北九州市立大学法政論集第 50 巻第1・2合併号(2022 年 10 月)抜刷

研究ノート

単調尤度比性(monotone likelihood ratio property)の基礎事項

上　條　諒　貴

北九州市立大学法政論集第 50 巻第 1・2 合併号（2022 年 10 月）

研究ノート

単調尤度比性（monotone likelihood ratio property）の基礎事項

上　條　諒　貴*

1　序

本稿は、特定の確率分布が満たす性質である、いわゆる「単調尤度比性（Monotone Likelihood Ratio Property, 以下 MLRP と略す）」について、政治学の数理モデルにおける利用の例などに触れつつ、その基礎的な事項について紹介することを目的とする。

MLRP は、情報の経済学、契約理論、本人-代理人論といったミクロ経済学の諸分野において、モデルの解やその比較静学が特定の（望ましい）性質を持つための条件として、各モデルの設定に即した形で具体的に特徴づけされつつ（或いは類似の条件が）かねてから用いられていたが、Milgrom の著名な論文において、一般的な定義づけや様々な応用分野における有用性の紹介がなされたことで（Milgrom 1981）、「ノイズを含んだシグナルからの（意思決定をする際に重要になる値についての）推論」という、不確実性下の意思決定では極めて頻繁に問題になる状況をモデル化する際に用いられる標準的な仮定の一つとして確立されることとなった。その結果、現在では契約理論（ex. 伊藤 2003）や本人-代理人論（ex. Laffont and Martimort 2002）の標準的なテキストにおいて（和書 / 洋書

＊本学法学部准教授

を問わず）MLRP についての解説がなされるようになっている。

ノイズを含んだシグナルを用いた推論が必要になる状況は、特定の意思決定に固有のものというわけでは当然なく、政治的意思決定においても頻繁に現れるため、政治学の数理モデルもしばしば MLRP を用いるようになっている。4.2 で扱う選挙アカウンタビリティのモデルにおける候補者の能力の高低に関する推論や、より一般的に有権者に好ましく評価される候補者の属性（いわゆる valence）に関する推論（政治家への献金を題材にした例として Prat (2002) など）、異なる政策手段の有効性に影響する要素に関する推論（専門化への委任を題材にした例として Li and Suen (2004)、政治的分極化を題材にした例として Dixit and Weibull (2007) など）といったように、その応用範囲は広い。

しかし、そうした状況にもかかわらず、政治学の数理分析のテキスト（それ自体数が少ないという問題もあるが）においては、一部の洋書において簡単に触れられる例があるのみ（McCarty and Meirowitz 2007, p.159）で、政治学者や政治学学習者を念頭に、平易且つある程度詳細に MLRP について解説したものは管見の限り見当たらない。こうした間隙を埋めるのが本稿の目指すところである。

本稿の構成は以下のとおりである。まず 2 節において、具体的な例を用いてなぜ MLRP のような仮定が必要になるのかという点をあらためて詳細に説明したうえで、MLRP の定義を行う。その後、3 節において、MLRP の有用性について検討したのち、4 節で政治学における応用例を紹介する。

2 MLRP とは何か

2.1 モチベーション

まずは（一部の人間にとってのみ、だが）卑近な例から、特定の意思決定をモデル化する上で、何故 MLRP なるものが必要となるのかについて考えることとしよう。

大学教員の重要な仕事の一つは、学期末に学生に対する評価を与えることである。ここでいう学生に対する評価とは、より正確に言えば、学生が単位を与えるに足る十分な能力を備えているか、あるいは十分な努力をしたかということへの評価である。こうした能力や過去の努力水準といった評価の際の関心事項 / 対象を θ で表そう。加えて、以下では話を単純にするため、学生の能力、或いは過去の努力水準は高い $(H, High)$ か低い (L, Low) かであるとしよう。すなわち、$\theta \in \{\theta_L, \theta_H\}, \theta_L < \theta_H$ である。

このように話を単純化してもなお、学生に対して評価を与えることがそれほど容易ではない最大の理由は、こうした θ が直接には観察できないためである。能力なるものは目に見えないものであるし、努力しているかということ自体は原理的には観察可能であっても、教員が個々の学生の生活を常に監視できるわけではない。

しかし、θ が直接観察できないからといって、教員がサイコロの目や答案を紙飛行機にして飛ばした際の飛距離などを用いて、いわばランダムに単位付与を決定するわけでは（おそらく）ないということも明らかであろう。評価をするために多くの場合行われるのがいわゆる試験（テスト）である。以下では試験におけるスコアを x で表す。

ここまでの議論について、以下の二つの点に注意を要する。

まず、θ が直接観察できないということは、ここで問題となるような意思決定（すなわち単位付与）は、例えば「能力が高ければ単位を与える」というような θ の値に直接条件付けた行動ルールによっては表現できないということを意味している。直接観察できるのはあくまでテストスコアである x であり、ここでの意思決定の根幹は、観察可能な x から観察不可能な θ について推論する過程となる。

二点目に、どれほど周到に設計しようとも、受験者の能力や過去の努力を完全に明らかにできる、すなわち θ が観察可能であるのと等しくなるような試験を作ることはできない。現実の試験というものを考えてみれば明らかなように、能力の高い学生、或いは徹底的に対策をしてきた学生が、緊張や体調不良によって当日に力が発揮できないというようなことがしば

しば起こることは想像に難くないだろう。つまりテストスコア x とは、θ の値についての、ノイズを含んだシグナルに過ぎない。したがって、観察した x に基づく θ の推論は、決定論的なものではなく不確実性を含んだ確率的なものとなる。

以上の議論は、テストスコア x が特定の値であることを観察した上で、ある学生の能力（努力水準）が高い（或いは低い）条件付確率 $\Pr(\theta = \theta_H|x), \Pr(\theta = \theta_L|x)$ が[1]、学生に対する評価形成を定式化する上で最も重要な要素になることを示唆している。

もし、観察した x という情報をどのように処理して学生の能力や努力水準に対する見込みを形成するのかということを表現するための標準的な手法が既に存在するならば[2]、残る問題はテストというものの性質をどのように定式化すればよいかということに専ら帰着する。直観的に言えば、これは、能力 / 過去の努力水準が高い（或いは低い）学生がこの試験を受けた場合に、どのようなスコアを取ることが見込まれるのかということを考えるということである。そこで、能力 / 過去の努力水準が高い（或いは低い）学生がこのテストを受けたときに特定のスコア x をとる見込み[3]をそれぞれ、$f(x|\theta = \theta_H), f(x|\theta = \theta_L)$ と表すこととしよう[4]。

もちろん、例えば、「$\theta = \theta_H$ の場合は、x は平均〇〇の××分布に従い、$\theta = \theta_L$ の場合は平均△△の××分布に従う」というように $f(x|\cdot)$ の形状を具体的に特定してしまうという方法もあるが、これはモデルに強い仮定を置くことを意味し、その適用可能性を狭めてしまうおそれがある。他方、$f(x|\cdot)$ に何らの制約も設けずに分析を進めようとしても、これは例えば能力や努力とスコアが全く相関しない、或いは負に相関してしまうような、

(1) 混乱が生じるおそれは小さいように思われるので、以下では $\Pr(\theta_H|x), \Pr(\theta_L|x)$ とそれぞれ略記する。

(2) 以下でも用いるが、これがいわゆるベイズ更新（Bayesian update）と呼ばれる枠組である。

(3) なぜここでは「確率」という用語を使わないかについては、2.2 で簡単に述べる。

(4) (1)と同様に、以下では $f(x|\theta_H), f(x|\theta_L)$ とそれぞれ略記する。

機能していない / おかしな試験を行う場合も一緒くたにして分析しているに等しいから、何らか意味のある分析上の含意が導かれることはあまり期待できないということは想像に難くないだろう。

それでは、モデルの射程範囲をできる限り狭めることなく、しかし何らか明確な含意を導くような、いわば一般性を保ち且つ意義のある分析モデルを構築するためには、$f(x|\cdot)$ にどのような制約をかけることが望ましいのであろうか。この点について具体的に検討するために、能力の高い / 努力した学生と能力の低い / 努力しなかった学生それぞれ 10 人ずつが受けたと仮定した場合に、以下のような成績分布になるような二つの試験について考えてみよう。

