あいまいさの認知における合理性

繁 榊 算 男

RATIONALITY IN PROBABILITY JUDGMENT

Kazuo SHIGEMASU

The purpose of this paper is to discuss what rationality in probability judgment is. We argued that rationality should be discussed by axioms which rational persons should obey. Coherent rules can be derived from the axioms and they can be used to check whether each judgment is coherent or not. It seems to us that only Bayesian approach has had an adequate axiomatic development, although the same kind of axiomization for other fuzzy measures seem to be possible.

Next, we examined ordinary and intuitive probability judgments, especially in terms of the results obtained from three evaluation problems, which are called "Alarm set and Theft", "Two Urns", and "Three Prisoners" Problems. Our conclusion is that, even when the subjects' responses appear to be very different from the numerical value calculated by Bayes theorem and the like, it can be explained why the subjects modified coherent probability judgment. We believe that by being educated suitably, they can be more coherent evaluator of probabilities.

1. あいまいさと合理性

合理性というとき我々がまず思い浮かべるのは論理的首尾一貫性である。たとえば、 A ならば B である、 B ならば C であるという 2 つの命題が成立するとき、 A ならば C であるという命題が真であることが論理的に要求される。ところが、人間の思考は生来このような合理性を持っているのではなく、学習によってこのような合理性を獲得していくことは、発達心理学が既に周知のものとしてきたとおりである。

論理学と心理学の関係について、ピアジェ(1967)は、「論理の方が思考の鏡であって、その逆ではない」、また、さらに「(1) 論理学は理性の公理学であり、(2) 知能心理学はこの論理的公理学に対応する経験科学である。」と言っている。すなわち、思考のメカニズムを説明するために常に論理的な法則の反映を読みとり、説明を全面的にそれに依拠するのは間違いであり、本質的に両者は別ものである。ただし、思考の最高の発達段階において、論理学と思考のメカニズムが「相互に

照応」することも一種の必然である。

ところが一方、あいまいさの心理学的とりあつかいに関しては、このようなことはあまり議論されていないようである。本論文で主張しようとするのは、あいまいさにおける認知判断にも心理学と論理学の関係と類似の関係が存在するということである。論理学と思考心理学との関係、もっと歴史的にさかのぼれば、演繹幾何学と物理幾何学との間の関係と類似の関係を見出すことがあいまいさの心理学の出発点になろう。あいまいさに関する科学としてあいまいさのとりあつかいに関する公理的展開としての規範モデルと、我々の日常的決定行動におけるあいまいさのとりあつかいを描写する記述モデルという 2 つの種類のモデルが存在するが、あいまいさにおける判断は規範モデルを目指して発達するであろう。本論文では、まず、あいまいさの認知における合理性とは何かを論じ、次に我々の日常的判断においては、どのような点においてこの合理性から逸脱するのか、また、そのような合理性を獲得するためにどのような教育が可能かを問題にする。

* 東京工業大学総合理工学研究科 (Tokyo Institute of Technology, Graduate School of Science and Engineering)

2. 規範モデル

ボレル (1967) の著書に、モリエールの喜劇「町人貴族」の登場人物、ムッシュ・ジョルダンが、「散文のことは何も知らずに散文を話すように、人間は確率を学習しなくても確率的な見方をしている」と言うくだりが引用されている。また戸田 (1982) は、「不確実事態での適応の決定の問題は、少なくとも環境のある程度の認知模型を作り得るほどの動物ならばすべてかなり能率的に処理していかなくてはならないものである。そしてその処理のしかたとしては、どうやら不確実性を確率でおきかえる以外に手はなさそうである。」と述べている。このように、不確実な事態における行動の背後に確率評価に伴う心理学的実体 (Psychological entity) が存在すると仮定することは自然であろう。問題は、この心理学的実体がどのようなものであるかである。これは、量的なもの (戸田 (1982) では、原主観確率と呼ばれている) であるか、質的なものであるかすらわからないが、一応、心理的確率 (Psychological probability) と呼んでおこう。(おそらくは、心理学的実体とは「きっと」「多分」「めったに」・・・の言語と対応した質的なものと思われる。しかし適当な変換機能を補助してやれば量的にもとりあつかうことができよう。)

記述モデルの対象であるところの心理的確率に対し、確率とは「かくあるべし (should)」の状態を記述するのが規範モデルである。規範モデルは、公理と公理から導かれる主観確率、及び主観確率が従うべき確率法則からなる。ここでいうところの公理とは、心理的測定理論における意味での公理 (あるいは前提条件) であり、数学的体系における任意の出発点としての条件という意味での公理 (たとえば、確率測度におけるコルモゴロフの公理系) とは異なる意味を持つので注意を要する。すなわち、ここでは、経験世界においてどのような条件を満たせば、抽象世界 (実数空間) においてどのような測度を得られるかを問題とする。

いま、議論を簡単にするために (といっても、本質的な意味ではこれで十分であるが)、問題とする事象の全体を、 $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ とする。 Θ は、標本空間、母集団、frame of discernment (後述する Shafer の用語である。 $\theta_1, \theta_2, \dots, \theta_n$ は外側世界の見方を表すものであり、その意味ではこの呼び方は適切である。) 等と呼ばれる。このとき、主観確率は、任意の事象 $A \subset \Theta$ に対して、① $0 \leq P(A) \leq 1$ ② $A \cap B = \phi$ のと

