佛教大学社会学部論集　第 60 号（2015 年 3 月）

四捨五入した％の合計が 100％にならないとき

山　口　　　洋

〔抄　録〕

相対度数分布表の四捨五入した％の加算値が 100 にならないとき，各カテゴリーおよび合計欄の％をどう表示すべきかが検討された。各カテゴリーの丸めた％を調整・変更して合計を 100 にすべしとする社会統計の教科書もあるが，そうした調整は必要最小限度であっても真の％の大小関係を歪曲する可能性があるため，特別な理由が無いかぎりしない方がよいと結論された。また相対度数分布表の合計欄の「100％」は，その欄の度数を分母に％が計算されたことを読み手に伝える役割を果たすから，丸めた％の加算値が 100 にならなくても常に「100％」と記した方が，読み手には分かりやすいと主張された。また合計を 100 とすべき特別な理由があるときは，いくつかの教科書で推奨された相対誤差を最小化する調整方法よりも，絶対誤差を最小化する方法が四捨五入の原理との一貫性の点で望ましいと結論された。ただし相対度数分布表の通常の使用目的や現在のコンピュータによる作成過程を考えれば，合計を 100 にすべき「特別な理由」は考えにくいと主張された。

キーワード　相対度数分布表，パーセンテージ，四捨五入，合計

1. 序　　論

相対度数分布表における真のパーセンテージ（以下％と略す）の合計が必ず 100 になることは，社会統計学を学ぶ者なら誰でも知っている。しかし現実データから作成された相対度数分布表の％のほとんどは四捨五入された％であり，その合計値は必ずしも 100 にはならない。このとき我々は各カテゴリーの％および合計欄の％をどう表示すべきだろうか？社会統計学を扱う教科書の中には，この事態について本文中で明確に言及しているもの，および明確に言及はしないが丸めた％の合計が 100 にならない分布表を掲載しているものがある。これらを検討すると，その対応法は大きく分けて 3 種類あることがわかる。

第 1 に，丸めた％の合計がちょうど 100 にならないときには，各カテゴリーのうちどれか

の丸めた％を修正・変更（以下「調整」という言葉でこれを表現）することによって合計をちょうど 100％ にすることを推奨する教科書がある。たとえば安田・原（1982： 237 頁）は「研究者が手許におく資料としてはそれでよいが，公表するばあいには合計が 100％ となるよう，調整をしなければならない」と述べる。その理由について安田・原（1982）は何も述べないが，安田（1969： 17 頁）は同様の趣旨を述べた箇所で「発表する際には，合計欄が 100.0％ とそろわないと見苦しいから」としている。また，西平（1985： 68 頁）は「ミスプリントでないことを示すために」自分が同様の調整を行っていることを認め，その例として表 1 の相対度数分布表を挙げている。

表 1 で調整が施されているのは「50 歳代」のカテゴリーである。度数から計算すると，このカテゴリーの四捨五入した％は本来 41.1％（≒100×(210/511)＝41.09589…）になる。ところが，このまま全カテゴリーの丸めた％を合計するとその値は 100.1％ になるので，「50 歳代」の％を 41.0％ に切り下げ，％の加算値をちょうど 100.0％ に調整している。第 3 節で述べるように安田・原（1982），西平（1985）は，調整を行う場合には表 1 のように度数の最も大きいカテゴリーで行うのがよいとする。上記の他にも合計を 100％ にすべく各カテゴリーの％を調整することを勧める教科書がいくつか存在する(1)。また本文中では勧めていないものの，明らかに調整が施されたとわかる相対度数分布表を掲載している教科書もある(2)。

表 1　相対度数分布表
（小数第 2 位以下四捨五入）

衆議院議員の年齢（1980 年）	度数	％
30 歳代	28	5.5
40 歳代	104	20.4
50 歳代	210	41.0
60 歳代	117	22.9
70 歳代	49	9.6
80 歳代	3	0.6
合計	511	100.0

表 2　相対度数分布表（N＝1473）
（小数第 2 位以下四捨五入）

信仰する宗教	％
プロテスタント	63.3
カトリック	25.5
ユダヤ教	1.8
信仰する宗教なし	7.3
その他の宗教	1.4
無回答	0.8
合計	100.1*

*まるめのために総計が 100％ にならない。

表 3　相対度数分布表
（小数第 2 位以下四捨五入）

50 m 走のタイム	度数	％
7 秒台	2	0.1
8 秒台	306	19.8
9 秒台	772	49.8
10 秒台	384	24.8
11 秒台	72	4.6
12 秒台	8	0.5
13 秒台	3	0.2
14 秒台	2	0.1
合計	1549	100.0

出典：表 1：西平（1985： 67 頁）より作成。数値はそのまま引用。
表 2：Bohrnstedt & Knoke（1988＝1990： 28 頁）より作成。数値と注記はそのまま引用。
表 3：日本統計学会（編）（2012 : 90 頁）より作成。数値はそのまま引用。

第 2 に各カテゴリーの丸めた％をそのまま示し，それらの加算値（99.9, 100.1 など）を合計欄に記している教科書もある。Bohrnstedt & Knoke（1988＝1990）に掲載された表 2 がその例である。このような分布表を掲載している教科書(3)は他にも散見されるが，表 2 のように合計が 100％ にならない理由を欄外に注記するものもあれば，しないものもある。

第 3 に最近の多くの教科書では，各カテゴリーの丸めた％をそのまま示し，それらの加算

値がたとえ 100％ にならなくても，合計欄には「100％」「100.0％」「100.00％」…などと表記している。表 3 がその例である。表 3 の各カテゴリーの丸めた％の加算値は 99.9％ であるが，合計欄では真の％の合計値である 100.0％ を記している。

このように，丸めた％の合計がちょうど 100 にならない場合の対応は，様々な教科書の間で明らかに一致していない。我々は一体どれをお手本にして相対度数分布表を作成したらよいのか？表 1～3 のようなケースが，ごく希にしか生じない現象であれば，この事態に対してアッドホックに対応していくことも許されよう。しかし，こうした現象は決してレアケースではなく，むしろ日常茶飯事と言ってよいことが既に明らかにされている。

