

ランダム化器具を用いないランダム回答法の 実験的調査結果について*

統計数理研究所 逆瀬川 浩 孝
" 高橋 宏 一
" 鈴木 達 三

(1977年1月 受付)

An investigation of a new randomized response model
Hirotaka Sakasegawa, Koiti Takahasi and Tatsuzo Suzuki
(The Institute of Statistical Mathematics)

The purpose of this validation study was to examine the possibility of a new randomized response technique proposed by Takahasi and Sakasegawa [5]. This technique uses some auxiliary questions instead of randomizing devices, and the procedure of this model may be somewhat complex.

The respondents are college students enrolled at such classes as statistics or survey research in several universities located at Tokyo and Okayama.

The respondents numbered about 900, and they were divided into three groups. The study requires these groups to eliminate the estimation bias caused by the relationship between auxiliary and main questions.

The respondents were asked ten items, which were divided into two halves. The one half served as validation study. For these questions it was possible to obtain precise information about joint response patterns of both auxiliary and main questions. The other half consisted of those items which were evasive or looked upon as morally condemnable by society.

The results indicated that this model was successfully applicable for well understanding respondents. But in case of general sample, it will be suspected that several percent of the sample who use this model will say it is confusing or troublesome.

The purpose of this study and a brief introduction of the method used are presented in Section 1. The main findings are shown in Section 2. In Section 3 further considerations and discussions are given. Optimal designs for models in particular cases are discussed in Section 4. Participant's views are presented in Section 5. Concluding note and some comments are given in Section 6.

§1. 目 的

高橋・逆瀬川 [5] はランダム化器具を用いないランダム回答法の一つのモデルを提案した。それは、ある母集団において特性 A をもつ人の割合 π を推定したいときに、その母集団から独立に3組の標本を抽出し、3肢 (C_1, C_2, C_3) の中から1つを選択する補助質問を用いて、一つの組では C_1A および $\bar{C}_1\bar{A}$ の人に、他の一つの組では C_2A および $\bar{C}_2\bar{A}$ の人に、残りの一つの組では C_3A および $\bar{C}_3\bar{A}$ の人に '1' と回答してもらい、それ以外の人には '2' と回答してもらい、その結果から π を推定しようとするものである。ここで C_iA ($i=1, 2, 3$) は、補助質問に対する選択が C_i であり、かつ特性 A をもつことを表し、 $\bar{C}_i\bar{A}$ は、補助質問に対する

*) この研究の一部は、昭和51年度特別研究「多面情報に基づく統計的推論」および文部省科学研究費補助金 試験研究 (2) 183001 「社会調査の標準化過程における回答誤差の研究」によるものである。

選択が C_i 以外であり、かつ特性 A をもたないことを表している。第 i 組の標本の大きさを n_i 、その中で '2' と回答した人数を y_i とすると、 π の推定量 $\hat{\pi}$ は

$$\hat{\pi} = \frac{y_1}{n_1} + \frac{y_2}{n_2} + \frac{y_3}{n_3} - 1 \quad (1)$$

で与えられる。このモデルの主要な特徴は

(i) 補助質問に対する母集団における C_1, C_2, C_3 の選択確率を p_1, p_2, p_3 とするとき、 $p_1 = p_2 = p_3 = 1/3$ は仮定していないし、 p_1, p_2, p_3 は未知であってよい；

(ii) 補助質問に対する選択と特性 A をもつかもたないかということの独立性は仮定していない；

という2つのことである。

一方このモデルの基本的な仮定は

(i) 3組の標本は独立であり、各組の標本はそれぞれ母集団からの無作為標本である；

(ii) $C_i A$ や $C_i \bar{A}$ ($i=1, 2, 3$) の割合の期待値は各組に共通である (後に詳述)；

ということである。

これらの基本的な仮定が実際場面においてどうなるかを調べるために大学生を対象にした実験的な調査をした。大学生を対象にした理由の一つは、調査実施面の制約にもよるが、著者らがこれまで、たびたび主張してきたように、ランダム回答法は統計学の教材として適当なものであると考えられ、学生の興味がどの程度であるかを知りたいということにある。もう一つの理由は、将来実際の目的のためにランダム回答法が応用されるとしたら、それはランダム回答法の意義を十分に理解できる対象に限るのが正当であるという立場から、学生に制限したことにある。

調査票は10問からなっており、最初の5問はいずれも A か \bar{A} かを直接きくことのできるものであり、ランダム回答法による回答のほか、補助質問で何を選択したかということと、 A であるか \bar{A} であるかということも記入してもらい、基本的な仮定の成立、非成立に関する情報と、ランダム回答法の手順が正しく行われるかどうかについての情報が得られるようにした。残りの5問はそれほど深刻ではないが、直接答えることにある程度の抵抗を感じるものであり、これらについてはランダム回答法による回答のみを記入してもらっている。これらの5問は、結果として得られる数値が妥当であるかどうかの検討にも使われるが、このような直接的に答えにくい質問にランダム回答法を用いた上で、回答者のランダム回答法に対する感想、意見を知るためにとり入れられている。なお、この調査は原則として氏名を記入することになっており、その点は通常の調査と同じ条件である。

以下に質問事項の全体を示す：

Q1. 桜、梅、桃のうち一つを心の中できめて下さい。きめたら、次の質問を読み、リストに従って1か2を答えて下さい。

質問：「あなたは早生まれですか？」

リスト：

“早生まれの人”で、きめたのが $\left\{ \begin{array}{l} \text{桜の人は} 2 [2, 1] \\ \text{梅の人は} 1 [2, 2] \\ \text{桃の人は} 2 [1, 2] \end{array} \right.$ と答えて下さい。

“早生まれでない人”で、きめたのが $\left\{ \begin{array}{l} \text{桜の人は} 1 [1, 2] \\ \text{梅の人は} 2 [1, 1] \\ \text{桃の人は} 1 [2, 1] \end{array} \right.$ と答えて下さい。

答 (1か2を記入)

きめたのは何でしたか? (桜, 梅, 桃を記入)

あなたの生まれたのは何月ですか? 月

Q2. ライオン, トラ, ヒョウのうち一つを心の中できめて下さい。きめたら次の質問を読み, リストに従って1か2を答えて下さい。

質問: 「あなたには男の兄弟がいますか?」

リスト:

“いる人” で, きめたのが $\begin{cases} \text{ライオンの人は} & 2 [2, 1] \\ \text{トラの人は} & 1 [2, 2] \\ \text{ヒョウの人は} & 2 [1, 2] \end{cases}$ と答えて下さい。

