

論文

紙媒体アンケートにおける複数回答形式の 設問の効率的な集計方法の開発

An effective aggregation scheme development for questions
with multiple answers allowed in a printed questionnaire

北川 晃・服部 裕一郎・遠藤 尚・加納 理成・柴 英里・鈴木 恵太
・武久 康高・幸 篤武・柳林 信彦 (高知大学教育学部)

KITAGAWA Akira, HATTORI Yuichiro, ENDO Nao, KANO Risei,
SHIBA Eri, SUZUKI Keita, TAKEHISA Yasutaka, YUKI Atsumu
and YANAGIBAYASHI Nobuhiko

Faculty of Education, Kochi University

ABSTRACT

We report an effective aggregation scheme for questions with multiple answers allowed in a printed questionnaire. The answer pattern is represented with a short code based on the octal digit that can be easily generated by means of a mental arithmetic and that is uniquely reconstructed with a general spreadsheet software. This coding procedure is very simple task and the code is generated with a little bit of help from the answerer in the present scheme. We also show an example of the question sheet. The present scheme is primarily designed for a questionnaire carried out in the Faculty of Education, Kochi University. However, the procedure is universal and that is widely applicable in various questionnaire.

1 はじめに

現代では高性能なコンピュータが安価に手に入るようになった。教育現場においても、ICT 機器の利点を取り入れた教育についての取り組みの他、日常業務の負担の軽減を目的として、様々な用途でコンピュータが導入されている。

また教育現場では、少しでも効率的な教育を提供することができるように、日々その教育方法の改善の努力がなされているが、その効果を検証するために、アンケートを実施する機会が多い。アンケートは、回答率を上げるために紙を媒体とした形式で行われることが一般的であるが、その集計は表計算ソフトを用いて行うことが多い。これは手間を少なくすることの他、集計上の誤りを少なくするという効果がある。

アンケートの設問には選択式、自由記述式など、さまざまな形式のものがある。選択式の設問のうち、複数回答を認めるものについては、その集計を効率的にするため、例えば三つまでといったように、回答数に上限を設けているのが普通である。しかしながら調査の種類によっては、回答数に上限を設けることができないものがある。例えば、教育学部における卒業要件免許以外の、いわゆる二枚目免許のうち、取得予定のものを答えさせる設問がこれに該当する。

現在高知大学教育学部においては、国語、社会、数学、理科、音楽、美術、保健体育、技術、家庭科、英語の中学校免許の取得が可能となっており、卒業要件のものを除いて、最小で0個、最大で9個の回答がなされる可能性がある。この種の設問では、回答される項目の数が不特定多数な上、その回答順序も任意であるため、コンピュータへの入力を効率的に行うことが困難であり、結果としてその集計作業にも大変手間がかかる。また回答者が100名を超す大人数となった場合、集計作業の繁雑さから入力作業に誤りが生ずることが予想され、正しい集計結果が得られない恐れがある。

効率の良い集計を行うためには、Web ページ上や、専用アプリケーションを用いたシステムを開発する手法が考えられるが、その開発には専門的な知識が必要となる。またコンピュータがなければアンケートに回答できないことから、その回答率が下がることが予想される。

そこで我々は、紙媒体で実施されるアンケートでの複数回答可の設問を、効率よく集計するための手法を開発した。今回の手法においては、集計すべき各項目に数字を割り当て、これより回答状況に対応した符号を生成する手順をとるが、集計をより効率的に行うために、一部の作業を回答者にも協力してもらう形式を前提とする。当然のことながら、この作業内容は集計精度に直結して

いるため、回答者に困難な手順を要求することは避けなければならない。

今回は、回答者に容易な作業（桁上がりのない足し算）を負担してもらうことで、回答状況に応じた符号を生成する手法を提案する。こうした作業内容を回答者に正しく理解してもらうための設問および解答用紙における工夫についても紹介する。また得られた符号より回答の組み合わせを一意に再現することができるが、一般的な表計算ソフトを用いて、符号を入力するだけで回答状況を再現することのできる関数の実装例についても紹介する。本手法では、一部の手順を回答者に負担してもらうことにより、その集計を著しく効率的なものとするすることができる大きな特長といえよう。

