

確率的順序関係と学習のプロセスについて

中井, 達

<https://doi.org/10.15017/4494403>

出版情報：経済學研究. 62 (1/6), pp. 465-486, 1996-03-10. 九州大学経済学会
バージョン：
権利関係：



確率的順序関係と学習のプロセスについて

中 井 達

1. 確率的順序関係

確率的決定問題、例えば最適停止問題・取り替え問題・スケジューリング・信頼性・待ち行列などを考えるとき、複数の確率分布を比較することは問題の性質を解析する上で必要となり、どのような分布の族で問題を扱うかを考える上でも有用なことが多い。また、不完備情報のもとで多段決定問題を考えるとき、これらの情報は集合の上の確率分布で表され、それらの情報の持つ価値を比較する必要がある。これらの確率分布の間に順序を導入して、最適政策をはじめその政策にしたがって得られる総期待利得の性質を考える。これらの確率的順序関係については Stoyan [20] でもまとめられ、待ち行列・信頼性など多くの確率モデルで応用されている。また Ross [17] などでも考えられている。

ここでは、確率分布の間の順序関係の中で基本的な確率順序・尤度比順序・故障率順序などを考える。確率順序は最も一般的な確率的順序関係で2節で考える。3節では故障率関数を用いた順序を考える。この順序関係は信頼性理論や取り替え問題などでよく用いられ、スケジューリング理論に関して Pinedo and Ross [15] でも扱われている。4節では尤度比順序を考える。この順序はこれまでの2つの順序に比べ、良い性質を持ち学習プロセスを考える上でも重要なもので、多くの問題で扱われている。特にベイズの理論に基づいた学習プロセスで、事前情報 (Prior Information) と事後情報 (Posterior Information) の関係については6節と7節で詳しく考える。この尤度比順序は Brown and Solomon [2] でも扱われている。5節ではそれまで考えた3つの順序を含め、関数類を用いて特徴付ける方法を説明する。この考え方は、Stoyan [20] によって体系的にまとめられ、新しい研究もある。

6節と7節では、ベイズ理論を用いた学習プロセスでの事前情報と事後情報の関係について述べる。この学習プロセスの性質を考える上で尤度比順序を導入する。6節では Nakai [7, 8, 10] などで扱われた結果をまとめて述べる。7節では6節とは異なり1度に複数の値を観測する場合の学習プロセスについて考える。この問題は中井 [14] で詳しく扱われ、学習プロセスについて得られた結果を述べる。これらの学習プロセスのもとでの決定問題は上記の論文で扱われている。

2. 確率順序

2つの確率変数 X と Y を考える。

定義 1 確率変数 X と Y に対し, x を任意の実数とすると

$$\Pr(X > x) \geq \Pr(Y > x) \quad (1)$$

のとき, 確率変数 X が確率変数 Y より確率的に大きいと言い $X \geq_s Y$ と表す。

X と Y がそれぞれ確率分布関数 $F_X(x)$ と $F_Y(x)$ を持つとき, 任意の実数 x に対し (1) 式は

$$\bar{F}_X(x) \geq \bar{F}_Y(x)$$

となることに等しい。ここで, $\bar{F}(x) = 1 - F(x)$ とする。このことは確率変数 X が x より大きい値を取る確率が確率変数 Y に比べて大きいことを表している。

2つの確率変数 X と Y に対し, 定義 1 の (1) 式は分布関数を用いれば任意の正整数 x に対し

$$F_X(x) \leq F_Y(x) \quad (2)$$

と表せる。いま, 関数 $I_x(t)$ を

$$I_x(t) = \begin{cases} 1 & t \geq x \\ 0 & t < x \end{cases} \quad (3)$$

とすれば, (2) 式は任意の $x \in \mathbf{R}$ に対し

$$\int_{-\infty}^{\infty} I_x(t) f_X(t) dt \leq \int_{-\infty}^{\infty} I_x(t) f_Y(t) dt \quad (4)$$

に等しい。ここで $f_X(t)$ と $f_Y(t)$ はそれぞれ確率変数 X と Y の確率密度関数とする。

定理 1 X と Y を 2つの確率変数とする。 t に関する任意の非減少関数 $g(t)$ に対し

$$\mathbf{E}[g(X)] = \int_{-\infty}^{\infty} g(t) f_X(t) dt \geq \int_{-\infty}^{\infty} g(t) f_Y(t) dt = \mathbf{E}[g(Y)] \quad (5)$$

となることと, $X \geq_s Y$ は等しい。

証明 $I_x(t)$ が t の非減少関数だから

$$\int_{-\infty}^{\infty} I_x(t) f_X(t) dt \geq \int_{-\infty}^{\infty} I_x(t) f_Y(t) dt$$

が成り立ち, (4) 式から $X \geq_s Y$ が導かれる。

逆に, $g(t)$ を微分可能な非減少関数で, $\lim_{t \rightarrow -\infty} g(t) = 0$ および $\lim_{t \rightarrow \infty} g(t) = K$ となる関数とする。 $g(t)$ は有界だから $\int_{-\infty}^{\infty} g(t) f_X(t) dt$ と $\int_{-\infty}^{\infty} g(t) f_Y(t) dt$ は定義可能で,

$$g(t) = \int_{-\infty}^t g'(x) dx = \int_{-\infty}^{\infty} I_x(t) g'(x) dx \quad (6)$$

となる。したがって, 確率変数 X の密度関数 $f_X(t)$ に対し

$$\int_{-\infty}^{\infty} g(t) f_X(t) dt = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} I_x(t) g'(x) dx \right) f_X(t) dt$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f_X(t) I_X(t) dt \right) g'(x) dx \\
&= \int_{-\infty}^{\infty} \left(\int_x^{\infty} f_X(t) dt \right) g'(x) dx \\
&= \int_{-\infty}^{\infty} (1 - F_X(x)) g'(x) dx
\end{aligned} \tag{7}$$

となる。確率変数 Y に対しても

$$\int_{-\infty}^{\infty} g(t) f_Y(t) dt = \int_{-\infty}^{\infty} (1 - F_Y(x)) g'(x) dx \tag{8}$$

となる。関数 $g(t)$ の単調性から $g'(t) \geq 0$ であり、 $X \geq_s Y$ より任意の t に対し $\bar{F}_X(t) \geq \bar{F}_Y(t)$ となる。したがって

$$\int_{-\infty}^{\infty} g(t) f_X(t) dt \leq \int_{-\infty}^{\infty} g(t) f_Y(t) dt$$

より、この定理が成り立つ。一般の有界でない非減少関数 $g(t)$ に対しても、この結果を用いて同様に求められる。また、 $g(t)$ が微分可能でなければ

$$g(t) = \int_{-\infty}^t dg(x)$$

として同様となる。

最後に、関数 $g(x)$ が減少関数で $X \geq_s Y$ となる任意の 2 つの確率変数に対し (5) 式を満足するものは存在しないことは次のことから示される。 $g(x_1) > g(x_2)$ となる $x_1 < x_2$ に対し $F_X(x_1) = I_{x_1}(x)$ および $F_Y(x) = I_{x_2}(x)$ となる確率変数 X と Y を考えれば良い。(すなわち確率変数 X は確率 1 で値 x_1 を取る確率変数である。) □

注 1 定理 1 で $f(x) = x$ とおけば $X \geq_s Y$ なら $\mathbf{E}[X] \geq \mathbf{E}[Y]$ となる。

このことは、つぎのように簡単に示される。 X と Y を 2 つの確率変数で $X \geq_s Y$ とすると、確率順序の定義 (定義 1) より

$$\begin{aligned}
\mathbf{E}[X] &= \int_0^{\infty} x f_X(x) dx = \int_0^{\infty} (1 - F_X(x)) dx \\
&\geq \int_0^{\infty} (1 - F_Y(x)) dx = \int_0^{\infty} y f_Y(y) dy = \mathbf{E}[Y]
\end{aligned}$$

となる。

確率変数 X の確率密度関数を $f(x)$ とし、分布関数を $F(x)$ とするとき

$$\lambda(t) = \frac{f(t)}{\bar{F}(t)} \tag{9}$$

とおき、この関数 $\lambda(t)$ を確率変数 X の故障率関数 (Failure Rate Function, Hazard Rate Function) と呼ぶ。(9) 式を積分すれば

$$\bar{F}(a) = e^{-\int_a^{\infty} \lambda(x) dx}$$

となる。

関数 $\lambda(t)$ が t に関する非減少関数のとき、この確率変数 X を IFR (Increasing Failure Rate) 確率変数と呼び、 $\lambda(t)$ が非増加関数のときこの確率変数を DFR (Decreasing Failure Rate) 確率変数と呼ぶ。