“いない人” で, きめたのが $\begin{cases} \text{ライオンの人は} & 1 [1, 2] \\ \text{トラの人は} & 2 [1, 1] \\ \text{ヒョウの人は} & 1 [2, 1] \end{cases}$ と答えて下さい。

答 (1か2を記入)

きめたのは何でしたか? (ライオン, トラ, ヒョウを記入)

あなたには男の兄弟がいますか? (いる, いないを記入)

Q3. 赤, 青, 黄のうち一つを心の中できめて下さい。きめたら次の質問を読み, リストに従って答えて下さい。

質問: 「あなたは北海道に行ったことがありますか?」

リスト:

“行ったことがある人” で, きめたのが $\begin{cases} \text{赤の人は} & 1 [2, 2] \\ \text{青の人は} & 2 [1, 2] \\ \text{黄の人は} & 2 [2, 1] \end{cases}$ と答えて下さい。

“行ったことがない人” で, きめたのが $\begin{cases} \text{赤の人は} & 2 [1, 1] \\ \text{青の人は} & 1 [2, 1] \\ \text{黄の人は} & 1 [1, 2] \end{cases}$ と答えて下さい。

答 (1か2を記入)

きめたのは何でしたか? (赤, 青, 黄を記入)

北海道に行ったことがありますか? (ある, ないを記入)

Q4. イギリス, フランス, ドイツのうち一つを心の中できめて下さい。きめたら次の質問を読み, リストに従って答えて下さい。

質問: 「あなたの朝食はパン食の方が多いですか?」

リスト:

“パン食の方が多い人” で, きめたのが $\begin{cases} \text{イギリスの人は} & 1 [2, 2] \\ \text{フランスの人は} & 2 [1, 2] \\ \text{ドイツの人は} & 2 [2, 1] \end{cases}$ と答えて下さい。

“パン食の方が少ない人” で, きめたのが $\begin{cases} \text{イギリスの人は} & 2 [1, 1] \\ \text{フランスの人は} & 1 [2, 1] \\ \text{ドイツの人は} & 1 [1, 2] \end{cases}$ と答えて下さい。

答 (1か2を記入)

きめたのは何ですか? (イギリス, フランス, ドイツを記入)

朝食はパン食の方が多いですか? (多い, 少ないを記入)

Q5. リンゴ, ナシ, ミカンのうち一つを心の中できめて下さい。きめたら次の質問を読み, リストに従って答えて下さい。

質問: 「地球以外の星に生物がいますか?」

リスト:

“いると思う人” で, きめたのが $\begin{cases} \text{リンゴの人は} & 2 [1, 2] \\ \text{ナシの人は} & 2 [2, 1] \\ \text{ミカンの人は} & 1 [2, 2] \end{cases}$ と答えて下さい。

“いないと思う人”で、きめたのが $\left\{ \begin{array}{l} \text{リンゴの人は} 1 [2, 1] \\ \text{ナシの人は} 1 [1, 2] \\ \text{ミカンの人は} 2 [1, 1] \end{array} \right.$ と答えて下さい。

答 (1か2を記入)

きめたのは何ですか? (リンゴ, ナシ, ミカンを記入)

地球以外に生物がいると思いますか? (思う, 思わないを記入)

(これからあとの質問は1か2だけを答えて下さい。)

Q 6. 松, 竹, 梅のうち一つを心の中できめて下さい。次の質問を読み、リストに従って答えて下さい。
質問: 「あなたはキセル乗車をしたことがありますか?」

“したことがある人”で、きめたのが $\left\{ \begin{array}{l} \text{松の人は} 2 [1, 2] \\ \text{竹の人は} 2 [2, 1] \\ \text{梅の人は} 1 [2, 2] \end{array} \right.$ と答えて下さい。

“したことがない人”で、きめたのが $\left\{ \begin{array}{l} \text{松の人は} 1 [2, 1] \\ \text{竹の人は} 1 [1, 2] \\ \text{梅の人は} 2 [1, 1] \end{array} \right.$ と答えて下さい。

答 (1か2を記入)

Q 7. 海, 山, 川のうち一つを心の中できめて下さい。次の質問を読み、リストに従って答えて下さい。
質問: 「あなたは現在交際中の恋人がいますか?」

リスト:

“いる人”で、きめたのが $\left\{ \begin{array}{l} \text{海の人は} 2 [1, 2] \\ \text{山の人は} 2 [2, 1] \\ \text{川の人は} 1 [2, 2] \end{array} \right.$ と答えて下さい。

“いない人”で、きめたのが $\left\{ \begin{array}{l} \text{海の人は} 1 [2, 1] \\ \text{山の人は} 1 [1, 2] \\ \text{川の人は} 2 [1, 1] \end{array} \right.$ と答えて下さい。

答 (1か2を記入)

Q 8. A, B, C のうち一つを心の中できめて下さい。次の質問を読み、リストに従って答えて下さい。
質問: 「あなたは高校時代にカンニングをしたことがありますか?」

リスト:

“したことがある人”で、きめたのが $\left\{ \begin{array}{l} \text{Aの人は} 1 [2, 2] \\ \text{Bの人は} 2 [1, 2] \\ \text{Cの人は} 2 [2, 1] \end{array} \right.$ と答えて下さい。

“したことがない人”で、きめたのが $\left\{ \begin{array}{l} \text{Aの人は} 2 [1, 1] \\ \text{Bの人は} 1 [2, 1] \\ \text{Cの人は} 1 [1, 2] \end{array} \right.$ と答えて下さい。

答 (1か2を記入)

Q 9. 風, 雲, 虹のうち一つを心の中できめて下さい。次の質問を読み、リストに従って答えて下さい。
質問: 「あなたは大学に入ってからカンニングをしたことがありますか?」

リスト:

“したことがある人”で、きめたのが $\left\{ \begin{array}{l} \text{風の人は} 1 [2, 2] \\ \text{雲の人は} 2 [1, 2] \\ \text{虹の人は} 2 [2, 1] \end{array} \right.$ と答えて下さい。

“したことがない人”で、きめたのが $\left\{ \begin{array}{l} \text{風の人は} 2 [1, 1] \\ \text{雲の人は} 1 [2, 1] \\ \text{虹の人は} 1 [1, 2] \end{array} \right.$ と答えて下さい。

答 (1か2を記入)

Q 10. 七, 五, 三のうち一つを心の中できめて下さい。次の質問を読み、リストに従って答えて下さい。

質問：「あなたはこれまでの質問に正しい手続きで正直に答えましたか？（たとえば、質問やリストを読んでから、きめたものを変更した人は正しい手続きで正直に答えたとはいえません）」

リスト：

“正直に答えた人”で、きめたのが $\begin{cases} \text{七の人は} 1[2, 2] \\ \text{五の人は} 2[1, 2] \\ \text{三の人は} 2[2, 1] \end{cases}$ と答えて下さい。