き、 $P(A \cup B) = P(A) + P(B)$ を満たす。このような測度を得るための前提条件を体系化する試みは数多くある。ここでは、ベイズ統計学についての標準的な教科書のひとつ (DeGroot, 1970) が示す公理系のうちから、 σ 集合体において確率測度を定義するための数学的要請 (現実世界において成立しているかどうかの検討は無理である) をのぞいて認知的に問題となる公理のみをとりあげて適当に書き直すと次のようになる。

(1) 全ての不確かさは相互に比較できる。

(2) 不確かさに関する判断は推移率を満たす。

(3) $[0, 1]$ の任意のある実数をとることと他の任意の実数をとることが同等におこりやすいような標準実験が存在する。

DeGroot の公理系は、理解しやすく、かつ、直接、生起の可能性や真実さの程度の比較から確率を扱っているので、賭けや効用の概念に言及することなく、主観確率を定義できる (Leonard, 1980)。上記の3つの公理を、(1) 比較可能性、(2) 推移律、(3) 標準実験の存在と呼んでおく。Phillips (1974) はベイズ統計学の教科書で、このような標準実験の例として、0 から 1 までの目盛がついたルーレット状のものにおいて、ほとんど摩擦のない針の指す点を例にだしている。また、Ramsey (1964) は、このような標準実験に対応するものとして、ethically neutral proposition を提案する。ethically neutral proposition とは文字通り訳せば、どんな考え方からも自由で中立的な命題となる。その命題が真であっても偽であっても世界の価値が変わらないような命題を意味する。このような中立的な命題に対する信念の度合 (主観的確率) は、次の状況において、確率 $1/2$ となる。すなわち、命題 A の直偽が明確でない場合に、2つの結果 α と β があるとき、命題 A が真のとき結果 α が生じ、命題 A が偽のとき結果 β が生じる“賭”と、逆に、命題 A が真のとき結果 β が生じ、命題 A が偽のとき結果 α が生じる“賭”とが同等な価値を持つとする。このことが全ての結果の対に対して成立するとき、命題 A が真である主観確率は $1/2$ である。ethically neutral proposition を基準として、各結果の効用を定めることができ、効用を所与とすると、 $1/2$ 以外の任意の主観確率 (この主観確率は条件付確率である) が定められる。

先にあげた3つの公理は、主観的確率が相互に矛盾のない体系であることを保証するものである故に、ベイズ的な整合性の基準であると言える。整合性に対応する通常の英単語は consistency であるが、consis-

tency は統計的推定の理論ですでに“一致性”の意味で使われているので、coherence という単語が用いられることがある。ここでも、ベイズ的な整合性のための基準が特別な意味を持つことを強調するために、上記の3条件をそのままコヒアランス基準と呼ぶことにする（コヒーレンスよりもコヒアランスのほうが若干英語の発音に近い）。

さて、コヒアランス基準の描く合理的人間としての確率評価者は次のようなことができるはずである。

① 評価者は全ての命題や現象に関し、自らの感情や要求に左右されずに独立に真実さの判断や起りやすさの判断ができる。

② 評価者は、全ての可能性を考慮に入れ、仮説をつくり、かつ各仮説の下における世界像（モデル）を構成し、以後、変更することはない。このモデルは、当該の現象に関しては、その一部始終を描写している。また、評価者は過去のデータを完全に記憶している。

③ 評価者の確率判断は推移律を満たす（コヒアランス基準を狭い意味でとらえるときは推移律のみを指すことがある）。

④ 各時点において、瞬間的に尤度やベイズの定理等の複雑な計算をこなす高速演算能力(?)を持っている。

以上の4点をもう少し詳しく解説し、現実の確率評価者の実際の対応のしかたを述べる。

① Ramsey のいう ethically neutral の概念に対応している。実際の評価者は、自分の要求に影響され、いわゆる“希望的観測”をしがちである。また、自分にとって重要な人生問題に関係する命題とサイコロをふって次に1の目がでる等のことを同列に比較することは困難である。

② ベイズ統計学の文脈において、客観性とはデータの累積的積み重ねによって、尤度が事前分布を“飲みこみ (swamp)”，事前分布がどのようなものであれ、事後分布は一定の分布へ収束することを指す。しかし、これが達成されるには、仮説は全てあらかじめ考えられており、途中で新しい仮説を導入しないこと、また、モデル分布（尤度）の骨格はデータの累積の過程において不変であること（すなわち、分布形は変化しないこと）が必要である。実際には、評価者は思いつきの下に仮説や尤度を変更しがちである。

③ 推移律を保つことは実際には非常に難しい。時間軸上の特定の時点において推移律を保つことは、まだしも可能であるにしても、時間軸上の異なる時点間の比較においても推移律を保つことは困難である。ベ

