報（information）を受け取り、利用する。健康状態を確かめたい人は、医者にかかって診断をもらい、明日の天気が気になる人はインターネットの天気予報サイトにアクセスして、明日の天気予報を確認する。情報は、医者の診断や明日の天気予報であり、医者やインターネットの天気予報サイトは情報源（information source）である。それゆえ情報源とは情報サービス（information service）を提供するものをいう。

情報の動きが経済に大きく影響することは、長い経験の歴史の中から明らかになっている。しかしながら、経験の長い歴史からすると、体系的な情報の研究は始まったばかりであると言っても過言ではない。急速なＩＴ技術進歩により、種々の新しい情報や情報源が出現し、情報環境が劇的に変化する今日においては、特に、様々な側面からの情報の研究は、社会科学においてもっとも重要な課題の一つとなる。

情報は、通常の経済財と異なる特殊な性質を有するものであるだけではなく、その動きの範囲も極めて広く複雑なものになっており、このことが情報の研究の進展を大きく妨げている。それにも関わらず、近年、「情報の経済学」の発展には目を見張るようなものがある。著しく発展を遂げている「情報の経済学」の議論では、情報は、主に、その基本的形式化において、確率論的概念に基づいたものとなっている。本稿では、情報の確率的形式化に焦点を当て、「情報の経済学」の議論を検討し、今後の研究の方向として、新しい「情報の経済学」のフロンティアについて考える。

＊　北九州市立大学経済学部

## １．確率論的思考における情報

古く１７世紀半ばまで遡り、賭けから編み出されたとされる確率論[1]は、大まかには、不確定な偶然現象の「確からしさ」について研究を行う数学の一分野である。ここでは、経済分析において、広く用いられている情報の確率論的捉え方について考える。

### 1-1「不確実性」の確率論的捉え方

一般に、人々にとって、現実の経済を取り巻く環境は確かではない性質に満ちたものである。経済学の分野においても、いち早く、２０世紀を代表する経済学者の一人であるF. Knight は、このような状況に対して先駆的な眼で光を当てている。Knight(1921)は、確かではない状況を確率的に測定可能なリスクと測定不可能な不確実性にも明確に区分している[2]。すなわちある状況に対して、確率の概念を用いて示すことのできる場合がリスク(risk)で、確率の概念を用いて示すことのできない場合が不確実性(uncertainty)となる。

確率とは、あるイベント（事象）が起こる可能性を数値で表したものであるが、確率論における確率については、基本的に、三つの異なる概念がある。

その一つは、数学的確率(mathematical probability)もしくは先験的確率(A priori probability)と呼ばれるものである。具体的に、これは、同じ程度に確からしく起こるすべての事柄の集合をΩとしたとき、Ωの部分集合として示されるイベント$E_i$が起こる確率$P(E_i)$を次のように表すものである$\left(i=1,2,\cdots,n\right)$。

$$P(E_i)=\frac{\text{イベント}E_i\text{が起こる事柄の数 }n_i}{\text{起こり得るすべての事 柄の数}N}$$

ただし、このような確率概念においては、同じ程度に確からしく起こり得ない事柄については確率の定義ができなくなる。それゆえ、この確率概念の下では、確率定義できるイベントが限定されたものにならざるを得ない。

もう一つは、統計的または経験的に捉えた確率の概念であるが、統計的確率(statistical probability)もしくは経験的確率(empirical probability)と呼ばれるものである。これは、同

[1] 確率論は、ブレーズ・パスカル（Blaise Pascal、1623-1662、フランスの哲学者）とピエール・ド・フェルマー（Pierre de Fermat, 1601-1665, フランスの数学者）の往復書簡によって始まったとされる（『岩波数学辞典第３版』、岩波書店、p.127）。ここで、往復書簡とは、シュヴァリエ・ド・メレ（1610-1684）が賭け事に関する問題、たとえば、２つのサイコロを 24 回振って 6 のゾロ目が 1 回以上出る確率に関する問題などをパスカルに問いかけたものである。（この問いの解などについては、林周二『統計学講義第２版』丸善、昭和４８年、p.73.を参照せよ）。

[2] 20 世紀を代表するもう一人の経済学者である J. M. Keynes も、Knight(1921)とほぼ同時期に、このような状況に対して、鋭い洞察をしている。Keynes(1921)では、確かではない状況を不確実性と蓋然性（probability）に峻別している。不確実性とは、まったく無知の状態であり、数値で示すこともできないし、比較可能でもない状態をいい、蓋然性とは、もっと狭義としてリスクから、一つの数値で示すことができなく、相互比較が難しい状態まで、より細分化された分類をしている。詳しくは、酒井 泰弘(2013)を参照せよ。

じ条件のもとで、独立な試行(trial)を繰り返して行い、導かれる確率である。すなわち、ある独立な試行を $N$ 回繰り返して行い、あるイベント $E$ が $n$ 回起こったとき、

$$\lim_{N\to\infty}\frac{n(E)}{N}=P(E)$$

であれば、$P(E)$ はイベント $E$ の起こる確率とする。これは、試行回数が増加するにつれて、収束する相対頻度の値に他ならない。それゆえ、このような確率概念の下では、同じ条件のもとで繰り返し試行できない状態においては、イベントの確率は定義できないとものなる。

上の二つの確率の概念は経験や実験、または論理を通じて求められるという意味で、客観的確率（objective probability）であると考えられる。最後の概念は、これらの客観的確率と異なり、確率をイベントの生起の不確実性の程度を主観的に評価したものとして捉える主観的確率（subjective probability）である。すなわち、あるイベント $E$ の発生に対する個人の確信の度合いを表した数値である $P(E)$ をイベント $E$ の主観的確率という。それゆえ、この確率概念の下では、あるスポーツ大会における優勝チームについての予想のように、同じイベントであっても、個々人によって、イベントの生起確率は異なり得る。しかし、現実の経済社会においては、実際、すべての事柄が同じ程度に確からしく起こり得ない状況や、何度も試行を繰り返すことができない状況が一般的である。それゆえ、このような主観的確率の概念は、より幅広い状況のイベントに対して確率定義が可能なものとなっている。