本稿では、教育学部エビデンス研究会により本学部生を対象として実施されているアンケート [1] へ今回提案する手法を実装した例について紹介している。しかしながら、この手法は一般的な複数回答可の設問の集計へも適用することが可能な汎用的なものであり、幅広く応用できるものである。

本稿の構成は以下の通りである。第2節において、アンケートでの回答項目の組み合わせを符号化するための手法について説明する。第3節では、得られた符号から、元の回答項目の組み合わせを再現する手法について説明する。続く第4章では、先の二つの節で紹介した手法を実際のアンケートに実装するために、設問および解答用紙の構成例を示し、一般的な表計算ソフトを用いて集計をするための関数の入力例を紹介する。最後に第5節において今回の提案内容をまとめるとともに、さらなる考察を加える。

2 アンケート項目の符号化

本稿においては、アンケートにおける複数回答可の設問の例として、以下のものを想定する。

Q. 「卒業要件以外の中学校教員免許のうち、取得予定であるものをすべて教えてください。」

回答の選択肢としては、国語、社会、数学、理科、音楽、美術、保健体育、技術、家庭科、英語の10種類の項目があり、この中から最小で0個、最大で9個の回答があるものとする（科目の一つは卒業要件として数えるため）。

本設問の回答状況の集計を容易にするため、回答者自身に回答項目からそれに対応する符号を生成してもらい、その符号を元に表計算ソフトを用いて復号化、項目を集計するという手順を考える。今回は紙媒体のアンケート

を想定しているため、符号化の手順は（例えば暗算などにより）容易に紙の上で行えるものとする。符号化の方法としては、1. 二進数に基づくもの、2. 十進数に基づくものが考えられるが、いずれの方法も問題点があることを指摘し、代わる方法として、3. 八進数に基づく符号化を提案する。

2.1 二進数に基づく符号化

最も簡単な符号化は、国語、社会、数学、…の各項目に0もしくは1の符号を割り当て、回答のあったものを1、なかったものを0として読み取る方法である。つまり、国語に1,000,000,000、社会に100,000,000、数学に10,000,000、…、英語に1の二進数を割り当て、回答のあった項目に対応する数字を足し合わせたものを、回答項目に対応する符号とする。この足し算においては、繰り上がりの手順は含まれないため、符号の生成自体は単純である。

計算の結果、例えば一つも項目の回答がなかった場合は0,000,000,000、国語のみの回答があった場合は1,000,000,000、数学と理科、技術の回答があった場合は0,011,000,100といった具合に符号が得られる。言うまでもなく、各桁がそれぞれの項目に対応した10桁の二進符号であり、1となっている桁に対応する項目が回答のあった項目である。この符号化において、任意の符号は項目の特定の組み合わせと一対一に対応しているため、この符号から一意的に元の回答状況を再現することができる。

しかしながらこの方法では、項目の数（ここでは10個）と同じだけの桁の符号が得られるため、復号化をする際にこの数字を覚えて表計算ソフトへ入力する作業は煩雑である。従って入力の誤りが頻発することが予想され、適しているとは言えない。

2.2 十進数に基づく符号化

符号の桁数を減らす工夫として、前節2.1の二進符号を十進符号に変換するという方法が考えられる。すなわち国語に $2^9 = 512$ 、社会に $2^8 = 256$ 、数学に $2^7 = 128$ 、…、英語に $2^0 = 1$ の数字を割り当て、回答のあった項目に対応する数字を足し合わせたものが、得られる符号ということになる。例えば一つも回答がなかった場合は0、国語のみの回答があった場合は512、数学と理科、技術の回答があった場合は $128+64+4=196$ といった具合である。今回のように10項目の場合では、3桁程度（最大でも4桁）の符号が得られる。得られた符号から回答のあった項目を復号化するためには、符号を二進数に変換すれば、2.1節の符号化の場合に帰着する。

