回答の信頼性

野　田　一　雄

(1966年11月受付)

Reliability of Responses in Social Surveys

Kazuo Noda

We have studied on a mathematical method of the analysis on the reliability of responses in social surveys. We made a new model in which individuals are supposed to answer questions with the probabilities depending on their reactions in the categories of the questionitems. In this case, we applied the idea of latent class analysis to this model which are given (2.1), (2.2). The definition of the reliability of responses is given by Def 2. It can be shown that, subject to two conditions, this definition of the reliability is equivalent to the concord of the responses in panel surveys (Def 3.), through the theorem Th 1.

Institute of Statistical Mathematics

まえがき

社会調査法において配慮しなければならない最も基本的な問題の一つとして，調査から得られる回答の信頼性をどのように考えればよいかということがあげられる．ふつう面接調査において採用されている方法は，連続2回の調査をある一定時間間隔をおいて繰返し，同一サンプルの前後の二つの回答の一致度をみて，その質問の信頼性の程度と考えるやり方である．しかるにこの方法で面接調査の信頼性を調べるためには，つぎの

a）調査対象の前後の調査時点における意見の変化

b）前回の調査時における回答の記憶の影響

c）personality からくる回答の変動

d）質問主旨の理解の度合

e）質問編成および回答のとり方からくる影響

f）調査員に原因のあるゆがみ

等について何らかの方法で識別する必要がある．回答の不変性を験めるという立場からは，a), b）についてはあらかじめ調査デザインを企画するときに，その影響が生じないように配慮すべきであろう．いずれにせよ，これらの問題を考えるには，一つの調査目標を定めたとき，それらの質問系とそれによる調査対象からの回答の収得について過程全体に適合したモデルを設計しなければならない． そのとき質問系を設定する場合に， 対応する調査対象間の 質的な差異によって生じる回答の変位を同時に考えておく必要がある．回答の各選択肢に反応する比率を解釈するにあたって，その信頼性を調査対象の分類された層についてただすことができるからである．かくして質問系の内容と編成に関して信頼性の限界が定められるならば，その範囲において信頼性の精度を高めるための方法を対処する問題へ移行することが可能であろう．

統計数理研究所では面接調査における回答誤差の研究を続けているが，今回上述のような目的のために，第2部第1研究室を中心に全国調査を企画し，回答の信頼性を研究することになった．そのために本調査にはいる前にプリ・テストを実施しているが，そこにおいては上述のf)「調査員に原因のあるゆがみ」の影響を調査から外すようにした．つまりサンプルに質問票を渡し直接回答を記入させる自記式の方法を採用したわけである．調査員の問題はこれまで何回かの研究がなされているし，これを質問編成と同時に処理すればいたずらに機構を複雑にするからである．結局プリ・テストの目的は回答のメカニズムを探知することと，それに対応する質問編成法を工夫することに集中することにした．

さて，c)，d)，e) の問題は互に関連していて，これらを切り離して議論すべきものではない．調査対象の質問に対する反応の過程を考えると，普通の状態では，まず質問文から受けたイメージのもとで自己の思想を集約し，与えられた回答の選択肢を判別決定するということになるであろう．そこで問題となるのは，かかる発想の過程が個人の質的な差異によって，それぞれ異ったメカニズムを生むことは当然であるにしても，ある類似性をもった部類に類別されるかどうかということである．例えば，質問内容のもつ論理性のもとに自己の判断を決定する調査対象の部類とか，質問の特定の語句あるいは全体のかもし出すイメージに反応する部類，質問系列の置き方で反応が異ってくるような言わばランダムな回答をする部類などが考えられるであろう．したがってこの仮定が成立するものであれば質問の反応の内容がこのような部類によって異っていることを質問系列の中で推定できる編成法をあらかじめ立案しておかねばならない．さらに質問の信頼性を結論するにはこの部類に対応する分析が必要になってくる．