このように、確率論では、異なる複数の確率の概念を取り上げているが、いずれの確率の概念においても、一般に、確率の満たす公理論的性質が存在する。同じ程度に確からしく起こるすべての事柄の集合を $\Omega_X$ とし、$\Omega_X$ の任意の二つの部分集合 $X_i$ と $X_j$ に対して、$X_i\cap X_j=\phi$ であり $(i,j=1,2,\cdots,n,i\neq j)$、$\Omega_X=\{X_1,X_2,\cdots,X_n\}$ であるとする。言い換えると、集合 $\Omega_X$ の分割（partition）における一つの部分すなわちセルが部分集合 $X_i$ である。ここで、ある集合の分割とは、その集合全体を互いに重ならないセルに分けることを言う。なお、これらの部分集合もしくはセルは事象もしくはイベントと呼ばれる。この場合、イベント $X_i$ に対応する実数を $P(X_i)$ とすると、次の 3 つの条件を満たす実数 $P(X_i)$ をイベント $X_i$ が起こる確率という。

（ｉ）任意のイベント $X_i$ に対して、$0\leq P(X_i)\leq 1$ である。

（ⅱ）$P(\Omega_X)=1$、$P(\phi)=0$ である。

（ⅲ）$X_i\cap X_j=\phi$ ならば $(i\neq j)$、$P(X_i\cup X_j)=P(X_i)+P(X_j)$ である。

ここで、$\Omega_X$ は同じ程度に確からしく起こるすべての事柄の集合、すなわち全イベント集合であり、$\phi$ はどのような事柄も存在しない空集合である。より一般的に、$\Omega_X$ 上で、それぞれの事象に確率を対応させる関数は確率分布（関数）と言われる。なお、それぞれの事象が可測で実数を用いて表すことができるとき、これを確率変数とする。

より一般的な状況における確率の概念を示すために、複数のサイコロを同時に振る場合

のように、複数の事象が同時に起こる状況を考える。たとえば、二つのサイコロを同時に振る場合、二つのサイコロの目の和が偶数である事象と奇数である事象の集合の中から一つの事象と、二つのサイコロの目の和が 6 未満となる事象と 6 以上となる事象の集合の中から一つの事象が同時に生起する状況を考えることができる。

## 1-2 ベイズ主義と情報

「ベイズ主義」的観点からすると、確率論的思考を用いて、より形式的に情報をとらえることができる。「ベイズ主義」とは、18 世紀のイギリスの数学者であるトーマス・ベイズ（Thomas Bayes、1702 年 - 1761 年）に遡る考え方である。これは、確率は個々人によって主観的に評価される信念の度合いとしてとらえられるので、信念の度合いは個人内的操作によって更新され得るというものである。言い換えれば、「ベイズ主義」とは、あるイベントについて最初に形成された信念の度合もしくは事前確率(prior probability)は、一定のルールによって、更新された信念の度合もしくは事後確率(posterior probability)へ確率変換が可能であるという考え方をいう。

ベイズ自身によって示された事前確率から事後確率への更新ルールはベイズの定理（Bayes' theorem）と呼ばれる。具体的には、ベイズの定理とは、ある確率的因果関係にある二つの事象、すなわち原因を示す事象$Y$と結果を示す事象$X$に対して、「事象$X$が起こる条件の下で、事象$Y$の起こる確率」について、次のような関係を示すものである[3]。

$$P(Y \mid X) = \frac{P(X \mid Y)P(Y)}{P(X)}$$

ここで、$P(\bullet)$はそれぞれのイベントの起こる確率であるが、$P(Y)$はイベント$Y$の起こる事前確率[4]、$P(Y \mid X)$はイベント$Y$の事後確率とも言われる。また$P(X \mid Y)$は、次のように定義される、イベント$Y$が生起する条件の下でイベント$X$が生じる条件付き確率である。

$$P(X \mid Y) = \frac{P(X \cap Y)}{P(Y)}$$

---

[3] ある二つのイベント $X$ と $Y$ がある場合を考える。この場合、イベント $X$ と $Y$ が同時に起こる確率（＝同時確率）$P(X \cap Y)$と$P(Y \cap X)$については、それぞれ、次の関係が成立する。

$$P(X \cap Y) = P(X \mid Y)P(Y)$$
$$P(Y \cap X) = P(Y \mid X)P(X)$$

ところで、$P(X \cap Y) = P(Y \cap X)$であるので、$P(X \mid Y)P(Y) = P(Y \mid X)P(X)$となる。この関係を$P(Y \mid X)$について整理すると、ベイズの定理が導かれる。より一般的に、$\Omega = \{Y_1, Y_2, \cdots, Y_n\}$とし$(i = 1,2,\cdots,n)$、任意の$i$と$j$に対して$(i \neq j)$、$Y_i \cap Y_j = \phi$である場合は、ベイズの定理は、次のように示される。

$$P(Y_i \mid X) = \frac{P(X \mid Y_i)P(Y_i)}{\sum_{i=1}^{n} P(X \mid Y_i)P(Y_i)}$$

[4] ここでは、情報という観点から、確率の概念を捉えていく。それゆえ、ここで議論する事前確率の概念としては、客観的確率と主観的確率をすべて対象としている。

また、この関係を書き直すと、$P(X \cap Y) = P(X \mid Y)P(Y)$ となる。

ただし、事象 $Y$ と事象 $X$ に間の因果関係を考慮すると、$P(X \mid Y)$ は結果を示す事象 $X$ が生起した場合における原因を示す事象 $Y$ の確からしさを示すもの（尺度）であるとみなす（考える）ことができる。このような見方の条件付き確率は尤度（ゆうど、likelihood）とも呼ぶ。より一般的に、$\Omega = \{Y_1, Y_2, \cdots, Y_n\}$ であり $(i = 1,2,\cdots,n)$、任意の $i$ と $j$ に対して $(i \neq j)$、$Y_i \cap Y_j = \phi$ である場合におけるベイズの定理は、次のように示される。

$$P(Y_i \mid X) = \frac{P(X \mid Y_i)P(Y_i)}{\sum_{i=1}^{n} P(X \mid Y_i)P(Y_i)}$$

