事実認定論の数理構造とその基礎について

弁護士　木本　茂樹

平成 29 年 7 月 8 日

概要

事実認定は，要証事実に対する裁判官の合理的な信念の度合いが一定の確信を超えるかという問題であるから，論理一貫した理論を構築するのであれば，確率論を用いるのは自然である．このとき，要証事実に対する内心は，その信念の度合いとしての確率とその分布として表現できる．そして，確率論から，証拠が加わると，オッズの期待値という意味では，心証度は真実に合致する方向に高まることが期待でき，一定の条件を満たす証拠が無限に存在すれば，真実に対する確信が得られることが示される．事実認定においては，このような確率論から導かれる性質からすれば，まずは，当事者の証拠へのアクセスを確保すべきであるという結論が導かれる．そして，それが確保されない場合には，木本 (2015) で示された反事実などの概念を用いてその事実を不利益に斟酌することで，心証度が証明度を超えるか否かという事実認定の基本的な判断枠組みを維持した形で，適切な結論を得ることができる．

Keywords: 事実認定，ベイズ論，確率的推論，解明度

0　用語・記法

確率変数は，U，E というようにイタリック体の大文字で，確率変数がとる値は，u，e のようにイタリック体の小文字で表す．確率変数が複数あり，これに番号を付して区別する場合，$U_1, U_2, \cdots$ や $e_1, e_2, \cdots$ というように右下に添字を付ける．確率変数 U がとる値を番号を付して区別するときは，右上に添字を付ける．例えば，当事者 X が証明すべき事実を u^X，確率変数 M_i が取りうる j 番目の値を m_i^j などと表す．また，確率変数がとる値の集合を $u^1, u^2, \cdots, u^n\}$ や $\{e\}$ というように $\{\}$ で表す．複数の確率変数や複数の確率変数のとる値をまとめて表すときは，ボールド体を用いる．$\boldsymbol{E} = \{E_1, E_2, \cdots, E_n\}$，$\boldsymbol{e} = \{e_1, e_2, \cdots, e_n\}$ などである．

1　はじめに

事実認定の問題とは，与えられた証拠，あるいは証拠などから導かれる裁判の前提となるべき事実（証拠的事実）のもとで，証明の対象となる事実（要証事実）についての裁判官の確信が一定の程度（証明度）を超えているかという問題である．したがって，証拠的事実を $\boldsymbol{e}$, 当事者 X が証明すべき要証事実を u^X とする[*1]と，確率を用いれば，要証事実に対する裁判官の確信の程度は，証拠的事実のもとでの要証事実についての条件付き確率確率，すなわち，$p(u^X|\boldsymbol{e})$ と表すことができるから，

[*1] 議論を単純にするため，本稿では，要証事実が単一の確率変数で表されるとするが議論に本質的な違いはない．確率変数が複数ある場合の処理については，木本 [7] 参照．

証明度を p^* とすると，事実認定の問題を，

$$p(\boldsymbol{u}^X|\boldsymbol{e}) > p^* \tag{1}$$

という不等式の成否を判断する問題として定式化するのは自然なことである．しかし，Finkelstein and Fairley[21] が事実認定に確率論を用いるべきとの問題提起をしてからすでに 40 年以上が経つが，現在においても，確率論の使用が，実務からも研究者からも全面的に支持されられているとはいいがたい[*2]．

確率論が受け入れられていない原因としては，確率論を現実の事案にどのように適用するかという実践的な問題のほかに，確率論の理論的な基礎についての批判や事実認定論の基本構造についての批判がある．後者は，事実認定は 1 回限りの歴史的な事実であり，確率論を用いて論じることはできないとか，事実認定の問題はそもそも単純な確率の多寡だけでは判断できないという批判である．しかしながら，これらの批判は，確率論についての誤った理解に基づくものであり，事実認定論を確率論の基礎の上に，これと整合する形で構築することは十分可能である．

本稿では，事実認定の問題に確率論を用いるにあたって検討すべき理論的な諸問題について考察し，事実認定論の数理構造とその基礎にある前提を明らかにすることを通じて，事実認定論の理論的な枠組みが確率論と整合的であることを示すことを目的とする．具体的には，次節で，Cox[20] の議論を紹介し，1 回限りの多様な出来事についても，なぜ事実認定に確率論を用いるのが適切なのかについて述べる．続いて，第 3 節では，確率的推論の基本的な構造から導かれる特質について論じ，証

[*2] 確率論を用いることについての否定的な見解としては，例えば，Tribe[34] のほか，Cohen[19]，長谷部 [11]．

拠を集めれば集めるほど真実への蓋然性が高まるという経験則が数学的に裏付けられることを示し，第 4 節では，確率論に対する批判の一つであるパラドックスの存在と事前確率の問題について考察する．第 5 節では，太田 [2] で提唱された解明度の問題を確率論の中でどのように位置づけるかや，心証度以外にどのような概念が必要であり，それが確率論でどのように定義づけられるかについて，考察し，解明度を心証度の分散を用いて定義するのが適切であることを明らかにする．最後に，第 6 節では，証明度と事実認定の基本的構造について論じ，解明義務についても証明度の増減ではなく，心証度の判断構造の問題と理解することで，事実認定の問題は，確率論の基本理論と整合する単純な理論的構造を持つことを明らかにする．

2　信念の度合いと確率

2.1　信念の度合いと確率

確率については，確率とは何かという観点から議論されることも少なくない．例えば，Gilles (2006) は，学説を確率を (1) 主観的な信念の度合いとする主観説，(2) 合理的な信念の度合いとする論理説，(3) 対象に内在する傾向であるとする傾向説，(4) 事象の発生する頻度であるとする頻度説の 4 つに分けている．

しかしながら，確率とは何かという議論の立て方をすると，信念の度合いに確率論を用いることの実用上の困難と理論上の困難とが区別されず，その実用上の困難から 1 回限りの事実についての信念の度合いを確率論で表すことは出来ないというような議論がなされることになってしまい妥当ではない．現実の問題への適用に困難がある場合がある

からといって，その理論が妥当ではないことにはならないからである．

確率論は，コルモゴロフ [29] によって定式化された数学（測度論）の一分野として理解した上で，確率論の枠組みを現実世界のどのような問題に適用できるかという観点から考えるべきである．このような観点から，信念の度合いについて，ブール代数などの一定の論理的な前提をもとに，理論を構築しようとすれば，それが必然的に確率論かこれと同等の理論を用いることになることを最初に示したのが Cox[20] である．Cox[20] は，

1. 命題の真偽に対する信念の度合いには強弱があり，
2. その信念の度合いについて，冪等則，交換則，結合則，吸収則，分配則といったブール代数の諸法則を満たし，
3. ある命題の否定の信念の度合いが，その命題の信念の度合いのみによって決まる，

などの条件を満たすとすると，そのような信念の度合いについての理論は，確率論かそれと同等の理論になることを示した．この意味で，確率論は，「ブール代数を不確実性を含む場合に拡張したもの」*3である．

信念の度合いに確率論を用いることに対する批判としては，確率が主観的になり，客観性が担保されないという批判がある．次節でみるように，確率的推論では，事前確率と得られた情報（証拠）に基づいて信念の度合いを更新していくから，事前確率や証拠が異なれば結論も異なりうるが，これらが同じであれば，同じ結論が得られる．事前確率をどのように定めるかという問題はあり，この点については，第 4 節で検討するが，確率的推論には客観性ないし間主観性を担保する基礎がある．

事実認定も，認定すべき事実についての裁判官の主観的な信念の度合

*3 Bishop[16]（村田監訳）p.21.