イズの合理的人間像は、ある種の一樣性 (uniformity) を仮定している。すなわち、将来も現在も、新しいデータの出現以外は過去と同じであるという仮定である。だからこそ、時間的に離れていても推移律を満たすと言えるのだが、これは現実の人間が時間軸上に変化する (dynamic) 存在であるから実際にはあてはまらない。(時間軸上の一樣性の仮定のもとにコヒアランスを考える立場を global coherence, 特定の時間軸上におけるコヒアランスに満足する立場を, local coherence として区別すべきだと思われる)。

④ 現実の現象、すなわちデータのふるまいをモデル分布（尤度）に表現しようとするとき、もし、細大もらさず現象を表現しようとするならば、その形は複雑にすぎるであろう。ベイズの定理を適用するとき、あまりに複雑な尤度（このモデル構成自体が実際には不可能であるが）をいれて計算することは時間がかかりすぎる。実際の統計的運用においては、尤度は、標準的な分布の中から選ばれることが多く、しかも、データ間に独立性を仮定するとか、計算が容易になるように尤度を簡潔することが多い。

上記のように実際には人間はベイズ的に振舞うことはできない。しかし、ベイズのコヒアランスを満たされないこと、即、非合理であるとは言えないかもしれない。逆にベイズ的な公理は、超人的な能力を要求するために、実際問題に、ベイズのアプローチを適用するとむしろ非合理的な結果を生むこともあり得よう。このことを考慮したうえで、しかし、確率評価や決定行動の合理性を問うには、ベイズの体系の出発点にさかのぼって吟味する必要がある。漠然と個々の決定にはそれぞれの合理性があると言っても意味がない（「盗人にも3分の理」というとおり、どんな決定にもなんらかの正当化は可能である）。合理性というからには、ルール体系があり、そのルール体系の背後には、いくつかの前提条件（公理、仮定、信念といったもの）があると仮定し、それがどのようなものであるかをあらかじめすることが大切である。本論文が最初に述べたように、規範モデルはそのままの形で記述モデルにはならないが、人間が最高の合理性に到達しようとするとき、人間のあいまい判断はそれに近似することが期待される。ただし、規範モデルとして、ベイズ理論（主観的期待効用最大化の理論）が唯一のものであるかについては、色々な意見があろう。次にそれを考察する。

3. ベイズ的アプローチ以外の正当化による合理性

あいまいさを扱う数学的言語としては、確率が代表的であるが、最近では、確率測度を一般化したいくつかのあいまい測度 (Fuzzy Measure; Sugeno Measure とも言われる) が提案されている。このような測度はベイズのコヒェランスを満たすであろうか。また、満たさないとすればどのような根拠づけが可能であろうか。ベイズ的ではないある種の合理性が存在するであろうか。たとえば、Dempster-Shafer 理論において、Belief Measure と、Plausibility Measure というあいまい測度を構成するものとして基本確率 (basic probability) というものがある (Shafer, 1976)。繁樹 (1988) は、この基本確率の認知論的正当化を次のように試みている。主観確率の3つの公理のうち、もっとも論争の余地のあるものは、(1)の比較可能性であろう。「知識が不足している」、「比較する次元が異なる」、「決定を先に延期したい」、「概念が明確でない」、等の理由によって、全ての事象が常に比較可能であるとは限らないでしょう。このとき全ての事象の比較可能性のかわりに、あいまいさの比較は次の3つの場合のどれかにあてはまるという仮定をおくことができる。すなわち、(1) 全てが互に比較できる。(2) 適当に基本事象を併合すれば、相互に比較可能である。(3) 相互に比較することはできない。

(1) の条件の下では、通常的主観確率が定義できる。(2) の条件の下では、基本事象の適当な組合せによって、新しい frame を構成すれば、やはり、確率を定義することができる。(3) の条件の下では、 $P(\theta)=1$ と定義できる。これらの条件付確率と、適当な組合せによって比較可能な frame を構成できるという信念とを結びつけて、基本確率が構成することを提案している。

基本確率を、このように理解するとき、少なくとも概念上、基本確率の前提条件が現実世界において成立しているかどうかを検証でき、一種の正当化がなされたと考えられる。いずれにしても、基本確率のようなファジィ測度の認知論的検討は未だほとんどなされていないと思われる。故に、ここでは、規範モデルとして、ベイズの理論を念頭において論を進める。

4. 日常的確率評価の規範モデルからの逸脱

先に述べたように日常生活において確率を評価するとき、我々は、ベイズの定理を使って整合的に評価するわけではない。多くの場合、ヒューリスティックルー

ル (heuristic rule) を用いて問題を解決する。これらのルールは実際たいいていの場合うまく機能するが、確率評価に一定の歪みをもたらすこともある。これらの歪みは、① 典型性 (representativeness) ② 入手可能性 (availability) ③ 固着性 (anchoring) の3つに大別して論じられる (Kahneman, Slovic and Tversky, 1982)。既によく説明されていることではあるが、この3つを簡単に説明する。

① 典型性: A が B に属する確率、 A が B の結果生じる確率、 A が原因となって B を生じる確率等の評価を A が B をどれほど代表しているか、 A と B がどの程度似ているかによって決めることを指す。このとき、確率評価において事前確率や標本の数は無視される傾向がある。特に、典型性を重視し、我々の確率評価がベイズの定理が示すほどにはデータによって改訂されていない現象を、確率評価の保守性 (conservatism) という。 (Edwards, 1968)。

