不完備情報マルコフ過程におけるベイズ学習について

中井, 達
九州大学経済学部 : 教授

https://doi.org/10.15017/4363005

出版情報：經濟學研究. 66 (1), pp.67-85, 1999-06-30. 九州大学経済学会

不完備情報マルコフ過程におけるベイズ学習について

中　井　　　達

1．序

部分観測可能なマルコフ過程における決定モデルで，状態に関する情報が多変量確率変数によって得られる場合のベイズ学習を考える。とくに状態についての情報が一般の確率分布で与えられる場合を中心に考察する。そのために，未知の状態に関する情報として状態空間上の確率測度を考える。これらの状態は，全順序が定義されている完備な可分距離空間に含まれ，情報全体に順序を定義する。そのために，MTP_2性（multivariate totally positivity of order two）または FKG 不等式として知られる尤度比を用いた性質を拡張し，一般化した FKG 不等式を定義する。さらに，この一般化した FKG 不等式の性質を求め，それらの結果を部分観測可能なマルコフ過程に応用する。また，このモデルではベイズ理論を用いた学習プロセスを仮定し，とくに事前情報のもとでの期待値が順序を保存する性質を求める。確率変数の間に，このような尤度比をもとにした順序を導入する場合についての研究は多くあるが，ほとんどは情報などの絶対連続性を仮定している。それに対して，ここでは状態についての情報を，一般的に確率測度とすることが異なっている。したがって，一般化された FKG 不等式を用いた仮定の下で，事前情報と事後情報の関係についても議論する。

はじめに第２章では，部分観測可能なマルコフ過程を定義し，未知の状態に関する情報が絶対連続な確率変数から得られる場合について，準備のために基本的な結果をまとめる。まず，尤度比を用いて情報の間にTP_2性を持つ順序を導入する。Brown and Solomon [1]，Nakai [5, 7, 8]， Ohnishi, Kawai and Mine [9]，Stoyan [12]などでもこの順序関係を用いて議論している。ここでは，この情報は多変量確率変数から得られるとし，MTP_2 性を用いた仮定の下で議論を進める。さらに，事前情報と事後情報の関係についても考える。

第３章では，部分観測可能なマルコフ過程における学習で，状態についての情報が一般の確率測度で与えられる場合を考えるために，完備な可分距離空間上の確率測度の集合に順序を定義する。この順序は，FKG 不等式で知られる不等式の一般化であり，一般化した FKG 不等式と呼ぶ。この場合に，部分観測可能なマルコフ過程における学習プロセスを考える上で必要な性質を求める。第４章では，状態に関する情報が，状態空間上の確率測度で表される場合を考え，一般的に絶対連続であることを仮定しない。そのために，確率測度の間に第３章で考えた順序関係を導入し，事前情報のもとでの非減少関数の期待値に関する性質を考える。また，第２章で考えたような部分観測可能なマルコフ過程

の学習プロセスにおいて，事前情報と事後情報の関係についてもあわせて考える。最後に例として，逐次決定モデルのなかで，簡単な最適停止モデルに応用する。

2．部分観測可能なマルコフ過程 ── 情報が絶対連続の場合 ──

状態を直接観測できない部分観測可能なマルコフ過程を考える。いま，$\mathfrak{S}$ を全順序が定義されている完備な可分距離空間とし，この確率過程の状態を表す。この空間で定義されている順序を $\leq$ で表す。状態空間 $\mathfrak{S}$ に対して，$\mathfrak{B}$ を $\mathfrak{S}$ の Borel 集合族とする。さらに，$P(\cdot|s)$ $(s\in\mathfrak{S})$ を推移規則とし，$(\mathfrak{S},\mathfrak{B})$ での stochastic kernel とする。すなわち，それぞれの $s\in\mathfrak{S}$ に対して，$P(\cdot|s)$ は $\mathfrak{S}$ での確率測度であり，それぞれの $B\in\mathfrak{B}$ に対して，$P(B|\cdot)$ は $\mathfrak{S}$ 上の絶対連続な Borel 可測な関数であり，密度関数 $p(\cdot|s)$ を持つ。この関数は，$\mathfrak{S}$ から $\mathfrak{S}$ への推移を表す。ここで，任意の可測集合 $B\in\mathfrak{B}$ に対して，

$$P(B|s)=\int_B p(t|s)\,dt$$

であり，$p:\mathfrak{S}\times\mathfrak{S}\to \boldsymbol{R}_+$ は P の可測な密度である。

また，この確率過程の任意の状態 s に対して，期待値が有限な非負の k 次多変量確率変数 $\boldsymbol{X}_s$ を考え，観測過程とする。任意の状態 $s\in\mathfrak{S}$ と可測集合 $C\subset\boldsymbol{R}^k$ に対して，この絶対連続な確率変数の確率密度関数 $f(\boldsymbol{x}|s)$ が存在して，

$$\Pr(\boldsymbol{X}_s\in C)=\int_C f(\boldsymbol{x}|s)d\boldsymbol{x}\quad(\boldsymbol{x}\in\boldsymbol{R}^k,\ s\in S) \tag{1}$$

とする。この多変量確率変数から得られる標本を用いて，この確率過程の状態についての情報を得る。ここでは，学習プロセスとしてベイズの定理を用いる。

次に，この確率過程の状態に関する情報は，状態空間 $\mathfrak{S}$ 上の絶対連続な確率測度 μ で表される。このとき，$\boldsymbol{P}(\mathfrak{S})$ を状態空間 $\mathfrak{S}$ 上の確率測度の集合とすれば，μ の密度関数を $\mu(s)$ とするとき

$$\boldsymbol{P}(\mathfrak{S})=\left\{\mu \mid \mu(s)\geq 0,\ \int_{\mathfrak{S}}\mu(s)ds=1\right\}$$

となる。これらの確率測度の間に，尤度比を用いて順序を次のように定義する。

定義 1　μ と ν を $\mathfrak{S}$ 上の確率分布とする。もし，$s<t, s, t\in\mathfrak{S}$ のとき

$$\mu(t)\nu(s)\geq\mu(s)\nu(t) \tag{2}$$

であり，$\mu(t)\nu(s)>\mu(s)\nu(t)$ が少なくとも 1 つの s と t の組み合わせに対して成り立つならば，$\mu>_l\nu$ とする。もし，任意の $s\in\mathfrak{S}$ に対して，$\mu(s)=\nu(s)$ のとき，$\mu=_l\nu$ とする。$\mu=_l\nu$ かつ $\mu>_l\nu$ ならば，$\mu\geq_l\nu$ である。

定義 1 で定義した順序が半順序になることは簡単にわかる。また，$\mathfrak{S}$ 上で定義された関数 $u(\mu)$ に対して，$u(\mu)\geq u(\nu)$ が任意の μ と ν（$\mu\geq_l\nu$）に対して成り立つとき，この関数を μ に関する非減少関数と呼ぶことにする。

任意の標本 $\boldsymbol{x}=(x_1,\cdots,x_k)$ $(\in \boldsymbol{R}^k_+=(0,\infty)^k)$ と事前情報 μ に対して，事後情報は存在しベイズの定理に従って学習を行い，次のようなプロセスで学習を行う。事前情報が μ のとき，まず始めにこの確率過程は推移法則 $P(\cdot|\cdot)$ に従って状態が推移する。その後での状態に関する情報 $\bar{\mu}$ は，任意の可測集合 $B\in\mathfrak{B}$ に対して

$$\bar{\mu}(B)=\int_B \bar{\mu}(t)dt, \tag{3}$$

となり，その確率密度関数は

$$\bar{\mu}(t)=\int_{\mathfrak{S}} p(t|s)\mu(s)ds \tag{4}$$

である。つぎに，標本 $\boldsymbol{x}$ を観測した後での情報はベイズの定理によって，$\overline{\mu(\boldsymbol{x})}$ と学習される。すなわち，

$$\overline{\mu(\boldsymbol{x})}(B)=\frac{\int_B f(\boldsymbol{x}|t)\bar{\mu}(t)dt}{\int_S f(\boldsymbol{x}|t)\bar{\mu}(t)dt} \tag{5}$$

がすべての可測集合 $B\in\mathfrak{B}$ について成り立つ。ただし，$\boldsymbol{x}\in\boldsymbol{R}^k$ および $\mu\in\boldsymbol{P}(S)$ である。また，絶対連続な確率測度 $\overline{\mu(\boldsymbol{x})}$ の密度関数は

$$\overline{\mu(\boldsymbol{x})}(t)=\frac{f(\boldsymbol{x}|t)\bar{\mu}(t)}{\int_S f(\boldsymbol{x}|t)\bar{\mu}(t)dt} \tag{6}$$

となる。次の２つの仮定の下で，これらの事前情報と事後情報に関する性質を考える。

仮定１　任意の状態 $s, t\in\mathfrak{S}$ に対して，$\boldsymbol{x}, \boldsymbol{y}\in\boldsymbol{R}^k_+$ のとき

$$f(\boldsymbol{x}\wedge\boldsymbol{y}|s\wedge t)f(\boldsymbol{x}\vee\boldsymbol{y}|s\vee t)\geq f(\boldsymbol{x}|s)f(\boldsymbol{y}|t) \tag{7}$$

である。ここで，$\boldsymbol{x}\wedge\boldsymbol{y}=(\min(x_1,y_1),\cdots,\min(x_k,y_k))$ および $\boldsymbol{x}\vee\boldsymbol{y}=(\max(x_1,y_1),\cdots,\max(x_k,y_k))$ とする。

仮定２　推移密度関数 $p(u|s)$ は TP_2 性を持つ $(s,u\in\mathfrak{S})$。すなわち，任意の s と t に対して $(s<t)$

$$p(u|s)p(v|t)\geq p(v|s)p(u|t)$$

がすべての $u<v$ について成立する $(u,v\in\mathfrak{S})$。

注１　状態空間 $\mathfrak{S}$ が一般的な半順序が定義されている空間であれば，仮定２は，

$$p(u\wedge v|s\wedge t)p(u\vee v|s\vee t)\geq p(v|s)p(u|t) \tag{8}$$

と表すことが出来る。ここで考えているように，$\mathfrak{S}$ が全順序集合のときは(8)式は，$s<t$, $u<v$ のとき $p(u|s)p(v|t)\geq p(v|s)p(u|t)$ と等しいことは簡単な計算からわかる。