表 1: 各項目に対する数字の割り当て

国語	社会	数学	理科	音楽
1,000	400	200	100	40
美術	体育	技術	家庭	英語
20	10	4	2	1

一旦符号化が完了すれば、表計算ソフトを用いて二進符号に変換すること（復号化）はたやすい。しかしながら、符号化の手順において、1~3桁の数字の足し算を行わねばならないが、これには繰り上がりの操作が含まれるため、暗算での計算には誤りが生ずることが予想される。これは回答者に負担してもらい手順としてはやや複雑であるといわざるを得ず、やはり適しているとは言えない。

2.3 八進数に基づく符号化

二進符号では、符号化が単純なものの、符号が長くなりすぎる。また十進符号では、符号は短くできるものの、符号化の手順がやや煩雑になる。そこでそれぞれの長所を生かしつつ、短所を克服できるよう、八進数に基づく符号化を考える。

ここでは、表1のように各項目に数字を割り当てる。先の方法と同じく、回答のあった項目に対応する数字を合計すれば、回答状況に合わせた符号が生成される。例えば国語のみの回答があった場合は1,000、数学と理科と保健体育の回答があった場合は $200+100+10=310$ といった具合である。この符号化の方法においても、項目の組み合わせ一つ一つに特定の符号が対応していることは同じである。

この方法の大きな特長は、数字を足し合わせる際に、繰り上がりの手順が出ないことである。三つの項目にそれぞれ1, 2, 4の数字を割り当てるが、十進数の場合のように四つ目の項目に8の数字を割り当ててしまうと、繰り上がりの手順が出てきてしまう。ここではそれを防ぐために、四つ目の項目の数字を意図的に次の桁のものとし、かつ下の位は0とすることで、それまでの項目に影響が出ないように工夫している。従って四つ目から六つ目は10, 20, 40の数字を割り当て、以下同じく七つ目から九つ目は100, 200, 400、最後の十個目には1,000を割り当てることになる。

繰り上がりがないということは、符号化の際に、各桁を独立して考えることができることを意味する。つまり各桁で、1, 2, 4が最大で一回ずつ出てくる足し算をすることで、符号が得られることになる。この操作は最も初歩的な算術に当たり、回答者自身に負担してもらって

業としては、十分許容範囲内であると考ええる。また生成される符号も、今回の項目数 10 の場合で 4 桁であるから、復号化の際に表計算ソフトに入力することも容易である。符号化と復号化の作業の繁雑さをともに適度に抑えた、人間工学に基づく方法であると言えよう。同様の考え方は、UNIX システムにおけるファイルの属性を与える方法 [2] にも見られる。

3 復号化の方法

本節では、前節の八進数に基づいた符号から、回答項目を復元する、復号化の方法について説明する。率直に考えれば、八進数として得られた符号を、二進数に変換することで、各桁が項目ごとの回答状況に対応する符号が得られる。表計算ソフトには ‘OCT2BIN’ といった八進数を二進数に変換する関数が実装されているため、これを用いることにより、表計算ソフト上で集計を行うことができる。

別の方法として、八進数として得られた符号を、我々になじみのある十進数と見なして復号化することが考えられる。これが可能となるのは、今回の符号化において、繰り上がりの手順を含んでいないためである。以下、この方法の原理について説明する。表計算ソフトへの実装の方法については、次の第 4 節で説明する。

今回得られている 4 桁の符号を A とし、その各桁の数字をそれぞれ a_3, a_2, a_1, a_0 とおく。 A を十進数の符号であるとすれば、

$$A = a_3 \times 10^3 + a_2 \times 10^2 + a_1 \times 10 + a_0 \quad (1)$$

と表すことができる。ここで a_3 は 0 もしくは 1, $a_i (i = 0, 1, 2)$ は $0 \leq a_i \leq 7$ を満たす整数である。符号 A について、それぞれの桁の評価を行う。まず一番上の桁を取り出すためには、 A の二桁目（つまり百の位）以降を切り捨てれば良い。これにより、

$$A \rightarrow A_3 = a_3 \times 1,000 \quad (2)$$

が得られる。もし A_3 が 1,000 の場合は、国語が選択に含まれており、0,000 の場合は含まれていなかったことを意味する。

次に A の二桁目を取り出すために、まず A を商が整数の範囲となるように 1,000 で割り算し、そのあまり A'_2 を求める。

$$A \rightarrow A'_2 = a_2 \times 10^2 + a_1 \times 10 + a_0 \quad (3)$$

ここで得られたあまり A'_2 の三桁目（つまり十の位）以降を切り捨てれば、二桁目を取り出すことができる。

$$A'_2 \rightarrow A_2 = a_2 \times 10^2 \quad (4)$$

表 2: 二桁目の数字と項目の選択状況

A_2	選択された項目
700	社会, 数学, 理科
600	社会, 数学
500	社会, 理科
400	社会
300	数学, 理科
200	数学
100	理科
000	選択項目なし