非負確率変数 X に対し、 X_t を

$$\begin{aligned}\bar{F}_t(x) &= P(X_t > x) \\ &= \Pr(X - t > x | X > t) \\ &= \frac{\bar{F}(t+x)}{\bar{F}(t)}\end{aligned}\quad (10)$$

を分布関数とする確率変数とする。 X をある製品の有効期間を表す確率変数とすれば、この確率変数 X_t は t まで有効のとき、残りの有効期間を表す条件付き確率変数となる。

補題 1 X_t が IFR となることと X_t が t に関し確率順序の意味で非増加となることは同値である。

また、 X_t が DFR となることと X_t が t に関し確率順序の意味で非減少となることは同値である。

証明 確率変数 X と X_t の故障率関数を、それぞれ λ と λ_t とする。 $F_t(x) = \frac{F(t+x) - F(t)}{\bar{F}(t)}$ より

$f_t(x) = \frac{f(t+x)}{\bar{F}(t)}$ だから、(10)式より

$$\begin{aligned}\lambda_t(x) &= \frac{f_t(x)}{\bar{F}_t(x)} \\ &= \frac{\frac{f(t+x)}{\bar{F}(t)}}{\frac{\bar{F}(t+x)}{\bar{F}(t)}} \\ &= \frac{f(t+x)}{\bar{F}(t+x)} \\ &= \lambda(t+x)\end{aligned}$$

となる。

$$\begin{aligned}\bar{F}_t(a) &= e^{-\int_0^a \lambda_t(x) dx} \\ &= e^{-\int_t^{t+a} \lambda(x) dx}\end{aligned}$$

より、 $\lambda(x)$ が非減少 (非増加) 関数なら $s < t$ のとき $\bar{F}_s(a) > \bar{F}_t(a)$ となり X_t は確率順序の意味で t の非増加 (非減少) 確率変数となる。

逆に、 X_t は確率順序の意味で t の非増加 (非減少) 確率変数なら、 $s < t$ のとき全ての正数 a に対し

$$\bar{F}_s(a) = e^{-\int_s^{s+a} \lambda(x) dx} \geq \bar{F}_t(a) = e^{-\int_t^{t+a} \lambda(x) dx}$$

が成り立つ。したがって

$$\int_t^{t+a} \lambda(x) dx > \int_s^{s+a} \lambda(x) dx$$

より、 $\lambda(t)$ は非減少 (非増加) 関数となる。□

例1 パラメータ $\lambda(>0)$ と $\mu(>0)$ の指数分布 X_λ と X_μ を考える。すなわち,

$$\Pr(X_\lambda > x) = e^{-\lambda x} \quad (0 \leq x < \infty)$$

とする。このとき, $\lambda \leq \mu$ なら任意の x に対して $\Pr(X_\lambda > x) \geq \Pr(X_\mu > x)$ だから $X_\lambda \geq_s X_\mu$ となる。

例2 自然数全体 $\{1, 2, \dots\}$ の上での確率分布は確率ベクトル $\mathbf{P} = (p_1, p_2, p_3, \dots)$ で表せる。ただし, $p_i \geq 0$ ($i=1, 2, 3, \dots$), $\sum_{i=1}^{\infty} p_i = 1$ とする。ここで, 自然数の上の確率分布全体を S とおくと,

$$S = \{\mathbf{p} | \mathbf{p} = (p_1, p_2, p_3, \dots), p_i \geq 0 (i=1, 2, 3, \dots), \sum_{i=1}^{\infty} p_i = 1\} \quad (11)$$

となる。確率分布 $\mathbf{p}$ と $\mathbf{q}$ に対し $\mathbf{p} \geq_s \mathbf{q}$ となることは, (1)式より

$$\sum_{i=h}^{\infty} p_i \geq \sum_{i=h}^{\infty} q_i$$

と表せる。

3. 故障率関数による順序

確率変数 X の故障率関数を $\lambda_X(t)$ とすると, (9)式より

$$\lambda_X(t) = \frac{f_X(t)}{F_X(t)} \quad (0 \leq t \leq \inf\{t | F(t) = 1\}) \quad (12)$$

となる。

定義2 確率変数 X と Y が任意の実数 $t \geq 0$ に対し

$$\lambda_X(t) \leq \lambda_Y(t) \quad (13)$$

となるとき, 確率変数 Y は確率変数 X より大きい故障率関数を持つと言い, $X \geq_h Y$ と表す。

(10)式から

$$\Pr(X > t+s | X > t) = e^{-\int_t^{t+s} \lambda_X(y) dy} \quad (14)$$

より, $\lambda_X(t) \leq \lambda_Y(t)$ なら

$$\Pr(X > t+s | X > t) \geq \Pr(Y > t+s | Y > t) \quad (15)$$

となり, $X_t \geq_s Y_t$ と等しい。

例3 (Stoyan [20])

(12)式から

$$\lambda_X(t) = -\frac{d}{dt} \log(1 - F_X(t))$$

より

$$F_X(t) = 1 - e^{-\int_0^t \lambda_X(u) du}$$

となる。 $\int_0^t \lambda_X(u) du \geq \lambda_0 t$ がすべての $t (\geq 0)$ に対し成り立てば,

$$F_X(t) = 1 - e^{-\int_0^t \lambda_X(u) du} > 1 - e^{-\lambda_0 t}$$

だから確率変数 Y をパラメータ λ_0 の指数分布とすると $X \leq_s Y$ となる。

例4 $\lambda(>0)$ と $\mu(>0)$ および $\alpha(>0)$ と $\beta(>0)$ に対し

$$\Pr(X>x)=e^{-\lambda t^\alpha} \quad (0\leq x<\infty)$$

$$\Pr(Y>x)=e^{-\mu t^\beta} \quad (0\leq x<\infty)$$

で定義される確率変数 X と Y を考える。これらの確率変数はワイブル分布である。

$$\begin{aligned}\lambda_X(t) &= \frac{\lambda \alpha t^{\alpha-1} e^{-\lambda t^\alpha}}{e^{-\lambda t^\alpha}} \\ &= \lambda \alpha t^{\alpha-1}\end{aligned}$$

$$\lambda_Y(t) = \mu \beta t^{\beta-1}$$

だから、 $\lambda \leq \mu$ および $\alpha \leq \beta$ とすると $X \leq_h Y$ となる。

ここで考えた故障率関数による順序は一般には次のように定義される。

定義3 すべての実数 s と u ($s \leq u$) に対し

$$\left| \begin{array}{cc} \bar{F}_X(u) & \bar{F}_X(s) \\ \bar{F}_Y(u) & \bar{F}_Y(s) \end{array} \right| = \left| \begin{array}{cc} \Pr(X>u) & \Pr(X>s) \\ \Pr(Y>u) & \Pr(Y>s) \end{array} \right| \geq 0$$

のとき、 $X \geq_h Y$ という。

補題2 確率変数 X と Y が連続ならば定義2と3は同値である。

証明 定義2を仮定する。 $X \geq_h Y$ とし、確率変数 X と Y の確率密度関数をそれぞれ $f_X(t)$ と $f_Y(t)$ とする。すべての t (≥ 0) に対し

$$\lambda_X(t) \leq \lambda_Y(t)$$

だから、(15)式より任意の t (≥ 0) に対し

$$\Pr(X>t+s|X>s) \geq \Pr(Y>t+s|Y>s)$$

となる。このことは、

$$\frac{\bar{F}_X(s+t)}{\bar{F}_X(s)} \geq \frac{\bar{F}_Y(s+t)}{\bar{F}_Y(s)}$$

に等しい。 $u=s+t$ とおくと

$$\frac{\bar{F}_X(u)}{\bar{F}_X(s)} \geq \frac{\bar{F}_Y(u)}{\bar{F}_Y(s)}$$

となるから、定義3が導かれる。

逆に、 $s \leq u$ のとき

$$\frac{\bar{F}_X(u)}{\bar{F}_X(s)} \geq \frac{\bar{F}_Y(u)}{\bar{F}_Y(s)}$$

とする。 $u=s+t$ ($t \geq 0$) とおくと

$$\frac{\bar{F}_X(s+t)}{\bar{F}_X(s)} \geq \frac{\bar{F}_Y(s+t)}{\bar{F}_Y(s)}$$

となり

$$\Pr(X>s+t|X>s) \geq \Pr(Y>s+t|Y>s)$$

が任意の $t \geq 0$ と s に対し成り立つ。(14)式よりすべての t (≥ 0) に対し

$$\int_s^{s+t} \lambda_X(x) dx \leq \int_s^{s+t} \lambda_Y(x) dx$$

となり、定義 2 が成り立つ。□

4. 尤度比順序

尤度比 (Likelihood Ratio) を用いて確率変数の間に順序を導入する。

定義 4 X と Y を連続な確率変数とし、 $f_X(x)$ と $f_Y(x)$ をそれぞれの確率変数の確率密度関数とする。

$x \geq y$ なら

$$\frac{f_X(y)}{f_Y(y)} \leq \frac{f_X(x)}{f_Y(x)} \quad \text{または} \quad \begin{vmatrix} f_X(x) & f_Y(y) \\ f_X(y) & f_Y(x) \end{vmatrix} \geq 0$$