いが問われている．特に，事実認定においては，例えば，違法収集証拠や時期に遅れた攻撃防御方法であるとして，証拠が排除され，その結果，真実とは異なる事実が認定されることがあるが，このような場合においても，裁判官の主観によってどのような結論でも許されるわけではなく，一定の合理性や客観性が必要である．この意味でも，事実認定においても合理的な信念の度合いとしての確率論を用いるのは自然である．

2.2　曖昧さ・情報の不完全性と確率

確率論は，標本空間とその部分集合族（σ-加法族），確率測度の 3 つからなる確率空間の上に成り立っている．つまり，議論の対象となる事象の集合としての標本空間と確率を定める対象となる部分集合族を定めた上で，その部分集合の確率を論じることになるから，これらが定義されていなければ，確率をそもそも定めることができない．例えば，ある部分集合 A の要素が定義されておらず，a が A に属するか否かがわからなければ，A の確率を厳密に定義することはできない．

ところが，日常生活の用語や法律用語は，必ずしもその範囲が明確に定められていたり，定義が明確にされていないことがしばしばある．例えば，「あの人は背が高い」という発言があったときに「背が高い」というのが何を意味するのかは明確には定められず，曖昧さが残る．同様に，法令の用語も必ずしも明確にその範囲が定義されているわけではなく，どこまで含まれるか争いが残ることが少なくない．例えば，窃盗は「他人の財物を窃取」すること（刑法 235 条）と定められているが，何が「他人の財物を窃取」にあたるのかは解釈に委ねられている．もっといえば，法解釈の問題は，本質的には，ある事実が法の定める特定の集

合の要素であるかを判断する問題にほかならないともいえる．

そうすると，そもそも用語の意味が厳密に定義されていない現実の諸問題について，確率論を適用することができないのではないかという批判が考えられる．例えば，ファジー理論[*4]などは，このような問題意識から作られた理論である．しかしながら，この点については，Pearl[32] や Pearl[33] で説明されているように，理論的には，確率の分布を考えることで解決される．例えば，ある事象 a の生起確率が別の事象 $B = \{b^1, b^2, \cdots, b^n\}$ の結果によって異なるとする．すると，全確率の公式から，a の確率 $p(a)$ は，$p(a) = \sum_{i=1}^{n} p(a, b^i) = \sum_{i=1}^{n} p(a|b^i)p(b^i)$ と表される．つまり，信念の度合いとしての a の確率 $p(a)$ は，b^i の値によってさまざまな値 $p(a|b^i)$ をとり，その期待値が $p(a)$ となるに過ぎない．このとき，$p(a) = \sum_{i=1}^{n} p(a|b^i)p(b^i)$ において $p(b_i)$ は，a である確率が $p(a|b^i)$ となる確率を表している，つまり，$p(b_i)$ は確率の確率であるとも理解できる．このように考えると，2 つの事象 a と α について，両者の確率が同じであっても，「確率の確率」は異なりうるから，単に確率が同じでも内心の度合いは異なることを説明できる．さらに，b^i についても，別の事象 $C = \{c^1, c^2, \cdots, c^m\}$ によって生起確率が異なるのであれば，$p(b^i) = \sum_{j=1}^{m} p(b^i, c^j)$ を考える．このようにして，ある事象について不確かさが残るのであれば，その不確かさの内容によって場合分けし，その確率を定めていくことで，より適切に内心の状態を表現できるようになる．

「確率の確率」という概念に対しては，2 次の確率で不十分であれば 3 次の確率，3 次で不十分であれば 4 次というように際限なく高次の確率を観念することになり，無限後退しかねないという批判がある．しか

[*4] ファジー理論については，Zadeh and Kacprizyk[36] 参照．

し，確率の確率は，ある確率事象について，別の事象を考慮の対象に加え，それによる場合分けしたときの同時確率であるから，考慮の対象となる他の事象の数が有限であれば，確率の確率が無限に続くということにはならないし，3 次以上の確率についても，同時確率を考えれば，結局，2 次までの確率に還元することができる[*5]．具体的には，上の例でいえば，$\sum_i p(a|b_i) \sum_j p(b^i|c_i^j) = \sum_{i,j} p(a|b^i, c^j)$ と表されるから，a についての条件付確率と b, c の同時分布の 2 つに還元される．

したがって，事実認定についても，内心の確信の度合いを示すには，信念の度合いとしての心証度だけでは足りず，その分布を考える必要があるが，これで足り，その分布が無限後退したりすることはない．

また，事実認定の問題は，不等式が成り立つか否かの問題であるから，必ずしも信念の度合いについて正確な値を求める必要はない．例えば，λ がある集合 Λ に含まれていたとしても，含まれていなかったとしても結論が変わらないのであれば，例えば，どちらか一方を仮定して推論をしても差しさわりはなく，含まれているかを確定させる必要はない．この意味では，事実認定の問題に確率論を適用するために必要なのは，ある用語が何を意味しているのか，その意味を常に確定しておくことではなく，必要があれば範囲を確定することができるということであり，この意味でも事実認定に確率論を用いる前提が欠けているということにはならない．

2.3　言語情報における暗黙の了解と確率

確率的推論を行う上で注意すべきこととして，人が伝える情報には様々な省略や暗黙の了解があり，これは確率的な表現とは異なるこ

[*5] 詳しくは，Pearl[32] や Kyburg[31]，Cheeseman[18] を参照．

とがあるということが挙げられる．例えば，a の確率（蓋然性）が，$B = \{b^1, b^2\}$ によって異なるとき，B について何ら言及することなく，a の確率といったときには，確率論では，$P(a) = \sum_{1,2} P(a, b^i)$ であり，$B = b^1$ である場合と $B = b^2$ である場合の両方を考えなければならない．しかしながら，一般の会話では，B について何の言及もされていないのに双方がそれぞれ B について特定の状態（例えば，$B = b^2$）であることを前提として話をしていることがしばしばある．実際，特定の事象の蓋然性を表すためには，それに関連する事実をすべて列挙し，そのもとでの条件付き確率として表さなければならなくなってしまうから，このような省略は，伝達する情報の符号長を短くするという観点から合理的である．

同様の方法は裁判においても用いられている．例えば，民事訴訟では，証明責任と呼ばれる責任法理によって，特定の事実の存否について当事者からの主張がない場合に，その事実について特定の状態（存在又は不存在）をとるという約束で裁判が進められる[*6]．したがって，確率論にしたがって，ある事実が記録に表れていないから，その事実について何も知らない（あらゆる可能性がある）と解釈することはできない．

しかしながら，このような違いは，表現方法の違いに過ぎない．日常生活や法廷での情報伝達の表現方法は，確率論の一般的な表現とは整合していないから，確率の数式で表現されたものをそのまま「逐語訳」で文章表現にしたり，あるいは，文章で表現されたものをそのまま数式にできるわけではないことに注意する必要はあるが，このことによって確率論を用いることが理論的にできないとか，確率論に本質的な変更を加

[*6] 例えば，貸金請求訴訟で，弁済の主張がいずれの当事者からもなされなければ，裁判所が証拠から弁済の事実があったと考えたとしても，弁済の事実がないものとして裁判がなされることになる．

えなければならないということではない．

3　心証度の挙動と収束定理

3.1　心証度の変化の度合いと証拠の価値

ある証拠的事実が加わったときに，心証度の度合いがある要証事実に対する心証度をどのように変えるかという問題を考える．

条件付き確率の定義から (1) 式は，オッズ形式を用いると，

$$\frac{p(u^X,\boldsymbol{e})}{p(u^Y,\boldsymbol{e})} > \frac{p}{1-p} \tag{2}$$

と表される．左辺は，

$$\frac{p(u^X,\boldsymbol{e})}{p(u^Y,\boldsymbol{e})} = \frac{p(u^X,\boldsymbol{e}_{n-1})}{p(u^Y,\boldsymbol{e}_{n-1})} \cdot \frac{p(e_n|u^X,\boldsymbol{e}_{n-1})}{p(e_n|u^Y,\boldsymbol{e}_{n-1})} \tag{3}$$

