

高校での統計教育、大学での確率統計教育

楠岡成雄（東京大学大学院数理学研究科）

奇妙なタイトルと思われると思うが、少し過激な主張を書かせて頂く。

初等中等教育において統計教育を強化すべきということは文科省でも強く主張されているようである。また、大学の教養教育では微積分学、線形代数は多くの大学で理系の必修科目となっているが、確率統計についてはあまり重視されていない。一方では、「確率」「統計」「データ解析」という言葉は社会では頻繁に登場する。しかし、地震学者が「大地震が**地方で今後10年以内に発生する確率」ということを述べる時、この人は「確率」という言葉の意味をどこまで理解して使っているのかと怪しむことがしばしばある。そもそも、「統計教育」「確率統計教育」という言葉を用いる時、そこでの「確率」「統計」という言葉が何を指し、その教育の目的は何か、何故そのような教育が必要か、ということが極めて曖昧なままに論じられている。ここでは、この問題を考えていきたい。

また、初等中等教育には過剰な期待がよせられている。例えば、初等中等教育においては「金融教育」が必要という話を文科省で聞いたことが、その理由は詐欺的な話に引っかかる人が多いので初等中等教育段階での教育が必要ということなのだそうである。しかし、あれもこれもという要求は、かえって教育効果を損ねる。それぞれの教育段階で何が本当に必要かということも込めて考えていきたい。

(1) 初等中等教育における確率統計教育

まず新指導要領における「確率統計」の内容を以下にまとめる。

小学校

D 数量関係

第3学年：資料を分類整理し、表やグラフを用いて表したり、読み取ったりする

第4学年：目的に応じて資料を集めて分類整理し、表やグラフを用いて表したり、
特徴を調べたりすることができるようにする。

ア 資料を二つの観点から分類整理して特徴を調べること。

イ 折れ線グラフの読み方やかき方について知ること。

第5学年：目的に応じて資料を集めて分類整理し、円グラフや帯グラフを用いて表したり、特徴を調べたりすることができるようにする。

第6学年：資料の平均や散らばりを調べ、統計的に考察したり表現する

ア 資料の平均について知ること

イ 度数分布を表す表やグラフについて知ること

具体的な事柄について、起こり得る場合を順序よく整理して調べる

中学校

D 資料の活用

第1学年（データの分析）：目的に応じて資料を収集し、コンピュータを用いたりするなどして表やグラフに整理し、代表値や資料の散らばりに着目してその資料の傾向を読み取ることができるようにする

ア ヒストグラムや代表値の必要性和意味を理解すること。

イ ヒストグラムや代表値を用いて資料の傾向をとらえ説明すること。

〔用語・記号〕 平均値、中央値、最頻値、相対度数、範囲階級

第2学年（確率）：不確定な事象についての観察や実験などの活動を通して、確率について理解し、それを用いて考察し表現することができるようにする。

ア 確率の必要性和意味を理解し、簡単な場合について確率を求めること。

イ 確率を用いて不確定な事象をとらえ説明すること。

第3学年（統計的推測）：コンピュータを用いたりするなどして、母集団から標本を取り出し、標本の傾向を調べることで、母集団の傾向が読み取れることを理解できるようにする。

ア 標本調査の必要性和意味を理解すること。

イ 簡単な場合について標本調査を行い、母集団の傾向をとらえ説明すること。

〔用語・記号〕 全数調査

高校

数学I（データの分析）：統計の基本的な考えを理解するとともに、それを用いてデータを整理・分析し傾向を把握できるようにする。

ア データの散らばり 四分位偏差、分散及び標準偏差などの意味について理解し、それらを用いてデータの傾向を把握し、説明すること。

イ データの相関 散布図や相関係数の意味を理解し、それらを用いて二つのデータの相関を把握し説明すること。

数学A 場合の数と確率（確率）：場合の数を求めるときの基本的な考え方や確率についての理解を深め、それらを事象の考察に活用できるようにする。

ア 場合の数

イ 確率

(ア) 確率とその基本的な法則、(イ) 独立な試行と確率、(イ) 条件付き確率

〔用語・記号〕 排反

数学B 確率分布と統計的な推測：確率変数とその分布、統計的な推測について理解し、それらを不確定な事象の考察に活用できるようにする。

ア 確率分布

(ア) 確率変数と確率分布（確率）

確率変数及び確率分布について理解し、確率変数の平均、分散及び標準偏差を用いて確率分布の特徴をとらえること。

(イ) 二項分布（確率）

二項分布について理解し、それを事象の考察に活用すること。

イ 正規分布

正規分布について理解し、二項分布が正規分布で近似できることを知ること。また、それらを事象の考察に活用すること。

ウ 統計的な推測

(ア) 母集団と標本(統計的推測)