簡単な例として、年初めに福袋を購入しようとするある買い物客の例を取り上げ、ベイズ主義の観点からの情報の概念について考えよ。ある店の店頭には外見上はまったく区別がつかない２種類の福袋 $X_1$ と $X_2$ がそれぞれ同じ数だけ並べられているとする。ただし、この店のインターネットホームページにはすでに、それぞれの種類の福袋に入っている品物の内容のみが（潜在的）買い物客に公開されており、この買い物客は、そこから、福袋 $X_1$ には自分の好みの品物が４点、また好みではない品物が２点入っており、福袋 $X_2$ には自分の好みの品物が２点、また好みではない品物が４点入っていることを知ることができたとする。この買い物客がこの店の店頭で一つの福袋を購入しようとしたとき、ある一つの福袋の隅から、偶然、自分の好みの品物の一点が覗かれた場合に、この福袋が自分の好みの品物が多く入っている福袋 $X_1$ である確率を求めよ。ここで、説明の簡単化のために、それぞれの品物が覗かれるのは同程度に確からしく起こるとする。

この場合、２種類の福袋 $X_1$ と $X_2$ についての事前確率はそれぞれ、$P(X_1) = \frac{1}{2}$、$P(X_2) = \frac{1}{2}$ となる。また $P(\text{自分の好みの品物}) = \frac{1}{2} \times \frac{4}{6} + \frac{1}{2} \times \frac{2}{6} = \frac{1}{2}$ であり、$P(X_1 \cap \text{自分の好みの品物}) = \frac{1}{2} \times \frac{4}{6} = \frac{1}{3}$ である。それゆえ、ベイズ定理を用いると、自分の好みの品物の一点が覗かれた福袋が自分の好みの品物が多く入っている福袋 $X_1$ である事後確率は、次のように導かれる。

$$P(X_1 \mid \text{自分の好みの品物}) = \frac{P(X_1 \cap \text{自分の好みの品物})}{P(\text{自分の好みの品物})} = \frac{1}{3} \div \frac{1}{2} = \frac{2}{3}$$

したがって、この場合、ある福袋が自分の好みの品物が多く入っている福袋 $X_1$ である事前確率が $\frac{1}{2}$ から、その福袋から自分の好みの品物の一点を覗いたこと、すなわち一種のメッセージを受けたことによって、その福袋が $X_1$ である事後確率が $\frac{2}{3}$ となる。このような例からすると、ある福袋が $X_1$ である確率は、その福袋から自分の好みの品物の一点を覗いたことによって、$\frac{1}{2}$ から $\frac{2}{3}$ に変わる。それゆえ、この場合には、覗かれた自分の好みの品物の

一点がその福袋が何であるかを判断する際の情報もしくはメッセージとなる。

## ２．情報と情報構造：情報構造論

ここでは、確率論的考え方の下で、様々な情報源と情報を統合的に捉えられる理論的枠組み、すなわち情報構造（information structure）に情報構造の特徴ついて説明する。

### 2-1　情報構造とは

情報構造とは、大まかにいうと、情報源の質的状態を形式・体系的に示したものである。生起し得る「自然の状態」（states of nature）もしくは「世界の状態」（states of world）の集合もしくはリストである状態空間を$S=\{s_1, s_2, \cdots, s_n\}$と表し、それぞれの状態と関連する特定の事柄の集合もしくはリストを$M=\{m_1, m_2, \cdots, m_m\}$と示すと、情報の「質的体系」とは、状態空間$S$から集合$M$への特定の関係を表すものである。なお、特定の事柄について、具体的に説明していくため、それぞれの状態と関連する特定の事柄をメッセージ(messages)であるとし、$M$をメッセージ空間と呼ぶことにする。

説明の簡単化のために、メッセージ空間を状態空間の分割（partition）概念を用いて示すこともできる。状態空間の分割は状態空間の空でない部分集合であるセル（cell）もしくはブロック（block）の集合として示される。例えば、状態空間$S$の分割の一つを$\Phi=\{S_1, S_2, \cdots, S_m\}$としよう[5]。このとき、メッセージ$m_j$を$S_j \equiv \left\{ s \in S \middle| \phi(s)=m_j \right\}$を満たすものとして定義すると$(j=1,2,\cdots,m)$、状態空間$S$とメッセージ空間$M$との関係は、$S$の分割$\Phi$のみを用いて単純な形で示すことができる。より具体的に、$S=\{s_1, s_2, s_3, s_4\}$とし、$S$の二つの分割$\Phi^A$と$\Phi^B$をそれぞれ、次のようなものであるとする。

$$\Phi^A=\{\{s_1\},\{s_2, s_3, s_4\}\}、\Phi^B=\{\{s_1\},\{s_2, s_3\}\{s_4\}\}$$

分割$\Phi^A$と$\Phi^B$を図示すると、＜図表１＞のようになる。＜図表１＞は、分割$\Phi^B$のすべてのセルは分割$\Phi^A$のセルの部分集合になっていることを示している。この場合、$S$の分割$\Phi^B$は分割$\Phi^A$の細分（refinement)という。すなわちより分割（もしくはメッセージ、情報）が細かいことを表す。

同時確率や条件付き確率の概念を用いると、より厳密に、情報構造$I_M$は次のように示される。

$$I^M=(M, v(s_i, m_j))、もしくは I^M=[L_{ij}] \quad (i=1,2,\cdots,n,\ j=1,2,\cdots,m)$$

ただし、$v(s_i, m_j)$は状態$s_i$とメッセージ$m_j$の同時確率であり、$[L_{ij}]$は、次のような、メッ

---

[5] より厳密には、集合$S$が空集合でない場合、$\Phi=\{S_1, S_2, \cdots, S_m\}$が次の条件を満たすとき、$\Phi$を$S$の分割という。

i) $S=S_1 \cup S_2 \cup \cdots \cup S_m$であること、ただし、すべての$j(j=1,2,\cdots,m)$に対して、$S_j$が空でないこと

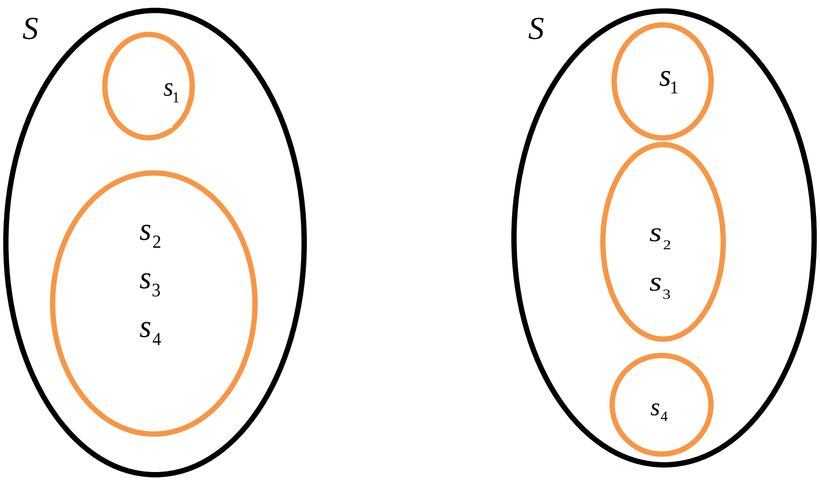
ii) 任意の$i, j(i \neq j)$について、$S_i \cap S_j=\phi$であること