仮定１は，MTP_2 性（multivariate totally positivity of order two）と呼ばれ，TP_2 性の一般化である。また，この確率過程の状態が経済の状態を表すとすれば，この状態は価格に反映することは明らかである。すなわち，この状態を直接知ることは出来ないのが，状態についての何らかの情報を持っていることは一般的である。仮定１から，s の値が大きくなれば，この経済の状態はよくなることになる。また，仮定２から s の値が大きくなればなるほど，大きな値の状態へ移りやすくなり，s の値が小さ

くなればなるほど，小さい値の状態へ移りやすくなる。

まず始めに，(3)式と(5)式で定義された事後情報に関する性質を見る。その前に，k 次元の標本空間に簡単な順序を定義する。情報が多変量確率変数から得られ得ることから，この順序は Nakai[8] で用いられた順序とは異なる。

定義２　$\boldsymbol{x}=(x_1,\cdots,x_k)$ と $\boldsymbol{y}=(y_1,\cdots,y_k)$ を k 次元確率変数 $\boldsymbol{X}$ から得られた標本とする。もし，$x_i \leq y_i$ $(i=1,2,\cdots,k)$ のとき，$\boldsymbol{x}$ は $\boldsymbol{y}$ より小さいといい，$\boldsymbol{x} \prec \boldsymbol{y}$ と表す。

１次元の確率変数から情報が得られるとき，Nakai[5, 7]でいろいろな性質が求められている。また，Nakai[8]においても，複数の独立な確率変数から情報が得られる場合を考察している。事後情報は次のような性質（定理１および２）を持ち，Nakai[5, 7, 8]で用いられた方法と同じようにして求められる。

補題１　任意の $\mu,\nu \in P(\mathfrak{S})$ に対して，$\mu \geq_l \nu$ ならば $\bar{\mu} \geq_l \bar{\nu}$ である。

証明. 任意の $s,t \in \mathfrak{S}$ に対して，$s>t$ のとき

$$\bar{\mu}(s)\bar{\nu}(t) \geq \bar{\mu}(t)\bar{\nu}(s)$$

すなわち

$$\bar{\mu}(s)\bar{\nu}(t) - \bar{\mu}(t)\bar{\nu}(s) \geq 0$$

を示す。

$$\bar{\mu}(s)=\int_{\mathfrak{S}} p(s|u)\mu(u)du,\ \ \bar{\nu}(t)=\int_{\mathfrak{S}} p(t|u)\nu(u)du$$

だから，

$$\begin{aligned}
&\bar{\mu}(s)\bar{\nu}(t) - \bar{\mu}(t)\bar{\nu}(s)\\
&\quad=\int_{\mathfrak{S}}\int_{\mathfrak{S}} p(s|u)\mu(u)p(t|v)\nu(v)dudv-\int_{\mathfrak{S}}\int_{\mathfrak{S}} p(t|u)\mu(u)p(s|v)\nu(v)dudv\\
&\quad=\int\int_{\{(u,v)|u\geq v\}}(p(s|u)p(t|v)-p(t|u)p(s|v))\mu(u)\nu(v)dudv\\
&\qquad+\int\int_{\{(u,v)|u<v\}}(p(s|u)p(t|v)-p(t|u)p(s|v))\mu(u)\nu(v)dudv\\
&\quad=\int\int_{\{(u,v)|u>v\}}(p(s|u)p(t|v)-p(t|u)p(s|v))(\mu(u)\nu(v)-\mu(v)\nu(u))dudv\\
&\quad\geq 0
\end{aligned}$$

となる。順序の定義（定義１）と仮定２より最後の不等式が導かれ，この性質が求められる。□

定理１　任意の $\mu \in \boldsymbol{P}(\mathfrak{S})$ に対して，$\boldsymbol{x} \prec \boldsymbol{y}$ ならば，$\overline{\mu(\boldsymbol{x})} \leq_l \overline{\mu(\boldsymbol{y})}$ である。

証明. $\mathfrak{S}$ における順序の定義より，任意の $s,t \in \mathfrak{S}$ に対して $s<t$ のとき不等式

$$\overline{\mu(\boldsymbol{y})}(t)\overline{\nu(\boldsymbol{x})}(s) \geq \overline{\mu(\boldsymbol{y})}(s)\overline{\nu(\boldsymbol{x})}(t)$$

が成り立つことを示せばよい。始めに

$$\overline{\mu(\boldsymbol{y})}(t)\overline{\nu(\boldsymbol{x})}(s) - \overline{\mu(\boldsymbol{y})}(s)\overline{\nu(\boldsymbol{x})}(t) \geq 0$$

を示す。分母を払って各項の被積分関数を整理すれば，

$$\bar{\mu}(t)f(\boldsymbol{y}|t)\bar{\mu}(s)f(\boldsymbol{x}|s)-\bar{\mu}(s)f(\boldsymbol{y}|s)\bar{\mu}(t)f(\boldsymbol{x}|t)$$
$$=\bar{\mu}(t)\bar{\mu}(s)(f(\boldsymbol{y}|t)f(\boldsymbol{x}|s)-f(\boldsymbol{y}|s)f(\boldsymbol{x}|t))\geq 0$$

となる。最後の不等式は，仮定1から導かれる。□

定理2　任意の $\boldsymbol{x}\in\boldsymbol{R}^k$ に対して，もし $\boldsymbol{\mu}\geq_l\boldsymbol{\nu}$ ならば $\bar{\boldsymbol{\mu}}(\boldsymbol{x})\geq_l\bar{\boldsymbol{\nu}}(\boldsymbol{x})$ である。

証明. 任意の $s,t\in\mathfrak{S}$ に対して，$s<t$ のとき不等式

$$\overline{\mu(\boldsymbol{x})}(t)\overline{\nu(\boldsymbol{x})}(s)\geq\overline{\mu(\boldsymbol{x})}(s)\overline{\nu(\boldsymbol{x})}(t)$$

を示す。

$$\overline{\mu(\boldsymbol{x})}(t)\overline{\nu(\boldsymbol{x})}(s)-\overline{\mu(\boldsymbol{x})}(s)\overline{\nu(\boldsymbol{x})}(t)$$

において，各項の被積分関数を整理すれば

$$\bar{\mu}(t)f(\boldsymbol{x}|t)\bar{\nu}(s)f(\boldsymbol{x}|s)-\bar{\mu}(s)f(\boldsymbol{x}|s)\bar{\nu}(t)f(\boldsymbol{x}|t)$$
$$=f(\boldsymbol{x}|s)f(\boldsymbol{x}|t)\left(\bar{\mu}(t)\bar{\nu}(s)-\bar{\mu}(s)\bar{\nu}(t)\right)\geq 0$$

となる。補題1と仮定1から定理1と同じように，この不等式が求まる。□

任意の標本 $\boldsymbol{x}$ に対して，簡単な計算から事後情報に関する次の性質が得られる。

定理3　$\boldsymbol{\mu}\in\boldsymbol{P}(\mathfrak{S})$ のとき，$\overline{\boldsymbol{\mu}(\boldsymbol{x})}$ はすべての状態 s と t に対して $s<t$ のとき

$$\overline{\mu(\boldsymbol{x}\vee\boldsymbol{y})}(s\vee t)\overline{\mu(\boldsymbol{x}\wedge\boldsymbol{y})}(s\wedge t)\geq\overline{\mu(\boldsymbol{y})}(s)\overline{\mu(\boldsymbol{x})}(t)$$

となる。

証明. 任意の $s,t\in S$ に対して，$s<t$ のとき

$$\overline{\mu(\boldsymbol{x}\vee\boldsymbol{y})}(s\vee t)\overline{\mu(\boldsymbol{x}\wedge\boldsymbol{y})}(s\wedge t)-\overline{\mu(\boldsymbol{y})}(s)\overline{\mu(\boldsymbol{x})}(t)\geq 0$$

が示されればよい。この不等式の分母を払って被積分関数を整理すると，

$$\begin{aligned}
&\bar{\mu}(s\vee t)f(\boldsymbol{x}\vee\boldsymbol{y}|s\vee t)\bar{\mu}(s\wedge t)f(\boldsymbol{x}\wedge\boldsymbol{y}|s\wedge t)-\bar{\mu}(s)f(\boldsymbol{y}|s)\bar{\mu}(t)f(\boldsymbol{x}|t)\\
&=\bar{\mu}(s\vee t)f(\boldsymbol{x}\vee\boldsymbol{y}|s\vee t)\bar{\mu}(s\wedge t)f(\boldsymbol{x}\wedge\boldsymbol{y}|s\wedge t)-\bar{\mu}(s)f(\boldsymbol{x}\vee\boldsymbol{y}|s\vee t)\bar{\mu}(t)f(\boldsymbol{x}\wedge\boldsymbol{y}|s\wedge t)\\
&\quad+\bar{\mu}(s)f(\boldsymbol{x}\vee\boldsymbol{y}|s\vee t)\bar{\mu}(t)f(\boldsymbol{x}\wedge\boldsymbol{y}|s\wedge t)-\bar{\mu}(s)f(\boldsymbol{y}|s)\bar{\mu}(t)f(\boldsymbol{x}|t)\\
&=\left(\bar{\mu}(s\vee t)\bar{\mu}(s\wedge t)-\bar{\mu}(s)\bar{\mu}(t)\right)f(\boldsymbol{x}\vee\boldsymbol{y}|s\vee t)f(\boldsymbol{x}\wedge\boldsymbol{y}|s\wedge t)\\
&\quad+\bar{\mu}(s)\bar{\mu}(t)(f(\boldsymbol{x}\vee\boldsymbol{y}|s\vee t)f(\boldsymbol{x}\wedge\boldsymbol{y}|s\wedge t)-f(\boldsymbol{y}|s)f(\boldsymbol{x}|t))\\
&\geq 0
\end{aligned}$$

となる。最後の不等式は仮定1と補題1より導かれる。□

これらの性質は事前情報と事後情報の関係を示すものであり，部分観測可能なマルコフ過程における決定問題を考える上で基本的なものである。

つぎに，$\boldsymbol{R}^k$ 上の関数を考える。

定義3　k 変数関数 $\varphi:\boldsymbol{R}^k\to\boldsymbol{R}$ が，$\boldsymbol{x}\prec\boldsymbol{y}$ となる $\boldsymbol{x}$，$\boldsymbol{y}$ に対して，$\varphi(\boldsymbol{x})\leq\varphi(\boldsymbol{y})$ $(\varphi(\boldsymbol{x})\geq\varphi(\boldsymbol{y}))$ となるとき，この関数は $\boldsymbol{x}$ に関して非減少関数（非増加関数）という。

