さいは投げられたのか
――確率論の応用の正当化と科学的経験の超越論性――

出口康夫

一　はじめに

確率論（theory of probability）は、今日、科学の広い分野で用いられている数学の理論である。それは、統計力学や量子力学において中心的な役割を果たしているし、また統計学の基礎となることで、数理科学全般に対して、実験・観察・調査の方法論を提供している。ここでいう数理科学とは、物理学・化学・生物学などの自然科学や、経済学・社会学・心理学などの社会科学、さらには歴史学・古文書学などの人文科学をも含んで、領域横断的に成立している学問分野である。数理科学では、現実の対象の何らかの数量的な性質を実際に測定することで、それについての仮説を経験的に検証するという作業が重要な位置を占めている。このような検証作業においては、確率論的な統計学における統計的仮説検定（tests of significance）が標準的な方法として採用されているのである。

また確率論が浸透しているのは、数理科学の諸分野にとどまらない。それは、好むと好まざるとにかかわらず、いろいろな形で、われわれの日常生活にも入りこんでいる。保険制度、工業製品の品質管理、気象予報、世論調査、治療の選択、各種の統計にもとづく経済・社会・環境政策の策定などは、そのような浸透のほんの一例であろう。

このように、ほとんど森羅万象と呼びたくなるような、広範囲の対象への確率論の応用を前にして、本論ではまず、

「そもそも、確率論を現実の対象へ適用するとは、いかなることか」、「ある対象が、確率論によって記述・分析されるにふさわしいものであることは、いかにして正当化されうるのか」を問うことから始めたい。

一口に確率と言っても、その数学的定式化や解釈を巡って、これまでさまざまな立場が唱えられてきた。数学的定式化に関しては、確率を、集合に実数を与える関数である測度（measure）の一種として定義する測度論的確率論が、科学において広く定着し、事実上の標準理論となっている。他方、確率解釈としては、確率を、「命題間の（演繹関係の拡張であるとされる）確証関係の度合い」と解釈する論理説、「相対頻度の極限値」と見なす頻度説、さらには「個人の信念の度合い」とする主観説などが代表的なものであった。これら三つのうち、測度論的確率の解釈として提案されているのは、頻度説と主観説であるが、ことに頻度説は、科学のさまざまな分野や日常生活の多様な側面への浸透という点では主流をなしている。そこで本論では、頻度解釈を施された測度論的確率論に焦点を絞って検討を加えていくことにしたい。

ところで、頻度的・測度論的確率論の現実への適用は、抽象的な数学理論一般の経験的現実への適用の平凡な一例と見なされることが、従来多かった（cf. Doob, 1976）。それに対して本論は、特に確率論の適用の正当化に関して、これまで見過ごされてきた、いくつかの特異性を明らかにすることを目指したい。

その上で、本論では最後に「それらの特異性が、いかなる哲学的含意を持つのか」を問うてみたい。その際、注目されるのは、統計的仮説検定における確率論の応用である。先に触れたように、確率論にもとづく仮説検定は、数理科学における仮説の、経験に照らした証拠立て、ないしは端的に言って、「科学における経験」の方法論的枠組みをなしている。科学における経験は、仮説の検証作業に対して確率論を適用することで、はじめて可能となるといえるのである。とすると、その適用の正当化にまつわる特異性は、科学的経験を可能とするこの理論的枠組みのあり方に対して、少なからぬ含意を持つと思われる。本論では、それらのうち、「科学的経験の超越論性」をはじめとする、いくつかの含意の

摘出を試みたい。

二　測度論的公理系

測度とは何か

以下では簡単に、コルモゴロフ（Kolmogorov, 1933）によって作られた確率論の測度論的公理体系を紹介しておこう。まず測度とは何か。測度とは、線分の長さ・平面図形の面積・立体の体積といった幾何学量に由来する概念である。一九世紀における幾何学や微積分学の発展にともない、幾何学量は、「集合に与えられた実数値」として一般化・抽象化され、その実数値を与える関数として「測度」が導入された。測度概念の一般化は、その後さらに進み、最終的に、ボレルやルベーグによって、「シグマ集合Sに対して0以上の実数値を与える可算加法的関数」として定義されるようになった。ここでいうシグマ集合とは、なんらの位相も設定されていない抽象空間Ωの、ある一定の条件を満たした部分集合として導入される集合である。ちなみに、この抽象空間は標本空間（sample space）、その要素は標本点（sample point）と、それぞれ呼ばれる。さらに、一定の標本空間Ωに含まれるすべてのシグマ集合を要素とする集合族は、シグマ集合族Σと呼ばれる。このシグマ集合族は、Ω自体をも含み、補集合を与えたり・集合の和をとったり・集合の交わりをとるといった、集合にたいする基本的な操作の無限回の適用に関して閉じているように定義されている。例えば、可算無限個の集合 $S_i(i\in N)$ の各々が、あるシグマ集合族に属している場合、それらの無限和集合 $\bigcup_{i=1}^{\infty} S_i$ もまた、そのシグマ集合族に属しているのである。シグマ集合族の例としては、標本空間Ωの可能なすべての部分集合からなる集合であるベキ集合や、Ωと空集合ϕだけからなる集合$\{\phi, \Omega\}$などがある。

さて、このようなシグマ集合に対して、いくつかの条件を満たしつつ、0から最大で正の無限大∞までの正の実数値を与える関数を測度μとよぶ。測度が満たすべき条件の一つに加法性（additivity）がある。いま、互いに共通の要素

を持たない二つのシグマ集合 S_1, S_2 があり、それら各々に対する一定の測度の関数値を、それぞれ $\mu(S_1)$、$\mu(S_2)$ とすると、両者の和集合 $S_1 \cup S_2$ に対する測度 $\mu(S_1 \cup S_2)$ が $\mu(S_1)+\mu(S_2)$ であることが、ここでいう加法性である。なお、有限個のシグマ集合の間に成り立つ加法性は「有限加法性」、可算無限個のシグマ集合の間に成り立つそれは「可算加法性」と呼ばれる。また測度は、この加法性以外にも、空集合には0を与えるという条件を満たさねばならない。

確率測度

コルモゴロフは、確率をΩに1を与える測度、すなわち確率測度 (probability measure) として定義した。ちなみに、標本空間Ωに有限の実数値を与える測度は有限測度と呼ばれる。測度論的確率論における確率とは、$\mu(\Omega)=1$ であるような、一種の可算加法的有限測度なのである。ところで、コルモゴロフの体系における確率測度は、つぎの三つの公理によって文脈的に定義されている。

(1) 任意のシグマ集合Sに対して、$0 \leqq P(S) \leqq 1$ が成り立つ。

(2) $P(\Omega)=1$。

(3) 可算無限個の、互いに共通の要素を持たないシグマ集合 $S_i (i \in N)$ に対して、$P(\bigcup_{i=1}^{\infty} S_i)=\sum_{i=1}^{\infty} P(S_i)$ が成り立つ。

これら三つの公理からなる公理系からは、単調連続性など、確率測度の重要な諸性質が定理として導かれる。コルモゴロフの公理体系におけるすべての定理は、それぞれの仕方で確率測度の性質を表しており、公理系はそれらの性質の集約的表現といえるのである。

上の公理系を満たしつつも、個々のシグマ集合に、さまざまな仕方で確率測度値を与えることが可能である。そしてその与え方によって、さまざまな確率測度が区別される。一定の確率測度を設定するさいには、その確率測度がその上

で定義される、標本空間Ωとシグマ集合族Σをつねに定めておかねばならない。一定の標本空間Ω・シグマ集合族Σ・確率測度Pからなる順序列$\langle\Omega, \Sigma, P\rangle$は、確率空間 (probability space) と呼ばれる数学的構造である。測度論的確率論とは、形式的に言えば、確率空間という集合論的な存在者についての研究に他ならないのである。

三　測度論的確率論の解釈

数学理論の解釈と三つの存在領域

集合論的構造についての研究とされていたコルモゴロフの測度論的確率論では、その公理体系のすべての基礎的な概念は、経験的な意味を一切剝奪され、専ら抽象的な集合論の術語で定義されていた。他方、測度論的確率論をさまざまな現象に適用する際には、それら基礎的な諸概念に対して、集合論的ではなく数学的でもない、何らかの意味を与えておかねばならない。ところで、一般の数学理論と同じく、測度論的確率論も、原理的に複数の異なった非数学的な意味付け、即ち解釈を許す。つまり、数学的確率論に対する唯一の正しい解釈などは、そもそも存在しないのである。その意味で、コルモゴロフをはじめクラメル・ドゥーブ・フェラーら、数学者の多くによって採用されている頻度説といえども、さまざまな可能な解釈のうちの一つに過ぎないといえるのである。ところで、一括りに頻度説といっても、論者の間にはある程度の立場の違いが認められる[1]。そこで、本論ではドゥーブの立場 (Doob, 1941) を代表的なものと認定し、頻度説の適用のあり方を検討する際に、つねに念頭に置くことにする。

頻度解釈の下で現実へ適用されるべく、ドゥーブが設定した確率空間は、可算無限個の点列からなる標本点 $\omega=\langle a_1, a_2, \ldots a_n, \ldots\rangle$を持つ、可算無限次元確率空間$\langle\Omega^N, \Sigma^N, P\rangle$であった。

また、数学理論を経験的現実に適用する際には、通常、現実をその理論による記述に適したように理想化・単純化した模型 (model) という存在領域が設定される。そして、その模型の上に仮構された存在者や事象が、数学理論の基礎

的な概念の非数学的な意味、即ち解釈だとされるのである。言いかえると、抽象的な数学理論は、模型上の理想化された存在者という意味を介して、現実に適用されるのである。このように、数学理論の現実への適用とは、理論・模型・現実という三つの異なる存在領域にまたがった概念操作なのである。測度論的確率論の適用も例外ではない。例えば、ドゥーブは可算無限次元確率空間を現実に適用する際に、その空間と完全に対応する模型として、可算無限回繰り返される「試行（trial）」という、現実を理想化・単純化した事象を設定したのである。

模型上の意味

それでは、確率空間〈Ω^N, Σ^N, P〉に含まれる諸概念に、無限回試行という模型において、どのような意味が与えられているかを確認していこう。まず、確率空間における最も基本的な概念である標本点には、「可算無限回繰り返された確率試行（stochastic trial）の結果がたどる一定の経過」という模型上の存在者が、その解釈として与えられる。ここでいう試行とは、その過程が人の手によって完全に制御可能で、また「われわれが、起こりうるすべての結果を、前もって網羅的に知っている」とされているような物理的出来事である。現実においてこの試行に対応するのは、サイコロふりや硬貨投げ、さらには加速器において素粒子同士を衝突させるといった実験的操作などである。ただし、われわれは、現実に行われる実験の可能な結果について、あらかじめ、ある程度は包括的な見通しを持っているが、そのすべてを網羅的に知っているわけではない。またわれわれは、実験の過程を、ある程度は制御できるが、その制御も決して完全なものではない。このように、ここでいう試行とは、現実における物理的な操作そのものではなく、それを単純化・理想化したものなのである。

一回ごとの試行の結果である物理的出来事は、気体分子の速度といった連続量であっても、サイコロの目の数といった離散的な値であっても、経験的に測定・同定可能な物理量・物理値として完全に表現されうるものでなければならな

い。またこの物理量・物理値は、通常、われわれの認識のありようとは無関係に確定している、物理的世界の客観的な性質であるとされている。そして、先に触れた、「ある試行を可算無限回繰り返して得られる一定の経過」という標本点の模型上の意味は、このような物理的出来事が、一定の順序で可算無限個並んだものに相当する。一つのサイコロを無限回投げて得られる、〈1・3・4・4…〉という出た目の一定の経過がその例である。もちろん、ここでいわれる、試行の「無限回の繰り返し」もまた、模型における現実の理想化の一つである。

