

統計的有意性に依存したリスク評価の弊害について

The adverse effects of risk assessment by statistical significance

永井宏幸

Koko Nagai

NPO 法人市民科学研究室

NPO Citizen's Science Initiative Japan

放射線疫学の分野で、放射線のリスクの有無を判定するのに「統計的有意性」の概念を使うのはあたりまえのようである。しかし、統計学が確率を扱う学問である以上、推論は本来連続量によって示されるものである。統計的有意性とは、これを2値判断に落とし込む自動的・機械的ルーチンであって、科学的思考になじむものか疑問であるといわざるをえない。

われわれが現在「数理統計学」として利用できるのはふたつの数学しかない。ベイズ統計学と「非ベイズ統計学」である。「非ベイズ統計学」はわれわれが共通してよく知っている統計学であるが、その推論と決定の関係はよく理解されているわけではない。歴史的にはこの問題を巡ってフィッシャーとネイマン-ピアソンの間で激しい論争があった。しかし、その論争の意味がわが国ではあまり理解されず、あいまいなまま「ハイブリッド統計学」となって使われていると統計学者は指摘する。この「ハイブリッド化」が推論と決定の間の混乱や誤解を助長する原因になっているようである。

その混乱や誤解を示す最たるものが「仮説検定」の「統計的有意性」であろう。

統計的有意を示す論文は帰無仮説を棄却するという。それでは、有意でない論文は「帰無仮説に関してなんの証拠能力も持たない」のか、それとも「帰無仮説が棄却されないという証拠能力を持つ」のか、ここが曖昧なのである。この曖昧さが、統計分析の結果をレトリック次第でどのようにでも誘導できる隙間をうみだしている。

2016年にアメリカ統計学会は「統計的有意性とP値に関する声明」を発表した^[1]。広く研究に使われているp値には誤解や誤用が多いばかりでなく、そもそも「二分割された一方の側で、結論が直ちに『真実』となったり、他方の側で『誤り』となったりすることはありえない」と指摘したうえで、「ポスト $p < 0.05$ 時代」に向けて新たな舵を切ると宣言した。そして2019年3月には、特集号「21世紀の統計的推論：P < 0.05を超える世界」に40を超える論文を掲載し、声明に沿った議論が盛んにおこなわれている。

講演では、放射線疫学の分野で「統計的有意性」が社会にどのように理解され利用されているかを裁判と教育の例で示して、その弊害について論じる。また「仮説検定」に代わる分析の要約方法について、ベイズ統計学の考え方に沿った試案を提示することにしたい。

1 Wasserstein RL, Lazar NA. Editorial: The ASA's statement on p-values: Context, process, and purpose. The American Statistician 2016; 70: 129-133. 主要部分については佐藤俊哉氏（日本計量生物学会）による和訳がある。
<http://www.biometrics.gr.jp/news/all/ASA.pdf>