となるとき、確率変数 X は確率変数 Y より尤度比の意味で大きいと言い $X \geq_l Y$ と表す。

定義 4 で導入された順序を尤度比順序という。この順序は確率過程を考えると、 TP_2 (Total Positivity of Order 2) と呼ばれる。(Karlin [3], Karlin and Taylor [4])

定理 2 X と Y を連続な非負確率変数で、それぞれの確率密度関数を $f_X(t)$ と $f_Y(t)$ とし、故障率関数 $\lambda_X(t)$ と $\lambda_Y(t)$ を持つとする。このとき、 $X \geq_l Y$ なら

$$\lambda_X(t) \leq \lambda_Y(t)$$

が任意の t に対し成り立つ。すなわち、 $X \geq_h Y$ となる。

証明 $X \geq_l Y$ より $x \geq t$ のとき $f_X(x) \geq \frac{f_Y(x)f_X(t)}{f_Y(t)}$ となる。よって、

$$\begin{aligned} \lambda_X(t) &= \frac{f_X(t)}{F_X(t)} \\ &= \frac{f_X(t)}{\int_t^\infty f_X(x) dx} \\ &\leq \frac{f_X(t)}{\int_t^\infty \frac{f_Y(x)f_X(t)}{f_Y(t)} dx} \\ &= \frac{f_Y(t)}{\int_t^\infty f_Y(x) dx} \\ &= \lambda_Y(t) \end{aligned}$$

となり、この性質が成り立つ。□

定義 5 確率密度関数 $f(x)$ を持つ連続な確率変数 X に対し $s \leq t$ なら

$$X_s \leq_l X_t$$

のとき、この確率変数は増加する尤度比を持つという。

補題3 確率変数 X が増加する尤度比を持つとき, $f(x)$ をこの確率変数の確率密度関数とすれば $\log f(x)$ は凸関数である。

証明 確率変数 X に対する X_t の定義から

$$\begin{aligned}\bar{F}_t(x) &= \Pr(X_t > x) \\ &= \frac{\bar{F}(t+x)}{\bar{F}(t)}\end{aligned}$$

だから, X_t の密度関数は $f_t(x) = \frac{f(t+x)}{\bar{F}(t)}$ となる。すべての $s \leq t$ に対し尤度比順序の定義から, $X_s \leq_t X_t$ が成り立つことは, $x \geq y$ のとき

$$\frac{f_{X_t}(x)}{f_{X_s}(x)} = \frac{\frac{f(t+x)}{\bar{F}(t)}}{\frac{f(s+x)}{\bar{F}(s)}} \geq \frac{\frac{f(t+y)}{\bar{F}(t)}}{\frac{f(s+y)}{\bar{F}(s)}} = \frac{f_{X_t}(y)}{f_{X_s}(y)}$$

すなわち

$$\frac{f(t+x)}{f(s+x)} \geq \frac{f(t+y)}{f(s+y)}$$

と同値である。このことは, 任意の $s \leq t$ と $y \leq x$ に対し

$$\log f(t+x) - \log f(s+x) \geq \log f(t+y) - \log f(s+y)$$

が成り立つことに等しい。よって $\log f(x)$ が凸関数となることが導かれる。□

補題4 確率変数 X が増加する尤度比を持つとき, この確率変数 X は IFR である。また, X が減少する尤度比を持つとき, X は DFR となる。

証明 $s \leq t$ のとき, $X_s \leq_t X_t$ なら定理2より $\lambda_{X_s} \geq \lambda_{X_t}$ となる。このことは $X_s \leq_h X_t$ に等しい。□

例5 正規分布の確率密度関数

$$f_k(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu_k)^2}{2\sigma^2}\right\}, \quad k=1, 2, \mu_1 \geq \mu_2,$$

を考える。 $X_k (k=1, 2)$ を確率密度関数 $f_k(x)$ をもつ確率密度関数とすると, $X_1 \geq_t X_2$ となることは簡単な計算から示される。

例6 確率分布が

$$P(X=n) = e^{-\lambda} \frac{(\lambda)^n}{n!}, \quad n=0, 1, 2, \dots$$

となる確率変数 X_λ , すなわちパラメータ λ のポアソン分布を考える。 $i < j$ のとき

$$\begin{aligned}\frac{f_{X_\mu}(i)}{f_{X_\lambda}(i)} \leq \frac{f_{X_\mu}(j)}{f_{X_\lambda}(j)} &\leftrightarrow \frac{e^{-\mu} \frac{\mu^i}{i!}}{e^{-\lambda} \frac{\lambda^i}{i!}} \leq \frac{e^{-\mu} \frac{\mu^j}{j!}}{e^{-\lambda} \frac{\lambda^j}{j!}} \\ &\leftrightarrow \left(\frac{\mu}{\lambda}\right)^i \leq \left(\frac{\mu}{\lambda}\right)^j\end{aligned}$$

$$\leftrightarrow 1 \leq \left(\frac{\mu}{\lambda}\right)^{j-i}$$

より, $\lambda \leq \mu$ のとき $X_\lambda \leq_i X_\mu$ となる。

例 7 例 2 で考えた自然数全体 $\{1, 2, \dots\}$ の上での確率分布全体 S に含まれる 2 つの確率分布 p と q に対し, $p \geq_i q$ は次のように定義できる。

定義 6 S に含まれる任意の p と q に対し, $p >_i q$ とは全ての i と j ($i \leq j$, $i, j = 1, 2, \dots$) に対し

$$p_j q_i \geq p_i q_j \quad \text{すなわち} \quad \begin{vmatrix} p_j & q_j \\ p_i & q_i \end{vmatrix} \geq 0 \quad (16)$$

が成立し, 少なくとも一つの i と j の組み合わせに対し

$$p_j q_i > p_i q_j,$$

が成り立つ場合とする。 $p_i = q_i$ が任意の $i = 1, 2, \dots$ に対し成り立つとき $p =_i q$ とする。 $p \geq_i q$ とは, $p =_i q$ で $p >_i q$ が成り立つときをいう。

尤度比順序の定義 (定義 4 と 6) は一般には次のように定義される。

定義 7 すべての半開区間 $(s, t]$, $(u, v]$ ($s \leq u$, $t \leq v$) に対し

$$\frac{\Pr(X \in (u, v])}{\Pr(Y \in (u, v])} \geq \frac{\Pr(X \in (s, t])}{\Pr(Y \in (s, t])} \quad (17)$$

または,

$$\left| \frac{\Pr(X \in (u, v])}{\Pr(X \in (u, v])} - \frac{\Pr(Y \in (s, t])}{\Pr(Y \in (s, t])} \right| \geq 0$$

のとき $X \geq_i Y$ と表す。

補題 5 確率変数 X と Y が連続な確率密度関数をもつとき, 定義 4 と 7 は同値である。

証明 定義 4 を仮定する。確率変数 X と Y の確率密度関数をそれぞれ $f_X(x)$ と $f_Y(x)$ とする。 $X \geq_i Y$ なら定義 7 より $x \geq y$ に対し

$$f_X(y)f_Y(x) \leq f_X(x)f_Y(y)$$

となる。 $s \leq t \leq u \leq v$ とすると

$$\begin{aligned} \Pr(X \in (u, v])\Pr(Y \in (s, t]) &= \int_u^v \int_s^t f_X(x)f_Y(y)dydx \\ &\geq \int_u^v \int_s^t f_X(y)f_Y(x)dydx \\ &= \Pr(X \in (s, t])\Pr(Y \in (u, v]) \end{aligned}$$

より, 定義 7 が導かれる。 $s \leq u \leq t \leq v$ ならば, 領域 (s, t) と (u, v) を, 3 つの領域 (s, u) , (u, t) と (t, v) に分ければ良いから, この補題は成り立つ。

同様にして, 逆も成り立つ。□

これらの 3 つの順序 (確率順序・故障率関数による順序・尤度比順序) には補題 4 と 5 から次の関

係がある。

定理 3 確率変数 X と Y に対し, $X \geq_l Y$ なら $X \geq_h Y$ となる。また $X \geq_h Y$ なら $X \geq_l Y$ となる。

つぎの 2 項関係を定義する。

仮定 1 ある集合 A の要素 $a, b, c, \dots$ の間に 2 項関係 $<$ が定義され, つぎの 3 つの条件を満足するとき, これを半順序 (Partial Order) と呼ぶ。