	低得点 x_1	中得点 x_2	高得点 x_3
$\theta = \theta_H$	1	2	7
$\theta = \theta_L$	7	2	1

表 1　良い試験

	低得点 x_1	中得点 x_2	高得点 x_3
$\theta = \theta_H$	2	5	3
$\theta = \theta_L$	3	5	2

表 2　悪い試験

単純化のため、この二つの試験において全ての受験者は、低得点、中程度の得点、高得点の三つのうちいずれかのスコアをとるとし、それぞれを x_1, x_2, x_3 と表記する。無論ここでは、$x_1 < x_2 < x_3$ が成り立つ。

表 1 で表される試験はいわばうまく作られた試験、“しっかりと差のつく” 試験である。能力が高い / 努力した学生の多くは高得点をとっており、能力の低い / 努力しなかった学生の多くは低得点をとっている。しかし他方、 H, L いずれにもそれ以外の得点をとった学生がいるという意味で当

然ノイズは含まれている。

これに対して、表 2 で表される試験は、いわばあまりうまく作られていない試験である。H, L いずれについても中くらいの得点をとった学生が最も多く、かつ高得点をとった学生も低得点をとった学生もそれぞれに複数いる。しかし他方、文字通りの “差のつかない” 試験ではないということにも注意が必要である。つまり、高得点と低得点の人数の違いに表れているように、平均としてはわずかに H の方の得点が高くなっており、ノイズは大きいが、試験として最低限の機能はしているといえよう。

このように、上記の二つの試験は、含まれるノイズの大きさが異なるという意味で多様だが、全く機能していない / おかしな試験ではないという意味では共通している。ここでの課題は上記の二つのような試験双方を含むような形で、いわば病的な振る舞いはしないような $f(x|\cdot)$ を一般的に特徴づけるにはどうすればよいかということである。

再び上記の二つの表に戻る。ここで、$f(x|\theta_H)$ の $f(x|\theta_L)$ に対する比、$\frac{f(x|\theta_H)}{f(x|\theta_L)}$ に注目してみよう。$f(x|\cdot)$ は、θ を所与とした場合にその学生がスコア x をとる見込みを表すものであったから、各表のセルの数字をその代理指標として用いてみる。すると、表 1 で表される試験については：

$$\frac{f(x_1|\theta_H)}{f(x_1|\theta_L)} = \frac{1}{7}, \frac{f(x_2|\theta_H)}{f(x_2|\theta_L)} = \frac{2}{2} = 1, \frac{f(x_3|\theta_H)}{f(x_3|\theta_L)} = \frac{7}{1} = 7$$

同様に、表 2 で表される試験については：

$$\frac{f(x_1|\theta_H)}{f(x_1|\theta_L)} = \frac{2}{3}, \frac{f(x_2|\theta_H)}{f(x_2|\theta_L)} = \frac{5}{5} = 1, \frac{f(x_3|\theta_H)}{f(x_3|\theta_L)} = \frac{3}{2}$$

となり、いずれの場合も：

$$\frac{f(x_1|\theta_H)}{f(x_1|\theta_L)} < \frac{f(x_2|\theta_H)}{f(x_2|\theta_L)} < \frac{f(x_3|\theta_H)}{f(x_3|\theta_L)}$$

が成り立っていることが分かる。すなわち、上記の二つの試験は、ノイズの程度が大きく異なるものでありながら、$\frac{f(x|\theta_H)}{f(x|\theta_L)}$ が x について単調増加（strictly increasing）であるという性質を共有しているのである。詳細については後ほど述べるが、実はこの性質がまさに、本稿の題材である「単

調尤度比性 （monotone likelihood ratio property, MLRP）」の具体的な例[5]となっている。

これは一見極めて技術的な性質に見えるが、上記のテストの例について考えれば、低いスコアに関して、能力の高い / 努力した学生がそのスコアをとる見込みは、能力の低い / 努力しなかった学生のそれに比して小さくなり、高いスコアに関してはその逆になるはずだ、というのは直観的にも十分首肯できるものであろう[6]。

MLRP は、ノイズを含むシグナル x から意思決定において重要となる θ の値について推論するという、政治的意思決定においても頻繁に問題になる場面をモデル化しようとする場合において、シグナル x がいわば自然な振る舞いをすることを保証するための条件を提供してくれる。

以下では、MLRP のより正確な定義についてあらためて述べたうえで、それが一般的な意思決定や政治的意思決定のモデル化にとっていかに有用であるのかについて説明していく。

2.2 累積分布関数と確率密度関数

冒頭でも述べたように、MLRP は特定の確率分布が満たす性質であるため、そもそも確率分布なるものを記述する手立てがなければその定義を与えることができない。そこでまずいくつかの用語を導入しよう[7]。

サイコロの目のように、様々な数値を確率的にとる変数 X を確率変数（random variable）と呼ぶ。Fair なサイコロであれば、各目が出る確率

---

(5) 後述するように、正確には「厳密な単調尤度比性（strict MLRP）」の例である。

(6) 中程度の得点に関しては、$\theta = \theta_H$ の学生がわずかに能力（努力）に比して低い点を取り、$\theta = \theta_L$ がわずかに能力（努力）に比して高い得点をとることは、能力（努力）からかけ離れた点数を記録することより起こりやすいはずだ、というふうに考えれば、中程度のスコアを記録する見込みの比が、低得点と高得点に関するそれの間の値をとるということも直観的に理解できるであろう。

(7) 以下の内容は、直観的な理解を優先した記述となっているため、確率論に関する諸事項のより正確な解説については、専門のテキスト（ex. Bertsekas and Tsitsiklis 2008 など）を参照されたい。

は同様に確からしい（すなわち $\frac{1}{6}$ ずつ）から、サイコロを一回振った時に出る目を表す確率変数 X は一様な（uniform）分布に従う、といったように、確率変数とそれが従う分布について考えることができる。

サイコロの目の場合は、X は $1, 2, 3, 4, 5, 6$ というとびとびの値しかとらない。こうした確率変数を離散型確率変数（discrete random variable）というのに対して、例えば区間 $[0,1]$ の間の任意の値をとる、というような確率変数は連続型確率変数（continuous random variable）と呼ばれる。

一般に、確率変数が従う分布は、より正確には以下で定義するような関数を用いて記述される。

定義 1

確率変数 X が特定の値 x 以下の値をとる確率を与える関数 $F: \mathbb{R} \to [0,1]$ を、累積分布関数（cumulative distribution function, CDF）と呼ぶ。すなわち：

$$F(x) = \Pr(X \leq x)$$

X が離散型確率変数の場合、X が各値をとる確率 $\Pr(X = x)$ を考えることができるから[8]、単純に $F(x) = \sum_{y \leq x} \Pr(X = y)$ が成り立つ。確率変数 X が従う分布を記述する $F(x)$ と X が各値をとる“見込み”（ここでは確率として表現される）の間に非常に直観的な関係が成立するのである。

しかし、X が連続型確率変数の場合は、話はそう単純ではない。例えば、区間 $[0,1]$ の間の任意の実数値をとり、かつ各値をとる“見込み”は同様に確からしいような X を考えてみよう[9]。X が $[0,1]$ の間の任意の実数 x をとる確率 $\Pr(X = x)$ はどうなるであろうか。もし、これが正の値であると仮定すると、$[0,1]$ の間の実数 x は無数に存在するから、$\sum \Pr(X = x)$ は無限大に発散してしまい、“何かが起こる”確率は 1 となるという確率

(8) これを与える関数は、確率質量関数（probability mass function）と呼ばれる。
(9) すなわち、X は区間 $[0,1]$ の一様分布に従う $(X \sim U[0,1])$。

の公理を満たさなくなってしまう。したがって、このとき、$\Pr(X = x)$ は厳密に 0 とならざるを得ないので、我々が慣れ親しんだ確率という概念をもって、X が各値をとる“見込み”を表現することはできないのである。

しかし、今考えているような区間 $[0,1]$ で一様に分布する X について、例えば、$0 \leq X \leq \frac{1}{2}$ となる確率は $\frac{1}{2}$ であることは直観的に明らかであろう[10]。すなわち、X が厳密に一点をとる確率は 0 となってしまうが、X がある区間に収まる確率は正の値をとりうる。そこで、連続型の確率変数については、以下のような関数を用いて、X が各値をとる“見込み”を表現する。