セージ $m_j$ の状態 $s_i$ に対する条件付き確率 $L_{ij} = L(m_j \mid s_i)$、すなわち尤度の（マルコフ）行列である。

$$I^M = [L_{ij}] = \begin{pmatrix} L(m_1 \mid s_1) & L(m_2 \mid s_1) & \cdots & L(m_m \mid s_1) \\ L(m_1 \mid s_2) & L(m_2 \mid s_2) & \cdots & L(m_m \mid s_2) \\ \vdots & \vdots & \ddots & \vdots \\ L(m_1 \mid s_n) & L(m_2 \mid s_n) & \cdots & L(m_m \mid s_n) \end{pmatrix} \text{---------------------------}(1)$$

このように、情報構造とは、メッセージもしくは情報を、「自然の状態」についての初期知識もしくは信念（belief）との関わりの中で、形式化し捉える分析的枠組みを提供するものである。

＜図表１＞状態空間の分割と細分




## 2-2　ノイズのない情報構造 vs. ノイズのある情報構造

ここでは、情報の質という観点から、ノイズのない情報構造(information structure without noise)とノイズのある情報構造(information structure with noise)に大別し、情報構造の特徴について説明する[6]。

### 2-2-1 ノイズのない情報構造

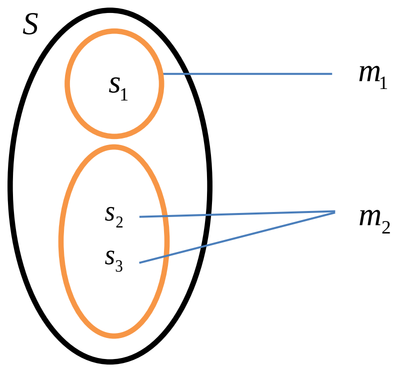
正確な時間を知らせる時計の例を考えよ。典型的な時計とは、時計の針の位置と１日

[6] ここでは、情報構造論の基本的解説のみを行っている。ノイズのない情報構造とノイズのある情報構造についてのより詳細な議論は、Laffont (1989)（第４章）や Marschak = Radner (1972) (第２章) などを参照せよ。また、ノイズのない情報構造とノイズのある情報構造を含むより一般的な枠組みを用いた情報構造関する議論については、Green (1981) や Milgrom = Stokey (1982) などを参照せよ。

における時間的状態とを関連付けるサービスを提供するものである。ここで、時計の針の示す時刻の集合はメッセージ空間であり、時間的状態の集合は状態空間であるとすることができる。正確な時間を知らせる時計においては、１日の中の時間的状態が決まれば、常にそれに対応する時計の針の位置が決まる。それゆえ、時計の針の観察者にとっては、その位置が分かれば、1 日の中のどのような時間帯にいるか分かることになる。このように、状態空間からメッセージ空間への「特定の関数」が定まっている情報構造をノイズのない情報構造という。ノイズのない情報構造下においては、特定のメッセージが得られると正確に状態を推論することができる。

より一般的に、ノイズのない情報構造とは、尤度の（マルコフ）行列におけるすべての成分(確率)が 1 かゼロとなっている情報構造をいう。たとえば、状態空間が $S=\{s_1,s_2,s_3\}$、またメッセージ空間が $M=\{m_1,m_2\}$ であり、$S$ と $M$ との関係を表す $f:S\to M$ が＜図表 1 ＞のように、関数として示される場合を考えよ。なお、$p(s_i)=\dfrac{1}{3}$ であるとする（$i=1,2,3$）。

＜図表２＞ノイズのない情報構造




＜図表２＞で示された関係を尤度の（マルコフ）行列で示すと、次のようになる。

$$[L_{ij}]=\begin{pmatrix} L(m_1\mid s_1) & L(m_2\mid s_1) \\ L(m_1\mid s_2) & L(m_2\mid s_2) \\ L(m_1\mid s_3) & L(m_2\mid s_3) \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \quad (i=1,2,3,\ j=1,2)$$

それゆえ、＜図表２＞の例で示される情報構造は、尤度の（マルコフ）行列におけるすべての成分（確率）が 1 かゼロとなっているので、ノイズのない情報構造となる。ノイズのない情報構造は、さらに、それぞれの状態とメッセージの対応状況によって、完全情報構造（perfect information structure）と不完全情報構造（imperfect information structure）に分類することができる。完全情報構造は一つの状態に一つのメッセージが対応している情報構造をいい、不完全情報構造は、＜図表１＞の例のように、複数の状態に一つのメッセージが対応している情報構造をいう。いずれにしても、これらの情報構造においては、

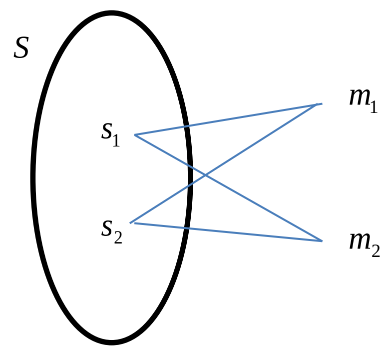
特定のメッセージが観察されると、それによって、真の状態がどのようなイベントに属しているかが正確に推論され得ることになる。

### 2-2-2 ノイズのある情報構造

おいしいワインを選ぶために、ソムリエ資格を持つ専門家にアドバイスを求める場合の例を考えよ。この例においては、おいしいワインかどうかが状態空間を表し、ソムリエのアドバイスがメッセージ空間を表すものとなる。有能なソムリエのアドバイスはおいしいワインを選ぶ際には重要な情報となるが、いくら資格を持つ専門家であっても間違ったアドバイスをするケースも少なくない。このケースでは、特定のワインに対して、おいしいワインであるメッセージを伝達する場合もあり、おいしくないワインであるというメッセージを伝達する場合もあることになる。このような状況下においては、メッセージにはノイズが存在することになり、この例における情報構造はノイズのある情報構造として分類される。