- (1) $a < a$ (反射法則)
- (2) $a < b$ かつ $b < c$ のとき $a < c$ となる。(推移法則)
- (3) $a < b$ かつ $b < a$ のとき $a = b$ となる。(反対称法則)

また, $b < a$ のとき, $a > b$ と表す。この 2 項関係は全ての要素間に成り立たなくても良い。 A に含まれるどの要素の間にも, $<$, $>$ または $=$ が成り立つとき全順序と呼ぶ。

補題 6 定義 1, 3 と 7 で定義した順序は半順序となる。

この性質の証明は簡単なので省略する。

例 8 状態空間が $\{i | i=0, 1, 2, 3, \dots\}$ のマルコフ連鎖を考える。推移確率行列を $\{p_{ij}\}_{i,j=0,1,2,3,\dots}$ とするとき, 尤度比順序を用いて次の関係を考える。任意の i, j ($i \geq j, i, j=0, 1, 2, \dots$) に対し

$$p_{mj}p_{ni} \geq p_{nj}p_{mi} \quad \text{すなわち} \quad \begin{vmatrix} p_{ni} & p_{nj} \\ p_{mi} & p_{mj} \end{vmatrix} \geq 0 \quad (18)$$

が $m \leq n$ ($m, n=0, 1, 2, \dots$) を満たす全ての m と n に対し成り立つ。

このマルコフ連鎖は取り替え問題などで劣化のある場合に用いられる。

5. 関数類による特徴づけ

これまでの確率的順序関係を別の角度から考える。Stoyan[20], Kijima and Ohnishi[5], Shaked and Shantikumar [18] など で扱われ, それらの結果をまとめて説明する。確率変数 X と Y の分布関数をそれぞれ $F_X(x)$ および $F_Y(x)$ とし, これらの確率変数は確率密度関数 $f_X(x)$ と $f_Y(x)$ を持つとする。一般の場合も同様の議論ができ前記の書物または論文に詳しい。

定義 8 G を $\mathbf{R}=(-\infty, \infty)$ 上の関数の類 (Class) とする。2 つの確率変数 X と Y の間に G で生成された半順序の関係 $X > Y$ が成り立つとは, 全ての $g \in G$ に対し

$$\mathbf{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx \geq \int_{-\infty}^{\infty} g(x)f_Y(x)dx = \mathbf{E}[g(Y)] \quad (19)$$

となることを言い, この関係を $x >_G Y$ と表す。ここでは積分が定義されている場合を考える。

このとき

$$G_a = \{g | g(x) \text{ は } x \text{ に関する非減少関数}\}$$

のときの半順序を $>_a$ で表せば, 定理 1 からこの順序は確率順序となる。すなわち, 2 つの確率変数 X と Y に対し $X >_a Y$ と $X \geq_s Y$ は同値である。

H を $\mathbf{R}$ に含まれる区間からなる類とする。確率変数 X と区間 D に対し、 $\Pr(X \in D) > 0$ のとき D 上の条件つき確率変数を $X(D)$ で表す。確率変数 X の密度関数を $f_X(x)$ としたとき、確率変数 $X(D)$ の密度関数は $x(\in D)$ に対し $\frac{f_X(x)}{\Pr(X \in D)}$ となる。

定理 4 X と Y を 2 つの確率変数とし

$$H = \{(s, t] | s < t, s, t \in \mathbf{R}\} \quad (20)$$

に含まれる任意の D に対し $X(D) >_a Y(D)$ となることと $X \geq_t Y$ は等しい。すなわち尤度比による順序となる。

証明 任意の実数 a に対し、次のことに注意する。任意の $D \in H$ に対し

$$\begin{aligned} X(D) >_a Y(D) &\leftrightarrow X(D) \geq_s Y(D) \\ &\leftrightarrow \Pr(X(D) \geq a) \geq \Pr(Y(D) \geq a) \\ &\leftrightarrow \frac{\int_{D \cap (a, \infty)} f_X(x) dx}{\Pr(X \in D)} \geq \frac{\int_{D \cap (a, \infty)} f_Y(x) dx}{\Pr(Y \in D)} \end{aligned} \quad (21)$$

となる。ここで、任意の $D \in H$ とし $D \cap (a, \infty) \in H$ となるから、以下の議論では、 $D \cap (a, \infty)$ の代わりに D で考える。

$s \leq t \leq u \leq v$ の場合を考える。 $D = (s, u]$ のとき (21) 式から

$$\int_D f_X(x) dx = \int_t^u f_X(x) dx = \Pr(X \in (t, u])$$

だから

$$\frac{\Pr(X \in (t, u])}{\Pr(X \in (s, u])} \geq \frac{\Pr(Y \in (t, u])}{\Pr(Y \in (s, u])} \quad (22)$$

となる。(22) 式を整理すれば

$$\begin{aligned} &\Pr(X \in (t, u]) (\Pr(Y \in (s, t]) + \Pr(Y \in (t, u])) \\ &\geq \Pr(Y \in (t, u]) (\Pr(X \in (s, t]) + \Pr(X \in (t, u])) \end{aligned}$$

より

$$\Pr(X \in (t, u]) \Pr(Y \in (s, t]) \geq \Pr(Y \in (t, u]) \Pr(X \in (s, t]) \geq 0 \quad (23)$$

となる。 $D = (t, u]$ も同様にして

$$\Pr(X \in (u, v]) \Pr(Y \in (t, u]) \geq \Pr(Y \in (u, v]) \Pr(X \in (t, u]) \geq 0 \quad (24)$$

となる。(23) 式と (24) 式から

$$\begin{aligned} &\Pr(X \in (t, u]) \Pr(Y \in (s, t]) \Pr(X \in (u, v]) \Pr(Y \in (t, u]) \\ &\geq \Pr(Y \in (t, u]) \Pr(X \in (s, t]) \Pr(Y \in (u, v]) \Pr(X \in (t, u]) \end{aligned}$$

となり

$$\Pr(Y \in (s, t]) \Pr(X \in (u, v]) \geq \Pr(X \in (s, t]) \Pr(Y \in (u, v]) \quad (25)$$

より (17) 式が導かれる。

$s \leq u \leq t \leq v$ の場合を考える。

$$\Pr(X \in (s, t]) = \Pr(X \in (s, u]) + \Pr(Y \in (u, t])$$

$$\Pr(Y \in (u, v]) = \Pr(Y \in (u, t]) + \Pr(Y \in (t, v])$$

で、3つの区間 $(s, u]$, $(u, t]$, $(t, v]$ のそれぞれの組み合わせに対し、(25)式と同様に

$$\Pr(Y \in (s, u])\Pr(X \in (u, t]) \geq \Pr(X \in (s, u])\Pr(Y \in (u, t])$$

などの不等式が成り立つから

$$\begin{aligned} \Pr(X \in (s, t])\Pr(Y \in (u, v]) &= \Pr(X \in (s, u])\Pr(Y \in (u, t]) + \Pr(X \in (s, u])\Pr(Y \in (t, v]) \\ &\quad + \Pr(X \in (u, t])\Pr(Y \in (u, t]) + \Pr(X \in (u, t])\Pr(Y \in (t, v]) \\ &\leq \Pr(Y \in (s, u])\Pr(X \in (u, t]) + \Pr(Y \in (s, u])\Pr(X \in (t, v]) \\ &\quad + \Pr(Y \in (u, t])\Pr(X \in (u, t]) + \Pr(Y \in (u, t])\Pr(X \in (t, v]) \\ &= \Pr(X \in (u, v])\Pr(Y \in (s, t]) \end{aligned}$$

より(17)式が導かれ、この順序は尤度比順序となる。逆もまた成り立つ。□

定理 5 X と Y を2つの確率変数とし

$$H = \{(t, \infty) | t \in \mathbf{R}\} \quad (26)$$

に含まれる任意の D に対し $X(D) >_d Y(D)$ となることと $X \geq_h Y$ は同値である。すなわち故障率関数を用いた順序となる。

証明 集合 H の定義から $D = (t, \infty)$ とおく。(21)式で

$$\int_D f_X(x) dx = F_X(t)$$

となることに注意する。定理4と同様に任意の $D \in H$ に対し $D \cap (a, \infty) \in H$ だから、以下の議論では $D \cap (a, \infty)$ の代わりに D で考える。よって、 $t \leq a$ なら

$$\frac{\Pr(X \geq a)}{\Pr(X \geq t)} \geq \frac{\Pr(Y \geq a)}{\Pr(Y \geq t)}$$

より

$$\Pr(X \geq a)\Pr(Y \geq t) \geq \Pr(Y \geq a)\Pr(X \geq t) \quad (27)$$

となる。(27)式で $a = t + h$ ($h > 0$) とおくと

$$\Pr(X \geq t + h)\Pr(Y \geq t) \geq \Pr(Y \geq t + h)\Pr(X \geq t)$$

より

$$\frac{\Pr(X \geq t + h) - \Pr(X \geq t)}{h} \Pr(Y \geq t) \geq \frac{\Pr(Y \geq t + h) - \Pr(Y \geq t)}{h} \Pr(X \geq t) \quad (28)$$