定義 2

確率変数 X が区間 $a \leq X \leq b$ の値をとる確率を以下のように与える関数 $f: \mathbb{R} \to \mathbb{R}_+$ を確率密度関数 （probability density function, pdf） と呼ぶ：

$$\Pr(a \leq X \leq b) = \int_a^b f(x)dx$$

このように、連続型確率変数には技術的に難しい面もあるが、他方、微積分[11]を用いることができるなど、かえって離散型より取り扱いやすい場合もある。例えば、2.1 のテストの例でいえば、テストスコアは、現実には（日本式には）0 から 100 までの整数値をとる離散型確率変数と考えるべきであろうが、101 通りの値に対して確率を与える確率質量関数を定義するのは煩雑であるから、便宜的に 0 から 100 の間の任意の実数値をとる連続型確率変数としてモデル化したほうが便利であるかもしれない。

MLRP は、シグナルを特徴づける確率変数が離散型である場合も連続型である場合も（すなわち確率質量関数、確率密度関数いずれに対しても）同様に定義することができるが、応用上、連続型確率変数を用いてシグナ

---

(10) なぜなら、それが各値の出る見込みは同様に確からしい、或いは一様分布に従うということが意味するところだからである。

(11) $F(x) = \int_{-\infty}^{x} f(t)dt$ であるので、累積分布関数と確率密度関数の間には、$F' = f$ の関係が成立する。

ルをモデル化することが多いため、以下では確率密度関数を用いて議論していく。

2.3 定義

ようやく MLRP の定義を与える準備が整った。MLRP は、最も一般的には任意の異なる確率密度関数に対して定義することができるが、応用上は一つのパラメータ $\theta \in \Theta \subset \mathbb{R}$ だけで違いが特徴づけられるような（条件付）確率密度関数に対する定義が用いられる。そこで本稿においても Milgrom （1981, p383） などに倣い、以下のように MLRP を定義する。

定義 3

条件付確率密度関数 $f(x|\cdot)$ が、任意の[12] $\underline{x} < \overline{x}$ および $\underline{\theta} < \overline{\theta}$ に対して、

$$\frac{f(\underline{x}|\overline{\theta})}{f(\underline{x}|\underline{\theta})} \leq \frac{f(\overline{x}|\overline{\theta})}{f(\overline{x}|\underline{\theta})}$$

を満たすとき、f は単調尤度比性 （monotone likelihood ratio property, MLRP）を持つという。

また、

$$\frac{f(\underline{x}|\overline{\theta})}{f(\underline{x}|\underline{\theta})} < \frac{f(\overline{x}|\overline{\theta})}{f(\overline{x}|\underline{\theta})}$$

を満たすとき、f は厳密な単調尤度比性（strict MLRP）を持つという[13]。

2.1 の具体例における説明に立ちかえれば、その意味するところは明らかであろう。この定義では、(具体例でいうところの) テストスコアを三つの水準に、能力 / 努力水準を二つの水準に限定しないという形で一般化が

---

(12) $\underline{x}, \overline{x}$ については、無論、確率変数 X がとりうる値、すなわち f のサポート（support） に属することが必要である。

(13) 3.1、4.2 における具体例も含め、応用上はこちらの strict MLRP がよく用いられる。そういった場合、単に「MLRP を仮定する」と説明されていても、実際には strict MLRP が仮定されていることも多い。

図られている[14]。

$\frac{f(x|\overline{\theta})}{f(x|\underline{\theta})}$ が x について微分可能であるとき、定義 3 は実用上より便利な形に言い換えられる。

定義 3.1

条件付確率密度関数 $f(x|\cdot)$ が、任意の $\underline{\theta} < \overline{\theta}$ に対して、

$$\frac{\partial}{\partial x}\left(\frac{f(x|\overline{\theta})}{f(x|\underline{\theta})}\right) \geq 0$$

を満たすとき、f は MLRP を持つという。

また、

$$\frac{\partial}{\partial x}\left(\frac{f(x|\overline{\theta})}{f(x|\underline{\theta})}\right) > 0$$

を満たすとき、f は strict MLRP を持つという。

2.4 例：正規分布

MLRP について学ぶ上で、特定の確率分布が MLRP の定義を満たすかを実際に確認してみることも有用であろう。指数分布やポアソン分布（平均を θ として）、一様分布 （区間を $[0,\theta]$ として）などの代表的な分布が MLRP を持つことが知られているが（cf. Milgrom 1981, p383）、以下では政治学の数理モデルでもしばしば用いられる正規分布が、分散既知（一定）、平均を θ とした場合に MLRP を持つことを実際に確認してみよう。

平均 μ 、分散 σ^2 である正規分布の確率密度関数 $f(x|\mu,\sigma^2)$ は：

$$f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \ (x \in \mathbb{R})$$

であることが知られている。このとき、$\underline{\theta} < \overline{\theta}$ とすると、$\frac{f(x|\overline{\theta},\sigma^2)}{f(x|\underline{\theta},\sigma^2)}$ は以

(14) 本稿では、政治学における応用例との関連から、以後も θ については最も単純な 2 値のケースのみを扱う。

下のようになる。

$$
\begin{aligned}
\frac{f\left(x\middle|\overline{\theta},\sigma^2\right)}{f\left(x\middle|\underline{\theta},\sigma^2\right)} &= \frac{\exp\left(-\dfrac{\left(x-\overline{\theta}\right)^2}{2\sigma^2}\right)}{\exp\left(-\dfrac{\left(x-\underline{\theta}\right)^2}{2\sigma^2}\right)} \\
&= \exp\left\{-\frac{\left(x-\overline{\theta}\right)^2}{2\sigma^2}-\left(-\frac{\left(x-\underline{\theta}\right)^2}{2\sigma^2}\right)\right\} \\
&= \exp\left[-\frac{1}{2\sigma^2}\left\{\left(x-\overline{\theta}\right)^2-\left(x-\underline{\theta}\right)^2\right\}\right] \\
&= \exp\left[-\frac{1}{2\sigma^2}\left\{\left(x-\overline{\theta}\right)+\left(x-\underline{\theta}\right)\right\}\left\{\left(x-\overline{\theta}\right)-\left(x-\underline{\theta}\right)\right\}\right] \\
&= \exp\left[\frac{\overline{\theta}-\underline{\theta}}{2\sigma^2}\left\{2x-\left(\overline{\theta}+\underline{\theta}\right)\right\}\right]
\end{aligned}
$$

よって、

$$
\frac{\partial}{\partial x}\left\{\frac{f\left(x\middle|\overline{\theta},\sigma^2\right)}{f\left(x\middle|\underline{\theta},\sigma^2\right)}\right\}=\frac{\overline{\theta}-\underline{\theta}}{\sigma^2}\exp\left[\frac{\overline{\theta}-\underline{\theta}}{2\sigma^2}\left\{2x-\left(\overline{\theta}+\underline{\theta}\right)\right\}\right]>0 \quad \because \underline{\theta}<\overline{\theta}
$$

より、分散一定の場合、平均を θ とする正規分布は確かに MLRP、それも strict MLRP を持つことが分かる。

3　何がうれしいのか

ここまでで、具体例の検討から MLRP を用いるモチベーションを示し、MLRP の定義を与えることを終えた。本項では、ノイズを含んだシグナルを MLRP によって特徴づけることの有用性について、まず 2.1 と同様の例を用いてあらためて明確にしたのち、より技術的・実用的な側面からも検討する。

3.1 例：能力を推論する

2.1 で用いた具体例においては、テストスコア x というシグナルから、学生の“能力もしくは過去の努力水準”について推論する、というように両者を合わせて述べたが、前者はあるプレイヤー（ex. 学生）の属性などの

パラメータ（ex. その学生の能力）についての情報を一部のプレイヤーのみ（ex. 当人である学生）が保有しており、他のプレイヤー（ex. 教員）はそのパラメータを直接には観察できないという「隠された情報（hidden information）」に起因する、いわゆる逆選択（adverse selection）の問題、後者は、あるプレイヤー（ex. 学生）の過去の行動（ex. 過去の努力水準）を、他のプレイヤー（ex. 教員）が直接観察できないという「隠された行動（hidden action）」に起因するいわゆるモラルハザード（moral hazard）の問題に対応する[15]。

