<紹　介>

経営のための不確実性理論入門

笹　井　　均

人間の能力の限界を前提とすれば，我々は将来生じ得る結果について確定的に予知することはできない．つまり，意思決定というものは，つねに不確実性を伴うものであるといえよう．

経済主体の行動を分析する場合，不確実性という視点を導入すると取り扱い得る問題領域とその分析成果は飛躍的に増大する．

不確実性と表裏一体の概念として情報という言葉がある．情報を不確実性を減らすための知識であると解釈すれば，不確実性の源泉は情報の不完全性にあると考えることができる．このことは，情報と不確実性が不可分の関係にあることを意味する．

不確実な事象が存在するとき，経済主体の対応として，不確実性それ自体の削減を目的とした能動的対応と不確実性を所与として，それに対する適応を目的とした受動的対応がある．前者は通常，情報の経済学，後者は不確実性の経済学と呼ばれている．

本稿は，経営学部に入学し，不確実性，或は，情報という視点から経営への接近を企ろうという学生にとって最小限必要と思われる基本的概念と理論的分析用具についての入門的解説である．尚，本稿は，統計学についてのごく初歩的な知識のある学生を対象としている．

1. 確率変数

統計学では，観察の結果，又は実験の結果生じ得る可能な事象の集まりを**標本空間**と呼ぶ．標本空間には，生じ得る事象について，その確からしさという意味の測度，すなわち確率が与えられているということが前提となる．

確率変数を標本空間上で定義された実数値関数であると定義する．偶然に支配される事象に対し，1つの数を対応させ，その数が特定の値をとる確率を定めることにより，現実の偶然事象のモデルを作ることになる．

いま，X をそのような確率変数とすると，X が特定の値をとる確率 $p(X=x)$ は，X の対応関係が分っていれば，標本空間上の確率測度によって得られる．離散的な値のみをとる確率変数は**離散確率変数**，ある区間内の任意の値をとることができる確率変数は**連続確率変数**と呼ばれる．

$F(x)=p(X\leqq x)$ で定義される関数 $F(\cdot)$ を確率変数 X の**分布関数**と呼ぶ．X が離散確率変数の場合には，

$$F(x)=\sum_{y\leq x}p(X=y)\equiv\sum_{y\leq x}f(y) \tag{1}$$

と書ける．一方，X が連続確率変数の場合には，
$p(X=x)=0$ となるが，

$$F(x)=\int_{-\infty}^{x}f(t)\,dt \tag{2}$$

と書ける場合がある．このとき，(1)(2)において，$f(x)$ を確率変数 X の**密度関数**と呼ぶ．

いくつかの代表的な分布関数は，実際に現れる不確実な現象を説明するモデルとして有用であることは広く知られているところである．

偶然に支配される事象を取り扱う時には，1つだけの確率変数より，むしろいくつかの確率変数を考えるのが普通である．このとき，2つの確率変数 X，Y を考え，X が特定の値 x をとり，同時に Y が特定の値 y をとる確率 $p(X=x,\ Y=y)$ を導入し，同様な考え方に基づいて定義された $F(x,y)$，$f(x,y)$は，X と Y の**同時分布関数**，**同時密度関数**と呼ばれる．同時密度関数 $f(x,y)$ に対して，$f(x)=\sum_{y}f(x,y)$ あるいは $f(x)=\int_{-\infty}^{+\infty}f(x,y)\,dy$ を X の**周辺密度関数**という．

