第1章　確率

本章では確率について述べる。しかし、確率に関する文献はそれこそ星の数ほどあり、大家による著書や良書といわれるものも数多出版されている。そこで、ここでは若干見方を変えて、確率のたどった歴史を簡単に追いながら、説明を加えていくことにしよう。

そうはいっても、基礎的な定義をなおざりにはできないので、それらについても必要に応じて触れていくことにする。

1.1　Laplaceの数学的確率

1.1.1　定義

この世で初めて確率を扱ったのは、1525年にGirolamo Cardanoによって記された「偶然のゲームの書」であるといわれている。Cardanoはその中で、カードやサイコロを用いた賭事に対する詳細な解析を行ったが、これを理論体系化するには、さらに100年近くの月日が必要であった。1812年にPierre-Simon Laplaceは「確率の解析的理論」を記したが、その中ではじめて確率は理論体系化される。Laplaceが「確率の解析的理論」の中で体系化した確率を、ここでは**数学的確率** (mathematical probability) ということにしよう。**算術的確率**という場合もあるが、同義である。

まず、硬貨を投げる、サイコロを振るといった操作を**試行** (trial) といい、試行の結果を**事象** (event) という。そして、起こりうる結果の集合を**標本空間** (sample space) という。また、標本空間においてこれ以上に分けられない事象を**根元事象** (elementary event) といい、複数の根元事象からなる事象を**結合事象** (joint event) という。

サイコロを例に取ると、前述のように「サイコロを振る操作」が試行、「1の目が出る」「偶数の目が出る」が事象となる。「1の目が出る」はこれ以上分けられないので根元事象であり、「偶数の目が出る」は「2の目が出る」「4の目が出る」「6の目が出る」の3つの根元事象からなる結合事象である。

以上の準備のもとに、数学的確率を定義しよう。

‘ 1.1 (数学的確率) 試行の結果の起こり方、すなわち事象の数をnとし、これらは「**全て同程度に確からしい**」とする。*1*つの事象Eにとって、「それが生じれば事象Eが生起するような」起こり方がr個あったとき、事象Eの生起する数学的確率$P(E)$を、

$$P(E) = \frac{r}{n}$$

と定義する。

この数学的確率は、順列・組合せの考え方が適用できるような事例に対して用いることができる。

1.1 サイコロを 2 つ振ったとき、生じ得るすべての組合せ、すなわち事象の総数は $n = 6 \times 6 = 36$ 通りである。また、「出た 2 つの目の和が 7 となる」という結合事象 E の生起の仕方は 6 通りある。よって、「出た 2 つの目の和が 7 となる」確率 $P(E)$ は $6/36 = 1/6$ で与えられる。

1.1.2　無差別の原理

さて、定義 1.4 では、「全て同程度に確からしい」という仮定を置いていることに注意してほしい。この意味ははっきりしているが、この仮定自体が確率の存在を前提としている。つまり、確率を定義するのに、確率を用いているという循環性の問題点がある。Laplace は、これについて以下のように説明している。

> 我々が無知であるがゆえに確率ということが問題になるのであり、同様に確からしいということは、それを判断する知識が欠けているということを意味する。例えば、硬貨を投げたとき、表が出るのと裏が出るのと、どちらがより確からしいかは全く分からない。そこで、表が出ることと裏が出ることは、同様に確からしい、とするのである。

すなわち、Laplace は、この問題点を克服するために、「**無差別の原理**」(principle of indifference) を用いた。「**理由不十分の原理**」とも言う。これは、「もし、ある事象がほかの事象よりも起こるべきであると考える十分な理由を持たないならば、それらの事象は同等の可能性を持つ」という原理であり、言い換えれば、「全て同様に確からしいことを明確に否定する十分な理由が無いので、全て同様に確からしいことにする」ということになる。

決定論者である Laplace は、確率を知識の欠如による主観的な理論と考えていた。よって、無差別の原理に基づいた数学的確率は、その循環性と主観的な立場ゆえに、厳しい批判にさらされることになる。しかし、現在多くの人が「確率」という用語からイメージするのも、この数学的確率である。

1.1.3　確率変数と確率分布

ところで、「コインの表が出る」「サイコロの 3 の目が出る」等の言葉ではなく、「コインの表が出る」の代わりに 1、「サイコロの 3 の目が出る」の代わりに 3 のように、各根元事象に対して数値を割り当てれば、扱いが容易になる。そこで、各根元事象に対して

割り当てられた数値のいずれかをとる変数 X を導入する。この X を**確率変数** (stochastic variable) といい、各根元事象に対して割り当てられた数値を X の**実現値**という。

1.2 コイン投げの場合は、「表が出る」に *1* を、「裏が出る」に *0* を割り当てることが多い。これを用いれば、「コインを投げたときに表が出る確率」という代わりに、「$X = 1$ である確率」ということができる。

1.3 サイコロ投げの場合は、しばしば出た目をそのまま確率変数と見なす。つまり、「*1* の目が出る」から「*6* の目が出る」をそのまま *1* から *6* の整数に割り当てる。これを用いれば、「サイコロを投げたときに *3* が出る確率」という代わりに、「$X = 3$ である確率」ということができる。

確率変数 X がサイコロの目のように飛び飛びの値をとるとき、X を**離散的確率変数** (discrete stochastic variable) といい、確率変数 X が電圧値や降雨量のように連続値をとるとき、X を**連続的確率変数** (continuous stochastic variable) という。これらの各確率変数 X に対して、数学的確率を付与することができる。以下、離散的な場合と連続的な場合で分けて考える。但し、基本的な考え方は変わらないことに注意しよう。

離散的な場合

離散的確率変数の場合、$X = x_i$ のときの数学的確率が p_i なら、その確率 $P(X)$ は関数の形で次のように表現できる。

‘ 1.2 (離散的確率変数の確率密度)

$$f(x) = \begin{cases} p_i & (x = x_i) \\ 0 & (\text{その他}) \end{cases}$$

このような $f(x)$ を**確率密度** *(probability density)* という。離散的な確率密度の場合を**確率関数** *(probability function)* ということもある。

確率密度の性質は次の 2 つである。まず、p_i は i 番目の実現値 x_i が起こる確率なので、

$$0 \leq f(x) \leq 1 \tag{1.1}$$

次に、確率の総和は 1 なので、

$$\sum_{i=1}^{n} f(x_i) = 1 \tag{1.2}$$

となる。X のとりうる値である実現値は根元事象に割り当てられているので、それぞれの事象は背反していることに注意しよう。

次に、確率変数 X のとる値が x 以下であるような確率 $F(x)$ について考えてみよう。

‘ **1.3 (離散的確率変数の確率分布関数)**

$$F(x) = P(X \leq x) \tag{1.3}$$

このような関数 $F(x)$ を**分布関数** *(distribution function)* という。

1.4 たとえば「*1* つのサイコロを投げたとき、出る目が *4* 以下である確率」は、

$$\begin{aligned} F(4) &= P(X \leq 4) \\ &= P(1) + P(2) + P(3) + P(4) \\ &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{2}{3} \end{aligned}$$

