「統計的評価論による誤差理論」の解説

伊藤先生が名古屋大学助教授時代に，統計学的問題の一つを考察された論文である．その問題とは，「ある量を偶然誤差を伴う観測によって得た観測値に基づいて評価するとき，その最良の評価値をどのように定めたらよいか」というものである．

先生は東京大学を卒業されてから名古屋大学に就職されるまでの間，内閣統計局に勤務され，確率論の統計学への応用に深い関心をお持ちであったと察せられる．京都大学では統計学の講義をされることは無かったが，確率論のセミナーに統計学の研究者を度々お招きになり，そのような折にはいつも，統計における標本値（統計量）は，確率変数の実現値と見なすべきだと強調しておられたことを思い出す．

先生が残された確率論の教科書には，統計学への応用に関する話題が散見されるが，最後の教科書というべき岩波講座基礎数学「確率論」[*1]の第 4 章，4.8 節では「 Gauss の誤差論」という表題のもと，観測誤差の問題が論じられ，例題 4.8 [*2]で次の具体的問題が取り上げられている：「三角形の内角 $\theta_1, \theta_2, \theta_3$ を同じ測定法で無関係に計って X_1, X_2, X_3 を得た． $\theta_1, \theta_2, \theta_3$ を評価する最良の式（X_1, X_2, X_3 の 1 次式）を求めよ.」 そしてこの問題が次のような確率論の問題として設定されている：$X_i, i = 1, 2, 3$ は平均値が θ_i で同じ分散 v をもつ互いに独立な正規確率変数系とし，θ_i の評価値 $\bar{X}_i$ は線形な不偏統計量，すなわち

$$\bar{X}_i = a_i + \sum_{k=1}^{3} a_{ik} X_k, \quad i = 1, 2, 3$$

と表され，その平均が θ_i と一致するものとする．評価値のうちで最小の分散をもつものを最良の評価値とする．それを求めよ．

このような設定のもとで，θ_i の最良の評価値は簡単な計算で求まり，それは

---

*1 I (1976), II (1977), III (1978)

*2 221 ページ

$$\bar{X}_1 = \frac{\pi}{3} + \frac{2}{3}X_1 - \frac{1}{3}X_2 - \frac{1}{3}X_3$$

$$\bar{X}_2 = \frac{\pi}{3} - \frac{1}{3}X_1 + \frac{2}{3}X_2 - \frac{1}{3}X_3$$

$$\bar{X}_3 = \frac{\pi}{3} - \frac{1}{3}X_1 - \frac{1}{3}X_2 + \frac{2}{3}X_3$$

で与えられる．また，評価値の精度はその分散の値（小さいほど精度が高いと理解して）で与えられ，いずれも $\frac{2}{3}v$ である．

このように，先生がお若いころ，ここで解説している論文で取り組まれた問題が，後年執筆された教科書で再現されている訳だが，もちろん論文のほうが問題がより一般的に論じられ，それによって教科書の内容の背後にひそむより深い理論を知る事ができる．特に上の三角形の内角の例でみた直接的な分散最小化の方法の他に，最小自乗法の原理に基づく方法が確立されており，一層理解が深まる．

例えば，上の三角形の内角の例の場合，結果として $\bar{X}_1 + \bar{X}_2 + \bar{X}_3 = \pi$ となっているが，それがたまたまそうなったのか，そうなる事に理論的理由があるのか，上の各 $\bar{X}_i$ の分散最小化の方法ではよくわからない．ところでこの論文で一般的に確立されている最小自乗法の原理に基づく方法をこの例に応用すると，$\bar{X}_i$ を変数と考え，X_i を定数と考えた二次形式

$$\frac{1}{v}\left[(X_1 - \bar{X}_1)^2 + (X_2 - \bar{X}_2)^2 + (X_3 - \bar{X}_3)^2\right]$$

を制限条件　$\bar{X}_1 + \bar{X}_2 + \bar{X}_3 = \pi$ のもとで最小とするものとして　$\bar{X}_i$ が得られるので，$\bar{X}_1 + \bar{X}_2 + \bar{X}_3 = \pi$ となることは，こちらの方法では初めから一般論として保障されるのである．

この論文でも，その典型的な応用例として，やはり三角形の内角の例が論じられているが，ここでは三つの内角と一つの外角，即ち，$\theta_1, \theta_2, \theta_3$ と $\pi - \theta_1$ を無関係に同じ精度で観測し，その値に基づいて $\theta_1, \theta_2, \theta_3$ を最良の精度で評価するという問題である．上と同様にこの問題を定式化する．即ち，$\theta_1, \theta_2, \theta_3, \pi - \theta_1$ の 観測値を X_1, X_2, X_3, X_4 とすると，それらは互いに独立で共通の分散 v をもち，平均がそれぞれ，$\theta_1, \theta_2, \theta_3, \pi - \theta_1$　となる正規確率変数系である．$\theta_1, \theta_2, \theta_3$ の評価値をそれぞれ，$\bar{X}_1, \bar{X}_2, \bar{X}_3$ とすると，それらはそれぞれ，平均が $\theta_1, \theta_2, \theta_3$ に等しく，X_1, X_2, X_3, X_4 の 1 次式として表される正規確率変数である．そして，そのような評価値で分散が最小となるものを，最良の評価値とする．

このように定式化すると, 上で述べた, より簡単な例の場合と全く同様に二通りの方法で最良の評価値が求まる．後者の最小自乗法の原理に基づく方法による場合は，$\bar{X}_1, \bar{X}_2, \bar{X}_3$ を変数とする二次形式

$$\frac{1}{v}\left[(X_1-\bar{X}_1)^2+(X_2-\bar{X}_2)^2+(X_3-\bar{X}_3)^2+(X_4-(\pi-\bar{X}_1))^2\right]$$

を制限条件　$\bar{X}_1+\bar{X}_2+\bar{X}_3=\pi$ のもとで最小とする問題となる．いずれの方法によっても，答は容易に求まって，

$$\bar{X}_1=\frac{3}{5}\pi+\frac{2}{5}X_1-\frac{1}{5}X_2-\frac{1}{5}X_3-\frac{2}{5}X_4$$

$$\bar{X}_2=\frac{1}{5}\pi-\frac{1}{5}X_1+\frac{3}{5}X_2-\frac{2}{5}X_3+\frac{1}{5}X_4$$

$$\bar{X}_3=\frac{1}{5}\pi-\frac{1}{5}X_1-\frac{2}{5}X_2+\frac{3}{5}X_3+\frac{1}{5}X_4$$

と与えられる．また，評価値の精度である分散の値（値が小さいほど精度が高いと理解する）は,

$$V(\bar{X}_1)=\frac{2}{5}v,\quad V(\bar{X}_2)=V(\bar{X}_3)=\frac{3}{5}v$$

となる.

先生の岩波講座基礎数学「確率論」で確率論を学ぶ人は多いと思うが，その第 4 章 4.8 節の内容は，上でみてきたように，先生の若い時代のお仕事がその土台となっている．その論文が今回，電子版公開されることは大変意義深いものがある.

渡辺 信三[*3]

[*3] 京都大学名誉教授