また，X，Y の密度関数を $f(x)$，$g(y)$ とすると，すべての x,y に対して，

$$f(x,y)=f(x)\cdot g(x)$$

が成立するとき，X と Y は**独立**であるという．

よく知られているように，事象 A_1, A_2 に対して，A_1 が起るという条件の下での A_2 の起る確率（**条件つき確率**）は，$p(A_1) \neq 0$ ならば

$$p(A_2 | A_1) = p(A_1 \cup A_2) / p(A_1) \tag{3}$$

で定義される．いま，確率変数 X が x の値をとる事象を A_1 とし，確率変数 Y が y の値をとる事象を A_2 とすると，$A_1 \cap A_2$ は X, Y がそれぞれ x および y を同時にとる事象であり，$p(A_2 | A_1)$ は X が x なる値をとるという条件のもとで Y が y という値をとる確率を示すことになる．結局(3)式から，X が x に固定されているときの Y の密度関数（**条件つき密度関数**）は $f(x) \neq 0$ なら，

$$f(y | x) = f(x, y) / f(x) \tag{4}$$

と考えることができる．もし，(4)式で条件つき密度関数を定義するとそれは(3)式の拡張になっていることに注意しよう．

確率変数 X の密度関数を $f(x)$ とする．X の実現値 x の関数 $\phi(x)$ について，**$\phi(X)$ の期待値**を，

$$E[\phi(x)] = \begin{cases} \sum \phi(x) f(x), & X \text{が離散のとき} \\ \int \phi(x) f(x)\, dx, & X \text{が連続のとき} \end{cases} \tag{5}$$

と定義する．$\phi(x) = x$ のときの期待値，すなわち X の期待値（**平均値**）を，

$$\mu = E[X] \tag{6}$$

と書くと，$g(x) = (x-\mu)^2$ のときの $g(X)$ の期待値，つまり，

$$\sigma^2 = \mathrm{Var}[X] = E[(X-\mu)^2] \tag{7}$$

は，X の**分散**（平均値の周りの2次の積率）と呼ばれる．

X を x に固定したときの $\phi(X, Y)$ の期待値，

$$\begin{aligned} E[\phi(X, Y) | x] &= \sum \phi(x, y) f(y | x) \\ &= \int \phi(x, y) f(y | x)\, dy \end{aligned} \tag{8}$$

は，$X = x$ が与えられたときの $g(X, Y)$ の**条件つき期待値**と呼ばれる．容易に検証されることであるが，次の等式は極めて高い利用価値をもつ結果である．

$$E[\phi(X, Y)] = E[E[\phi(X, Y) | X]] \tag{9}$$

(問)　ある捕虜が3つのドアをもつ独房に収容されている．1番目のドアを出ると直ちに自由が得られ，2番目のドアを出ると1日間費やした後にトンネルを通らされ再び独房にもどり，3番目のドアについては3日間費やした後に独房にもどるものとする．捕虜は3つのドアのどれも等しい確率で選ぶものとするとき，彼の自由を得るまでの期待時間（日数）を(9)式を用いて計算せよ．

2. 期 待 効 用

意思決定問題をモデル化する際に用いられる基本的概念は，**行動**または決定の集合，**結果**または**成果**の集合，**状態**の集合のつである．確定的な世界においては，ある1つの行動に対してただ1つの結果が対応するが，不確実な世界では，ある1つの行動に対し，状態と呼ばれる意思決定主体が制御できない外的環境に依存して複数の結果が生じ，その対応関係は複雑になる．このような不確定性を含んだ状況において，意思決定主体は自己の**決定規準**を設定し，最適な行動の選択を行わなければならない．

上記の3つの概念と決定規準によって構成されるこのような考え方が，不確実性下における意思決定モデルの基本的な理論構造である．以下では，行動，状態，結果はすべて実数によって規定されるものとする．

最もよく知られた決定規準として，**期待値基準**がある．期待値の大小によって行動の選択を順序づけるという考え方である．しかしながら，期待値規準は，しばしば不合理な選択を導くことがある．聖ペテルスブルグのパラドックスとして知られている例を紹介することにしよう．1枚のコインを表が出るまで投げ続け，n 回目で始めて表が出れば 2^n 円受けるという賭を考える．この賭から得られる期待値は $\sum 2^n \left(\frac{1}{2}\right)^n = \infty$ となる．したがって，例えば10億円確実に受けとるという行動とこの賭に参加するという行動を期待値規準によって比べた場合，つねに後者が選択されるということになる．明らかにこのような選択をする人は多くない．期待値規準は意思決定主体の不確実性に対する態度や価値観を十分に反映した基準ではないと推論できる．