Mosteller, Youtz, & Zahn（1967）および Diaconis & Freedman（1979）によれば，総度数が比較的大きく，カテゴリーの数が 3 個の相対度数分布表の約 4 分の 1 は，丸めた％の合計が 100 にならない。同じくカテゴリーの数が 4 個の場合，同様の表の約 3 分の 1 は，丸めた％の合計が 100 にならない。そしてカテゴリー数(ℓ)が増大するにしたがって，合計が 100 にならない確率は増大し，ℓ が十分大きければ，その確率は「$1-\{6/(\pi\ell)\}^{(1/2)}$」で近似できるとされている。この式に $\ell=47$，すなわち日本の都道府県の数を代入すると，都道府県別の相対度数分布表において，上記の問題が生じる確率はなんと約 79.8％ になる。この場合，丸めた％の合計がちょうど 100 になってくれることの方が，むしろ幸運な偶然である。

常識的に言えば，こうした日常茶飯事に対してアッドホックな対応は許されない。何か統一的な指針が必要だと思われる。では本当のところ我々はどう対応すればよいのだろうか？

本稿はこの問題について基本的に「表 3 の形式でよい」と結論したい。まず表 1 のような合計を 100％ にするための調整は特別な理由が無いかぎり行わない方がよい。第 4 節でみるように，そうした調整は（たとえ必要最小限度のものであっても）％の大小関係を歪曲する危険を免れないからだ。また合計欄には丸めた％の加算値ではなく，真の％の加算値である「100％」を常に記してよいであろう。第 5 節で述べるように合計欄の「100％」という表示は各カテゴリーの構成比の分母を読み手に示す機能を果たす（クロス表ではこの機能が特に重要）。この機能を重視すれば情報価値に乏しい丸めた％の加算値を示すよりも，真の％の加算値である「100」を常に記した方が読み手には親切である。丸めた％の加算値が 100 にならないときは，必要に応じて（また可能なら）分布表の欄外にその旨，記せばよいだろう。

ただし，しかるべき理由で合計をジャスト 100％ にする必要がある場合には，第 4 節で記す不合理を覚悟の上で，何らかの方法で各カテゴリーの％を調整することになる。本稿は第 3 節で安田・原（1982），西平（1985）が示す方法の問題点を明らかにし，対案として「最小絶対誤差法」を示した。ただしこの方法を用いるとしても，それは合計を 100％ とすべき「特別な理由」がある場合に限られる。そして第 6 節では，相対度数分布表の一般的な用途および今日におけるコンピュータでの作表過程を考えると，誰もが納得しうる「特別な理由」は想定困難であると主張したい。

なお丸めた％の合計が100にならない現象は，議席の人口比例配分の際に生じるやっかいな問題（Balinski & Young, 1982＝1987，一森，2006），つまり整数に丸めた各州（各国）の配分議席の合計値が所定の全議席数に一致しなくなる問題と，数学的構造は基本的に同じである(4)。よって本稿は議席の人口比例配分の方法に関する研究を適宜参考にしつつ論を進める。ちなみに，本稿が第3節で示す最小絶対誤差法は Balinski & Young（1982＝1987）が「ハミルトン方式」と，一森（2006）が「最大剰余法」と呼ぶ人口比例配分の方法をそのまま応用したものである。

2. 基 本 概 念

本稿は相対度数分布表の各カテゴリーの構成比を四捨五入した％で表示する方法について論じる。ここではそのための基本概念を定義したい。相互に排他的なカテゴリーの総数を ℓ とし，カテゴリー番号を i とする（$i = 1, \ldots, \ell$）。各カテゴリーの度数（絶対度数）を n_i，総度数を N とするとカテゴリー i の構成比，すなわち％は $p_i = 100 \times (n_i / N)$ である。これを本稿では「真の％」と呼ぶ。真の％の小数第 m 位以下を切り捨てた値を d_i と表記し，本稿では「切り捨て値」と呼ぶことにする。同じく真の％の小数第 m 位以下を切り上げた値は u_i と表記し，「切り上げ値」と呼ぶことにする。

・本稿の基本概念の呼称・記号と数値例

真の％　　：p_i……例：$p_3 = 15.315\cdots$％（カテゴリー3の真の％）

切り捨て値：d_i……例：$d_3 = 15.3$％（p_3 の小数第2位以下を切り捨てた値）

切り上げ値：u_i……例：$u_3 = 15.4$％（p_3 の小数第2位以下を切り上げた値）

丸め後の％：q_i……例：$q_3 = 15.3$％（p_3 の小数第2位以下を四捨五入した値）

調整後の％：r_i……例：$r_3 = 15.2$％（合計を100とするため q_3 を調整した値）

・四捨五入の原理

$p_i - d_i < u_i - p_i$ ならば $q_i = d_i$

$p_i - d_i \geqq u_i - p_i$ ならば $q_i = u_i$

・単位：真の％の小数第 m 位以下を四捨五入した％の最小単位

$= (1/10^{m-1})$ ％

p_i の小数第 m 位以下を四捨五入した％を q_i と表記し，「丸め後の％」と略称する。なお本稿では特に断らないかぎり「丸め」を四捨五入と同義に用いるが，正確に言えば四捨五入は広い意味での丸めの方法のひとつに過ぎない（次節参照）。周知のように「四捨五入する」とは，上述の切り捨て値と切り上げ値のうち，真の％との差が小さい方に丸めることを指す。なお本

稿では，真の％が切り捨て値と切り上げ値のちょうど真ん中にあるとき，すなわち小数第 m 位が「5」で $m+1$ 位以下に数字が無いときは，一律に切り上げるルールを採用する(5)。

そして q_i の合計（$\sum_i q_i$）が 100 にならないとき，合計を 100 にする何らかの調整を経た後のカテゴリー i の％を r_i と表記し「調整後の％」と呼ぶ。こうした調整を行う場合，カテゴリー i の％が変更される場合もあれば，変更されない場合もある。変更された場合は $r_i \neq q_i$，変更されなかった場合は $r_i = q_i$ となる。

