経済学研究 *59−3*
北海道大学 2009. 12

古典的統計理論における三つの問題点

園　信太郎

1．はじめに

ここで古典的と呼ぶのは，classical とか traditional とか orthodox とか形容される統計学的理論のことであり，頻度論的理論とほぼ一致する。人物名で示唆すれば，J. Neyman, E. S. Pearson, A. Wald, E. L. Lehmann と言った雰囲気であり，権威的である。これに対峙するのがベイズ統計学（Bayesian statistics）であるが，伝統的にはベイジアンの方が古いので，「古典的」という形容詞は文字通りに受け取れば誤解を招きやすい。

実は古典的理論には少なくとも三つの本質的難点があり，しかし，内外の教科書の多くを見ればわかるように，これらの難点は言及されることが極端に稀である。そこで筆者はこの場をかりて，これらの問題点をコンパクトにまとめておくこととした。例えば数学科や，さらには理系出身だからとか，文系だが計量的方法を採っているとかで，統計学の授業を担当することとなってしまった諸氏にはぜひとも授業中に，これらの難点に，言及して頂きたいのである。

2．「確率」の定義の失敗

第一の難点は，「確率」の定義に失敗していると言うことである。例えば，「このコインを投げ上げて裏となる確率は二分の一である」という陳述を取り上げよう。「投げ上げ実験を多数回繰り返すことによって，裏が現れる相対的頻度は，限りなく二分の一に近づく」というのがその「定義」だが，ここで「限りなく近づく」とはいかなることなのかを確定する必要がある。もし，「将来現れるであろう相対的頻度が 1/2 からずれる確からしさが，小となる」と言うことで，確率論での確率収束に相当する事柄を持ち出そうとするのならば，その際の「ずれる確からしさ」は紛れもなく「確率」への言及なので，その際の「確率」を定義する必要がある。

一方，実際の実験に訴えて，日常的な尺度からして「多数」回だと思われてしまう回数にわたって実験を遂行して，「わずかに」二分の一からずれている結果を得た場合，これに基づいて「裏の確率は 1/2 として良い」と結論しようとするのならば，その「わずかな」ずれが「有意，significant」ではなく，ひとまず「無視できる」と判断する際の根拠が問われることとなる。なにしろ「多数」回の試行の末に苦労して得た実験結果であるので，「わずかな」ずれでも「無視できない」はずだからなのである。しかし，この「有意か否か」の見極めの為に，あたかも「検定」の様に，「確率」を持ち出すのならば，やはりその際の「確率」を定義する必要がある。

いずれにしても「暗黙の内の論理の悪循環」か，無定義の「確率」の天下り的導入を行うこととなり，「確率」の定義に失敗することとなる。なお，数値実験で「相対的頻度の極限としての確率」をもっともらしく例示するやり方の教育上の効果には二面性があることに注意すべきである。つまり，「わずか」ではあっても「ずれている」のであるから，この「わずか」は本当に「わずか」と見なして良いのかという論理的「問いかけ」を無視して，「雰囲気的なもっともらしさ」という教育上の便利さのみを採るのならば，科学的な思惟を育てることとは