と分解できるから，心証オッズ O_i，（$\cdot$ が真実であるときの）e^i の尤度 $l^i_\cdot$，尤度比 $r_{X/Y}$ をそれぞれ，

$$O_i := \frac{p(u^X,\boldsymbol{e}_i)}{p(u^Y,\boldsymbol{e}_i)}(i=0,1,\cdots,n.\text{ ただし，}\boldsymbol{e}_0=\phi\text{とする．}), \tag{4}$$

$$l^i_\cdot := p(e_n|u^\cdot,\boldsymbol{e}_{n-1}), \tag{5}$$

$$r_i^{X/Y} := \frac{l^i_X}{l^i_Y} \tag{6}$$

と定めると，$n-1$ 個の証拠的事実 $\boldsymbol{e}_{n-1}$ のもとでの心証オッズ O_{n-1} と，ここに n 個目の証拠 e_n が加わった後の心証オッズ O_n との間の関係は，

$$O_n = O_{n-1}\cdot r_n^{X/Y} \tag{7}$$

で表される．これは，さらに，両辺の対数をとると，

$$\log O_n = \log O_{n-1} + \log r_n^{X/Y} \tag{8}$$

となるから, $\log r_n^{X/Y}$ が正であれば心証オッズは大きく, また, $\log r_n^{X/Y}$ が負であれば心証オッズは小さく，$\log r_n^{X/Y}$ が 0 であれば心証オッズは変化しないことになり，$\log r_n^{X/Y}$ は，新しい証拠が加わったことによって心証オッズがどのように変化したかを表す．そこで，これを（証拠 $\boldsymbol{e}_{n-1}$ のもとでの u^X についての）証拠の価値と定める．(8) 式からも明らかなように，証拠価値は，要証事実だけでなく他の証拠の値にも依存するから，単に要証事実との関係である証拠的事実の持つ価値というのを定めることはできない[*7]．

3.2 平均証拠価値の非負性

今，ある未知の事実 $E = \{e^1, e^2, \cdots, e^m\}$ について，その値を知った場合の証拠価値を考える[*8]．前節でみたように $e^i (i = 1, 2, \cdots, m)$ の尤度を l_X^i とすると，e^i の証拠価値 w^i は，$\log r_n^{X/Y} = \log \frac{l_X^i}{l_Y^i}$ で与えられ，$l_X^i > l_Y^i$ であれば $w^i > 0$ であるし，$l_X^i < l_Y^i$ であれば $w^i < 0$ である．

今，$w^i > 0$（$l_X^i > l_Y^i$）となる i が存在したとすると，$\sum_i l_X^i = \sum_i l_Y^i = 1$ であるから，必ず $l_X^j < l_Y^j$ となる j も存在しなければならない．したがって，すべての $k = \{1, 2, \cdots, m\}$ について，$w^k > 0$ とな

*7 例えば，倉田 [8] では,「経験則の確率」という概念を用いているが，証拠の強さは他の証拠的事実にも依存するものであり，他の証拠的事実を抜きにして「経験則の確率」を要証事実との関係だけから定めて議論するのは適切ではない．

*8 簡単のため E が離散の値をとる場合を考えるが，連続の場合でも和を積分に変えれば議論に本質的な違いはない．

ることはなく，$w^i > 0$ となる i が存在すれば $w^j < 0$ となる j も必ず存在する．つまり，未知の事実について知った結果，その結果がどのようなものであったとしても，必ず一様にある要証事実の心証度を高めるということはあり得ず，結果によっては心証度が低くなる可能性が必ず存在する．このことは，その要証事実が真実であったとしても同様であるから，未知の事実を知ることによって，真実についての心証度が必ず高くなる（真実に対する確信が高まる）とは限らず，真実についての心証度が低くなるような結果が生じることもありうるのであり，この意味では，証拠を集めて未知の事実を知ることが，確実に真実の発見に繋がるということはできない．

しかしながら，証拠価値の期待値を考えると，以下に見るように要証事実が真実である場合とそうでない場合とではその期待値に違いがあり，証拠を知ることによって「平均的に近づくことが期待できることが示される．すなわち，u^X が真実であったとしたときの w^i の期待値 $E[w]$ は，

$$E[w] := \sum_i l_X^i w^i = \sum l_X^i \log \frac{l_X^i}{l_Y^i} \tag{9}$$

で表されるが，

$$E[w] \geq 0 \tag{10}$$

であり，かつ，等号が成立するのは，$\{l_X^i\} = \{e_Y^i\}(i = 1, 2, \cdots m)$ のときだけである[9]

証明は，カルバック・ライブラー情報量 （KL 情報量）の性質を考えればほとんど自明である．KL 情報量 $D(p||q)$ は，2 つの確率分布

[9] Hawthorne[23].

$p = \{p_i\}$ と $q = \{q_i\}$ の差異を計る尺度であり，

$$D(p||q) := \sum_i \log p_i \frac{p_i}{q_i} \tag{11}$$

で定義され，$D(p||q) \geq 0$，かつ，2 つの確率分布が等しい（$\{p_i\}$=$\{q_i\}$）ときに限って等号が成立するという性質を持つ．$\{l_X^i\}$ と $\{l_Y^i\}$ はいずれも確率分布だから，u^X が真実であるとした場合の平均証拠価値 $E[w]$ は，$\{l_X^i\}, \{l_Y^i\}$ の KL 情報量 $D(l_X||l_Y)$ にほかならない．同様に，u_Y が真実である場合には，

$$\begin{aligned} E[w] = \sum_i l_Y^i w^i &= \sum l_Y^i \log \frac{l_X^i}{l_Y^i} \\ &= -D(l_Y||l_X) \\ &\leq 0 \end{aligned}$$

となる．

このことから，u^X が真実であれば，証拠価値の期待値は正，u^X が真実でなければ（u^Y が真実であれば），証拠価値の期待値が負になることがわかる．つまり，期待値をとれば，前者の場合には心証オッズは大きくなり，u^X が真であるという確信が高まり，後者の場合には心証オッズは小さくなり，u^X が真であるという確信が低くなり，いずれにしても，期待値としては，真実に合致する方向に確信が強まることが期待できる．

3.3　ベイズの収束定理と心証度の収束

平均証拠価値の非負性は，真実についての心証オッズが期待値としては強まるというものであるから，証拠が積み重なっていけば，いずれは真実に対する確信が得られる，つまり，心証度が 1 になることが予想

される．実際，一定の条件を満たす証拠が無限にあるという前提のもとでは，以下にみるように証拠を無限に積み重ねることにより真実の命題に対する心証度がほとんど確実に 1 になることが示される[*10]．以下では，Hawthorne[23] に従い，これを「ベイズの収束定理」(Bayesian Convergence Theorem) と呼ぶ.

証拠が複数ある場合には, $\log O_i$ に (8) 式を繰り返し適用することで,

$$\log O_n = \log O_0 + \sum_{i=1}^{n} \log r_i^{X/Y} \tag{12}$$

を得る．つまり，(12) 式は，n 個の証拠的事実のもとで，心証オッズが事前オッズ O_0 に u^X についての $e_1 \sim e_n$ の証拠価値の和で表されることを示している．ここで，各証拠について，その期待値をとると，複数の証拠についてその値を知った時に心証オッズの変化がどのようになるかを知ることができる.

Hawthorne [23] では，証拠が互いに独立であると仮定して，総証拠価値はほぼ確実に（確率 1 で）無限大に発散し，したがって，真実の命題の事後確率（心証度）はほぼ確実に 1 に収束することを示している．しかし，ベイズの収束定理においては，証拠の独立性についての条件はさらに緩めることができる．科学法則のような一般的事実の探求と異なり，特定事実（歴史的事実）の探求では，証拠が独立ではない場合が少なくない上，前の証拠の結果によって取り調べる証拠を選択する場合には，証拠が互いに独立にならない．そこで，以下では証拠が互いに独立であるという仮定はせず，単に，前の結果に完全には従属しないという条件でもベイズの収束定理が成立することを示す.

今，h_i を i 個の証拠的事実 $\boldsymbol{e}$ の結果を並べた履歴とする．そして，任

[*10] Hawthorne [23].