具体例の説明において、学生の能力と過去の努力水準どちらを推論の対象と考えても説明が成立した（であろう）ことからもわかるように、逆選択、モラルハザードいずれに関する研究においても（能力もしくは過去の努力水準に関する）ノイズを含んだシグナルから情報を得るという状況は共通しているので、どちらの場合も（strict）MLRP は標準的な仮定として用いられる。

とはいっても、両者は厳密には異なるクラスのモデルであり、正確な用語で両者を同時に語ることには困難もあるため[16]、以下では政治学における応用例としてより数が多い逆選択の状況を念頭において話を進めることとする。

それでは、2.1 における例の検討に戻ろう。あらためて述べると、今、教員はある学生の能力 $\theta \in \{\theta_L, \theta_H\}$ $(\theta_L < \theta_H)$ をテストスコア x から推論することを考えている。テストスコアの下限を l 、上限を u とし、x は $[l, u]$ の間の任意の実数をとる連続型確率変数 X の実現値であるとしよう。

---

(15) 逆選択やモラルハザードについては、上述の伊藤（2003）、Laffont and Martimort（2002）などを参照されたい。

(16) 逆選択は通常、パラメータの値について私的情報を持つプレイヤーが存在する不完備情報のモデルとして扱われるが、モラルハザードの場合、過去の行動は厳密にはパラメータでなく、それ自体が選択変数であり、形式的には不完全情報のモデルであるという違いがある。不完備情報や不完全情報については、例えば Tadelis（2013）などのゲーム理論のテキストを参照されたい。

学生の能力が高いとき、その学生がテストスコア x をとる見込みは、条件付確率密度関数 $f(x|\theta_H)$ によって表される（同様に能力の低い場合は $f(x|\theta_L)$ によって表される）ものとする。

そして、テストスコアを特徴づける条件付確率密度関数 $f(x|\cdot)$ は、strict MLRP を持つと仮定しよう。すなわち、$\frac{f(x|\theta_H)}{f(x|\theta_L)}$ は x について単調増加であるものとする。

2.1 における具体例の検討では、テストが最低限、全く機能しない / おかしなものでないということ、言い換えれば意思決定者である教員が行う推論において有益な情報をわずかにでも含む、という制約をテストにかける必要があるという観点から、（strict） MLRP をいわば帰納的に導いたのであった。では今、テストスコア x が strict MLRP を持つ分布に従うと仮定した場合に、教員の行う推論はどのような性質を持つのであろうか。実際に確認してみよう。

当該学生が H である（能力 $\theta = \theta_H$ を持つ）という事前確率（テストスコア x を観察する前に、教員が学生の能力に対して持っていた主観的見込み）を $\Pr(\theta = \theta_H) = p \in (0,1)$ とすると[17]、テストスコア x を観察したうえで（所与として）学生が H である条件付確率 $\Pr(\theta_H|x)$ は、ベイズの定理 （Bayes' Theorem）を用いて[18]、以下のように求められる[19]。

$$\Pr(\theta_H|x) = \frac{f(x|\theta_H)\Pr(\theta_H)}{f(x|\theta_H)\Pr(\theta_H) + f(x|\theta_L)\Pr(\theta_L)}$$

---

(17) これまでと同様に、以下では $\Pr(\theta_H)$ と略記する。

(18) 事象 A, B について、B が起きたことを所与として A が起きる条件付確率 $\Pr(A|B)$は、$\Pr(A|B) = \frac{\Pr(B|A)\Pr(A)}{\Pr(B)}$ によって求められるというのがベイズの定理である。ただし、とりわけ政治学の論文においては明確に言及されないが、このように定式化されたベイズの定理を、確率密度を用いる場合（つまり確率変数が厳密に一点の値をとることが事象として定義できない場合）に拡張してよいのかは、本来は自明ではない。しかし幸いなことに、単に確率（質量関数）を確率密度関数に置き換える形で計算をしても問題がないことが知られている（詳細は Bertsekas and Tsitsiklis 2008, ch.3 などを参照）。

(19) 諸々の前提から、$pf(x|\theta_H) > 0$ であることに注意。

$$= \frac{pf(x|\theta_H)}{pf(x|\theta_H) + (1-p)f(x|\theta_L)}$$
$$= \frac{1}{1 + \frac{1-p}{p} * \frac{f(x|\theta_L)}{f(x|\theta_H)}}$$

p は定数であるから、x の値が変化したときに $\Pr(\theta_H|x)$ がどう変化するかは、$\frac{f(x|\theta_L)}{f(x|\theta_H)}$ の変化にのみ依存する。$\frac{f(x|\theta_L)}{f(x|\theta_H)} > 0$ は分母にあり、$\frac{1-p}{p} > 0$ であるから、$\frac{f(x|\theta_L)}{f(x|\theta_H)}$ が x について単調減少であれば、$\Pr(\theta_H|x)$ は x について単調増加である。一般に、関数 $k(x)$ が単調増加、すなわち $k'(x) > 0$ のとき、$\left(\frac{1}{k(x)}\right)' = -\frac{k'(x)}{\{k(x)\}^2} < 0$ が成り立つから、strict MLRP より、$\Pr(\theta_H|x)$ は x について単調増加である。

また、上式を変形すると $\frac{f(x|\theta_L)}{f(x|\theta_H)} = \frac{p}{1-p}\left(\frac{1}{\Pr(\theta_H|x)} - 1\right)$ となるから、$\Pr(\theta_H|x)$ が x について単調増加であるとき、$\frac{f(x|\theta_L)}{f(x|\theta_H)}$ が x について単調減少、すなわち $f(x|\cdot)$ が strict MLRP を持つことは明らかである。

以上の議論は、$f(x|\cdot)$ が strict MLRP を持つことは、$\Pr(\theta_H|x)$ が x について単調増加であるための必要十分条件であることを意味している。$\Pr(\theta_H|x)$ が x について単調増加であるということは、具体例でいえば、テストスコア x が高くなるほど、学生の能力が高いという見込みが大きくなることを意味しており、我々がテストというものに当然に期待する性質であろう。

テストスコアが従う分布が strict MLRP を持つことを仮定するという、一見極めて技術的な制約を設けることは、点が良い学生ほど能力が高い見込みが大きいという関係が成立することを単に保証するだけではなく、そのような現実に即した性質をテストが持つことと全くの同値なのである。より一般的に言えば、ノイズを含むシグナルが従う分布が strict MLRP を持つということは、値が大きなシグナル（ex. 良い試験の成績、良い政策パフォーマンス etc.）を観察することが、（意思決定者にとっても推論の対象であるパラメータの値が大きいこと（ex. 良い学生、良い候補者である etc.）が望ましいという前提の下で）“良い知らせ （good news, cf. Milgrom 1981）”であるということと同値である。

3.2 分布のその他の重要な性質に対する十分性 [20]

前項で述べた基本的な有用性に加えて、MLRP はその他いくつかの重要な確率分布の性質の十分条件になっていることが知られている [21] 。以下ではそうした性質のうち、一つを例にとって紹介する。

ここでは、MLRP を持つ $f(x|\cdot)$ に従う、$[l,u]$ の間の任意の実数をとる連続型確率変数 X とその実現値 x について考えよう [22]。f を特徴づけるパラメータは、$\theta \in \Theta \subset \mathbb{R}$ とする。定義 3 より、$[l,u]$ 内の任意の $\underline{x} < \overline{x}$ および $\underline{\theta} < \overline{\theta}$ に対して：

$$\frac{f(\underline{x}|\overline{\theta})}{f(\underline{x}|\underline{\theta})} \leq \frac{f(\overline{x}|\overline{\theta})}{f(\overline{x}|\underline{\theta})}$$

$$\Leftrightarrow f(\underline{x}|\overline{\theta})f(\overline{x}|\underline{\theta}) \leq f(\overline{x}|\overline{\theta})f(\underline{x}|\underline{\theta}) \tag{1}$$

が成り立つ。上式は、$\underline{x} = \overline{x}$ の場合にも自明に成り立つから、結局任意の $\underline{x} \leq \overline{x}$ および $\underline{\theta} < \overline{\theta}$ について成り立つ。

まず、式 (1)の両辺を、l から $\overline{x}$ まで、$\underline{x}$ で積分すると、以下を得る。

$$\int_l^{\overline{x}} f(\underline{x}|\overline{\theta})f(\overline{x}|\underline{\theta})d\underline{x} \leq \int_l^{\overline{x}} f(\overline{x}|\overline{\theta})f(\underline{x}|\underline{\theta})d\underline{x}$$

$$\Leftrightarrow f(\overline{x}|\underline{\theta})F(\overline{x}|\overline{\theta}) \leq f(\overline{x}|\overline{\theta})F(\overline{x}|\underline{\theta}) \tag{2}$$