ここでは上述の目的のために，処理法上の数学的な formulation を考えたい．一つの試みとして次ぎのようなモデルを設定する．まず個々の調査対象から回答を得る現象を確率現象と仮定し，その確率はある幾つかの類別された（調査対象の）部類のうちでは一定であるとする．類別には幾つの質問系によって内在的な関係によって定まる方法を考える．その試みとして latent class analysis の適用を考える．このようにして推定された部類別の確率から reliability の定義の一方法を示し，それが二つの条件のもとで，従来のパネル調査における回答の一致度と同値であることを導こう．

§1.　サンプリングの立場から

ここでは簡便のために二項選択の質問についてのみ考える．ある質問 Q について N 人の調査対象が考えられるとき，その要素たる回答がそれぞれ一定の確率でもって得られるものとする．このとき母集団は，質問に対する回答が「賛成」もしくは「反対」によって，

(1.1)
$$x_\alpha=\begin{cases}1\\0\end{cases}$$
$$\left.\begin{aligned}P\{x_\alpha=1\}&\equiv p_\alpha\\P\{x_\alpha=0\}&\equiv q_\alpha=1-p_\alpha\end{aligned}\right.\qquad(\alpha=1,2,\cdots\cdots,N)$$

な確率変数 x_α で表わされる標識の組と考えてよい．このときある全数調査における平均

(1.2)
$$\tilde{x}\equiv\frac{1}{N}\sum_{\alpha=1}^{N}x_\alpha$$

の期待値，分散は

(1.3)
$$\boldsymbol{E}(\tilde{x})=\frac{1}{N}\sum_{\alpha=1}^{N}p_\alpha\equiv\bar{p}$$

(1.4)
$$\boldsymbol{V}(\tilde{x})=\frac{1}{N^2}\sum_{\alpha=1}^{N}p_\alpha q_\alpha$$

で与えられる．

さてこの母集団から単純サンプリングによって n 個の標本 $\{x_{\alpha(1)}, x_{\alpha(2)}, \cdots, x_{\alpha(n)}\}$ を with replacement で抽出するときの分布を考えるとつぎのようである．標本値を値とする確率変数 $X_j(j=1,2,\cdots,n)$ を考え，$\delta_\alpha=\begin{cases}1\\0\end{cases}$ として

(1.5)
$$P\{X_j=\delta_\alpha|x_\alpha=\delta_\alpha\}=\frac{1}{N}$$

(1.6)
$$\begin{aligned}P\{X_j=1\}&=\sum_{\alpha=1}^{N}P\{X_j=1|x_\alpha=1\}P\{x_\alpha=1\}\\&=\frac{1}{N}\sum_{\alpha=1}^{N}p_\alpha\\&=\bar{p}\end{aligned}$$
$$P\{X_j=0\}=1-\bar{p}$$
$$(j=1,2,\cdots,n;\quad \alpha=1,2,\cdots,N)$$

したがって $\sum_{j=1}^{n}X_j$ は二項分布 $B(n,\bar{p})$ をもち，標本平均 $\bar{X}=\frac{1}{n}\sum_{j=1}^{n}X_j$ については，

(1.7)
$$\boldsymbol{E}\bar{X}=\bar{p}$$

(1.8)
$$\boldsymbol{V}\bar{X}=\frac{1}{n}\bar{p}(1-\bar{p})$$

が成立する．つまりこのモデルの過程では，質問 Q についての調査の結果は Q に賛成する調査対象のそれぞれの確率の算術平均 $\bar{p}$ を推定するものとなり，いわゆる固定した賛成の比率の推定値ではない．

さてこのような場合，質問 Q の回答の信頼性を問うならば，各調査対象について，その賛成する確率 p_α が時間変化に対して不変であるという仮定をおくとき，$\left|p_\alpha-\frac{1}{2}\right|$ をその一つの尺度をみることができよう．そこで母集団については，質問 Q の reliability を

Def 1

(1.9)
$$\begin{aligned}\text{rel }Q&\equiv\frac{4}{N}\sum_{\alpha=1}^{N}\left(p_\alpha-\frac{1}{2}\right)^2\\&=\frac{4}{N}\sum_{\alpha=1}^{N}\left(\frac{1}{4}-\boldsymbol{V}x_\alpha\right)\end{aligned}$$