となる。

分布関数には次のような性質がある。まず、定義より、

$$F(x) = \sum_{x_i \leq x} f(x_i) \tag{1.4}$$

よって、$F(x)$ は非減少の階段関数となる。これから、次が導かれる。

$$F(-\infty) = \sum_{x_i \leq -\infty} f(x_i) = 0 \tag{1.5}$$

$$F(\infty) = \sum_{x_i \leq \infty} f(x_i) = 1 \tag{1.6}$$

$$F(a < x \leq b) = \sum_{a < x_i \leq b} f(x_i) = F(b) - F(a) \tag{1.7}$$

連続的な場合

離散的な場合と同様に、連続的確率変数の場合には

1.2　頻度確率

頻度確率 (frequency probability) とは、無差別の原理に対する批判的立場から生まれた考え方である。Laplace の「同様に確からしい」という立場に対して、Cramér-von Mises は「何をもって確からしいとするか」と疑問を投げかけ、「もっと客観的な立場から確率は論じられるべきだ」とした。Mises はまず、対象とする事象について、観測結果や理論的考察に基づいて、全試行に対するその事象の相対頻度を考えた。そして、試行回数を、実験上では大きくとったり、理論的考察上では無限大にすることによって、その極限として得られた相対頻度を、その事象の頻度確率としたのである。頻度確率は、客観的に得られた観測結果や理論をもとにしているので、**客観確率** (objective probability) ともいう。

‘ **1.4 (頻度確率)** 事象 E が生じ得る実験を n 回繰り返したとき、E が n_E 回生起したとし、$P_n(E) = \dfrac{n_E}{n}$ と置く。試行の回数 n を増やし、$n \to \infty$ のとき、事象 E の生起する頻度確率 $P(E)$ を、

$$P(E) = \lim_{n \to \infty} P_n(E) \tag{1.8}$$

と定義する。

この頻度確率は、Ronald Aylmer Fisher や Jerzy Neyman らによって支持された。しかし、この頻度確率にも以下の問題がある。

- 式 1.8 の右辺はいつも収束するのか？
- n をどのくらい大きくとればよいのか？

1.3 Kolmogorov の公理主義

A. N. Kolmogorov が 1933 年に記した「確率論の基礎概念」に述べられている。

根元事象全体からなる集合を標本空間といい、Ω で表す。標本空間 Ω の部分集合からなる族 F で次の条件を満たすものを、**Borel 集合体**という。

1. Ω は F に属する。
2. A が F に属すれば、A の補集合も F に属する。
3. A_k $(k = 1, 2, \ldots)$ が F に属するとき、それらの和集合も F に属する。

このとき、F 上で定義された実数値関数 $P(A)$ で、次の 3 条件を満たすものを確率という。

1. $P(A) \geq 0$
2. $P(\Omega) = 1$
3. F の要素 A_k $(k = 1, 2, \ldots)$ において、互いに排反のとき、A_k の和集合に対応する値は、$P(A_k)$ の和になる。

このようにして定義される (Ω, F, P) を確率空間といい、P を確率測度という。また、$P(A) \leq 1$、$P(\phi) = 0$、$P(A\text{ の補集合}) = 1 - P(A)$ も簡単に示すことができる。

Laplace の確率を、Kolmogorov の定義から見直す。いま、理想的なコイン投げについて考え、表が出れば 1、裏なら 0 とする。$\Omega = \{0, 1\}$、$F = \Omega$の部分集合全体、$P(A) = (A$ の要素の個数$)/2$ とおくと、(Ω, F, P) は確率空間をなす。

確率を定義するとき、必ず Ω、F、P を組として考える。従って、標本空間が同じであっても、確率空間が違えば、異なる確率となる。

ある集合から 1 つの要素を取り出すとき、どの要素がとられるか同程度に期待されない場合もある。その場合には、確率密度関数 $F(X)$ という実数値関数を考える。

確率密度関数 $F(X)$ は次の条件を満たすものとする。

1. $F(X) \geq 0$
2. 定義域における $F(X)$ の定積分 $= 1$

このとき、確率 $P(A)$ は A における $F(X)$ の定積分で定義される。

1.4　主観確率

「同程度の確からしさで生起」の仮定下での計算や、「生起回数の相対頻度」は、誰が計算しても同一。一方、人が主観的にある確率を与えて分析を行う場合がある。これらの確率は、人の持つ情報、知識、経験等で大きく異なる。これらの確率を**主観確率**という。例えば、ひいきの野球チームが次の試合で勝つ確率や、選挙の当落。

ベイズ推定、選好理論、効用分析理論、ファジィ理論が主観確率を有効に採り入れている。

では、なぜ確率が金科玉条のごとく扱われているのか？

1. 自然科学者に顕著に見られるが、すべての統計的推移には、「単純かつ美しい法則」が背景にあると信じられている。
2. その確率的法則は、試行を増やせば増やすほど、近似的に確認できる。
3. 逆に、理論的な確率値を前提とした時、観測値の妥当性が確認できる。

これを裏付けるのが、Bernoulli と Kolmogorov によって示された**大数の法則**である。

1.5　大数の法則

$\{X_n\}$ を確率変数の列とする。以下の仮定を設ける。

- X_n は全て同じ確率法則に従う。
- X_n は独立試行。
- 分散 $V(X_n)$ は有界。

$S_n = \sum_{k=1}^{n} X_k$ としたとき、

$$\lim_{n\to\infty} \frac{S_n - E(S_n)}{n} = 0$$

これを **Kolmogorov の大数の強法則**という。

$$\forall\varepsilon, \quad \lim_{n\to\infty} P\left(\frac{|S_n - E(S_n)|}{n} > \varepsilon\right) = 0$$

ともかける。

同一分布の場合には、

$$\frac{S_n}{n} \to \mu \text{ (平均値)}$$

これを **Bernoulli の大数の弱法則**という。

()　$Cebyšev$ の不等式：

$$\forall \alpha, \quad P\left(|X(\gamma) - E(\gamma)| > \alpha\right) = \frac{V(X)}{\alpha^2}$$

より、

$$\forall \alpha, \quad P\left(|S_n - E(S_n)| > \alpha\right) = \frac{V(S_n)}{\alpha^2}$$

$\{X_n\}$ は独立試行なので、

$$V(S_n) = \sum_{k=1}^{n} V(X_k) \leq n \max_k V(X_k)$$

よって、

$$\forall \alpha, \quad P\left(|S_n - E(S_n)| > \alpha\right) \leq \frac{n}{\alpha^2} \max_k V(X_k)$$

$\alpha = n\varepsilon$ と置くと、

$$\forall \varepsilon, \quad P\left(\frac{|S_n - E(S_n)|}{n} > \varepsilon\right) \leq \frac{1}{n\varepsilon^2} \max_k V(X_k)$$

$n \to \infty$ とすると、$V(X_k)$ の有界性より、右辺 $\to 0$ となり、

$$\forall \varepsilon, \quad \lim_{n\to\infty} P\left(\frac{|S_n - E(S_n)|}{n} > \varepsilon\right) = 0$$

が導かれる。

第2章　ベイズ推定

ベイズ理論は、18 世紀後半のスコットランドの長老派教会の牧師だった T. Bayes によって生み出された。20 世紀始めに、R. A. Fisher や J. Neyman によって激烈な避難にさらされ、一度は葬り去られたが、1954 年に、L. J. Savage が記した著書「The Foundation of Statistics」をきっかけとして復権する。Bayes の生み出した理論を基本とした推定をベイズ推定 (Bayes estimation) といい、現在は、Fisher や Neyman の創出した統計的推定よりも広く支持されている。

