

○日立製作所 正 員 岡田 明
 北海道大学 正 員 山形 耕一
 北海道大学 正 員 佐藤 馨一

1. はじめに

交通調査の実施に際しては、多くの種類の誤差が、付随して発生することが経験的に知られている。これらの誤差は、多くの場合推定値の偏りをもたらすが、実際のデータ利用に際しては、無作為抽出誤差にもとづく推定値の分散のみを考慮している場合が多く、実査過程において発生する偏り誤差の補正に配慮している例は、少ない。本研究の目的は、交通調査の設計実施の各段階における種々の誤差の発生過程を吟味し、これらを複合化してデータ誤差を算出することにより、調査結果の利用者がデータを扱う上でその信頼性を評価し得るモデルを提案することにある。

2. 確信度分析

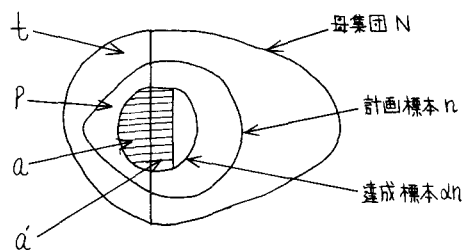
大きさ N の母集団においてある特定の属性をもつ要素の部分集団の比率 τ が調査対象であるものとする。サンプル調査により調査を行うものとし、計画標本の大きさ n とする。実査過程においては調査不能や無回答拒否が生じるため、調査を達成した標本の規模は an (a : 回収率, $0 \leq a \leq 1$) となる。母集団における未知の母比率は τ であるが、計画標本および達成標本が着目した属性をもつ要素を含む比率 P および a は一般に τ には一致しない。前者の食い違いはフレームの不適切さや標本抽出の偶然性に、後者は無回答層と回答層では一般に着目する要素を含む割合が異なることに起因する。さらに、達成標本において真の比率が a であるとしても、意識を調べる調査等では観測の誤差のため、測定結果としての比率は a とは異なり a' となる。しかしながら多くの調査では a' をそのまま真の母比率 τ として用いている。

本研究では、誤差の発生を上記の各段階に分解して詳しく観測し、それらの誤差を合成することにより対象とする統計量の推定誤差をより客観的に評価するモデルを提案するものである。すなわち、測定値 a' の調査対象の真の値に対する比率を次式の如く乗法的に分解する。

$$\frac{a'}{\tau} = \frac{a'}{a} \times \frac{a}{P} \times \frac{P}{\tau} \quad (\text{記号は図1参照})$$

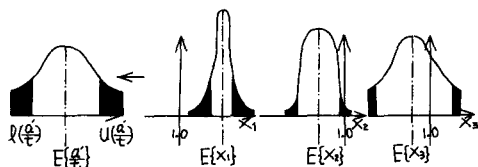
ここで、 a'/a , a/P , P/τ はそれぞれ、測定時、調査不能、標本抽出時における誤差の発生に対応している。これらの比率は、例えば測定時では、実査を通じての測定結果 a' と測定対象の真の値 a の比であり、測定値比と呼ぶことにする。測定値比を調査から直接知ることは不可能であり、より詳しい補足調査や実査過程をつづさに観察した評価者の経験的判断により定めなければならない。本研究では、測定値比を点推定的に定めるのではなく補足調査や実査過程の吟味にもとづく評価者のもつ測定値比の実現に対する主観的確率分布として表し、これを確信度分布と呼ぶことにする。

確信度分析は、調査対象の統計量の誤差を上式の如く、調査の各段階に対応して分解し、各段階における確信



a' : 測定比率 P : 計画標本の真の比率
 a : 達成標本の真の比率 τ : 母集団の真の比率

図1 調査段階別の真の比率と測定値の関係



($\frac{a'}{\tau}$)の確信度分布 ← 測定値の確信度分布

図2 合成確信度分布

度分布を求める分解過程と、それぞれの確信度分布を合成して α' および α の確信度分布を求める合成過程とからなる。分解過程は本研究では3段階に分けたが、それぞれの調査の性格により、誤差の発生過程に対応して評価者が最も適切と考える段階に分解することが重要である。合成過程では $(\frac{\alpha'}{\alpha})$ の合成確信度分布の期待値と分散を定めることが必要である。本研究では確信度分布の関数型は正規分布と考え、合成確信度分布の信頼上限および下限を2 σ の域としている。上述の3段階における測定値比を X_1, X_2, X_3 とし、それぞれの確信度分布における期待値、分散を $E\{X_i\}, V\{X_i\}$ で表わすと、合成確信度分布および調査対象値では、