ならないであろう。

3．変量の実現値による置き換え

この「置き換え」が最も露骨に現れるのは，いわゆる区間推定においてである。論点を明確にするために，母集団分布が平均θ分散1の正規分布$N(\theta, 1)$に従い，θを未知固定（unknown but fixed）の母数として，これを推定する状況を想定する。但し，母数空間は非現実的だが実数の全体としておく。ここで問題の母集団からの大きさnの無作為標本（$X_1, \cdots, X_n$）を想定しておくと，未知固定母数θに対する信頼区間$\left[\overline{X}-\frac{z(\alpha/2)}{\sqrt{n}}, \overline{X}+\frac{z(\alpha/2)}{\sqrt{n}}\right]$が呼び出される。ここで$z(\alpha/2)$は，標準正規分布の上側50$\alpha$%点である。これは信頼係数$1-\alpha$の信頼区間と呼ばれることがあるが，実際には区間を値として取る変量であり，$1-\alpha$は，未知固定母数θがこの可変的区間で覆われる頻度論的な「確率」である。所が，実際にデータ（$x_1, \cdots, x_n$）が得られると，このデータの平均$\bar{x}$による変量$\overline{X}$の「置き換え」が行われて，区間$\left[\bar{x}-\frac{z(\alpha/2)}{\sqrt{n}}, \bar{x}+\frac{z(\alpha/2)}{\sqrt{n}}\right]$が算出されて，これをも信頼係数$1-\alpha$の信頼区間と呼ばれるのである。$\theta$は未知固定であるので，それがこの固定されている既知の区間に属するか否かは，未知固定で定まっており，信頼係数$1-\alpha$を頻度論的に解釈する道は絶たれている。つまり事後的に(データが得られた後に)$1-\alpha$の内訳を述べようとしても，それができなくなるのである。古典的理論では，ただ信頼係数という言葉のみが導入されて，事後的$1-\alpha$は無定義なのである。

ここでよく持ち出される状況説明がある。データが得られる前に，類似の標本抽出が限りなく反復され，各抽出において事後的な信頼区間が形成されるとして，このように得られる未知固定の事後的信頼区間の系列を想定すれば，未知固定の母数を含む区間の相対的頻度が$1-\alpha$に「近づく」と想定しても不自然ではないであろうから，そこで，実際に得られる「その」区間を，この未知固定区間系列の冒頭の区間が「実現した」ものと見なすというのである。だが,「確率」の定義の問題を脇におくとしても，このような状況説明から，実現していない残りの未知固定区間系列に対して導入される極限値$1-\alpha$と，実際に正に実現した「その」区間とが，いかにして「結びつく」のであろうか。極限値$1-\alpha$は，冒頭の「値」には影響されないのであり，この「結びつき」の様式こそできれば透徹した形式によって表記すべきなのである。

変量を何の断りもなしに「実現値，realized value」で置き換えるという作業は，そのままでは(数学的論理では)正当化できないものである。一方，実際の現場では,「その」区間が未知固定の母数を含む「確からしさ」こそが問題なのである。古典的理論は現場の要求にこたえることに失敗しているのである。

なお「確率」の定義の問題がここでも絡んで来る。$1-\alpha$を，大数の強法則に訴えて，この値へと概収束するある種の確率変数列の極限として理解する場合に，概収束の概念を定義する際に導入されている確率測度の(解釈的な)定義が述べられなければならないはずである。その際「確率とは何か」という本質的な問と対峙しなければならない。

4．「検定」における「確率」の無定義性

古典的理論に忠実である限り,「検定」における「確率」は無定義であり,「解釈」は利用者にまかされている。従って，第一種及び第二種の過誤に対する「確率」の大きさの選択は，各自の選択に(理論上は)任されることとなる。特に，いかなる値を「小さな確率」と見なすのかは，まさに一人一人の問題なのである。古典的理論は，いかなる基準に基づいて「小さい」と判断すべきなのかについては，沈黙するのである。所が，一方では,「過誤たち」によって惹き起される「損失」について古典的理論は言

及したりもする。さらには,「損失」と「確率」との積の和として,「期待される損失」にも言及したりする。だが,「損失」と「確率」との積がいかなる根拠に基づいて導入され，それが何故「不確定性下の選択」のための合理的な規準をもたらすに至るのかについては，やはり沈黙するのである。

古典的理論は「不確定性下の選択」の定式化の理論的重要性に気づいているのだが，定式化を行う際にいかなる行為規範に基づいてそれを遂行すべきかについては，言わば捉え損ねているのである。「世界」に直面している「個人」は「未知固定の真の状態」を問題とするのだが,「世界の状態」を表現する母数θに対しては「本来の確率」を導入できず，一方で,「θが与えられている情況での，過誤がもたらされる確率」を問題とするにもかかわらず,「過誤がもたらされる確率」そのものは導入できないのである。A. Waldの枠組では「事前分布」と呼ばれるものは，抽象的な荷重であって「本来の確率」ではなく，それ故に「結局いかなる行為規範に基づいて,「その」決定を下すのか」という問に答えられないのである。