“正直に答えなかった人”で、きめたのが $\begin{cases} \text{七の人は} 2[1, 1] \\ \text{五の人は} 1[2, 1] \\ \text{三の人は} 1[1, 2] \end{cases}$ と答えて下さい。

答 (1か2を記入)

以上で終了します。なお、この調査法についての意見や感想があったら表紙に書いて下さい。またプライバシーに関する質問の調査法として考えられるものがあつたらそれも書いて下さい。

リストの [] の中の最初の数字は第2組の標本用であり、後の数字は第3組の標本用であり、第1組の標本用の調査票には [] の部分は印刷されていない。

ここで前に述べた基本仮定 (ii) について少し詳しく説明する。たとえば Q1. を例にとってみよう。まず桜、梅、桃のうち一つを心の中で回答者にきめてもらう（ついでながら、好きなものを選ぶようにいっていないことに注意する必要がある。非常にデリケートな内容の質問になった場合、回答者が補助質問で何を選んだかがわからないことによって回答者の秘密は守られ得るのであるから、好き嫌い他人が知り得ることもあり、補助質問としては避けた方が無難であろう）。この段階で基本仮定は、桜、梅、桃の選択確率は各組に共通であることを含んでいる。こう仮定することは無理がないように思われるが（3組はいずれも同一母集団から無作為に抽出されるから）、調査実施の場面では予期せぬ原因で、このことが成り立たない恐れがある。たとえば、この調査では各組の調査票の識別を容易ならしめるために、表紙の色を第1組は桃色、第2組は空色、第3組は黄色にした。おそらく、このためであろうと思われるが後述するように、Q3. の赤、青、黄の選択確率は組によってちがっているようである。しかし、こうした場合を除けば、選択確率は通常各組に共通であることを確かめておく必要がある。さて次に、目的としている質問を読んでもらい、リストに従って‘1’あるいは‘2’と答えてもらうことになる。重要なことは、最初にきめた桜、梅、桃の選択を変更せずに、正直に間違いずにリストに従って答えてもらうことである。もし、質問を読み、リストを見てから選択結果を変更することになると、調査結果は信用できないものになってしまう。すなわち、選択、質問の部分は3組ともまったく同じであるが、リストの部分は3組別々であるから、選択がリストを見てから行われると、基本仮定 (ii) の成立は期待できにくくなる（実際に応用する場合には、したがって勿論、選択項目の部分と質問およびリストの部分とを別個に提示するなどの配慮が必要であろう）。

基本仮定 (i) については、この調査が主として統計学の講義の時間に演習を兼ねて行われているので、1組の調査票は3人おき、あるいは3列おきなどに配られており、その成立を疑う理由はとくにない。

以下、調査結果およびその特徴的な事項を示すと同時に、基本仮定に関する検討を行う。

§2. 調査対象と集計結果

調査は昭和51年10月から11月にかけて行われた。この調査は統計数値を得ることが目的でなく、補助質問使用のランダム回答法モデルの適用可能性を検討するのが主目的であったか

ら、その調査対象は著者らにとって調査実施面で協力の得やすい東京、岡山の一部の大学において統計学を履修している学生が主として選ばれた。したがって、ここに得られた数値が、調査されていない学生を含むいかなる母集団をも代表するものではないことを断っておく。

2.1 調査対象

調査対象は、岡山の2大学7クラス、東京の7大学10クラス（内2つは女子大学、1つは東京近郊）、計9大学17クラスの主として統計履修学生で、調査時点に出席していた学生893名である。それぞれのクラスで講義時間、あるいは演習時間の一部を調査に当て、調査票を配り、説明をした後、学生に回答を記入させた。説明は調査票に添付した2頁からなる解説記事にそってなされたが、説明の時間と回答時間には、それぞれのクラスでバラツキがある。後述するように、主要な集計は不明な回答を含まない調査票のみを採用して行ったが、その意味での標本数は823であった。それらの大学別、性別、組（調査票）別、学年別の内訳を表1に示す（性別や学年に不明があるのは、この調査は原則として記名式であったが、とくに希望する場合は無記名でもよいことにしたためである）。

表 1

大学名	標本数	男	女	不明	第1組	第2組	第3組	1年	2年	3年以上	不明
A (中央大学)	122	119	3	0	39	42	41	0	111	11	0
B (日本大学)	24	20	1	3	9	9	6	0	21	1	2
C (東京大学)	54	39	15	0	19	17	18	0	47	6	1
D (東京女子医科大学)	187	0	187	0	66	63	58	108	79	0	0
E (東京農工大学)	53	51	2	0	17	18	18	0	42	11	0
F (千葉大学)	18	12	5	1	8	5	5	0	0	18	0
G (聖心女子大学)	16	0	16	0	6	5	5	0	5	11	0
H (岡山大学)	230	179	50	1	76	76	78	108	68	53	1
I (岡山理科大学)	119	75	26	18	36	42	41	50	61	8	0
計	823	495	305	23	276	277	270	266	434	119	4

2.2 集計結果

§1に示した調査票の質問事項を見ればわかるように、最初の5問はランダム回答法としての回答（1か2で答える）のほかに、補助質問で何を選んだか（Q1なら桜、梅、桃のいずれか）ということも、本質問に対する直接の答（Q1なら生れ月）も記入させている。したがって、この5問については、回答者がリストの指示に従って正しく回答しているか否かを調べることができる。延4465問（＝5問×893人）のうちに、無記入が18、誤答が53あった。したがって、無記入を除いた中での誤答率は1.2%（＝ $\{53/(5 \times 893 - 18)\} \times 100$ ）である。なお、誤答の6割はQ1に集中し、残りは他の4問にほぼ同数ずつにわかれている（53の誤答の内訳はQ1からQ5にそれぞれ32, 5, 5, 6, 5）。また、Q1からQ10までについて、誤答や無記入を含む調査票は第1組が23、第2組が19、第3組が28で、これらを除く完全な調査票は823（＝893－70）である。以下の集計にはこの823の調査票のみを使用する。

表6の各質問に対する各組ごとの回答‘2’の人数から、(1)式を用いて推定した $\hat{\pi}$ の値が全体、性別、地域別、学年別に、表2に示されている。表2のQ1からQ5までの $\hat{\pi}$ は、直接質問に対する答を集計したものである。 $\hat{\pi}$ が推定値である。Q1は“早生れ”，Q2は“男の兄弟がいる”，Q3は“北海道に行ったことがある”，Q4は“朝食ではパン食の方が多い”，Q5は“地球以外の星に生物がいると思う”，Q6は“キセル乗車の経験がある”，Q7は“恋