$$\begin{aligned} \text{合成確信度分布；期待値 } E\{\frac{\alpha'}{\alpha}\} &= E\{X_1\} \times E\{X_2\} \times E\{X_3\} \quad \text{調査対象値 } t; \text{期待値 } E\{t\} = \alpha' / E\{\frac{\alpha'}{\alpha}\} \\ \text{；信頼下限 } l(\frac{\alpha'}{\alpha}) &= E\{\frac{\alpha'}{\alpha}\} - 2\sqrt{V\{\frac{\alpha'}{\alpha}\}} \quad u(\frac{\alpha'}{\alpha}) = E\{\frac{\alpha'}{\alpha}\} + 2\sqrt{V\{\frac{\alpha'}{\alpha}\}}; \quad l(t) = \alpha' / u(\frac{\alpha'}{\alpha}) \\ \text{；分散 } V\{\frac{\alpha'}{\alpha}\} &= \sum_{i=1}^3 V\{X_i\} / E\{X_i\}^2; \quad u(t) = \alpha' / l(\frac{\alpha'}{\alpha}) \end{aligned}$$

となる。調査対象値 t の合成された確信度分布のもつ意味は、「評価者は $E\{t\}$ という値が最も真の値に近いと見ているし少なくとも区間 $(l(t), u(t))$ に真の値があると確信している。」と理解される。

3. 意識調査への適用例と考察

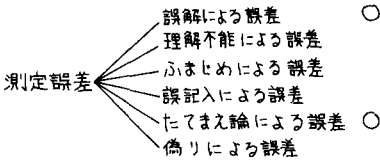
本研究では、昭和55年12月に実施された航空機と鉄道間のモード選択に関する意識調査について確信度分析を適用した。この調査では、石勝線等の開通により鉄道側のサービス水準がアップした場合、札幌と釧路・帯広間で航空機と鉄道のいずれを使用するかをその理由と共に調査した。実験計画法を使用した自記式アンケート調査である。

この調査における主な誤差発生要因は図3に示した誤差要因に基づく、測定誤差であると考えられ、測定誤差の除去が主要な目的となった。測定誤差 α'/α の確信度分布を求めるため、アンケート調査の対象者の約5%（80名）を再抽出し、面接調査を行なった。面接調査では、質問文に対する回答者の解釈および回答内容、日頃の交通機関選択の実態とその理由等を聞き取り調査した。その結果調査上の概念の伝達が不十分であり、質問の意図を誤解したことによって生じた誤差と国鉄赤字線廃止問題の渦中にあって国鉄を実際には利用しないにもかかわらず利用すると答えるためえ論的な回答による誤差が生じていることが判明した。本調査の結果は分散分析の入力データとして使用されたが、表1は回答に確信度分析を施したデータを入力した場合と回答をそのまま集計したデータを入力した場合の分散分析の結果を対比したものである。前者の場合において誤差項の変動が極めて小さくなり、ているし、各要因の寄与率の大きさおよび寄与率の構造も変り、ている。従来、分散分析では誤差項の大きさおよび寄与率の構造の妥当性により調査の成否が判断されてきた。すなわち誤差項の大きさに換返し誤差と要因設定の不適切さが反映されると考えこおり、調査自体の測定誤差は極く小さいと考えられていた。しかしながら、意識調査等における調査では、回答自体に誤差が生じることが予想されるし、事実本例では確信度分析を行なった場合と行なわなない場合に明らかな相違を示している。

4. 結 論

確信度分析では推定誤差とくに偏りの発生に対する評価者の主観的判断を誤差の各発生段階毎に明示的に表すことができる。それゆえ、推定値の信頼性に関して評価者間における見解の相違を具体的に指摘し合うことができ、るため、より真の値に近い推定値を求めたりくことを可能とするアプローチである。

またアンケート法を用いた交通調査等では、測定誤差の発生に常に細心の注意を払うことおよび測定誤差の発生傾向を把握する補足調査を同時に実施することの重要性を示すとともに、これらの誤差を補正する方法として確信度分析が有用であることが示された。



○は今回の調査で目立、た誤差
図3 測定誤差（アンケート調査）

表1 分散分析表（累積法による）

要因	確信度分析前				確信度分析後			
	変動	df	F	寄与率	変動	df	F	寄与率
A	2225.3	6	6.6	29%	1977	6	58.06	208%
B	3448.7	2	30.7	512%	7046	2	620.78	75.0%
C	175.0	2	1.6	1%	7.9	2	0.694	
D	553.3	2	4.9	6.8%	344	2	30.307	3.6%
e	112.5	2		12%	11.4	2		0.6%
計	6514.9	14		100%	9389.3	14		100%