同様に、式 (1)の両辺を、$\underline{x}$ から u まで $\overline{x}$ で積分すると、以下を得る。

$$\int_{\underline{x}}^{u} f(\underline{x}|\overline{\theta})f(\overline{x}|\underline{\theta})d\overline{x} \leq \int_{\underline{x}}^{u} f(\overline{x}|\overline{\theta})f(\underline{x}|\underline{\theta})d\overline{x}$$

$$\Leftrightarrow f(\underline{x}|\overline{\theta})\{1 - F(\underline{x}|\underline{\theta})\} \leq f(\underline{x}|\underline{\theta})\{1 - F(\underline{x}|\overline{\theta})\} \tag{3}$$

式 (2)を変形すると、

---

(20) 本項の内容についての詳細は、Wolfstetter（1999, ch.4）などを参照。

(21) 技術的に厳密なまとめとして、Hopkins and Kornienko（2003） がある。

(22) strict MLRP は MLRP の十分条件であるので、ここでの十分性の議論は strict MLRP にも当然妥当する。

$$(2) \Leftrightarrow \frac{F(\overline{x}|\overline{\theta})}{F(\overline{x}|\underline{\theta})} \leq \frac{f(\overline{x}|\overline{\theta})}{f(\overline{x}|\underline{\theta})}$$

$\underline{x} = \overline{x}$ が許容されるので、これは任意の $\overline{x}$ について成り立ち、よって以下が成り立つ。

$$\frac{F(x|\overline{\theta})}{F(x|\underline{\theta})} \leq \frac{f(x|\overline{\theta})}{f(x|\underline{\theta})} \tag{2$'$}$$

同様に、

$$(3) \Leftrightarrow \frac{f(\underline{x}|\overline{\theta})}{f(\underline{x}|\underline{\theta})} \leq \frac{1-F(\underline{x}|\overline{\theta})}{1-F(\underline{x}|\underline{\theta})}$$

より以下が成り立つ[23]。

$$\frac{f(x|\overline{\theta})}{f(x|\underline{\theta})} \leq \frac{1-F(x|\overline{\theta})}{1-F(x|\underline{\theta})} \tag{3$'$}$$

式 (2)′と式 (3)′を併せて、

$$\frac{F(x|\overline{\theta})}{F(x|\underline{\theta})} \leq \frac{1-F(x|\overline{\theta})}{1-F(x|\underline{\theta})}$$

$$\Leftrightarrow F(x|\overline{\theta}) \leq F(x|\underline{\theta}) \tag{$*$}$$

式 (*)は、1 次確率支配 （first-order stochastic dominance, FOSD）と呼ばれる性質であり[24]、$f(x|\cdot)$ が MLRP を持つことは、任意の $\underline{\theta} < \overline{\theta}$ に対して、$F(x|\overline{\theta})$ が $F(x|\underline{\theta})$ を 1 次確率支配することの、十分条件となっている。しかし他方、FOSD は MLRP を必ずしも含意しない、すなわち、MLRP は FOSD の十分条件ではあるが必要条件ではないことが知られている。これはつまり、MLRP が FOSD よりも強い条件であることを意味している[25]。

---

(23) (3)′ 式を変形して得られる $\frac{f(x|\overline{\theta})}{1-F(x|\overline{\theta})} \leq \frac{f(x|\underline{\theta})}{1-F(x|\underline{\theta})}$ は単調ハザード率（monotone hazard rate）と呼ばれ、MLRP が十分条件となるその他の重要な分布の性質の一つである。

(24) 正確には少なくとも一つの x について、厳密に不等号が成立する必要がある。

(25) この点については、5 節でもあらためて言及する。

FOSD は、$F(\cdot|\overline{\theta})$ と $F(\cdot|\underline{\theta})$ にそれぞれ従う 2 つの確率変数がある場合に、任意の x について、$F(\cdot|\overline{\theta})$ に従う確率変数の方が x 以下の値をとる確率が小さいことを意味している。より具体的には、$F(\cdot|\overline{\theta})$ と$F(\cdot|\underline{\theta})$ にそれぞれ従い、x を獲得額とする二つのくじ（或いはギャンブル）のうち、どちらが望ましいかという観点から考えると理解がしやすい。つまり、獲得額が x 以下である確率は、任意の x について、$F(\cdot|\underline{\theta})$ に従うくじの方が大きいので、$F(\cdot|\overline{\theta})$ が $F(\cdot|\underline{\theta})$ を 1 次確率支配するとき、もしどちらのくじを引くかを選択できるならば、$F(\cdot|\overline{\theta})$ に従うくじを引く方が得になる。FOSD はこのような形でくじ（つまり累積分布関数）の間の順序付けを行うのである。

FOSD の概念を導入すると、前項で述べた、「『シグナル x が大きいことが θ がより大きな方の値をとる見込みが大きいことを示唆すること』の必要十分条件は strict MLRP である」という関係に対してより技術的な裏付けを与えることができる。

実は、3.1 の具体例における「$\Pr(\theta_H|x)$ が x について単調増加であること」は、「x に条件づけられた θ の累積分布関数を $G(\cdot|x)$ とするとき、任意の $\underline{x} < \overline{x}$ について、$G(\cdot|\overline{x})$ が $G(\cdot|\underline{x})$ を一次確率支配すること」という形に一般化できる(26) （Milgrom 1981, p.383 Proposition 1.）。すなわち、シグナル x が大きいことが θ がより大きな方の値をとる見込みが大きいことを示唆する、という性質を $G(\cdot|x)$ を FOSD によって順序付けることでより一般的に記述できるのである。

「任意の $\underline{x} < \overline{x}$ について、$G(\cdot|\overline{x})$ が $G(\cdot|\underline{x})$ を 1 次確率支配することの必

(26) この点について、3.1 の具体例のように、θ が 2 値の場合については簡単に確認できる。任意の $\underline{x} < \overline{x}$ について、$G(\cdot|\overline{x})$が $G(\cdot|\underline{x})$ を一次確率支配することは、具体例においては、$G(\theta_H|\overline{x}) \leq G(\theta_H|\underline{x})$ かつ、$G(\theta_L|\overline{x}) \leq G(\theta_L|\underline{x})$ が成り立ち、かつどちらかは厳密な不等号が成立する（注（24））ことに等しい。$G(\theta_H|\overline{x}) = G(\theta_H|\underline{x}) = 1$ は明らかであり、θ_L 以下の θ は θ_L のみであるため、結局この条件は $\Pr(\theta_L|x)$ が x について単調減少であること、すなわち $\Pr(\theta_H|x)$ が x について単調増加であることに一致する。

要十分条件が $f(x|\cdot)$ が strict MLRP を持つことである」ということと、「$f(x|\cdot)$ が MLRP を持つことは、任意の $\underline{\theta} < \overline{\theta}$ について $F(x|\overline{\theta})$ が $F(x|\underline{\theta})$ を 1 次確率支配することの十分条件ではあるが必要条件ではない[27]」ということを併せ考えると、「任意の $\underline{\theta} < \overline{\theta}$ について $F(x|\overline{\theta})$ が $F(x|\underline{\theta})$ を 1 次確率支配することは、任意の $\underline{x} < \overline{x}$ について、$G(\cdot|\overline{x})$ が $G(\cdot|\underline{x})$ を 1 次確率支配すること（これは上述のようにシグナル x が大きいことが θ がより大きな方の値をとる見込みが大きいことを示唆する、という性質に等しい）の十分条件ではない」ということが分かる。