意の履歴 h_{i-1} のもとでの i 番目の証拠の証拠価値 $w_i(h_{i-1})$ について，以下の 2 つを仮定する．

1. 任意の履歴 h_{i-1} のもとで，証拠価値 w_i の期待値 $E[w_i(h_{i-1})]$ は 0 に収束しない（すべての $E[w(h_i)]$ について，$E[w_i(h_{i-1})] > \mu$ となる $\mu(> 0)$ が存在する）．
2. 任意の履歴のもとで，証拠価値 $w_i(h_{i-1})$ の分散 $V[w_i(h_{i-1})]$ は無限大に発散しない（すべての $V[w_i(h_{i-1})]$ について，$V[w_i(h_{i-1})] \leq v$ となる $v(> 0)$ が存在する）．

このとき，無限の証拠を取り調べる，すなわち，n を $+\infty$ に発散させると，以下にみるように，総証拠価値 $S = \sum_i^n w_i(h_{i-1})$ は，確率 1 で $+\infty$ に発散する．

任意の履歴 h_{i-1} のもとで得られた結果 $e_i^j(h_{i-1})$ の証拠価値 $w_i^j(h_{i-1})$ から適当な値を引いて，分散 $V[\overline{w}_i(h_{i-1})] \leq v$ であり，かつ，期待値 $E[\overline{w}_i(h_{i-1})] = \mu$ となるようにした確率変数 $\{\overline{w_i}\}$ を考える．

このとき，$\overline{w_i}$ の和の期待値 $E[\overline{S}_n]$ については，

$$E[\overline{S}_n] := \sum_i^n \overline{w}_i(h_{i-1}) = n\mu \tag{13}$$

が成り立つから，$n \to \infty$ とすると， $E[\overline{S}_n] \to \infty$ となる．

また，分散については，証拠価値 w_i について $i = 1$ から $n-1$ までの和 S_{n-1} と n 個目の証拠価値 $\{\overline{w}_n\}$ の共分散が 0 になることから，

$$V[\overline{S}_n] = \sum_i [\overline{w}_n] < nv \tag{14}$$

が成り立つ．今，チェビシェフの不等式により，任意の k について，

$$p(|\overline{S}_n - n\mu| k\sqrt{V[\overline{S}_n]}) \leq \quad \frac{1}{k^2} \tag{15}$$

が成り立つから，$n\mu > M$ となるような十分大きい n のもとで $k = (n\mu - M)/\sqrt{V[\overline{S}_n]}$ とし，さらに，(14) 式を用いると，

$$\frac{p(|\overline{S}_n - n\mu| \geq n\mu - M) \leq nv}{(n\mu - M)^2} \tag{16}$$

となる．$p(\overline{S}_n \leq M) \leq p(|\overline{S}_n - n\mu| \geq \mu - M)$ であり，また，$\overline{S}_n$ の定義から，明らかに，$p(S_n \leq M) \leq p(\overline{S}_n \leq M)$ であるから，結局，

$$p(S_n \leq M) \leq \frac{nv}{(n\mu - M)^2} \tag{17}$$

となり，

$$\lim_{n\to\infty} p(S_n \leq M) = 0 \tag{18}$$

を得る．

つまり，いかなる正の値 M に対しても，総証拠価値 S_n の値が M より小さくなる確率は n が無限大に発散すると 0 に収束する．よって，総証拠価値はほぼ確実に無限大に発散し，u^X の確率は 1 に収束する．このように，ベイズの収束定理によって，証拠が無限に積み重なっていけば，確実に真実の命題について確信（確率 1）が得られることが保証される．

3.4　考察

ベイズの収束定理は，(1) 尤度の値が正確であること，(2) 証拠の平均証拠価値が 0 に収束しないこと，(3) 証拠の平均証拠価値の分散が無限大に発散しないこと，(4) 前 2 者を満たす証拠が無限に存在することを前提としている．以下では，これらの条件がどのような場合に満たされるかや，満たされない場合の影響について考察する．

3.4.1 誤差の影響

まず，現実の裁判では，尤度の値について正確な値が求められないことも少なくない．今，推論に用いている $\{l_X^i\}$ や $\{l_Y^i\}$ が正確ではなく，その真の値は $\{\hat{l}_X^i\}$，$\{\hat{l}_Y^i\}$ である場合，つまり，$\{l_X^i\}$ や $\{l_Y^i\}$ を用いて事後確率を計算しているが，実際には，u_X が真実であるときの e^i の条件付確率が $\{\hat{l}_X^i\}$ である場合を考える．$\hat{l}_X = \{\hat{l}_X^i\}, \hat{l}^Y = \{\hat{l}_Y^i\}$ とすると，この場合における（u^X が真であるときの）平均証拠価値は，

$$
\begin{aligned}
E[w] &= \sum_i \hat{l}_X^i \log \frac{\{l_X^i\}}{\{l_Y^i\}} \\
&= \sum_i \hat{l}_X^i \log \frac{\{l_X^i\}}{\hat{l}_X^i} + \sum_i \hat{l}_X^i \log \frac{\{\hat{l}_X^i\}}{\{l_Y^i\}} \\
&= D(\hat{l}_X || l_X) - D(\hat{l}_X || l_Y) \qquad (19)
\end{aligned}
$$

となる．平均証拠価値が正になる，つまり，真実についての確信に近づくことが期待できるのは，(19) 式の右辺が正である，すなわち，真の分布 $\hat{l}_X$ と l_X との KL 情報量の方が $\hat{l}_X$ と l_Y との KL 情報量よりも小さいときであることがわかる．このように，誤差がある場合には，u^X が真実である場合でも，u^X についての平均証拠価値は理論的には負の値もとり得る，つまり，期待値としても真実の心証オッズが小さくなることがあり得，真の尤度の分布との KL 情報量の大小がそのメルクマールとなることがわかる．

3.4.2 条件付独立と平均証拠価値の収束

次に，(2) と (3) について考えると，まず，(3) については，それまでに取り調べた証拠の価値を無にするような決定的な証拠が後に必ず出

てくることはないというものであって，何ら無理な前提ではない．

他方，(2) は簡単に満たされるものではない．例えば，U に対して，$\boldsymbol{M}$ という事実があり，$\{E_i\} = \{E_1, E_2, \cdots\}$ という一連の証拠が，M のもとで U と条件付独立である，すなわち，$\{E_i\}$ が M を介して U に影響を与えているとする．そうすると，

$$p(u|\boldsymbol{e}) = \sum_M p(u|m)p(m|\boldsymbol{e}) \tag{20}$$

であるから，証拠 $\boldsymbol{e}$ が増えていっても，$p(u|\boldsymbol{e})$ の値は，$p(u|m)$ により制約され，その上限や下限を超えることはない．より一般的に言えば，証拠が十分にある場合でも，それらが直接要証事実に影響を与えるのではなく，何らかの中間事実を介して影響を与えている場合には，その中間事実が要証事実に与える影響を超えることはできない．逆に言えば，証拠的事実が無限にあったときに，要証事実の確率が 0 または 1 に収束するためには，要証事実と直接関係のある証拠的事実が無限になければならない．

3.4.3　平均証拠価値の挙動

一方，(10) 式は，(2) や (3) を前提としていないから，(1) が成り立つか，誤差がある場合でも，「(19) 式 > 0」であれば，平均証拠価値の非負性は，必ず成り立つ．つまり，どのような証拠であっても，その証拠の尤度の正確性が一定程度が担保されれば，真実により近くなる蓋然性が（平均オッズの意味で）高くなることが期待できる．

また，証拠価値 $w = \log l_X^i / l_Y^i$ であるから，w は，l_X^i を大きくし，l_Y^i を小さくして 0 に近づけることでどんな大きい値でもとれるし，逆に，l_X^i を小さくして 0 に近づけ，l_Y^i を大きくするとすることでどんな小さな値でもとれる．したがって，u^X が真実であるとすると，w が大

きく，かつ，l_X^i は 1 に近いような l_X^i と l_Y^i の組み合わせがとれる．他方，$w > \log l_X^i$ であるから，w の値を小さくするには，l_X^i を 0 に近づけて小さくしなければならない．つまり，高い確率で真実に対する確信を大きく高めるような証拠は存在しうるが，真実に反するものに対する確率を大きく高めるような証拠は，高い確率では存在し得ない．この意味でも，ベイズの収束定理を満たすような証拠が無限に存在しない場合であっても，証拠がある程度集まれば，真実に対する心証度が十分大きくなることが説明しうる．

4　事前確率と無差別の原理

4.1　事前確率

前節でみたように，一定の平均証拠価値を持つ証拠的事実が無限にあれば，事前確率がどのような値であっても，真実の心証度が 1 に収束することになる．しかし，現実には，証拠は無限にあるわけではないし，無限の証拠を取り調べるということもできない．このため，ベイズ更新は高々有限回しか行われないことになり，(8) 式の右辺の第 2 項は有限の値をとるため，事前確率の比として得られる第 1 項の $\log O_0$ の値によって結論が左右されることになるから，事前確率を任意に定めてよいとなると，客観性が担保されなくなってしまう．そこで，事前確率をどのように定めるかということが問題となる．

4.2　無差別の原理とパラドックス

事前確率の定め方については，当初，複数の事象について，ある一つが他よりも大きいと決める理由が知られていなければ，これらの事象は

それぞれ同じように確からしいと考えるべきであるという「理由不充分の原理」がラプラスやベルヌイによって提唱され，その後，Keynes[28] や Jaynes[24] は，対称性に注目して，これを無差別の原理として発展させた．　しかしながら，理由不充分の原理や無差別の原理に対しては，円に内接する正三角形の弦を本無作為に選んだときの弦が正三角形の辺よりも長くなる確率を求めるという Bertrand のパラドックスなどの例を用いて確率を一意に定められないようなパラドックスの存在などの不備が指摘されてきた*11.