となる。

$$\frac{\Pr(X \geq t + h) - \Pr(X \geq t)}{h} = \frac{\int_t^{t+h} f_X(x) dx}{h}$$

に注意して、(28)式で $h \rightarrow 0$ とすれば

$$f_X(t)\Pr(Y \geq t) \geq f_Y(t)\Pr(X \geq t)$$

となり、

$$\frac{f_X(t)}{\Pr(X \geq t)} \geq \frac{f_Y(t)}{\Pr(Y \geq t)}$$

が成り立つ。よって、この順序は故障率による順序となる。逆もまた成り立つ。□

関数類の中で、応用上重要と考えられるものは

$$G_{ix} = \{g | g(x) \text{ は } x \text{ に関する非減少凸関数 (convex)}\}$$

で、この関数類で定義される順序を $>_{ix}$ と表す。特に関数類

$$G_x = \{g | g(x) \text{ は } x \text{ に関する凸関数}\}$$

で定義される順序 $>_x$ と表し、この順序を凸順序という。

$X >_{ix} Y$ のとき、定義より x に関する任意の非減少凸関数 $g(x)$ に対し

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \geq \int_{-\infty}^{\infty} g(x) f_Y(x) dx = E[g(Y)] \quad (29)$$

が成り立つことに等しい。

補題 7 X と Y を 2 つの非負確率変数とする。 $X >_{ix} Y$ となることと、 $\int_x^{\infty} \bar{F}_X(t) dt \geq \int_x^{\infty} \bar{F}_Y(t) dt$ が全ての $x \geq 0$ に対し成り立つことは同値である。

証明 つぎの関数 $J_x(t)$ を考える。

$$J_x(t) = (t-x)^+ = \max\{t-x, 0\} = \int_{-\infty}^t I_x(u) du \quad (30)$$

$J_x(t) \in G_{ix}$ で

$$\begin{aligned} \int_{-\infty}^{\infty} J_x(t) f_X(t) dt &= \int_{-\infty}^{\infty} (t-x)^+ f_X(t) dt \\ &= \int_x^{\infty} (t-x) f_X(t) dt \\ &= \int_x^{\infty} (1-F_X(t)) dt \\ &= \int_x^{\infty} \bar{F}_X(t) dt \end{aligned} \quad (31)$$

だから、(29)式より求める性質が成り立つ。

逆に、 G_{ix} に含まれる関数 $g(t)$ で $\lim_{t \rightarrow -\infty} g(t) = 0$ となるものを考える。 $g(x)$ は非減少凸関数だから $\lim_{t \rightarrow -\infty} h(t) = 0$ となる関数 $h(t)$ を用いて

$$g(t) = \int_{-\infty}^{\infty} J_x(t) dh(x) = \int_{-\infty}^t h(x) dx \quad (32)$$

と表せる。(31)式から

$$\begin{aligned} \int_{-\infty}^{\infty} g(t) f_X(t) dt &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} J_x(t) dh(x) \right) f_X(t) dt \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} J_x(t) f_X(t) dt \right) dh(x) \\ &= \int_{-\infty}^{\infty} \left(\int_x^{\infty} \bar{F}_X(t) dx \right) dh(x) \end{aligned} \quad (33)$$

より、(29)式が成り立つ。□

$G_{iv} = \{g|g(x) \text{ は } x \text{ に関する非減少凹関数 (concave)}\}$

のとき、この関数族で定義される順序を $>_{iv}$ と表す。特に、関数族

$G_v = \{g|g(x) \text{ は } x \text{ に関する凹関数}\}$

で定義される順序を凹順序 (Concave Order) と呼び $>_v$ と表す。

$X >_{ix} Y$ のとき、定義より x に関する任意の非減少凸関数 $g(x)$ に対し

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx \leq \int_{-\infty}^{\infty} g(x)f_Y(x)dx = E[g(Y)] \quad (34)$$

となる。補題7と同様につきの性質が成り立つ。

補題8 X と Y を2つの非負確率変数とする。 $X >_{ix} Y$ となることと、 $\int_x^\infty F_X(t)dt \leq \int_x^\infty F_Y(t)dt$ が全ての $x \geq 0$ に対し成り立つことは同値である。

この関数 $f(x)$ を効用関数とすれば、これら2つの順序は確率順序が一般的な選好関係を表し、凸順序は危険愛好的な (Risk Share) 選好を、凹順序は危険回避的な (Risk Averse) 選好を表す。

注2 Stoyan [20] では、この節で考えたような確率的順序関係 $>$ の持つ性質と考えられる次のものを挙げている。

- (1) $X > Y$ のとき、任意の確率変数 Z に対し $X+Z > Y+Z$ となる。
- (2) $a, b \in \mathbf{R}$ で $a \geq b$ のとき $a > b$ である。ここで、 a は確率1で値 a をとる確率変数とする。
- (3) $X > Y$ のとき、任意の実数 a に対し $aX > aY$ となる。
- (4) $X > Y$ のとき、 $E[X] \geq E[Y]$ となる。

(1)はたたみこみ (Convolution) で順序が保存されることを示している。ここで考えた順序に関し、つぎの性質を持つ。

補題9

- (1) $>_d$, $>_{ix}$ と $>_{iv}$ は、(1), (2), (3)と(4)の性質を持つ。
- (2) $>_x$ と $>_v$ は(1), (3)と(4)の性質を持つ。

注3 定理3から $X \geq_l Y$ なら $X \geq_h Y$ であり、 $X \geq_h Y$ なら $X >_d Y$ だから、これらの順序(尤度比と故障率による順序)は(1), (2), (3)と(4)の性質を持つ。

6. 部分観測可能なマルコフ連鎖

状態空間が可算集合 $\{0, 1, 2, \dots\}$ のマルコフ連鎖で、その状態を直接知ることは出来ないが、情報過程を通して情報が得られるとき、これを部分観測可能なマルコフ連鎖と呼ぶ。この部分観測可能なマルコフ連鎖の基本的な性質は Nakai [7, 8] などでは考えられている。Nakai [10] では部分観測可能なマルコフ連鎖のもとでの決定問題が扱われている。

状態空間が $\{0, 1, 2, 3, \dots\}$ で推移確率行列が $\mathbf{P} = (p_{ij})_{i,j=0,1,2,\dots}$ のマルコフ連鎖を考える。このマルコ

フ連鎖の状態は観測できないが、それらの状態に依存する確率変数を観測し情報を得る。このプロセスを情報過程と呼び、マルコフ連鎖の状態が i のときに観測できる非負確率変数を X_i とおけば、この確率変数は

$$\Pr\{X_i \leq x | Y_t = i\} = F_i(x) \quad (x \in \mathbb{R}, i, t \in \{0, 1, 2, \dots\}), \quad (35)$$

で定まり、これらの分布関数 $F_i(x)$ は確率密度関数 $f_i(x)$ を持つとする。ここで Y_t は時刻 t でのマルコフ連鎖の状態を表す確率変数とする。一般の場合にも以下と同様の議論ができるが、ここでは省略する。

部分観測可能なマルコフ連鎖の状態についての情報は、状態空間上の確率分布 $\mathbf{p} = (p_0, p_1, p_2, \dots)$ ($p_i \geq 0, \sum_{i=0}^{\infty} p_i = 1$) で表され、それ全体 S は例の 2 の (11) 式のように定義される。この集合に例 7 で考えた尤度比順序を仮定する。

このマルコフ連鎖の状態についての学習プロセスを考える。情報過程から得られた値 x と事前分布 $\mathbf{p}$ に対し、事後分布が存在し学習プロセスはベイズの定理を用いる。

部分観測可能なマルコフ連鎖の状態に関する事前情報が $\mathbf{p}$ のとき、値 $x \in R(0, \infty)$ を観測したとする。この値に対し、マルコフ連鎖の状態の事後情報を $\overline{T(\mathbf{p}, x)}$ とする。ベイズの定理から事後分布 $T(\mathbf{p}, x)$ はすべての $j=0, 1, 2, \dots$ に対し

$$\begin{cases} T_j(\mathbf{p}, x) = \frac{p_j f_j(x)}{\sum_{i=0}^{\infty} p_i f_i(x)} \\ T(\mathbf{p}, x) = (T_0(\mathbf{p}, x), T_1(\mathbf{p}, x), T_2(\mathbf{p}, x), \dots) \end{cases} \quad (36)$$