事後情報のもとでの条件付き期待値に関して，Preston[10], Karlin and Rinott[3]等により，次の性質を持つ。

命題1　(*Preston 1974*) μ と ν を $\boldsymbol{R}^k$ 上の絶対連続な確率測度で，それらの確率密度 $\mu(\boldsymbol{x})$ および $\nu(\boldsymbol{y})$ が任意の $\boldsymbol{x}$, $\boldsymbol{y} \in \boldsymbol{R}^k$ に対して

$$\mu(\boldsymbol{x} \vee \boldsymbol{y})\nu(\boldsymbol{x} \wedge \boldsymbol{y}) \geq \mu(\boldsymbol{x})\nu(\boldsymbol{y})$$

とする。このとき，もし $h: \boldsymbol{R}^k \to \boldsymbol{R}_+$ が有界で可測な非減少関数のとき

$$\int_{R^k} h(x)d\mu(x) \geq \int_{R^k} h(x)d\nu(x)$$

である。

系1　μ と ν を $\mathfrak{S}$ 上の絶対連続な確率測度で，それらの確率密度 $\mu(s)$ および $\nu(s)$ が任意の s, $t \in \mathfrak{S}$, $s \leq t$ に対して

$$\mu(s)\nu(t) \geq \mu(t)\nu(s)$$

とする。このとき，もし $\Phi: \mathfrak{S} \to \boldsymbol{R}_+$ が有界で可測な非減少関数のとき

$$\int_{\mathfrak{S}} \Phi(s)d\mu(s) \geq \int_{\mathfrak{S}} \Phi(s)d\nu(s)$$

である。

まずはじめに，関数 $\Phi: \mathfrak{S} \to \boldsymbol{R}$ を，

$$\Phi(s) = \mathrm{E}[\varphi(X_s)]$$

とおく。もし，$\varphi(\boldsymbol{x})$ が $\boldsymbol{x}$ の非減少関数であれば，仮定1と命題1より，$\Phi(s)$ は s に関して非減少関数である。このことと，系1より次の性質が導かれる。

補題2　$\mu \geq_l \nu$ (μ, $\nu \in \boldsymbol{P}(\mathfrak{S})$) ならば

$$\begin{aligned}\mathrm{E}_{\mu}[\varphi(\boldsymbol{X})] &= \int_{\mathfrak{S}}\left\{\int_{R^k}\varphi(\boldsymbol{x})f(\boldsymbol{x}|s)\right\}d\boldsymbol{x}\mu(s)ds = \int_{\mathfrak{S}}\Phi(s)\mu(s)ds \\ &\geq \int_{\mathfrak{S}}\Phi(s)\nu(s)ds = \int_{\mathfrak{S}}\left\{\int_{R^k}\varphi(\boldsymbol{x})f(\boldsymbol{x}|s)d\boldsymbol{x}\right\}\nu(s)ds = \mathrm{E}_{\nu}[\varphi(\boldsymbol{X})]\end{aligned}$$

が $\boldsymbol{x}$ に関する非滅文少関数 $\varphi(\cdot)$ に対して成り立つ。

3．一般化された FKG 不等式とその性質

全順序が定義された完備で可分な距離空間を $\mathfrak{S}$ とし，$\mathfrak{S}$ 上の確率測度を考える。いま，$\mathfrak{S}$ 上の確率測度を μ で表し，$\boldsymbol{P}(\mathfrak{S})$ を距離空間 $\mathfrak{S}$ 上の確率測度の集合とすれば

$$\boldsymbol{P}(\mathfrak{S}) = \left\{\mu \,\middle|\, \int_{\mathfrak{S}}\mu(ds) = 1,\ \mu(B) \geq 0 (B \in \mathfrak{B})\right\}$$

となる。ここでは，これらの確率測度 $\mu(s)$ は必ずしも絶対連続である必要はない。これらの確率測度の間に，前節と同様の順序を定義する。

定義4　$\mathfrak{S}$に含まれる集合AとBに対し，$a \leq b$がすべての$a \in A$と$b \in B$について成り立つとき，そのときに限り$A \prec B$とする。

定義5　μとνを$\mathfrak{S}$上の確率測度とする。互いに背反な*Borel*集合AとB（$A, B \in \mathfrak{B}$）で$A \prec B$のとき

$$\mu(A)\nu(B) \leq \mu(B)\nu(A),\ i.e., \begin{vmatrix} \mu(B) & \nu(A) \\ \mu(B) & \nu(A) \end{vmatrix} \geq 0, \tag{9}$$

であり，少なくとも1つの組 A と B に対して$\mu(B)\,\nu(A) > \mu(A)\,\nu(B)$となるとき，$\mu >_l \nu$とする。もし，任意の$A \in \mathfrak{B}$に対して確率1で$\mu(A) = \nu(A)$ならば$\mu =_l \nu$とする。$\mu =_l \nu$で$\mu >_l \nu$のとき$\mu \geq_l \nu$とする。

まず，状態空間$\mathfrak{S}$で定義されている距離をdとする。このとき，$\mathfrak{S}$の任意の部分集合Aに対して，$d(A)$を

$$d(A) = \sup_{a_1,\ a_2 \in A} d(a_1,\ a_2)$$

と定義する。この値は集合Aの大きさを測る尺度である。

つぎに，μとνを全順序が定義された距離空間$\mathfrak{S}$上の確率測度で，互いに背反なBorel集合AとB（$A, B \in \mathfrak{B}$）で$A \prec B$のとき

$$\mu(A)\nu(B) - \mu(B)\nu(A) \leq 0 \tag{10}$$

とする。これら2つの確率測度μとνに対し，任意の$\varepsilon > 0$について，$\mathfrak{S}$の部分集合S_εを

$$S_\varepsilon = \{A \in \mathfrak{B} \mid \mu(A) \geq \nu(A), d(A) < \varepsilon,\ A\text{は連結である}\}$$

とする。この集合はμとνによって決まるものであるが，以下の議論では簡単のためにμとνを省くことにする。

補題3　$A, B \in S_\varepsilon$のとき$A \cap B \in S_\varepsilon$である。また，$\mu(A \cup B) \geq \nu(A \cup B)$である。

証明． $A, B \in S_\varepsilon$とすると$A \cap B \in S_\varepsilon$であることを示す。簡単のために，$U = A \cap B^c$, $V = A \cap B$, $W = A^c \cap B$とおく（ここで，A^cは集合Aの補集合を表す）。このとき，$A, B \in S_\varepsilon$だから，集合A, Bは連結である。$V \neq \emptyset$であれば一般性を失うことなく$U \prec V \prec W$と仮定してもよい。

このとき仮定より，

$$\mu(A) = \mu(U) + \mu(V) \geq \nu(U) + \nu(V) = \nu(A)$$
$$\mu(B) = \mu(V) + \mu(W) \geq \nu(V) + \nu(W) = \nu(B)$$

であり，$U \prec V \prec W$より

$$\mu(U)\,\nu(V) - \mu(V)\,\nu(U) \leq 0$$
$$\mu(V)\,\nu(W) - \mu(W)\,\nu(V) \leq 0$$

となる。従って，簡単な計算によって$\mu(V) = \mu(A \cap B) \geq \nu(V) = \nu(A \cap B)$を示すことが出来る。

最後に，$\mu(A \cup B) \geq \nu(A \cup B)$であることも，同じように，$U \prec V \prec W$としたとき$\mu(W) \geq \nu(W)$となることに注意すれば明らかである。□

つぎに，集合S_εに対して$u_\varepsilon = \inf_{s \in S_\varepsilon} s$とおく。この$u_\varepsilon$に対し集合族$E_\varepsilon$を

$$E_\varepsilon = \{A \mid \{u_\varepsilon\} \prec A,\ \mu(A) = \nu(A) = 0, A \in \mathfrak{B}\}$$

とし，$\widetilde{E}_\varepsilon$を$\widetilde{E}_\varepsilon=\cup_{A\in E_\varepsilon}A$とする。このとき集合$G_\varepsilon$を

$$G_\varepsilon=(\cup_{A\in S_\varepsilon}A)\cup\widetilde{E}_\varepsilon$$

とおく。

補題4　集合$G_\varepsilon=(\cup_{A\in S_\varepsilon}A)\cup\widetilde{E}_\varepsilon$は連結である。

証明． この集合G_εが連結でないとする。すなわち，1点集合$\{u_\varepsilon\}$に対して，$\{u_\varepsilon\}\prec B$であって$B\cap G_\varepsilon=\emptyset$であり，さらに$A\prec B$となる$A\in S_\varepsilon$が存在すると仮定する。また，この$B$は$d(B)<\varepsilon$となるように取ることが出来る。集合$G_\varepsilon$の定義より，$\{u_\varepsilon\}\prec B$であり，$B\cap\widetilde{E}_\varepsilon=\emptyset$だから$\mu(B)=\nu(B)=0$とはならない。

まず，$A\prec B$だからμとνが(10)式を満たすことから，

$$\mu(A)\nu(B)-\mu(B)\nu(A)\leq 0$$

である。一方，$A\in S_\varepsilon$より$\mu(A)\geq\nu(A)$である。したがって，$\mu(B)=\nu(B)=0$ではないから$\mu(B)\geq\nu(B)$となる。このことから，$B\in S_\varepsilon$となり矛盾となる。従ってこの性質が成り立つ。□