なお真の％の小数第 m 位以下を四捨五入すると，丸め後の％は（$1/10^{m-1}$）％の整数倍として表される。たとえば真の％の小数第 2 位以下を四捨五入するとき，q_i =53.1% ならば，それは 0.1% の 531 倍であり，3.7% ならば 0.1 の 37 倍である。そこで，この（$1/10^{m-1}$）％のことを，本稿では「四捨五入の単位」または略して「単位」と呼ぶことにする。たとえば小数第 1 位以下を四捨五入して整数部分のみの％に丸めるときの単位は 1% であり，「2 単位」とは 2%，「3 単位」とは 3% のことである。小数第 3 位以下を四捨五入するなら単位は 0.01%，「2 単位」は 0.02%，「3 単位」は 0.03% を表す。

3. 丸めた％の合計を 100% にする調整方法

3.1　先行文献の方法とその問題点

本節では丸めた％の合計が必ず 100% になるように調整したい場合（どんな場合が想定可能かは結論部で改めて論じる），どんな調整の方法が妥当なのかを考える。過去の社会統計の教科書は，丸めた％の修正・変更をどのカテゴリーで行うかに関して，大きく分けて 2 種類の指針を示す。第 1 の指針は，各カテゴリーの内容からそれを判断するもので，西平（1985：68 頁）や辻・有馬（1987：148 頁）は「その他」や「わからない」などの重要でないカテゴリーで調整を行うのがよいとする。しかし表 1 や表 3 のようにその種のカテゴリーが存在しない相対度数分布表もあるから，この方法は汎用的なものとは言えない。第 2 の指針は，各カテゴリーの度数や％の数値からそれを判断するものであり，これはどんな相対度数分布表にも適用可能である。以下，本節ではこうした汎用的方法にしぼって検討したい。

安田・原（1982）の調整方法は，表 4 のように「この調整によって生ずる相対誤差が最小になるようなところ，換言すればもっとも数字の大きいカテゴリーで行う」（安田・原 1982：237 頁）というものである。また西平（1985：68 頁）も「‘その他’のような重要でない」カテゴリーがあればそこで調整を行うが，それが無ければ，「誤差が相対的にいちばん小さくなるように，いちばん大きな数字のところでする」と述べ，安田・原（1982）と同様の方法を勧める。

安田・原（1982），西平（1985）は相対誤差の概念を明確に定義していない。しかし，両テキストとも「数字の大きい」カテゴリーで調整せよ，としていることから，相対誤差は丸め後

表 4　合計を 100％ とする調整の例（安田・原，1982：237 頁より作成）

カテゴリー（i）	度数（n_i）	真の％（p_i）	丸め後の％（q_i）		調整後の％（r_i）
1	70	47.619…％	47.6	⇒	47.7
2	54	36.734…	36.7		36.7
3	23	15.646…	15.6		15.6
計	147	100	99.9		100.0

小数第 2 位以下を四捨五入

の％（q_i）と調整後の％（r_i）の比を用いて定義できそうである。本稿はこの概念を「相対誤差Ⅰ」として下のように定義した。比を用いた定義の仕方は他にも考えられるが，それは注 7 に記した。下の相対誤差Ⅰは，丸め後の％と調整後の％のうち大きい方を分子に小さい方を分母にとった値で，そのカテゴリーで全く調整をしなければ最小値 1 をとり，丸め後の％と調整後の％が乖離すればするほど，その値は大きくなる。誤差の大きさをこう定義すれば，どこかのカテゴリーの丸め後の％を 1 単位変更するとき，度数が最大のカテゴリーでそれを行えば，安田・原（1982）と西平（1985）が述べるとおり誤差が最も小さくなる(6)。

・調整後の％の相対誤差Ⅰ＝丸め後の％と調整後の％の比

$= \max(r_i, q_i)/\min(r_i, q_i)$

・調整後の％の相対誤差Ⅱ＝真の％と調整後の％の比

$= \max(r_i, p_i)/\min(r_i, p_i)$

$\max(x, y)$・・・x と y のうち大きい方の数，$\min(x, y)$・・・同じく小さい方の数

ただしカテゴリーの数が増えれば，丸め後の％の合計値が 100.2 とか 99.7 といったように，100 から 2 単位以上乖離する可能性が出てくるが，この場合の対処について安田・原（1982）と西平（1985）には言及がない。しかし 1 カテゴリーにつき 1 単位のみの修正を認めるならば（第 4 節で述べるように％の修正は最小限にとどめるべきなので），やはり度数の最も大きいカテゴリーから順に，合計した％が 100 になるまで調整を繰り返すことにより，表全体における相対誤差Ⅰの総計は最も小さくなるはずである。

しかしこの相対誤差Ⅰには難点がある。この概念は誤差の大きさを丸め後の％を基準に定義するが，その基準自体が誤差を含む値だからである。本来，誤差の大きさは真の％を基準に定義すべきであろう。すなわち相対誤差は調整後の％と真の％の比で定義した方がよい。上に示した相対誤差Ⅱがそれである。ただし相対誤差Ⅱを最小化するには，各カテゴリーを調整した場合の相対誤差Ⅱを逐一計算する必要があり，安田・原（1982）と西平（1985）が示した方法よりもずっと手間がかかる。

しかしⅠにせよⅡにせよ，こうした相対誤差の概念は四捨五入の原理それ自体と矛盾している。言い換えれば，ある数値に含まれる誤差を正しい数値との「比」で定義する考え方は，四捨五入という丸めの方法が前提とする誤差の捉え方と矛盾しているのである。

第2節で示した定義から明らかなように，四捨五入の原理は，真の％と丸め後（未調整）の％の「比」ではなく「差」によって誤差を定義し，それを最小化する方法と言ってよい。つまり，四捨五入という手続きは誤差を「相対的」にではなく「絶対的」に捉え，それを最小化するものである。丸め後の％の絶対誤差および相対誤差を，数式で示せば以下のとおりである。

