

[最終講義]

統計的推論—基礎と資源解析から—

東京海洋大学 名誉教授 山田 作太郎

Statistical Inference—From Its Basis and Fish Population Analysis —

Sakutaro YAMADA

0. 本報告は筆者に与えられた最終講義，統計的推論—基礎と資源解析から—，の再録である。

1. 論理と解釈—統計の基礎から—

初等統計学を教えていて，時々 " 教えるににくいなあ " と思うところがある。歴史的にはそれらがどのように考えられてきたのか，少し紹介したい。

まず区間推定から。 $X_1, X_2, \dots, X_n$ を正規母集団 $N(\mu, \sigma^2)$ からランダムサンプルとし， σ は簡単のため既知としよう。

$$P(\bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}}) = 0.95 \quad (1)$$

から， μ の信頼係数 95 % の信頼区間は

$$(\bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}}) \quad (2)$$

で与えられる。この時 $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$ なる観測値を (2) に " 代入 " して作られた区間

$$(\bar{x} - 1.96 \times \frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96 \times \frac{\sigma}{\sqrt{n}}) \quad (3)$$

は μ を含んでいるのだろうか？

歴史的には，Fisher と Neyman の間で，ここでの区間推定や，次に紹介する " 仮説検定 " に関して論争があった。Neyman は確率の頻度説の下で (1) について通常の説明 (4 行下) をしたが，しかし区間 (3) が μ を含むかどうかについては何も言えないとした。Fisher は確率の頻度説を否定し，しかし区間 (3) は高い確率で μ を含むと考え，fiducial probability なる考えを提案した。しかし，この Fisher の考えは多くの統計学者の理解を得られなかった。(1) と (3) がなにも関係なければ，勿論 (3) について何も発言することは出来ないが，そうではなく，もし確率の頻度説を採用するならば，(1) が成立する時，多数の繰り返し実験で，その都度 (3) でもって μ の区間推定を繰り返したならば，そのうちの 95 % ぐらいは正しく推定しているので，今手元にある (3) なる実現値に対して，それも μ を含むだろうと推論するのは自然である。0.95 という数字はその期待に対して与える信頼を表す数字であると考えられる。これが現在のおおかたの教科書における説明である。後に Fisher は，確率の頻度説の下で，(3) を含む異なる相対度数を持つどんな実験系列族も認識出来なければ (Fisher のいう認識可能族の否定) (3) も μ を含むと期待していいとした。この事は数学的確率論の創始者コルモゴロフも述べている。

Fisher と Neyman は確率の理解に違いがあったが，それをのぞいても，上記の問題は " 数学の理論を応用するにあたっては，数学的な論理で押し通せる部分とその後の解釈する部分がある。 " ことを述べていると考えられる。ひとつの論理で最後まで進むことは出来ないのだろうか？ 数学者で生物学も研究しているグロモフは，「21 世紀数学への期待」で，「…新しい論理を数学にどのように受け入れるかが問題になってくるのではないのでしょうか。ここでいう「論理」とは，厳密でないが，説得力のあるもの，問題を見れば正しいが数学的には正しくない，その「習慣」は正しい，というものです。」と述べ新しい数学の論理の出現に期待している。

仮説検定は歴史的には最初有意性検定として現れた。それは，" データと整合的でない (帰無) 仮説を (一定の水準で) 棄却する " ことを目的としたものである。データと仮説の非整合性は p -値，

$$P(|\bar{X} - \mu_0| \geq |\bar{x} - \mu_0|) \quad (4)$$

で測る。一定の水準として Fisher は 5 % を用いた。有意性検定は、対立仮説が不要でありデータに基づく判断であるが、検定（統計量）の選択基準があいまいであるという欠点を持つ。この欠点を克服するために、Neyman と Pearson は仮説検定を考えた。それは、現今の教科書の大半がこの考えに従っているのであるが、仮説のみでは検定統計量の選択は一般になし得ず、仮説に対して対立仮説を考え、一定の有意水準の下で、検出力（＝ 1－第Ⅱ種の過誤の確率）を最大にする検定統計量（とそれに基づく棄却域の決定）を選択するものであると定式化された。しかし、この仮説検定は仮説か対立仮説のどちらを取るかの決定問題としての性格が強く、この点も Fisher が異論を唱えたところである。特に仮説が棄却されないときでもそれが真であるという主張は証明されるべくも無く、仮説を棄却する証拠は見つからなかったと言う意味である。有意性検定では（データに基づいて）データが仮説から乖離している程度をはかっており、仮説検定の一定有意水準の考えより多くの情報を含むという優れた点も持っている。参照の枠組みという点からは、有意性検定は仮説に示された分布のみを扱うのに対して、仮説検定では仮説と対立仮説をあわせた分布の集まりを考察の対象としているという相違がある。