つぎに，集合Gを

$$G=\cap_{\varepsilon>0}G_\varepsilon \tag{11}$$

とすると，Gは集合族$\{G_\varepsilon\mid\varepsilon>0\}$の射影的極限となる。

補題5　$A\subset G$となる$A\in\mathfrak{B}$に対し，$\mu(A)\geq\nu(A)$である。

証明． $A\subset G$となる$A\in\mathfrak{B}$で$\mu(A)\leq\nu(A)$であるようなAが存在したとする。Aに含まれる点aのε近傍$U_\varepsilon(a)$を考え，集合族$\{U_\varepsilon(a)\cap A\mid a\in A\}$を考える。もし，すべての$U_\varepsilon(a)\cap A$に対して，$\mu(U_\varepsilon(a)\cap A)\geq\nu(U_\varepsilon(a)\cap A)$であれば，簡単な計算から$\mu(A)\geq\nu(A)$となり仮定に反する。したがって，$\mu(U_\varepsilon(a)\cap A)<\nu(U_\varepsilon(a)\cap A)$となるような点$a$とその近傍$U_\varepsilon(a)$が存在する。ところが，$d(U_\varepsilon(a)\cap A)<\varepsilon$だから$U_\varepsilon(a)\cap A\notin S_\varepsilon$となる。したがって，$U_\varepsilon(a)\cap G_\varepsilon=\emptyset$となり，集合$G$の定義から$A\subset G$であることに矛盾する。従って，この性質が成り立つ。□

さらに，集合Hを

$$H=\mathfrak{S}\cap G^c$$

すなわち$G\cup H=\mathfrak{S}$とおく。また，集合Gの定義より$0<\mu(G)$, $\nu(G)<1$となることは明らかである。

補題6　μとνを全順序が定義された距離空間$\mathfrak{S}$上の確率測度で，互いに背反な*Borel*集合AとB $(A, B\in\mathfrak{B})$で$A\prec B$のとき

$$\mu(A)\nu(B)-\mu(B)\nu(A)\leq 0 \tag{12}$$

とする。このとき，$(\mathfrak{S}\times\mathfrak{S},\ \mathfrak{B}\times\mathfrak{B})$上の確率測度$\delta$で

$$\delta(A\times\mathfrak{S})=\mu(A)\qquad A\in\mathfrak{B} \tag{13}$$

$$\delta(\mathfrak{S}\times B)=\nu(B)\qquad B\in\mathfrak{B}, \tag{14}$$

および

$$\delta(\{(s,t)\mid(s,t)\in\mathfrak{S}\times\mathfrak{S},\ s\geq t\})=1 \tag{15}$$

となるものが存在する。

証明. (12)式を満たす2つの確率測度μとνに対し，上で考えた集合GとHを考える。直積空間$\mathfrak{S}\times\mathfrak{S}$での直積測度を考えるので，確率測度$\delta$を矩形集合$A\times B\in\mathfrak{B}\times\mathfrak{B}$で考えればよい。いま，$\mathfrak{S}\times\mathfrak{S}$上の確率測度$\delta$を

$$\delta(A\times B)=\nu(G\cap(A\cap B))+\mu(L\cap(A\cap B))+\frac{(\mu(G\cap A)-\nu(G\cap A))(\nu(L\cap B)-\mu(L\cap B))}{\mu(L)-\nu(L)}$$

と定義する。このとき，(13)式，(14)式と(15)式が成り立つことは簡単な計算からわかる。たとえば，$\mu(G)-\nu(G)=\mu(L)-\nu(L)$だから

$$\begin{aligned}\delta(\mathfrak{S}\times B)&=\nu(G\cap B)+\mu(L\cap B)+\frac{(\mu(G)-\nu(G))(\nu(L\cap B)-\mu(L\cap B))}{\mu(L)-\nu(L)}\\&=\nu(G\cap B)+\nu(L\cap B)=\nu(B)\end{aligned}$$

より(14)式が導かれる。また，$A\times B\subset\{(s,t)\mid s<t\}$の場合に$\delta(A\times B)=0$となることは，$\delta$の定義より明らかである。□

注2　*Preston*[10]では，確率測度μとνが絶対連続で密度関数$\mu(s)$と$\nu(s)$を持つ場合を議論している。補題6で定義した確率測度δが*Preston*[10]で用いられた手法を，確率測度μとνが絶対連続でない場合に拡張したものであることは容易に確認できる。

命題2　μとνを全順序が定義された可測距離空間$\mathfrak{S}$上の確率測度とし，互いに背反な*Borel*集合Aと$B(A,\ B\in\mathfrak{B})$で$A\prec B$のとき

$$\mu(A)\nu(B)-\mu(B)\nu(A)\leq 0 \tag{16}$$

とする。いま，$h:S\to\boldsymbol{R}_+$が有界で可測な非減少関数とするとき，

$$\int_{\mathfrak{S}}h(s)d\mu(s)\geq\int_{\mathfrak{S}}h(s)d\nu(s)$$

である。

証明. いま$h:S\to\boldsymbol{R}_+$を有界で可測な非減少関数とするとき$E=\{(s,\ t)\mid(s,\ t)\in\mathfrak{S}\times\mathfrak{S},\ s\geq t\}$とおけば$(s,\ t)\in E$のとき$h(s)-h(t)\geq 0$だから

$$\begin{aligned}\int_{\mathfrak{S}}h(s)d\mu(s)-\int_{\mathfrak{S}}h(t)d\nu(t)&=\iint_{\mathfrak{S}\times\mathfrak{S}}(h(s)-h(t))d\delta(s,t)\\&=\iint_{E}(h(s)-h(t))d\delta(s,t)\geq 0\end{aligned}$$

となる。□

つぎに，補題6を状態空間が全順序空間の直積空間で，自然な半順序が定義されている場合に拡張する。いま，

$$\mathfrak{S}^n=\prod_{i=1}^{n}\mathfrak{S},\ \mathfrak{B}^n=\prod_{i=1}^{n}\mathfrak{B}$$

とし，$\mathfrak{S}^1=\mathfrak{S}$, $\mathfrak{B}^1=\mathfrak{B}$とおく。また，$\mu_n$と$\nu_n$を$(\mathfrak{S}^n,\ \mathfrak{B}^n)$上の2つの確率測度とする。このとき補題6と同様の性質を示す。まず$\boldsymbol{P}(\mathfrak{S}^n)$を$(\mathfrak{S}^n,\ \mathfrak{B}^n)$での確率測度の集合とすれば，

$$P(\mathfrak{S}^n)=\left\{\mu_n \middle| \int_{\mathfrak{S}^n}\mu\,(ds^n)=1,\ \mu^n(B^n)\geq 0(B^n\in\mathfrak{B}^n)\right\}. \tag{17}$$

となる。ここでは，直積空間$\mathfrak{S}^n$での直積測度を考え，$\mathfrak{B}^n$に含まれる矩形集合について議論を進める。したがって，以下で得られる性質は一般の Borel 集合Aに対しても成り立つ。

まず始めに，Borel 集合$\mathfrak{B}^n$に含まれる矩形集合の間に半順序を定義し，さらに直積空間$\mathfrak{S}^n$上の測度の間に定義5を一般化した順序を定義する。

定義6　$\mathfrak{S}^n$に含まれる集合$\boldsymbol{A}^n=\Pi_{i=1}^n A_i$と$\boldsymbol{B}^n=\Pi_{i=1}^n B_i$に対し$(A_i, B_i\subset\mathfrak{S}, i=1,\cdots,n)$，$A_i\cap B_i=\emptyset$で$A_i<B_i$がすべての$i=1,\cdots,n$について成り立つとき，そのときに限り$\boldsymbol{A}^n\prec\boldsymbol{B}^n$とする。

定義7　$\mathfrak{S}^n$に含まれる集合$\boldsymbol{A}^n=\Pi_{i=1}^n A_i$と$\boldsymbol{B}^n=\Pi_{i=1}^n B_i$を，$A_i, B_i\subset\mathfrak{S}$で$A_i\cap B_i=\emptyset$とする$(i=1,\ \cdots,n)$。もし，$A_i\prec B_i$ならば$A_i\vee B_i=B_i$および$A_i\wedge B_i=A_i$とする。このとき，$\boldsymbol{A}^n\vee\boldsymbol{B}^n=\Pi_{i=1}^n A_i\vee B_i$, $\boldsymbol{A}^n\wedge\boldsymbol{B}^n=\Pi_{i=1}^n A_i\wedge B_i$とする。

定義8　μ_nとν_nを$(\mathfrak{S}^n,\mathfrak{B}^n)$上の確率測度で，互いに背反な*Borel* 集合 $\boldsymbol{A}^n=\Pi_{i=1}^n A_i$と$\boldsymbol{B}^n=\Pi_{i=1}^n B_i$ $(\boldsymbol{A}^n, \boldsymbol{B}^n\in\mathfrak{B}^n, A_i, B_i\subset\mathfrak{S}, A_i\cap B_i=\emptyset, i=1,\cdots,n)$のとき，

$$\mu_n(\boldsymbol{A}^n\vee\boldsymbol{B}^n)\nu_n(\boldsymbol{A}^n\wedge\boldsymbol{B}^n)-\mu_n(\boldsymbol{A}^n)\nu_n(\boldsymbol{B}^n)\geq 0 \tag{18}$$

とする。このとき，μ_nはν_nより大きいといい，$\mu_n>_l\nu_n$とあらわす。

注3　μとνが全順序が定義された距離空間$\mathfrak{S}$上の確率測度で，互いに背反な*Borel*集合Aと$B(A, B\in\mathfrak{B})$で$A\prec B$に対して(10)式を仮定する。もし，μとνが絶対連続で密度関数$\mu(s)$，$\nu(s)$を持つとすれば，(10)式は，任意の$s, t\in\mathfrak{S}$に対して

$$\mu(s\vee t)\nu(s\wedge t)\geq\mu(s)\nu(t)$$

となり，TP_2性と呼ばれ，(10)式は一般化されたTP_2性と言える。

また，(18)式によって$(\mathfrak{S}^n,\mathfrak{B}^n)$上の確率測度$\mu_n$と$\nu_n$の間に順序を定義した。ここにおいても同じように，絶対連続性を仮定すれば，FKG 不等式（[2], [4], [10]など）または，MTP_2性と呼ばれるものである。従って，(18)式は FKG 不等式を拡張したものであり，一般化された FKG 不等式(generalized FKG inequality)と呼ぶ。