　上で見た標本点の模型上の意味から、それを可算無限個、要素として持つ標本空間Ω^Nの意味を確定することができる。即ち、「無限回試行のすべての可能な経過の集合」というのがそれである。さらに、Ω^Nの部分集合であるシグマ集合には、「無限回試行の経過において繰り返し登場する部分列としての物理的な出来事（event）」という模型上の意味が与えられる。サイコロ投げでいえば、「1が出る」とか「6が三回続けて出る」といった出来事が、この部分列に当たる。いま、「i回目・$i+1$回目・$i+2$回目で6が出る可能な経過からなる集合」ϕ_iを考えてみよう（$i \in \mathbb{N}$）。つぎにϕ_iの、すべてのiに関する無限和$\bigcup_{i=1}^{\infty} \phi_i$を取ると、この無限和集合が、「6が三回続けて出る」という模型上の出来事に対応するシグマ集合となる。このように、ここでいう「出来事」とは、「無限回試行の可能な経過から構成された一定の集合」として、数学的に表現されるのである。

　さて、試行を有限回繰り返した後、いったん止め、それまでに得られた経過において、ある出来事が起こった相対頻度（relative frequency）を数えるとしよう。この相対頻度もまた、経験的に測定できる客観的な物理量である。つぎに試行を再開し、繰り返し回数を一つづつ増やすごとに、同じ出来事の相対頻度を数えていくことにする。すると最終的には、相対頻度の無限列が得られることになる。模型上で確率測度に与えられる意味とは、この相対頻度の無限列が、それに対して収束していく極限値に他ならない。

　通常、確率試行とは、一回の試行で何が起こるか、前もって予測できるような規則性が見あたらない試行であるとさ

れ、その意味で不規則（random）であると見なされる。「この予測不可能性は、われわれの知識不足に由来するのか、それとも事柄そのものの性質なのか」については意見が分かれる。が、いずれにしても、一回ごとの結果には規則性はないが、出来事の相対頻度が一定の値に収束するという意味での規則性を持つことが、試行が確率的であるとされるための要件なのである。ところで、ここでいわれる収束とは、可算無限個の点列において成り立つ数学的な性質である。従って、相対頻度の収束もまた、模型において仮構される理想化された現実にすぎない。

さらに、後で確認するように、模型上の無限回確率試行といえども、そのすべての可能な経過において、出来事の相対頻度の収束が起こることは保証されていない。そこで特に、そのような収束が起こる経過を「収束事象列」と呼ぶことにする。収束が起こらない可能な事象の経過は、確率測度に対して模型上の意味を与えない。すると、収束事象列が模型上で存在することが、何らかの意味で保証されなければ、確率測度は模型上の意味を持ち得ず、従ってまた、それを介して現実に適用されることもできないのである。

四　大数の強法則と確率測度の頻度解釈

大数の強法則

模型上の収束事象列の存在を、ある条件の下で保証しているのが、「大数の強法則（strong laws of large number）」である。以下で、この存在保証のあらましを見てみよう。

形式的に見れば、大数の強法則とは、確率変数の和の収束に関する一群の定理であるといえる。ちなみに確率変数（random variable）とは、各々の標本点に対して、一定の仕方で実数値を与える関数である。（標本点には値を与えない確率測度と異なることに注意。）いま、無限次元標本空間Ω^Nの上に、一定の順序で並んだ可算無限個の確率変数

$$X_1, X_2, \cdots, X_n, \cdots \quad （1）$$

を定義したとする。すると、任意の一つの標本点ωに対して、これら無限個の確率変数が与える関数値の無限列

$$\langle X_1(\omega), X_1(\omega), \cdots, X_1(\omega), \cdots\rangle \ (\omega \in \Omega^N) \quad （2）$$

が得られる。さらに、これらの確率変数の有限和を含む関数の無限列

$$X_1, (X_1+X_2)/2, \cdots, (1/n)\sum_{i=1}^{n} X_i, \cdots \quad （3）$$

を設定する。この有限和関数の各々も、確率変数の条件を満たす。いま、この関数の無限列（3）をΩ^N上に設定すると、任意のωに対して、こんどは

$$\langle X_1(\omega), (X_1(\omega)+X_2(\omega))/2, \cdots, (1/n)\sum_{i=1}^{n} X_i(\omega), \cdots\rangle \quad （4）$$

という関数値の無限列が得られる。大数の強法則は、最初の確率変数列（1）が、一定の条件を満たせば、ωに対する二番目の関数値列（4）が、一定の値に「確率1で」収束することを示している。そして、この場合の収束の極限値は、確率変数列（1）に属する任意の確率変数$X_i(\omega)$のルベーグ積分値、即ち期待値$E[X_i(\omega)]$であるとされる。つまり、（1）に関する一定の条件下で、

$$P(\lim_{n\to\infty}(1/n)\sum_{i=1}^{n} X_i(\omega) = E[X_i(\omega)]) = 1$$

が主張されているのである。ここでの一定の条件とは、ボレルによってこの定理が初めて証明された際には、「確率変

数列（1）が互いに独立で、かつ同一の分布に従う」というものであった。後にこの条件は、ドゥーブによって、非独立な確率変数列であるマルチンゲールにまで緩和されたが、本論では一貫して、ボレルの古典的な条件の下で話を進める。

いまΩ^Nの標本点$\omega = \langle a_1, a_2, \ldots a_3, \ldots \rangle$の各点$a_i$が、1から6までのいずれかの値しかとらないとする。そして、確率変数列（1）のi番目の確率変数$X_i(\omega)$を、ωのi番目の点a_iが取る値に応じて、つぎのような関数（単関数）として定義する。

$$X_i(\omega) = 1 \quad \text{iff} \quad a_i \in \{1, 3, 5\}$$

$$X_i(\omega) = 0 \quad \text{iff} \quad a_i \in \{2, 4, 6\}$$

すると、確率変数$(1/n)\sum_{i=1}^{n} X_i(\omega)$は、$\omega$の最初の$n$個の点列において、奇数が出現する相対頻度を表すことになる。いま、i番目の点列が奇数となるすべての標本点を含むシグマ集合をS_i、偶数となるすべての標本点を含むシグマ集合を$\neg S_i$とする。（ちなみに、両者はΩ^Nを二分割する集合である。）すると、単関数$X_i(\omega)$のルベーグ積分値、即ち期待値$E[X_i(\omega)]$は、定義上、

$$E[X_i(\omega)] = 1 \times P(S_i) + 0 \times P(\neg S_i) = P(S_i)$$

となる。つまり$X_i(\omega)$が単関数である場合、$E[X_i(\omega)]$はS_iに与えられた確率測度値$P(S_i)$に等しくなるのである。大数の強法則は、確率変数列（1）の各々が単関数であり、かつまた「互いに独立で、同一の分布に従う」という先の条件を満たす場合、$(1/n)\sum_{i=1}^{n} X_i(\omega)$が確率測度値$P(S_i)$に収束することを、「確率1で」保証しているのである。

ちなみに、ここでいう「確率1での保証」とは、関数値列（4）の$P(S_i)$への収束がその上で起こるすべての標本

点ωからなる集合、

$$\Theta=\{\omega\in\Omega^N: (1/n)\sum_{i=1}^{n}X_i(\omega) \text{ が } P(S_i) \text{ に収束する}\}$$

の確率値が1、即ち$P(\Theta)=1$であることを意味するのである。

大数の強法則の模型上の意味

前節で見た、確率空間上の数学的な事実は、以下のように模型の上に翻訳できる。まず、確率変数列（1）が満たすべき「独立性・分布の同一性」という条件に対しては、「模型上の一回ごとの試行が、互いに因果的に独立に、即ち、互いの試行結果に因果的な影響を及ぼし合うことなく、かつ同じ状況下で行われる」という物理的意味が与えられる。これを以下では、「試行の独立・同一性」と呼ぼう。また確率変数$(1/n)\sum_{i=1}^{n}X_i(\omega)$には、例えば「サイコロを無限回投げた場合、最初の$n$回の試行において奇数が出る相対頻度」という意味が与えられる。さらに、確率測度$P(S_i)$には、「任意の一回の試行で奇数が出る確率」を表す。

これらの模型上の意味を踏まえると、大数の強法則の模型上の意味はつぎのようになる。即ち、「無限回のサイコロ投げ試行が、同一・独立性を満たして行われた場合、〈奇数の出る相対頻度が、任意の一回の試行で奇数が出る確率値へと収束するような出た目の経過〉からなる集合、即ち収束事象列の集合が持つ確率値が1であること」。このことはまた、「独立・同一性の条件の下で、無限回試行の経過のうち、収束事象列でない経過からなる集合の確率値が0であること」を意味する。

一般に、Ω^Nのような可算無限集合において、確率値0が与えられる部分集合の要素は、高々有限個しかない。他方、同様の集合において、確率1の部分集合は、必ず無限集合となる。つまり大数の強法則は、試行の独立・同一性を、模

型上において収束事象列が無限個存在するための十分条件として設定しているのである。模型上において収束事象列が無限個存在していれば、少なくともその事象列に関しては、確率の頻度解釈はなりたつ。結局、大数の強法則は、確率測度の頻度解釈が（その解釈を満たす模型上の存在者が存在するという意味で）空虚でないための十分条件を、試行の独立・同一性として設定しているのである。

五　確率論適用の正当化

現実的近似態と確率試行仮説

たびたび触れてきたように、模型とは現実の単純化・理想化であり、その上で仮構された存在者は現実には存在しない。逆にいえば、現実の存在者は、模型上のそれの、何らかの意味での近似態（approximation）に過ぎないのである。例えば、現実の試行は有限回しか繰り返すことができず、その経過である物理事象列の長さも、高々有限である。場合によっては、この有限事象列において一定の部分列が出現する相対頻度が、試行の回数が増えるにつれ、ある一定の値に近づいたり、その値から離れなくなるという意味で、安定傾向を示すこともあるだろう。しかしこの安定傾向も、無限に続く漸近傾向としての収束ではないのである。

このような現実の有限事象列のあるものが、模型上の収束事象列の初めの有限部分と見なされることで、その現実的近似態であると解釈されることになる。模型上に可算無限個存在する収束事象列のうち、その最初の有限部分列が、ある特定の現実の物理事象列と一致するものは無限個ある。現実の物理事象列が長くなるにつれ、それに対する理想形としての収束事象列の候補は絞られていくが、最終的に一つに確定することはない。しかし模型上の無限事象列は、どの二つをとっても、確率試行の無限回繰り返しがたどる経過としては異なり、現実において互いに並び立つことはない。従って、現実の事象列の理想的な無限事象列となるものは、どれか一つに絞られなければならない。現実の事象列は、

不特定だが、あくまで単一の模型上の無限経過の近似態なのである。

さて以上までで、模型上の収束事象列を真ん中にして、それぞれ、有限物理事象列がその現実的近似態、標本点がその数学的表現となるという三項関係が見て取れた。測度論的確率論を頻度解釈の下で現実に適用するとは、現実・模型・確率空間という三つの異なった存在領域にまたがる、この三項関係を設定する作業に他ならないのである。そして、確率論の現実の有限事象列への適用は、「その事象列は、模型上の確率試行の試行結果である収束事象列の、現実的近似態である」という仮定のもとで行われるのである。（今後、この仮定を確率試行仮説（stochastic trial hypothesis）、ないし*ST*仮説と呼ぶ。）すると、ある事象列が、頻度的で測度論的な確率論による記述・分析の対象としてふさわしいといえるのは、それに対する確率試行仮説が真である場合であり、かつその場合に限られることになる。結局、この種の確率論の現実の対象への適用が正当かどうかは、その適用に際して前提されている確率試行仮説の正当性いかんにかかっているのである。

