

統計データ解析センター

健康の数量化

駒澤 勉

1. はじめに

今日ほど人の健康管理対策が医学的、社会的に重要な時代はない。健康を害してからの病院における診断・治療の守備型の個人医学を更に充実させねばならないことは言うまでもないが、これからの健康問題を考える時、特に病気にさせない各種集団健康診断を中心とした攻撃型の集団医学は、積極的に健康の予知・予防を進める上で、また保健経済面から見ても重要な課題である。集団健康診断で得た各々の健康に関わる基本的な各種検査項目の結果を的確にかつ総合的に健康指標化（パラメーター化）することは予後の判断に役立てるのに必要である。この健康指標作りには、数量化の方法を適用した。また、実際に脳血管損傷疾患や虚血性心疾患に関する大量な集団健康検診データをもとに実証的研究を行った。

2. 外的基準のない数量化（アイテム・カテゴリー・データの分析）

数量化III類についてはいろいろな性質が調べられているが、ここでは、アイテム・カテゴリー・データの時の基準尺度を相関係数と分散での相異について循環器系の集団健康検診データをもちいて論ずる。

$$\begin{aligned} \langle \text{データ} \rangle = D &= \{ \delta_{i(jk)} \}, & \text{個体} : i &= 1, 2, \dots, n, \\ \text{アイテム} : j &= 1, 2, \dots, m, & \text{カテゴリー} : k &= 1, 2, \dots, l_j, \\ \delta_{i(jk)} &= \{ 1; C_{(jk)} \text{に該当}, 0; C_{(jk)} \text{に非該当} \}, \end{aligned}$$

ただし、 $C_{(jk)}$; アイテム j ・カテゴリー k の内容。数量 $\{ \chi_{(jk)} \}$ の解法は次式で解く。

$$(2.1) \quad A\mathbf{x} = \lambda\mathbf{x}$$

ただし、

$$\begin{aligned} A &= \{ a_{(jk)(uv)} \}, & \mathbf{x} &= \{ \chi_{(jk)} \}, & n_{(jk)} &= \sum_{i=1}^n \delta_{i(jk)}, \\ n_{(jk)(uv)} &= \sum_{i=1}^n \delta_{i(jk)} \delta_{i(uv)}, & a_{(jk)(uv)} &= \alpha \left(n_{(jk)} - \frac{n_{(jk)}n_{(uv)}}{n} \right) \\ & & & (j, u = 1, 2, \dots, m, k = 1, 2, \dots, l_j, v = 1, 2, \dots, l_u). \end{aligned}$$

- 1) 基準尺度が相関係数 γ の数量化III類の場合: $\lambda = \gamma^2_{XY}$, $\alpha = 1 / (m \cdot n_{jk})$
- 2) 基準尺度が分散 σ^2 の主成分的分析の場合:

- (a) 生データ; $\lambda = \sigma^2_Y$, $\alpha = 1/n$
- (b) 正規化データ; $\lambda = \sigma^2_Y$,

$$\alpha = n \sqrt{(n - n_{(jk)})(n - n_{(uv)}) / (n_{(jk)}n_{(uv)})}.$$

3. 実 例

第1の実例は検診の検査項目が血圧因子中心、第2の実例は血圧と脂質因子中心から成る。前者は単独因子と思われる場合で、分散の散布図は全体的に広がった。後者は多因子の場合で、散布図に明らかな布置構造が現れた。