金銭表示による利得の期待値そのものを比較するのではなく，利得のもつ決定主体の主観的価値を効用と定義し，効用の期待値の大小によって選択を結論づけるという考え方がある．不確実性に直面した意思決定主体は，ある種の効用関数の期待値を最大化するように行動しているという前提に立脚するいわゆる**期待効用規準**と呼ばれるものである．

フォン・ノイマン＝モルゲンシュテルンは，行動に

よって実現する結果(成果)が不確実な状態に依存して決まるとき，意思決定主体の行動が合理性に関するいくつかの公理を満たすなら，各行動の結果に対する効用を特定化するような効用関数が存在し，それは正の1次変換の範囲で一意的に定まるということを示した．このような(基数的)効用関数はノイマン＝モルゲンシュテルン形の**効用関数**と呼ばれる．多くの場合，行動によって実現される結果としては金銭的利得を考えることが多い．

いま，効用関数の存在を前提とすれば，行動の選択，同じことであるが，結果集合上の確率分布の選択が実数値である期待効用値の大小によって表現されるということになる．すなわち，$f^a(x)$，$f^b(x)$ を行動 a, b によって確率変数である結果 $X=x$ が実現する密度関数，$u(x)$ を結果 x に対する効用関数とすれば，

$$E^a[u(x)]=\sum u(x)f^a(x)\geq E^b[u(x)]$$
$$=\sum u(x)f^b(x) \iff a\succ b \qquad (10)$$

である．ここで，$a\succ b$ は a は b より選好されるという意味の記号である．

3. 効 用 関 数

意思決定主体の不確実性に対する態度を規定する代表的な効用関数 $u(x)$ について説明することにする．以下では，簡単のため，離散分布，とくに2つの結果のみをとる場合，例えば確率 p で結果 x_1 を，確率 $1-p$ で結果 x_2 を実現する**くじ**をとりあげて説明するが，そこで得られる結論は容易に一般化可能である．更に現実性という観点から，$u(x)$ は単調増加関数とする．

まず，意思決定主体の効用関数が凹関数となる場合を考える．確率 p で結果 $X=x_1$ を確率 $1-p$ で結果 $X=x_2$ を実現するくじについて，その期待値は

$$\bar{x}=px_1+(1-p)x_2, \qquad x_1<x_2$$

であり，期待効用値は

$$E^l[u(X)]=pu(x_1)+(1-p)u(x_2)$$

となる．また，確率変数 X により得られる期待効用水準と同一水準の効用をもたらす確定的な結果，すなわち

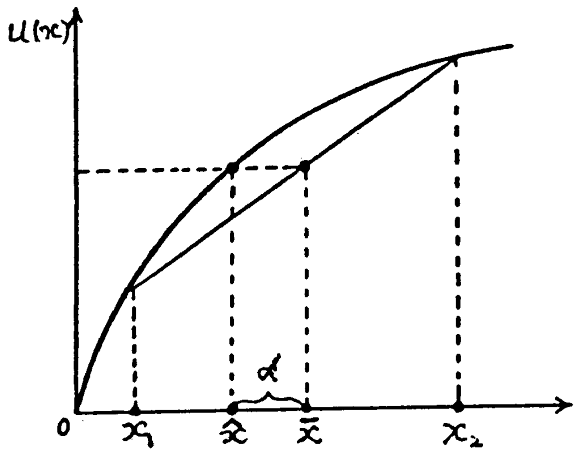
$$u(\hat{x})=E^l[u(X)] \qquad (11)$$

となる $\hat{x}$ は，**確実同値量**と呼ばれる．$u(x)$ は凹関数であるから，図から明らかなように，つねに $\hat{x}<\bar{x}$ となる．したがって，凹形効用関数を有する意思決定者は，もし確定的に $\bar{x}$ が得られるなら，不確実に変動するくじ(期待$=\bar{x}$)より，その方を選好する．この意味で**危険回避的**であるという．