実用主義的正当化

しかし一方、確率論の現実への適用は、確率試行仮説が主張する「適用対象の適格性」を正当化することによってではなく、確率論を用いた理論の科学的・技術的成功によって、実用主義的（pragmatic）に正当化されるという見解が、ドゥーブ（Doob, 1976）を始め、少なからぬ論者によって表明されている。ところが以下で述べるような理由で、このような考えは成り立たないのである。

そもそも、実用主義的正当化でいわれる理論の科学的・技術的成功とは、理論による予測の成功と見なしてよいだろう。ところで、確率論を含む理論の予測は、「ある実験結果のみが起こり、その他の結果は全く起こらない」といったたぐいのものではなく、すべての実験結果に対して一定の出現頻度を与えるものとなる。そのような頻度的な予測は、

どのような結果とも矛盾することは、さしあたってない。すると、頻度的予測の成否の判定は、実験結果との単純な照合によって一致・不一致を確認するという形では、もはやなされないのである。確率論が幅広く用いられている現代科学において、第一節で触れたように、理論の予測の判定、ひいてはその判定による理論の検証に対しては、それ自体、確率論にもとづいた統計的仮説検定を用いざるを得ない所以である。

ところが次節で検討するように、仮説検定を用いた予測の成否の確認は、ある対象に対する確率論の適用を前提した上で、確率値、即ち頻度極限値や、その分布形を特定する作業に他ならない。確率論を含め、数学理論一般の予測の成功・不成功は、そもそも確率論の適用を前提した上でないと、判定できないのである。すると、もし予測の成功が何事かを正当化するとしても、それは特定の確率値なのであって、確率論の適用そのものではないことになろう。このように、確率論の現実への適用は、それを用いた理論の成功によっては正当化されない。結局、その適用の正当性の根拠は、その適用において前提されているST仮説の正当性に求めざるを得ないのである。

六　相対頻度の観察による正当化

従来の頻度説批判と二つの論点

それでは、「ST仮説は、いかにして正当化され得るのか」を、「サイコロ投げ」という試行に即して考えていこう。ちなみに、近代の数学的確率論の源流の一つは、サイコロ投げに関わる問題に対して、一六五四年にパスカルが与えた解答にあった。サイコロ投げは、確率論が成立するきっかけとなった、最初の適用対象の一つなのである。

また、近代確率論の成立時から、確率に対する解釈の一つとして与えられていたとされる「頻度解釈」[2]は、「さまざまな現実の現象において、相対頻度が安定する傾向がある」という確認がきっかけとなって成立した。この頻度安定現象の具体例として、死亡率統計とならんで、サイコロ投げや硬貨投げといった「偶然事象」がある。[3]

このような歴史的経緯から言っても、サイコロ投げなどの繰り返し試行において、相対頻度の安定傾向が実際に観察されたならば、「その試行は確率試行である」とするST仮説が「正当化される」と考えるのは、一見もっともなように思える。すると、例えば「硬貨を数千回、ないし数万回投げることによって、表の出る相対頻度を観察する」というビュフォンやカール・ピアソンの実験は、硬貨投げに関するST仮説の経験的な検証であることになろう。

しかし一方で、少なからぬ論者が、「この頻度の安定傾向の経験によっては、ある事柄が頻度的意味での確率事象であるという主張は確証も反証もされない」と指摘している（cf. Jeffreys, 1931, Fine, 1973, Wheatherford, 1982）。彼らは、実際に観察される相対頻度の安定値を、頻度説が確率概念に与える経験的意味であると、まず見なす。その上で、その安定値が確率の経験的意味として不適格であると指摘することで、頻度説に対する批判を展開するのである。彼らの論拠の要点は、「有限個の事象列における相対頻度の安定傾向の経験によっては、極限値を特定できないのみか、そもそも極限が存在するかどうかすら決定できない」というものである。この指摘は、「最初の有限個の経過は観察された有限事象列と似ているが、その後は収束せず発散したり、またたとえ収束しても異なった極限値をとるような無限系列」、さらには逆に、「最初の有限個の値では頻度の安定は観察されないが、その後収束するような（収束の速度が極端に遅い）無限系列」をも、数学的に構成できるという事情にもとづいている。このような数学的事実を踏まえて、「頻度の安定傾向の観察という手段によっては、相対頻度の極限値として解釈された確率に言及している命題を、決定的に反証も支持もできない」という主張が導かれるのである。さらにそこから、「「経験的で客観的な物理量としての確率解釈」を標榜する頻度説は、実はその主張とは裏腹に、確率を経験的に確証も反証もできない概念としてしまっている」という頻度説批判がなされるのである。

このような批判と並んで、私も本節で、「相対頻度の経験によっては、ST仮説、ひいては確率論の現実への適用は正当化されえない」と論じたい。しかし以下での私の議論は、上で紹介したこれまでの頻度説批判と軌を一にするもの

ではない。むしろそれは、従来の批判が、頻度説に対するいささか的外れな批判であることをも、明らかにするだろう。

まず議論の出発点として、相対頻度の収束の有無の判定と、その収束の極限値、即ち頻度確率値の特定とが別個の作業であることを確認しておこう。例えば、コーシーの収束の判定条件は、極限値を特定せずに、コーシー列と呼ばれる種の数列が収束することを保証している。このように、極限値を特定せずに収束の有無を論ずることは、数学的には十分に意味のあることなのである。(4)

この区別を念頭に置いて、以下ではつぎの二つの主張を論証したい。(一)相対頻度の観察という作業は、あくまで相対頻度の収束を前提した上で、極限値の特定を目指したものである。(二)相対頻度の収束さえ何らかの仕方で確保されれば、確率値としての相対頻度の極限値を特定せずとも、頻度的確率論は現実へ適用され得る。即ち、確率論の経験的適用において、頻度極限値の特定は本質的ではない。

第一の論点：仮説検定としての頻度の観察

本節では、第一の論点を、サイコロを、一、〇〇〇回投げることで、1の出る相対頻度の極限値の特定を目指す、ネイマン＝ピアソン型の統計的仮説検定に即して確認する。

そもそも仮説検定とは、一九世紀末にカール・ピアソンによって開発された統計的手法である。その後一九二〇年代に、フィッシャーが、(後で説明する)「確率化(randomisation)」という概念とともに、特定の標本の大きさ(例えば、サイコロを投げた回数)が与えられた場合の、一定の測定値が出現する頻度の分布(標本分布(sampling distribution))を導出する方法論(いわゆる厳密標本論(strict sample theory))を仮説検定に導入した。しかし、フィッシャー流の仮説検定は、一つの仮説のみを検定対象とし、検定結果も、その仮説の棄却が生じるか否かのいずれかしかなかった。この場合、検定対象の棄却が生じなかったことは、その対象が確証されたことを意味しない。フィッシャー流の仮

説検定は、経験的反証の方法ではあっても確証の方法ではなかったのである。それに対して、ネイマンとエゴン・ピアソン（カール・ピアソンの息子）は、経験的確証をも扱える方法を考案した。現在、標準となっているのは、このネイマン＝ピアソン型の仮説検定である。なお、これらの仮説検定は、いずれも古典統計学（classical statistics）の技法であり、今日では頻度的・測度論的確率論にもとづいて展開されている。

ネイマン＝ピアソン型の仮説検定は、二種類の相対立する仮説、即ち帰無仮説H_0と対立仮説H_1を立てることから始まる。ここでは帰無・対立仮説として、それぞれH_0「1の出る相対頻度の極限値は1/2である」、H_1「1の出る相対頻度の極限値は1/6である」を採用しよう。

さてここで、サイコロを可算無限回投げて得られた無限個の結果、即ち無限個の「目」からなる集合が、仮説無限母集団（hypothetical infinite population）として設定される。ちなみに、この無限母集団は、先の無限次元標本空間Ω^Nの標本点ωを構成する無限個の点列$a_1, a_2, \ldots a_n, \ldots$からなる集合である。先の$H_0$は、この無限母集団における「1の目」の存在頻度が1/2であることを、またH_1はそれが1/6であることを主張していると解釈される。

つぎに、サイコロを一、〇〇〇回投げるという操作が、この母集団から一、〇〇〇個の目を抽出する作業であると見なされる。いま抽出された一、〇〇〇個の目を元の母集団に戻し、改めて一、〇〇〇個の目を抽出し直すという操作を、これまた可算無限回繰り返したとしよう。すると、無限母集団に属する無限個の目の各々は、一定の頻度で抽出されることになる。（この抽出頻度は、無限母集団での「一定の目の存在頻度」とは別物であることに注意されたい。）

ここで「各々の目の抽出頻度は一定の値に収束し、その極限値の全体は一定の確率分布（例えば一様分布）に従う」という仮説が立てられる。この仮説は「確率標本（random sample）仮説」と呼ばれ（以後、RS仮説と略称する）、それが成り立っているとされる場合、一回の抽出で得られた一、〇〇〇個の目からなる結果は「確率標本」であると見なされる。ちなみに、この「仮説無限母集団からの確率標本の抽出」という考えも、フィッシャーによって提案されたも

のであり（Fisher, 1925）、ネイマン＝ピアソン型の検定も含め、現代の古典統計学の主要な部分である厳密標本論の根幹をなしている。今日、確率標本という概念はさまざまな分野に登場するが、その背後には、このような二重の無限集合からなる概念枠がひかえているのである。そして、その概念枠の中で、帰無仮説・対立仮説と確率標本仮説という三つの仮説が設定されるわけである。

つぎに、帰無・対立仮説の各々から、サイコロを一、〇〇〇回投げた場合、1の目が出る回数mのすべての可能な値（即ち0以上、一、〇〇〇以下）に対して、一定の頻度分布、即ち標本分布が数学的に導出される。ところで、母集団における分布についての仮説であった帰無・対立仮説のみからは、この観察値の分布、即ち標本分布を導くことができない。標本分布を導出するためには、それら以外に、母集団の要素の抽出頻度を特定するRS仮説が不可欠である。結局、帰無仮説とRS仮説、対立仮説とRS仮説の連言がとられ、それら二つの連言から、互いに異なった標本分布が導かれることになる。例えば、標本の大きさをn、帰無・対立仮説がそれぞれ主張している「1の目が出る相対頻度」をpとし、「標本抽出頻度は一様分布に従う」というRS仮説を前提すると、すべての可能なmに対して、その出現頻度$P(m)$が、典型的には、

$$P(m)=C_n^m p^m(1-p)^{n-m}$$

という二項分布として求められるのである。

仮説検定において検定の直接の対象となるのは、これら二つの標本分布である。そして、帰無仮説とRS仮説から得られた標本分布が退けられた場合、帰無仮説が反証されたと見なされ、対立仮説とRS仮説から導かれた標本分布が受け入れられた場合、対立仮説が確証されたとされる。このように、通常、仮説検定の検定対象となるのは帰無ないし対立仮説であり、RS仮説は、専ら、検定の不可欠の前提として用いられるのである。