によって定める．rel Q を推定するには，各 p_α の推定値を得ればよいが，同じサンプルに幾回もの調査を当てることは事実上不可能であるし，また可能であるとしても何らかの偏向を生じる要因が大きくなる．さらにまえがきにも述べたような諸問題を判別するにはこのままの形では解決されない．そこでこのモデルの上に潜在構造分析法を適用する．

§ 2 latent class analysis の適用

いまある質問系 $Q_i(i=1,2,\cdots,p)$ について § 1 で述べた母集団を調査する場合，Q_i に反応するメカニズムのパターンによって同一視された調査対象の類別が可能であるとして，それを

$$Y\in I_\alpha\qquad(\alpha=1,2,\cdots,m)$$

で表わす．サンプルとしてひとりの調査対象を抽出した場合，質問 Q_i に反応するパターンが部類 I_α に属す確率を

(2.1)
$$P\{Y\in I_\alpha\}\equiv\nu_\alpha\qquad(\alpha=1,2,\cdots,m)$$

で表わし，そのとき質問 Q_i に賛成する確率を

(2.2)
$$P\{X_i=1|Y\in I_\alpha\}\equiv\lambda_{i\alpha}\qquad(\alpha=1,2,\cdots,m;\quad i=1,2,\cdots,p)$$

で表わすことにする．つぎにある調査対象を抽出したとき，質問 Q_i に賛成する確率を π_i で表すとき，

$$\pi_i=\boldsymbol{E}X_i=\sum_{\alpha=1}^{m}\lambda_{i\alpha}\nu_\alpha \tag{2.3}$$

となり，この π_i は §1 における $\bar{p}$ に一致する．

この場合 Q_i の reliability は

Def 2.

$$\mathrm{rel}\,Q_i=4\,\boldsymbol{E}\left[\frac{1}{4}-\boldsymbol{E}\{X_i|Y\}\right]^2=4\sum_{\alpha=1}^{m}\left(\lambda_{i\alpha}-\frac{1}{2}\right)^2\nu_\alpha \qquad (i=1,2,\cdots p) \tag{2.4}$$

によって与えられ，Def 1. と一致する．

さて，latent parameter ν_α, $\lambda_{i\alpha}$ を推定するにあたっては，潜在構造分析における条件付独立の仮定が成立するものとする．つまり部類 $\{I_\alpha;\ \alpha=1,2,\cdots,m\}$ について homogeneous であって，

$$P\{X_1=1,\cdots,Xp=1|Y\in I_\alpha\}=\prod_{i=1}^{p}\lambda_{i\alpha} \tag{2.5}$$

等が成立するものとする．心理学的な内容でいえば，latent characteristic が決定されると behavior は統計的な意味で random であること，つまり同一の latent characteristic をもつ個人の集団は homogeneous とみなすわけである．われわれの場合には，あらかじめ質問系 $\{Q_i;\ i=1,2,\cdots,p\}$ を編成するとき，同じ部類 I_α に属すであろうと思われる調査対象には，$Q_i, Q_j (i\neq j)$ に対する“賛成”の確率が独立となるように設定しておかなければならない．実際的に latent class について homogeneous であることを判定する方法としては，モデルの験証について Lazarsfeld [4] の文献を参照されたい．

latent parameter ν_α, $\lambda_{i\alpha}$ を推定するには，Mc Hugh, R.B. [6] が maximum likelihood による推定を行っているが，その方法に従って実際に計算するには容易でないために，調査法に関する推定方式として Anderson, T.W. [1,2] による方法を選ぶことにしよう．