② 入手可能性: 確率や頻度の評価が、心理的記憶ファイルから検索されやすいかどうかにかんして依存することを指す。

③ 固着性: 確率評価をある特定の値から出発して最終的な答えを見出そうとすると、確率評価が最初特定の値に引きずられることを言う。

5. 3つの例題によるベイズ的合理性の検討

確率評価の歪みは、先にあげた ① 典型性 ② 可能性 ③ 固着性だけでは説明できない場合もある。次の3つの問題をととして、直観的確率評価における合理性とは何かをさらに検討する。

問題1 防犯装置の問題

「事象 A を、本年中に大村さんが防犯装置をつける事象、また、事象 B を、大村さんが本年中に盗難事故に合う事象とする A, B を各々の否定とすると、

(a) $P(A|B)$ と $P(A|\bar{B})$ ではどちらが高いか。

(b) $P(B|A)$ と $P(B|\bar{A})$ ではどちらが高いか。」

大学生 91 名にこの確率評価実験をしてもらった結果を表 1 に示す (1988 年 8 月実験)。

上記の問題は、因果関係を推論する際の確率評価に関するものである。データ D が事象 X の原因と想定されている場合 D を原因データ (Causal Datum) と呼ぶ。逆に、データ D が、事象 X を原因として生起するものであるとき、 D を診断データ (Diagnostic Datum)

表 1. 防犯装置の問題

(例: * (a) において $P(A|B)$, (b) において $P(B|A)$ を高いとした者の人数)

(a)	(b)	
$P(A B)$	$P(B A)$	4*
$P(A B)$	$P(B \bar{A})$	34
$P(A \bar{B})$	$P(B A)$	29
$P(A \bar{B})$	$P(B \bar{A})$	16
不明・無答	不明・無答	8
計		91 (人)

という (Tversky and Kahneman, 1982). この例題では A (防犯装置) と B (盗難事故) は, 原因データにも診断データにも解釈できるが, このような場合, 我々はデータ (即ち, 所与とする事象) を原因データとみなして用いる傾向があるとされている. (a), (b) においていずれも条件付ける事象を原因データとみなすならば, 矛盾する確率評価をもたらす. すなわち, 表 1 によれば, $P(A|B)$ を $P(A|\bar{B})$ より大きいと評価し, かつ $P(B|\bar{A})$ を $P(B|A)$ より大きいと評価する者が一番多い (34 名) が, 確率法則, すなわち, 規範モデルから導かれるルールによれば, $P(A|B) > P(A|\bar{B})$ と判断するのであれば, $P(B|A) > P(B|\bar{A})$ でなければならない. 何故ならば, 次のような変形をおこなってみればわかる. すなわち

$$\begin{aligned} P(A|B) > P(A|\bar{B}) &\rightarrow P(A|B) > P(A) \\ \rightarrow P(A|B)P(B) &= P(AB) > P(A)P(B) \\ \rightarrow P(AB)/P(A) &> P(B) \rightarrow P(B|A) > P(B) \\ \rightarrow P(B|A) &> P(B|\bar{A}) \end{aligned}$$

実験の結果は, (a), (b) ともに診断データとみなす傾向が意外に多かったが, この場合も確率法則に従わない結果を生じる. すなわち, $P(A|\bar{B})$ を $P(A|B)$ より大きいとみなすならば, 確率法則によれば $P(B|\bar{A})$ は, $P(B|A)$ より大でなければならない.

条件付確率を問題とすると, 日常的確率判断では, 過去のデータから将来を予測するという形をとることが多い. 本来の条件付確率の定義から言えばこのことは誤りであるが, 広く, 人間の適応ということから考えればこのような態度はあながち不適當とは言えないかもしれない. しかし, 上記のような例で示される非整合性はやはり賭け等をする場合には問題にされるべきであろう. 本例は, 確率評価が往々にして中立的に

はなされないことを示すものであるが, 以上のような説明によって, 歪みを持たらず傾向の原因を知り, また, 条件付確率の定義を学べば, 歪みを直す (debiasing) は可能であろう.

問題 2: 2 つの壺の問題

「いまここに 2 つの壺があるとする.

一つの壺には, 赤い玉が 50 個, 青い玉が 50 個, 計 100 個入っていることがわかっている. この壺を K (known の略と考える) と呼ぶ.

もう一つの壺には, 赤い玉と青い玉が計 100 個入っていることはわかっているが, その割合はわからない. この壺を U (Unknown の略) と呼ぶ.

(1) 次の 2 つの賭け (a) と (b) のうちどちらかに参加しなければならないとすると, あなたはどちらに参加しますか.

(a) K の壺から赤い玉がでれば 1000 円貰えるが, 青い玉がでれば 1,000 円払わなければならない.

(b) U の壺から赤い玉がでれば 1,000 円貰えるが, 青い玉が出れば 1,000 円払わなければならない.

(2) 次も同じような問題です. 次の 2 つの賭け (a) と (b) のうちどちらかに参加しなければならないとすると, あなたはどちらに参加しますか.