2.1　統計的推定

2.1 A と B のコインがある。投げたときに表と裏が出る割合の比は、A は $5:5$、B は $6:4$ である。*100* 回投げて *49* 回表が出るコインはどちらか？

簡単に考えると、*100* 回投げたら A は *50* 回、B は *60* 回表が出る。*49* 回という数字は *50* 回に近いので、このコインは A。

もう少し詳細な議論をする。A を *100* 回投げて *49* 回表が出る確率 P_A は

$$P_A =_{100} C_{49} \left(\frac{5}{10}\right)^{49} \left(\frac{5}{10}\right)^{51}$$

これを**二項分布** *(binomial distribution)* という。また、B を *100* 回投げて *49* 回表が出る確率 P_B は

$$P_B =_{100} C_{49} \left(\frac{6}{10}\right)^{49} \left(\frac{4}{10}\right)^{51}$$

以上より $P_A \gg P_B$ となり、A と判断できる。

つまり、最尤推定とは、

> 現実に生じた現象を説明するモデルが複数あるとき、各モデル毎の生起確率を計算し、それが最も大きくなるモデルを採用する。

という推定法である。特徴は、

- 推測は入らない。
- 多くの試行・例に基づいた判断となり、判断材料の乏しいケースには適用が困難である。
- 多くの試行が観測可能で、変化しない「静的」な対象のみに適用可能。

2.2 ベイズ推定の原理

ベイズ推定は、ベイズの逆確率 (invers probability) とも呼ばれ、以下のプロセスで行われる。

1. 適当な確率 (事前確率、prior probability) を付加する。
2. データに基づいて確率を更新する。

2.2 狼少年の物語をモデルしてみよう。

$$\begin{aligned}\Omega &= \{\text{ 正直}, \text{嘘つき }\} \equiv \{\theta_1, \theta_2\} \quad \text{(状態、ステイト)}\\ X &= \{\text{ 狼が来た}, \text{狼が来なかった }\} \equiv \{A, B\} \quad \text{(現象)}\end{aligned}$$

ここで、正直者なら、*10* のうち *9* は本当のことを言い、嘘つきなら、*10* のうち *8* は嘘を言うことが判っているとする。

最初はその少年が正直か嘘つきか判らないので、とりあえず、

$$\theta_1 : \theta_2 = 1 : 1$$

とする。

	θ_1	θ_2
A	$\frac{1}{2} \times \frac{9}{10}$	$\frac{1}{2} \times \frac{2}{10}$
B	$\frac{1}{2} \times \frac{1}{10}$	$\frac{1}{2} \times \frac{8}{10}$

実際に狼が来なかった場合、現象は B となるので、その後の少年の状態の比は、

$$\theta_1 : \theta_2 = \frac{1}{2} \times \frac{1}{10} : \frac{1}{2} \times \frac{8}{10} = 1 : 8$$

ここで、少年の状態の比は更新される。この比、またはこの比から得られる確率を**事後確率** *(posterior probability)* と言う。さらに少年が同じことを言うと、

	θ_1	θ_2
A	$\frac{1}{9} \times \frac{9}{10}$	$\frac{8}{9} \times \frac{2}{10}$
B	$\frac{1}{9} \times \frac{1}{10}$	$\frac{8}{9} \times \frac{8}{10}$

再度実際に狼が来なかった場合、現象は B となるので、その後の少年の状態の比は、

$$\theta_1 : \theta_2 = \frac{1}{9} \times \frac{1}{10} : \frac{8}{9} \times \frac{8}{10} = 1 : 64$$

また、少年の状態の比は更新される。

つまり、少年が嘘つきである確率は、

$$\frac{1}{2} \to \frac{8}{9} \to \frac{64}{65}$$

と変遷している。

2.3　ベイズの定理

ベイズ推定は、次のベイズの定理 (Bayes' Theorem) をベースとしている。
事象 $A_1, \ldots, A_n$ に対し、以下を仮定する。

1. $A_1 \cup \cdots \cup A_n = X$
2. $A_i \cap A_j = \phi \quad (i \neq j)$

A_i の生起確率を $P(A_i)$、A_i が生起したときの B の生起する条件付き確率を $P(B|A_i)$ とすると、B が生起したときの A_i の生起する条件付き確率を $P(A_i|B)$ は、

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\displaystyle\sum_{k=1}^{n} P(A_k)P(B|A_k)}$$

で与えられる。

()　条件付き確率の定義より、

$$P(A_i|B) = \frac{P(A_i \cap B}{P(B)} \tag{2.1}$$

同様に、

$$P(A_i \cap B) = P(A_i)P(B|A_i) \tag{2.2}$$

一方、

$$\begin{aligned} B &= B \cap X \\ &= (A_1 \cup \cdots \cup A_n) \cap B \\ &= (A_1 \cap B) \cup \cdots \cup (A_n \cap B) \end{aligned}$$

より、

$$\begin{aligned} P(B) &= P(A_1 \cap B) + \cdots + P(A_n \cap B) \\ &= \sum_{k=1}^{n} P(A_k)P(B|A_k) \end{aligned} \tag{2.3}$$

式 *2.2*、式 *2.3* を式 *2.1* に代入して、

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\displaystyle\sum_{k=1}^{n} P(A_k)P(B|A_k)}$$

を得る。

$P(A_i)$ は、A_i の生起の仕方に対してあらかじめ持っていた確率であり、**事前確率** (prior probability) と呼ばれる。それに対して、$P(A_i|B)$ は、B であることを知った後で、A_i の生起の仕方に対して持つことになった確率であり、**事後確率** (posterior probability) と呼ばれる。

2.3 (医学的意志決定) ある年齢の人が健康である *(この事象を* θ_1 *とする)* 確率を *0.995*、ガンである *(この事象を* θ_2 *とする)* 確率を *0.005* とする。また、ガンの陽性 *(この事象を* A *とする)* を陰性と誤診する確率、陰性 *(この事象を* B *とする)* を陽性と誤診する確率を共に *0.05* とすると、

表 *2.1:* ガンの罹患に関する確率

	θ_1	θ_2
A	0.995×0.05	0.005×0.95
B	0.995×0.95	0.005×0.05

検査の結果、もしも陽性と出たら、表 *2.1* の上の行に対応するので、健康である確率 $P(\theta_1 \mid A)$ は、

$$P(\theta_1 \mid A) = \frac{0.995 \times 0.05}{0.995 \times 0.05 + 0.005 \times 0.95} = \frac{199}{218} \approx 0.91$$

0.91 という値をどう見るかは人によって異なる。例えば、

- もともと検査の信頼性は 95% なので、ガン罹患の可能性が 9% ということはあまり高いとは言えない。
- 通常のガン罹患の可能性は 0.5% なので、それと比較すると 9% という値は *19* 倍で、非常に高いと言える。