まず latent equation は (2.5) によって

$$\begin{aligned}
&\pi_i=\boldsymbol{E}X_i=\sum_{\alpha=1}^{m}\lambda_{i\alpha}\nu_\alpha\\
&\pi_{ij}=\boldsymbol{E}X_iX_j=\sum_{\alpha=1}^{m}\lambda_{i\alpha}\lambda_{j\alpha}\nu_\alpha \qquad (i\neq j)\\
&\pi_{ijk}=\boldsymbol{E}X_iX_jX_k=\sum_{\alpha=1}^{m}\lambda_{i\alpha}\lambda_{j\alpha}\lambda_{k\alpha}\nu_\alpha \qquad (i\neq j\neq k)
\end{aligned} \tag{2.6}$$

となる．この方程式を解くために，つぎのような行列を考える．

$$\Pi_{VH}\equiv\boldsymbol{E}\begin{pmatrix}1\\X_1\\\vdots\\X_{m-1}\end{pmatrix}(1X_m\cdots X_{2m-2}) \tag{2.7}$$

$$=\begin{pmatrix}1 & \pi_m & \pi_{m+1} & \cdots\cdots & \pi_{2m-2}\\ \pi_1 & \pi_{1m} & \pi_{1,\,m+1} & \cdots\cdots & \pi_{1,\,2m-2}\\ \vdots & & \cdots\cdots\cdots & & \\ \pi_{m-1} & \pi_{m-1,\,m} & \pi_{m-1,\,m+1} & \cdots\cdots & \pi_{m-1,\,2m-2}\end{pmatrix}$$

$$\Pi_{VHk}\equiv\boldsymbol{E}\begin{pmatrix}1\\X_1\\\vdots\\X_{m-1}\end{pmatrix}(1X_m\cdots X_{2m-2})X_k \tag{2.8}$$

(2.9)
$$\Lambda_V \equiv \begin{pmatrix} 1 & \cdots\cdots & 1 \\ \lambda_{11} & \cdots\cdots & \lambda_{1m} \\ & \cdots\cdots & \\ \lambda_{m-1,1} & \cdots\cdots & \lambda_{m-1,m} \end{pmatrix}$$

(2.10)
$$\Lambda_H \equiv \begin{pmatrix} 1 & \cdots\cdots & 1 \\ \lambda_{m,1} & \cdots\cdots & \lambda_{m,m} \\ & \cdots\cdots & \\ \lambda_{2m-2,1} & \cdots\cdots & \lambda_{2m-2,m} \end{pmatrix}$$

(2.11)
$$N \equiv \begin{pmatrix} \nu_1 & & & \\ & \nu_2 & & 0 \\ & & \ddots & \\ 0 & & & \nu_m \end{pmatrix}$$

(2.12)
$$\Delta_k \equiv \begin{pmatrix} \lambda_{k1} & & & \\ & \lambda_{k2} & & 0 \\ & & \ddots & \\ 0 & & & \lambda_{km} \end{pmatrix}$$

このとき

(2.13) $$\Pi_{VH} = \Lambda_V N \Lambda_H'$$

(2.14) $$\Pi_{VHk} = \Lambda_V \Delta_k N \Lambda'_H \qquad (k > 2m-2)$$

が成立する．(2.6) を解くには (2.13), (2.14) を解けば十分である．その方法はつぎのように述べられる．

(2.13), (2.14) において $|\Lambda_V|, |\Lambda_H|, |N| \neq 0$ で，かつ k を固定したとき Δ_k の対角元 $\lambda_{k\alpha}(\alpha=1, \cdots, m)$ がすべて異る値であれば

(2.15) $$\Pi_{VHk}\, z = \lambda_{k\alpha} \Pi_{VH}\, z$$

を満足するベクトル z_α が存在する．このとき

(2.16) $$Z = (z_1 z_2 \cdots z_m)$$

と置くとき，Λ_V の元は次のようにして求められる．すなわち，

(2.17) $$(\Lambda_V \text{ の } \alpha \text{ 列}) = c_\alpha^{(VH)} \times (\Pi_{VH} Z \text{ の } \alpha \text{ 列})$$

ここで $c_\alpha^{(VH)}$ は

(2.18) $$c_\alpha^{(VH)} = \frac{1}{\{\Pi_{VH} Z \text{ の } \alpha \text{ 列の第1行の元}\}}$$