(a) (略, (1) (a) と同じ.)

(b) U の壺から赤い玉がでれば 1,010 円 (すなわち 1,000 円 + 10 円) 貰えるが, 青い玉がでれば 1,000 円払わなければならない.

(3) (略, (1) と同じ. ただし, (a), (b) とも, 賭け金が 1,000 円から, 10,000 円に増える.)

Tversky (1974) の実験 (オリジナルは, Elsber, D (1961) とされている) では赤い玉をひいたときに, 100 ポンドの賞金がでるという設定である. この実験の結果, (a) も (b) も赤玉をひく確率を等しく $1/2$ と評価するにもかかわらず, (a) を好む者が多いという結果を報告している (1974 年の論文ではこれ以上の詳しいことは書いていない). 問題を上記のように少し修正して, 同じ大学生 91 名に実験した結果を表 2 と図 1 に示す (1988 年 8 月実験). (1), (2), (3) ともに (a) を好む者が多い. ただし, (a) を好む者の割合は (3) において最も多く, (2) では, (b) を好む者が他と比べれば相対的に多い. このような結果に対して, Lindley

表2. 壺の問題結果
((a)あるいは(b)を好む者の人数)

	男		女		全 体	
	(a)	(b)	(a)	(b)	(a)	(b)
(1)	32	8	39	12	71	20
%	80.0	20.0	76.5	23.5	78.0	22.0
(2)	30	10	34	17	64	27
%	75.0	25.0	66.7	33.3	70.3	29.7
(3)	35	5	41	10	76	15
%	87.5	12.5	80.4	19.6	83.5	16.5

図1. 2つの壺の問題: 賭aを好む者の割合

(Shafer (1982) にたいするコメント) は、ベイズ的合理性に従うならば、(1) と (3) では、(a) と (b) は同等であり、(2) では、(b) を好むべきだという(であろう)。理由は、「効用」という概念を持ちださなくても、期待金額を比較すればすぐわかる。(2) の問題において、(a) の賭の期待金額は、1000 円、(b) の賭の期待金額は 1005 円となり、(b) のほうが有利である。このことを少し詳しく見てみよう。つぼの中の赤い玉の占

める割合を θ としよう。この θ は、外界の性質であり、主観確率ではない。

この θ に関する事前の知識を、事前確率密度 $p(\theta)$ によって表明することができる。(a) の場合、 $p(\theta)$ は、 $\theta=1/2$ における一点分布である。(b) の場合、 $\theta=1/2$ をモードとする対称な分布である (一様分布を含む)。 θ を所与として、次に赤玉をひく事象 R の確率は、 $P(R|\theta)=\theta$ である。故に (a) の場合、 $p(\theta)$ の期待値は $1/2$ なので

$$P(R)=\int_0^1 \theta p(\theta) d\theta=1/2$$

となる。すなわち、ベイズ的合理性を持つ者は、(1) と (3) においては賭 (a) と (b) は同等であり、(2) においては、(b) を選ぶべきであるということになる。しかし、実際には、表2に示されるように、そうはならない。この結果は次のように説明されよう。まず、被験者は無知の状態を畏れたと思われる。賭ならば、相手のほうがより多くの知識を得る可能性を避けたいためである。第2の理由として、被験者は、このような賭を何回も行なうとすると無知の状態 (b) のほうが、赤玉 (R) や青玉 (B) のでる回数のばらつきが大きくなり、ミニマックス戦略の考え方からすれば、(a) のほうがより好ましいと判断することがあげられる。このことは (1) よりも (3) において、即ち危険が大きいときに (a) を好む者が多いことにも表れている。

この2番目の理由を吟味してみる。いま、 n 回このような賭に参加するとして、 θ を所与として x 回 R がでる確率は、

$$P(x|\theta)=\binom{n}{x}\theta^x(1-\theta)^{n-x}$$

である。故に、 $\theta=1/2$ のとき、

$$P(x)=\binom{n}{x}\left(\frac{1}{2}\right)^n$$

となる。また、 $p(\theta)$ を対称なベータ分布とすると、

$$\begin{aligned} P(x) &= \int_0^1 \binom{n}{x} \theta^x (1-\theta)^{n-x} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} d\theta \\ &= \binom{n}{x} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(x)\Gamma(n-x+\alpha)}{\Gamma(n+\alpha+\beta)} \end{aligned}$$

となる。 $\theta=1/2$ と比較して θ の値が特定できない時のほうが分散の大きい分布となる。たとえば、 $n=10$ の

図2. 2つの分布の比較

場合、(a) $\theta=1/2$ の場合の2項分布 (binomial distribution), (b) $\alpha=5$ のベータ分布による混合分布 (mixed distribution) の2つの分布を図2に示す。故にある意味では、表2の結果は正当化される。すなわち、問題では1回限りの賭を想定しているが、10回このような賭に参加するとすれば、(b)の賭のほうが大勝ちするかもしれないかわりに大負けするかもしれない。危険回避的 (risk averse) な者は当然 (a) を選ぶことになる。問題2においては、ベイズ的記述が案外と人間の直観的合理性を説明するのに役立った例であり、ベイズの確率評価に矛盾するものではない。また、赤池 (1984) は、情報量の立場から直観的判断を弁護している。