$\bar{x}-\hat{x}=\alpha>0$ と置くと，

$$u(\hat{x})=u(\bar{x}-\alpha)=E^l[u(X)]$$

と書ける．したがって，α は，$\bar{x}$ を中心に考えて x_1 または x_2 と不確実に変動する成果 X を避けて，確定的な成果 $\hat{x}$ を得ることができるなら，$\bar{x}$ を得ることを断念し犠牲にしてもよい（支払ってもよい）と考える最大限度量と解釈できる．この α は**保険プレミアム**或いは，危険コストという意味で（マイナスの）**危険プレミア**と呼ばれる．



もし，効用関数が凸関数であれば事態はまったく逆転し，その場合には**危険愛好的**であるという．更に，効用関数が線形の場合には，明らかに期待値のみで選好が定まり，**危険中立的**と呼ばれる．

(問)　次の2つの行動があり，$u(-200)=0, u(100)=0.4, u(200)=0.6$ である．

a_1：確実に100万円を得る

a_2：確率 $\frac{3}{4}$ で200万円を得るか，確率 $\frac{1}{4}$ で200万円を失う．この意思決定者の危険に対する態度を規定せよ．

アロー＝プラットは（増加効用関数に対して）危険回避の尺度として**絶対的危険回避関数**[1)]

$$r(x)=-\frac{u''(x)}{u'(x)}=-\frac{d}{dx}\{\log u'(x)\} \qquad (11)$$

を導入した．任意の x に対して $r(x)>0$ なら危険回避的，$r(x)<0$ なら危険愛好的，$r(x)=0$ なら危険中立的となる．

4. 応　用　例

(1) 保険契約[2)]

ある事故のリスクを保険会社Bに移転しようと考えている企業Aがある．企業Aは危険回避的(効用関数$u(x)$)であり，保険会社Bは多数の顧客と保険契約を結んでいるため危険中立的 (効用関数 $v(x)$) であるものとする．A社が被る事故の発生確率は0.001ときわめて小さいが，その損害額は10億円と膨大であるため(勿論，事故が起らない場合の損失は0)，A社はB社にc万円の保険料を支払って保険契約を締結するという状況を想定しよう．したがって，契約前は，A社は確率0.001で−10億円，0.999で0円のくじに，B社は確率1で0円のくじに直面し，契約後は，A社は確率1で $-c$ 万円，B社は確率0.001で −10億円$+c$万円，0.9999でc万円のくじに直面することになる．契約前におけるA社の期待利得(損失)は−100万円であるが，危険回避的であるため，くじの確実同値量 $-\hat{x}$ は，$-\hat{x}<-100$万円$=\bar{x}$ となる．もし，$-\hat{x}<-c$ であれば契約後に得られるくじの期待効用値($=u(-c)$)は契約前の期待効用値 ($=u(-\hat{x})$) より大きくなり，保険契約の締結を選好する．一方，B社にとっては，もし$c>100$万円ならば契約後の期待利得は$c-100$万円であるが，危険中立的であるため期待利得は確実同値量に一致し，契約後の期待効用値 ($=v(c-100$万円$)$) は契約前の期待効用値($=v(0)$)より大きくなり，B社も保険契約の締結を望むことになる．

この問題は，視点を変えれば，契約前の状態も，契約後の状態も，不確実な期待値 −100 万円の将来利得(損失) をあらかじめ A 社と B 社で分配している状況と見なすことができよう．そして，c万円，$-\hat{x}<-c<-100$万円，の保険料金の支払いにより，すべてのリスクを移転した後者の分配方式が前者の分配方式よりA社にとってもB社にとっても有利であるということを意味している．