それでは、標本分布はいかにして検証されるのか。まず、帰無仮説とRS仮説から導かれた標本分布が、一定の小さな出現頻度――例えば、0.05や0.01――を与えている、可能な観察値の一定の領域が、棄却域（critical region）Cとして設定される。ちなみに、帰無仮説の下で棄却域に与えられる出現頻度$P_{H_0}(C)$は、有意水準（significance level）と呼ばれ、その領域に入る観察値に与えられた相対頻度の総和であり、m_1とm_2をそれぞれ棄却域の最小・最大値とした場合、

$$P_{H_0}(C)=\sum_{m=m_1}^{m_2}P(m)$$

として計算される。さて、もし実際に観察された1の出る回数がこの棄却域に入れば、それに小さな出現頻度を与えた標本分布、ひいては帰無仮説が偽として棄却され、対立仮説が真として採用される。この場合重要なのは、有意水準、即ち、「もし帰無仮説（とRS仮説）が真なら、実験を無限回繰り返した場合、棄却域に入る観測結果が得られる相対頻度」が知られていることである。この頻度はまた、無限回の実験のうちで、「帰無仮説が実際は真であるにもかかわらず退けられてしまう（偽なる対立仮説が受け入れられてしまう）」という、「第一種の誤り」が起こる相対頻度でもある。他方、もし観測値が棄却域に入らなければ（棄却域の補集合$\neg C$に入れば）、今度は、帰無仮説が真であると見なされ、対立仮説が退けられる。ここでもまた、対立仮説とRS仮説から導かれた標本分布が$\neg C$に与える確率$P_{H_1}(\neg C)$が求められていなければならない。この確率は、「実際には正しい対立仮説が採用されず、偽なる帰無仮説が採用される」という「第二種の誤り」が起こる相対頻度を表している。

一般に、決定的でない実験結果にもとづいて、二つの対立する仮説の一方を採用しようとする限り、ここで挙げた二種類の誤りの可能性はつきものである。問題は、これらの誤りの確率をいかにして最小限に止めるような棄却域を選ぶかである。ところがやっかいなことに、第一種の誤りの確率を下げようとすると、帰無仮説が棄却される機会、即ち対

立仮説が採用される相対頻度が下がり、結果として第二種の誤りの確率を高めてしまうことになる。そこで、先に触れたように、第一種の誤りの確率を、一定の低い値に固定した上で、第二種の誤りの確率を最小とするような棄却域を選ぶのが次善の策であろう。ここでの例のように、関心の対象となっている物理量に特定の数値を与えている帰無・対立仮説を前提した場合、一定の第一種の誤りの確率に対して、第二種の誤りの確率を最小とするような棄却域が存在することは、ネイマン＝ピアソンの基本定理で保証されている。われわれは、この定理を用いて棄却域を設定し、実際にサイコロを一、〇〇〇回投げてみて、帰無・対立仮説のどちらを採用するかを決定すればよいのである。(5)

以上のように、仮説を反証ないし確証するさいに、二種類の誤りの確率を明示することが、統計的仮説検定の本質であり、かつ最大の長所でもある。ところで、今の場合、この誤りの確率が算出できたのは、帰無・対立の両仮説が一回の試行で1の出る確率値を特定していたからであった。さらに、確率の頻度解釈の下では、これらの主張は、「1の目が出る相対頻度が収束すること」を含意していることになる。つまり、帰無・対立仮説のいずれもが、頻度収束を含意したものでなければ、そもそも仮説検定を行うことができないのである。例えば、相対頻度の収束を主張する帰無仮説と、その発散を主張する対立仮説という組み合わせに対しては、仮説検定を行うことができないのである。仮説検定とは、収束自体の有無の決定なのではない。それはあくまで、相対頻度の収束を前提した上での、特定の極限値の間の選択に過ぎないのである。結局、有限事象列における頻度の観察にもとづく統計的仮説検定によっては、頻度の収束を経験的に正当化することはできないのである。

第二の論点：極限値の特定か収束の確保か

第二の論点に移ろう。そもそも科学における典型的な確率論の使用において、われわれが記述し予測しようとしているのは、ある特定の出来事の相対頻度の極限値ではなく、平均や分散や共分散といった、多数の事象の集団的性質に関

わる量である。例えば、統計力学において現実との一致が問題となるのは、気体分子個々の運動量の相対頻度の極限値ではなく、それらの平均の関数としての気体の熱・圧力・体積である。またサイコロ投げの場合における興味の対象は、しばしば「出現する期待値が最大の目は何か」であろう。すると、現実の出来事の記述や予測に用いられる確率模型においても、平均値や分散値は特定されねばならないが、個々の確率値、即ち頻度極限値は不特定のまま、前者と両立するものと見なしておけばよいのである。われわれが、確率論の応用において、経験的意味を持った特定の値を与えておかねばならないのは、形式的に言えば、確率測度値なのではなく、可測関数の確率測度に関するルベーグ積分値（期待値・平均）や、その関数値（分散・共分散など）なのである。この意味で、確率論の経験的適用において頻度極限値の特定は本質的な問題ではないと言える。

それを反映して、科学の現場で実際に用いられる仮説検定において検定対象となるのは、例えば、平均値に異なった値を与える帰無・対立仮説であることが多い。これらの仮説は、特定の平均値と両立可能なさまざまな頻度分布を主張する仮説からなる集合である。サイコロ投げに即していえば、例えば、「3が出る期待値が最も高い」という主張を共有する仮説の集合と、「5が出る期待値が最も高い」という主張を含意する仮説の集合との間で、検定が行われるのである。

しかし、確かに頻度極限値としての個々の確率値は未規定のまま放置できるとしても、期待値を特定する以上、相対頻度の収束は前提しておかねばならない。さもなくば、確率値は存在せず、その関数として定義される期待値もまた、存在しないことになるからである。頻度説にもとづく確率論が現実へ適用され得るためには、確率値としての頻度極限値は特定される必要はないが、相対頻度の収束は前提されねばならないのである。

このような理解を踏まえて、前節で触れた、確率論の適用において常に前提されるST仮説に関しても、そこで含意されているのは相対頻度の収束だけであって、特定の極限値ではないことを、ここで改めて確認しておこう。このこと

はまた、仮説検定の前提となる、異なった抽出頻度の分布を主張するさまざまなRS仮説においても、例外なくST仮説が含意されていることを意味する。

二つの論点を受けて

以上二つの論点を踏まえることで、確率論の現実の現象への適用が正当なものだと言い得るために決定的に重要な事柄は、相対頻度の極限値をいかに特定するかではなく、その収束をいかに確保するかにあり、従ってまた、頻度の収束を前提した上でその極限値（ないしその期待値や分散値）を判定する相対頻度の観察は、確率論の現実への適用を正当化する手段とはいえないと結論づけることができる。結局、頻度安定現象の経験は、頻度解釈を生み出すきっかけにはなったとしても、その適用を正当化する根拠とはならないのである。

また以上のことは、先に紹介した頻度説批判が的外れであることも示唆している。というのも、相対頻度の収束の極限値としての確率を特定せずとも、期待値や分散値といった指標に経験的に観測可能な数値を与えてさえおけば、頻度的確率論は経験的意味を持ちうるからである。また期待値・分散値の経験的意味を問題にしたいのであれば、それらが前提としている頻度収束の有無の確証・反証可能性を論ずるべきであり、収束の有無の判定を意図していない頻度の観察という手続きを取り上げても無意味なのである。

実用主義的正当化再論

ところで本節で概観した仮説検定は、「観測値は$\neg C$に入る」「観測値はCに入る」という言明を、それぞれ帰無・対立仮説による観測結果に関する予測と見なすことで、これら両仮説の予測の成否を判定する作業として捉え直すことができる。この場合、観察値が$\neg C$に入れば、帰無仮説の予測が成功し、対立仮説のそれは外れたとされ、またCに入れ

ば、対立仮説が予測に成功し、帰無仮説は失敗したと判定される。このような仕組みを採用することで、すべての観測結果とさしあたっては両立可能な、頻度的な主張である両仮説の予測の成否を、一定の誤りの確率を伴って、一意的に判定することが可能となるのである。

上で指摘したように、帰無・対立の両仮説とも、ある試行における相対頻度の収束を含意していた。すると、これらの仮説のどちらが予測に成功しようと、その試行に対して確率論が適用されることには変わりがないし、またどちらの予測が失敗しようと、それに対する確率論の適用が中止されることはない。仮説検定としての予測の成否の判定とは、確率論のある特定の対象への適用を当然視した上で、確率値や期待値を選択する作業に過ぎないのである。このように見てくると、「確率論を含んだ理論の予測の成功によって何かが正当化されるとしても、それは当の対象についての確率論の適用そのものではない」という先節で見た論点が、改めて確認できるだろう。

七　確率化による正当化

模型上の十分条件の実現

それでは、頻度の収束を確保する手段とは何か。コルモゴロフはつぎのように述べている。

……確率とは、実験の数が無際限に増えていくにつれて頻度がとる極限値であると見なされている。ところでこの実験が確率的である、即ち頻度がある一定の値のまわりに集まっていく傾向を持つというこの仮説は、ある条件が無限回の繰り返しにおいて、つねに完全に精確に満たされ続ける場合にのみ妥当であり続ける。

(Kolmogorov, 1963, pp. 253-4)

ここで言及されている、「頻度がある一定の値のまわりに集まっていく傾向を持つという……仮説」とは、われわれの言うST仮説に相当する。そしてコルモゴロフの言う、ST仮説の妥当性を保証する条件とは、先にわれわれが模型上

の収束事象列が存在するための十分条件として触れた「試行の独立・同一性」に他ならない。つまりST仮説は、模型上の収束事象列が存在する十分条件を、現実において実現することで正当化されると考えられているのである。先に確認したように、無限回試行という模型上の仮構存在が独立・同一性の条件を満たせば、その可能な経過のうちに収束事象列が無限個存在することは、大数の強法則によって保証されていた。この理解にもとづいて、現実の試行における独立・同一性の確保は、「その試行が結果としてもたらす有限事象列を、模型上の収束事象列の最初の有限部分であると見なすこと」を正当化すると考えられているのである。この立場から言えば、ST仮説の正当化のためには、試行の独立・同一性さえ確保できれば、有限事象列において、特定の数値のまわりでの頻度の安定現象が観察される必要もないことになる。頻度の安定現象が観察されない独立・同一な試行の有限の経過は、有限個の観察では安定現象が気づかれないくらい遅い収束速度を持った収束過程を理想形としていると解釈すればよいのである。

近似的実現の手段：確率化

言うまでもなく、話を現実の有限回の繰り返しに限っても、すべての試行において独立・同一性を「完全に精確に」満たすことは、実際上不可能である。試行の独立・同一性もまた、現実においては近似的にしか実現することができない、一つの理想に過ぎない。しかしこの正当化の戦略にとって好都合なことに、この模型上の十分条件は、実験状況の制御・乱数表などの乱数発生装置を用いてなされる「確率化」を施した標本抽出といった、人為的操作によって近似的に達成可能であると目されている。加えて、これらの人為的操作によって達成される近似の程度は、数学理論を現実へと適用する他の事例と比べても、遜色のないものであると理解されている。例えば物理実験の場合、一回の観測が他の観測に影響を与えないように、また観測が同じ条件で繰り返されるように実験装置を工夫しさえすれば、ある程度満足のいく形で、模型上の十分条件が実現されたと見なされるのである。その結果、なにも観測を繰り返して、観察値の相

対頻度の安定傾向を経験的に確かめなくとも、その観測値の系列を収束事象列の近似態と見なし、確率論を用いてそれを記述・分析することは、数学理論一般の現実への適用の平凡な一例として容認されるわけである。このように、実験や標本抽出における人為的操作によって比較的容易に達成できるという利点があるため、模型上の十分条件の近似的実現によるST仮説の正当化（以降、これを典型的な人為的操作の名を冠して「確率化による正当化」と呼ぶ）は、科学の現場で日常的に採用されており、頻度的・測度論的確率論の多方面への応用を支えている。