より一般的に、ノイズのない情報構造と異なり、尤度の（マルコフ）行列における成分が０と１以外の成分を含むものとして示される情報構造をノイズのある情報構造という。たとえば、状態空間が$S=\{s_1,s_2\}$、またメッセージ空間が$M=\{m_1,m_2\}$であり、$S$と$M$との関係を表す$f:S\to M$が＜図表２＞のように、関数ではなく、対応として示される場合を考えよ。なお、$p(s_i)=\frac{1}{2}$であるとし（$i=1,2$）、$L(m_2|s_1)=\frac{1}{3}, L(m_1|s_2)=\frac{1}{4}$であるとする[7]。

＜図表３＞ノイズのある情報構造




[7] ワイン選びの例からすると、状態空間の元（element）については、$s_1$＝おいしいワイン、$s_2$＝おいしくないワインをそれぞれ示す。またメッセージ空間の元については、ソムリエのアドバイスもしくはメッセージは、おいしいワインについては、平均的に、３回に１回は「おいしくないワインである」と間違ったアドバイスをし（また３回に２回は正しいアドバイスをする。）、おいしくないワインについては、平均的に、４回に１回は「おいしいワインである」と間違ったアドバイスをする（また４回に３回は正しいアドバイスをする。）状況として考えることができる。それゆえ、ワイン選びの例におけるソムリエのメッセージはノイズもしくは誤差が介在しているものとなっている。

＜図表３＞で示された関係を尤度の（マルコフ）行列で示すと、次のようになる。

$$[L_{ij}]=\begin{pmatrix} L(m_1 \mid s_1) & L(m_2 \mid s_1) \\ L(m_1 \mid s_2) & L(m_2 \mid s_2) \end{pmatrix}=\begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix} \quad (i=1,2,\ j=1,2)$$

このように、＜図表３＞の例で示される情報構造は、尤度の（マルコフ）行列における成分が０と１以外の成分を含むものとして示される情報構造であるので、ノイズのある情報構造となる。それゆえ、ノイズのある情報構造におけるメッセージにはノイズが内在しているものとして考えることができるが、いずれにせよ、ノイズのある情報構造においては、特定のメッセージが観察されると、それによって、真の状態がどのようなイベントに属しているかは確率的に推論され得ることになる。

2-3 同質的情報構造 vs. 異質的情報構造

ここでは、情報の内容という観点から、同質的情報構造(homogeneous information structure)とノイズのある情報構造(heterogeneous information structure)に大別し、情報構造の特徴について説明する。

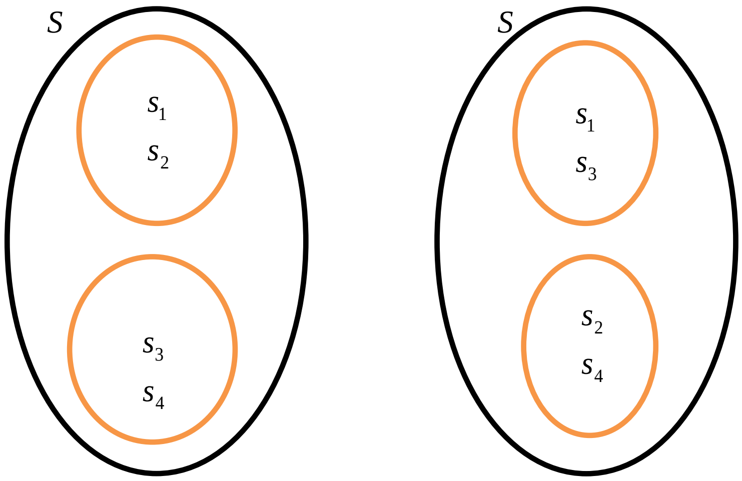
ある状態空間を $S=\{s_1,s_2,\cdots,s_n\}$ と表し、状態空間 $S$ の二つの分割を $\Phi^l=\{S_1^l,S_2^l,\cdots,S_{m^l}^l\}$ としよう（$l=A,B$）。このとき、状態空間 $S$ の分割 $\Phi^B$ は、分割 $\Phi^A$ の任意のセル $S_j^A$ を部分集合として含むあるセール $S_k^B$ が常に存在する場合（その逆の場合も同様である。）、分割 $\Phi^A$ と $\Phi^B$ として示される情報構造は、同質的情報構造と呼ばれる（$j=1,2,\cdots,m^A,k=1,2,\cdots,m^B$）。これに対して、分割 $\Phi^B$ は、分割 $\Phi^A$ のあるセル $S_j^A$ を部分集合として含むあるセール $S_k^B$ が存在しない場合（その逆の場合も同様である。）、分割 $\Phi^A$ と $\Phi^B$ として示される情報構造は、異質的情報構造と呼ばれる。

このような情報構造の分類は、直感的には、二つの分割におけるセルに重なる元が存在するかどうかによって行われる。それゆえ、＜図表１＞のように示される二つの情報構造は同質的情報構造となる。これに対して、＜図表４＞のように示される二つの情報構造は異質的情報構造となる[8]。ここで、同質的情報構造から得られるメッセージもしくは情報は同質的情報と呼ばれる。それゆえ、「情報の経済学」で議論されている非対称的情報

[8] ＜図表４＞の例においては、左の例は二つのセルはそれぞれ、下添え字のナンバーが小さい数字か大きな数字かを示すメッセージである。これに対して、右の例は、下添え字のナンバーが奇数か偶数かを示すメッセージである。それゆえ、これらの情報構造は異質的情報構造となる。

（asymmetric information）は同質的情報として分類される。一方、異質的情報構造から得られるメッセージもしくは情報は異質的情報と呼ばれる。

＜図表４＞異質的情報構造




３．情報と情報サービスの価値

一般に、メッセージの受け手にとって、伝達されるメッセージの情報としての評価は、情報構造だけではなく、情報もしくはメッセージをどのように利用するかによって異なるものとなる。ここでは、メッセージの受け手が受け取ったメッセージを情報として利用し意思決定を行う場合、メッセージの受け手にとってのメッセージとメッセージ・サービスの価値について説明を行う。