問題3: 3囚人問題

「3人の囚人が幽閉されているとしよう。3人の囚人の名前は、アラン (Alan), バーナード (Bernard), チャールズ (Charles) とする。アランは、翌日3人のうち2人が処刑され、1人が釈放されることを知っているが、誰が釈放されるかについては全くわからない。アランが釈放される事象 A , バーナードが釈放される事象を B , チャールズが釈放される事象を C とする。

A, B, C の可能性に差がないので、 $P(A)=P(B)=P(C)=1/3$ としよう。このような状況において、アランが看守に対し、3人のうち2人が処刑されるのだから、バーナードとチャールズのうち少なくとも1人は処刑されるのは確実である。バーナードとチャールズのうち処刑される者の名前を1人だけ教えてくれても、アランの釈放については全く情報を与えないはずだから、その名前を教えてほしいと言ったところ、看守はアランの言い分を納得し、バーナードは処刑されると答えた。アランはこれを聞いて、釈放される可能性のあるのは、自分の他はチャールズのみに変わったので、 $P(A)=P(C)=1/2$, すなわち、自分が釈放される可能性が増えたと喜んだとする。

直感的には正しいように見えるが、この確率評価は根拠があるのだろうか。」

これは、Lindley (1985) が紹介している興味深い例である。ベイズの定理では、看守の言葉を D とするとき、尤度 $P(D|A)=P(D|A)=1/2$ とするとき、 $P(A|D)=1/3$, $P(C|D)=2/3$ となり、 A に関する限り、データ D は、影響を与えない。しかし、実は、このことは、 B と C が、釈放の可能性に関して交換可能 (exchangeable) という特殊な条件においてのみ成立することが指摘されている (市川, 1988)。市川 (1988) は、 $P(A)=1/4$, $P(B)=1/4$, $P(C)=1/2$ と事前確率を変えてみた。ベイズの定理に従うならば、 $P(A|D)=1/5$ となる。釈放される可能性のある者の数 (競争者) は減ったにもかかわらず釈放される可能性は減じている。これは多くの人にとって納得し難い結果である。

市川 (1988) は、多くの場合は妥当だが時々誤りに導く確率評価ルールが存在することがこの困難さの原因であるとしている。(通常のヒューリスティックは、いわば、手続き型であるのに反し、このルールは、いわば宣言型である故に、彼らは、これを主観的定理の名で呼んでいる。) もし、ベイズの定理を正しいとするならば、確率評価者にベイズの定理の結果をどのように納得させるかが問題となる。このために、たとえば、視点の変換 (佐伯, 1986), ルーレット表現 (市川, 1988), 問題の表現の書き換え (守, 1988), 具体的実験演示 (繁樹, 1987) 等の試みがなされている。

筆者は、この問題に関して、現在のところ、次のように考えている。

このような人工的な例において、もし世界の全てが上記のような設定によって表現されるのだと仮定すれ

ば、やはりベイズの定理を信用すべきであろう。何故ならば、事態が明確である以上、合理的人間が決定問題に対処するときと同様に通常の人間でもふるまえると思うからである。そして、この結果を納得させるには、我々の定性的確率推論を数値的に変換させる補助装置として、ルーレット表現が最も有望なのではないかと思われる。しかし、さらに重要なことは、ベイズの定理を含むベイズ的帰納推論体系を理解することであり、このことが、しぶしぶながらも、この反直感的な結果を納得させる原動力ではないかと思われる。このときベイズ的体系から導かれることを基にして、“主観的定理”をより妥当な“主観的定理”の体系にかえることが必要であろう。例えば、“選択肢の数が減れば確率は増える”という主観的定理は、ベイズ的推論体系から導かれませんが、“起りうることを全てに喜ぶことはできない”という定理は、ベイズの体系から導くことができる。これは、事前確率評価と比べて、起りうることを全て(標本空間)の場合に、事後評価が好ましい方向へ変化(たとえば、全ての確率評価が大きくなる)ということはある程度を示す。もしそうならば、全確率の定義によって事前確率自体が修正されなければならないことになるからである。しかし、直観的確率判断にも同情すべき(?)点がある。事前確率の評価における不確実さ、尤度の設定における不確実さを考えるとき、 $P(A|D)=1/5$ という数値を見せられても納得しないのでまた考え直し、その結果が、事前確率と似ているということもおおいにありそうである。このことは次の例(Solmaon, 1988)と共通点をもつ。誰かが別室でコイン投げをしている。そのコインが両面とも表である(Two-headed,つまりインチキのコイン)であることを疑う理由があったとして、その確率を $1/10$ としよう。今、別室からの報告で4回とも表がでたとすると、ベイズの定理によれば、Two-headedである事後確率は 0.64 になる。ここで評価者は、これは大きすぎる値であると考えて、何とかこの確率を下げようとする。この場合尤度については議論の余地が少ないので事前確率に着目し、実はTwo-headedである事前確率は $1/100$ であったと考え直す。この場合の事後確率は 0.14 であり、評価者は一応満足するであろう。このコイン投げの例の場合、評価者は、時間軸上を通じて大域的な coherence は保っていないが、局所的には辛うじて coherence を保っていることができる。ただし、繁樹(1987)の実験的演示の場合には同情すべき余地は少なくなる。こ