として与えられる．

証明はつぎのようにして示される．

(2.19) $$|\Pi_{VHk} - \theta \Pi_{VH}| = 0$$

の根は Δ_k の対角元である．何故なら

(2.20)
$$\begin{aligned} 0 &= |\Pi_{VHk} - \theta \Pi_{VH}| \\ &= |\Lambda_V \Delta_k N \Lambda_H' - \theta \Lambda_V N \Lambda_H'| \\ &= |\Lambda_V| \cdot |\Delta_k - \theta I| \cdot |N| \cdot |\Lambda_H'| \end{aligned}$$

となるからである．したがって $\lambda_{k\alpha}$ がすべて異ることから定数比を無視すれば，$\lambda_{k\alpha}$ を固有値とする (2.15) を満足する固有ベクトル z_α が決定される．これを

(2.21) $$\Pi_{VHk} Z = \Pi_{VH} Z \Delta_k$$

と表わせば，

$$\Pi_{VHk} \Pi_{VH}^{-1} (\Pi_{VH} Z) = (\Pi_{VH} Z) \Delta_k$$

となり，これに (2.13)，(2.14) を代入して

(2.22) $$\Lambda_V \Delta_k \Lambda_V^{-1}(\Pi_{VH}Z)=(\Pi_{VH}Z)\Delta_k$$

を得る．

(2.23) $$W=\Lambda_V \Delta_k \Lambda_V^{-1}$$

と置けば，(2.22) は $\Pi_{VH}Z$ の第 α 列ベクトルは Δ_k の $\lambda_{k\alpha}$ を固有値とする W の固有解に比例することを意味する．一方 (2.23) は Λ_V の第 α 列ベクトルは $\lambda_{k\alpha}$ を固有値とする W の固有解に比例していることを示す．したがって Λ_V の各列は対応する $\Pi_{VH}Z$ の各列に比例し，Λ_V の各列の第1行の元は1であることから (2.17)，(2.18) が成立する．($q.e.d.$)

定義 (2.7)～(2.14) において Λ_V と Λ_H をとりかえるならば（したがって (2.13) では Π_{HV}，(2.14) では Π_{HVk} となる)，Λ_H の解を与える．かくして (2.13)，(2.14) より N および Δ_k が求められる．

Π_{VH}，Π_{VHk} の各元の代りに，調査からの推定値を代入するときは，ν_α，$\lambda_{i\alpha}$ の推定量は consistent である．Anderson T.W. [2] によれば，それらの推定量は asymptotically normal であるが，efficient ではなく，また admissible でもないことが報告されている．すなわち latent parameter は0と1の間になくてはならないが，必ずしもそのように求められず，したがってその場合は Λ_V，Λ_H の作り方を工夫しなければならない．

かくして各部類 $I_\alpha(\alpha=1,2,\cdots,m)$ についての latent parameter $\nu_\alpha, \lambda_{i\alpha}$ が推定されると，まずその部類における質問系 $\{Q_i\}$ の reliability が求められる．つまり

(2.24) $$\mathrm{rel}(Q_i|Y\in I_\alpha)=4\left(\lambda_{i\alpha}-\frac{1}{2}\right)^2$$
$$(i=1,2,\cdots,p;\ \alpha=1,2,\cdots,m)$$

となる．各部類についての reliability の平均は (2.4) で与えられる．このように質問系の信頼性が調査対象の類別された部類に応じて考察することが可能になり，まえがきに述べた提案の一つがこの意味で解決されることになる．

§ 3. パネル調査における一致度

社会調査においては，ある時点に調査された調査対象を二度同一の質問で再調査する場合が多く見受けられる．その際調査対象が前回の回答と同じものを繰返すかどうか，つまり回答の一致度を験かめることが一つの大きな問題となる．この二つの調査期間内に質問に対する意見の変動が生じると，後期の調査において当然回答も前回時点のときと異ってくるわけで調査の目的にそぐわないものとなる．したがって一致度をみる調査では，二つの調査時点間の意見変動が起らないという前提のもとに実施されなければならないが，この仮定を験証する方法の一つとして § 3 に述べた latent class の分布を比較することが考えられる．つまり ν_α の前期の調査における推定値と後期調査における推定値との間に有意差があれば，意見変動が生じなかったという仮説は棄却されねばならない．