先に確認したように、確率化によって実現が目指される十分条件を設定しているのは、大数の強法則であった。大数の強法則は、（一）測度論的確率論の頻度解釈が、模型上において空虚でないことを保証し、かつ（二）その確率論が、頻度解釈の下で現実へ適用される際に広範に採用されている正当化の方法を提供しているという二点において、頻度的・測度論的確率論の現実への適用にさいして、重要な役割を果たしているのである。

八　確率化による正当化と科学的仮説一般の正当化の共通点

正当化の不確実性

以下では、確率化によるST仮説の正当化と、科学において提案されている、さまざまな事実に関する仮説一般の正当化との異同を明らかにしていきたい。本節では両者の共通点に着目するが、そのためにまず、「確率化による正当化は、仮説が真であることを確実に保証しないこと」から確認していこう。「確率化による正当化が確実なものではない」という事態は、以下で論ずるように、その正当化の基礎となっている大数の強法則による確率1での頻度収束の保証そのものに由来するものである。従って、それは、その法則が設定する条件の実現が近似的にしか達成できないことを度外視しても成り立つ事柄なのである。

さて、大数の法則による確率1での収束の保証とは、「独立・同一性の条件を満たした無限回試行のすべての可能な

経過のうち、相対頻度の収束が起こるものからなる集合Θの確率値が1であること」を意味していた。ところで、現実の繰り返し試行の経過である有限事象列が、模型上に理想形として持つ無限事象列とは、無限回試行の不特定だが単一の経過であった。また、模型上の単一の経過の確率空間上での数学的表現は単一の標本点ωである。そして、この単一の標本点からなる単元集合$\{\omega\}$の確率値は、例外的な場合を除きつねに0である。(6) 一般に、あるシグマ集合Sが0以外の確率値αを持つとすると、このSは確率1を持つとされた、先に言及した集合Θに必ず含まれなければならない。(そうでなく、ΘとSの交わりが空集合だとすれば($\Theta \cap S = \phi$)、確率測度の加法性より、両者の和集合の確率値が1を越えてしまう($P(\Theta \cup S) = 1 + \alpha > 1$)という矛盾が生ずる。)一方、確率値0のシグマ集合については、このようなことは言えない。すると確率値0を持つ、(例外的な場合を除く)単元集合$\{\omega\}$は、Θに含まれてもよいし、Θに含まれなくともよいのである。言いかえると、任意の単一の標本点ωは、たとえΘの確率が1であったとしても、Θに属するとは限らないのである。ところで、Θは可算無限個の標本点からなるのに対して、その補集合$\neg\Theta$は高々有限個の標本点からなる集合であった。たとえ確率変数列が独立・同一性の条件を満たしたとしても、大数の強法則は、「単一の標本点が、その標本点に対して定義された確率変数列の和の収束が起こる無限個の標本点の一つとしてΘに属するのか、それともそのような収束が起こらない有限個の例外として$\neg\Theta$に属するのか」については、何も含意しないのである。このことを模型上で言えば、たとえ独立・同一性を満たした無限回試行であっても、その単一の経過に関しては、大数の強法則を以ってしても、相対頻度の収束は保証されていないことになる。さらにこれを現実に戻せば、たとえ有限回の試行において、独立・同一性が完全に達成されたとしても、その結果生ずる事象列が、模型上の収束事象列の最初の有限部分であるかどうかは分からないことを意味するのである。言いかえると、たとえ独立・同一性の条件を完全に満たした試行を、そのまま無際限に繰り返しても、その結果得られる無限列が、頻度収束が生じない高々有限個の例外の一つでない保証は、どこにもないのである。従って、ある現実の試行に対して、大数の強法則が設定する条件の近似的実現を目

指して確率化を行ったとしても、その試行の経過において頻度が発散する可能性、即ちその試行に関するST仮説が偽である可能性は、つねに排除できないことになる。確率化によるST仮説の正当化は、その仮説をつねに真とすることができないという意味では、不確実なものに過ぎないのである。

一般的特徴としての不確実性

しかし正当化がいま述べた意味で確実でないことは、何も確率化によるST仮説の正当化に限った話ではない。そもそも一般に、「正当化される」という概念は「真である」という概念からは独立であり、ある仮説が何らかの手続きに則って真であると正当化されることは、その仮説が実際に真であることを必ずしも含意しないのである。ある仮説を正当化するためには、大雑把に言って、それが真であることの一定の根拠を与えればよいのであり、それが必然的に真であることを保証する必要はない。その意味で、確率化による正当化がST仮説を必然的に真としないからといって、それだけでは、それが正当化に当たらないとする理由にはならない。さらに統計的仮説検定において、検定結果にはつねに一定の第一種・第二種の誤りの確率が伴っていたという事実を思い起こそう。「たとえ正当化がなされたとしても、仮説が偽である可能性が残る」という、確率化による正当化が抱える問題は、科学的仮説を正当化する一般的な手段であった仮説検定においても当てはまる、普遍的で平凡な事態に過ぎないといえるのである。

三つの類似点

確率化による正当化と、仮説検定による科学的仮説一般の正当化の間に成り立つ類似性はこの点に止まらない。仮説検定において、正当化の成立にもかかわらず残る誤りの可能性が、第一種・第二種の誤りの確率として表現されることで頻度的な意味を与えられ、公認されているように、確率化による正当化においても、偽なる仮説が正当化される可能

性は、無限回試行における有限回の出現頻度を表す「確率0」という形で表現され、いわば理論的に織り込み済みの事態と見なされているのである。さらに仮説検定においては、誤った結論へと導く可能性を表現している確率値が、何らかの意味で最低であるということが、その検定の「よさ」を表す指標とされていた。他方、確率化による正当化においても、誤りの確率が0という最低値に抑えられていることは、その正当化の「よさ」を表していると見なせるのである。

結局、ST仮説の確率化による正当化と、仮説検定による科学的仮説一般の正当化の類似点は、(一)仮説の正当化が、必ずしもその仮説が真であることを保証しない、即ち、偽なる仮説が正当化される可能性が存在すること、(二)その偽なる仮説を正当化してしまう可能性が、一定の確率値として頻度的に表現され、いわば理論的に公認されていること、(三)その頻度としての確率値が、何らかの意味で最低であるということが、それら正当化の「よさ」を示していることの三つである。

九　確率化による正当化と科学的仮説一般の正当化の相違点

撤回可能な正当化と訂正可能な誤り

しかし一方で、両者は、いったんなされた正当化が撤回可能であるかどうか、そして、それぞれの意味で最低の値に抑えられた確率値としてその可能性が表現されている誤りが、将来において経験的な証拠にもとづいて訂正可能かどうかという点に関して、大きく異なっている。

まず仮説検定における誤りの改訂可能性から確認していこう。一般に、同じ実験値に対して異なる検定方法(例えば、適合度検定におけるカイ二乗検定とコルモゴロフ＝スミルノフ検定)が相反する検定結果を与えた場合、両者のうち、より低い第二種の誤りの確率値を持った結果を優先するという判断は、一つの有力な選択方針として確立されている。(7) この選択基準に従えば、同じ実験値に対して標本の大きさだけが異なる二つの検定が、相反する検定結果を与えている場

合も、両者のうち、より低い第二種の誤りの確率値を持つ結果が優先されてしかるべきであろう。ところで、他の条件がすべて同じならば、仮説検定の標本の大きさが大きいほど、その検定の結果に伴う第二種の誤りの確率は小さくなることが知られている（cf. Henkel, ibid., p. 82）。例えば、サイコロを一〇、〇〇〇回投げて行われる検定の第二種の誤りの確率は、一、〇〇〇回投げる検定のそれよりも小さいのである。するともし、同じサイコロに対する一、〇〇〇回投げ検定と一〇、〇〇〇回投げ検定で、採用される仮説が異なった場合、第二種の誤りの確率がより低い後者の結果が優先されることになる。また中心極限定理（central limit theorem）によれば、標本の大きさが増せば、一定の第一種の誤りの確率を持つ棄却域の範囲は広くなるので、一、〇〇〇回投げ検定と一〇、〇〇〇回投げ検定の結果が異なるのは、前者において帰無仮説が受容され、後者において棄却される場合に限られる。この場合、一、〇〇〇回投げ検定で採用された帰無仮説は、サイコロをさらに投げ続けることによって得られた頻度の観察結果という新たな経験的証拠に照らして覆されたのである。これはまた、一、〇〇〇回投げ検定で帰無仮説に与えられた正当化が、一〇、〇〇〇回投げ検定において撤回された、ということでもある。

もちろんこのことを以って、単純に、「一、〇〇〇回投げ検定の誤りが、一〇、〇〇〇回投げ検定によって訂正された」とはいえない。確かに、一〇、〇〇〇回投げ検定における帰無仮説の棄却が持つ第一種の誤りの確率は、一、〇〇〇回投げ検定における帰無仮説の保持に伴う第二種の誤りの確率に比べて低い(8)。しかし、このことは、あくまで「誤りの相対頻度が低いこと」しか意味せず、特定のサイコロ投げ検定において、「前者における仮説の棄却が正しく、後者におけるその受容が偽であること」までは含意しない。しかしもし、一、〇〇〇回投げ検定における帰無仮説の保持が間違っていた場合、サイコロを投げ続けることによって標本の大きさを増し、新たに得られた経験的な証拠にもとづいて、いったんなされた正当化を撤回し、その誤りを訂正することが、われわれにとって、理論的にも実践的にも可能なのである。

撤回不可能な正当化と訂正不可能な誤り

これに対して、確率化によるST仮説の正当化においては、仮説検定における、標本の大きさを増加させることによる新たな経験的証拠の獲得に対応する、「正当化を撤回し、誤りを訂正する方策」は存在し得ない。われわれは、いったん、あるST仮説を確率化によって正当化したならば、その正当化がもたらす誤りの可能性を、理論的には十分認識しながらも、その誤りを原理的に訂正不可能なものとして放置せざるを得ないのである。

「確率化による正当化の撤回不可能性（unrevocability）」、ないし「可能な誤りの訂正不可能性（incorrigibility）」とでも呼べるこの事態を、ここではつぎのような比喩で説明してみよう。私がある人に電話をかけたとする。私は、その人の正しい番号を押したつもりである。ただ、もちろん、押し間違えをしている可能性はつねにある。（かつ、遺憾ながら、私の電話機は旧式で、かけた番号を表示する機能を持たない。）ここで電話に相手が出てくれれば、私は、「正しい番号を押したのか、そうでないのか」を確かめることができる。しかし、電話は相手を延々と呼び出してはいるものの、いっこうに誰も出てくれない。相手が出てくれない以上、私は、これが間違い電話かどうかを確認する術を持たないし、また、間違い電話であると確認した上で、電話をかけ直すこともできない。

確率化によって確率論の適用を正当化しようとする科学者は、決して相手が出てくれず、自分が間違い電話をかけているかどうか分からないまま、延々と電話をかけつづけているようなものである。いや、科学者がおかれている状況の方が、さらに悪い。科学者はST仮説を真とすべく、確率化を行うために最善を尽くす。一生懸命、正しい番号を押そうとする私も同様である。にもかかわらず、一方は確率化に、他方は正しい番号を押すことに失敗する場合も十分あるのである。しかし電話の場合、正しい番号を押せば、正しい相手につながることに（一応）なっている。ところが、科学者が確率化を満足すべき程度まで見事やりおおせても、問題となっている事象列が、確率試行の近似態であるとは限