この節以降の議論では，$\mathfrak{S}$上の関数の2つの確率分布μとνに関する積分の値を比較する。すなわち，$f(s)\geq 0$を$\mathfrak{S}$の非負関数で，μとνを$\mathfrak{S}$上の確率測度とするとき，$\int_A f(s)\,\mu(ds)$と$\int_A f(s)\,\nu(ds)$の比較である$(A\in\mathfrak{B})$。この場合，ルベーグ積分の定義から，これらの積分は単関数

$$\sum_{j=0}^n\alpha_j\chi_{A_j}(s)$$

の極限として定義されている（ただし，Aは有限個の集合の直和で$\chi_{A_j}(s)$は集合A_jの定義関数とし，$A=A_0+A_1+\cdots+A_n$, $\alpha_0=0<\alpha_j$, $j=1,2,\cdots,n$とする)。いま，単関数$f_n(s)\geq 0$の単調増加列で$f(s)$に近づく関数列$\{f(s)\}$が存在して，

$$\int_A f(s)\mu(ds)=\lim_{n\to\infty}\int_A f_n(s)\mu(ds)$$

である。したがって，これらの単関数$f_n(s)$に対して$\int_A f_n(s)\mu(ds)$と$\int_A f_n(s)\nu(ds)$を比較すればよい。

このとき単関数の定義より，$\{A_j\}$それぞれの定義関数ごとに比較することになる。しかし，ここでは議論を見通しよくするため，積分 $\int_A f(s)\,\mu(ds)$ において記号である ds を便宜上，定義関数の領域 A_j に含まれる部分集合を表す変数のように扱う。

補題 7　μ_n と ν_n を $(\mathfrak{S}^n, \mathfrak{B}^n)$ 上の確率測度で，互いに背反な *Borel* 集合 $A^n = \Pi_{i=1}^n A_i$ と $B^n = \Pi_{i=1}^n B_i$ $(A^n, B^n \in \mathfrak{B}^n, A_i, B_i \subset \mathfrak{S}, A_i \cap B_i = \emptyset, i=1, \cdots, n)$ のとき，

$$\mu_n(A^n \vee B^n)\,\nu_n(A^n \wedge B^n) - \mu_n(A^n)\,\nu_n(B^n) \geq 0 \tag{19}$$

とする。いま，確率測度 μ_{n-1}, ν_{n-1} を，μ_n と ν_n の周辺確率測度とし，$\mu_{n-1}(A^{n-1}) = \mu_n(A^{n-1} \times \mathfrak{S})$ および $\nu_{n-1}(B^{n-1}) = \nu_n(A^{n-1} \times \mathfrak{S})$ とする $(A^{n-1} \times \mathfrak{S}, B^{n-1} \times \mathfrak{S} \in \mathfrak{S}^{n-1} \times \mathfrak{S} = \mathfrak{S}^n)$。このとき，

$$\mu_{n-1}(A^{n-1} \vee B^{n-1})\,\nu_{n-1}(A^{n-1} \wedge B^{n-1}) - \mu_{n-1}(A^{n-1})\,\nu_{n-1}(B^{n-1}) \geq 0$$

である。

証明．まず，全順序集合 $\mathfrak{S}$ の直積集合 $\mathfrak{S} \times \mathfrak{S}$ を 3 つの領域 $D = \{(s,t) \mid s < t, s, t \in \mathfrak{S}\}$，$E = \{(s,t) \mid s = t, s, t \in \mathfrak{S}\}$，$F = \{(s,t) \mid s > t, s, t \in \mathfrak{S}\}$ にわける。このとき，

$$\mu_{n-1}(A^{n-1})\,\nu_{n-1}(B^{n-1}) = \int_{\mathfrak{S} \times \mathfrak{S}} \mu_n(A^{n-1}, ds)\,\nu_n(B^{n-1}, dt)$$

$$= \int_D \mu_n(A^{n-1}, ds)\,\nu_n(B^{n-1}, dt) + \int_E \mu_n(A^{n-1}, ds)\,\nu_n(B^{n-1}, dt) + \int_F \mu_n(A^{n-1}, ds)\,\nu_n(B^{n-1}, dt)$$

$$= \int_D \mu_n(A^{n-1}, ds)\,\nu_n(B^{n-1}, dt) + \int_E \mu_n(A^{n-1}, ds)\,\nu_n(B^{n-1}, dt) + \int_D \mu_n(A^{n-1}, dt)\,\nu_n(B^{n-1}, ds)$$

$$= \int_D \{\mu_n(A^{n-1}, ds)\,\nu_n(B^{n-1}, dt) + \mu_n(A^{n-1}, dt)\,\nu_n(B^{n-1}, ds)\} + \int_E \mu_n(A^{n-1}, ds)\,\nu_n(B^{n-1}, dt)$$

となる。同じように，

$$\mu_{n-1}(A^{n-1} \vee B^{n-1})\,\nu_{n-1}(A^{n-1} \wedge B^{n-1})$$

$$= \int_D \{\mu_n(A^{n-1} \vee B^{n-1}, ds)\,\nu_n(A^{n-1} \wedge B^{n-1}, dt) + \mu_n(A^{n-1} \vee B^{n-1}, dt)\,\nu_n(A^{n-1} \wedge B^{n-1}, ds)\}$$

$$+ \int_E \mu_n(A^{n-1} \vee B^{n-1}, ds)\,\nu_n(A^{n-1} \wedge B^{n-1}, dt)$$

である。

まず，$(ds, dt) \in E$ のとき $(A^{n-1} \vee B^{n-1}, ds)$，$(A^{n-1} \wedge B^{n-1}, dt) \in \mathfrak{S}^n$ だから

$$\mu_n(A^{n-1} \vee B^{n-1}, ds)\,\nu_n(A^{n-1} \wedge B^{n-1}, dt) \geq \mu_n(A^{n-1}, ds)\,\nu_n(B^{n-1}, dt)$$

となることは，領域 E の定義と (19) 式より明らかである。

つぎに，$(ds, dt) \in D = \{(s, t) \mid s < t, s, t \in \mathfrak{S}\}$ の場合を考える。このとき

$$\begin{aligned} &\mu_n(A^{n-1} \vee B^{n-1}, ds)\,\nu_n(A^{n-1} \wedge B^{n-1}, dt) + \mu_n(A^{n-1} \vee B^{n-1}, dt)\,\nu_n(A^{n-1} \wedge B^{n-1}, ds) \\ &\geq \mu_n(A^{n-1}, ds)\,\nu_n(B^{n-1}, dt) + \mu_n(A^{n-1}, dt)\,\nu_n(B^{n-1}, ds) \end{aligned} \tag{20}$$

を示せばよい。一方，$(ds, dt) \in D$ だから仮定より

$$\mu_n(A^{n-1}\vee B^{n-1}, dt)\,\nu_n(A^{n-1}\wedge B^{n-1}, ds)\geq \mu_n(A^{n-1}, ds)\,\nu_n(B^{n-1}, dt)$$

であり，

$$\mu_n(A^{n-1}\vee B^{n-1}, dt)\,\nu_n(A^{n-1}\wedge B^{n-1}, dt)\geq \mu_n(A^{n-1}, dt)\,\nu_n(B^{n-1}, dt)$$

である。また

$$\mu_n(A^{n-1}\vee B^{n-1}, ds)\,\nu_n(A^{n-1}\wedge B^{n-1}, ds)\geq \mu_n(A^{n-1}, ds)\,\nu_n(B^{n-1}, ds)$$

および，

$$\mu_n(A^{n-1}\vee B^{n-1}, dt)\,\nu_n(A^{n-1}\wedge B^{n-1}, ds)\geq \mu_n(A^{n-1}, dt)\,\nu_n(B^{n-1}, ds)$$

から

$$\mu_n(A^{n-1}\vee B^{n-1}, ds)\,\nu_n(A^{n-1}\wedge B^{n-1}, dt)\,\mu_n(A^{n-1}\vee B^{n-1}, dt)\,\nu_n(A^{n-1}\wedge B^{n-1}, ds)$$
$$\geq \mu_n(A^{n-1}, ds)\,\nu_n(B^{n-1}, dt)\,\mu_n(A^{n-1}, dt)\,\nu_n(B^{n-1}, ds)$$

となる。ここで，

$$a:=\mu_n(A^{n-1}\vee B^{n-1}, ds)\,\nu_n(A^{n-1}\wedge B^{n-1}, dt)$$
$$b:=\mu_n(A^{n-1}\vee B^{n-1}, dt)\,\nu_n(A^{n-1}\wedge B^{n-1}, ds)$$
$$c:=\mu_n(A^{n-1}, ds)\,\nu_n(B^{n-1}, dt)$$
$$d:=\mu_n(A^{n-1}, dt)\,\nu_n(B^{n-1}, dt)$$

とおけば，$(a-c)(a-d)\geq 0$と $ab\geq cd$ を組み合わせれば簡単に $a+b\geq c+d$ すなわち(20)式が導かれる。この性質は Preston[10]の Lemma 2 で知られている性質である。□

注 4　状態空間 $\mathfrak{S}^n$ に含まれる２つの集合 A^n と B^n に対し，ここで考えた順序ではなく次のような順序を考えることができる。すなわち，$\mathfrak{S}^n$ に含まれる集合 A^n と B^n に対し，$a^n\leq b^n$ がすべての $a^n\in A^n$ と $b^n\in B^n$ について成り立つとき，そのときに限り $A^n\prec B^n$ とする。ただし，$a^n\leq b^n$ であるとは，任意の$1\leq i\leq n$ に対して $a_i^n\leq b_i^n$となることである（ただし，$a^n=(a_1^n, \cdots a_n^n)$, $b^n=(b_1^n, \cdots, b_n^n)$ とする)。

しかし，状態空間 $\mathfrak{S}^n$ に含まれる２つの集合 A^n と B^n に対し，このような順序を定義した場合に，状態空間 $\mathfrak{S}^n$ 上の確率測度の間にはここで考えた一般化された FKG 不等式とは異なる順序を考える必要があり，この場合には補題７で得られた性質を求めることができない。

補題 8　μ_n と ν_n を$(\mathfrak{S}^n, \mathfrak{B}^n)$上の確率測度で，互いに背反な *Borel* 集合 $A^n=\Pi_{i=1}^n A_i$ と $B^n=\Pi_{i=1}^n B_i$ $(A^n, B^n\in\mathfrak{B}^n, A_i, B_i\subset\mathfrak{S}, A_i\cap B_i=\emptyset, i=1, \cdots, n)$のとき，

$$\mu_n(A^n\vee B^n)\,\nu_n(A^n\wedge B^n)-\mu_n(A^n)\,\nu_n(B^n)\geq 0 \tag{21}$$