ここでは意見変動が生じないう仮定のもとで一致度を考えることにする．つまり調査対象の部類 I_α に属する分布は不変であると仮定する．しかるに特に短期間においてパネル調査を実施するときは，意見変動は生じなくても調査対象が前回の自身の回答を記憶していることも考えられ，一般には前後の調査で部類 I_α を固定したときの質問に反応する確率に相関があるとみてよい．そこで § 3 の latent class model をパネル調査に対してつぎのように拡張する．

第1回の調査において，質問 Q_i について，ある調査対象がそれに賛成もしくは反対するという事象を表わす確率変数を $X_i^{(1)}=1,0$；第2回調査における確率変数を $X_i^{(2)}=1,0$ で表わす．その分布は，

(3.1) $$P_r\{X_i^{(1)}=\delta_1;\ X_i^{(2)}=\delta_2\}\rightleftarrows\pi_i^{(\delta_1\delta_2)}\qquad \delta_1,\delta_2=1 \text{ or } 0$$
$$P_r\{X_i^{(1)}=\delta_1, X_j^{(1)}=\delta_1;\ X_i^{(2)}=\delta_2, X_j^{(2)}=\delta_2\}\rightleftarrows\pi_{ij}^{(\delta_1\delta_2)}$$

等々．つぎに latent class を固定したときの条件付確率を

(3.2) $$P_r\{X_i^{(1)}=\delta_1, X_i^{(2)}=\delta_2|Y\}\rightleftarrows\pi_i^{(\delta_1\delta_2)}(Y)$$

で表わし，$Y\epsilon I_\alpha$ のとき

$$\pi_i^{(\delta_1\delta_2)}(Y\epsilon I_\alpha)\rightleftarrows\lambda_{i\alpha}^{(\delta_1\delta_2)}$$

なる値をとるものとする．この場合にも条件付独立を仮定する．つまり，

(3.3) $$P_r\{X_i^{(1)}=\delta_1, X_j^{(1)}=\delta_1;\ X_i^{(2)}=\delta_2, X_j^{(2)}=\delta_2|Y\epsilon I_\alpha\}=\lambda_i^{(\delta_1\delta_2)}\cdot\lambda_j^{(\delta_1\delta_2)}$$

かくして latent equation

(3.4) $$\pi_i^{(\delta_1\delta_2)}=\sum_{\alpha=1}^{m}\lambda_{i\alpha}^{(\delta_1\delta_2)}\nu_\alpha$$
$$\pi_{ij}^{(\delta_1\delta_2)}=\sum_{\alpha=1}^{m}\lambda_{i\alpha}^{(\delta_1\delta_2)}\lambda_{j\alpha}^{(\delta_1\delta_2)}\nu_\alpha\qquad(i\neq j)$$
$$\pi_{ijk}^{(\delta_1\delta_2)}=\sum_{\alpha=1}^{m}\lambda_{i\alpha}^{(\delta_1\delta_2)}\lambda_{j\alpha}^{(\delta_1\delta_2)}\lambda_{k\alpha}^{(\delta_1\delta_2)}\qquad(i\neq j\neq k)$$

が成立する．latent parameter の推定は § 2 の場合とまったく同様である．

第１回調査における条件付確率は，

(3.5) $$\lambda_{i\alpha}^{(1)}=\lambda_{i\alpha}^{(11)}+\lambda_{i\alpha}^{(10)}$$
$$1-\lambda_{i\alpha}^{(1)}=\lambda_{i\alpha}^{(01)}+\lambda_{i\alpha}^{(00)}$$

第２回調査における条件は確率は

(3.6) $$\lambda_{i\alpha}^{(2)}=\lambda_{i\alpha}^{(1,1)}+\lambda_{i\alpha}^{(0,1)}$$
$$1-\lambda_{i\alpha}^{(2)}=\lambda_{i\alpha}^{(1,0)}+\lambda_{i\alpha}^{(00)}$$