3-1 情報の価値

メッセージの受け手にとってのメッセージの価値は、具体的に、メッセージを利用する場合に得られる利得とそれを利用しない場合の利得との差として求められる、具体的に、ある意思決定者に直面する発生可能な状態の集合である状態空間を $S$ 、意思決定者の選択可能な行動就集合を $A$ とする場合を考えよ。ここで、$a \in A$ とし、状態空間上に、状態 $s_i$ の発生についての意思決定者の初期知識もしくは信念（belief）を示す、確率分布 $p(s_i)$ が定義されているとする $(i = 1, 2, \cdots, n)$。

まず、どのようなメッセージも利用せずに、意思決定行うケース、すなわち無情報のケースにおいて、意思決定者が行動 $a(\in A)$ を選択して得られる期待利得は

$EU_0(a,s)=\sum_i U(a,s_i)p(s_i)$ であるので、無情報のケースにおいて得られる最大期待利得 $V^0$ は、(2)式のように示すことができる。

$$V^o = EU^0(a^0,s)=\sum_i U(a^0,s_i)p(s_i) \qquad \text{--------------------------------}(2)$$

ここで、$a^o = \underset{a}{Max}\, EU^0(a,s)$、すなわち $a^o$ は無情報のケースにおける最適行動である。

次に、メッセージ空間を $M$ とし、メッセージ $m_j \in M$ を利用して意思決定行うケースにおいては、意思決定者が行動 $a$ を選択して得られる期待利得は次のように示される。

$$EU^{m_j}(a,s)=\sum_i U(a,s_i)p(s_i|m_j)$$

それゆえ、メッセージ $m_j$ を利用するケースにおいて得られる最大期待利得 $V^{m_j}$ も、(3)式のように表す $(j=1,2,\cdots,m)$。

$$V^{m_j} = EU^{m_j}(a^{m_j},s)=\sum_i U(a^{m_j},s_i)p(s_i|m_j) \qquad \text{---------------------------}(3)$$

ここで、$a^{m_j} = \underset{a}{Max}\, EU^{m_j}(a,s)$、すなわち $a^m$ はメッセージ $m_j$ を利用するケースにおける期待利得を最大化する行動である。また $p(s_i|m_j)$ は、メッセージ $m_j$ を用いてベイズ定理によって更新された、起こり得る状態 $s_i$ についての事後確率を示す。

以上のことから、メッセージ $m_j \in M$ を利用して意思決定行うケースにおけるメッセージ $m_j$ の価値は、メッセージ $m_j$ を利用するケースにおいて得られる最大期待利得 $V^{m_j}$ と無情報のケースにおいて得られる最大期待利得 $V^0$ との差として、(4)式のように示される。

$$V(m_j)=V^{m_j}-V^0=\sum_i U(a^{m_j},s_i)p(s_i|m_j)-\sum_i U(a^0,s_i)p(s_i) \qquad \text{-----------}(4)$$

3-2 情報サービスの価値

メッセージの受け手にとってのメッセージ・サービスの価値も、同様に、メッセージ・サービスを利用する場合に得られる利得とそれを利用しない場合の利得との差として求められる。メッセージ・サービスとは発生可能なメッセージを受け取ることができるサービスである。それゆえ、メッセージ・サービスの価値は、受け取ることができるそれぞれの

メッセージについての最大期待利得の期待値として考えることができる。

より具体的に、メッセージ・サービスを用いて意思決定をする場合の最大期待利得$V^M$は、次のように示すことができる$(j=1,2,\cdots,m)$。

$$V^M = EV^{m_j} = E\left\{EU^{m_j}(a^{m_j},s)\right\} = \sum_j \left\{\sum_i U(a^{m_j},s_i)p(s_i|m_j)\right\}p(m_j) \qquad \text{-------(5)}$$

ここで、$p(m_j)$はメッセージ$m_j \in M$を受け取る（周辺）確率である。それゆえ、無情報のケースにおける最大期待利得$V^0$を用いると、メッセージ・サービスの価値は、(2)式と(5)式を用いて、$V(M)=V^M - V^0$として示される。ただし、ここでの情報と情報サービスの価値は、説明の簡単化のために、明示的に情報費用を考慮に入れない、もしくは情報と情報サービスの純価値として取り扱ったものである。

ところで、メッセージ・サービスを用いて意思決定をする場合の最大期待利得$V^M$は、最適行動$a^{m_j}$を選択する場合の期待利得であり、$a^{m_j}$以外の行動を選択する場合は期待利得が$V^M$を上回ることはない。それゆえ、このことから、常に、次の関係が成り立つ。

$$\sum_j \left\{\sum_i U(a^{m_j},s_i)p(s_i|m_j)\right\}p(m_j) \geq \sum_j \left\{\sum_i U(a^0,s_i)p(s_i|m_j)\right\}p(m_j) \qquad \text{-------(6)}$$

ここで、$a^0$は無情報のケースにおいて選択する最適行動である。さらに(5)式の右辺は、ベイズ定理の関係を用いると、次のように書き直される。

$$\sum_j \left\{\sum_i U(a^0,s_i)p(s_i|m_j)\right\}p(m_j) = \sum_j \left\{\sum_i U(a^0,s_i)\frac{L(m_j|s_i)}{p(m_j)}p(s_i)\right\}p(m_j)$$

$$= \sum_i U(a^0,s_i)p(s_i) = V^0 \qquad \text{----------------(7)}$$

ただし、(7)式の第２行目の式は、$\sum_j L(m_j|s_i)=1$であることを利用し、書き直したものである。

したがって、(5)式と(7)式の関係を用いると、メッセージ・サービスの価値は、常に非負、すなわち$V(M)\geq 0$となることが導かれる。

### 3-3 情報サービスの比較：比較情報性の理論（theory of comparative informativeness）

前項で示したように、一般に、情報サービスの価値は、情報構造だけではなく、意思決定者の利得関数によって大きく異なるが、情報サービスの価値は、常に非負となる。しかし、このことは、情報の量と情報もしくは情報サービスの価値（もしくは生産性）が常に正の関係を持つことを示すものではない。