このように，ある分配方式より，他者を不利にすることなく，少くとも1人がより有利になるような分配方式はパレート優位な分配方式であるという．ある分配方式より，パレート優位な分配方式が存在しないとき，パレート最適な分配方式であるという．

(2) 危険のプーリング

n 個の経済主体があり，i 番目の主体の生み出す利得を確率変数 X_i とする．X_i の期待値を $\mu_i=E[X_i]$，分散を $\sigma_i^2=E[(X_i-\mu_i)^2]$ と置く．n 個の経済主体がすべての利得をプールした場合，$X=\sum_1^n X_i$ の期待利得は，

$$E[X]=\sum E[X_i].$$

分散は簡単な計算から，

$$E[(X-E[X])^2]=\sum_i \sigma_i^2+2\sum_{i\neq j}\mathrm{cov}(X_i, X_j) \tag{13}$$

となる．ここで，X_i と X_j の共分散は，

$$\begin{aligned}\mathrm{cov}(X_i, X_j)&=E[(X_i-E[X_i])(X_j-E[X_j])]\\&=E[X_i\cdot X_j]-E[X_i]\cdot E[X_j]\end{aligned} \tag{14}$$

で与えられる．もし，$\mathrm{cov}(X_i, X_j)<0$ なら利得のプーリングにより，期待利得は変化しないが分散は減小する．このことは危険回避的意思決定主体にとって保険プレミアムを減小させる[1)] 効果をもつ．

5. 情報の価値

意思決定を情報という側面から分析しようとする場合，情報のもつ価値の計量という問題は，避けて通ることのできない重要な問題となる．このテーマは，情報の特性をいかに把握するかというきわめて，本質的でかつ遠大な課題であり，現在なお統括的理論構築を目指して発展しつつある分野である．ここでは，情報の価値の計量のための1つのアプローチについて紹介する．

まず，このことを議論するために，重要な役割をになうベイズの定理について簡単に触れることにしよう．

標本空間を互いに素な集合 $A_1, A_2, \cdots\cdots A_k$ に分割する．すると，どんな結果にしろある結果が生じたとするとその結果が起因する可能な原因は $A_1, A_2, \cdots\cdots, A_k$ であると考えることができる．いまある実験或いは観察が行われ事象Bが起ったとき，この結果がある特別な原因 A_i にもとづくものである確率はよく知られたベイズの公式：

$$p(A_i\,|\,B)=\frac{p(B\,|\,A_i)p(A_i)}{\sum_{j=1}^{k}p(B\,|\,A_j)p(A_j)}=\frac{p(B\,|\,A_i)p(A_i)}{p(B)} \tag{15}$$

によって計算できる．確率変 X のとり得る値を $x_1, x_2, \cdots\cdots, x_k$ と考え，X が x_i の値をとる事象を A_i，確率変数 Y が y の値をとる事象を B とすると(15)式

は，密度関数を用いて，$g(y)>0$ なら，

$$f(x_i\mid y)=\frac{g(y\mid x_i)f(x_i)}{\sum_{j=1}^{k}g(y\mid x_j)f(x_j)}=\frac{g(y\mid x_i)f(x_i)}{g(y)} \tag{16}$$

と書くことができる．連続確率変数の場合は和が積分に変わる．これは，ベイズの公式の拡張になっていることに注意されたい．$f(x_i)$ は**事前確率密度**，$f(x_i\mid y)$ は**事後確率密度**と呼ばれる．

さて，意思決定問題において，行動または決定の集合を $A=\{a_1,a_2,\cdots\cdots,a_n\}$，$a_i\in R'$，確率変数 S である状態のとり得る集合 $\{s_1,s_2,\cdots\cdots,s_m\}$ を，その離散密度関数を $p(S=s_j)=f(s_j)$，$s_j\in R'$，結果の集合を $\{u(a_i,s_j)\}$，$i=1,\cdots\cdots,n$，$j=1,\cdots\cdots m$，とする．$u(a_i,s_j)$ は状態が s_j であるとき，行動 a_i を行ったときの効用値であるが，ここではモデルを具体的に把握するため，決定主体は危険中立的であり，$u(a_i,s_j)$ は利得関数を表わすものとする．したがって，意思決定主体は期待利得規準にもとづいて行動する．