このことは、数理的な分析がいかに我々の直観を規律してくれるかということ（そして （strict）MLRP という仮定がいかに重要かつ有用であるかということ）を示してくれているように思われる。というのも冒頭の具体例において、低いスコアに関して、能力の高い / 努力した学生がそのスコアをとる見込みは、能力の低い / 努力しなかった学生のそれに比して小さくなり、高いスコアに関してはその逆になるはずだ、といった形で MLRP の意味するところを言葉での説明によっても直観的に理解することが可能であることを述べた。このような説明を受けた事後においては、これがテストが“機能していない / おかしな試験でない”ための条件であることは自明の理のように思えてくる。しかし他方、「『もし任意の点数について、その点数以下をとる確率は、能力が高い場合より低い場合の方が大きい（これはすなわち FOSD である）』ということを仮定すれば、テストを“機能していない / おかしな試験でない”ものとして定式化できる」というふうに説明を受けたとしたら、それもまた至極当然なことのように思えてくるのではないだろうか。しかし、上述の説明から明らかなように、“機能していない / おかしな試験でない”ことを、「シグナル x が大きいことが θ がより大きな方の値をとる見込みが大きいことを示唆すること」というふうに定義すると、これは必ずしも成り立たない。このように数理

(27) したがって、strict MLRP が FOSD の十分条件であるが必要条件ではないことも明らかである。

分析は、分析者が意図したことを明確に表現し、様々な性質の間の関係を厳密に検討することを可能にすることで、しばしば我々を欺く（自然言語的）直観を正しく導いてくれるのである。

最後に[28]、本項では、MLRP がその他いくつかの重要な確率分布の性質の十分条件になっていることを紹介したが、これは単なる技術的な問題ではなく実用上も大きな利点をもたらしてくれる。というのも、実際の意思決定の分析においては、それぞれの選択が意思決定者にもたらす利得（あるいは効用）の大小を比較することで最も合理的な選択は何かということを検討するわけであるが、この作業はモデルの設定が抽象的である（一般性が高い）ほど困難であることが多い。MLRP は、少なくとも政治学の数理分析においては比較的一般性が高い仮定の部類に入るが、MLRP を仮定することで、FOSD や単調ハザード率など確率密度関数や累積分布関数の間の様々な不等式が得られるため、利得（効用）の評価が容易になる場合がある[29]。MLRP はいわば、モデルの一般性と tractability の間のバランスが良い仮定ということもできよう。

4　政治学における具体例

ここまでは、ノイズを含むシグナル x から意思決定において重要となるパラメータ θ の値について推論するという状況に着目して MLRP について説明してきたが、未知のパラメータ θ についての推論は、意思決定そのものではない（テストの例でいうと、学生の能力の推論ではなく、単位

(28)　極めて技術的な内容となるため本稿では紹介しないが、MLRP は、最適な選択がパラメータの値の変化と共に単調に変化するための十分条件を検討するいわゆる「単調比較静学 （monotone comparative statics）」の手法を不確実性下の意思決定に拡張する際にも大きな役割を果たしている。詳細は Athey (2002) を参照（政治学者向けの平易な解説として Ashworth and Bueno de Mesquita 2006, pp. 227-229.）。

(29)　MLRP から派生する様々な不等式関係を駆使して利得を評価する例としては、後述の Li (2021) などがある。

を付与するか否かが意思決定である)。本節では、政治学における具体的な問題を例に、意思決定まで含めてモデル化する場合に MLRP が果たす役割について検討する。

以下では、政治学における中心的な研究対象である選挙、とりわけ選挙アカウンタビリティの問題を具体例として扱うが、その前に、選挙での投票を含め政治学が分析するような政治的意思決定の特質について簡単に述べる。

4.1　経済学と政治学における本人 - 代理人問題

政治学の数理分析において、MLRP が用いられるのは多くの場合、有権者と政治家或いは政治家と官僚のような、委任 - 責任関係（本人 - 代理人関係）における委任側（本人側）の意思決定である。本稿でも何度か言及したように、経済学の研究においても MLRP は逆選択やモラルハザードなど本人 - 代理人関係の分析において広く用いられてきたため、政治学においてもいわば自然な形で応用が進んだということができよう。

しかし（もちろん全てにおいて当てはまるわけではないが一般的な傾向として)、経済学と政治学がそれぞれ問題とする本人 - 代理人関係には違いも存在する。

例えば、経済学におけるモラルハザードの初期の代表的な研究である Holmström（1979）は、労働者が選択する連続変数である努力水準（雇用者に利益をもたらす成果（アウトプット）は、労働者の努力水準の増加関数だが、ノイズを含む）が観察できない場合に、雇用者は成果に応じてどのような報酬を与えればよいか、という問題を検討した。そして、労働者の最大の努力を引き出すという意味で最適な報酬体系が、成果についての増加関数になる（すなわち、成果が大きいほど報酬が高いといういわば常識的な関係が成立する）ための必要十分条件は、シグナルである成果が従う（努力水準に条件付けられた）確率密度関数が strict MLRP をもつことである、すなわち strict MLRP を仮定することが、未知の量に対する推論のみならず意思決定も単調性を持つことと同値であることを示したのであ

る[30]。

ここで重要なのは、経済学で問題となる本人 - 代理人関係においてはしばしば、雇用者などの本人は、金銭報酬など、シグナルの値に対して連続的に変化させることのできる、代理人に対する柔軟な事後コントロールの手段を持つことが仮定されているということである。

これに対して政治学で問題となる本人 - 代理人関係において、本人は、再選させる / させない、再任する / 解任する、といったような（しばしば 2 値の）“ 荒い ” コントロール手段しか持ち合わせていないことが多いという特徴がある。

このような状況において、本人の意思決定が単調性を持つということは、観察したシグナルの値が一定の閾値（threshold/cutoff）以上であれば再選、或いは再任するといったような、いわゆる閾値戦略をとることが本人にとって最適な選択となることを意味する。

以下では、政治学における選挙アカウンタビリティの数理モデルを具体例として、シグナルが従う分布に strict MLRP を仮定することが、本人の閾値戦略を導くことを見てみよう。

4.2　例：選挙アカウンタビリティのモデル

狭義の選挙アカウンタビリティのモデルは、現職の再選動機を通じて最大限の努力を引き出すという、モラルハザードのクラスに属するものだが、近年は、こうした狭義のアカウンタビリティの確保だけでなく、より良い（ex. 能力の高い）政治家を選抜するというのも選挙の重要な役割であり、狭義のアカウンタビリティと選抜（selection）双方でもって選挙アカウンタビリティという問題を構成するという理解が広く浸透している[31]。

---

(30) すなわち、累積分布関数の FOSD 条件では十分ではないということであり、3.2 における議論と全くのパラレルである。

(31) 選挙アカウンタビリティに関するレビューとしては、曽我（2015）, Ashworth（2012）などを参照。

その結果、近年のアカウンタビリティのモデルはモラルハザードと逆選択双方の要素を含むことが多いが、本稿は全体を通じて能力に対する推論の問題に着目して議論を進めてきたため、以下でも選抜の問題に限定して説明できるようモデルを簡略化して紹介する。

政治家（現職）の選抜の問題の基本的な構造は、本稿が扱ってきた能力の推論の問題と同じである。すなわち、現職の能力の（ノイズを含んだ）シグナルとして、何らかの政策パフォーマンス（ex. 経済指標）を観察し、それに基づいて現職の能力を推論する。有権者の目的を、より能力が高い政治家を選挙後に戴くことに限定すれば、有権者はシグナルを用いて更新した現職の能力への評価と、挑戦者（野党候補者）の潜在的な能力への見込みを比較して、より高い方に投票する。

ただし、こうしたシグナルとしての政策パフォーマンスをモデル化する方法として、(1) シグナルを、3.1 における具体例のように能力によって条件付けられる確率密度関数に従うとしてモデル化する場合と、(2) シグナルの生成過程をより具体的に設定し、能力などに依存する体系的な部分とノイズの部分を分けてモデル化する場合があるので、以下ではその双方について検討する。

4.2.1　シグナルが従う条件付確率密度関数を用いたモデル化

以下では、Li（2021）のモデルで用いられた設定を簡略化して[32]紹介する。(代表的) 有権者が、現職政治家を再選させるか、挑戦者[33]（野党候補）を当選させるかを選択しようとしているとしよう。現職政治家および挑戦者の能力は高いか低いかの 2 値であるとし ($\theta \in \{\theta_L, \theta_H\},\ \theta_L < \theta_H$)、現職、挑戦者いずれも能力が高い事前確率は $\Pr(\theta_H) = p \in (0,1)$ であるとす

(32) Li（2021）のモデルは、正確には政治家と、官僚などの政治的任用者の間の関係を検討したものであり、選挙アカウンタビリティのモデルではないが、全く同様の設定を選挙に応用することができるため、ここでは選挙の文脈で説明する。