となる。次に、このマルコフ連鎖は推移確率行列 $\mathbf{P}$ にしたがって状態が推移し、マルコフ連鎖の状態についての新しい情報 $\overline{T(\mathbf{p}, x)}$ は

$$\begin{cases} \overline{T_j(\mathbf{p}, x)} = \sum_{i=0}^{\infty} T_i(\mathbf{p}, x) p_{ij}, \\ \overline{T(\mathbf{p}, x)} = (\overline{T_0(\mathbf{p}, x)}, \overline{T_1(\mathbf{p}, x)}, \overline{T_2(\mathbf{p}, x)}, \dots) \end{cases} \quad (37)$$

となる。

確率分布関数と推移確率行列 $\mathbf{P}$ に 2 つの仮定を設け、事前情報と事後情報の関係を考える。

仮定 2 確率変数 $\{X_i\}_{i=0,1,2,3,\dots}$ に対し、 $i \leq j$ ($i, j=0, 1, 2, \dots, n$) なら $X_i \geq_t X_j$ とする。すなわち $x \leq y$ のとき

$$f_j(x) f_i(y) \leq f_j(y) f_i(x) \quad \text{すなわち} \quad \begin{vmatrix} f_i(x) & f_i(y) \\ f_j(x) & f_j(y) \end{vmatrix} \geq 0 \quad (38)$$

とする。

仮定 3 推移確率行列 $\mathbf{P}$ に例 8 で考えた尤度比順序を仮定する。

注 4 状態空間が $\{1, 2\}$ の 2 状態空間なら、この仮定は不等式

$$p_{11} \geq p_{21},$$

と等しい。この仮定は Ross [16], Monahan [6] など で用いられ、仮定 3 はこの一般化となる。

情報全体の集合 S 上の関数 $u(\mathbf{p})$ が $\mathbf{p} \geq_i \mathbf{q}$ を満たす $\mathbf{p}$ と $\mathbf{q}$ に対し $u(\mathbf{p}) \geq u(\mathbf{q})$ のとき、この関数を $\mathbf{p}$ に関する非減少関数と呼ぶ。

この学習プロセスに関し、仮定2と3のもとで(36)式と(37)式で与えられた事後情報の性質を考える。

補題10 $x \leq y$ を満たす任意の x と y に対し

$$T(\mathbf{p}, x) \leq_i T(\mathbf{p}, y)$$

となる。($\mathbf{p} \in S$)

証明 S での順序の定義から、任意の $j < i$ ($i, j = 0, 1, 2, \dots$) に対し

$$T_j(\mathbf{p}, x) T_i(\mathbf{p}, y) \leq T_i(\mathbf{p}, x) T_j(\mathbf{p}, y)$$

が示されれば補題が得られる。そのため、

$$T_j(\mathbf{p}, x) T_i(\mathbf{p}, y) - T_i(\mathbf{p}, x) T_j(\mathbf{p}, y)$$

を考える。この式の分母を払えば

$$p_j f_j(x) p_i f_i(y) - p_i f_i(x) p_j f_j(y) = p_j p_i (f_j(x) f_i(y) - f_i(x) f_j(y)) \leq 0$$

となる。ここで、最後の不等式は仮定2から得られる。□

定理6 $x \leq y$ を満たす任意の x と y に対し

$$\overline{T(\mathbf{p}, x)} \leq_i \overline{T(\mathbf{p}, y)}$$

となる。($\mathbf{p} \in S$)

証明 任意の $j < i$ ($i, j = 0, 1, 2, \dots$) に対し

$$\overline{T_j(\mathbf{p}, x) T_i(\mathbf{p}, y)} \leq \overline{T_i(\mathbf{p}, x) T_j(\mathbf{p}, y)}$$

が示されれば定理が成り立つ。いま

$$\overline{T_j(\mathbf{p}, x) T_i(\mathbf{p}, y)} - \overline{T_i(\mathbf{p}, x) T_j(\mathbf{p}, y)}$$

を考える。分母を払うと

$$\begin{aligned} & \left(\sum_{l=0}^{\infty} T_l(\mathbf{p}, x) p_{li} \right) \left(\sum_{k=0}^{\infty} T_k(\mathbf{p}, y) p_{ki} \right) - \left(\sum_{l=0}^{\infty} T_l(\mathbf{p}, x) p_{li} \right) \left(\sum_{k=0}^{\infty} T_k(\mathbf{p}, y) p_{kj} \right) \\ &= \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} T_l(\mathbf{p}, x) T_k(\mathbf{p}, y) (p_{li} p_{ki} - p_{li} p_{kj}) \\ &= \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} (T_l(\mathbf{p}, x) T_k(\mathbf{p}, y) - T_k(\mathbf{p}, x) T_l(\mathbf{p}, y)) (p_{li} p_{ki} - p_{li} p_{kj}) \leq 0 \end{aligned}$$

となり、補題10と仮定3からこの定理は導かれる。□

補題11 S に含まれる任意の $\mathbf{p}, \mathbf{q}$ に対し、 $\mathbf{p} \geq_i \mathbf{q}$ なら

$$T(\mathbf{p}, x) \geq_i T(\mathbf{q}, x)$$

となる。($x \in \mathbf{R}$)

証明 任意の $j < i$ ($i, j = 0, 1, 2, \dots$) に対し

$$T_i(\mathbf{p}, x) T_j(\mathbf{q}, x) \leq T_j(\mathbf{p}, x) T_i(\mathbf{q}, x)$$

が示されれば良いから、補題10と同じように

$$T_j(\mathbf{p}, x) T_i(\mathbf{q}, x) - T_i(\mathbf{p}, x) T_j(\mathbf{q}, x)$$

を考える。この分母を払えば

$$p_j f_j(x) q_i f_i(x) - p_i f_i(x) q_j f_j(x) = f_j(x) f_i(x) (p_j q_i - p_i q_j) \geq 0$$

となる。ここで、最後の不等式は情報全体の集合に導入した尤度比順序の定義（定義1）から導かれる。□

定理1 S に含まれる任意の $\mathbf{p}, \mathbf{q}$ に対し, $\mathbf{p} \geq_i \mathbf{q}$ なら

$$\overline{T(\mathbf{p}, x)} \geq_i \overline{T(\mathbf{q}, x)}$$

となる。($x \in \mathbf{R}$)

証明 任意の $j < i$ ($i, j = 0, 1, 2, \dots$) に対し

$$\overline{T_j(\mathbf{p}, x)} \overline{T_i(\mathbf{q}, x)} \leq \overline{T_i(\mathbf{p}, x)} \overline{T_j(\mathbf{q}, x)}$$

が示されれば良いから,

$$\overline{T_j(\mathbf{p}, x)} \overline{T_i(\mathbf{q}, x)} - \overline{T_i(\mathbf{p}, x)} \overline{T_j(\mathbf{q}, x)}$$

を考える。補題10より定理3と同様に求められる。□

つぎの補題は、部分観測可能なマルコフ連鎖の上での多段決定問題を考えるときに必要となる。

補題12 $\{f_i(x)\}_{i=0,1,2,\dots}$ を確率密度関数の有限列とする。 $\mathbf{p}$ と $\mathbf{q}$ を状態空間上の確率分布で $\mathbf{p} \geq_i \mathbf{q}$ とし, $\{a_i\}_{i=0,1,2,\dots}$ を $a_i = p_i - q_i$ とする。

$$g(x) = \sum_{i=0}^{\infty} a_i f_i(x)$$

とおくと、任意の非減少関数 $h(x)$ に対し

$$\int_0^{\infty} h(x) g(x) dx \geq 0$$

となる。

証明 定理3より $\mathbf{p} \geq_i \mathbf{q}$ なら $\mathbf{p} \geq_s \mathbf{q}$ となる。定理1から

$$\int_0^{\infty} h(x) \sum_{i=0}^{\infty} p_i f_i(x) dx \geq \int_0^{\infty} h(x) \sum_{i=0}^{\infty} q_i f_i(x) dx \geq 0$$

となるから,

$$\int_0^{\infty} h(x) g(x) dx \geq 0$$

が導かれ、この性質が成り立つ。□

補題13 関数 $u(\mathbf{p})$ が $\mathbf{p}$ に関する非減少凸関数なら

$$u^*(\mathbf{p}) = \int_0^{\infty} u(\overline{T(\mathbf{p}, x)}) dF_p(x)$$

も $\mathbf{p}$ に関する非減少凸関数となる。

証明 凸性は Aström [1] と同様に求められる。ここでは、非減少性を示す。

$\mathbf{p}, \mathbf{q} \in S$ で $\mathbf{p} \geq_i \mathbf{q}$ のとき

$$\begin{aligned} u^*(\mathbf{p}) - u^*(\mathbf{q}) &= \int_0^{\infty} u(\overline{T(\mathbf{p}, x)}) dF_p(x) - \int_0^{\infty} u(\overline{T(\mathbf{q}, x)}) dF_q(x) \\ &\geq \int_0^{\infty} u(\overline{T(\mathbf{q}, x)}) d(F_p(x) - F_q(x)) \end{aligned}$$

$$= \int_0^\infty u(\overline{T(\mathbf{q}, x)}) \sum_{i=0}^\infty (p_i - q_i) dF_i(x)$$