意思決定主体が状態 S に関する事前密度関数 $f(s_j)$ しか利用できない場合（**ゼロ情報**）には，彼は期待利得，

$$E[u(a_i,S)]=\sum_{j=1}^{m}u(a_i,s_j)f(s_j) \tag{17}$$

を最大にする行動 a_i^* を A の中から選択する．この場合，意思決定主体は行動の期待値は知り得るものの行動を行った後で実際に実現する利得については不確実である．

もし**状態に関連する**なんらかの情報が，**決定が行われる前に**利用或いは観測できるとした場合には，その情報が状態について，いくばくかの情報を提供する以上，行動は入手できる情報に依存して選択されるべきものになるはずである．いま，その情報を確率変数 Θ，その可能な実現値を $\theta_1,\theta_2,\cdots\cdots,\theta_m$ で表わすと，意思決定主体は，期待利得，$E[u(a_i(\Theta),S)]$ を最大にする行動 $a_i^*(\theta_k)$，$k=1,2,\cdots\cdots,m$ を集合 A の中から選択することになる．

ところで，

$$\begin{aligned}E[u(a_i(\Theta),S)]&=\sum_j\sum_k u(a_i(\theta_k),s_j)g(\theta_k\mid s_j)f(s_j)\\&=\sum_k\sum_j u(a_i(\theta_k),s_j)f(s_j\mid\theta_k)g(\theta_k)\\&=E[E[u(a_i(\Theta),S)\mid\Theta]]\end{aligned} \tag{18}$$

が成立しているから，最適な行動は θ_k 毎に，

$$\sum_j u(a_i(\theta_k),s_j)f(s_j\mid\theta_k)$$

を最大にする $a_1^*(\theta_k)$，$k=1,2,\cdots\cdots,n$ となる．

いま，情報の信頼度とでも呼ぶべき，$g(\theta_k\mid s_j)$，k，$j=1,\cdots\cdots,n$ が所与とすると，ベイズの公式，(16)式を用いて，事前確率密度 $f(s_j)$ から事後確率密度 $f(s_j\mid\theta_k)$ が計算でき，この問題は完全に解決できる．換言すれば，ゼロ情報の場合の最適意思決定は，事前確率密度を用いた期待利得の最大化問題であり，関連する情報の場合の最適意思決定は，事後確率密度を用いた条件つき期待利得の最大化問題であるということができる．ここにおいて，関連情報の価値を

$$E[u(a_i^*(\Theta),S)]-E[u(a_i^*,S)] \tag{19}$$

と定義することは，きわめて説得力をもつことになろう．もし，

$$g(\theta_k\mid s_j)=\begin{cases}1, & k=j\\0, & k\neq j\end{cases},\ j,\ k=1,\cdots\cdots,m$$

であるなら，そのとき，情報 θ_k は**完全情報**を提供するものとなることはいうまでもない．

(問)　選択可能な行動 a_1,a_2 があり，各行動をとったときの利得は状態に依存してつぎのように定まるものとする．

行動＼状態	s_1	s_2
a_1	3	2
a_2	8	−2

事前確率密度は $f(s_1)=0.6$，$f(s_2)=0.4$ とする．このとき，最適意思決定において，ゼロ情報の場合の期待利得と信頼度80%の情報が利用可能な場合の期待利得を計算せよ．

注

1)　保険プレミアムについて，$\alpha\doteqdot r(x)\cdot\frac{\mathrm{Var}[X]}{2}$ の近似式がよく使用される．

2)　小林孝雄「リスク分担と意思決定」，ビジネスレビュー，vol. 28, No. 2, 1980. から引用した．

〔ささい ひとし　横浜国立大学経営学部教授〕