2.4　ベイズ推定の長所

1. フィッシャーの頻度主義とは異なり、多数のデータを収集しなくても推定可能である。また、データを多くするにつれて大数の法則が働き、初期値に関係なく、理想値に近づく。但し、初期値に関係ないと入っても、0 : 1 のようなオッズでは意味がない。
2. 確率の更新は、それまでに収集した全てのデータをその都度呼び出す必要はなく、直前のデータからだけで計算可能である (逐次合理性)。

例えば、コインを 2 回投げて真偽の判定を行う。本物なら { 表, 裏 } が {5 : 5}、偽物なら { 表, 裏 } が {6 : 4} であるとし、2 回投げたところ、2 回とも表が出たとしよう。

頻度主義に基づいた計算 コインが真であり、かつ 2 回とも表が出る確率 $P(\text{真},\text{表},\text{表})$ と、コインが偽であり、かつ 2 回とも表が出る確率 $P(\text{偽},\text{表},\text{表})$ は次のようになる。

$$
\begin{aligned}
P(\text{真},\text{表},\text{表}) &= \frac{1}{2}\times\frac{5}{10}\times\frac{5}{10}=\frac{25}{200}\\
P(\text{偽},\text{表},\text{表}) &= \frac{1}{2}\times\frac{6}{10}\times\frac{6}{10}=\frac{36}{200}
\end{aligned}
$$

よって、$\{\text{真},\text{偽}\}=\{25:36\}$ となる。

ベイズ推定に基づいた計算 逐次更新で考えてみる。

1 回目 $\{\text{真},\text{偽}\}=\{1:1\}$ とする。

表 2.2: コインの真偽判定：1 回目

	真	偽
表	$\frac{1}{2}\times\frac{5}{10}$	$\frac{1}{2}\times\frac{6}{10}$
裏	$\frac{1}{2}\times\frac{5}{10}$	$\frac{1}{2}\times\frac{4}{10}$

1 回目で表が出たので、表 2.2 の上の行より、

$$
\begin{aligned}
P(\text{真},\text{表}) &= \frac{1}{2}\times\frac{5}{10}=\frac{5}{20}\\
P(\text{偽},\text{表}) &= \frac{1}{2}\times\frac{6}{10}=\frac{6}{20}
\end{aligned}
$$

よって、$\{\text{真},\text{偽}\}=\{5:6\}$ となる。

2 回目 1 回目が終わった時点で $\{\text{真},\text{偽}\}=\{5:6\}$ となっている。

表 2.3: コインの真偽判定：2 回目

	真	偽
表	$\frac{5}{11}\times\frac{5}{10}$	$\frac{6}{11}\times\frac{6}{10}$
裏	$\frac{5}{11}\times\frac{5}{10}$	$\frac{6}{11}\times\frac{4}{10}$

2 回目でも表が出たので、表 2.3 の上の行より、

$$
\begin{aligned}
P(\text{真},\text{表}) &= \frac{5}{11}\times\frac{5}{10}=\frac{25}{110}\\
P(\text{偽},\text{表}) &= \frac{6}{11}\times\frac{6}{10}=\frac{36}{110}
\end{aligned}
$$

よって、$\{\text{真},\text{偽}\}=\{25:36\}$ となる。

2.5 信憑性の獲得

ベイズ推定の結果として、対象の性質 (例えばコインの真偽) が「確実に真実である」、すなわち「100% 近く間違いない」と認識されたとき、それを「**信憑性** (credibility) を獲得した」、「**評価** (reputation) が確立した」と言う。では、「100% 近く」にはどのようなオーダーがあるのか？この逆の希少現象に関して、エミール・ボレルが次のような分類を行っている。

$\frac{1}{10^6}$：人間的尺度において無視できる割合

例えば、国内で交通事故は 1 件/秒 の割合で起きている。また、1 分間に事故に遭う確率は $\frac{1}{2.7\times10^6}$ である。

$\frac{1}{10^{15}}$：地球的尺度において無視できる割合

遺伝的疾患を持つ確率がこのオーダーに相当する。

$\frac{1}{10^{50}}$：宇宙的尺度において無視できる割合

観測可能な全ての恒星は 10^9 個であり、これら全てについて、数世紀に渡って観測できた全ての物理量は多く見積もって 10^{20} 個である。全ての物理法則はこの中でのみ観測可能である。

$\frac{1}{10^{90}}$：超宇宙的尺度において無視できる割合

1 リットル中に酸素分子と窒素分子が同量存在している場合、これらの分子が完全に 2 つに分離する確率は $\frac{1}{10^{10^{22}}}$ である。この分母を数字で記した場合、必要な 0 の個数は、1 ページ 10000 文字、1 巻 1000 ページ、1 揃 100 巻が 1 つの図書館に 100 揃あったとして、その図書館 100 棟でやっと 1 億分の 1 にしかならない程の量であり、人間の想像の限界を超えている。

第3章　期待効用理論

期待効用理論 (expected utility theory) は、von Neumann と Morgenstern によって 1947 年に提唱された。

3.1　期待値基準

不確実性を伴う現象への行動の指針としては、一般に、確率論をもとにした期待値が用いられる。これを**期待値基準**と言う。

クジの購入を例に考えてみる。

ケース 1　クジ A は確率 $\frac{1}{2}$ で 2 万円、クジ B は確率 $\frac{1}{2}$ で 3 万円が手に入る。この場合、多くの人は賞金額の多少でクジ B を選択する。

ケース 2　クジ A は確率 $\frac{2}{3}$ で 2 万円、クジ B は確率 $\frac{1}{3}$ で 2 万円が手に入る。この場合、多くの人は確率の大小でクジ A を選択する。

ケース 3　クジ A は確率 $\frac{1}{2}$ で 4 万円、確率 $\frac{1}{2}$ で 2 万円が手に入り、クジ B は確率 $\frac{1}{4}$ で 8 万円、確率 $\frac{3}{4}$ で 1 万円が手に入る。この場合、多くの人は期待値の大小でクジ A を選択する。すなわち、A の期待値は $\frac{1}{2} \times 4$ 万円 $+ \frac{1}{2} \times 2$ 万円 $= 3$ 万円、B の期待値は $\frac{1}{4} \times 8$ 万円 $+ \frac{3}{4} \times 1$ 万円 $= 2.75$ 万円 で、A の期待値の方が大きい。

期待値基準の背景には大数の法則がある。

3.2　決定基準

しかし、ギャンブルでは胴元が存在し、手数料や利益を得ているので、どんな場合でも期待値は賭け金より小さくなる。また、保険会社は掛け金から利潤を得ているので、保険に加入することは期待値基準に背く。では、なぜ人々はギャンブルに手を出したり、保険に加入したりするのか？それは、期待値基準以外の別の基準が働いているからと考えられる。

いま、行動 a_i で m 個の状態 s_j $(1, \ldots, m)$ が想定され、各状態で利得 x_{ij} が得られるとする。

期待値基準 (Laplace の基準) 期待値基準は Laplace の基準とも言う。期待値基準に基づいた場合、人は

$$\arg_i \max_i \sum_{j=1}^{n} \frac{x_{ij}}{n}$$

なる行動 a_i を取る。

Ward の基準 (マックスミン利得基準) Ward の基準はマックスミン利得基準とも言う。Ward の基準に基づいた場合、人は

$$\arg_i \max_i \min_{j=1\sim n} x_{ij}$$

なる行動 a_i を取る。これは、最小利得の最大化であり、最悪の状況の中で最善を取る。

Hurwiz の基準 Hurwiz の基準に基づいた場合、人は

$$\arg_i \max_i \{\alpha \max_{j=1\sim n} x_{ij} + (1-\alpha) \min_{j=1\sim n} x_{ij}\} \quad (0 < \alpha < 1)$$