人がいる”, Q8は“高校時代にカンニングをしたことがある”, Q9は“大学入ってからカンニングをしたことがある”, Q10は“この調査で正直に答えた”という人の比率である。

表 2

区 分	標 本 数				Q 1			Q 2			Q 3			Q 4			
	1組	2組	3組	計	π	$\hat{\pi}$	2s	π	$\hat{\pi}$	2s	π	$\hat{\pi}$	2s	π	$\hat{\pi}$	2s	
全 体	276	277	270	823	0.282	0.205	0.101	0.564	0.608	0.104	0.205	0.284	0.102	0.516	0.570	0.104	
性 別	男	158	171	166	495	0.281	0.229	0.131	0.574	0.572	0.134	0.196	0.242	0.131	0.420	0.514	0.135
	女	111	99	95	305	0.282	0.148	0.165	0.544	0.612	0.172	0.223	0.342	0.168	0.685	0.662	0.170
地域別	東京	164	159	151	474	0.291	0.187	0.134	0.559	0.582	0.138	0.276	0.304	0.136	0.578	0.509	0.137
	岡山	112	118	119	349	0.269	0.222	0.153	0.570	0.638	0.159	0.109	0.258	0.153	0.433	0.651*	0.160
学年別	一年	95	89	82	266	0.301	0.185	0.178	0.530	0.509	0.183	0.120	0.291	0.178	0.549	0.499	0.184
	二年	145	144	145	434	0.283	0.209	0.140	0.574	0.652	0.143	0.244	0.286	0.141	0.509	0.597	0.143
	三年以上	36	42	41	119	0.235	0.204	0.260	0.597	0.678	0.271	0.261	0.207	0.269	0.471	0.635	0.274

(続き)

区 分	Q 5			Q 6		Q 7		Q 8		Q 9		Q 10		
	π	$\hat{\pi}$	2s	$\hat{\pi}$	2s	$\hat{\pi}$	2s	$\hat{\pi}$	2s	$\hat{\pi}$	2s	$\hat{\pi}$	2s	
全 体	0.938	0.963	0.099	0.681	0.104	0.337	0.103	0.293	0.102	0.552	0.104	0.906	0.098	
性 別	男	0.943	0.984	0.127	0.695	0.133	0.282	0.131	0.428	0.134	0.665	0.133	0.848	0.127
	女	0.931	0.963	0.160	0.682	0.171	0.399	0.171	0.011	0.152	0.401	0.171	1.006**	0.158
地域別	東京	0.928	1.019**	0.127	0.833	0.134	0.310	0.136	0.239	0.133	0.572	0.137	0.896	0.129
	岡山	0.951	0.887	0.154	0.469	0.160	0.368	0.156	0.362	0.159	0.514	0.160	0.918	0.149
学年別	一年	0.925	0.950	0.174	0.523	0.184	0.231	0.176	0.295	0.179	0.426	0.184	0.937	0.170
	二年	0.945	0.970	0.136	0.790	0.140	0.383	0.143	0.299	0.141	0.659	0.143	0.880	0.136
	三年以上	0.941	0.977	0.259	0.636	0.274	0.395	0.269	0.272	0.270	0.487	0.263	0.889	0.250

表 3

組		第1組	第2組	第3組	χ^2 -値 (自由度2)
標本数		276	277	270	
Q1	A	77 (77.8)	90 (78.1)	65 (76.1)	4.80
	\bar{A}	199 (198.2)	187 (198.9)	205 (193.9)	
Q2	A	163 (155.6)	146 (156.2)	155 (152.2)	2.44
	\bar{A}	113 (120.4)	131 (120.8)	115 (117.8)	
Q3	A	60 (56.7)	58 (56.9)	51 (55.4)	0.72
	\bar{A}	216 (219.3)	219 (220.1)	219 (214.6)	
Q4	A	141 (142.5)	148 (143.0)	136 (139.4)	0.56
	\bar{A}	135 (133.5)	129 (134.0)	134 (130.6)	
Q5	A	257 (258.9)	258 (259.8)	257 (253.3)	1.32
	\bar{A}	19 (17.1)	19 (17.2)	13 (16.7)	

推定量 $\hat{\pi}$ の標準誤差の推定値は次式によって計算される：

$$s \equiv \hat{\sigma}(\hat{\pi}) = \sqrt{\frac{u_1(1-u_1)}{n_1} + \frac{u_2(1-u_2)}{n_2} + \frac{u_3(1-u_3)}{n_3}} \quad (2)$$

ただし、 $u_j = y_j/n_j$, $j = 1, 2, 3$ である。各質問に対する標準誤差の推定値を2倍したものを、 $2s$ が表2に示されている。

π の信頼区間として、 $\hat{\pi} \pm 2s$ をとるとき、 π を含まない場合は岡山の大学の Q4 に対する推定（表2の中で*を付した所）だけであるが、Q1とQ3については信頼区間の端近くに π の値がくる場合が多い。なお、表2の中で**を付した所は、推定値を当然1にすべきである。

§1で述べたモデルの基本仮定を検討するために、次の集計を行なった。表3は、Q1からQ5までについて、各組ごとに A , \bar{A} の人数を集計したものである。なお、 A は表1の説明にある各質問の選択肢（たとえば、Q1なら“早生まれ”，Q2なら“男の兄弟がいる”）を表している。（ ）内の数字は3組における A の比率が同一であると仮定したときの期待値である。

表4は、Q1からQ5までについて、各組ごとに補助質問の選択率（百分率）を示したものである。

表 4

質問	選択肢	1組	2組	3組	全体	χ^2 (自由度4)
Q1	桜	46	52	54	51	3.90
	梅	24	20	21	21	
	桃	30	28	25	28	
Q2	ライオン	43	42	45	43	0.97
	ト ラ	32	35	31	33	
	ビ ョ ウ	24	23	24	24	
Q3	赤	40	30	28	33	15.68
	青	46	54	49	50	
	黄	14	16	23	17	
Q4	イギリス	35	32	31	33	2.65
	フランス	36	33	37	35	
	ドイツ	29	34	32	32	
Q5	リンゴ	45	38	41	41	3.12
	ナ シ	21	24	24	23	
	ミカン	34	38	36	35	

表5は、Q1からQ5までについて、各組ごとに補助質問の選択と特性 A の関係の有無をクロスした実数（上段）と比率（下段、百分率）を示したものである。表の最後の列の χ^2 は3つの組における比率が同一かどうかを検定するためのものであり、（ ）内に示したものは、全体として補助質問の選択と特性 A の独立性を調べるためのものである。