(33) 単純化のため、ここでは候補者は現職を含めた 2 人のみであるとする。

る[34]。有権者は投票の前に、現職政治家の政策パフォーマンス x を観察する。x は任意の実数値をとるものとし $(x \in \mathbb{R})$ 、能力 θ に条件付けられた確率密度関数 $f(x|\cdot)$ に従うものとする。そして（当然最も重要なことだが）f は strict MLRP を持つと仮定しよう。

ここまでで明らかなように、この設定は、3.1 の具体例とほぼ全く同じものである。したがって、これまでと異なる部分は、現職の能力に対する推論を経ていかに意思決定（投票先の決定）を行うかということであるが、上述のように、有権者が（選挙に勝利した）政治家の能力の高低のみに関心があるというふうに単純化すれば、政策パフォーマンスを観察した後に現職の能力が高い確率が、挑戦者の能力が高い確率[35] 以上[36] であれば、すなわち $\Pr(\theta_H|x) \geq p$ であれば有権者は現職に投票することとなる。

$\Pr(\theta_H|x)$ はベイズの定理を用いて 3.1 の例と全く同様に求めることができるから、$\Pr(\theta_H|x) \geq p$ という条件は以下のように変形できる：

$$\begin{aligned} &\Pr(\theta_H|x) \geq p \\ \Leftrightarrow\ &\frac{pf(x|\theta_H)}{pf(x|\theta_H) + (1-p)f(x|\theta_L)} \geq p \\ \Leftrightarrow\ &\frac{f(x|\theta_H)}{pf(x|\theta_H) + (1-p)f(x|\theta_L)} \geq 1 \\ \Leftrightarrow\ &\frac{f(x|\theta_H)}{f(x|\theta_L)} \geq 1 \end{aligned} \tag{4}$$

$f(x|\cdot)$ は strict MLRP を持つので、式 （4） の左辺は x について単調増

(34) すなわち、両者の間に事前の能力の違いはないと有権者が考えているということである。これは政策パフォーマンスというシグナルの影響をよりはっきりさせるための仮定であり、当然、より現実的に現職と挑戦者それぞれに異なる能力の事前分布を想定するモデルもある（例として、Ashworth et al.（2018）など）。

(35) 挑戦者（野党候補）は政策に影響を与える立場にないと考えれば、有権者が挑戦者の能力に関する見込みを更新する機会はないから、これは p のままである（選挙キャンペーンなど、挑戦者の能力のシグナルが存在する場合について検討したものとしては、Dewan and Hortala-vallve（2019）などがある）。

(36) 単純化のため、両者が等しい場合は現職に投票するものとする。

加である。今 x は任意の実数値をとるから、単調性と非有界性から、$\lim_{x\to-\infty}\frac{f(x|\theta_H)}{f(x|\theta_L)}=0, \lim_{x\to\infty}\frac{f(x|\theta_H)}{f(x|\theta_L)}=\infty$ が言える。したがって、$\frac{f(x^*|\theta_H)}{f(x^*|\theta_L)}=1$ を満たす x^* は必ず存在し、かつ一意である。

すなわち、$f(x|\cdot)$ が strict MLRP を持つとき[37]、有権者は政策パフォーマンスが一定以上（$\frac{f(x^*|\theta_H)}{f(x^*|\theta_L)}=1$ を満たす x^* 以上）であれば現職を再選させ、一定より小さければ挑戦者を当選させる。

このように、シグナルの分布の形状を具体的に特定せずとも、strict MLRP を仮定することで、閾値戦略という直観に合致した意思決定が、本人にとって最も合理的な選択として導かれるのである。

4.2.2　シグナルを分解してモデル化

それでは、今度はシグナルを体系的な部分とノイズの部分に分けてモデル化する方法を紹介しよう。以下の例は Ashworth and Bueno de Mesquita（2014）におけるモデルの一つを簡略化したものである。

モデルの説明に入る前に、4.2.1 における例のように、シグナルの生成過程の詳細には立ち入らない形で抽象的にモデル化をすることも可能であるにもかかわらず、シグナルの構造をより具体的にモデル化するというのは奇異に思われるかもしれない。しかし、政策パフォーマンスというシグナルに影響を与えるのは能力だけとは限らず、例えば政治家の過去の努力水準なども当然影響すると考えられるであろう。このように複数の要因がシグナルに体系的な影響を与える状況をモデル化したい場合、本稿で扱ってきたような 1 パラメータの確率密度関数について定義された MLRP では、シグナルが従う確率密度関数の性質を十分に特定することができない。そのため、技術的に過度に難解にならない形で、複数の体系的要因＋ノイズからなるシグナルをモデル化したい場合は、より具体的にその生成過程

(37)　正確には、x が任意の実数値をとるといういわゆる full support の仮定と併せてという限定がつく。

を特定する必要がしばしば出てくるのである。

以下の例においてもシグナルの構造を除く基本的な設定は 4.2.1 の例と同様である。大きな違いは、今回はシグナルの観察前に、現職政治家が努力 $a \in \mathbb{R}_+$ を投じたことを有権者が観察しており[38]、こうした努力 a は政策パフォーマンスを高めると考える。より具体的には、政策パフォーマンス x は以下のように定まるものとする。

$$x = g(a, \theta) + \varepsilon$$

関数 g はパフォーマンス x の体系的部分を定める関数であり、努力 a 、能力 θ の増加関数である。これに対して、ε はノイズを表す。パフォーマンスの体系的な部分は既に g によってモデル化されているから、ε の性質としては、バイアスがない（平均が 0）と仮定するのがよく、かつ大きいノイズほど生じにくいとした方が現実的であろう。そこで、ε は標準正規分布（平均 0 分散 1）に従うものと仮定しよう[39]。

有権者は a の大きさを観察し[40]、それを所与としてさらに x を観察する。あとは 4.2.1 における例と同様に、$\Pr(\theta_H | x, a) \geq p$ であれば、現職を再選させることとなる。

$\Pr(\theta_H | x, a)$ は、標準正規分布の確率密度関数を ϕ とすると、ベイズの定理により以下のように求められる[41]。

---

(38) Ashworth and Bueno de Mesquita（2014）は、現職の努力が観察できる / できない場合の比較検討を行うことがその主たる内容であるが、ここではより単純なケースについて検討する。なお、現職の努力が観察できないモデルは、モラルハザードと逆選択の問題を同時に分析することになる。

(39) 分散を 1 と仮定するのは単に簡便のためである。

(40) Ashworth and Bueno de Mesquita（2014）のモデルは、現職自体も自らの能力が正確には分からないといういわゆる career-concern モデルをベースにしているため、a を観察したことによるシグナリングの効果については考慮する必要がない。career-concern とシグナリングについては、上條（2020, pp.5-9.）などを参照されたい。

(41) a, θ を所与とした場合に特定の値 x が観察される確率密度は、ε が $x - g(a, \theta)$ となる確率密度に等しいことに注意。

$$\Pr(\theta_H|x,a) = \frac{p\phi\{x - g(a,\theta_H)\}}{p\phi\{x - g(a,\theta_H)\} + (1-p)\phi\{x - g(a,\theta_L)\}}$$

したがって、

$$\Pr(\theta_H|x,a) \geq p$$

$$\Leftrightarrow \frac{p\phi\{x - g(a,\theta_H)\}}{p\phi\{x - g(a,\theta_H)\} + (1-p)\phi\{x - g(a,\theta_L)\}} \geq p$$

$$\Leftrightarrow \frac{\phi\{x - g(a,\theta_H)\}}{\phi\{x - g(a,\theta_L)\}} \geq 1 \tag{5}$$

となる。式（5）は 4.2.1 の例における式（4）に似ているが、先ほどは $f(x|\cdot)$ が strict MLRP を満たすことが仮定されていたから、左辺が x について単調増加であることは明らかだったが、今回は一見即座には判断できない。しかし幸いなことに、2.4 で確認したように、ノイズが従う正規分布が strict MLRP を持つことによって、式（5）の左辺が x について単調増加であることが保証される(42)。