とする。このとき，$(\mathfrak{S}^n\times\mathfrak{S}^n, \mathfrak{B}^n\times\mathfrak{B}^n)$上の確率測度 δ で

$$\delta(A^n\times\mathfrak{S}^n)=\mu_n(A^n)\qquad A^n\in\mathfrak{B}^n \tag{22}$$

$$\delta(\mathfrak{S}^n\times B^n)=\nu_n(B^n)\qquad B^n\in\mathfrak{B}^n, \tag{23}$$

および

$$\delta(\{(s^n, t^n)\mid(s^n, t^n)\in\mathfrak{S}^n\times\mathfrak{S}^n, s^n\succ t^n\})=1 \tag{24}$$

となるものが存在する。

証明.　μ_n と ν_n の周辺確率測度 μ_{n-1}, ν_{n-1} を $\mu_{n-1}(A^{n-1})=\mu_n(A^{n-1}\times\mathfrak{S})$, $\nu_{n-1}(B^{n-1})=\nu_n(A^{n-1}\times\mathfrak{S})$と

する$(\boldsymbol{A}^{n-1}\times\mathfrak{S}, \boldsymbol{B}^{n-1}\times\mathfrak{S}\in\mathfrak{S}^{n-1}\times\mathfrak{S}=\mathfrak{S}^n)$。このとき補題7より，帰納法の仮定を用いて

$$\mathfrak{S}^{n-1}=\prod_{i=1}^{n-1}\mathfrak{S},\quad \mathfrak{B}^{n-1}=\prod_{i=1}^{n-1}\mathfrak{B}$$

のとき，$(\mathfrak{S}^{n-1}, \mathfrak{B}^{n-1})$上の確率測度$\mu_{n-1}$, ν_{n-1}に対して$(\mathfrak{S}^{n-1}\times\mathfrak{S}^{n-1}, \mathfrak{B}^{n-1}\times\mathfrak{B}^{n-1})$上の確率測度$\tilde{\delta}$で

$$\tilde{\delta}(\boldsymbol{A}^{n-1}\times\mathfrak{S}^{n-1})=\mu_{n-1}(\boldsymbol{A}^{n-1})\quad \boldsymbol{A}^{n-1}\in\mathfrak{B}^{n-1} \tag{25}$$

$$\tilde{\delta}(\mathfrak{S}^{n-1}\times\boldsymbol{B}^{n-1})=\nu_{n-1}(\boldsymbol{A}^{n-1})\quad \boldsymbol{B}^{n-1}\in\mathfrak{B}^{n-1} \tag{26}$$

および

$$\tilde{\delta}(\{(\boldsymbol{s}^{n-1}, \boldsymbol{t}^{n-1})\mid(\boldsymbol{s}^{n-1}, \boldsymbol{t}^{n-1})\in\mathfrak{S}^{n-1}\times\mathfrak{S}^{n-1}, \boldsymbol{s}^{n-1}\succ\boldsymbol{t}^{n-1}\})=1 \tag{27}$$

となるものが存在する。

つぎに，$\boldsymbol{A}^{n-1}\prec\boldsymbol{B}^{n-1}$ となる $\boldsymbol{A}^{n-1}, \boldsymbol{B}^{n-1}\in\mathfrak{B}^{n-1}(\boldsymbol{A}^{n-1}=\Pi_{i=1}^{n-1}A_i, \boldsymbol{B}^{n-1}=\Pi_{i=1}^{n-1}B_i)$の組を１つ定め，任意の$A, B\in\mathfrak{B}$対し，$\mu_{n-1}(\boldsymbol{A}^{n-1})\neq0$, $\mu_{n-1}(\boldsymbol{B}^{n-1})\neq0$のとき

$$\hat{\mu}_{\boldsymbol{A}^{n-1}}=\frac{\mu_n(\boldsymbol{A}^{n-1}\times A)}{\mu_{n-1}(\boldsymbol{A}^{n-1})}, \tag{28}$$

$$\hat{\nu}_{\boldsymbol{B}^{n-1}}=\frac{\nu_n(\boldsymbol{B}^{n-1}\times B)}{\nu_{n-1}(\boldsymbol{B}^{n-1})} \tag{29}$$

とおけば

$$\mu_n(\boldsymbol{A}^{n-1}\times A)=\hat{\mu}_{\boldsymbol{A}^{n-1}}(A)\mu_{n-1}(\boldsymbol{A}^{n-1}), \tag{30}$$

$$\nu_n(\boldsymbol{B}^{n-1}\times B)=\hat{\nu}_{\boldsymbol{B}^{n-1}}(B)\nu_{n-1}(\boldsymbol{B}^{n-1}) \tag{31}$$

であり，さらに$\hat{\mu}_{\boldsymbol{A}^{n-1}}$, $\hat{\nu}_{\boldsymbol{B}^{n-1}}$は$(\mathfrak{S}^1, \mathfrak{B}^1)=(\mathfrak{S}, \mathfrak{B})$上の確率測度である。

この２つの確率測度が，(16)式を満たすことを示す。すなわち，$A\prec B$ となる$A, B\in\mathfrak{B}$に対して

$$\hat{\mu}_{\boldsymbol{A}^{n-1}}(A)\hat{\nu}_{\boldsymbol{B}^{n-1}}(B)-\hat{\mu}_{\boldsymbol{A}^{n-1}}(B)\hat{\nu}_{\boldsymbol{B}^{n-1}}(A)\leq0 \tag{32}$$

を示す。(32)式に(28)式と(29)式を代入し分母を払って整理すれば

$$\mu_n(\boldsymbol{A}^{n-1}\times A)\nu_n(\boldsymbol{B}^{n-1}\times B)-\mu_n(\boldsymbol{A}^{n-1}\times B)\nu_n(\boldsymbol{B}^{n-1}\times A)\leqq0$$

であればよい。ところで，$A\prec B$ だから$(\boldsymbol{A}^{n-1}\times B)\vee(\boldsymbol{B}^{n-1}\times A)=\boldsymbol{A}^{n-1}\times A$ および$(\boldsymbol{A}^{n-1}\times B)\wedge(\boldsymbol{B}^{n-1}\times A)$ $=\boldsymbol{B}^{n-1}\times B$ となり，(21)式より

$$\mu_n(\boldsymbol{A}^{n-1}\times A)\nu_n(\boldsymbol{B}^{n-1}\times B)-\mu_n(\boldsymbol{A}^{n-1}\times B)\nu_n(\boldsymbol{B}^{n-1}\times A)\leq0$$

となる。

このことが示されれば補題6より，$(\mathfrak{S}\times\mathfrak{S}, \mathfrak{B}\times\mathfrak{B})$上の確率測度$\hat{\delta}_{\boldsymbol{A}^{n-1}, \boldsymbol{B}^{n-1}}$で，

$$\hat{\delta}_{\boldsymbol{A}^{n-1}, \boldsymbol{B}^{n-1}}(A\times\mathfrak{S})=\hat{\mu}_{\boldsymbol{A}^{n-1}}(A)\quad A\in\mathfrak{B} \tag{33}$$

$$\hat{\delta}_{\boldsymbol{A}^{n-1}, \boldsymbol{B}^{n-1}}(\mathfrak{S}\times B)=\hat{\nu}_{\boldsymbol{B}^{n-1}}(B)\quad B\in\mathfrak{B}, \tag{34}$$

であり

$$\hat{\delta}_{\boldsymbol{A}^{n-1}, \boldsymbol{B}^{n-1}}(\{(s, t)\mid(s, t)\in\mathfrak{S}\times\mathfrak{S},\ s\geq t\})=1$$

となるものが存在する。

このとき，$(\mathfrak{S}^n\times\mathfrak{S}^n,\ \mathfrak{B}^n\times\mathfrak{B}^n)$上の確率測度$\delta$を

$$\delta((\boldsymbol{A}^{n-1}, A)\times(\boldsymbol{B}^{n-1}, B))=\tilde{\delta}(\boldsymbol{A}^{n-1}\times\boldsymbol{B}^{n-1})\hat{\delta}_{\boldsymbol{A}^{n-1}, \boldsymbol{B}^{n-1}}(A\times B)$$

によって定義すれば，この δ が条件を満たすことがわかる。すなわち，(25)式より，

$$\tilde{\delta}(A^{n-1}\times\mathfrak{S}^{n-1})=\mu_{n-1}(A^{n-1})$$

であり，(33)式より

$$\tilde{\delta}_{A^{n-1},B^{n-1}}(A\times\mathfrak{S})=\mu_{A^{n-1}}(A)$$

だから

$$\begin{aligned}\delta(A^n\times\mathfrak{S}^n)&=\delta((A^{n-1},A)\times(S^{n-1},\mathfrak{S}))\\&=\tilde{\delta}(A^{n-1}\times\mathfrak{S}^{n-1})\,\hat{\delta}_{A^{n-1},\mathfrak{S}^{n-1}}(A\times\mathfrak{S})\\&=\mu_{n-1}(A^{n-1})\,\hat{\mu}_{A^{n-1}}(A)\\&=\mu_n(A^{n-1}\times A)\\&=\mu_n(A^n)\end{aligned}$$

となり，(22)式が示された。□

4．部分観測可能なマルコフ過程 ── 情報が確率測度で与えられた場合 ──

第2節と同様な部分観測可能なマルコフ過程を考える。$\mathfrak{S}$ を状態空間とし，$\mathfrak{S}$ の Borel 集合族 $\mathfrak{B}$ に対し，$P(\cdot|s)\ (s\in\mathfrak{S})$ を推移規則とする。このとき，任意の可測集合 $B\in\mathfrak{B}$ に対して

$$P(B|s)=\int_B p(dt|s)$$

とする。この確率過程の状態 s に対して，期待値が有限の非負の k 次元多変量確率変数 $\boldsymbol{X}_s$ の存在を仮定し，観測過程を表す。この確率変数の確率分布を，任意の状態 $s\in\mathfrak{S}$ と可測集合 $C\subset\boldsymbol{R}^k$ に対して

$$\Pr(X_s\in C)=\int_C f(d\boldsymbol{x}|s)\quad(\boldsymbol{x}\in\boldsymbol{R}^k,\ s\in\mathfrak{S}),\tag{35}$$