・丸め後の％の絶対誤差と相対誤差

絶対誤差 $= \max(q_i, p_i) - \min(q_i, p_i)$　　相対誤差 $= \max(q_i, p_i) / \min(q_i, p_i)$

両者の違いを具体例で説明しよう。たとえば真の％が1.45% だったとする。四捨五入によって小数点以下を四捨五入すれば，丸め後の％は2% でなく1% になる。これは「1.45－1＜2－1.45」という事実に基づいている。つまり2% に丸めるよりも1% に丸めた方が，絶対誤差が小さくなるという事実に基づいている。ところが，この例では1% に丸めるよりも2% に丸めた方が相対誤差は小さくなる。1.45は1の1.45倍，2は1.45の約1.38倍であり1.45＞1.38だからである。このように相対誤差を最小化するという考え方と四捨五入の原理は矛盾する。

したがって，合計を100% に調整する際，相対誤差を小さくすることを目指すならば，そもそも最初から丸めの方法として四捨五入ではなく，別の特殊な方法を採用した方がよい(7)。この特殊な方法の検討は本稿の射程を超えるので，これ以上深入りはしない。しかし，少なくとも四捨五入という丸めの原理を採用するなら，誤差を「比」ではなく「差」で定義し，それを最小化する調整の方法を採る方が，方法論的な一貫性の点で望ましい。

3. 2　最小絶対誤差法

そこで丸め後の％を100% にしたい場合の調整方法として，絶対誤差の合計を最小化する方法を下に示した。本稿はこの方法を「最小絶対誤差法」と呼んでおく。なお，ここでも1カテゴリーにつき1単位の修正のみを認めるルールを採用している。また，この方法は本稿の独創ではなく，慣習的な比例配分の方法を単に応用したものである。序論で述べたように，議席の人口比例配分の問題を扱う文献において，この方法はハミルトン方式（Balinski & Young, 1982＝1987），最大剰余法（一森，2006）などと呼ばれている。

・最小絶対誤差法（小数第 m 位以下を四捨五入する場合）

（1）まず全カテゴリーの真の％の小数第 m 位以下を一律に切り捨てて合計する。つまり各カテゴリーの切り捨て値の合計を求める。この値は必ず 100 未満となる。

（2）切り捨てた端数（小数第 m 位以下）が最も大きいカテゴリーから順番に，切り捨て値に 1 単位（$1/10^{m-1}$）を足す操作を施す。

（3）上の（2）の操作を合計が 100％ に達するまで繰り返す。

このように全カテゴリーの％の端数を一旦切り捨て，端数の大きいものから順に切り上げ操作を施す方法をとれば，調整後の絶対誤差の合計値（$\sum_i \max(r_i, p_i) - \min(r_i, p_i)$）が最小となるのは明らかである。この方法のメリットは，四捨五入の原理を自然に拡張した調整方法だという点に求められよう。この方法の適用例を表 5 に示したので参照されたい。

ただし，このような調整を行う必要がある場合でも，実際にそれを行った場合には，表の読み手に調整箇所が分かるような注記が必要かもしれない。たとえば，安田・原（1982：237 頁）は「調整した数字には＊印を打つことが望ましい」としている。

表 5　最小絶対誤差法による調整の例（仮想例：小数第 1 位以下四捨五入）

カテゴリー (i)	度数 (n_i)	真の％ (p_i)	丸め後の％ (q_i)	切り捨て値 (d_i)	切り捨てた端数の大きさ（順位）	調整後の％ (r_i)
1	971	20.994…％	21％	20％	1	21％
2	950	20.540…	21	20	4	21
3	795	17.189…	17	17	7	17
4	721	15.589…	16	15	3	16
5	579	12.518…	13	12	5	12*
6	347	7.502…	8	7	6	7*
7	262	5.664…	6	5	2	6
計	4625	100	102	96	–	100

*は調整箇所

4. 合計を 100％ にするリーズナブルな方法は存在するか？

4.1　リーズナブルであるための規準－必要最小限，同一性，整合性，単調性－

前節では丸めた％の合計が 100％ になるように調整する方法について論じた。しかし，こうした調整を推奨する安田・原（1982：237 頁）は「ただし，比率構成の全体が問題になるのでなく，特定のカテゴリーの構成比だけが問題になっているばあいは，少なくてもそのカテゴリーにおいて調整することは，避けねばならない」とも述べ，調整によって何らかの問題が生じる可能性を暗に認めている。

実際，各カテゴリーの丸めた％を調整して合計を 100％ にする完全にリーズナブルな方法

は，前節で述べた諸方法を含めて存在しない。このことを示すため，本節ではまず「リーズナブルな方法」と言えるための 4 つの規準を提示する。そしてこの 4 つの規準を常に，すなわちいかなる相対度数分布表（クロス表を含む）においても常に満たす方法は存在しないことを示す。

まず，そうした調整は「必要最小限」にとどめなくてはならない。各カテゴリーの％の正確性を大きく損なう形で％の合計を 100 にするのは本末転倒だからである。そこで本稿は「必要最小限の規準」として，以下の数式で表現・定義されるルールを定めておく。

・規準（1）必要最小限の規準（小数第 m 位以下を四捨五入する場合）

$\sum_i q_i = 100$ ならば $r_i = q_i$

$\sum_i q_i > 100$ ならば $r_i = q_i - (1/10^{m-1})$ または $r_i = q_i$

$\sum_i q_i < 100$ ならば $r_i = q_i + (1/10^{m-1})$ または $r_i = q_i$

この規準の含意を具体的に説明しよう。第 1 に，この規準によれば小数第 m 位以下を四捨五入する場合，必要に応じて，ひとつのカテゴリーにつき 1 単位の変更だけが許される。たとえば小数第 1 位以下を四捨五入して，整数部分のみの％に丸める場合には，必要に応じて 1 カテゴリーにつき 1% の変更だけが許される。第 2 に，この規準によれば丸めた％の合計値が 100 と s 単位異なるときは，s 個のカテゴリーの変更だけが許される。たとえば整数部分のみの％に丸める場合，丸め後の％の合計が「102」であれば，ふたつのカテゴリーの％を 1 ％ずつ切り下げることで調整するものとし，これ以外の調整法，たとえばあるカテゴリーを 1 ％切り上げ，他の 3 個のカテゴリーを 1% ずつ切り下げるといった方法は許されない。ちなみに前節で示した最小絶対誤差法はこの規準（1）を満たす。