事前確率の問題は，統計学の分野でも，情報がない時点での事前分布をどのように定めるかという問題として議論され，パラメータの変換に対して不変となるよう事前分布を定めるべきだとする Jeffereys[?] の事前分布や，実験における情報の欠落を最小化する Berger-Bernardo の方法*12や Haar 測度を用いる方法*13などが提唱されている．しかしながら，Kass and Wasserman[26] によれば，事前分布の設定方法について意見の一致はみられていないようである．もっとも，このことは，事前確率がいかなるものであってもよいということを意味しているわけではない．無情報下での事前分布は，何らかの対称性を反映していなければならず，何の理由もないのに極端な事前確率や不均質な事前分布を定めることが正当化されるわけではない．そして，統計学においては，データが十分に集まれば，事前分布が結論に与える影響は小さくなり，意思決定において当初の事前分布の設定の違いが結論を左右しないことも少なくない．同様に，裁判においても，事前確率の定め方には一定

*11 パラドックスについては，Jaynes[24]p.451 以下，Keynes[28]（佐藤訳）47 頁以下参照.
*12 Berger and Benardo[15].
*13 Chang and Eaves[17].

の制約があり，前節でみたベイズの収束定理から，どの事前確率を用いても結論が変わらないということは考えられるし，結論が左右されうる場合でも，事前分布がどうあるべきかを考えるよりも，さらに考慮すべき事実はないかを探した方が適切な結論に早く辿り着くことも十分考えられる．この意味では，事前分布の定め方は，理論上は，理論の基礎に関わる本質的な問題ではあるが，実践上は，事前分布が一意に定まるか，ひいては確率が一意に定まるかという問題を考える実益は少ないと思われる．

5　心証度以外の概念の必要性

5.1　心証度のばらつきとしての解明度

太田 [2] は，ベイズの定理を中心とした確率論を証明論に持ち込むに際し，審理結果の確実性を示す概念として「解明度」と「尽証度」いう概念を提唱した．解明度は，「新たな証拠で証明主題の蓋然性が変動することのない程度」，尽証度は，「情報価値のある証拠が出尽くした程度」 とされ，要証事実（証明主題）の蓋然性（確率）を示す心証度とは別の概念であり，心証度が同じ値（例えば，80%）であっても，解明度は異なりうるとされている ．

第 2.2 節で述べたとおり，ある事象に対する確信の度合いを表現するには，確率分布として捉える必要があり，点（実数）として捉えるのでは不十分である．したがって，事実認定の確率的理論を考えるにあたっても，信念の度合いの期待値としての心証度だけでは十分でない．そこで，心証の状態を表現する概念として，解明度や尽証度がどのようなものであり，それを確率論でどのように定義するかが問題となる．

5.2 相互情報量，信頼区間と解明度

5.2.1 相互情報量と解明度

太田 [4] は，解明度が情報理論における相互情報量によって正当化できるとしている．しかし，以下に述べるように相互情報量を解明度の大小の基準とすることには疑問がある．

2 つの離散確率変数 U と E についての相互情報量 $I(U;E)$ は，

$$I(U;E) = \sum_{e\,\in\,E}\sum_{u\,\in\,U} p(u,e)\log\frac{p(u,e)}{p(u)p(e)} \tag{21}$$

により定義される．(21) 式はさらに，U のエントロピー $H(U)$ と E のもとでの U の条件付きエントロピー $H(U|E)$ を用いて，

$$I(U;E) = H(U) - H(U|E) \tag{22}$$

と表すことができる[*14]．

U を要証事実のとりうる事実の確率変数，E を証拠のとりうる事実の確率変数と考えると，相互情報量 $I(U;E)$ は，証拠を取り調べることによる要証事実のエントロピーの減少の期待値を示している．したがって，相互情報量の増加が解明度の増加を意味するのであれば，それは証拠の取り調べによる要証事実のエントロピーがどれほど減少するかの期待を意味することにほかならない．そうすると，まず，相互情報量は，あくまでもエントロピーの減少の期待値を示すものであるから，過去の証拠調べによる現在の状態を示すものではなく，今後の証拠調べによってエントロピーがどれほど減少するかという期待を示すものと

*14 相互情報量，エントロピーについては，有本 [1]87 頁以下参照．

なってしまう．また，要証事実が存在するか否かという場合，すなわち，u^X と u^Y の 2 値しかとらない場合には，エントロピーは，

$$H = p(u^X)\log p(u^X) + (1-p(u^X))\log(1-p(u^X)) \tag{23}$$

と表され，$p(u^X)$ が定まれば H は一意に定まることになる．そうすると，解明度は心証度に完全に従属した概念になり，同じ心証度のもとで異なる値をとることにはならない．

また，(22) 式において，U について要証事実以外の事実も考慮する場合には，心証度に従属することにはならないが，そもそもどのような事実を考慮するのかということが問題となる．この場合，他の事実をより多く考慮すればするほど，エントロピーの値は大きくなる．したがって，それらの事実を特定しうる証拠が得られれば，相互情報量の値も大きくなる．したがって，事件とは無関係な事実を考慮の対象とすると，その事実を特定する証拠は，要証事実と全く無関係でも相互情報量が高いことになってしまい，妥当とは言えない．

5.2.2 信頼区間と心証度

これに対して，三木 [9] は，心証度を点ではなく，幅のある分布（信頼区間）で捉え，その揺らぎの幅（信頼度）が心証度の確実性を意味するという考え方を提唱している．実数としての心証度をその幅の中の特定の値（例えば，期待値や最頻値等）と考えると，同じ心証度であっても異なる幅を取りうるという意味では，心証度とは別に解明度という概念が必要であるとする問題意識とは合致する．しかしながら，信頼水準をどの程度に設定すべきかが定まらない上，三木 [9] では，信頼区間の下限を基準として証明の成否を判断する考え方をとっている[*15] が，

[*15] 三木 [9]461 頁以降．

なぜ，下限を基準とするのか合理的な根拠がない．

5.3 解明度の定義と性質

5.3.1 定義

心証度を確率分布として考えると，心証度が変動することのない程度としての解明度は，心証度分布が期待値からどれだけ散っているかによって測ることができる．そして，確率変数が期待値からどれほど離れて分布しているかの指標としては分散があり，分散は値が小さい方が散らばり具合が小さいことを考えると，解明度は，分散（$V[p(u^X|e)]$）の逆数（$1/V[p(u^X|e)]$），あるいは，分散に 1 を加えた数の逆数（$1/1+V[p(u^X|e)]$）と定義するのが自然であろう．心証度の変動が完全にない状態の解明度を 1 になるようにするのであれば，後者の方が適切であり，以下では，これを解明度とする．

5.3.2 性質

i 個の証拠的事実を考慮した後の要証事実 u^X の心証度の分布を $\{p_i\}$ とすると，このとき，$n-1$ 個の証拠調べを行う前の心証分布の分散 $V[p_{n-1}]$ と，そこからさらに 1 個の証拠調べを行った後の心証分布の分散 $V[p_n]$ を比較したときに，後者の方が前者以下になる（$V[p_n] \leq V[p_{n-1}]$ となる）保証はない．