とし，これらの確率変数が絶対連続のとき，その密度関数を $f(\boldsymbol{x}|s)$ とすれば

$$\Pr(X_s\in C)=\int_C f(\boldsymbol{x}|s)d\boldsymbol{x}$$

となる。これらの確率変数から得られる標本を使って未知の状態に関する情報を得る。第2節と同様に，学習プロセスとしてベイズの定理を用いる。

この確率過程の状態に関する情報は，状態空間 $\mathfrak{S}$ 上の確率測度 μ で表し，第3節で考えた $\boldsymbol{P}(S)$ が情報の集合とする。ここでは第2節とは異なり，これらの確率測度 μ は必ずしも絶対連続である必要はない。これらの確率測度の間に，第3節で考えた順序を定義する（定義5）。以下の議論では，推移法則 $P(\cdot|s)\ (s\in\mathfrak{S})$ と確率変数 $\boldsymbol{X}_s$ は便宜上，絶対連続である場合を扱うが，一般の場合にも同様に示すことができる。

任意の標本 $\boldsymbol{x}=(x_1,\cdots,x_k)\ (\in\boldsymbol{R}_+^k=(0,\infty)^k)$ と事前情報 μ に対して，事後情報は存在し，ベイズの定理を用いて学習し，そのプロセスは第2節で考えたと同様に次のように進む。事前情報を μ とするとき，まず始めに推移法則 P に従って推移する。推移した時点での状態空間上の確率分布 $\bar{\mu}$ は，任意

の可測集合 $B\in\mathfrak{B}$ に対して

$$\bar{\mu}(B)=\int_{\mathfrak{S}}p(B|s)\mu(ds) \tag{36}$$

となる。

標本 $\boldsymbol{x}$ を観測した後で，ベイズの定理に従って，情報を $\overline{\mu(\boldsymbol{x})}$ のように改良する。すなわち，$\boldsymbol{x}\in\boldsymbol{R}_+^k$ および $\mu\in\boldsymbol{P}(\mathfrak{S})$ のとき，任意の可測集合 $B\in\mathfrak{B}$ に対して

$$\overline{\mu(\boldsymbol{x})}(B)=\frac{\int_B f(\boldsymbol{x}|t)\bar{\mu}(dt)}{\int_{\mathfrak{S}} f(\boldsymbol{x}|t)\bar{\mu}(dt)} \tag{37}$$

となる。

ここで，第２節と同様の仮定（仮定１と２）の下で，(36)式と(37)式で与えられた事後情報の性質を考える。

補題９　任意の μ，$\nu\in\boldsymbol{P}(\mathfrak{S})$ に対して，$\mu\geq_l\nu$ ならば $\bar{\mu}\geq_l\bar{\nu}$ である。

証明．任意の部分集合 A と $B(A, B\in\mathfrak{B})$ で，$A\prec B$ のとき

$$\bar{\mu}(A)\bar{\nu}(B)\leq\bar{\mu}(B)\bar{\nu}(A)$$

すなわち

$$\bar{\mu}(A)\bar{\nu}(B)-\bar{\mu}(B)\bar{\nu}(A)\leq 0$$

を示す。

$$\bar{\mu}(A)=\int_{\mathfrak{S}}p(A|s)\mu(ds) \qquad \bar{\nu}(B)=\int_{\mathfrak{S}}p(B|s)\nu(ds)$$

だから，

$$\begin{aligned}
&\bar{\mu}(A)\bar{\nu}(B)-\bar{\mu}(B)\bar{\nu}(A)\\
&\quad=\int_{\mathfrak{S}}p(A|s)\mu(ds)\int_{\mathfrak{S}}p(B|t)\nu(dt)-\int_{\mathfrak{S}}p(B|s)\mu(ds)\int_{\mathfrak{S}}p(A|t)\nu(dt)\\
&\quad=\int_{\mathfrak{S}\times\mathfrak{S}}p(A|s)p(B|t)\mu(ds)\nu(dt)-\int_{\mathfrak{S}\times\mathfrak{S}}p(B|s)p(A|t)\mu(ds)\nu(dt)\\
&\quad=\int_{\mathfrak{S}\times\mathfrak{S}}\{p(A|s)p(B|t)-p(B|s)p(A|t)\}\mu(ds)\nu(dt)
\end{aligned}$$

となる。一方，$A\prec B$ だから，順序の定義（定義５）と仮定２より

$$\begin{aligned}
&p(A|s)p(B|t)-p(B|s)p(A|t)\\
&\quad=\int_A p(u|s)du\int_B p(v|t)dv-\int_B p(v|s)dv\int_A p(u|t)du\\
&\quad=\int_{A\times B}(p(u|s)p(v|t)-p(v|s)p(u|t))dudv\\
&\quad\leq 0
\end{aligned}$$

である。したがって

$$\bar{\mu}(A)\bar{\nu}(B)-\bar{\mu}(B)\bar{\nu}(A)\leq 0$$

が導かれる。□

定理 4　任意の $\mu\in\boldsymbol{P}(\mathfrak{S})$ に対して，$\boldsymbol{x}\prec\boldsymbol{y}$ ならば $\overline{\mu(\boldsymbol{x})}\leq_l\overline{\mu(\boldsymbol{y})}$ である。

証明. $\mathfrak{S}$ における順序の定義より，任意の A と B $(A,B\in\mathfrak{B},A\prec B)$ に対して次の不等式が成り立てばよい。

$$\overline{\mu(\boldsymbol{y})}(A)\overline{\mu(\boldsymbol{x})}(B)\leq\overline{\mu(\boldsymbol{x})}(B)\overline{\mu(\boldsymbol{y})}(A)$$

始めに

$$\overline{\mu(\boldsymbol{y})}(A)\overline{\mu(\boldsymbol{x})}(B)-\overline{\mu(\boldsymbol{y})}(B)\overline{\mu(\boldsymbol{x})}(A)$$

を考え，分母を払って整理すれば

$$\int_A f(\boldsymbol{y}|s)\bar{\mu}(ds)\int_B f(\boldsymbol{x}|t)\bar{\mu}(dt)-\int_B f(\boldsymbol{y}|t)\bar{\mu}(dt)\int_A f(\boldsymbol{x}|s)\bar{\mu}(ds)$$

$$=\int_A f(\boldsymbol{y}|s)\bar{\mu}(ds)\int_B f(\boldsymbol{x}|t)\bar{\mu}(dt)-\int_A f(\boldsymbol{x}|s)\bar{\mu}(ds)\int_B f(\boldsymbol{y}|t)\bar{\mu}(dt)$$

$$=\iint_{A\times B} f(\boldsymbol{y}|s)f(\boldsymbol{x}|t)\bar{\mu}(ds)\bar{\mu}(dt)-\iint_{A\times B} f(\boldsymbol{x}|s)f(\boldsymbol{y}|t)\bar{\mu}(ds)\bar{\mu}(dt)$$

$$=\iint_{A\times B}\{f(\boldsymbol{y}|s)f(\boldsymbol{x}|t)-f(\boldsymbol{x}|s)f(\boldsymbol{y}|t)\}\bar{\mu}(ds)\bar{\mu}(dt)$$

となる。最後の不等式は仮定 1 より導かれる。□

定理 5　任意の $\boldsymbol{x}\in\boldsymbol{R}^k$ に対して，$\mu\geq_l\nu$ ならば $\overline{\mu(\boldsymbol{x})}\geq_l\overline{\nu(\boldsymbol{x})}$ である。

証明. 任意の部分集合 A と B $(A,B\in\mathfrak{B},A\prec B)$ に対して，不等式

$$\overline{\mu(\boldsymbol{x})}(A)\overline{\nu(\boldsymbol{x})}(B)\leq\overline{\mu(\boldsymbol{x})}(B)\overline{\nu(\boldsymbol{x})}(A)$$

を示せばよい。

$$\overline{\mu(\boldsymbol{x})}(A)\overline{\nu(\boldsymbol{x})}(B)-\overline{\mu(\boldsymbol{x})}(B)\overline{\nu(\boldsymbol{x})}(A),$$

において，分母を払って分子を整理すれば

$$\int_A f(\boldsymbol{x}|s)\bar{\mu}(ds)\int_B f(\boldsymbol{x}|t)\bar{\nu}(dt)-\int_B f(\boldsymbol{x}|t)\bar{\mu}(dt)\int_A f(\boldsymbol{x}|s)\bar{\nu}(ds)$$

$$=\iint_{A\times B} f(\boldsymbol{x}|s)f(\boldsymbol{x}|t)\bar{\mu}(ds)\bar{\nu}(dt)-\iint_{A\times B} f(\boldsymbol{x}|s)f(\boldsymbol{x}|t)\bar{\nu}(ds)\bar{\mu}(dt)$$

$$=\iint_{A\times B} f(\boldsymbol{x}|s)f(\boldsymbol{x}|t)\{\bar{\mu}(ds)\bar{\nu}(dt)-\bar{\nu}(ds)\bar{\mu}(dt)\}$$

$$\leq 0$$

となる。補題 9 と仮定 1 より，定理 4 と同じように，この不等式が示される。□

これらの性質は，事前情報と事後情報の関係についてのもので，部分観測可能なマルコフ過程での決定問題を考える上で，基本的なものである。最後に，事前情報のもとでの期待値に関する性質（命題2）を示す必要がある。これは，情報が絶対連続な確率測度である場合の命題1の一般化となっている。したがって，$\Phi(s)=\int_{R^k}\varphi(x)f(x|s)dx$とおけば，事後情報のもとでの条件付き期待値に関して，命題2から次の性質が得られる。

補題10　任意の x の非減少関数 $\varphi(\cdot)$に対して，$\mu \geq_l \nu$ $(\mu,\nu\in \boldsymbol{P}(\mathfrak{S}))$ならば

$$\mathrm{E}_{\mu}[\varphi(X)]=\int_{\mathfrak{S}}\left\{\int_{R^k}\varphi(x)f(x|s)dx\right\}\mu(ds)=\int_{\mathfrak{S}}\Phi(s)\mu(ds)$$

$$\geq\int_{\mathfrak{S}}\Phi(s)\nu(ds)=\int_{\mathfrak{S}}\left\{\int_{R^k}\varphi(x)f(x|s)dx\right\}\nu(ds)=\mathrm{E}_{\nu}[\varphi(X)]$$