となる。ここで不等式は定理7から得られる。つぎに尤度比による順序の定義から不等式

$$\frac{p_k}{q_k} > 1 \geq \frac{p_{k+1}}{q_{k+1}}$$

を満足する $1 < k < n$ が存在する。

$$dF_i(x) = f_i(x) dx$$

より

$$g(x) = \sum_{i=1}^\infty (p_i - q_i) f_i(x)$$

とおく。定理6から $\overline{T(\mathbf{q}, x)}$ は x に関する非減少関数で、 $u(\overline{T(\mathbf{p}, x)})$ は $\mathbf{p}$ の非減少関数だから、補題13から不等式

$$\int_0^\infty u(\overline{T(\mathbf{q}, x)}) g(x) dx \geq 0$$

が成り立ち、この補題が成立する。□

7. 複数の値を観測する場合の学習プロセス

6節の部分観測可能なマルコフ連鎖の状態に依存する確率変数の列を観測し、一度に複数の観測値が得られるときに、事前情報と事後情報の関係を考える。ここでも前節と同様の仮定と記号を用いる。

部分観測可能なマルコフ連鎖の状態 i に依存した n 個の確率変数 $\{X_k\}_{k=1, \dots, n}$ から値 $\{x_k\}_{k=1, \dots, n}$ ($x_k \in R = (0, \infty)$, $k=1, \dots, n$) を観測したとき、これらの値から学習を行いその結果を $\overline{T(\mathbf{p}, x)}$ と表す。マルコフ連鎖の状態についての事後情報はベイズの定理を用いて(36)式と(37)式と同様に、(39)式と(41)式から求められる。ベイズの定理により事後確率分布は

$$\begin{cases} T_j(\mathbf{p}, \mathbf{x}) = \frac{p_j f_j(\mathbf{x})}{\sum_{i=0}^\infty p_i f_i(\mathbf{x})} \\ T(\mathbf{p}, \mathbf{x}) = (T_0(\mathbf{p}, \mathbf{x}), T_1(\mathbf{p}, \mathbf{x}), T_2(\mathbf{p}, \mathbf{x}), \dots) \end{cases} \quad (39)$$

となる。 $(j=0, 1, 2, \dots)$ ここでは、 $\mathbf{x} = (x_1, \dots, x_n)$ と表す。また、 $f_j(\mathbf{x})$ はマルコフ連鎖の状態が j ($j=0, 1, 2, \dots$) のとき n 個の確率変数 $\mathbf{X}$ の同時分布関数とし

$$f_j(\mathbf{x}) = \prod_{k=1}^n f_j(x_k) \quad (40)$$

である。次に、推移確率行列 $\mathbf{P}$ に従って状態が推移し、このマルコフ連鎖の状態についての事後情報 $\overline{T(\mathbf{p}, \mathbf{x})}$ は

$$\begin{cases} \overline{T_j(\mathbf{p}, \mathbf{x})} = \sum_{i=0}^\infty T_i(\mathbf{p}, \mathbf{x}) p_{ij}, \\ \overline{T(\mathbf{p}, \mathbf{x})} = (\overline{T_0(\mathbf{p}, \mathbf{x})}, \overline{T_1(\mathbf{p}, \mathbf{x})}, \overline{T_2(\mathbf{p}, \mathbf{x})}, \dots) \end{cases} \quad (41)$$

となる。

ここで、観測値の組に対しつぎの順序を考える。

定義 9 2つの k 個の観測値の組 $\mathbf{x}, \mathbf{y} \in R^n$ に対し $x_{(i)} \leq y_{(i)} (i=1, 2, \dots, k)$ のとき、またそのときに限り $\mathbf{x} < \mathbf{y}$ と表す。

補題 14 $\mathbf{x} < \mathbf{y}$ を満たす任意の $\mathbf{x}$ と $\mathbf{y}$ に対し

$$f_j(\mathbf{y})f_i(\mathbf{x}) \geq f_i(\mathbf{y})f_j(\mathbf{x}) \quad \text{すなわち} \quad \begin{vmatrix} f_i(\mathbf{x}) & f_j(\mathbf{x}) \\ f_i(\mathbf{y}) & f_j(\mathbf{y}) \end{vmatrix} \geq 0 \quad (42)$$

($j < i, i, j=1, 2, \dots$) となる。

証明 n に関する帰納法を用いる。 $n=1$ の場合は仮定 2 より明かである。

$n-1$ までを仮定する。マルコフ連鎖の状態が i のとき、確率変数 $X_1, \dots, X_n$ は互いに独立だから $X_1, \dots, X_n$ と $X_{(1)}, \dots, X_{(n)}$ の同時確率密度関数は等しい。したがって $\mathbf{x}=(x_1, \dots, x_n) (x_1 \geq \dots \geq x_n)$ と $\mathbf{y}=(y_1, \dots, y_n) (y_1 \geq \dots \geq y_n)$ とする。

$\mathbf{x} < \mathbf{y}$ で 1 つの要素を除きすべて等しいとする。 $x_1 \leq y_1, x_k = y_k (k=2, \dots, n)$ のとき、仮定 2 より

$$f_j(\mathbf{y})f_i(\mathbf{x}) - f_i(\mathbf{y})f_j(\mathbf{x}) = \left\{ \prod_{k=2}^n f_j(x_k)f_i(x_k) \right\} (f_j(y_1)f_i(x_1) - f_i(y_1)f_j(x_1)) \geq 0$$

となる。 $x_k \leq y_k, x_l = y_l (l=1, \dots, k-1, k+1, \dots, n)$ の場合も同様に成り立つ。

$x_1 \leq y_1, x_2 \leq y_2, x_k = y_k (k=3, \dots, n)$ とする。次の 2 つの場合を考える。

1) $y_1 > x_1 > y_2 > x_2 > x_3 = y_3 = \dots = x_n = y_n$

2) $y_1 > y_2 > x_1 > x_2 > x_3 = y_3 = \dots = x_n = y_n$

1) の場合は、

$$\mathbf{w} = (y_1, x_2, \dots, x_n)$$

と考えると、 $\mathbf{x} < \mathbf{w} < \mathbf{y}$ となり

$$f_j(\mathbf{w})f_i(\mathbf{x}) \geq f_i(\mathbf{w})f_j(\mathbf{x})$$

$$f_j(\mathbf{y})f_i(\mathbf{w}) \geq f_i(\mathbf{y})f_j(\mathbf{w})$$

となる。これらの不等式を整理すれば、それぞれ

$$\frac{f_j(\mathbf{w})}{f_i(\mathbf{w})} \geq \frac{f_i(\mathbf{x})}{f_j(\mathbf{x})},$$

$$\frac{f_j(\mathbf{y})}{f_i(\mathbf{y})} \geq \frac{f_i(\mathbf{w})}{f_j(\mathbf{w})}$$

となり、

$$\frac{f_j(\mathbf{y})}{f_i(\mathbf{y})} \geq \frac{f_i(\mathbf{x})}{f_j(\mathbf{x})}$$

が成り立つ。よって、

$$f_j(\mathbf{y})f_i(\mathbf{x}) \geq f_i(\mathbf{y})f_j(\mathbf{x})$$

となる。

2) の場合も同様に

$$w = (x_1, y_2, \dots, x_n)$$

を考える。 $x_k \leq y_k, x_l \leq y_l, x_k = y_k (k \neq l)$ の場合も同様に成り立つ。

一般の場合は

$$1) y_1 > x_1 > y_2 > x_2$$

$$2) y_1 > y_2 > x_1 > x_2$$

の2つの場合が考えられる。1) の場合は、

$$w = (y_1, x_2, \dots, x_n)$$

を考えると $x < w < y$ となる。

つぎに、 $x' = (x_1, x_2, \dots, x_{n-1})$ と $w' = (y_1, x_2, \dots, x_{n-1})$ を考える。仮定から $x' < w'$ は明かだから、帰納法の仮定より

$$f_j(w') f_i(x') \geq f_i(w'') f_j(x') \quad (j < i)$$

となる。したがって簡単な計算から

$$f_j(w) f_i(x) \geq f_i(w) f_j(x) \quad (j < i)$$

が成り立つ。

$w'' = (x_2, \dots, x_n)$ と $y' = (y_2, \dots, y_n)$ を考えれば同様に $w'' < y'$ だから、帰納法の仮定より

$$f_j(y') f_i(w'') \geq f_i(y') f_j(w'') \quad (j < i)$$

が成り立つ。簡単な計算から

$$f_j(y) f_i(w) \geq f_i(y) f_j(w) \quad (j < i)$$

が成り立ち1)の場合にこの補題が成り立つ。

2) の場合も同様に成り立つ。□

補題15 $x < y$ を満たす任意の x と y に対し $T(p, x) \leq_l T(p, y)$ が全ての $p \in S$ に対し成り立つ。

証明 S での順序の定義から、任意の $l < k (k, l = 1, 2, \dots)$ に対し

$$T_l(p, x) T_k(p, y) \leq T_k(p, x) T_l(p, y)$$

が示されればこの補題が成り立つ。したがって

$$T_l(p, x) T_k(p, y) - T_k(p, x) T_l(p, y)$$

を考える。この式の分母を払えば、

$$\begin{aligned} & p_l f_l(x) p_k f_k(y) - p_k f_k(x) p_l f_l(y) \\ &= p_l p_k (f_l(x) f_k(y) - f_k(x) f_l(y)) \leq 0 \end{aligned}$$