しかし，証拠価値と同様，期待値で考えると，心証度の分布の分散は，事実を知ることで小さくなる．すなわち，n 個目の証拠のもとでの条件付き分散の期待値 $E[V[p_n]]$ と元の分散 $V[p_{n-1}]$ については，

$$V[p_{n-1}] = E[V[p_n]] - V[E[p(u^X|e_i)]], \tag{24}$$

という関係が成り立つが，分散は常に非負であるから，

$$V[p_{n-1}] \geq E[V[p_n]] \tag{25}$$

が成立する．つまり，ある事実を知った後の心証度の分散は，期待値で考えれば，元の心証度の分布の分散以下となる．解明度は，分散に 1 を加えたものの逆数であるから，このことは，解明度の調和平均の意味での期待値（逆数の期待値の逆数）が証拠を積み重ねることで大きくなることを意味している．

さらに，心証度は 0 から 1 までの値しかとれないから，心証度が高い（あるいは低い）場合には，心証度の分布も必然的に高い（あるいは低い）ところに偏ることになる．例えば，心証度の期待値 μ であるときの心証度の分布の分散の最大値 V_{max} を考えると，分散が最大になるのは，心証度が両端の 0 と 1 に集中しているときであるから，

$$\begin{aligned} V_{max} &= \mu(1-\mu)^2 + (1-\mu)\mu^2 \\ &= \mu(1-\mu) \end{aligned} \tag{26}$$

となり，V_{max} は，$\mu = 1/2$ のとき最大値 1/4 をとる一方，μ が 0 や 1 に収束すれば必然的に 0 に収束する．したがって，証拠が無限に蓄積することで，ベイズの収束定理により心証度は 0 か 1 に収束すれば，解明度も 1 に収束することになる．

5.4 尽証度

太田 [2] は，解明度とは異なる概念として，尽証度を「情報価値のある証拠が出尽くした程度」として定義している*16．これは，さらに証

*16 太田 [2]105 頁．

拠が追加されたとしても結論が左右されない可能性の程度と考えることができる．そうすると，今，$\boldsymbol{e}$ という証拠のもとで，心証度 $p(u^X|\boldsymbol{e})$ が証明度 p^* を超えているとすると，尽証度は，追加の証拠 $\boldsymbol{e}'$ を取り調べることで結論が変動しない，すなわち，心証度 $p(u^X|\boldsymbol{e},\boldsymbol{e}')$ が証明度 p^* を超えている（下回っている）確率として定義できる．

この定義からも明らかなように，尽証度は証明度の値に依存する．これに対して，解明度は，あくまでも心証度の分布の形にしか依存しない．したがって，心証度のばらつきとしての解明度が同じであっても，証明度が異なれば尽証度も異なり得る．また，証明度が 1/2 でなければ，心証度と証明度が近い場合には，解明度が同じでも尽証度は小さくなるが，心証度と証明度が離れている場合には解明度が同じでも尽証度は大きくなる[*17]．

第 5.3.2 節でみたように，解明度は，内心の状態を結論の判断基準とは独立して表現でき，確率論から一般的な性質が得られる一方，裁判で審理を尽くしたと言いうるかを判断するには，結論が左右されるかを考える必要があるから，この意味では，解明度と尽証度を区別して，両概念を考えることには意味がある．

5.5　解像度

要証事実に関連して内心の状態を示す概念としては，要証事実の具体性の程度というものも考えられる．例えば，覚せい剤の自己使用につい

[*17] 例えば，u^X を被告人が犯人である確率とすると，$p(u^X)$ が小さい（犯人である可能性が低い）場合には，解明度が小さければ，尽証度は高くなる（新たな証拠調べで犯であると認定される可能性は低い）が，$p(u^X)$ が大きく，証明度のボーダー近くにある場合には，解明度が小さくても，新たな証拠調べによって結論が左右される可能性は高く，尽証度は低くなる．

て，いつ，どこで，どのような方法で身体に摂取したのかという事実である．これらの事実は，要証事実を含む一連の事実経過を想像する上では有用な事実であり，その意味では，要証事実に関連する事実が判明している程度を解像度とでも呼ぶのが適切かもしれない．しかし，覚せい剤の自己使用については，尿から覚せい剤成分が検出されれば，特段の事情のない限り，その故意が推認される．この意味では，いつ，どこで，どのような方法で身体に摂取されたのかについてより詳しく特定されたとしても，心証度が変化するとは限らず，その意味で事実認定に直接は関係しない．もっとも，例えば，刑事裁判では，検察官は，訴因の明示にあたって，日時，場所，方法などをできる限り特定することを求められている（刑事訴訟法 256 条）から，このような概念が事実認定において全く不要であるとも言い切れない．要証事実に関連する事実の具体性の程度としての解像度が事実認定の確率的推論において，どのような位置づけられ，どのような意味を有するかについてはなお検討が必要であるように思われる．

6　証明度と解明義務

6.1　証明度の必要性

第 3 節でみたように，ベイズの収束定理に従えば，一定の証拠価値のある証拠が無限に存在すれば，真実についての心証度は必ず 1 に収束するから，心証度の基準を定める必要はない．しかしながら，現実には，証拠が十分に存在しなかったり，存在しても集まらないことがあり得る．このような不確実性が残る状況下で何らかの判断をするためには，その判断の基準が必要となる．それが証明度である．

証明度の問題は，(1)（原則的な）証明度の値をどう定めるかという問題と，(2) 事案によって証明度の変動を認めるのかという判断構造の問題の 2 つがある．前者は，刑事訴訟においては，「合理的な疑いを超える程度」の証明とはどの程度の証明を指すのかという問題であり，民事訴訟においては，証明の程度は，「高度の蓋然性」が必要か「優越的な蓋然性」で足りるかという問題である．後者は，刑事訴訟においては，例えば，「合理的な疑いを超える程度」の証明とは，単なる確率の問題ではないという批判*18を巡る議論であり，民事訴訟においては，証明責任の軽減法理の問題などである．これらの根底には，いずれも，法廷に実際に顕出された証拠を所与として，そのもとで判断すれば足りるのではなく，法廷に顕出されるべき，あるいは，考慮されるべき証拠ないし事実があり，それを踏まえた上で証明の成否が判断されるべきだから，単純に心証度の大小で証明の成否を判断する確率的推論では，それが適切に考慮されないという問題意識がある．

6.2　厚生経済学的アプローチ

これまでの事実認定に確率的推論を持ち込もうとする考え方は，(1) と (2) のいずれについても，厚生経済学的なアプローチに依拠していた*19．例えば，Kaplan[27] は，事実認定を社会的意思決定の一つであり，その目的を社会的期待便益の最大化であると考えると，事実認定にあたっては，正しい判断をしたときの社会的便益や誤った判断をしたと

*18 例えば，Cohen[19] は，刑事訴訟では，被告人は有罪となるのであれば，法廷に顕出された証拠のもとでの「条件付確率」として有罪かを認定されるべきではなく，関連する事実をすべて考慮した上で「無条件」確率として，有罪かを判断されるべきであるとする．

*19 Kaplan[27]，太田 [2].

きの社会的費用によって導かれる一定の値（証明度）を超えるか否かによって事実の存否を認定すべきであるというものである．具体的には，事実認定に際しての社会的便益が，判断が正しかったか否かのみに依存し，判断の際の確信の度合いによらないとし，$u^{\cdot}$ という認定が正しかったときの社会的便益を $B(\cdot)$ ，間違っていたときの社会的損失を $L(\cdot)$，証拠 $\boldsymbol{e}$ のもとで $u^{\cdot}$ が真である確率を $p(u^{\cdot}|\boldsymbol{e})$ とすると，$u^{\cdot}$ を認定をした場合の社会的便益 $W_{\cdot}$ の期待値 $E[W_{\cdot}]$ は，

$$E[W_{\cdot}] = p(u^{\cdot}|e)B_{\cdot} - (1 - p(u^{\cdot}|e))L_{\cdot} \tag{27}$$