また合計を 100% とするための調整は，真の％の間の「大小関係」をできるだけ維持・保存するものでなくてはならない。なお％の大小関係の比較には主にふたつのタイプがあり，本稿では表 6 で示したように，それらを「タイプ 1」「タイプ 2」と呼んで区別する。

タイプ 1 の比較は，ある時点（または集団）での度数分布表において，ふたつのカテゴリーの構成比（％）を比較することである。ただしタイプ 1 の比較だけなら％を用いずに絶対度数で行うことも可能である。一方，タイプ 2 の比較は複数時点（集団）において，あるカテゴリーの％を

表 6　％の比較（小数第 2 位以下を四捨五入）

カテゴリー	時点（または集団）	
	1	2
a	↕ 29.8%	⟷ 25.1
b	52.3	43.5
c	17.9	31.4
合計 （N）	100.0 （413）	100.0 （577）

↕…タイプ 1 の比較　　⟷…タイプ 2 の比較

比べることである。時点（集団）ごとに総度数が異なるとき，この種の比較は相対度数（％）を用いないとできないから，％を示すことの意義はこの比較にあると言ってよい。我々は表 6 のような一連の度数分布表を通常，クロス表と呼ぶ。そこでクロス表の用語で表現すれば，タイプ 2 の比較は，列計（行計）を 100％ として計算した％を行方向（列方向）に比較することに相当する。

調整後の％の間の大小関係（タイプ 1 および 2 の）は，真の％のそれをできるかぎり維持・保存すべきである。本来（微妙に）異なる真の％が丸めにより同じ値になることは避けられないにせよ，本来同じものが違ってしまうこと，また本来の大小関係が逆転することは最低限回避したい。丸め後（未調整）の％では，そうしたことは起こり得ないからである。

こう考えると，丸め後の％の合計を 100％ にするための調整方法は以下の（2）～（4）の規準を「常に」，すなわちいかなる度数分布表・クロス表においても満たすべきである。手短に言えば，（2）は本来同じものが違ってしまってはいけないことを，（3）と（4）は真の大小関係を逆転させてはならないことを述べている。（2）（3）は表 6 のタイプ 1 の比較に，（4）はタイプ 2 の比較に関わる。

・規準（2）同一性の規準

ある相対度数分布表のふたつのカテゴリーの真の％（構成比）が同一ならば，それらの丸め・調整後の％も同一でなければならない。

・規準（3）整合性の規準

ある相対度数分布表のカテゴリー i の真の％がカテゴリー $h(i \neq h)$ のそれを上回るならば，丸め・調整後の％において i が h を下回ってはならない。

・規準（4）単調性の規準

時点（集団）j におけるカテゴリー i の真の％が，時点（集団）$k(j \neq k)$ におけるそれを上回るならば，時点（集団）j における i の丸め・調整後の％が，時点（集団）k におけるそれを下回ってはならない。

4.2　同一性規準の不満足

度数の等しいカテゴリーが複数あり，それらの％のすべてではなく，いずれかのみを調整しなくてはならないとき，必要最小限の規準を満たすどんな調整法を採用しても，同一性の規準は満たされない。

そんな場合の簡単な例として，3 カテゴリーの度数分布表で，全カテゴリーの度数が同じだった場合が考えられる。この場合，真の相対度数（％）は全カテゴリーで 33.33…％になり，％の小数第 2 位以下を四捨五入すると，全カテゴリーとも 33.3％ となって，合計は 99.9％ になる。ここで必要最小限の規準を満たす調整を行うならば，3 カテゴリーのいずれかを 33.4％

とするしかないが，いずれも同一性の規準に反する。ちなみにこの場合，小数第何位以下を四捨五入したとしても合計は 99.9…9% になり，調整を行えば，やはり必ず同一性規準に反する結果になる。

6 カテゴリーの度数分布表で，全カテゴリーの度数が等しかった場合にも同じことが起きる。このとき，全カテゴリーの真の％は 16.66…％となる。ここで小数第 2 位以下を四捨五入すれば，全カテゴリーにおいて 16.7% となり，その合計は 100.2% になる。必要最小限の規準に従って合計を 100 にするには，どこか 2 つのカテゴリーを 16.6% にせねばならないが，そうすれば同一性規準に反する。この場合も，小数第何位以下を四捨五入したとしても，合計は 100 にならず，合計を 100 とする調整を行えば必ず同一性規準に反する結果となる。

上のようなケースが現実データで生じるとすれば，総度数がごく小さい場合に限られよう。しかし教育用の仮想例として全カテゴリーの度数が等しい分布表を用いることは十分考えられる。この種の度数分布表の作成に際し，教員が丸め後の％の合計を 100 とする調整を奨励したならば，学生から次のような質問が出るのは必至である。「真の％はどのカテゴリーも同じなのに，わざわざそのいくつかを別の％にするような調整は本当に必要なのですか」と。本稿の答えは特別な理由が無いかぎり「必要なし」である。

4.3　整合性および単調性規準の不満足

どんな場合でも整合性規準を満たし，同時に必要最小限の規準も満たすような調整法は存在する。たとえば丸め後の％の合計が 100 を超える場合，真の％が最も小さいカテゴリーから順に丸め後の％を 1 単位ずつ切り下げる操作を行えば，真の％の大小関係を逆転させずに，つまり整合性規準を満たしつつ合計を 100% にできる。丸め後の％の合計が 100 を下回るなら，真の％が最大のカテゴリーから順に，丸め後の％を 1 単位切り上げる操作を合計が 100 になるまで繰り返せばよい。この調整法を採れば必要最小限の規準を満たしつつ，整合性だけは常に満足される。