となる。ここで、最後の不等式は補題14より求められる。□

定理8 $x < y$ を満たす任意の x と y に対し $\overline{T(p, x)} \leq_l \overline{T(p, y)}$ となる。 ($p \in S$)

証明 S の順序の定義から、任意の $l < k (k, l = 1, 2, \dots)$ に対し

$$\overline{T_l(p, x) T_k(p, y)} \leq \overline{T_k(p, x) T_l(p, y)}$$

が示されればこの定理が成り立つ。したがって

$$\overline{T_l(\mathbf{p}, \mathbf{x}) T_k(\mathbf{p}, \mathbf{y})} - \overline{T_k(\mathbf{p}, \mathbf{x}) T_l(\mathbf{p}, \mathbf{y})}$$

を考える。この値は

$$\begin{aligned} & \left(\sum_{l=0}^{\infty} T_l(\mathbf{p}, \mathbf{x}) p_{li} \right) \left(\sum_{k=0}^{\infty} T_k(\mathbf{p}, \mathbf{y}) p_{ki} \right) - \left(\sum_{l=0}^{\infty} T_l(\mathbf{p}, \mathbf{x}) p_{li} \right) \left(\sum_{k=0}^{\infty} T_k(\mathbf{p}, \mathbf{y}) p_{kj} \right) \\ &= \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} T_l(\mathbf{p}, \mathbf{x}) T_k(\mathbf{p}, \mathbf{y}) (p_{li} p_{ki} - p_{li} p_{kj}) \\ &= \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} (T_l(\mathbf{p}, \mathbf{x}) T_k(\mathbf{p}, \mathbf{y}) - T_k(\mathbf{p}, \mathbf{x}) T_l(\mathbf{p}, \mathbf{y})) (p_{li} p_{ki} - p_{li} p_{kj}) \leq 0 \end{aligned}$$

となり、この不等式は補題15と推移確率行列に関する仮定2から導かれる。□

補題16 任意の $\mathbf{p}, \mathbf{q}$ に対し、 $\mathbf{p} \geq_l \mathbf{q}$ なら $T(\mathbf{p}, \mathbf{x}) \geq_l T(\mathbf{q}, \mathbf{x})$ である。 ($\mathbf{x} \in R^n$)

証明 S の順序の定義から、任意の $l < k$ ($k, l = 1, 2, \dots$) に対し

$$T_k(\mathbf{p}, \mathbf{x}) T_l(\mathbf{q}, \mathbf{x}) \leq T_l(\mathbf{p}, \mathbf{x}) T_k(\mathbf{q}, \mathbf{x})$$

が示されれば良いから、補題15と同様に

$$T_l(\mathbf{p}, \mathbf{x}) T_k(\mathbf{q}, \mathbf{x}) - T_k(\mathbf{p}, \mathbf{x}) T_l(\mathbf{q}, \mathbf{x})$$

を考える。この分母を払えば

$$\begin{aligned} & p_l f_l(\mathbf{x}) q_k f_k(\mathbf{x}) - p_k f_k(\mathbf{x}) q_l f_l(\mathbf{x}) \\ &= f_l(\mathbf{x}) f_k(\mathbf{x}) (p_l q_k - p_k q_l) \geq 0 \end{aligned}$$

となる。ここで、最後の不等式は S での順序の定義 (定義6) より導かれる。□

定理9 任意の $\mathbf{p}, \mathbf{q}$ に対し、 $\mathbf{p} \geq_l \mathbf{q}$ なら $\overline{T(\mathbf{p}, \mathbf{x})} \geq_l \overline{T(\mathbf{q}, \mathbf{x})}$ である。 ($\mathbf{x} \in R^n$)

証明 S での順序の定義から、任意の $l < k$ ($k, l = 1, 2, \dots$) に対し

$$\overline{T_l(\mathbf{p}, \mathbf{x}) T_k(\mathbf{q}, \mathbf{x})} \leq \overline{T_k(\mathbf{p}, \mathbf{x}) T_l(\mathbf{q}, \mathbf{x})}$$

が示されれば良い。定理8と同じように

$$\overline{T_l(\mathbf{p}, \mathbf{x}) T_k(\mathbf{q}, \mathbf{x})} - \overline{T_k(\mathbf{p}, \mathbf{x}) T_l(\mathbf{q}, \mathbf{x})}$$

を考える。補題16より定理8と同様に求められる。□

参 照 文 献

- [1] K. J. Aström, Optimal Control of Markov Processes with Incomplete State Information, II: The Convexity of the Lossfunction, *Journal of Mathematical Analysis and Applications*, vol. 26, 403-406, 1969.
- [2] M. Brown and H. Solomon, Optimal Issuing Policies under Stochastic Field Lives, *Journal of Applied Probability*, vol. 10, 761-768, 1973.
- [3] S. Karlin, *Total Positivity*, Stanford University Press, Stanford, California, 1968.
- [4] S. Karlin and H. M. Taylor, *A Second Course in Stochastic Processes*, Academic Press, New York, New York, 1981.
- [5] M. Kijima and M. Ohnishi, Portfolio Selection Problems via the Bivariate Characterization of Stochastic Dominance Relations, submitted.
- [6] G. Monahan, Optimal Stopping in a Partially Observable Markov Processes with Costly Information,

- Operations Research*, vol. 28, 1319-1334, 1980.
- [7] T. Nakai, Optimal Stopping Problem in a Finite State Partially Observable Markov Chain, *Journal of Information & Optimization Sciences*, vol. 4, 159-176, 1983.
 - [8] T. Nakai, The Problem of Optimal Stopping in a Partially Observable Markov Chain, *Journal of Optimization Theory and Applications*, vol. 45, 425-442, 1985.
 - [9] T. Nakai, A Sequential Stochastic Assignment Problem in a Partially Observable Markov Chain, *Mathematics of Operations Research*, vol. 11, 230-240, 1986.
 - [10] T. Nakai, A Stochastic Ordering and Related Sequential Decision Problems, *Journal of Information & Optimization Sciences*, vol. 11, 49-65, 1990.
 - [11] 中井 達, 尤度比を用いた二つの順序と不完備情報の多段決定問題の性質について、経済学研究(九州大学経済学会), vol. 57, 251-280, 1991.
 - [12] 中井 達, 尤度比を用いた順序と多段決定問題について、京都大学数理解析研究所講究録「数理計画モデルにおける最適化理論」, vol. 798, 191-201, 1992.
 - [13] T. Nakai, A Partially Observable Decision Problem under a Shifted Likelihood Ratio Ordering, *Proceedings of the Australia-Japan Workshop on Stochastic Models in Engineering, Technology and Management* (Eds. S. Osaki and D. N. Pra Murthy), World Scientific Publishing, 413-422, 1993.
 - [14] 中井 達, 部分観測可能なマルコフ連鎖の上での最適選択問題について、経済学研究(九州大学経済学会), vol. 59, 392-424, 1993.
 - [15] M. L. Pinedo and S. M. Ross, Scheduling Jobs Subject to Nonhomogeneous Poisson Shocks, *Management Science*, vol. 26, 1250-1257, 1980.
 - [16] S. M. Ross, Quality Control under Markovian Deterioration, *Management Science*, vol. 17, 587-596, 1971.
 - [17] S. M. Ross, *Stochastic Processes*, John-Wiley & Sons, New York, New York, 1983.
 - [18] M. Shaked and J. G. Shanthikumar, *Stochastic Orders and their Applications*, Academic Press, San Diego, California, 1994.
 - [19] J. G. Shanthikumar and D. D. Yao, The Preservation of Likelihood Ratio Ordering under Convolution, *Stochastic Processes and their Applications*, vol. 23, 259-267, 1986.
 - [20] D. Stoyan, *Comparison Methods for Queues and Other Stochastic Models*, John Wiley & Sons, New York, New York, 1983.