では、コップ A の中にサイコロがある確率を $P(A)$ 、同じく B の中にある確率を $P(B)$ 、コップ C の中にある確率を $P(C)$ とする。コップ C の口の面積は A, B の 2 倍であるという設定で、1 m ぐらい離れたところから、サイコロがどれかのコップに入るまで投げつづけることを想像させる。たいていの被験者は、 $P(A)=1/4$, $P(B)=1/4$, $P(C)=1/2$ と評価する。次に、スクリーンの上に、B, C を隠し、B が空であることを示す。そこで $P(A|D)$ を評価させるというものである。この場合、D を得る前後に状況の変化はなく、また尤度 $P(D|A)$ を $2/3$ 未満であればどんな値であれ、 $P(A|D) < P(A)$ となる。このことを考えると、この問題は、問題 1 や問題 2 よりも反直観的であり、結果を納得させることも難しいであろう。しかし、この困難さは間違った(定性的)確率法則を持っているためであり、ベイズの体系から正しく導かれる定性的ルール体系を、きちんと教育しているならばこの結果も理解し得るはずである。なお、情報理論の立場から直観的判断の一種の正当化を試みている論文もある(井原, 1988)。

以上の 3 つの例題、特に問題 3 は、納得させることが困難なものではあっても、直観的確率評価が、ベイズ的なものと敵対的な関係にあるのではなく、説明をきちんとすれば納得が得られるものと考えられる。

6. 確率評価の“評価”と改善

主観的確率が信用できるものであるためには、評価値が整合的でなければならない。それでは、整合的でありさえすれば、個々の確率評価が良いとか悪いとかいう議論は無意味であろうか。

もし、主観的確率評価の“真の値を”知ることができれば、この評価値の良さの評価は簡単である。“真の値”を公認された理論から導かれる場合や、過去の統計データの累積からまず正確な値をつかんでいる場合がこれに当たる。しかし、このような場合の主観的確率の評価自体があまり意味のないもので、確率評価をする前に、理論やデータを調べたほうが早い。

次の方法は、評価した事象の主観的確率と実際に生じた割合とが近いものであればその確率評価は良いものであるし、両者が遠いものであればその確率評価は悪いとする方法である。たとえば、天気の数値予報において、30% の降雨確率予報をした日の実際の天候のデータを何年間かまとめて、そのうち雨が降った日の割合を求め、30% に近いかどうかを問うものであ

る。Lindley, Tversky and Brown (1979) では、前者を pragmatic な方法、後者を semantic な方法 (calibration) と呼んでいる。

これらに対して、コヒアランスの基準を持っているかどうかによる確率評価の“評価”は、syntactic な方法と呼ばれている。コヒアランスを満たさない確率評価者は長い目で見て損をすると言える。コヒアランスを満たさないことの代価として、確実に負ける賭け (Dutch book) を設定できるということがある。現実の世界でも「ジャックと豆の木」の中で主人公が交換をくり返して結局つまらないものを得るというような事態はしばしば起り得る。故に、この意味では syntatic な評価は、確率評価が良いか悪いかの評価よりも基本的であり重要であると言える。

次に、それでは、どのようにすれば、コヒアランスを回復できるかと考える。実際に、心理的確率評価が教育や訓練によってベイズ理論に近づくといういくつかのデータが得られている。たとえば、心理的確率評価が学習によってより正確になるというデータ (Wallsten, 1983) とか、医学部の学生と経験を積んだ臨床医の診断を比較すると臨床医のほうが適切に事前確率を用いていること (Wallsten, 1981) が報告されている。また、コンピュータと対話することによって、このような歪みを少なくし、整合的な確率評価を可能にしようとするシステムも考えられる。

具体的にどうすれば確率評価が改善されるかについて次の論文が参考になる。Shafer and Tversky (1982) は、確率評価を一種のメンタル試行 (mental experiment) と考える。解を見出すためのメンタル試行のデザインには優劣があると彼らは考える。たとえば、全米での一日の卵の生産量はどのくらいに答えるために、全てのにわたりの数を考え、その数に、一日あたりの産卵量をかけるというデザインが考えられる。それに対して、アメリカの全人口を考え、それぞれの人が平均何個卵を食べるかを想像し、この2つの数を乗ずるというデザインもある。あきらかに後者のほうがすぐれている。確率評価において、整合性になるべく満たすことを可能にするデザインとして次の2つがあげられている。

Shafer and Tversky は、ベイズ的な確率評価として、

(1) 関連事象による細分化デザイン (原著では Total Evidence Design とされている) これは事象 A の確率を評価する際に、他の関連する事象 $B_1, B_2, \dots, B_m$ を