なる行動 a_i を取る。ここでは、最善が実現したときと最悪が実現したときの割合を α で与えている。

Savage の基準 まず、

$$r_{ij} = \max_{j=1\sim n} x_{ij} - x_{ij}$$

とおく。この r_{ij} は機会損失といい、損失、後悔の量を表す。Savage の基準に基づいた場合、人は

$$\arg_i \min_i \max_{j=1\sim n} r_{ij}$$

すなわち、最大の後悔を最小化するような行動 a_i を取る。

3.3 聖ペテルブルクのパラドクス

これは、Bernoulli により示されたパラドクスである。コイン投げ：

$$\begin{cases} 1\text{回目に初めて表} & \cdots \quad 2\text{円の賞金} \\ 2\text{回目に初めて表} & \cdots \quad 4\text{円の賞金} \\ & \vdots \\ \text{n回目に初めて表} & \cdots \quad 2^n\text{円の賞金} \end{cases}$$

を考える。プレーヤが賞金を得たらゲームは終了となる。このゲームを平等にするためには、最初にいくら胴元に払えばよいか? もしも期待値基準で行動するなら、このゲームの期待値は

$$\sum_{n=1}^{\infty} \frac{1}{2^n} \cdot 2^n = \infty$$

なので、いくら払っても胴元不利のゲームと言うことになる。しかし、これは直感に反する。そこで、 Bernoulli は、

> 人は金額の期待値をもとにして行動しているのではなく、金額に対応する効用をもとにして行動している。

と考えた。

1. では、金額に対応する効用はどのように決めるのか？
2. また、なぜこのような考え方が人の行動の規範になるのか？

これについて、von Neumann と Morgenstern が 1947 年に期待効用仮説という 1 つの答えを示した。

3.4　選好順序

まず、準備として、順序について定義しよう。

‘ **3.1 (順序)** 集合 X に属する $x,\ y$ に対して、

(1) $x \leq x$　*(反射律)*

(2) $x \leq y,\ y \leq x \Rightarrow x = y$　*(反対称律)*

(3) $x \leq y,\ y \leq z \Rightarrow x \leq y$　*(推移律)*

が成り立つとき、「$\leq$」を**順序**といい、「$\leq$」が存在する X を**半順序集合**という。さらに、任意の $x,\ y \in X$ に対して

(4) $x \leq y$ または $y \leq x$

が成り立つとき、X を**全順序集合**という。

以上の準備のもとに、選好順序について定義しよう。

‘ **3.2 (選好順序)** 集合 Y に属する任意の A、B において：

(1) $A \preceq A$　*(少なくとも同等か良い、無差別)*

(2) $A \preceq B,\ B \preceq C \Rightarrow A \preceq C$　*(選好半順序)*

(3) $A \preceq B$ または $B \preceq A$　$(\forall A, B)$　*(選好全順序)*

(4) $\{A \mid A \preceq B\}$, $\{A \mid A \succeq B\}$ は閉集合　*(連続性)*

が成り立つとき、「$\preceq$」を選好順序という。

選好順序は、反対称律「$A \preceq B,\ B \preceq A \Rightarrow A = B$」がないので、順序関係より緩い関係となる。

3.4.1　期待効用仮説

A、B が確率的に得られるとし、$A \succ B$ とする。この時、$A \succ B$ だけでは、どちらを選ぶかは決められない。

例えば、$A = 100$ 万円、$B = 99$ 万円 と、$A = 100$ 万円、$B = 20$ 万円 とでは、「$\succ$」の意味が異なり、A と B に「どの程度の価値があるか」が重要となる。

その価値を実数で表したものを、$U(A)$、$U(B)$ と書き、「A の効用」「A の基数表現」という。U は効用関数と呼ばれ、「表現した」、または「A の基数表現」とは、

$$U(A) \geq U(B) \Leftrightarrow A \succeq B$$

の対応がつくことを言う。これは、「A を B より好むことの表現として、A の効用が B より大きいとする」という意味となる。

効用関数：

$$A \succ B \Leftrightarrow U(A) > U(B)$$
$$A \prec B \Leftrightarrow U(A) < U(B)$$

を表現する $U(\cdot)$ は一意ではないが必ず存在する。

選び得る方策 A について生じる結果 x_j と x_j を得る確率 p_j、すなわち、

$$A = [x_1, x_2, \ldots, x_n; p_1, p_2, \ldots, p_n]$$

が与えられているとき、von Neumann と Morgenstern は、その方策の効用関数は

$$U(A) = \sum p_j U(x_j)$$

で表すことができ、人間はこの U を最大化するように選好を行っているとした。これを**期待効用仮説**という。

ここで、前述の聖ペテルブルクのパラドクスを、効用関数を用いて考えてみよう。金額 x に対応する効用関数を

$$U(x) = k \log x \quad (k > 0)$$

とすると、

$$\sum_{n=1}^{\infty} \frac{1}{2^n} \cdot U(2^n) = \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot k \log 2^n = k \log 4$$

となり、収束する。これで聖ペテルブルクのパラドクスについては、出口が見えたかに思えるが、実は、アレイのパラドクス (Allais's Paradox) という反例がある。例えば、次のようなクジを考えよう。

$$\begin{aligned} a_1 &= [10000, 0; 0.1, 0.9] \\ a_2 &= [15000, 0; 0.09, 0.91] \\ a_3 &= [10000, 0; 1.0, 0.0] \\ a_4 &= [15000, 0; 0.9, 0.1] \end{aligned}$$

このクジに関しては、多くの人が、

$$
\begin{aligned}
a_1 &\prec a_2 \\
a_3 &\succ a_4
\end{aligned}
$$

とする。ここで、

$$a_0 = [*, 0; 0, 1.0]$$

を導入して上を書き換えると、

$$
\begin{aligned}
a_1 &= 0.1a_3 + 0.9a_0 \\
a_2 &= 0.1a_4 + 0.9a_0
\end{aligned}
$$

ここで、

$$k[a_1, a_2; p_1, p_2] = [a_1, a_2; kp_1, kp_2]$$
$$[a_1, a_2; p_1, p_2] + [a_1, a_2; q_3, q_4] = [a_1, a_2; p_1 + q_1, p_2 + q_2]$$

とする。よって、

$$a_3 \succ a_4 \Rightarrow a_1 \succ a_2$$

が導かれ、矛盾する。

3.5　リスク回避形

起こり得る全ての結果 x_j について $p_j \neq 1$ であるような任意のクジ a (これを**非退化クジ**という) において、意志決定者 (decision maker, DM) が、クジよりもその期待結果 $\bar{x}$：

$$\bar{x} = \sum_{j=1}^{n} p_j x_j$$

を好み、

$$U(\bar{x}) > U(a) = \sum_{j=1}^{n} U(x_j)$$

となるとき、DM の行為を**リスク回避形**、

$$U(\bar{x}) = U(a)$$

を**リスク中立形**、

$$U(\bar{x}) < U(a)$$

を**リスク志向形**という。U は、リスク回避形では上に凸、リスク中立形では線形、リスク志向形では下に凸となる。

第4章　不確実性理論

4.1　エルスバーグのパラドクス

不確実性理論は、期待効用仮説への反論として、Frank H. Knight によって提唱された。彼が示したエルスバーグのパラドクス (Ellsberg's paradox) とは次のようなものである。