したがって、4.2.1 の場合と同様に、$\frac{\phi\{x^* - g(a,\theta_H)\}}{\phi\{x^* - g(a,\theta_L)\}} = 1$ を満たす x^* が必ず存在し、かつ一意であることが分かる。より具体的には、

$$\frac{\phi\{x^* - g(a,\theta_H)\}}{\phi\{x^* - g(a,\theta_L)\}} = 1$$

$$\Leftrightarrow \phi\{x^* - g(a,\theta_H)\} = \phi\{x^* - g(a,\theta_L)\}$$

となるが、標準正規分布は $x = 0$ を軸に対称な単峰形の分布であり、$g(a,\theta_H) \neq g(a,\theta_L)$ であるから、

---

(42) 2.4 で示した正規分布の確率密度関数に $\mu = 0, \sigma^2 = 1$ を代入すれば ϕ が得られるから、$\frac{\phi\{x - g(a,\theta_H)\}}{\phi\{x - g(a,\theta_L)\}}$ を計算して x で微分すれば、これが単調増加であることは容易に示すことができる。しかし、そのような計算をせずとも、シグナルの構造 $x = g(a,\theta) + \varepsilon$ についてあらためて考えると、能力が高い（低い）場合のシグナルは結局、平均 $g(a,\theta_H)$（$g(a,\theta_L)$）、分散 1 の正規分布にしたがうことが分かる。実際、平均 $g(a,\theta_H)$、 分散 1 の正規分布に従うシグナルが x をとる確率密度と、標準正規分布に従うノイズが $x - g(a,\theta_H)$ をとる確率密度は等しくなる。したがって、$g(a,\theta_L) < g(a,\theta_H)$ であることと正規分布が strict MLRP を持つことから、式（5）の左辺が x について単調増加であることが分かる。

$$\Leftrightarrow x^* - g(a, \theta_H) = -\{x^* - g(a, \theta_L)\}$$

$$\Leftrightarrow x^* = \frac{g(a, \theta_H) + g(a, \theta_L)}{2}$$

となる。すなわち、この例においては、政策パフォーマンス x が $x^* = \frac{g(a,\theta_H)+g(a,\theta_L)}{2}$ 以上であれば、現職を再選させるという形で、やはり先ほどの例と同様に strict MLRP（正確にはノイズの正規性）によって閾値戦略の最適性が導かれる。

他方、こちらの例においては、（正規分布という具体的な分布を仮定する等、モデルの抽象度は多少下がるものの）現職の努力水準がシグナルに与える影響を明示的にモデル化できているので、こうした有権者の選択を念頭に置いて、現職政治家がどのように努力水準を決定するかといった、戦略的な相互作用に分析の射程を広げることができる。

5　終わりに

本稿では、（厳密な）単調尤度比性 （（strict） MLRP） という確率分布の性質の解説を通じて、ノイズを含んだシグナルを用いた推論という、政治的意思決定も含めた様々な意思決定において現実に問題になる状況が、いかに数理的に分析できるかについて検討してきた。そこではシグナルが従う分布に （strict） MLRP を仮定することで、我々が現実に行っているような、（成果給や閾値戦略といった）単調性を持つ意思決定ルールが導かれることが示された。

MLRP の数理分析における有用性についてはこれまでの各節で様々な観点から繰り返し述べてきたため、ここでは最後に一点の留保を述べて本稿の締めくくりに代えたい。というのも、本稿では、MLRP が、できるだけ一般性を保ちつつ、不確実性下の意思決定に単純な性質を与えることができる仮定である旨を強調してきたが、他方、FOSD との関係について述べた 3.2 でも少し言及したように、MLRP はそれなりに “ 強い ” 仮定であるということもしばしば指摘されている。こうした懸念から、経済学に

おいては、常に MLRP を仮定することから生じる弊害について既に検討がなされている[43]。

幸か不幸か、政治学の数理分析においては、MLRP の仮定が過度に制約的であるという指摘がなされるのを目にする機会は今のところないが、本稿で扱った選挙アカウンタビリティなど、非常に重要なテーマに関するモデルで MLRP は用いられているから、今後、MLRP を仮定した標準的なモデルの理論的 / 実証的な不整合が広く指摘されれば、政治学においてもさらなる技術的な検討が進むかもしれない。

（そもそもそのようなものが存在し得ないのは明らかであるから、あらためて述べるほどのことでもないのだが）MLRP は、ノイズを含むシグナルからの推論を伴う意思決定をモデル化する際に、常にそれを仮定すればうまくいくというような魔法の杖ではない。しかし、何が理論的 / 実証的アノマリーなのかということは、ベンチマークとなるモデルがあって初めて判断できることである。MLRP はそのための適切な出発点を提供してくれるものといえるだろう。

本稿が扱った MLRP は、数理分析における一つの技術的トピックに過ぎないが、極めて抽象的な演繹的手法でありながら、具体的な現実の意思決定を分析するものであるという難しさ / 面白さを持つ（社会科学における）数理分析を学ぶ / 用いる上で、大きな示唆を持つものと言えよう。

〔付記〕本研究は JSPS 科研費（22K13337）の助成を受けたものである。

(43) このような指摘の例として、サラニエ（2010, p.77）などを参照。

参考文献

Ashworth, Scott. 2012. “Electoral Accountability: Recent Theoretical and Empirical Work.” *Annual Review of Political Science* 15（1）: 183–201.

Ashworth, Scott, and Ethan Bueno de Mesquita. 2006. “Monotone Comparative Statics for Models of Politics.” *American Journal of Political Science* 50（1）: 214–31.

Ashworth, Scott, and Ethan Bueno de Mesquita. 2014. “Is Voter Competence Good for Voters?: Information, Rationality, and Democratic Performance.” *American Political Science Review* 108（3）: 565–87.

Ashworth, Scott, Ethan Bueno de Mesquita, and Amanda Friedenberg. 2018. “Learning about Voter Rationality.” *American Journal of Political Science* 62（1）: 37–54.

Athey, Susan. 2002. “Monotone Comparative Statics under Uncertainty.” *The Quarterly Journal of Economics* 117（1）: 187–223.

Bertsekas, Dimitri P., and John N. Tsitsiklis. 2008. *Introduction to Probability, Second Edition.* Athena Scientific.

Dewan, Torun, and Rafael Hortala-Vallve. 2019. “Electoral Competition, Control and Learning.” *British Journal of Political Science* 49（3）: 923–39.

Dixit, Avinash K., and Jörgen W. Weibull. 2007. “Political Polarization.” *Proceedings of the National Academy of Sciences* 104（18）: 7351–56.

Holmström, Bengt. 1979. “Moral Hazard and Observability.” *The Bell Journal of Economics* 10（1）: 74–91.

Hopkins, Ed, and Tatiana Kornienko. 2003. “Ratio Orderings and Comparative Statics.” working paper.

伊藤秀史 2003.『契約の経済理論』有斐閣

上條諒貴 2020.「首相の “ 責任 ” は追及すべきか① ― 対称情報モデルによる検討 ― 」『北九州市立大学法政論集』48 （1/2） pp. 1-38

Laffont, Jean-Jacques, and David Martimort. 2002. *The Theory of Incentives: The Principal-Agent Model.* Princeton University Press.

Li, Christopher. 2021. “Indirect Accountability of Political Appointees.” *Journal of Theoretical Politics* 33（3）: 383–96.

Li, Hao, and Wing Suen. 2004. “Delegating Decisions to Experts.” *Journal of Political Economy* 112（S1）: S311–35.

McCarty, Nolan, and Adam Meirowitz. 2007. *Political Game Theory: An Introduction.* Cambridge University Press

Milgrom, Paul R. 1981. “Good News and Bad News: Representation Theorems and Applications.” *The Bell Journal of Economics* 12（2）: 380–91.

Prat, Andrea. 2002. “Campaign Advertising and Voter Welfare.” *The Review of Economic Studies* 69（4）: 999–1017.

サラニエ, ベルナール（細江守紀・三浦功・堀宣昭 訳）2010.『契約の経済学（第二版）』勁草書房

曽我謙悟 2015.「選挙アカウンタビリティの構造」高橋百合子編『アカウンタビリティ改革の政治学』有斐閣

Tadelis, Steven. 2013. *Game Theory: An Introduction.* Princeton University Press.

Wolfstetter, Elmar. 1999. *Topics in Microeconomics: Industrial Organization, Auctions, and Incentives.* Cambridge University Press.

Reprinted from

KITAKYUSHU SHIRITSU DAIGAKU HOU-SEI RONSHU

Journal of Law and Political Science. Vol. L No. 1/2

October 2022

A Note on Monotone Likelihood Ratio Property: For Political Scientists

KAMIJO Akitaka