標本調査の考え方について理解し、標本を用いて母集団の傾向を推測できることを知ること。

(イ) 統計的な推測の考え

母平均の統計的な推測について理解し、それを事象の考察に活用すること。

これを見ると、小学校においては「確率」という言葉が現れない。また、資料の整理におけるグラフも比例関係を表すグラフと共に教えられているようであり、私は小学生に対して適切な内容になっていると思う。また、中学1年生の段階で、平均値、中央値などの統計における基本的な概念が教えられており、これも適切と感じている。気になるのは中学2年における「確率」である。実は中学校の教科書を見ると、先験的確率(数学的確率とも呼ばれるが、現代から見れば不適當な言葉遣いである)、統計的確率(経験的確率、確率頻度説もここではこれに含めることとする)などを紹介し、これも確率、あれも確率といった述べ方がされている。しかし、これらの確率の考え方は、真っ向から対立した考え方であり、歴史上多くの論争があり、未だ解決がついていない問題を含んでいる。「数学」という教科でこういった形で教えるのが妥当であろうか。まず、何に対して確率が問えるかがこれら論争の焦点の一つであるということに注意すべきである。初等中等教育において「確率」を教える時は、万人が認める「(確率の)考察の対象となるものは一定条件で何回でも実験できるか、又は少なくとも実験できると考えられる事柄」(丸山儀四郎: 確率および統計 共立出版 1956年 よりの引用)に限って教えるべきである。

「確率」の概念を本当に理解するのは、高校生でも不可能である。「確率」の概念を巡る論争で現れた「頻度確率」や「数学的確率」といった言葉は初等中等教育では用いることを避けるべきである。「確率」の概念を教えることは、エプシロン・デルタ論法で微積分を教えるより難しいことを自覚すべきである。「確率」を教える時は正確にできたさいころを振る、玉を箱から取り出すといった純粋な試行をベースにしていけばよい。その上で確率概念は広い分野で用いられているという事実と、それが正しい使い方であるかははっきりとしていないということに少し触れる程度にとどめるのが良いと思う。

また初等中等教育においては、「確率」を「同様に確からしい」といった言葉を用いて教えられており(この言葉は先験的確率につながるので、本来は望ましくないが、初等中等教育の段階ではこのような曖昧な言葉使いも仕方がないと思う)、正規分布のような連続的な値を取る確率変数の話題に突然飛ぶのは問題があると思う。正規分布はお話にとどめるべきで、「発展」に位置づけるべきである。

また、統計教育は「データの整理」を教えることを重点にすべきである。統計的推測も、もし高校数学の中で教えるならば、論理的に理解可能なベイズ推定を教えるのがよい(題材は慎重に選ぶべきだが)。統計的推測には「正解」はないはずで、母平均の統計的な推

測の持つ意味は、高校生には理論的理解が難しく、統計手法についてはコンピュータによる統計推定の実験などを通しての理解にとどめるべきと考える。

(2) 大学における確率統計教育

大学初年時においてこそ、確率統計を教えるべきであるが、教える側がまず「確率」とは何かを真剣に考えるべきである。数学科で「数学としての確率論」を教えるわけではないので、(難しいことだが) コルモゴロフの公理化された理論から離れて概念を説明すべきである。「確率変数」を Ω 上の関数とのみ理解しているようでは、一般教養としての確率統計を教えることは無理である。また、経済学における「不確実性下での意思決定理論」にも注意を払うべきである。「確率」を知ることは、それが目的ではなく、どのような「行動」あるいは「政策」をとるべきかを決定する過程に必要な事柄にすぎない。また、このようなことを教える時、「確率の客観性」の度合いは適用対象によって大きく異なることなども教えるべきである。

統計についても大学初年時には、「統計推測の考え方」について教えるべきである。確率モデル、あるいはその族である統計モデルの考え方、水準5%で有意であるとは何を意味するのかといったことをまず教える必要があるように思う。今日では、よい統計ソフトが存在し、データを入れさえすればかってにコンピュータが水準5%で有意である等と判断してくれるが、その意味することをまったく理解しないまま、「有意であることが統計学的に示されている」等々と述べている人が非常に多いように感じる。

一方、統計的手法は様々なところで用いられるようになってきているので、早い時期にどのような手法があるかを教える必要があるのかもしれない。

(3) おわりに

統計学と確率論は本来別なものであったが、19世紀に統計の基礎は確率論であるという認識が完全に定着し、20世紀に急速に発展していった。しかし、今日のデータ科学への統計手法の応用を見る時、そのデータは自然科学的なデータでないことも多く、「確率」を裏付けにした統計手法の正当化は難しくなりつつあるように見える。データに対する統計モデルも科学に裏打ちされたはっきりしたものではなく、方便のような所がある。(数理ファイナンスでは、証券価格の確率過程モデルを前提としながら、それをあまり信用せずに様々なリスク管理手法を行っており、状況は似ている。) 極めて「工学的」(利用できそうなものは理屈抜きに何でも使う) に統計学的手法が用いられているという印象がある。このようなことまでも教える時、それは数学教育という枠にはもはや入らない。統計教育にはそのような側面があることも考えて大学でのカリキュラムを考える必要がある。