§3. 結果の考察

まず、延回答に対する誤答率は1.2%であったが、そのうちの6割がQ1に集中していた。このことは、説明を十分にして、練習として簡単な質問を一つか二つ回答するようにすれば、理解不足による誤答は十分に防ぐことが可能であることを示唆している。[3]によると、シ

表 5

質問	組	C_1A	C_2A	C_3A	$C_1\bar{A}$	$C_2\bar{A}$	$C_3\bar{A}$	χ^2 (自由度 10(2))
Q1	1	36 13.0	26 9.4	15 5.4	92 33.3	39 14.1	68 24.6	22.27
	2	47 17.0	13 4.7	30 10.8	98 35.4	42 15.2	47 17.0	
	3	39 14.4	12 4.4	14 5.2	107 39.6	44 16.3	54 20.0	
	全体	122 14.8	51 6.2	59 7.2	297 36.1	125 15.2	169 20.5	(0.83)
Q2	1	72 26.1	56 20.3	35 12.7	48 17.4	33 12.0	32 11.6	5.97
	2	60 21.7	57 20.6	29 10.5	55 19.9	40 14.4	36 13.0	
	3	64 23.7	53 19.6	38 14.1	57 21.1	32 11.9	26 9.6	
	全体	196 23.8	166 20.2	102 12.4	160 19.4	105 12.8	94 11.4	(4.37)
Q3	1	18 6.5	35 12.7	7 2.5	92 33.3	92 33.3	32 11.6	19.11
	2	14 5.1	36 13.0	8 2.9	70 25.3	114 41.2	35 12.6	
	3	10 3.7	26 9.6	15 5.6	65 24.1	107 39.6	47 17.4	
	全体	42 5.1	97 11.8	30 3.6	227 27.6	313 38.0	114 13.9	(6.45)
Q4	1	45 16.3	53 19.2	43 15.6	53 19.2	46 16.7	36 13.0	10.81
	2	55 19.9	44 15.9	49 17.7	35 12.6	48 17.3	46 16.6	
	3	49 18.1	41 15.2	46 17.0	36 13.3	58 21.5	40 14.8	
	全体	149 18.1	138 16.8	138 16.8	124 15.1	152 18.5	122 14.8	(3.07)
Q5	1	116 42.0	49 17.8	92 33.3	8 2.9	8 2.9	3 1.1	7.60
	2	96 34.7	62 22.4	100 36.1	9 3.2	4 1.4	6 2.2	
	3	103 38.1	61 22.6	93 34.4	9 3.3	3 1.1	1 0.4	
	全体	315 38.3	172 20.9	285 34.6	26 3.2	15 1.8	10 1.2	(6.27)

カゴにおけるランダム回答法（無関連質問法 [1]）の調査では、回答者のうち 5% だけが *confusing*, *silly* または *unnecessary* であると述べたようである。

モデルの基本的仮定は §1 で述べたように二つある。今回の調査において、これらの仮定が満足されたかどうかを検討する必要がある。その第 1 の仮定 (i) は、各組の標本が同一母集団からの無作為標本であるということであった。3 種類の調査票は各クラスで着席順に 3 人おき、あるいは 3 列おきに配布されているので問題はないと思われるが、念のため、Q1 から Q5 までについて、「A の比率と {C_i} の比率が各組共通である」という仮説を検定してみる。表 3 によると、A の比率に関しては、いずれの質問についても仮説は棄却されない（有意水準 10%）。

表 6

区 分	組	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	
全 体	1	90	140	134	149	168	164	102	149	148	137	
	2	107	153	136	152	171	142	148	105	146	177	
	3	133	148	83	130	199	155	117	101	132	208	
性 別	男	1	57	74	72	78	92	97	49	84	88	75
		2	62	95	83	92	114	87	92	75	105	102
		3	84	91	50	80	122	95	72	75	82	129
	女	1	30	61	60	68	72	63	49	62	57	59
		2	41	52	47	58	54	53	51	25	41	72
		3	44	51	31	44	73	55	42	19	45	71
地域別	東京	1	56	86	83	87	105	110	67	90	89	79
		2	66	85	70	85	95	90	78	55	92	107
		3	65	79	54	67	118	90	62	52	69	112
	岡山	1	34	54	51	62	63	54	35	59	59	58
		2	41	68	66	67	76	52	70	50	54	70
		3	68	69	29	63	81	65	55	49	63	96
学年別	1年	1	27	42	49	49	58	48	29	53	44	46
		2	40	45	44	43	53	45	52	32	39	62
		3	37	46	23	41	61	42	28	31	43	62
	2年	1	51	81	70	81	90	97	61	78	85	76
		2	54	83	70	84	87	73	74	58	79	86
		3	70	75	46	66	108	89	65	52	76	110
	3年 以上	1	12	17	14	19	20	19	12	18	19	15
		2	12	23	20	25	31	24	20	14	28	27
		3	24	27	14	21	28	22	24	18	12	34

表4によると、 $\{C_i\}$ の比率に関しては、Q3を除くと仮説は棄却されない（有意水準10%）。Q3では有意水準0.1%でも仮説は棄却される。すなわち、補助質問の赤、青、黄の選択率が組によって異なっている。全体としての選択率と比較すると、第1組では赤、第2組では青、第3組では黄色の選択率が高くなっている。§1で触れたように、今回の調査では、各組の調査票の識別を容易にするため、調査票の表紙を第1組は桃色、第2組は空色、第3組は黄色と色分けした（実際の調査ではその必要はない）。このことが原因となって、各組の表紙の色に近い色の選択率がそれぞれ高くなったと考えられる。

第2の基本仮定(ii)は、 $\{C_1A, \dots, C_3\bar{A}\}$ の比率の期待値が各組に共通であるということであった。一寸考えると、各組の標本が同一母集団からの無作為標本であるならば、この仮定は自動的に保証されているように思える。しかし、§1でも述べたように、リストにおける数字の順序が異なっており、実際に調査票を開き、回答を記入する時点で組ごとに異なった影響が現われなとは限らない。表5によると、Q2、Q4、Q5については、3組の比率が共通であるという仮説は、有意水準10%で棄却されない。Q1、Q3については、有意水準5%で棄却される（有意水準1%では棄却されない）。Q3については、 $\{C_i\}$ の選択比率が前述のように異なっていることで、その原因が説明されよう。Q1については、その原因は不明であり、偶然によるのか、それとも何らかの理由があるのかを更に検討する必要がある。