らないのである。科学者は、正しい番号を押しても、必ずしも正しい相手にかかるかどうか分からない、いわば欠陥品の電話を使っているようなものである。

話を戻そう。確率化によるST仮説の正当化は、確かにその仮説が偽である確率を劇的に低下させるが、同時にまた、確率0で残ったその誤りを、訂正不可能なものとするのである。言いかえると、確率化による正当化の撤回不可能性は、この、最低値に抑えられた確率値として、その可能性が理論的に認知されている誤りに対して、将来における訂正可能性の道を閉ざしているのである。また確率化による正当化が、訂正不可能なものとして放置するのは、この種の誤りだけではない。それは、われわれが「そもそも確率化を、満足いく程度になしえなかった」という意味で誤っていたとしても、その誤りをも訂正不可能なものとしてしまうのである。このように、その撤回不可能性によって、二重の意味での可能な誤りを訂正不可能なものとすることこそ、科学的仮説一般の仮説検定による正当化から、ST仮説の確率化による正当化を際立たせている特徴なのである。

一〇　まとめ

確率論を現実の対象に適用するとはいかなることか

以上論じてきたことを、改めてまとめてみよう。頻度的・測度論的確率論の現実への適用のあり方に関して確認されたのは、つぎの二点である。

(一) 現実の有限事象列が、数学的構造における標本点に対応する、模型上の無限収束事象列の、現実的近似態とみなされていること。(第五節)

(二) 大数の強法則によって、標本空間中の集合論的存在者に対応する事象が模型上に存在することが保証され、かつ、その模型上の事象を現実に適用する際に広く用いられている正当化の方法が提供されていること。(第七節)

これらはいずれも、数学的構造・模型・現実という三つの存在領域にわたる概念操作の一環である。確率論の現実への適用とは、端的に言って、このような領域横断的な作業に他ならないのであった。ちなみに、他の数学理論の現実への適用に際しても、通常、同様の三つの存在領域が設定され、それらにまたがった概念操作がなされている。もちろん細部を見れば、上の（一）・（二）のような特徴がひかえているが、その適用という操作の内実に関して、確率論が他の数学理論と比べ、取りたてて異なっているというわけではないのである。

正当化にまつわる三つの特異性

ところが、いったん適用の「正当化」に目を転じ、「頻度的・測度論的確率論の現実への適用の正当化は、確率試行仮説（*ST*仮説）の正当化に帰着すること」（第五節）を確認した上で、その正当化のあり方を問うていくにつれ、確率論の特異性が際立つことになる。そしてその特異性は、以下の三点に整理できる。

（一）まず、サイコロ投げ試行に即して考える限り、*ST*仮説、ひいては確率論の適用は、経験的証拠にもとづいて正当化できないことが分かった。*ST*仮説を経験的な証拠に照らして検証する、現時点で標準的と見なされている方法は、ネイマン＝ピアソン型の統計的仮説検定である。その仮説検定は、頻度的・測度論的確率論を理論的枠組みとし、一定の頻度収束を前提した上で、頻度極限値である確率値や期待値などを判定していた。結局、サイコロ投げにおける頻度収束を主張する*ST*仮説は、検定の前提ではあっても、その対象とはなりえないのであった。（第六節）

また、科学的仮説一般の実用主義的な正当化とは、仮説検定における確証の蓄積にもとづくものであった。従って、上記の経験的正当化の不可能性は、そのまま、その実用主義的な正当化の不可能性に直結するのである。（第五・六節）

他方、統計的仮説検定は、*ST*仮説以外のさまざまな科学的仮説、ひいては、それらの仮説で用いられているさまざまな数学理論の現実への適用を、経験的証拠に照らして検証する装置である。それらの仮説や、数学理論の適用は、仮

説検定に合格することで経験的に正当化されうるし、またその合格の蓄積によって、実用主義的にも正当化されることになるのである。

このように、少なくともサイコロ投げ試行に即して見る限り、ネイマン＝ピアソン型の統計的仮説検定の理論的前提となっている、頻度的・測度論的確率論の現実への適用には、経験的・実用主義的正当化の途が閉ざされているのに対して、その検定の前提とならない数学理論の現実への適用は、仮説検定によって、経験的・実用主義的に正当化されうる。結局、サイコロ投げ試行において確認された、経験的・実用主義的正当化の不可能性が、確率論の現実への適用にまつわる第一の特異性なのである。

（二）　経験的・実用主義的に正当化されえないST仮説、従って確率論の現実への適用は、確率化に代表される人為的操作によってのみ正当化されうる（第七節）。これもまた、確率論の適用が持つ特異性であるといえる。

（三）　さまざまな科学的仮説や数学的理論の現実への適用が、統計的仮説検定によって正当化されたとしても、その正当化は、例えば、より大きな標本を用いた新たな検定によって撤回可能である。同時に、このことは、正当化された後も残りうる誤りが、経験的証拠に照らして訂正可能であることを意味していた。それに対して、ST仮説ひいては確率論の適用に対して、いったん与えられた確率化による正当化は、その後、撤回することができない。このことは、確率化によってもなお残りうる誤用の、経験的な改訂可能性が閉ざされていることを意味する（第九節）。これらが第三の特異性である。

サイコロ投げ再論

これら三つの特異性を踏まえ、頻度的・測度論的確率論の現実への適用の正当化のあり方を、近代確率論の成立の端緒の一つとなったサイコロ投げ試行に即して、改めて振り返っておこう。確率論を、サイコロの目が出る頻度や期待値

の予測に用いることは、実際にサイコロを投げた後で、結果として生じた有限事象列における頻度安定現象を観察することで、経験的に正当化される事柄ではない。むしろそれは、われわれが、独立・同一性の条件を（近似的に）満たしつつ、サイコロを投げることで、言いかえると、収束事象列（の現実的近似態）を、われわれ自身が「作り出す」ことで、正当化されるのである。ところで、われわれは、たとえそのような正当化がなされたとしても、サイコロ投げが確率論の適切な対象ではない可能性が残ることを公認している。しかしながら、われわれは、経験的証拠に照らして、その正当化を撤回したり、誤用を訂正することができないまま、その試行に対して確率論を適用しつづけざるを得ないのである。この意味で、サイコロ投げの確率化によって、それに対する確率論の適用を正当化することは、われわれの退路を断ち切ることをも意味するのである。

確率化という投げられた「さい」

確率化は、サイコロ投げのみならず、さまざまな試行に対して用いられることで、頻度的・測度論的確率論の適用対象を大幅かつ急速に広げた。その結果もあって、冒頭に述べたように、確率論の多分野への浸透は、時には「確率論的帝国主義」と評されるほど目覚ましいものがある（cf. Gigerenzer et al., 1989, p. 271）。しかし一方で、確率化は、なお残っている可能性がある誤用を、（仮説検定における、標本の大きさの拡大に相当するような）何らかの経験的証拠にもとづいて改めることを許さないまま、確率論の適用領域を更にまた一歩広げる尖兵の役割を果たしている。この意味で、新たな試行を、独立・同一性条件の下で繰り返す時点で、まさに「さいは投げられた」といえる。確率化は、確率論の新たな領域への適用へ向けて、われわれをして、もはや引き返すことのできないルビコン河を渡らせてしまうのである。

二　特異性はどこまで成り立つのか

対象系と観察系

前節で再確認された特異性（一）（従って、それから導かれる（二）・（三）も含めたすべての特異性）は、サイコロ投げ試行に即して明らかにされ、また（以下で述べるように）事実上、ほとんどすべてのこれまでの頻度的・測度論的確率論の適用について確認できると思われる。しかしながら、それは必ずしも、原理的にそのすべての適用において成り立つわけではない。それでは、確率論の正当化は、いかなる場合に特異なものとなり、いかなる場合にそうならないのか。そのことを見るために、ここで〈対象系〉と〈観測系〉の区別を導入しよう。サイコロの一定の目が出る相対頻度の観察には、特段の装置を必要としないのに対して、例えば、コバルトのγ線放射の時間間隔の相対頻度の測定には、それ自体一つの物理系である複雑な観測装置、即ち〈観測系〉を必要とする。それに対して、この〈観測系〉によって測定される対象としての物理系（例えば、コバルトのγ線放射現象）は〈対象系〉と呼ばれる。

確率論の適用を検定する仮説検定

さて今、対象系に対して何らかの非確率的、即ち決定論的な仮説を設定し、観測系に対しては、例えば「一定の大きさの観測誤差の出現頻度は収束し、その頻度分布は正規分布となる」という確率論的な仮説（この種の仮説は「確率誤差仮説」と呼ばれる）を立てたとしよう。ここで、これら二つの仮説の連言を取り、両仮説が主張している効果を一定の仕方で加算することで、観測結果に対して特定の頻度分布を与える、一つの確率論的仮説を作ることができる。この確率論的仮説の下では、一定の「誤りの確率」を求めることができるので、この仮説は統計的仮説検定の対象となりうる。例えば、決定論的な古典力学の仮説は、このような仕方で一定の確率誤差仮説と組み合わされることで、統計的に

検定されるのである。

同じことは、「対象系も観察系も確率論的」、「対象系は確率論的で、観察系は決定論的」、という二種類の仮説についても言える。そこで、これら三種類の仮説から二つの仮説を——例えば「対象系は決定論的だが、観測系は確率論的」という仮説と「対象系は確率論的だが、観測系は決定論的」という仮説——選び、これら計三個の仮説の組に対して、それぞれ仮説検定を行うとしよう。このような仮説検定において、一方の仮説が退けられ、他方が採用されることで、対象系ないしは観察系に対する確率論の適用は、経験的に反証または確証されることになる。その場合、前節でまとめた、確率論の適用の正当化に関する特異性は成り立たない。

確率論適用に対する検定の事実上の不可能性

しかし、私の知る限り、このような仮説の組に対する仮説検定は、これまでなされてこなかった。その理由は二つある。一つは、ある程度以上の精密な測定において、一定の条件下でつねに同じ測定値を与えつづけるような精度の高い装置が、これまで開発されたことがなかったという事情である。(ちなみに、そのような装置が開発されることは、これからも望み薄であろう。)その結果、同じ条件下でも、測定値がばらつくのが常態となっている。この場合、少なくとも、観測系に対して「同一の条件下でつねに同じ観測値を与える」という決定論的な仮説を立てることはできない。また、個々の測定値の出現頻度が本当に収束するかどうかは別にして、このばらつき現象を説明しうる仮説として、確率論的仮説以外の仮説は、現在のところ、考案されていないのである。結果として、「対象系は確率論的で、観察系は決定論的」という仮説を設定し、それを帰無ないし対立仮説として仮説検定を行う可能性は、事実上、閉ざされているのである。

残る、「対象系は決定論的だが観測系は確率論的」という仮説と、「対象系も観測系も確率論的」という仮説の間の検

定も、二つめの理由によって、少なくとも現在のところ実行不可能である。第三節で触れたように、確率論は、サイコロの目のように、「つぎに何が起こるのかが予測不可能なほど不規則な事象」に対して適用されてきた。そして、確率論以外に、このような不規則事象を扱える数学理論はながらく登場しなかった。ところが、ようやく二〇世紀後半になって、一つの有力な対抗馬としてカオス理論が提案された。カオス理論の登場によって、「不規則な振る舞いを見せる一つの対象系が確率論的か決定論的か」を論ずること、即ち、「それに対する確率論の適用が妥当かどうかを論ずること」が、現実的な意味を持ち始めたのである。ただし、現在でもなお、主として統計学上の技術的困難から、同一の対象系に対して、一方ではカオス理論を当てはめる仮説と、他方では確率論を適用する仮説を検定対象とする統計的仮説検定法は開発されていない。現在、用いられている統計的検定法は、せいぜい、線形的な確率論的仮説を帰無仮説とし（そして対立仮説を設定せず）、それを棄却することで、対象系の非線形性を推測する、サロゲート法だけである[9]。もちろんこのサロゲート法によって、確率論の適用が反証されるわけでもないし、カオス理論の適用が確証されるわけでもない。非線形的な確率論的仮説も、カオス的ではない非線形的な決定論的仮説も存在するからである。このように、適切な統計的手法の欠如によって、確率論適用の検定の最後に残った可能性も、現在のところ封じられているのである。