しかし，どんな場合でも整合性規準と単調性規準の両方を満たすような調整方法は，必要最小限の規準を満たす調整方法の中には存在しない。たとえば，表 7 および 8 のような相対度数分布表（クロス表）の場合，必要最小限の規準を満たすいかなる調整方法を採用しても，整合性と単調性のどちらか一方に必ず反する結果となる。

表 7　真の％

カテゴリー	時点 1		時点 2		時点 3
a	44.01%	＞	43.65	＞	43.41
b	43.51	＜	43.55	＜	43.88
c	12.48	＜	12.8	＞	12.71
計	100		100		100

表 8　小数点以下を四捨五入した％

カテゴリー	時点 1	時点 2	時点 3
a	44%	44	43
b	44	44	44
c	12	13	13
計	100	101	100

表7は真の％の推移を示す架空例である。表7の小数点以下を四捨五入すると表8のようになり，時点2の合計は「101％」となる。必要最小限の規準を満たしつつ，時点2の％の合計を100とする調整の方法は下の表9〜11の3種類しかない。ちなみに，安田・原（1982）と西平（1985）の言う最大カテゴリーにおける調整を行えば結果は表9になり，本稿が示した最小絶対誤差法によれば調整結果は表10のようになる。

表9　調整後の％（1）

カテゴリー	時点1	時点2	時点3
a	44％	43	43
b	44	44	44
c	12	13	13
計	100	100	100

しかし表9の調整の仕方では整合性規準が満たされない。時点2の真の％（表7）を見ると，aがbを上回っているにもかかわらず，調整後の％（表9）ではbがaを上回ることになり，大小関係が逆転するからである。

また，表10の調整の仕方では，単調性規準が満たされない。カテゴリーbの真の％（表7）は時点1から2にかけて上昇しているにもかかわらず，bの調整後の％（表10）は，時点1から2にかけて下降しているからである。同様に，表11の調整の仕方では，単調性が満たされない。カテゴリーcの真の％（表7）は時点2から3にかけて下降しているにもかかわらず，cの調整後の％（表11）は時点2から3にかけて上昇しているからである。

このように表7，表8のような場合，合計を100％とする必要最小限の調整をどんな形で行ったとしても，整合性と単調性のいずれか一方が満たされず，％の大小関係が事実と逆になることを避けられない。

表10　調整後の％（2）

カテゴリー	時点1	時点2	時点3
a	44％	44	43
b	44	43	44
c	12	13	13
計	100	100	100

表11　調整後の％（3）

カテゴリー	時点1	時点2	時点3
a	44％	44	43
b	44	44	44
c	12	12	13
計	100	100	100

もちろん，上の例は筆者が説明のため意図的に作成したものである。したがって「このような不運が実際に生じることは滅多にないから，できるかぎり調整は行うべきだ」との意見があるかもしれない。本稿は，そうした意見に対して2点反論を述べておこう。

第1に，表7の時点の数は3だが(8)，毎年あるいは毎月，同様のフォーマットで行われる

調査の場合，同じカテゴリー構成の度数分布表を，数多くの時点で作成することが可能であり，それら多数の時点間で時系列の比較を行うならば，表7のような事例が生じる可能性はかなり高まるはずである。第2に，表7のカテゴリーの数は3だが，カテゴリーの数が増せば冒頭で述べたように丸めた％の合計が100にならない可能性が高まるだけでなく，複数のカテゴリーの丸めた％が同じ値をとる可能性（表8のように）も高まるだろう。もちろん無難な調整を行いうるカテゴリーの数も増えるから，表7のようにどうにもならない事例が増えるか否かは即断できない。しかし多くのカテゴリーを持つ分布表で，整合性・単調性を満たす必要最小限の調整法を見つける作業は，確実な方法が存在しないだけにやっかいなものになろう。加えて多数の時点が存在すれば，その作業は手に負えないものになるだろう。

このように，四捨五入した％の合計が100にならないとき，これを100とする必要最小限の調整を行うならば，どんな方法を用いても真の％の大小関係を歪曲する可能性がある。したがって合計を100％としなくてはならない特別な理由が無いかぎり，このような調整は行うべきではない。また特別な理由があって，何らかの方法（たとえば第3節で示した方法）で調整を行うとしても，その結果得られた％には本節で述べた問題が生じる可能性があることを覚悟すべきである。そして相対度数分布表のごく普通の使用目的や，現在のコンピュータによる作表過程を考えると，四捨五入した％を捻じ曲げてまで合計を100％に揃えねばならない理由はほとんど存在しないのではないだろうか。この点は，本稿末尾でより詳しく論じたい。

5．四捨五入した％の一般的表示方法

以上の考察をふまえ，本節では，相対度数分布表における四捨五入した％の合計が100にならないとき，それをどう表示して公表すべきかについての見解を述べる。合計を100％としなくてはならない特別な理由がある場合については第3節で既に述べた。以下では，そうした理由が特に存在しない場合についてまとめておこう。ポイントは3点である。

第1に，四捨五入した％の合計が100％にならなくても各カテゴリーの％はそのまま表示すればよい。理由は前節で述べたとおりである。

第2に，四捨五入した％の加算値が100％にならなくても，相対度数分布表の合計欄には常に「100％」と記してよいと考える。その理由は，我々が合計欄に100％と記す意味を考えることで明らかになる。そもそも各カテゴリーの真の％の合計が100％になることは記すまでもなく周知の事実である。しかし合計欄の「100％」を見て，表の読み手は各カテゴリーの構成比（％）が，その合計欄の総度数を分母にして計算されたことを知る。

合計欄に「100％」と記すことのこうした役割は，クロス表の場合に特に重要となる。クロス表の場合，各カテゴリーの％は行側の周辺度数を分母にしても，列側のそれを分母にしても求められるので，我々は行側と列側のどちらの合計欄に「100％」表示があるかを見て，各カ