ケース 1 赤玉と黒玉が合わせて 100 個入った 2 つのツボ A と B がある。ツボ I には赤玉と黒玉が各 50 個づつ入っているが、ツボ II の赤玉と黒玉の個数はわからない。赤玉が出たら賞金を手に出来る。どちらのツボを選ぶか？

ケース 2 赤玉 30 個と、青玉と緑玉が合わせて 60 個の計 90 個の玉の入った 2 つのツボ C と D がある。赤玉が出たら賞金を手に出来る場合 I と青玉が出たら賞金を手に出来る場合 II、どちらの場合が好ましいか？

ケース 3 やはりツボ C と D に関して、赤玉または青玉が出たら賞金を手に出来る場合 III と青玉または緑玉が出たら賞金を手に出来る場合 IV、どちらの場合が好ましいか？

どのケースに関しても、主観確率は同じなので、期待効用仮説に基づくと、どちらのツボ、どちらの場合も同じはずである。しかし、多くの人は、ケース 1 ではツボ A、ケース 2 では場合 I、ケース 3 では場合 IV を選ぶ傾向がある。ではなぜこのような傾向が出てくるのか？

Knight は、「確率のわかっている不確実性とわかっていない不確実性は分けなければならない。」と主張した。

4.2　Knight の定義

Knight は、リスクを、確率計算可能な不確実性、不確実性を、確率計算不能な不確実性と定義している。彼によれば、

- 確率がわかっている不確実性はもはや不確実ではない。
- 同一の独立試行の反復による大数の法則を背景にした確率論は、現実社会の不確実性を描写していない。
- 人は、確率のわからない不確実性を避ける傾向にある。これを**不確実性回避**という。

4.3　非加法的確率とマルチプル・プレーヤ

前出のエルスバーグのツボ A と B をモデル化してみよう。

ツボ A 標本空間は $\Omega = \{R, B\}$。ただし、R と B はそれぞれ赤玉と黒玉を表す。また、それぞれの事象のオッズは $\{1:1\}$ となる。

ツボ B 標本空間は $\Omega = \{R, B\}$。それぞれの事象のオッズは不明なので、$\{? :?\}$。

4.3.1　非加法的確率

とりあえず、ツボ B のオッズを $\{1:1\}$ とする。すると、ツボ A で R と B が出る確率は、$P(R) = P(B) = \frac{1}{2}$、しかし、ツボ B で R と B が出る確率は不明なので、$F(R) = F(B) =?$。A $\succ$ B、すなわち、多くの人はツボ A の方が色を当てやすいと思っているので、

$$\begin{cases} P(R) > F(R) \\ P(B) > F(B) \end{cases}$$

両辺をそれぞれ加えると、

$$1 > F(R) + F(B)$$

確率における加法性が破綻している。そこで、非加法的確率を導入する。

‘ 4.1 (非加法的確率) *X を状態空間、F を事象の集まり、(X, F) を可測空間とする。このとき、$\theta : F \to [0, 1]$ が次の 3 つを満たせば、非加法的確率測度と呼ばれる。*

1. $\theta(\phi) = 0$
2. $\theta(X) = 1$
3. $A \subset B \Rightarrow \theta(A) \leq \theta(B) \quad (\forall A, B \in F)$

4.3.2　マルチプル・プレーヤ

オッズがわからないので、考え得る限りのオッズを付加してみる。ツボ B は、赤玉と黒玉合わせて 100 個なので、(赤玉の個数, 黒玉の個数) は $(0, 100)$ から $(100, 0)$ までの 101 通り存在する。すなわち、オッズ $\{P(R) : P(B)\}$ も、$S_1 = \{0 : 100\}$ から $S_{101} = \{100 : 0\}$ の 101 通り存在する。これは、ツボを選ぶ人の心の中の違いを表していると考えられる。これをマルチプル・プレーヤという。

全ての S_i の赤玉が出る期待値 $a_1 \sim a_{101}$ を計算すると、

$$a_i = 1 \frac{i-1}{100} + 0 \frac{101-i}{100} = \frac{i-1}{100}$$

同様に、黒玉が出る期待値 b_i は、

$$b_i = 0\frac{i-1}{100} + 1\frac{101-i}{100} = \frac{101-i}{100}$$

このマルチプル・プレーヤの各期待値 a_i, b_i の最小値 $\min_i a_i, \min_i b_i$ を**マルチプル期待値**という。それぞれのマルチプル期待値を計算すると、

$$\min_i a_i = \min_i b_i = 0$$

一方、ツボ A では、$a_i = b_i = 0.5$ なので、マルチプル期待値も 0.5 となる。それぞれのツボのマルチプル期待値のうち、大きい方はツボ A なので、人はツボ A を選ぶ傾向にあると考えられる。

すなわち、人は、確率のわからない状態では複数のマルチ・プレーヤを持ち、マルチプル期待値の大くなるような行動を選択する。これは、マックスミン原理であり、最悪を最小にする行動選択と考えられる。

では、なぜ人は迷うのか？これは、人は意志と反対のことを過大に見積もる傾向があるからと言える。

4.4　不確実性回避

エルスバーグのパラドクスの示唆しているところは、「人は不確実性をリスクより嫌う」、すなわち、不確実性回避である。

しかし、確率がわかっていても、生起する事象を自ら選べるわけではないし、どの事象が実際に生起するかは相変わらずわからない。それにも関わらず、ぜ不確実性回避の傾向があるのか？

意志の問題 人は自分の運命を自分自身で決めたいと考えている。例えば、3 本のうち 1 本が当たりであるようなクジがあり、3 人が順番にクジを引くとき、何番目で引こうが当たりとなる確率は同じであるにもかかわらず、多くの人は最初にクジを引こうとする。

情報量の問題 たとえ運命に身を委ねるとしても、何らかの情報を持った上で、すなわち運命の正体を知った上で委ねたい。つまり、例え最悪な事態が生起しても、もっとも小さな傷になるようにしたいというマックスミン原理が働いている。

この考え方に基づくと、「自分の意志で自分の運命を決めることへの自信のなさの表れ」が、不確実性回避の根底にあるともいえる。

例えば、前述のツボ B に関して、自分が決めた色が出たら賞金が貰えるという賭けの場合、赤玉が出ると思った瞬間に、「黒玉が多いかもしれない」と考え、逆に、黒玉が出ると思った瞬間に、「赤玉が多いかもしれない」と考える。賭けると決めた方の可能性が賭けなかった方の可能性より低く見積もられてしまう。