すでに、Radner = Stiglitz (1984)は、意思決定問題においては、情報の限界費用が正で

ある限り、少量の情報は負の限界純価値（もしくは期待利得）をもたらし得ることを示している。これは、正の純価値をもたらす情報の量が存在する場合は、情報の価値は情報の量の凹関数として示されないことを意味する[9]。したがって、情報構造もしくは情報サービス間の一般的比較を行う際には、十分な注意を払わなければならない。

情報構造もしくは情報サービスを一般的に比較することは、極めて限られた状況においてのみ可能である。ここでは、D. Blackwell にまで遡る、情報構造もしくは情報サービス間比較の条件についての基本的考え方を紹介する[10]。ここでは任意の初期信念と利得関数の下で、異なる二つの情報構造の比較問題は、大まかには、ある情報構造とこの情報構造に一種のノイズを追加して得られるもう一つの情報構造との比較の問題、すなわち比較情報性（comparative informativeness）の問題として捉えて議論を行っている。

状態空間を $S=\{s_1,s_2,\cdots,s_n\}$ とする場合における二つの情報構造 $I^A$ と $I^B$ について考える。議論を単純化するために、情報構造 $I^B$ は $I^A$ にノイズを追加したものであるとする[11]。なお、二つの情報構造におけるメッセージ空間をそれぞれ、$M^A=\{m_1,m_2,\cdots,m_{m(A)}\}$ と $M^B=\{m_1,m_2,\cdots,m_{m(B)}\}$ と示し、情報構造 $I^A=[L^A_{ij}]$ は $n\times m(A)$ 行列であり、$I^B=[L^B_{ik}]$ は $n\times m(B)$ 行列であるとする。

まず、議論の準備として、ある行列 $R$ が存在して、二つの情報構造 $I^A$ と $I^B$ について、次の関係が満たされる場合を考える。

$$I^A\cdot R=I^B \qquad \text{------------------------------------------------------------(8)}$$

ここで、行列 $R=\left[r_{ik}\right]$ は $m(B)\times m(A)$ 行列である。(8)式の関係における行列 $R$ とは、情報構造 $I^A$ に追加されるノイズを定めるものとして考えることができる[12]。具体的には、$R$ の成分 $r_{jk}$（$r_{jk}\geq 0$）は $m^A_j$ を $m^B_k$ として受け取る確率という意味を持つので、$\sum_j r_{jk}=1$ である。なお、状態 $s_i$ の下でメッセージ $m^B_k$ を受け取る条件付き確率（尤度）である $I^B$ の成分 $L^B_{ik}$ は、次のように示される。

---

9 それゆえ、ここでは、情報の需要は情報の価格に対して連続関数として示されない可能性もあることも指摘している。

10 ここでの説明内容は Blackwell's Theorem を利用したものとなっている。Blackwell's Theorem についての具体的な議論については、Laffont (1989) や Cremer (1982) などを参照せよ。

11 このような状況を示す用語として「ガーブル（garble）」という言葉が使われている。大ざっぱには、「誤って伝える」とか、「真実をゆがめる」などの意味があり、本稿での説明における「ノイズが追加される」状況と同様な状況において使われている。ここでは、情報構造 $I^B$ は $I^A$ をガーブリングしたものとして考える。

12 それゆえ、ノイズが追加された情報構造におけるメッセージは元も情報構造におけるメッセージの線形結合として示めされれるものとして考える。すなわち、一次変換されたものである。ここで、行列 $R$ は $I^A$ から $I^A$ への一次変換における対応の規則を定めるものである。

$$L_{ik}^{B} \equiv L^{B}(m_{k}^{B}|s_{i}) = \sum_{j} r_{jk} L^{A}(m_{j}^{A}|s_{i}) \qquad \text{-----------------------------------------(9)}$$

これらを用いると、簡単に情報構造$I^A$と$I^B$の価値についての比較を行うことができる。もし情報構造$I^A$が$I^B$より情報的である、すなわち情報性が高い、もしくは価値があるとすると、次の関係が成り立つことになる。

$$V^{A} = \sum_{j} \left\{ \sum_{i} U(a^{m(A)}, s_{i}) p^{A}(s_{i}|m_{j}) \right\} p^{A}(m_{j}) \geq$$

$$\sum_{k} \left\{ \sum_{i} U(a^{m(B)}, s_{i}) p^{B}(s_{i}|m_{k}) \right\} p^{B}(m_{k}) = V^{B}$$

$$\text{----------------------------(10)}$$

ここで、$U(\cdot)$は意思決定者の利得関数である。また$V^A$と$a^{m(A)}$は、メッセージ・サービス$I^A$を用いて意思決定をする場合の最大期待利得と最適行動をそれぞれ表し、$V^B$と$a^{m(B)}$は、メッセージ・サービス$I^B$を用いて意思決定をする場合の最大期待利得と最適行動をそれぞれ表す。

ところで、(10)式の右辺（第２行目）の$V^B$は、(9)式の定義と$\sum_{j} r_{jk} = 1$の関係、およびベイズ定理を用いると、次のように書き直すことができる。

$$V^{B} = \sum_{k} \left\{ \sum_{i} U^{B}(a^{m(B)}, s_{i}) \frac{L^{B}(m_{k}|s_{i})}{p^{B}(m_{k})} p(s_{i}) \right\} p^{B}(m_{k}) = \sum_{k} \left\{ \sum_{i} U^{B}(a^{m(B)}, s_{i}) \sum_{j} r_{jk} L^{A}(m_{j}|s_{i}) p(s_{i}) \right\}$$

$$= \sum_{j} \left\{ \sum_{i} U(a^{m(B)}, s_{i}) p^{A}(s_{i}|m_{j}) p^{A}(m_{j}) \right\}$$

$$\text{----------------------------------(11)}$$

しかし、(11)式で示される期待利得関数は、$V^A$の期待利得関数であり、定義により、このような期待利得関数においては、$a^{m(A)}$が最適行動になる。したがって、常に、次の関係が示される。

$$V^{A} \geq \sum_{j} \left\{ \sum_{i} U(a^{m(B)}, s_{i}) p^{A}(s_{i}|m_{j}) p^{A}(m_{j}) \right\} = V^{B}$$

これは、ノイズが追加された情報構造を利用する場合における最大期待利得は元の情報構造の最大期待利得を上回ることはないことを示しており、常に(10)式が満たされることを導くものである。この場合、情報構造$I^A$が$I^B$より情報的（informative）であるいい、任意の初期信念と利得関数の下で、常に、情報サービス$I^A$の情報的価値が$I^B$の情報的価値より高く評価される。