表 3: 三桁目の数字と項目の選択状況

A_1	選択された項目
70	音楽, 美術, 保健体育
60	音楽, 美術
50	音楽, 保健体育
40	音楽
30	美術, 保健体育
20	美術
10	保健体育
00	選択項目なし

ここで A_2 は社会, 数学, 理科の選択状況に対応した数字であり, 400, 200, 100 を最大一回ずつ含む足し算の結果となっているはずである。逆に A_2 から, 表 2 のように各項目の選択状況が一意に特定できる。

さらに A の三桁目について調べる。 A を商が整数の範囲内となるように 100 で割り算し, そのあまり A'_3 の四桁目（つまり一の位）を切り捨てる。

$$A \rightarrow A'_1 = a_1 \times 10 + a_0 \rightarrow A_1 = a_1 \times 10 \quad (5)$$

こうして得られた A_1 からは, 音楽, 美術, 保健体育の選択状況を特定することができる (表 3)。

最後に, 四桁目の数字について調べる。 A を商が整数の範囲内となるように 10 で割り算し, そのあまりを A_0 とする。

$$A \rightarrow A_0 = a_0 \quad (6)$$

この数字は, 技術, 家庭科, 英語の選択状況に対応する (表 4)。

以上をまとめれば, 得られた符号の各桁についてそれぞれの表に照らし合わせて, 全体の選択状況を特定することができる。例えば $A = 1,101$ の場合は, 国語, 理科, 英語が, また $A = 0,624$ の場合は, 社会, 数学, 美術, 技術が選択されているということになる。

表 4: 四桁目の数字と項目の選択状況

A ₀	選択された項目
7	技術, 家庭科, 英語
6	技術, 家庭科
5	技術, 英語
4	技術
3	家庭科, 英語
2	家庭科
1	英語
0	選択項目なし

卒業要件以外で取得予定の免許について、当てはまるものに○をつけてください。桁ごとに教科の下の数字を加えたものを、所定の欄に記入してください。

国語	社会	数学	理科	音楽	美術	保健	技術	家庭	英語
1	4	2	1	4	2	1	4	2	1
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

例：数学, 理科, 保健体育を選択する場合→0310

国語	社会	数学	理科	音楽	美術	保健	技術	家庭	英語
1	4	2	1	4	2	1	4	2	1
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

図 1: アンケート用紙の例

4 アンケート集計用紙の工夫と表計算ソフトへの実装

本節では、ここまでで説明した符号化および復号化を、紙を媒体としたアンケート用紙および一般的な表計算ソフトを用いて実装する方法について説明する。

図 1 に、アンケート用紙に記載する設問の例を示す。ここでは回答項目に丸を付けることに加えて、それぞれの項目に対応する数字の足し算を回答者に負担してもらい、4 桁の符号を生成してもらっている。また、回答者がこの作業のルールを容易に理解できるように、回答例を併記している。これまでに説明したとおり、今回の符号化の方法では繰り上がりの手順が含まれないため、項目の右端から三つずつ区切って、それぞれを独立して扱うことができる。結果として、回答者に負担してもらう作業としては、4, 2, 1 を最大で一回ずつ含む足し算を 4 回行ってもらうだけであり、作業内容としては容易なものであると考えられる。

次に、得られた 4 桁の符号を表計算ソフト上へ入力して、集計する手法について論ずる。ここで紹介する方

```
=IF(ROUNDDOWN(B2,-3)=1000,1,0)
```