テゴリーの％がどちらの周辺度数を分母にして求めたのかを確認する。合計欄の「100％」の機能をこう理解すれば，冒頭に掲げた表 2 のように，合計欄に丸めた％の加算値を「100.1％」などと表示するのは，読み手を混乱させるだけではないだろうか。そもそも合計欄の「99.8％」「100.1％」といった数値は読み手には情報価値がほとんど無い。真の％の合計は常に正確に「100％」だからである。したがって，その正確な値（100％）を合計欄に記すことで，その欄の総度数が構成比の分母であることを，表の読み手に分かりやすく示すことが重要であろう。

合計欄は一律に「100％」でよいと考える理由を％表示の基本的目的に照らしてもうひとつ挙げよう。そもそも％とは「総度数を仮に 100 に揃えたとすると各カテゴリーの度数は幾らに相当するか」を表すものである。したがって表 8 のように，クロス表の合計欄に様々な数値が並ぶことは，「総度数を仮に 100 に揃える」という％表示の趣旨に反するのではないだろうか。

第 3 のポイントは，丸めた％を加算した値が 100％ にならないとき，分布表の欄外にそのことを注記するか否かであるが，これはケースバイケースだと考える。なお，ここで言う「注記」とは合計欄にはすべて「100％」と表示した上で，その表の欄外に「丸めのため各カテゴリーの％の加算値はちょうど 100 にはならない」といった注記を入れることである。

まず調査報告書等で，多くのカテゴリーから成る相対度数分布表を大量に作成して公表する場合を考える。このとき丸めた％の加算値が 100％ にならない表について，そのことを逐一欄外に注記するのは現実的に不可能である。冒頭で述べたように，カテゴリーの数が増大すれば四捨五入した％の合計がちょうど 100％ にならない確率は高まる。数十のカテゴリーが存在する分布表の場合，丸めた％の合計は 100％ にならないことの方が多い。したがって報告書に掲載する大量の分布表について逐一％の加算値が 100 になるかどうかを確かめ，ならないものにいちいち上のような但し書きを加える作業はあまりにわずらわしい。

したがって，こうした報告書類で注記をつけるとすれば，掲載するすべての相対度数分布表についての説明として一括して記せばよいだろう。具体的には①一連の相対度数分布表の各カテゴリーの％が小数第 m 位以下（m には具体的な正の整数を入れる）を四捨五入したものであること，②よって％の加算値がちょうど 100 にならない場合がありうること（各表の合計欄は一律に 100％ と表記するが）の 2 点について，報告書のどこかに記せばよいだろう。

次に少数のカテゴリーからなる相対度数分布表を数枚作成する場合を考える。この場合，丸めた％の加算値が 100％ にならない場合にのみ，それを欄外に注記することは可能だが，それを実際にするか否かは，公表した分布表の読み手として，どんな人々が想定されているかによるだろう。普段から度数分布表に親しんでいる専門家が読者なら，丸めた％の合計が 100％ にならないことがあるのを理解しているだろう。そういう読者に対しては特にそのことを注記する必要はあるまい。しかし，そうでない一般の読者にとって，合計欄に「100％」とあるの

に各カテゴリーの％の加算値が 100 にならないならば（カテゴリー数の少ない分布表の場合，読み手がこうした検算を行う可能性はある），ミスタイプや誤植を疑うかもしれない。そうした疑いを解消するために，欄外に「丸めのため各カテゴリーの％の加算値はちょうど 100 にはならない」などと記すことは無駄ではないだろう。

6. 結論と考察
－丸めた％の合計をちょうど 100 にせねばならない理由はあるのか？－

過去の社会統計のテキストの中には，公表を前提に相対度数分布表を作成するとき，四捨五入した％の合計値が 100％ にならないならば，各カテゴリーの％を微調整（変更）して，合計を 100 にすることを奨励するものも存在する。しかし，合計を 100％ にしなくてはならない特別な理由がないかぎりそうした調整は不要であり，むしろしない方がよいというのが本稿の主な結論である。

なぜならそうした調整は，たとえ必要最小限の調整（定義は第 4 節参照）であっても，％の大小関係を歪曲してしまう可能性があるからだ。第 4 節で述べたように，そうした調整を行うことで％の同一性，整合性，単調性が損なわれる。すなわち，本来同一であるはずの％が違う％になったり（同一性の不満足），真の％の大小関係が逆転してしまう（整合性または単調性の不満足）可能性がある。

このような認識に基づき第 5 節では丸めた％の合計が 100 にならないとき，そしてそれを 100 に調整する特別な理由が見当たらないとき，相対度数分布表をどのように作成・公表すべきかについての本稿の見解を述べた。第 1 に，各カテゴリーの％は四捨五入したものをそのまま（調整せずに）表示する。第 2 に，合計欄はいかなる場合でも一律に 100％（あるいは 100.0％，100.00％ など）とする。第 3 に，公表する相対度数分布表のカテゴリー数や枚数を考慮して可能であれば，または分布表の読者層を考慮して必要であれば，欄外に「丸めのため各カテゴリーの％の合計は正確に 100 にならない」といった注意書きを添える。以上である。

一方，第 3 節では，合計をちょうど 100 にすべき特別な理由がある場合の調整法として，安田・原（1982），西平（1985）の提示した相対誤差を最小化する方法の問題点を指摘した。また対案として，四捨五入の原理とより整合的な方法（最小絶対誤差法）を示した。ただし先行文献の方法にせよ，本稿が推奨する方法にせよ第 4 節で示した問題を免れない点は変わりがない。

最後に，ここまで再三言及してきた「丸めた％の合計を 100％ としなくてはならない特別な理由」が実際にありうるのか，あるとすれば具体的には何なのかを考えたい。冒頭で述べたように，合計を 100％ とする調整を勧める過去の教科書において，この点の説明はあまりにも簡単に済まされており，第 4 節で述べた不合理を覚悟してまで調整を行う理由の説明には