次に、Aの比率の推定値 $\hat{\pi}$ の値について考察してみる。表2によるとQ1からQ5までに

については、 $\hat{\pi} \pm 2s$ の区間をつくる時、それらの区間は全体で見ると、いずれも π を含んでいる。したがって、とくに述べることはない。Q6からQ10については、 π の値がわからないので、上記のような検討はできないから、各質問について予想される傾向を拾い上げ、推定結果がそれらの傾向を満たしているかどうかを調べることにした。なお、すでに注意したように、調査対象はあくまで調査実施上の便宜によってきまったものであるから、推定値そのものに意味を見出そうとすることは避けなければならない。たとえば、表2で、男子学生、東京の大学と書いてあっても、それらは一般の男子学生、一般の東京の大学を代表するものではない。さて、Q6からQ10について予想される傾向として次のものをあげることができよう。

(Q6) 岡山より東京の方がキセル乗車の誘因が多いであろう。

(Q7) 男子学生より女子学生の方が、一年生より二年生の方が、恋人のいる率が高いであろう。

(Q8とQ9) 高校時代より大学の方が、女子学生より男子学生の方が、カンニングが多いであろう。大学に限ると、当然、一年生より二年生の方がカンニング経験が多いであろう。

(Q10) 二年生より一年生の方が、男子学生より女子学生の方が調査に対して協力的であろう。

表2を見ると推定結果は、いずれも上記の予想と同じ傾向を示していることがわかる。これらのことに、あまり信をおくことは危険であるが、調査結果が妥当なことを一面から裏付けているといえよう。

§1で述べたモデルの特徴は、一つが「補助質問の三つの選択肢の選択率がそれぞれ $1/3$ であるということは仮定しない」ということであり、一つが「補助質問の選択と特性 A の間の独立性を仮定しない」とことであった。これらのことが、果して必要であるかどうかということについて考察しよう。まず、前者については、表4から明らかなように、Q4を除くと、三つの選択肢の選択率はいずれにおいても異なっている。いずれの質問においても、三つ並べた中で、最初に並べたものの選択率は $1/3$ 以上であることが一つの特徴といえそうであるが、そのほかに、Q3においては、青の選択率がかなり高くなっているといったこともある。この意味で、「選択率が $1/3$ であることを仮定しない」ということはモデルの有用性を示しているといえてよからう。

後者については、表5によると、Q3を除くと有意水準 5% で独立性の仮説は棄却されない。しかしながら、[4]で考察したように、独立性を仮定したモデルによる推定を行なえるためには、独立性が積極的に裏付けされる必要がある。もし独立性が保証されているならば、標本を2群に分け、一つの群では補助質問のみを行ない、選択比率を推定し、他の群で補助質問をランダム化器具の代りに使って、たとえば Warner 法による調査を実施することができる ([4] 参照)。いま、今回の調査結果を利用して、Q1からQ5までについて、独立性を仮定したモデルによって π の推定値 $\hat{\pi}$ を求めてみる。たとえば、Q1についていうと、補助質問の選択が桜か桃である比率を第2組と第3組のデータから推定すると、表5からわかるように、 $436/547 = 0.797$ である。第1組では、補助質問の選択が桜か桃で、かつ A の人か、補助質問の選択が梅で、かつ \bar{A} の人が '2' と回答し、それ以外の人 '1' と回答している。そこで、第1組のデータを 0.797: 0.203 の割合のランダム化器具を使用する Warner 法の実験結果とみなすと、

$$\hat{\pi} = \frac{(90/276) - 0.203}{0.7971 - 0.203} = 0.207$$

を得る。

すなわち、Warner 法で用いられるランダム化器具の割合 p が未知である時、別の独立な標本

から推定された \hat{p} を p の代りとして使うというもので、 π の推定量 $\hat{\pi}$ は次の式で表される。

$$\hat{\pi} = \left(\frac{y}{n} - \hat{q} \right) / (\hat{p} - \hat{q})$$

p を推定するために用いた標本の大きさを m とすると、 m が十分に大きければ、

$$E(\hat{\pi}) = \pi$$

$$E(\hat{\pi} - \pi)^2 = \frac{1}{(p - q)^2} \cdot \frac{\lambda(1 - \lambda)}{n} + \left(\frac{\pi - \bar{\pi}}{p - q} \right)^2 \frac{pq}{m}$$

となる。但し、 $\lambda = p\pi + q\bar{\pi}$ 、 $\bar{\pi} = 1 - \pi$ である。 $\hat{\pi}$ の推定の精度を $\sigma(\hat{\pi}) = \{E(\hat{\pi} - \pi)^2\}^{1/2}$ で測ることにして上式の p 、 λ 、 π にそれぞれの推定値を代入すると、Q1 の場合では、 $2\sigma(\hat{\pi}) = 0.101$ となる。このようにして Warner 法を使った場合の推定値とその精度が計算されるが、どの組のデータを Warner 法の実験結果として選ぶか、ということが問題である。ここでは、その規準として、上記の 0.797 に当る値が 0.5 からもっとも遠いものを採用することにする。これは推定の精度を最良にする組み合わせ方である。この時 Q2 では第2組、Q3、Q4、Q5 では第3組のデータを Warner 法の実験結果とみなし、他の2組で p を推定することになる。それぞれの質問での計算結果を表7にまとめる。

表 7

質問	Q1	Q2	Q3	Q4	Q5
\hat{p}	0.797	0.760	0.852	0.685	0.778
π	0.207	0.601	0.226	0.450	0.927
$2\sigma(\pi)$	0.101	0.116	0.083	0.164	0.111

Warner 法の精度は、標本の分割の仕方、 π 、 p の値によって異なり、この試算結果だけから判断するのは公正でないが、推定量 $\hat{\pi}$ と同等の精度を得るためには、 p が 0.5 から相当離れている必要がありそうで、これは、回答者の秘密保護に大きな疑問をいだかせる。逆に、 p を $\frac{2}{3}$ とした時、標本の数は p を推

定するのに必要な m を別として、約 900 必要であり ($\pi = \lambda = \frac{1}{2}$ の場合)、Warner 法の推定効率はかなり落ちる。また、あらかじめ沢山の補助質問を考えて、それぞれの p を推定した後、その推定値を使って Warner 法による調査を行なり、という2段階の調査を実施すれば適当な割合を持った補助質問を選ぶことも可能であるが、もし、補助質問と本質問の調査を同時に行った場合(手間、費用のことを考えると、多くの場合はこうなるであろう)、適当な割合を持つであろう補助質問を作り出す(あるいは選ばれた補助質問の選択率を事前に推定する)必要がある。これが困難な作業であることは、今回の調査で使われた5問の補助質問選択比率をデータを見ないで推量し、実際のデータと比べてみたらわかるであろう。このことは、($\hat{\pi}$ の精度が p の値に大きく影響される為)必要な精度を確保するための標本の大きさが前もって決定できないということを意味している。このように、たとえ、補助質問と特性の間の独立性が仮定できたとしても、ランダム化器具を使わない Warner 法によって π を推定する方法には、解決を要する多くの問題点が残されている。