導入し、

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \dots + P(A|B_m)P(B_m)$$

として評価するものである。ただし、このとき $B_i \cap B_j = \emptyset, \cup_i B_i = \mathcal{Q}$ ($\mathcal{Q}$ は全事象) である (すなわち、 B_i は分割 (partition) である)。たとえば、次の大統領に A があるという確率を直接評価するよりも、 B_1 : 経済諸条件が良くなる、 B_2 : 現状維持、 B_3 : 悪化するという3つの事象に分けて評価するほうがより正確な評価が可能になる。関連する分割は他にもあるので、いくつかの評価値がでてくる。これらの値を相互に比較し、もし違いがあればそれが何故かを検討し、もう一度評価し直すことによってコヒアランスを獲得することができる。

(2) 条件づけデザイン (conditioning design)

これはベイズの定理を利用する確率の評価方法に他ならない。すなわち、関連する事象 D があるとき、評価したい不確定事象 A について、 $P(D|A)$ の評価が、 $P(A|D)$ を直接評価するよりも容易であり、かつ、 A よりも先に D が既知となるような場合に、

$$P(A|D) = D(D|A)D(A)/P(D)$$

によって、 A の確率を評価することを指す。 $P(A)$ の評価よりも、 D の持つ情報によってより程度の高い評価が可能となる。

7. ま と め

本論文で主張したかったことは、だいたい次のようにまとめられる。

(1) 規範モデルは、現実世界で検証されるべき公理的体系をもつべきである。

(2) 実際の決定行動が規範モデルに従って行なわれないことは、人間の認識能力の不十分さが原因である。

(3) しかし、規範モデルと一致しない直観的判断も、なぜそうなるかについては説明可能である。

(4) また、いろいろな手法をつかうことによって、決定行動を規範モデルに近づけること、すなわち、コヒアランスを回復することは必要であり、可能でもある。

参 考 文 献

- 赤池弘次 (1984) 確率の解釈における困難さについて。統計数理研究所彙報, 32, 117-128.
- エミール・ボレル (1967) 確率と生活。文庫クセジュ, 白水社。

- de Groot, M.H. (1970) Optimal statistical decisions. New York, McGraw Hill.
- Edwards, W. (1968) Conservatism in human information processing, in Kleinmuntz, B. (ed.) Formal Representation of Human Judgement. John Wiley, 17-52.
- 市川伸一 (1988) 3 囚人問題の解決と理解の過程をめぐって. 認知科学の発展 I., 講談社.
- 井原二郎 (1988) 3 囚人問題の理論的探究. 認知科学の発展 I., 講談社.
- Kahneman, D., Slovic, P. and Tversky, A. (eds.) (1982) Judgement under uncertainty: Heuristics and biases. Cambridge University Press.
- Leonard, T. (1979) The roles of inductive modelling and coherence, in "Bayesian statistics." (J. M. Bernardo et al., eds.), University Press, 537-555.
- Lindley, D.V. (1985) Making decisions, 2nd Ed. John Wiley and Sons, Ltd.
- Lindley, D. V., Tversky, A. and Brown, R.V. (1979) On the reconciliation of probability assessments. J.R.S.S., 142, 146-180.
- 守 一雄. (1988) 「3 囚人問題」は何故難しいか. 信州大学教育学部紀要, 62, 45-50.
- Phillips, L.D. (1974) Bayesian statistics for social scientists. T.Y. Crowell.
- ピアジェ (Piaget J.) (1967) (オリジナル, 1947) 波多野完治, 滝沢武久訳 知能の心理学. みすず書房.
- Ramsey, F.P. (1964) (オリジナル, 1926) Truth and probability, in "Studies in subjective probability." (H.E. Kyburg, Jr. and H.E. Smokeler, eds.) John Wiley and Sons, Inc.
- 佐伯 胖 (1986) 認知科学の方法. 東京大学出版会.
- Salmon, W.C. (1988) Dynamic rationality: Propensity, probability, and credence, in "Probability and causality." (J.H. Fetzer ed.). D. Reidel, 3-40.
- Shafer, G. (1976) A mathematical theory of evidence. Princeton Univ. Press.
- Shafer, G. (1982) Lindley's paradox. J.A.S.A., 77, 325-334.
- Shafer, G. and Tversky, A. (1985) Languages and designs for probability Judgement. Cognitive Science, 1985, 9, 309-339.
- 繁榎算男 (1987) 3 囚人問題の具体化について. 日本心理学会第 51 回大会発表論文集.
- 繁榎算男 (1988) あいまい測度の認知論的基礎. 日本心理学会第 52 回大会発表論文集.
- 戸田正直 (1982) 主観確率の測定法について. 行動計量学, 9, 1-11.
- Tversky, A. (1974) Assessing uncertainty. J.R.S.S. (Series B), 36, 148-159.
- Tversky, A. and Kahneman, D. (1982) Causal schemas in judgements under uncertainty, in "Judgement under uncertainty: Heuristics and biases." (Kahneman, D. et al. eds.), Cambridge University Press, 117-128.
- Wallsten, T.S. (1981) Physician and medical student bias in evaluating diagnostic information, Medical Decision Making, 1, 145-164.
- Wallsten, T.S. (1983) The theoretical status of judgement heuristics, in "Decision making under uncertainty." (R.W. Scholz ed.), North-Holland.