	A	B	D	E	F	
1	氏名	符号	国語	社会	数学	理科
2	Alice	1000	1	0	0	0
3	Bob	0310	0	0	1	1
4	Clare	0000	0	0	0	0

```
=IF(OR(ROUNDDOWN(MOD($B2,1000),-2)=400,
ROUNDDOWN(MOD($B2,1000),-2)=500,
ROUNDDOWN(MOD($B2,1000),-2)=600,
ROUNDDOWN(MOD($B2,1000),-2)=700),1,0)
```

図 2: 表計算ソフトへの入力例

法は、商用の Microsoft Excel の他、無償で利用できる LibreOffice[3] に含まれる Calc などの表計算ソフトを利用して、実装することが可能である。符号を十進数として扱う場合、必要な演算は (i) 割り算のあまりの計算と (ii) 特定の桁以降の切り捨ての二つである。

割り算のあまりは、'MOD' 関数を用いて実装できる。例えば整数 a を整数 b (<a) で割ったあまりは、一般的に a を第一引数、b を第二引数として 'MOD(a,b)' のように記述することで計算できる。それぞれの引数には特定のセルを指定することも可能である。

特定の桁以降の切り捨ては、'ROUNDDOWN' 関数を用いることで実装できる。第一引数に処理したい数字を、第二引数に切り捨てる桁を指定する。第二引数は、0 で小数点第一位を切り捨てるが、-1 で一の位を、-2 で十の位を、-3 で百の位を切り捨てる。

以上の関数を用いて取り出した各桁の数字を、'IF' 関数を用いて評価する。例えば一番目の桁 (千の位) が 1 であるか 0 であるかを場合分けすることにより、国語が選択されているか否かを判定することができる。また二桁目の数字では、例えば社会が含まれている場合は、4, 5, 6, 7 のいずれかの場合であるから、条件式は 'OR' 関数を用いて実装している。他の項目についても、同様に評価することができる。

図 2 に、表計算ソフトへの入力の例を示す。セルに入力する書式の先頭の = (イコール) 記号は、そのセルに自動計算させることを表している。また入力の便宜上、社会の項目 (D2 セルなど) を数学や理科、それ以降の項目に対応するセルにコピーしながら入力することも考慮に入れて、セルが絶対参照されていることに注意されたい。

図 1 のアンケート用紙上で回答者が生成した符号を、表計算ソフト上に集計者が入力することにより、項目の選択状況が入力される仕組みになっている。

5 まとめ

本稿では、アンケートにおける複数回答可である設問の集計方法を提案した。今回の方法では、紙を媒体としたアンケートを実施することを想定するが、集計を効率的に行うため、回答状況に対応する符号の生成を、回答者自身に行ってもらおう形式とした。回答者に負担してもらおう作業としては、1, 2, 4 を最大で一回ずつ含む足し算であるが、これは繰り返り上がりの手順が含まれないため、暗算でも間違いにくい容易な操作であるとみなしてよい。

また得られた符号を一般的な表計算ソフトへ入力することで、その選択状況を自動的に入力するための関数の実装例について述べた。アンケートの回答者ごとに本手法を適用するが、最終的にまとめて全体の集計を行うことは容易である。

今回は特に、高知大学教育学部における、卒業要件以外に習得する中学校免許状の種類についての設問を例にとったが、他の一般的な複数回答可とする設問についての集計に使うこともできる。今回の方法では、各桁を3つの項目に当てているため、 k 個の項目に対しては、 $k \leq 3^n$ を満たす最小の n だけの桁の符号を扱う必要がある。今回の例では $k = 10$ であるから、 $n = 4$ であるが、12 個までの項目についてのアンケートであれば、同じく 4 桁の符号で処理することができる。しかしこれ以上の項目を含む場合は、得られる符号の桁数が増加し、集計作業はやや煩雑になるといわざるを得ない。その場合は設問自体を分割するなど、さまざまな工夫をすることも合わせて検討すべきであろう。

謝辞

本研究は平成 25～27 年度高知大学教育学部学部長裁量経費の支援を受けて行われている。

参考文献

- [1] 高知大学教育学部エビデンス研究会, 「2015 年度版 高知大学教育学部データブック」(2015 年)
- [2] 山口 和紀 監修, 「The UNIX Super Text 上・下」, 技術評論社 (1992 年).
- [3] LibreOffice の公式サイト :
<https://ja.libreoffice.org/>