で与えられる．また条件付共分散は

(3.7) $$\boldsymbol{E}[\{X_i^{(1)}-\boldsymbol{E}(X_i^{(1)}|Y\epsilon I_\alpha)\}\{X_i^{(2)}-\boldsymbol{E}(X_i^{(2)}|Y\epsilon I_\alpha)\}|Y\epsilon I_\alpha]=\lambda_{i\alpha}^{(11)}-\lambda_{i\alpha}^{(1)}\cdot\lambda_{i\alpha}^{(2)}$$
$$=\lambda_{i\alpha}^{(11)}\cdot\lambda_{i\alpha}^{(00)}-\lambda_{i\alpha}^{(10)}\cdot\lambda_{i\alpha}^{(01)}$$

となり，これより調査の前後における回答の相関が求められる．

さて，通常パネル調査における回答の一致度は

Def. 3

(3.8) $$\mathrm{con}(Q_i)\leftrightarrows\frac{\pi_i^{(11)}+\pi_i^{(00)}-\dfrac{1}{2}}{\dfrac{1}{2}}$$

で定義されている．比較される数値 1/2 は「回答が random になされる確率」ということを意味する．この場合 latent class model の立場から考えると (3.4) より，

(3.9) $$\mathrm{con}(Q_i)=2\sum_{\alpha=1}^{m}\left(\lambda_{i\alpha}^{(11)}+\lambda_{i\alpha}^{(00)}-\frac{1}{2}\right)\nu_\alpha$$

となる．したがって latent class の分布 ν_α を固定して考えると，各 latent class における $\lambda_{i\alpha}^{(11)}+\lambda_{i\alpha}^{(00)}-\dfrac{1}{2}$ が大であるほど，質問 Q_i の一致度は高いといえる．いま

(3.10) $$\mathrm{con}(Q_{i\alpha})\leftrightarrows 2\left(\lambda_{i\alpha}^{(11)}+\lambda_{i\alpha}^{(00)}-\frac{1}{2}\right)$$

(3.11) $$\mathrm{rel}_{(l)}(Q_{i\alpha})\leftrightarrows 4\left(\lambda_{i\alpha}^{(l)}-\frac{1}{2}\right)^2\qquad(l=1,2)$$

と置く．(3.11) は第 α-class に属する調査対象の質問 Q_i に対する reliabllity であってこれと $\mathrm{con}(Q_{i\alpha})$ との関係を求めてみよう．

latent class を固定するとき Q_i に賛成する確率は前後の調査において不変であると仮定する．つまり周辺分布を意味する $\lambda_{i\alpha}^{(1)}$ と $\lambda_{i\alpha}^{(2)}$ と相等しいと置くわけである． このとき (3.5), (3.6) より，

(3.12) $$\lambda_{i\alpha}^{(10)}=\lambda_{i\alpha}^{(01)}$$

が成立する．

さらに，latent class を固定したときの前後の調査における Q_i に賛成する確率が独立であると仮定する．このときは (3.7) より

(3.13) $$\lambda_{i\alpha}^{(11)}\cdot\lambda_{i\alpha}^{(00)}-\lambda_{i\alpha}^{(10)}\cdot\lambda_{i\alpha}^{(01)}=0$$

が成立する．さて，

(3.14) $$\lambda_{i\alpha}^{(11)}+\lambda_{i\alpha}^{(10)}+\lambda_{i\alpha}^{(01)}+\lambda_{i\alpha}^{(00)}=1$$

であるから，これら3式より

(3.15) $$\lambda_{i\alpha}^{(11)}\lambda_{i\alpha}^{(00)}=(\lambda_{i\alpha}^{(10)})^2$$

(3.16) $$\sqrt{\lambda_{i\alpha}^{(11)}}+\sqrt{\lambda_{i\alpha}^{(00)}}=1$$

がしたがう．

$\mathrm{rel}(Q_{i\alpha})$ を $\lambda_{i\alpha}^{(11)}$ のみで表わすと

(3.17) $$\mathrm{con}(Q_{i\alpha})=1-4(\sqrt{\lambda_{i\alpha}^{(11)}}-\lambda_{i\alpha}^{(11)})$$