まとめおよびディスカッション

本稿では、情報を情報構造の確率論的枠組みの中で捉えて、情報や情報サービスを評価する議論を紹介した。具体的に、ここでは、複数の議論の単純化のための仮定の下で、情報や情報構造の精度が高いほど、これらの情報的価値は高くなることを示した。しかしながら、一般に、情報や情報サービスの価値は、情報や情報サービスの精度（precision）だけではなく、情報利用者の初期信念や選好などを含む、さまざまな要因によって大きく異なるものである。それゆえ、他の要因如何によっては、常に、情報や情報構造の精度と情報的価値は正の関係が成立するとは限らないものとなることにも注意を払う必要がある。

確率論的知識は強力な分析的ツールを提供する。本稿での議論においても、意思決定者がベイズ定理を用いてもっとも効率的に情報や情報サービスを評価し、効率的な意思決定に導く状況が議論の対象となっている。現実の経済においては、人々は、何者かから、様々な情報を受け取り、利用する。ここでは、本稿で想定している意思決定者とは異なり、人々は、常に、情報的意思決定に際し、何者から受け取った情報なのかという外部要因と複雑に絡み合う状況の下でどのように情報を正確に評価し利用するかという内部要因によって悩まされる。

情報源の信頼性の問題は、情報的意思決定に際しての外部要因に関する要因の一つである。友人からマスコミやインターネットに至るまでの種々の情報源に対する信頼性は個々人によって異なる。情報源間の競争をも考慮に入れる場合には[13]、個々人にとっての情報的環境はさらに複雑なものとならざるを得ない。一方、情報や情報サービスの評価に際しての心理的要因は内部要因に関する要因の一つである。この要因に対しては、近年の「行動経済学」の発展が示唆するものが大きいが、情報や情報サービスの評価に際して、認知的バイアス（cognitive bias）が存在する場合には[14]、合理的情報的行動は歪められたものにならざるを得ない。

このような要因について研究は、ある情報や情報サービスをノイズを追加することによって、他の情報や情報サービスとして示す単純な確率論的枠組みを大きく超える範囲の研究内容にならざるを得ない。今後、「情報の経済学」のフロンティアにおいては、情報自体よりは情報利用者の方向へ、議論が広がっていくことが期待される。

---

[13] 情報源の問題は情報を持つ専門家（expert）から情報を持たない非専門家（nonexpert）への情報伝達の問題として捉えることができる。このような問題についての幅広く要領のあるサーベイについては、Valsecchi (2013)を参照せよ。

[14] この場合についての議論の一つとしては、現実の経済主体の確率判断についての議論を挙げることができる。ここでは、被験者がサンプリング原理もしくはベイズ・ルールに反する確率判断を行うことを示す実験結果を提示し、これを導くヒューリスティック・メカニズム（heuristic mechanisms）が存在し得るという知見を示している。この分野についての広範囲にわたる、要領のあるサーベイとして、Camerer = Loewenstein（2004）(pp. 3–51)を参照せよ。

＜参考文献＞

Akerlof, G. A. and W. D. Dickens(1982), The Economic Consequences of Cognitive Dissonance, *American Economic Review*, Vol. 72, pp. 307-19.

Bierman, S. H. and L. L. Ferandez (1995), *Game Theory with Economic Application*, Addison-Wesley.

Birchler, U. and M. Butler (2007), *Information Economics*, Routledge.

Blackwell, D. (1953), Equivalent Comparisons of Experiments, *Annals of Mathematical Statistics*, Vol. 24, pp. 265 - 272.

Cremer, J. (1982), A Simple Proof of Blackwell's "Comparison of Experiments" Theorem, *Journal of Economic Theory*, Vol. 27, pp. 439 – 443.

Camerer. C. F. and G. Loewenstein (2004), Behavioral economics: Past, Present, Future, In C. F. Camerer, G. Loewenstein, & M. Rabin ( Eds.), *Advances in Behavioral Economics*, Princeton: Princeton University Press.

Green, J. (1981), Value of Information with Sequential Futures Markets, *Econometrica*, Vol. 49, pp.335 – 358.

Hirshleifer, J. and J. G. Riley (1992), *The Analytics of Uncertainty and Information*, Cambridge University Press.

Keynes, J. M. (1921), *A Treatise on Probability*, Macmillan.（佐藤隆三訳（2010）『ケインズ全集８巻　確率論』東洋経済新報社。）

.Keynes, J. M. (1936), *The General Theory of Employment, Interest and Money*, Macmillan. （間宮陽介訳（2008）『雇用、利子および貨幣の一般理論（上下 2 巻）』岩波文庫。）

Kihlstrom R. (1984), A Bayesian Exposition of Blackwell's Theorem on the Comparison of Experiments, in M. Boyer and R. Kihlstrom (eds.), *Bayesian Models in Economic Theory*, North-Holland.

Knight, F. (1921), *Risk, Uncertainty and Profit* , Univ., of Chicago Press.

Laffont, J. (1989), *The Economics of Uncertainty and Information*, The MIT Press.

Lawrence, D. B. (1999), *The Economic Value of Information*, Springer-Verlag New York,Inc.

Marschak, J. and R. Radnar (1972), *Economic Theory of Team*, Yale University Press.

Milgrom, P. and N. Stokey (1982), Information, Trade and Common Knowledge, *Journal of Economic Theory*, vol. 26, pp.17 – 27.

Philps, L. (1988), *The Economics of Imperfect Information*, Cambridge University Press.

Radner, R. and J. Stiglitz (1984), A Nonconcavity in the Value of Information, in M.

Boyer and R. Kihlstrom (eds.), *Bayesian Models in Economic Theory*, North-holland, pp.33 – 52.

Valsecchi, I. (2013), The Expert Problem：A Survey, *Economic Governance*, Vol. 14, pp. 303–331.

酒井 泰弘(2013) 「ケインズとナイト―不蓋然性と確実性を中心として」Discussion Paper No. J-36.

林周二（1973）『統計学講義第２版』丸善

小針晛宏（1973）『確率・統計入門』岩波書店（2000 年、第３０刷）

繁桝算男（1985）『ベイズ統計入門』東京大学出版会

宮沢光一（1971）『情報・決定理論序説』岩波書店