特異性が成り立つ範囲

以上見てきたように、対象系と観測系が区別でき、その中でも特に、確率論の適用の是非を統計的検定によって経験的に検証することが可能な場合は原理的に存在する。しかし、完全な観測装置や、測定値のばらつきを説明する対抗理論、さらには、カオス理論か確率論かを択一的に選択する統計的手法の不在という、つまるところ偶然的な理由によって、その可能性は、現在のところ事実上閉ざされているのである。

他方、そもそも対象系と観測系のどちらか一方しか存在しない場合は、確率論の適用を経験的に（ひいては実用主義

的に）正当化する可能性は原理的にない。このような場合、確率論の適用の正当化に関する三つの特異性は、例外なく成り立つのである。

一二　哲学的含意

サイコロ投げ試行は、近代確率論の最初の適用対象の一つである。だが、その適用事例は、経験的にも実用主義的にも正当化されえず、またそれが誤用であった可能性を現在でも払拭できていないばかりか、もはやその適用を経験的証拠に照らして撤回することもできない。さらに同様の事態は、すべての確率論の適用についても、現在のところ事実上成り立っている。これらが、本論のさしあたっての主張である。

しかし、本論が指摘してきた「三つの特異性」は、統計的仮説検定一般の前提である「確率標本仮説（RS仮説）」（そしてそれが含意する「確率試行仮説（ST仮説）」）に則して考えられた場合、また別の哲学的含意を持つように思われる。以下、その含意のいくつかについて、ごく簡単に見通しを述べておきたい。

RS仮説に含意されるST仮説

まず、RS仮説に含意されているST仮説の内容を、ある物体の質量を測定するという事例に即して、改めて確認しておこう。先にサイコロ投げについて見たように、この場合でも、無限個の測定値からなる仮説無限母集団が設定される。さらに、その中で、一定の測定値（測定限界が一〇〇分の一グラムである場合は、例えば 1.01g から 1.02g までの値）が持つ存在頻度が、例えば、真の値を平均とする分散1の正規分布に従うと想定される。その上で、この無限母集団から一つの値を抽出しては戻すという操作が無限回繰り返された場合、一定の値が抽出される相対頻度が収束すると見なされるのである。[10] これらの前提の下で、「その物体の質量の測定という物理的な操作は、「それを繰り返した場合、

一定の測定値の出現頻度に関して、収束事象列の近似態を生み出す」という意味で、一種の確率試行である」と主張するのが、ここでのST仮説なのである。ちなみに、このST仮説は、頻度収束の極限値やその分布を特定していない。従って、現実の繰り返し測定における一定の値の出現頻度が、仮説無限母集団におけるそれの存在頻度を、どの程度まで反映しているのかは、未規定のままである。

RS仮説に含意されるST仮説の経験的正当化は不可能である

第六節で紹介したように、仮説検定において検定対象となるのは、通常、科学者の興味の的である帰無ないし対立仮説である。RS仮説は、それらの仮説から、測定値である標本の出現頻度を導くために、不可欠の前提として用いられているに過ぎない。このような場合、検定結果のいかんにかかわらず、RS仮説が、経験的に確証されたり反証されたりすることは、原理的になかったのである。

他方、実験・観察・標本調査で用いられている「確率化」などの標本抽出法が、本当に求められている確率標本を与えるものかどうかを確かめるために仮説検定がなされることもある。この場合、例えば、仮説無限母集団についてのある一定の仮説が既知のものとして前提され、標本抽出に関する二種類のRS仮説が帰無・対立仮説として立てられた上で、確率化を施された標本抽出が行われる。このような場合でも、サイコロ投げについての検定で確認したように、帰無・対立いずれの仮説が選ばれたとしても、両者がともに含意するST仮説が保持されることに変わりはない。また、この場合のST仮説は、測定という操作について、かつそれについてのみの仮説であり、その中で対象系と観測系を区別することができない。したがって、前節で確認したように、この場合のST仮説に対しては、サイコロ投げにおけるST仮説と同様、それを検定対象とする仮説検定を行うことができないのである。結局、このような仮説検定は、RS仮説が主張している抽出頻度の分布形を検証する作業に過ぎず、ST仮説を検証するものではないのである。

以上のように、通常の仮説検定の場合は、RS仮説の経験的正当化がなされず、またRS仮説が検定対象となる場合でも、それが含意するST仮説の経験的正当化はなされない。ちなみに、前者の場合でも、RS仮説が含意するST仮説は検定対象となっていないので、結局のところ、いずれの場合でも、ST仮説が統計的仮説検定によって経験的に正当化されることは、原理的にないのである。

科学的経験の超越論性

すると、サイコロ投げ試行についてのST仮説と同様に、RS仮説に含意されるST仮説にも、先に指摘した三つの特異性が成り立つことになる。まず、経験的正当化の不可能性という第一の特異性が持つ哲学的含意を見てみよう。

そもそも、科学の特徴の一つはその実証性、即ち、その仮説が実験や観察の結果に照らして検証されるという点にある。そして、これまで繰り返し触れてきたように、数理科学における仮説の経験的な検証作業において、少なくとも現在のところ、統計的仮説検定がほとんど不可欠といってよい役割を果たしている。実際、少なからぬ科学者が、仮説検定が利用可能な場合は、それが、仮説の検証の方法として「最も望ましい」、ないしは「必要不可欠」であると主張しているのである。[11] このように、さまざまな科学的仮説に対して経験的検証の機会を与えている仮説検定であるが、それ自体は、上で見たように、ST仮説を、「検証されざる前提」としているのである。結局、仮説検定の不可欠の前提としてのST仮説は、現代数理科学における仮説の証拠立て作業、端的に言って、「科学における経験」を可能にする条件の一つであるといえる。

またこのST仮説は、測定という物理的出来事についての仮説であるという点で、論理的真理を表わす命題などとは異なり、事実についての主張（factual claim）であるといえる。いま、経験的に検証されえない命題を「アプリオリな命題」、事実に関する命題を「総合的な命題」と呼べば、ここでのST仮説は「アプリオリで総合的な命題」であるこ

とになろう。さらに、経験の可能性の条件となっているアプリオリで総合的な命題を、カントにならって「超越論的な命題」と呼ぶことにすれば、ここでの*ST*仮説は、また、超越論的な命題であるともいえるだろう。結局、統計的仮説検定として遂行される、科学における仮説の経験的な証拠立て作業は、その作業における測定を一種の確率試行であるとする、超越論的な前提の下で行われていることになる。頻度論・測度論的確率論と古典統計学を、仮説の検証の方法論的枠組みとして採用し続ける限り、科学的経験は、このような意味での、一種の超越論的な構造を持たざるをえないのである。このことはまた、このような超越論的構造を科学的経験において認めない経験主義的ないし自然主義的な立場（例えば、クワインの方法論的一元論（methodological monism））は、科学の現状に則さない、単なる画餅に過ぎないことを意味するのである。(12)

超越論的前提の行為による正当化

経験的に正当化されえない*ST*仮説は、確率化によって正当化するしかなかった。先に見たように、*RS*仮説が仮説検定の前提となる場合でも、検定の対象となる場合でも、確率化が行われていた。そして、*RS*仮説が含意する*ST*仮説は、その確率化によって正当化されると見なされていたのである。ところで、確率化とは*ST*仮説がその存在を主張している収束事象列（の近似態）を作り出す人為的操作、即ち「行為」であった。ここで注意すべきは、*ST*仮説を正当化するのは、あくまで、この行為の所産である有限事象列における頻度安定現象の観察や記述ではないという点である。実際、通常の測定においては、この行為は一回しか行われない（即ち、標本の数は通常一である）。この場合、有限事象列を構成する事象の数はわずか一である。科学的経験の超越論的前提を正当化するのは、むしろ「行為そのもの」なのである。言いかえると、ここでは、「実際にそのような「作り出す」行為がなされたかどうか」が焦点となるのである。

科学的仮説を正当化するのは、「科学者（たち）が、実際に確率化という行為を行ったという事実」であるとしても、この事実そのものを正当化する手立ては、もはや、少なくとも統計的方法論の中には存在しない。統計的方法論に話を限れば、行為を行ったという事実は、科学的仮説の最終的な正当性付与者（justifier）なのである。この限りでは、「科学における正当化の連鎖は、一定の最終点で止まる」ともいえる。しかし、哲学的分析はここで止まらない。われわれは、統計的方法論を越えて、さらに「この事実がいかにして正当化ないしは受容されているのか」を問うことができるのである。このような問いはまた、「一定の統計的方法論自体は、いかにして科学者の共同体において、さらには社会一般において受け入れられているのか」という問題へもつながるであろう。これらについては稿を改めて論じたい。

不確実だが訂正不可能な超越論的前提

科学的経験の超越論的前提としてのST仮説は、たとえ確率化という行為によって正当化されたにしても、偽である可能性が残る。しかし、再三述べているように、いったん確率化がなされた以上、われわれは、その正当化を撤回できず、もしその前提が偽であったとしても、それを経験的証拠に照らして訂正できないのである。ここで、超越論的前提が「疑いえないほど確実」であるから、それが訂正不可能なのではないことに留意すべきである。われわれは、正当化された前提が、なお二重の意味で疑わしいことを公認しつつ、それを「証拠によっては訂正不可能なもの」とせざるをえないのである。

訂正不可能なST仮説が撤回され、サイコロ投げや仮説の検定に対する確率論の適用が見直されることも、将来十分ありうるであろう。しかしその見直しは、確率論の適用を保証していた、行為によって正当化されたST仮説が、証拠に照らして訂正された結果起こる事態ではあるまい。では確率論の適用の見直しは、もし起こるとすれば、いかなる仕方で生ずるのか。これは、「現在の確率論的な統計学がいかにして受け入れられているのか」という先に挙げた問題と

表裏一体の問いである。従って、これについても別稿に譲りたい。

さらなる一般化への見通し

以上、本論で述べてきたことは、さしあたって、頻度的・測度論的確率論と、それにもとづいた古典的統計学の仮説検定において一般的に成り立つと考えられる。しかし同様の事柄は、現代におけるいま一つの有力な確率論と統計理論、即ち、主観的解釈を施された測度論的確率論と、それに依拠するベイズ統計理論についても妥当する、と私は考えている。具体的には、ベイズ的な種々の仮説検定においても、一定の「事実主張」が「検定されざる仮説」として前提されており、その仮説は経験的な証拠に照らしてではなく、科学者の一種の行為によってのみ正当化される。そして、いったん正当化された前提は、もはや経験的には訂正不可能となる、と考えるのである。しかし、この話もまた、別の機会に委ねたい。

いずれにせよ、ここで素描した哲学的な諸含意は、科学が、古典的理論・ベイズ的理論を含め、現在利用可能な、いかなる確率論的な統計的方法論を採用しても、好むと好まざるとにかかわらず、普遍的に成り立つ事柄であろう、というのが現在の私の見通しである。[13]

（了）

参照文献

合原一幸編、池口・山田・小室著（二〇〇〇）『カオス時系列解析の基礎と応用』、東京、産業図書。
Blalock, H.M., (1979) *Social Statistics*, 2nd edition, Singapore, McGraw-Hill.
Bernoulli, J., (1713) *Ars conjectandi*, reprinted in his *Ars conjectandi : opus posthumum : accedit Tractatus de seriebus infinitis, et Epistola Gallicè scripta de ludo pilae reticularis*, 1968, Bruxelles, Culture et Civilisation.
Chalmers, T., (1967) "The Ethics of Randomization as a Decision-Making Technique, and the Problem of Informed Consent"

in *Report of the Fourteenth Conference of Cardiovascular Training Grant Program Directors*, Washington D. C., U. S. Department of Health, Education, and Welfare.