となる．

一方 (3.15), (3.16) より

$$\lambda_{i\alpha}^{(1)}=\lambda_{i\alpha}^{(2)}=\lambda_{i\alpha}^{(11)}+\lambda_{i\alpha}^{(10)}$$
$$=\sqrt{\lambda_{i\alpha}^{(11)}}$$

であるから

(3.18) $$\mathrm{rel}_{(l)}(Q_{i\alpha})=4\left(\sqrt{\lambda_{i\alpha}^{(11)}}-\frac{1}{2}\right)^2 \qquad (l=1,2)$$

を得る．したがって (3.17) を併せ考えると

(3.19) $$\mathrm{con}(Q_{i\alpha})=\mathrm{rel}_{(l)}(Q_{i\alpha}) \qquad (l=1,2)$$

が従う．ところで，

(3.20) $$\mathrm{con}(Q_i)=\sum_{i=1}^{m}\mathrm{con}(Q_{i\alpha})\nu_\alpha$$

(3.21) $$\mathrm{rel}_{(l)}(Q_i)=\sum_{i=1}^{m}\mathrm{rel}_{(l)}(Q_{i\alpha})\nu_\alpha \qquad (l=1,2)$$

であるから，(3.19) より

(3.22) $$\mathrm{con}(Q_i)=\mathrm{rel}_{(l)}(Q_i) \qquad (l=1,2)$$

が結論される．以上まとめると次ぎのようになる．

Th. 1. latent class を固定するとき調査対象の質問 Q_i に賛成する確率が前後の調査において不変でありかつ独立であるならば，質問 Q_i に対する一致度 Def. 3 と reliability. の Def. 2 とは相等しい．

この結果は二つの仮定の何れが欠けても成立しない．周辺分布の一致をはずすと，例えば

$$\alpha_{i\alpha}^{(11)}=\alpha_{i\alpha}^{(00)}=\alpha_{i\alpha}^{(01)}=0,\ \ \alpha_{i\alpha}^{(10)}=1$$

なるとき，$\mathrm{rel}_{(l)}(Q_{i\alpha})$ は $l=1,2$ について 1 であるが，$\mathrm{con}(Q_{i\alpha})=0$ となる． また独立性の仮定をはずすと

$$\lambda_{i\alpha}^{(11)}=\lambda_{i\alpha}^{(00)}=\frac{1}{2}, \qquad \lambda_{i\alpha}^{(10)}=\lambda_{i\alpha}^{(01)}=0$$

なる場合には，$\mathrm{rel}_{(t)}(Q_{i\alpha})=0$ であるが，$\mathrm{con}(Q_{i\alpha})=1$ である．この独立性の仮定は前にも述べたように，調査対象が自身の前回の回答を記憶していれば成立しないわけで，パネル調査で最も注意しなければならない点であろう．

（統計数理研究所）

文　　献

［1］ Anderson, T.W.: On estimation of parameters in latent structure analysis, Psychometeika, 19, 1954, p.p. 1～10.

［2］ Anderson, T.W.: Some Scaling models and estimation procedures in the latent class model, Probability and Statistics, The Harald Cramér Volume, p.p. 9～49, 1959, John Wiley.

［3］ 林知己夫： 態度数量化の一方法 II，統計数理研究所彙報，第6巻，第1号，1958.

［4］ Lazarsfeld, P.F.: The logical and mathematical foundation of latent structure analysis, in Chap. 10, Mesurment and Prediction, Priceton, New Jursey Princeton Univ. Press, 1950.

［5］ Mc Hugh, R.B.: Efficient estimation and local identification in latent class analysis. Psychometrika, 21, 1956, pp. 331～347.

［6］ Meredith, W.: Some results based on a general stochastic model for mental tests, Psychometrika, Vol. 30, 1965. p.p. 419～440.