象徴的な表記法を用いれば,「検定仮説が通用する確率」×「検定仮説の下で検定仮説が棄却される確率」及び「対立仮説が通用する確率」×「対立仮説の下で検定仮説が採択される確率」が導入できないので，決定を下す者にとっての「本来の期待損失」が計算できないのである。

なおここでも「確率」の定義の問題が絡んでくる。つまり，例えば第一種の過誤に対する「確率」αを,「何らかの限りなく繰り返される決定に関る極限」とする場合，この「極限」αを定義する際に導入されるであろう確率測度の(解釈的な)定義とはいかなるものかということである。しかも，決定を下す者にとっての「本来の損失」とはいかなるものなのかという問が,「確率」の定義の試みに絡んでくることとなる。この厄介な状況に対して古典的理論は,「無定義」という態度をとって，沈黙するのである。

5．解決策としてのベイズ統計学

Leonard Jimmie Savage（愛称Jimmie)はSavage（1954, 1972)において,「個人」の行為に関する七つの規範を公準系(postulates)として導入して，個人的確率(personal probability)の概念を基礎づけ，さらに効用(utility)の概念を個人論的な立場から基礎づけ，副産物として，経済学の基礎に関る原理である，期待効用最大化の原理を導いたのである。数学的には，彼は「存在定理」を提示しているのであり，P１から(第６公準の特別な場合である)P６′までの六つの公準に基づいて,「合理的な」個人が一意的に定まる個人的確率を保持するに至ることを示したのである。さらにP７の導入によって，期待効用最大化の原理が正当化され，これによって統計的決定理論の精確な基礎づけが提示されたのである。

つまり,「確率」に対しては個人的確率を提示し,「変量の実現値による置き換え」に対しては,(観測の値によって条件づけられる)条件つき個人的確率を提示し,「検定」に対しては個人的確率及び効用に基づく期待効用の最大化を提示したのである。なお,「実現値による置き換え」だが，これは象徴的には，Prob(・|X=x)を考えることで，事後的に(変量Xに)値xが代入されるに至るということである。

今日から視れば，サヴェジは，彼の個人的確率に基づくベイズ統計学を提示したこととなる。彼は始めは，自身の立場から，古典的理論の諸手法が正当化できるはずだと思っていたのだが，結局，古典的理論と対決する理論を構築することとなる。彼は，統計学の基礎に関る深い思索を通して，主観確率(subjective probability)に基づくベイズ統計学こそ「正しい」統計学であると，固く信じるに至るのである。

6．付記

サヴェジの立場からすれば,「確率」も「効用」も「個人」が無くなれば消え去るのであり,「科学」における「事実」とは，元来個人のオ

ピニオンであり，「客観的な」事実などは存在しないこととなる。客観論的な立場からすれば，これは認め難い見解であり，そこで今日でも厳しい対立が続いている。なおサヴェジの理論及び思索については園（2001，2007）がある。Webcat で検索すれば所蔵している最寄の図書館がわかるはずである。とにかく，「「確率」とは何か」とか，さらには「「確率」が「ある」とはいかなることか」という問は，多分永遠の課題だが，「統計学において」という制約のもとでは，大体の決着がつくかもしれないのである。ぜひとも諸賢の明察をまちたい。

参考文献

Savage, Leonard Jimmie, *The Foundations of Statistics, Second Revised Edition, Dover Publications*, New York, 1972. 第一版は，John Wiley & Sons, New York, より 1954 年に出ている。この「基礎論」への一つの「読み」として，園（2001）がある。またサヴェジ氏の思索への一つの「読み」として，園（2007）がある。

園 信太郎，『サヴェジ基礎論覚書』，岩波出版サービスセンター，東京，2001 年 12 月 20 日。

園 信太郎，『サヴェジ氏の思索』，岩波出版サービスセンター，東京，2007 年 8 月 31 日。

2009 年 5 月 5 日（火）