である。

補題6では Preston[10]で用いられた手法を，確率測度μとνが絶対連続でない場合に拡張したものであった。つぎに，状態空間が全順序空間の直積空間となっていて，自然な半順序が定義されている場合にも補題8ように拡張できた。したがって，命題2と同じように事前情報のもとでの非減少関数の期待値に関して次の性質が成り立つ。証明は命題2と同様の方法で求められるのでここでは省略する。

命題3　μ_n と ν_n を$(\mathfrak{S}^n,\mathfrak{B}^n)$上の確率測度で，互いに背反な *Borel* 集合 $\boldsymbol{A}^n=\Pi_{i=1}^n A_i$ と $\boldsymbol{B}^n=\Pi_{i=1}^n B_i$ $(\boldsymbol{A}^n,\boldsymbol{B}^n\in\mathfrak{B}^n, A_i, B_i\subset\mathfrak{S}, A_i\cap B_i=\emptyset, i=1,\cdots,n)$のとき

$$\mu_n(\boldsymbol{A}^n\vee\boldsymbol{B}^n)\nu_n(\boldsymbol{A}^n\wedge\boldsymbol{B}^n)-\mu_n(\boldsymbol{A}^n)\nu_n(\boldsymbol{B}^n)\geq 0$$

とする。いま，$h:\mathfrak{S}^n\to\boldsymbol{R}_+$が有界で可測な非減少関数とするとき，

$$\int_{\mathfrak{S}^n}h(s)d\mu_n(s)\geq\int_{\mathfrak{S}^n}h(s)d\nu_n(s)$$

である。

この命題において，$\mathfrak{S}=\boldsymbol{R}$ とすれば次の性質が成り立つ。

系2　μ_k と ν_k を$(\boldsymbol{R}_+^k,\mathfrak{B}^k)$上の確率測度で，互いに背反な *Borel* 集合 $\boldsymbol{C}^k=\Pi_{i=1}^k C_i$ と $\boldsymbol{D}^k=\Pi_{i=1}^k D_i$ $(\boldsymbol{C}^k,\boldsymbol{D}^k\in\mathfrak{B}^k, C_i, D_i\subset\boldsymbol{R}, C_i\cap D_i=\emptyset,\ i=1,\cdots,k)$のとき

$$\mu_k(\boldsymbol{C}^k\vee\boldsymbol{D}^k)\nu_k(\boldsymbol{C}^k\wedge\boldsymbol{D}^k)-\mu_k(\boldsymbol{C}^k)\nu_k(\boldsymbol{D}^k)\geq 0$$

とする。いま，$h:\boldsymbol{R}_+^k\to\boldsymbol{R}_+$が有界で可測な非減少関数とするとき，

$$\int_{R_+^k}h(x)d\mu_k(x)\geq\int_{R_+^k}h(x)d\nu_k(x)$$

である。

第2節と同じように，関数$\Phi:\boldsymbol{R}\to\boldsymbol{R}$を，

$$\Phi(s)=\mathrm{E}[\varphi(X_s)]$$

とおく。仮に観測プロセスの状態 s に依存する k 次元多変量確率変数 X_s が絶対連続な確率変数でなくても，系２が成り立つ。したがって，$\varphi(\boldsymbol{x})$ が $\boldsymbol{x}$ の非減少関数であれば，仮定１と系２より，$\Phi(s)$ は s に関して非減少関数である。このことと命題２から，次の性質が得られる。

補題11　$\mu \geq_l \nu$ $(\mu, \nu \in \boldsymbol{P}(\boldsymbol{R}))$ならば

$$\mathrm{E}_{\mu}[\varphi(X)]=\int_R\left\{\int_{R^k}\varphi(\boldsymbol{x})f(d\boldsymbol{x}|s)\right\}\mu(ds)=\int_R\Phi(s)\mu(ds)$$

$$\geq\int_R\Phi(s)\nu(ds)=\int_R\left\{\int_{R^k}\varphi(\boldsymbol{x})f(d\boldsymbol{x}|s)\right\}\nu(ds)=\mathrm{E}_{\nu}[\varphi(X)]$$

が $\boldsymbol{x}$ に関する非減少関数 $\varphi(\cdot)$ 対して成り立つ。

最後に簡単な決定問題に応用する。ここで扱った部分観測可能なマルコフ過程での最適停止問題を考える。n 個の k 次の多変量確率変数が１度に１つずつ現れる。これらの確率変数から得られる標本値 $\boldsymbol{x}=(x_1, \cdots, x_k)$ を観測して，この確率過程の状態に関する情報を得る。これらの確率変数は確率過程の状態に依存し，状態についての情報は状態空間上の確率測度で表される。また，あわせて観測を続けるかどうかを決定する。観測をやめれば，標本値 $\boldsymbol{x}$ に依存した利得 $\varphi(\boldsymbol{x})$ が得られる。停止しなければ，状態について学習し，次の標本値を観測する。この場合の利得はない。利得関数 $\varphi(\boldsymbol{x})$ は，$\boldsymbol{x}$ に関して非減少とする。例えば，$\varphi(\boldsymbol{x})=\max_{1\leq i\leq k}x_i$ とおけばよい。この問題の目的は，n 期の間に総期待利得を最大にするように，１つの標本を選ぶことである。

この確率過程の状態についての情報が μ のとき，$v_n(\mu)$ を n 期間最適攻策に従ったときに得られる総期待利得とする。このとき，最適性の原理から次の再帰方程式が得られる(Ross[8])。

$$v_n(\mu)=\mathrm{E}_{\mu}\left[\max\left\{\varphi(\boldsymbol{x}), v_{n-1}\left(\overline{\mu(\boldsymbol{x})}\right)\right\}\right] \tag{38}$$

であり，便宜上

$$v_n(\mu\,|\boldsymbol{x}):=\max\left\{\varphi(\boldsymbol{x}), v_{n-1}\left(\overline{\mu(\boldsymbol{x})}\right)\right\} \tag{39}$$

とおく。これらの再帰方程式と定理４より，定理５と補題10から次の性質が得られる。

命題４　$v_n(\mu\,|\boldsymbol{x})$ は μ と $\boldsymbol{x}$ の非減少関数である。すなわち，$\mu \geq_l \nu$ および $\boldsymbol{x} \prec \boldsymbol{y}$ のとき，$v_n(\nu\,|\,\boldsymbol{x}) \leq v_n(\mu\,|\boldsymbol{x})$, $v_n(\mu\,|\boldsymbol{x}) \leq v_n(\mu\,|\boldsymbol{y})$ となる。$v_n(\mu)$ は μ 非減少関数である。

証明． n に関する帰納法を用いればよい。$n=1$ の場合は明らかである。これらの性質が $n-1$ 以下の場合に成り立つとする。定理５より，$\mu \geq_l \nu$ ならば $\overline{\mu(\boldsymbol{x})} \geq_l \overline{\nu(\boldsymbol{x})}$ である。このことから，$v_n(\mu\,|\boldsymbol{x})$ が μ に関して非減少であることが帰納法の仮定よりわかる。一方，定理４から，$\boldsymbol{x} \prec \boldsymbol{y}$ ならば $\overline{\mu(\boldsymbol{x})} \geq_l \overline{\mu(\boldsymbol{y})}$ である。したがって帰納法の仮定より，$v_n(\mu\,|\boldsymbol{x})$ は $\boldsymbol{x}$ に関して非減少である。補題10より，$v_n(\mu)$ は μ に関して非減少関数である。□

次に，$\boldsymbol{R}^k$ に含まれる領域 $S_n(\mu)$ を $S_n(\mu)=\{\boldsymbol{x}\,|\,\varphi(\boldsymbol{x}) \geq v_{n-1}(\overline{\mu(\boldsymbol{x})})\}$ と定義すれば，$S_n(\mu)$ はこの最適停止問題の停止領域を表す。したがって，この問題の最適政策を表す。

命題５　$\mu \geq_l \nu$ ならば $S_n(\mu) \subset S_n(\nu)$ である。

証明. 命題４より，$v_n(\mu\,|\,x)$はμに関する非減少関数である。さらに，$\mu \leq \nu$ならば，$\varphi(x) \geq v_{n-1}(\overline{\mu(x)}) \geq v_{n-1}(\overline{\nu(x)})$である。したがって，この性質が示される。□

References

[1] M. Brown and H. Solomon, Optimal Issuing Policies under Stochastic Field Lives, *Journal of Applied Probability,* **10**, 761-768, 1973.

[2] R. Holley, Remarks on the FKG Inequaliteis, *Communications on Mathematical Physics,* **36**, 227-231, 1974.

[3] S. Karlin and Y. Rinott, Class of Orderings of Measures and Related Correlation Inequalities I: Multivariate Totally Positive Distributions, *Journal of Multivariate Analysis,* **10**, 467-498, 1980.

[4] J. H. B. Kemperman, On the FKG-Inequality for Measures on a Partially Ordered Space, *Indagationes Mathematicae,* **39**, 313-331, 1977.

[5] T. Nakai, The Problem of Optimal Stopping in a Partially Observable Markov Chain, *Journal of Optimization Theory and Applications,* **45**, 425-442, 1985.

[6] T. Nakai, An Optimal Selection Problem with a Random Number of Applicants per Period, *Operations Research,* **34**, 478-485, 1986.

[7] T. Nakai, A Sequential Stochastic Assignment Problem in a Partially Observable Markov Chain, *Mathematics of Operations Research,* **11**, 230-240, 1986.

[8] T. Nakai, An Optimal Selection Problem on a Partially Observable Markov Chain, *Stochastic Modelling in Innovative Manufacturing* (Eds. A. H. Christer, S. Osaki and L. C. Thomas), Lecture Notes in Economics and Mathematical Systems **445**, Springer-Verlag, Berlin, 140-154, 1996.

[9] M. Ohnishi, H. Kawai and H. Mine, An Optimal Inspection and Replacement Policy under Incomplete State Information, *European Journal of Operations Research,* **27**, 117-128, 1986.

[10] C. J. Preston, A Generalization of the FKG Inequaliteis, *Communications on Mathematical Physics,* **36**, 233-241, 1974.

[11] S. M. Ross, *Applied Probability Models with Optimization Applications,* Holden-Day,San Francisco, California, 1970.

[12] D. Stoyan, *Comparison Methods for Queues and Other Stochastic Models,* John Wiley & Sons, New York, 1983.

〔九州大学経済学部教授〕