§4. 標本数の割当てについて

これまででは、端数的なちがいは別として、原則として、大きさ n の標本を三つの組に分けるとき、 $n/3$ ずつ割当てることにしてきた。これは調査実施の容易さが一つの理由である。しかしながら、調査票を作成する段階において、ある程度の労をいとわなければ、各質問ごとに、3つの組への割当てを変更することができる。推定の精度の面からだけ考えれば、 $n/3$ ずつ割当てることは最適ではない。

(1) 式で与えられる推定量 $\hat{\pi}$ の分散は

$$\sigma^2(\hat{\pi}) = \frac{\theta_1(1-\theta_1)}{n_1} + \frac{\theta_2(1-\theta_2)}{n_2} + \frac{\theta_3(1-\theta_3)}{n_3} \quad (3)$$

である。ただし、 θ_i は第 i 組における回答 '2' の出現率の期待値、すなわち、 θ_1 は $C_1\bar{A}UC_2AUC_3A$ 、 θ_2 は $C_1AUC_2\bar{A}UC_3A$ 、 θ_3 は $C_1AUC_2AUC_3\bar{A}$ の母比率である。

$n_1+n_2+n_3=n$ のもとで (3) 式を最小にする n_i は

$$n_i = n \frac{\sqrt{\theta_i(1-\theta_i)}}{\sum_{j=1}^3 \sqrt{\theta_j(1-\theta_j)}}, \quad i = 1, 2, 3 \quad (4)$$

である。このときの (3) の値 $\sigma_0^2(\hat{\pi})$ は

$$\sigma_0^2(\hat{\pi}) = \frac{1}{n} \left\{ \sum_{i=1}^3 \sqrt{\theta_i(1-\theta_i)} \right\}^2 \quad (5)$$

である。(3) 式で $n_1=n_2=n_3=n/3$ とおいたものを $\bar{\sigma}^2(\hat{\pi})$ とすると

$$e \equiv \frac{\bar{\sigma}^2(\hat{\pi})}{\sigma_0^2(\hat{\pi})} = \frac{3 \sum_{i=1}^3 \theta_i(1-\theta_i)}{\left\{ \sum_{i=1}^3 \sqrt{\theta_i(1-\theta_i)} \right\}^2} \quad (6)$$

である。 θ_i の代りにその推定値 y_i/n_i を表 6 から求め、(6) 式に代入して e を求めてみると、たとえば、Q1 では 1.00007、Q10 では 1.005 というように、どの質問でもほとんど効果のないことがわかる。

予備調査などで特に効果のあることが明きらかにならない限り、標本数は同数ずつ割当てるのが適当であろう。

なお、最近、ランダム回答法一般について、回答者のさらされる危険（ランダム回答法で回答したときの特性 A をもつと判断される事後確率に関連したもの）を考慮した上での、モデルの推定の精度に関する議論が [2] で行なわれている。

§5. 回答者の感想

調査を実施した条件が各大学、各クラスで異なっていたので、回答者の感想を求めるのに十分な時間がない場合もあった。したがって、以下の感想は回答者全員からのものではないが今後この種の調査研究を実施する際の参考として以下 4 つの項目にまとめた。

a. 手法そのものに対する判断

a-1 評価するもの

- ① 調査がやり易い
- ② 答えにくいことを答えるのにより正確になる
- ③ 正直に回答し得た
- ④ 有効な方法である

a-2 疑念を表明するもの

- ① 本当に求めたい比率が得られるかだまされたような感じがする
- ② 心の中にきめるものがランダムになるのかどうか
- ③ やっぱり偽答、回答拒否はさげられないのではないか
- ④ 正確に出来るとは考えられない
- ⑤ でたらめ回答になりがちではないか
- ⑥ やっぱりどこかで秘密がもれてしまうのではないか

b. この手法の手續き上の問題点として提案されたものは

- ① やり方が複雑でまちがえる可能性があること
- ② 心の中できめるもの(補助質問)と本来の質問を列記してあるので混乱する
- ③ 補助質問の3選択肢に品物, 植物, 事象等同一種類のものを利用するのは選択に迷う
- ④ 補助質問の選択肢が3つであるのに本来の質問の回答肢が2つで形式上バランスしていない

c. 手法の改善として提案されたもの

- ① 補助質問を各問ごとにしなくて共通にするのはどうか
- ② 補助質問に利用する3つの回答選択肢を記号のような無関係なものにする
- ③ 回答時間を適当にきめる

d. 実施上の問題として提案されたこと

- ① 手法をよく納得させる必要がある
- ② 質問項目のとり上げ方, 質問内容等を吟味する必要がある
- ③ 通常の方法で正確を期するよう努力する方がよい
- ④ 例題を用いて練習させる必要がある

さらに, この手法の問題点として考えられることは推定の精度のことである。

推定の精度については, §4の(3)式からわかるように $n_1 = n_2 = n_3 = n/3$ のとき推定量 $\hat{\pi}$ の分散は

$$\sigma^2(\hat{\pi}) = \frac{3}{n} \sum_{i=1}^3 \theta_i (1 - \theta_i)$$

である。これは, $\theta_1 = \theta_2 = \theta_3 = 1/2$ のとき最大になる。

このときの $2 \times \sigma(\hat{\pi})$ を通常の単純ランダムサンプリングによる調査における $2 \times$ 標準誤差の大きさとくらべると3倍になる(分散で9倍)。

これを, いくつかの標本数の場合について示すと表8のようになる。

たとえば, 通常の調査では標本数1000で上記の値は約3%であるが, この方法で同様の精度を得るためには標本数を9000にする必要があり, とくに標本数の小さいときには注意が必要である。

疑念を表明するもののうち, とくに補助質問の利用について多くの学生はこれをカード等のランダム化器具の代用としてとらえているので, 「その選択比率をどのように推定するのか?」「選択比率が平等でないところに何かからくりがあるのではないか?」——たとえば最初が選択されやすい等の一定の傾向を利用して, どの回答をしたか判断を下す等のこと——「補助質問の選択と本来の質問の回答との関連が高いとき問題があるのではないか?」等の疑問点が多くみられた。これらの点は今回の実験調査の調査実施前に利用した説明にも原因があるが, 問題点を一般の人々に納得できる形で事前に説明することは, 実際上かなり困難なことと考えられる。この手法を実際に利用するとき一つの問題点といえるであろう。

手法の手続き上の問題点としてあげられたものは, 今回の実験調査の方式を変更して, まず補助質問の選択をおこない, 次にその選択に応じて, 本来の質問の回答を一定形式で回答させる方式に改める方式をとるとほとんど解消される。すなわち,