出口康夫（一九九八）「統計学から見たクワインの科学論」『アルケー（関西哲学会年報）』第六号、六〇―七〇頁。

―――（一九九九）「現代科学論カント風―超越論的でアプリオリな命題は科学において存在するか―」『理想』第六六三号、一一―一七頁。

Doob, J., (1941) "Probability as Measure", *Annals of Mathematical Statistics*, 37, pp. 206-14.

――― (1976) "Foundations of Probability Theory and Its Influence on the Theory of Statistics", in Owen, D. B. (ed) *On the History of Statistics and Probability*, New York, M. Dekker, pp. 197-204.

Fine, T., (1973) *Theories of Probability*, New York, Academic Press.

Fisher, R., (1925) *Statistical Methods for Research Workers*, Edinburgh, Oliver & Boyd.

Gigerenzer, G. et al., (1989) *The Empire of Chance*, Cambridge, Cambridge UP.

Hacking, I., (1975) *The Emergence of Probability*, Cambridge, Cambridge UP.

Henkel, R., (1976) *Tests of Significance*, California, Sage Publications.

一松信・竹之内脩編（一九九一）『新数学事典』改訂増補版、大阪、大阪書籍。

Jeffreys, H., (1931) *Scientific Inference*, Cambridge, Cambridge UP.

Kolmogorov, A., (1933) *Grundbegriff der Wahrsheinlichkeitsrechnung*, Berlin, Springer.

――― (1963) "The Theory of Probability", in Kolmogorov et al. (eds) *Mathematics, Its Contents, Methods, and Meaning*, vol. 2, pp. 229-64, Mass., M. I. T. Press, (Russian original 1956).

日本数学会編（一九八五）『岩波数学辞典』第三版、東京、岩波書店。

Quine, W. V., (1981) *Theories and Things*, Mass., Harvard UP.

Weatherford, R., (1982) *Philosophical Foundations of Probability Theory*, London, Routledge.

注

(1) 例えば、ドゥーブは、確率を「相対頻度の極限値」であると明確に解釈するのに対して、コルモゴロフやクラメルは、確率の解釈から、無限個の事象の存在を含意する「極限値」を排除しようとする。

(2) cf. Hacking, 1975.

(3) 例えば、頻度の収束を示す「大数の弱法則」が最初に証明された、J・ベルヌーイの『推論法』（一七一三）第四部では、サイコロ投げが、頻度収束が生ずる具体例の一つとして挙げられている。(cf. Bernoulli, 1713)

(4) コーシーの収束判定条件については、『岩波数学辞典』四三六頁や『新数学事典』四四五頁を参照されたい。

(5) 以上は、仮説検定をかなり強引に単純化して説明したものである。実際に、二項分布を帰無・対立仮説とする場合には、ここで示したような「観測値が棄却域に入れば必ず帰無仮説を棄却し、そうでなければ必ず保持する」という非確率化検定ではなく、観察値に応じて帰無仮説の棄却の確率が0から1までのさまざまな値をとる、確率化検定が用いられる。またネイマン＝ピアソンの基本定理については『新数学事典』七〇一頁、『岩波数学辞典』八二五頁を参照されたい。

(6) 例外とは、例えば一回のサイコロ投げにおいて1の目が出る確率が1で、かつ無限回の繰り返しにおいてすべての結果が1であったような場合である。このような場合、単元集合の確率値は1となる。

(7) もちろんすべての場合において、この方針が有効であるわけではない。例えば、単一の値を仮定しているような対立仮説がないような場合には、第二種の誤りの確率は特定できず、従ってこのような選択方針も意味をなさない (cf. Henkel, 1976, p. 50)。

(8) 同一の検定において、第二種の誤りの確率は、常に第一種のそれと等しいかより大きく (Henkel, ibid., p. 81)、またこの場合、一、〇〇〇回投げ検定と一〇、〇〇〇回投げ検定は、同じ第一種の誤りの確率を持つ。

(9) サロゲート法に関しては、合原編（二〇〇〇）第六章を参照されたい。

(10) ここでいう「繰り返し」とは、標本の数を増すことを意味する。ここで、標本の数と、標本の大きさの区別に注意されたい。標本の大きさとは、「そこから一つの標本統計量（例えば平均）が導かれる観測値の数」である。例えば、質量を一〇〇回測って平均をとった場合、標本の大きさは一〇〇だが、標本の数は1である。この作業を再度「繰り返して」二つの平均値を得た場合、標本の数は2となる。本文の質量測定の例では、話を単純にするため、一回の測定における標本の大きさを1としてある。

(11) 例えば、「統計的検定でない臨床治験は、科学的に無意味であり、結果として患者に有害無益の負担を強いるので、一切、行

ってはならない」といった類の主張がなされることがある（cf. Chalmers, 1967）。この種の主張は、科学の他の分野においても決して珍しいものではない。

(12) ここで論じた立場からの、クワイン流の経験論・自然主義に対する批判については、拙論（一九九八）、（一九九九）を、またクワインの方法論的一元論については、Quine, 1981, pp. 71-72を、それぞれ参照されたい。

(13) 本論は、現在刊行準備中の、ある論文集のために、一九九七年から一九九八年にかけて執筆された原稿に、加筆・修正をほどこしたものである。元の原稿の執筆にさいしてさまざまな助言を頂戴し、また、本誌に掲載することをお認め頂いた、論文集の編者の方々に、ここで改めてお礼申し上げたい。

（筆者　でぐち・やすお　京都大学大学院文学研究科助教授／哲学）

dieser Form lebendige Wirklichkeit werden.

Die Einzelheit von Kierkegaard und die Privatheit sind begrifflich von verschiedener Art, aber sie werden in der jetzigen Welt gleichgesetzt. Für die Puritaner des siebzehnten Jahrhunderts war die private Sphäre sehr viel öffentlicher als die politische Sphäre, weil sie glaubten, daß die Freiheit und die Demokratie in der Privatsphäre in wesentliches Weise ausgeführt werden können. Die private Sphäre hat jedoch heute solche Tiefe verloren.

Th. W. Adorno beachtet nicht Kierkegaards "Innerlichkeit" selbst, sondern ihre Erscheinungsweise in der ersten Hälfte des zwanzigsten Jahrhunderts. Die Innerlichkeit verbirgt sich, um einen Ort zu schützen, wo das Selbst sich zu Gott verhält. Adorno erkennt, daß dieses Verbergen und das Geheimhalten der Privatsache in gleicher Weise erscheinen, daß die Innerlichkeit sich in das eingeschlossene Bewußtsein verwandelte, und daß die subjektive Wahrheit verkleinert wurde. Zwischen Kierkegaard und Adorno gibt es eine zweifache Verdrehung der Wirklichkeit, und damit ein Verbergen dieses Wirklichkeit. Die Versachlichung der Innerlichkeit geschieht nicht so, daß die darin beinhalteten Verhältnisse zur Transzendenz ausgestoßen werden, sondern daß diese Verhältnisse als Waren im Vordergrund des Warentäuschgeschäfts in Umlauf kommen. Wir können zwischen der religiösen Freiheit und der wirtschaftlichen nicht mehr eine Grenze ziehen.

Die Privatheit ist kein Begriff, die Sache der Religion als solche zu bestimmen. "Ist die Religion etwas Privates?" ist eine Leitfrage dafür, den Zusammenhang zwischen Religion und Öffentlichkeit zu erforschen. Sie wird dann mit der anderen Frage, "ist die Religion etwas Öffentliches?", verbunden und an sie weiter geleitet.

Is the Die Cast?

How to Justify the Use of Probability and a Transcendental Characteristic of Scientific Experience

by

Yasuo DEGUCHI

Associate Professor of Philosophy
Graduate School of Letters
Kyoto University

Probability theory is now applied in many fields of science as well as in various aspects of our everyday life. In the wake of its extraordinarily wide use, this paper raises the following questions: what it is that we apply it to reality, and how, if possible, we can justify its applications? In answering these questions, this paper identifies the distinctive features of justification for such applications which are not shared by justifications for the use of other mathematical theories. The discussion then draws some philosophical implications from the distinctive features that hold, in particular, for the use of the theory in statistical tests of significance.

Although many mathematical formulations and non-mathematical interpretations of probability have been proposed in the past, nevertheless this paper confines its considerations exclusively to Kolmogorovian measure-theoretic theory of probability with limit-of-frequency interpretation, on the grounds that that theory has acquired the status of the *de facto* standard in scientific and ordinary uses of probability.

The distinctive features mentioned above are disclosed through detailed analyses of *strong law of large number* and *Neyman-Peasonian test of significance.* And those can be summarised in the following three points.

(1) **Impossibility of empirical and pragmatic justification**: One might argue that applications of probability to some actual events could be justified *empirically* by means of observations, if there are any, of stability of their relative frequency. Or it would be claimed that predictive successes of some probabilistic models could serve as pragmatic justification for their applications. But either of these sorts of justification is impossible in so far as empirical tests of its applicability or of predictions of the models should be made within the framework of one or another form of significance tests in classical statistics, as that of any scientific hypothesis is expected to be done.

(2) **Justification by means of randomisation**: Applications of probability can only be justified by such human interventions on their objects as *randomisation.* Randomisation or its ilk is intended to make the objects *random* in the sense that they meet a sufficient condition for the applications set by the strong law of large number.

(3) **Unrevocability of justification and incorrigibility of possible misapplications**: There still remains the possibility that randomisation would fail to make the objects random. However, once the application of probability has been justified by randomisation, no choice is left open to revoke the justification and to correct potential misapplications in response to some new empirical evidence. So once randomised, *the die is cast.* We are then destined to apply the probability theory to another new sphere.

It is also claimed that any significance test in classical statistics is built on a *stochastic trial hypothesis* or *ST hypothesis,* as I call it, that measurement in the test is a stochastic trial or random process. The three distinctive features also hold for justification of the ST hypothesis. The ST hypothesis cannot be justified empirically and pragmatically but only by randomisation. However, once randomisation has been made, it can never be revoked on the basis of some empirical evidence even if it is false.

Wherever it is available, statistical hypothesis testing is regarded as the most reliable or the only appropriate method for an empirical test of any scientific hypothesis. Given the dominance of this testing, the distinctive features with the ST hypothesis seem to have some significant philosophical implications that can be outlined as follows.

(1) **A transcendental characteristic of scientific experience**: The ST hypothesis is *a priori* in that it is not susceptible of any empirical test, although it can be justified as true by means of randomisation. It is also *synthetic* in that it is a factual claim, that is, a statement of measurements as physical processes rather than that of logic or mathematics. Being a priori and synthetic, it is *transcendental* in that it constitutes, as an indispensable assumption of statistical testing, a necessary condition for the possibility of an empirical test of any scientific hypothesis, or, in short, for that of scientific experience. In other words, experience in science should be made under a transcendental assumption.

(2) **Justification of the transcendental assumption by an act**: Randomisation is an act taken by a scientist or scientists. The transcendental assumption or the ST hypothesis cannot be justified by any product of the act, namely, the stability of relative frequency of a particular value in repeated measurements, but rather only by the act itself, or the fact that she or they actually did it. The act is also the final justifier of any scientific hypothesis so far as statistics is concerned.

(3) **Incorrigibility of the uncertain assumption**: The transcendental assumption becomes incorrigible by empirical evidence once justified by an act although we still recognise the possibility that it might be false. Its incorrigibility does not mean its certainty at all. Rather it is incorrigible even though it is uncertain.