これに関しても、非加法的確率を用いた説明がある。「赤玉を選択」という事象を R、「黒玉を選択」という事象を B、「赤玉を選択しない」という事象を $\bar{NR}$、「黒を選択しない」という事象を $\bar{NB}$ とする。$\Omega = \{R, B\}$ なので、R の余事象は B となり、通常の確率では $P(B) = P(NR)$ が成り立つが、この場合は非加法的確率なので、必ずしもそうではない。今、

$$F(R) = 0.4$$
$$F(B) = 0.4$$

と考えているとする。非加法的確率なので、$F(R) + F(B) \neq 1$ でよい。すると、

$$F(\bar{NR}) = 0.6 > F(B)$$
$$F(\bar{NB}) = 0.6 > F(R)$$

となり、余事象より、「選択しない」方が高い確率となる。

実は、非加法的確率とマルチプル・プレーヤのマックスミン原理は等価であることがわかっている [1]。

第5章　コモン・ナレッジ

コモン・ナレッジ (common knowledge) は、R. Aumann が 1976 年に定式化した理論である [2]。

コモン・ナレッジによれば、

$$\text{「}A\text{ が }X\text{ を知っている」}\cap\text{「}B\text{ が }X\text{ を知っている」}$$
$$\neq\text{「}X\text{ は }A\text{ と }B\text{ の共有の知識である。」}$$

例えば、サッカーにおけるオフサイド・トラップ (オフサイド・ラインを故意に引き上げ、相手選手を反則に誘い込む戦略) のように、オフサイド・トラップを実行することがディフェンダーの共有の知識になっていないと成立しない。

また、e-mail を使った場所の待ち合わせも同じ。e-mail では「自分が知っていることを相手も知っている」ことの確認は出来ない。

「X が A と B との共有の知識である」とは、「A が X を知っている」、「B が X を知っている」だけではだめで、「「A が X を知っている」ことを B が知っている」、「「B が X を知っている」ことを A が知っている」、更に、「「「A が X を知っている」ことを B が知っている」ことを A が知っている」、「「「B が X を知っている」ことを A が知っている」ことを B が知っている」、更に…と、これが無限に続いて初めてそうなる。この共有の知識を**コモン・ナレッジ** (common knowledge) という。

5.1　定式化

標本空間 $\Omega = \{a_1, \ldots, a_n\}$ とし、A の情報分割を $P_A = \{G_1, \ldots, G_m\}$ とする。但し、

$$G_1 \cup \cdots \cup G_m = \Omega$$
$$G_i \cap G_j = \phi \quad (i \neq j)$$

である。情報分割は A の知識の状態を表しており、「A にとって G_i に含まれる a は区別できない」ことを意味する。

情報 $X \subset \Omega$ に対して、「A が X を知っているために生起する必要がある事象」を $K_A(X)$ と表すとき、

$$K_A(X) = \{\bigcup_k G_k \mid G_k \subset X\}$$

となり、$K_A(X)$ は「A が X を知っている」ことと等価となる。

5.1 いつも A に電話をする人の集合を $\Omega = \{a_1, a_2, a_3, a_4, a_5\}$ とする。a_1 と a_2、a_3 と a_4 は同姓とする。すると、A の情報分割は

$$P_A = \{(a_1, a_2), (a_3, a_4), a_5\} = \{G_1, G_2, G_3\}$$

となる。今、誰かが A に電話をかけてきたが、A はその場におらず、他人が取り次いだ。ここで、$X =$ 電話をしてきたのは同窓生である とし、同窓生が $a_1,\ a_2,\ a_3$ であったとすると、

$$X = \{a_1, a_2, a_3\}$$

であり、A が「電話をしてきたのは同窓生である」、つまり X を知っているために生起する必要がある事象は、

$$K_A(X) = \{a_1, a_2\}$$

である。なぜなら、a_3 が生起する、すなわち a_3 が電話をかけてくると、a_4 と区別が付かない。a_4 は同窓生ではないので、a_3 を $K_A(X)$ に含めることは出来ない。

5.2　共有されない情報

これは Littlewood によって示された。いま、A と B がおり、A は B が単位を落とすかどうか知っているが自分のことはわからない、B は A が単位を落とすかどうか知っているが自分のことはわからない、とする。標本空間を $\Omega = \{0, a, b, z\}$ とする。0 はどちらも合格、a は A が単位を落とす、b は B が単位を落とす、z は A と B2 人ともに単位を落とす、とする。このとき、A と B の情報分割はそれぞれ

$$\begin{aligned} P_A &= \{(0, a), (b, z)\} \\ P_B &= \{(0, b), (a, z)\} \end{aligned}$$

となる。対称とする情報 X を「少なくともどちらかは単位を落とす」とすると、

$$X = \{a, b, z\}$$

となる。X は A と B のコモン・ナレッジになるか？

X がコモン・ナレッジになるためには、双方が X を知っており、そのことをやはり双方がしっており、そのことをやはり双方が知っており…という重層構造が生じなければならない。

A も B も単位を落とすことになったとする。A は「B が単位を落とす」ことは知っている。しかし、もし「A は「B が単位を落とす」ことを知っている」と B が知れば、すなわち B は自分自身が単位を落とすことを知ることになり、仮定に矛盾する。それを定式化してみよう。

$$\begin{aligned}
\Omega &= \{0, a, b, z\} \\
X &= \{a, b, z\} \quad (\text{少なくとも 1 人は落第}) \\
P_A &= \{(0, a), (b, z)\} \\
P_B &= \{(0, b), (a, z)\}
\end{aligned}$$

のとき、

$$\begin{aligned}
K_A(X) &= \{\bigcup_k G_k^A \mid G_k^A \subset X\} = \{b, z\} \quad (A \text{ が } X \text{ を知っている}) \\
K_B(X) &= \{\bigcup_k G_k^B \mid G_k^B \subset X\} = \{a, z\} \quad (B \text{ が } X \text{ を知っている})
\end{aligned}$$

更に同様にして、

$$\begin{aligned}
K_A(K_B(X)) &= \phi \quad (A \text{ が「} B \text{ が } X \text{ を知っている」ことを知っている}) \\
K_B(K_A(X)) &= \phi \quad (B \text{ が「} A \text{ が } X \text{ を知っている」ことを知っている})
\end{aligned}$$

よって、両者のコモン・ナレッジは ϕ、すなわち存在しないことになる。

ここで、大学から全員に、「A と B の少なくとも 1 人は落第」という情報 (**公的情報**という) がアナウンスされたとしよう。このとき、相手が単位を落としてなければ、自分が落としたということが判るので、情報分割は次のように変化する：

$$\begin{aligned}
P_A &= \{(0), (a), (b, z)\} \\
P_B &= \{(0), (b), (a, z)\}
\end{aligned}$$

よって、

$$\begin{aligned}
K_A(X) &= \{\bigcup_k G_k^A \mid G_k^A \subset X\} = \{a, b, z\} \quad (A \text{ が } X \text{ を知っている}) \\
K_B(X) &= \{\bigcup_k G_k^B \mid G_k^B \subset X\} = \{a, b, z\} \quad (B \text{ が } X \text{ を知っている})
\end{aligned}$$

更に同様にして、

$$\begin{aligned}
K_A(K_B(X)) &= \{a, b, z\} \quad (A \text{ が「} B \text{ が } X \text{ を知っている」ことを知っている}) \\
K_B(K_A(X)) &= \{a, b, z\} \quad (B \text{ が「} A \text{ が } X \text{ を知っている」ことを知っている})
\end{aligned}$$