- 補助質問例 ○, ×, △のうち一つを心の中できめて下さい
 本質問例 あなたはAですか Bですか 下表に従ってアカイで回答してください

表 8

調査方式 標本数 (n)	2 × 標準誤差の値	
	通常の調査 (単純ランダム サンプリング)	この論文の ランダム回答法
50	14.1	42.4
100	10.0	30.0
150	8.2	24.5
200	7.1	21.2
250	6.3	19.0
300	5.8	17.3
400	5.0	15.0
500	4.5	13.4
1000	3.2	9.5
2000	2.2	6.7
3000	1.8	5.5
5000	1.4	4.2
10000	1.0	3.0

$\pi=1/2$ としたときの数値

$$\begin{array}{l}
 \text{きめたのが } \bigcirc \text{の人は } \begin{cases} A & \text{ならば}\cdots\text{ア} \\ B & \text{ならば}\cdots\text{イ} \end{cases} \\
 \text{きめたのが } \times \text{の人は } \begin{cases} A & \text{ならば}\cdots\text{ア} \\ B & \text{ならば}\cdots\text{イ} \end{cases} \\
 \text{きめたのが } \triangle \text{の人は } \begin{cases} A & \text{ならば}\cdots\text{イ} \\ B & \text{ならば}\cdots\text{ア} \end{cases}
 \end{array}$$

しかし、手続きの複雑さの問題点は、技法上解消されない。手法の改善として提案されたものも、この範囲のことである。補助質問を共通にすることは複雑さを軽減するが、選択比率の偏りの問題と関連するので、この方式による調査実施までにはかなりの準備が必要であろう。d については、次節に含めてのべる。

§6. 結 び

実際の調査の場において、調査が匿名であることを十分理解してもらって正確な回答を求めているにもかかわらず、調査の場における信頼関係の確立に失敗するなどのことから、これが偽答、回答拒否に結びついてくる。

とくに、個人的にみて秘密にしておきたいことなどが調査目的となる場合には、信頼関係の確立に相当の努力をしても、疑惑の念を完全には打ち消すことが出来ない場合も考えられるし、被調査者にとっては、面前にいる調査員に自分の考えを“知られてしまう”ということが不測の悪影響を及ぼすのではないかという疑念が絶えず存在する。調査の実施上の倫理とでもいべきものを十分に考え、特定個人の回答内容をそれと分るそのまゝの形では外部に一切出さない、という確固とした手続きをふんで調査を実施する場合においても、このような疑念が存在するので、偽答や回答拒否が発生する可能性が考えられる。このような状況のとき個人の正確な回答を一定の手続きにより確率化し、個人を特定できない形にして、一定の形式で正確な情報収集をはかるのがランダム回答法といえる。

したがって、ランダム回答法にとっては、それぞれの回答者に、回答の手続きを十分に理解した上で、正確な情報を提供してもらうことが必須の条件といえる。

調査員と回答者との間の信頼関係の確立に十分の努力をほらいその上で、質問項目の特殊性からランダム回答法を利用するという立場をはっきり示して回答者に了解してもらおうという手続きをおろそかにしてはこの方法は成立しない。これらの点から考えて、学生調査における調査後の感想は大変有用である。感想をのべた学生も、これらの点を強調している。

本研究は、ランダム化器具を利用しないで、補助質問の選択により、回答を確率化する手法の実用の可能性を探ることにある。その目的は前述のごとく大略達成されたものと考えられるが、この調査に附随して生じたことあるいは最近の動向にかんがみて、一、二の注意をのべておく。

ランダム回答法は、この方法だけで独立して利用するのは非常に危険であると考えられる。上述のごとく通常的手法でとことんまでつめて調査を実施し、その上での補助手段として考えるべきものと思う。アメリカにおいても実験的段階であり、たとえば、私生児に関する調査においても、それまで主として得られている情報を補う意味の利用の仕方が主なものとみられる。また、ランダム回答法は、社会的にみてよくない事項の調査に利用したとき、面接法、電話法、自記式調査法等に比較して客観的データと対照したとき、より一層偏りが少なく出るといふ傾向は認められるが、それでも偏りを解消させるまでには到らないことが、実験調査として報告されている [3]。

すなわち、統計的モデルとしては、有効性がみとめられるとしても、この手法の中に入りこ

む調査誤差（偽回答，デタラメ，回答拒否，手法を理解しないための誤回答など）が考えられ，その大きさが不明であるから，推定されたものが時として誤った情報源となりかねないからであり，これによって引き起される実害の方が大きくなる危険性がある．このような誤差は，ランダム回答法の手法そのものだけからは見出すことが不可能であるから，他の方法との併用あるいは通常の方法の補助手段としての利用を通じて，この手法の有効性を吟味検討していくことが必要である．

さらに，調査誤差（偏りなど）があるばかりでなく，前述のとおり推定の誤差（推定量の標準誤差の2倍）を考えると，通常の方法にくらべ3倍になり（ $\pi=1/2$ の場合），精度の面からみてかなり悪くなる．普通の調査を実施するような感覚で，安易に目標精度の高い調査を企画することは出来ないことに留意する必要がある．

本研究には多くの方のご協力を得た．実験調査の実施には統計数理研究所の水野欽司，大隅昇，村上征勝，馬場康維，清水良一，早川 毅の各氏，および，岡山大学の脇本和昌氏，岡山理科大学の一村 稔の各氏に種々の面でお世話になった．

集計整理の過程では統計数理研究所の高橋和子さんにお世話になった．

これらのご協力いただいた諸氏ならびに有益な助言を頂いたレフェリーの方に深謝する．

参 考 文 献

- [1] Horwitz, D.G., Shah, B.V. and Simmons, W.R. (1967) "The unrelated question randomized response model", Proceedings of the American Statistical Association, Social Statistics Section, 67-72.
- [2] Leysieffer, F.W. and Warner, S.L. (1976) "Respondent jeopardy and optimal designs in randomized response models", JASA, vol. 71, No. 355, 649-656.
- [3] Locander, W., Sudman, S. and Bradburn, N. (1976) "An investigation of interview method, threat and response distortion," JASA, Vol. 71, No. 354, 269-275.
- [4] 鈴木達三，高橋宏一，逆瀬川浩孝 (1977) "ランダム回答法における二，三の注意——クロス集計にもとづく推定の精度，偽答・D.K.の影響，補助質問使用の問題——"，統計数理研究所彙報，第24巻，第1号，No.45, 1-13.
- [5] Takahasi, K. and Sakasegawa, H. (1977) "A randomized response technique without making use of any randomizing device," Ann. Inst. Statist. Math., Vol. 29, No. 1, 1-8.