となる．事実認定の問題は，社会的効用の最大化という観点からは，$E[W_X]$ と $E[W_Y]$ のうち大きい方を選択することであるということができるから，u^X を認定すべきであるのは，$E[W_Y] > E[W_Y]$ である場合，すなわち，

$$p(u^X|e) > \frac{B_y + L_x}{B_x + L_y + B_y + L_x} \tag{28}$$

である場合であり， u^Y を認定すべきであるのは，(28) 式の不等号が逆の場合ということになるから，(28) 式の右辺が証明度ということになる．

しかし，この考え方に従うと，事案によって判決が当事者に与える影響を検討し，その帰結に応じて証明度を変えるべきであるという結論が導かれることになる．

確かに，刑事裁判と民事裁判とは，条文上も，その目的や構造においても，大きく異なる点がある*20から，刑事裁判と民事裁判で証明度を

*20 例えば，刑事訴訟法は，「事案の真相を明らか」にすることをその目的の一つとしていることを謳っている（刑事訴訟法 1 条）のに対し，民事訴訟法では，このような規定はなく，「公正かつ迅速」に行われることが求められている（民事訴訟法 2 条）．

変えることは正当化されよう[21]．しかし，刑事裁判や民事裁判それぞれの訴訟において，事案に応じて証明度を変えることについては，基本的には法的な基礎付けがない．

また，例えば，原告が訴訟の目的物を被告よりも必要としていれば，原告が証明すべき要証事実の証明度を下げるべきだということになれば，その結論が出されたときの結果が証明度の決定を通じて審理結果に影響することになるから，当事者はこれらについても主張立証をし，裁判所はこれらについても判断する必要が出てくる．そうすると，訴訟において問題となる事実が無制限に拡大することになりかねず，当事者の負担が増し，また，訴訟の遅延を招きかねない．

当事者間の利益衡量や社会厚生の問題は，法の要件を定める場面で一度考慮されている．その上で，法は，ある規範が適用されるかについて，要件事実を定めることで考慮の対象とする事実を絞っている．これによって，訴訟において問題となる事実が限定されることになる．しかし，証明度の決定において，再度社会厚生を考慮することになると，重複となる上，考慮すべき対象が無制限に広がることになり，判断過程が複雑になる．裁判実務をみても，原告の請求が認められると被告は甚大な損害を被ることになる一方，請求が棄却されても原告の損失はわずかだから，証明度を著しく高くすべきであるというような考え方が受け入れられているとはいいがたく，証明度に対する厚生経済学的アプローチは，現実にも支持を得られていないし，理論的にも妥当であるとは思われない．

[21] 民事裁判について，さらに行政事件や家事事件，民事事件に分けることもあり得よう．

6.3 対称性と証明度

証明度をどう定めるか，あるいは，事実認定の基本構造をどう定めるかは，自然科学などが依拠する合理的な思考様式とは何かという観点から導かれるべきである．Wigner[35] が指摘しているように「自然の法則と思うものから不変性を導くのではなく，逆に，不変性によって自然法則を導き，その有効性を試そうとする」というのが現代の自然科学における「ごく自然な」考えであり，これを事実認定論にも応用するならば，証明度についての不変性によって事実認定の理論を導き，その有効性を試すという考え方が自然に導かれよう．つまり，証明度は不変（一定）であるとして，事実認定の理論を構築し，その上で，その有効性を検証するのである．

自然科学においては，理論の有効性を検証するのは実験結果である．法学のような規範的な理論においては，実験を行うことはできなが，やはり，具体的な事例から理論的に導かれる結論が，規範的に妥当かという観点から理論の有効性を試すことができる．そして，自然科学においても，実験結果が理論から得られる理論値と整合しない場合においても，特定の命題が直ちに否定されるわけではなく，その命題のほか，当該理論を構成する前提命題が全体として否定されるにすぎない*22のと同様に，規範理論においても，ある命題を前提とした理論から導かられる結論が，検証対象となる結果と一致しない場合にも，当然にその命題が否定されるわけではなく，他の前提を変えることで理論と検証対象が整合すれば，命題は依然として支持される．

事実認定論について言えば，従来，裁判上の証明とは，単なる客観的

*22 Kuhn[30].

な事実の確定ではないとか，証明度は一定ではなく，事案によって証明度を増減させる必要があるとされていた要因について，右辺の心証度を動かしたり，あるいは，確率的推論そのものを捨てるのではなく，左辺の心証度の決定のあり方を変えることで問題が解消されるかを検討することになる．前者の例としては，違法収集証拠排除法則や合理的意思解釈などの問題であり，後者の例としては，証拠が偏在する場合や，原因を特定する場合に証明度を軽減させるという議論があり得る．前者については，木本 [6] で，また，後者のうち原因の特定については，浜上 [12] に基本的な考え方が示されているほか，木本 [7] で論じられているので，以下では証拠の偏在などにより証拠が十分に集まらない場合について論じる．

6.4　証拠の欠如と解明義務

第 3.2 節で見たように，提出された証拠が多ければ多いほどより正しい結論が得られやすくなる．したがって，当事者双方が，証拠に等しいアクセスを有しており，合理的に行動することが期待できるならば，つまり，証拠に対する完全情報と完全合理性の仮定のもとでは，勝訴すべき当事者は，その時点での証拠のみでは敗訴しそうであれば，（訴訟を続けることのコストが十分小さければ）勝訴するために必要な証拠を追加するから，勝訴すべき当事者が必ず勝訴することになる[*23]．

しかし，現実には，当事者双方が証拠へに等しくアクセスを有しているわけではない．その時点での証拠のみでは敗訴する危険を負ってい

[*23] 民事訴訟では，敗訴すべき当事者の側の方で，そもそも無駄に争うことなく自白（民事訴訟法 179 条）や擬制自白（民事訴訟法 159 条）によって負けを受け入れることになるであろう．

る側が必要な証拠へのアクセスを有していない場合には，あるべき結論を導くための証拠が法廷に顕出されず，事案が十分解明されず，誤った判断（証拠が十分顕出されていればなされたであろう判断と異なる判断）がなされることが起こりうる．このような証拠が偏在する場合においてどのようにして事案を解明するかというのは，事案解明義務の問題として議論されてきた[*24]．事案解明義務の問題は，当事者が事案の解明に協力する法的義務を負うか（負うとするとその根拠は何か）という問題と，当事者がその義務に違反した場合にどうするかという 2 つに分けられる。そして，例えば，松本 [13] は，証明責任を負わない当事者についても，事案解明義務を負わせ，これに違反した場合には，証明度の引き下げがなされるべき場合があるとしている．文書や検証物の提出命令（民事訴訟法 223 条，232 条）の制度やこれらの違反や当事者尋問における不出頭等に対する真実擬制（208 条，224 条，232 条）の規定からすると，民事訴訟法が当事者に事案解明に協力する義務を一定程度化しており，その違反に対して，当該訴訟において不利益に働きうる効果を認めていることは明らかである．そして，木本 [6] で論じたように，証明妨害について証明度を変えるのではなく，特定の反事実を擬制することで当事者に不利益な効果をもたらすことができるように，事案解明義務違反一般についても，同様の議論が可能である．このように民事訴訟においては，両当事者が対等な立場にあることを前提に，双方に事案解明義務を一定程度認め，その義務違反を真実擬制と同様の形で不利益を課していると理解できる．

これに対して，刑事裁判においては，両当事者は決して対等な関係にはない．捜査段階においては，当事者の一方である検察官（捜査機関）

[*24] 事案解明義務については，春日 [5]，松本 [13]，畑 [10]，右田 [14] 参照.