よって、a, b, z のどれが生起しても、X は A と B とのコモン・ナレッジになる、ということが判る。

5.3　大恐慌の数理モデル

大恐慌とは、1929 年から 1993 年にかけて世界中の資本主義諸国を襲った、史上最大規模の世界恐慌のことである。1929 年 10 月 24 日 (木曜日) にニューヨーク、ウォール街の株式市場で起こった株価の大暴落に端を発し、全資本主義諸国に波及した。木曜日に起きたので、「暗黒の木曜日」と呼ばれている。コモン・ナレッジを用いた大恐慌の数理モデルが、1997 年に Hart, Tauman によって提案された。

大恐慌は、ある日、なんの前触れもなく起こった。つまり、市場の投資家達が一斉に持ち株の売りに出た結果、買い手が付かなくなり、株価の大暴落を引き起こしたのだが、これは偶然だろうか？

まず、

$$\Omega = \{\tilde{a}, b, c, d, \tilde{e}, f, g, h, \tilde{i}\}$$

とし、以下の仮定を設ける。

- 各ステイトは様々な経済状況を表す。また、$\tilde{a}$、$\tilde{e}$、$\tilde{i}$ は、生起していれば景気後退となるような経済状況である。
- 各ステイトの生起確率は等しい。
- 市場は A と B の 2 人のみで、情報分割は次の通り：

$$\begin{aligned} P_A &= \{(\tilde{a}, b, c), (d, \tilde{e}, f), (g, h, \tilde{i})\} \\ P_B &= \{(\tilde{a}, b, c, d), (\tilde{e}, f, g, h), (\tilde{i})\} \end{aligned}$$

  この情報分割はお互いに公開されている。

- A と B の売買戦略は共に、「景気後退の可能性が 1/3 以上になったら売り、以下だったら買い」。

では、日を追って見ていくことにする。

1 日目の市場

開始前 現実に生起している経済状況は $\tilde{a}$。「A は 1 日目に売り」を Z とし、「B は 1 日目に買い」を W とする。W が生じるために生起すべき事象は

$$W = \{\tilde{a}, b, c, d, \tilde{e}, f, g, h\}$$

A が W を知っているために生起すべき事象 $K_A(W)$ は

$$K_A(W) = \{\tilde{a}, b, c, d, \tilde{e}, f\}$$

以下、

$$K_B(K_A(W)) = \{\tilde{a}, b, c, d\}$$
$$K_A(K_B(K_A(W))) = \{\tilde{a}, b, c\}$$
$$K_B(K_A(K_B(K_A(W)))) = \phi$$
$$K_A(K_B(K_A(K_B(K_A(W))))) = \phi$$

また、Z が生じるために生起すべき事象は

$$Z = \{\tilde{a}, b, c, d, \tilde{e}, f, g, h, \tilde{i}\}$$

B が Z を知っているために生起すべき事象 $K_A(W)$ は

$$K_B(Z) = \{\tilde{a}, b, c, d, \tilde{e}, f, g, h\}$$

以下、

$$K_A(K_B(Z)) = \{\tilde{a}, b, c, d, e, f\}$$
$$K_B(K_A(K_B(Z))) = \{\tilde{a}, b, c, d\}$$
$$K_A(K_B(K_A(K_B(Z)))) = \{\tilde{a}, b, c\}$$
$$K_B(K_A(K_B(K_A(K_B(Z))))) = \phi$$
$$K_A(K_B(K_A(K_B(K_A(K_B(Z)))))) = \phi$$

よって、W, Z ともにコモン・ナレッジではない。

売買中 A にとっては、生起事象が $(\tilde{a}, b, c)$ のどれかまでは分かるので、$\tilde{a}$ の確率は 1/3。よって売り。

B にとっては、生起事象が $(\tilde{a}, b, c, d)$ のどれかまでは分かるので、$\tilde{a}$ の確率は 1/4。よって買い。

クローズ後 W と Z は公的情報となるので、それらはコモン・ナレッジとなる。すると、次のように情報分割が変化する。

$$\begin{aligned} P_A &= \{(\tilde{a}, b, c), (d, \tilde{e}, f), (g, h), (\tilde{i})\} \\ P_B &= \{(\tilde{a}, b, c, d), (\tilde{e}, f, g, h), (\tilde{i})\} \end{aligned}$$

なぜなら、もし $\tilde{i}$ が生起したら B は売り出たはずだが、実際に生じたのは W だったので、A は「$\tilde{i}$ は生起していない」ことを知る。

2 日目の市場

開始前 現実に生起している経済状況は $\tilde{a}$。

売買中 A は売り、B は買い。

クローズ後 同様に情報分割が変化。

$$\begin{aligned} P_A &= \{(\tilde{a}, b, c), (d, \tilde{e}, f), (g, h), (\tilde{i})\} \\ P_B &= \{(\tilde{a}, b, c, d), (\tilde{e}, f), (g, h), (\tilde{i})\} \end{aligned}$$

3 日目の市場

開始前 現実に生起している経済状況は $\tilde{a}$。

売買中 A は売り、B は買い。

クローズ後 同様に情報分割が変化。

$$\begin{aligned} P_A &= \{(\tilde{a}, b, c), (d), (\tilde{e}, f), (g, h), (\tilde{i})\} \\ P_B &= \{(\tilde{a}, b, c, d), (\tilde{e}, f), (g, h), (\tilde{i})\} \end{aligned}$$

4 日目の市場

開始前 現実に生起している経済状況は $\tilde{a}$。

売買中 A は売り、B は買い。

クローズ後 同様に情報分割が変化。

$$\begin{aligned} P_A &= \{(\tilde{a}, b, c), (d), (\tilde{e}, f), (g, h), (\tilde{i})\} \\ P_B &= \{(\tilde{a}, b, c), (d), (\tilde{e}, f), (g, h), (\tilde{i})\} \end{aligned}$$

5 日目の市場

開始前 現実に生起している経済状況は $\tilde{a}$。

売買中 A, B ともに売りとなり、市場は買い手不在となる。結果として株価の暴落を引き起こす。

関連図書

[1] Itzhak Gilboa, David Schmeidler : Maxmin Expected Utility with Nonunique Prior, J. Math. Econ. Vol.18, pp.141–153 (1989).

[2] Robert Aumann : Agreeing to Disagree, Annals of Statistics Vol.4, No.6, pp.1236-1239 (1976).

[3] Lotfi Asker Zadeh : Fuzzy Sets, Inform. Control, Vol.8, pp.338–353 (1965).

[4] Pierre-Simon Laplace 著, 伊藤清訳 : 確率の解析的理論, 共立出版 (1986).

[5] Andrey Nikolaevich Kolmogorov 著, 根本伸司訳 : 確率論の基礎概念, 東京図書 (1971).

[6] Frank Hyneman Knight : Risk, Uncertainty And Profit, Dover Pubns (2006, first published on 1921).

[7] 小島寛之 : 確率的発想法, 日本放送出版協会 (2004).

[8] 柴田文明 : 確率・統計, 岩波書店 (2001).

[9] 薩摩順吉 : 確率・統計, 岩波書店 (1989).

[10] Peter Bernstein 著, 青山護訳 : リスク (上下), 日経ビジネス人文庫 (2001).