この後 Neyman の資源解析に関する研究を紹介し、また私自身の数理統計学の研究についてこれも簡単に紹介した。

2. 不完全データの持つ情報—資源解析から—

水産資源の変動様式はその生残、成長、再生産を通して記述される。資源解析の研究を難しくしている最大の原因は資源量の観測値がとれないことである。資源を取り巻く環境に関するデータでも取りにくいものがある。従って漁業や環境の変動の影響を受けた資源の変動を正しく記述する法則は存在せず、モデルに基づく研究が中心となる。

上に述べた観測の難しさから、" 不完全なデータ " でもって資源解析はなされる場合が多い事になる。

若い時から、" 不完全データはどのような情報をどのくらい失っているのか、逆にどのくらいの情報は所持しているのか " という問題には関心があつた。この問題に関して Paulik と Seber による標識魚の不完全な報告に基づく資源量推定の問題があり、その推論法の統計的正当化の研究を少し紹介したが、ここでは字数の制限の関係で省略する。

資源解析にとって資源の直接推定は漁業データに依存せず、また資源変動のモデルをも用いないで資源量を推定するという点で、また比較的速く資源量が推定されるという点で大変重要なテーマと思われる。最終講義では、標本抽出単位における資源量の不完全な観測をする Counting Sampling (CS) と、ただ魚がいるかないかだけを観測する Presence—Absence Sampling (PAS) という、ともに不完全なデータを生むサンプリング法、しかし前者の方が資源量推定に関して後者より多くの情報を含むだろうことが予想される、に関して比較検討を行った研究を紹介した。なお、サンプリングについてもうひとつ大事な点は観測費用である。

各観測ステーションにおける魚（卵、貝）の数 X は負の二項分布 $NB(m, k)$ にしたがうとし、平均 m を推定する問題を考えよう。 $X = x$ の時、サンプリング法が CS の時は、観測される Y は二項分布 $B(x, p)$

$$P(Y = y \mid X = x) = {}_x C_y p^y (1-p)^{x-y} \quad (5)$$

に従うとする。一方 PAS による観測 U は値 1 = 存在、0 = いない を取り、

$$P(U = 1 \mid X = x) = a(1 - \exp(-bx)) = 1 - P(U = 0 \mid X = x) \quad , \quad 0 < a \leq 1, b > 0 \quad (6)$$

を満たすとする。

観測モデル (5), (6) を、 m の推論に関して比較するために、サンプラーの性能を揃えておく必要がある：

$$1 - \exp(-b) = p \quad (7)$$

を仮定する。この時

(A) 各サンプリング法に基づく観測 1 個が、 m に関して持っている平均情報量は、Fisher 情報量で測ると

$$I_C(m) > I_{PA}(m) \quad (8)$$

が成立し、予想どおり CS の方が PAS より多くの情報を含む。

しかし、PAS の方が CS より観測費用が安いので、PAS の観測ステーション数を増して推測精度を高めうることが期待される。全部 N 個の観測ステーションがあり、そのうち n 個は PAS の $N - n$ 個は CS に当てられるとしよう。すると、

(B) $\frac{N}{2} < n_0 < N$ をみたす（ただひとつの） n_0 が存在し、 $n \geq n_0$ なるすべての n に対して

$$MSE(\hat{m}_{PA}) \leq MSE(\hat{m}_C) \quad (9)$$

が成立し、この n_0 を具体的に書き下す事が出来る。(9) で MSE は平均 2 乗誤差、 $\hat{m}_{PA}$ は n 個の PAS による観測に基づく m の最尤推定量を表す。

一方観測費用に関する考察をすると別の知見が得られる。 c_1, c_2 はそれぞれ P A S, C S を 1 標本抽出単位実施するのにかかる費用とし, $c_1 < c_2$ とする。観測のための総予算 C_0 は $c_1 N < C_0 < c_2 N$ を満たすとする。 m のそれぞれのサンプリング法からの最尤推定量の観測個数による重み付き平均

$$\hat{m} = \frac{n}{N} \hat{m}_{PA} + \frac{N-n}{N} \hat{m}_C$$

の M S E を最小にする n は

$$n = \frac{c_2 N - C_0}{c_2 - c_1}$$

で与えられ, この値が $\frac{N}{2}$ より大きいための必要十分条件は $C_0 < \frac{N(c_1 + c_2)}{2}$ である。

以上はすべて条件 (7) の下でのものである。この条件をはずせばいろいろなことが考えられるが, それは省略する。

(参考文献)

Cox,D.R.; Principles of Statistical Inference, Cambridge University Press,2006.

グロモフ;「論理」と「構造」,数学セミナー,2003年4月号.

コルモゴロフ;確率論,数学通論Ⅲ,東京図書,1958.

Lehmann,E.L.; The Fisher, Neyman-Pearson Theories of Testing Hypotheses :One Theory or Two?

Journal of American Statistical Association,Vol.88,1993.

竹内啓;数理統計学的方法的基礎,東洋経済新報社,1973.

Yamada,S.,Zenitani,H.; Efficiency of presence-absence sampling used to estimate the mean abundance of fish eggs. Fisheries Science ,Vol.71,2005.