に対してのみ捜索や差押などの強力な証拠収集方法を認める一方，被告人の独自の証拠収集手段は限られており，検察官の保有する証拠へのアクセスが一定程度確保されているに過ぎない（刑事訴訟法 316 条の 14 以下）．このため，刑事裁判においては，民事裁判における自白のような制度はなく，被告人が罪を認める場合でも，検察官は証拠により証明をしなければならないし，立証責任は常に検察官にあり，正当防衛や責任能力などを争う場合でも，被告人側に争点形成義務はあるものの，それらの事由が存在しないことの立証責任を検察官が負うとされている．

民事訴訟における事案解明義務との対比からしても，このような立証責任の意味は，検察官が一定の心証度を超えて証明をしなければならないというのみならず，事案を解明する義務を負っており，その違反がある場合には，解明がなされなかった事実を検察官に不利に考慮する趣旨を含んでいると考えるべきである．

例えば，証拠の資料を廃棄することで鑑定ができなくなるなどして事案の解明ができなくなった場合についても，検察官がその不利益を負担すべきである[*25]．

6.5 考察

証明度の軽減などの問題を (1) 式の左辺の問題として捉えると，事実認定のモデルが単純になるというだけでなく，事実認定の基本原理とも整合的である．そもそも，事実認定の問題の前提には，特定の法的な効果がもたらされるためにどのような事実が揃う必要があるかが確定さ

[*25] 証拠を捜査機関が廃棄してしまったことを検察官に不利に斟酌したと思われるものとして，大阪高判平 29・3・2（判例 (2)）がある．また，民事訴訟において，事案解明義務を果たさなかったことを不利に斟酌したと思われるものとして，東京高判昭和 54・10・18（判例 (1)）がある．

れている必要がある．そして，その事実の確定こそが法解釈であり，法が体現しようとする価値が問われている場面である．事実認定の手続において，真実の探求が妨害された場面においても，特定の法的効果を得るために必要な事実が変わると考える方が，事実認定の基準である証明度が変わると考えるよりも自然なことであるとも考えられる．

7　結語

事実認定の問題は，価値の問題と不可分であり，物理学や統計学，情報科学などの自然科学の理論をそのまま用いればよいというものではない．しかしながら，事実認定が価値の問題を不可避的に含むから，事実認定論については，自然科学で使われている確率論を用いることができず，別の理論体系が必要であるとか，そもそも事実認定論は複雑で自然科学のような美しい理論体系が構築できないということにはならない．信念の度合いについて論理一貫した理論を構築しようとすれば，必然的に確率論を用いることになるという議論は，事実認定の問題においても妥当する．確率論には，事前確率の定め方などなお検討すべき問題は残っているが，事前確率を一意に定められなかったとしても，収束定理の存在によって事実上の支障は大きくない．また，事実認定の判断要素が要証事実の確率のみではないとする批判についても，内心の状態を表すのには，解明度を考えれば十分であり，解明の問題も解明義務違反に対する不利益を証明妨害や違法収集証拠排除法則などと同様，何に対する心証度をどのように定めるかという問題として理解することで，確率論の基本的な枠組みを維持したまま妥当な結論を導くことができる．このように，事実認定論においても，確率論を用いることで物理学や情報科学などの自然科学との滑らかな接合を図りつつ，事実認定において

生じる価値の問題を心証度の定め方の問題として理解することで，経済学における均衡理論などに比肩しうる単純で美しい理論の構築が可能である.

引用判例

(1) 東京高判昭 54・10・18 判時 942 号 17 頁.
(2) 大阪高判平 29・3・2 裁判所ウェブサイト.

参考文献

[1] 有本卓（1980）『確率・情報・エントロピー』，森北出版.
[2] 太田勝造（1982）『裁判における証明論の基礎－事実認定と証明責任のベイズ論的再構成』，弘文堂.
[3] 太田勝造（1985）「『訴訟カ裁判ヲ為スニ熟ストキ』について」，法学教室（58 号），69-75.
[4] 太田勝造（1986）「民事訴訟法と確率・情報理論：証明度・解明度とベイズ決定方式・相互情報量」，判例タイムズ 1986（598 号），203-220.
[5] 春日偉知郎（1991）『民事証拠法研究－証拠の収集・提出と証明責任』，有斐閣.
[6] 木本茂樹（2015）「事実認定において真実の発見を目的としない諸法理のベイズ論的モデル」，URL=http://www.jlea.jp/2015zy_zr/ZR15-13.pdf.
[7] 木本茂樹（2016）「グラフィカルモデルの活用による裁判実務における 確率的推論の実践可能性」，URL= http://www.jlea.jp/2016zy

_zr/ZR16-07.pdf.

[8] 倉田卓次（1969）「交通事故訴訟における事実の証明度」鈴木忠一・三ケ月章（監修）『実務民事訴訟講座（３） 交通事故訴訟』，日本評論社，101-137.

[9] 三木浩一（2013）『民事訴訟における手続運営の理論』，有斐閣.

[10] 畑瑞穂（2002）「模索的証明・事案解明義務論」，鈴木正洋先生古希『民事訴訟法の史的展開』，有斐閣，607.

[11] 長谷部恭男（1986)「訴訟上の事実認定と確率理論ー太田勝造（名古屋大学）助教授の批判に答えてー』，判例タイムズ 616，15-17.

[12] 浜上則雄（1976)「製造物責任における証明問題（一一・完）判例タイムズ（335)，14-34.

[13] 松本博之（2015）「民事訴訟における事案の解明」，日本加除出版社.

[14] 右田堯雄（1985）「証明責任を負わない当事者の事案解明義務」，実務民事法，日本評論社，21.

[15] Berger, J.O. and Bernardo, J.M. (1989). ”Estimating a product of means: Bayesian analysis with reference priors,” 84 Journal of American Statistic Association, pp.200-207.

[16] Bishop, C.M. (2007). ”Pattern Recognition and Machine Learning,” Springer. 元田浩ほか（監訳）（2012）『パターン認識と機械学習 上』，丸善出版.

[17] Chang, T. and Eaves, D. (1990). ”Reference Priors for the Orbit in a Group Model,” The Annals of Statistics, vol.18, No.4, pp.1595-1614.

[18] Cheeseman, P. (1985). ”In Defense of Probability,” Proceedings of the Ninth International Joint Conference on Artificial

Intelligence, Los Angeles, pp.1002-1009.

[19] Cohen, L.J., (1986). "The Probable and the Provable," Clarendon Library of Logic and Philosophy.

[20] Cox, R.T. (1961). Algebra of Probable Inference, Baltimore, the Johns Hopkins Press.

[21] Finkelstein, M.O. and Fairley, W. (1970). "A Bayesian Approach to Identification Evidence," 83 Harvard Law Review, pp.489-517.

[22] Gillies, D. (2000). "Philosophical Theories of Probability," London:Routledge. （中山智香子（訳）（2004）『確率の哲学理論』，日本経済評論社.）

[23] Hawthorne, J.(2013). "Inductive Logic", in Edward N. Zalta ed., The Stanford Encyclopedia of Philosophy, Winter 2013 Edition,
URL = http://plato.stanford.edu/archives/win2013/entries/logic-inductive/.

[24] Jaynes, E.T. (2003). Probability, The Logic of Science, Cambridge, Cambridge University Press.

[25] Jeffreys. H. (1998). The Theory of Probability, 3rd ed., Oxford, Oxford University Press.

[26] Kass, R.E. and Wasserman, L. (1996). "The Selection of Prior Distributions by Formal Rules," Journal of the American Statistical Association, Vol.91, No.435, pp.1343-1370.

[27] Kaplan, J. (1968). "Decision Theory and the Factfinding Process," 20 Stanford Law Review, pp.1065-1092.

[28] Keynes, J.M. (1921). A Treatise on Probability, London,

Macmillan.（佐藤隆三（訳） (2010)『確率論』東洋経済新報社.）

[29] Kolmogorov, A.N. （1933). Grundbegriffe der Wahrscheinlichkeitrechnung, Ergebnisse Der Mathematik.（坂本實（訳））（2010）『確率論の基礎概念』，ちくま学芸文庫）

[30] Kuhn, T.S. (1962). The Structure of Scientific Revolutions, Chicago, University of Chicago Press. （中山茂（訳）(1971)『科学革命の構造』，みすず書房.）

[31] Kyburg, H. E. (1987). ”Higher Order Probabilities,” Proc. of AAAI Workshop on Uncertainty in AI, Seatle, WA, pp.30-38.

[32] Pearl, J. (1987). ”Do we need higher-order probabilities and, if so, what do they mean?,” Proceedings, AAAI Workshop on Uncertainty in AI, Seattle, July 1987, pp.170-179.

[33] Pearl, J. (1988). ”Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference,” revised second printing, Morgan Kaufmann Publishers, Inc.

[34] Tribe, L. (1971). “Trial by Mathematics: Precision and Ritual in the Legal Process,” 84 Harvard Law Review, pp.1329-1393.

[35] Wigner, E.P. (1967). ”Symmetries and Reflections,” Bloomington, Indiana University Press. （岩崎洋一他（訳）(1974)『自然法則と不変性』ダイヤモンド社.）

[36] Zadeh, L. and Kacprzyk, J. (1992), ”Fuzzy Logic for the Management of Uncertainty,” Wiley-Interscience.