なっていないからである。

まず，相対度数分布表の作成過程に注目して，丸めた％の合計を 100 としなければならない理由があるかどうかを考えてみたい。主に筆算・電卓・算盤による手計算および手書きで相対度数分布表を作成していた時代には，各カテゴリーの丸めた％を調整して合計を常に 100 にしておくことには，それなりの効用があったと思われる。すなわち，そうすることによって表に記された％の誤記・誤植の見直しを非常にスムーズに行うことができたであろう。なぜなら丸めた％に調整を施して合計を常に 100 とすることにしておけば，公表するすべての分布表の合計欄がきっかり 100 になっていることを目で見て確かめるだけで，ある程度までチェックができたことになる。そして表中の％を足し算して実際に 100 になることを確かめれば，各カテゴリーの％にもある程度の確信を抱くことができる。誤記・誤植等がなければ％の合計は必ず 100 になるはずだからである。

一方，丸めた％の合計が 100.1 や 99.9 などになる可能性があれば，検算者が同程度の確信に至るには，もっと手間が必要になる。まず合計欄を眺めるだけでは，ほとんど何もチェックしたことにならない。さらに，各カテゴリーの％を足し算した結果が，表の合計欄に記されたとおりになったとしても，各カテゴリーの％に誤記等があるのではないかという疑いは晴れない。合計は「100.1」「100.0」「99.9」など，様々な数値である可能性が開かれているからである。したがって各カテゴリーの度数と総度数から％を逐一計算し直してみないかぎり，各カテゴリーの％に確信を抱くことはできない。そして手軽に使えるパソコンの無い時代において，こうした再計算は現在とは比較にならないほど手間がかかったであろう。こう考えると，本稿の第 4 節で示したような不合理が起きる可能性と，誤記・誤植等が生じる可能性を天秤にかけ，後者を重く見るという判断も，当時としてはあり得ない判断ではなかったであろう(9)。

しかし相対度数分布表の多くが，パソコンの表計算ソフトウェアやワープロ等で半ば自動的に作成されるようになった現代，丸めた％を調整して合計を 100％ になるようにしておくことの，上記の効用の大半は失われてしまったのではないだろうか？

次に，相対度数分布表の実質的な用途を考えたときに，丸めた％の合計を 100 としなければならない理由があるかどうかを考えてみたい。人文・社会科学における相対度数分布表のごく普通の用途を考えるかぎり，そのような理由はほとんど存在しないと言ってよいであろう。したがってほとんどの場合，合計を 100％ とする調整はしない方がよい。

筆者の認識の範囲で敢えて例外を挙げるなら以下のふたつが考えられる。いずれも比率として％ではなく 0 以上 1 以下の小数を用いる方が一般的なので「四捨五入した比率の合計がちょうど 1 にならないと困る場合」と言い換えた方がよいだろう。第 1 に，決まった量の資源を諸集団にその構成員数に比例して分配する場合が考えられる。この分配が慣習的に丸めた構成比率に基づいて行われていたとすると，各集団の丸め後の比率の合計がちょうど 1 にならないならば，資源の不足や余剰が生じることになる。第 2 に，公表された相対度数分布表を，

繰り返し計算を含む特殊な分析に 2 次利用する場合が考えられる。たとえば産業連関分析や移動現象のマルコフ連鎖分析では，クロス表のセルの比率同士を繰り返し掛け算する。よって四捨五入した比率の合計が正確に 1 でないと誤差が大きくなりやすい。たとえば初期の比率の合計の中に 1 を若干上回るものがあれば，最終結果の比率の合計は 1 を大きく上回ることになるかもしれない。しかしこれらの場合でも大雑把に丸めた比率を使わず，各カテゴリーの度数(10)から求めた分数表示の比率を使ったり，計算機・ソフトウェアの許す最大限の桁数の小数を使ったりすれば，そう大きな問題は生じないのではなかろうか(11)。

このように考えると今日，人文・社会科学で用いられる相対度数分布表において，丸めた％の合計を 100％ にしなくてはならない特別な理由というのは，筆者の認識の範囲ではほとんど思いつかない。したがってやはり丸めた％の合計を，同一性・整合性・単調性を犠牲にする覚悟で 100 にするような調整は，ほとんどの場合しない方がよいという結論に至る。

〔注〕

(1) 辻・有馬（1987：148 頁）は「①そのままの数字（99.9％ や 100.1％）を示す　②いずれか 1 つの回答選択肢（通常は「その他」や「わからない」など）で調節して丁度 100.0％ になるようにする」のどちらかにすべきだと述べる。また篠原・清水・榎本・大矢根（編）（2010：148 頁）は「相対度数が 99.9％ になってしまう場合には，相対度数の中で数値を改訂する必要がある」と述べるが，改訂の方法は述べていない。

(2) たとえば佐伯・松原（編）（2000：150 頁）に掲載された相対度数分布表（カテゴリー数＝10，総度数＝794）では，度数が 319 のカテゴリーの％が 40.1％ となっている。100×(319/794)＝40.1763…なので，小数第 2 位以下を四捨五入すれば 40.2％ になるが，そのまま全カテゴリーの丸めた％を合計すると 100.1 になるため，度数 319（10 カテゴリー中，最大）のカテゴリーの％を 0.1％ 切り下げたものと思われる。

(3) 岩永・大塚・高橋（1996：78 頁），辻・有馬（1987：169 頁）の分布表がそうである。

(4) アメリカ合衆国の憲法は下院（連邦議会）の全議席を各州に人口比例配分するよう定めている。しかしある州の人口の合衆国全体に占める割合と，その州への配分議席数の全議席数に占める割合とは，正確には一致しない。各州への配分議席は当然ながら「整数」に丸められねばならないからである。さらに各州への配分議席数を単純に四捨五入で整数に丸めた値で決定すると，配分議席の合計値は所定の全議席数と多くの場合一致しなくなる。したがって四捨五入以外の特殊な丸めの方法を用いて，合計を所定の総議席数に一致させる必要が生じ，その方法を巡って建国以来，論争が続いている（Balinski & Young, 1982＝1987）。この問題の「各州の配分議席」を「各カテゴリーの丸めた％」に「総議席数」を「丸めた％の合計値」に類比すれば，問題の構造が同じだとわかるだろう。実際，Balinski（1996）は議席の人口比例配分の方法を応用して数値データの特殊な丸めの方法を提案している。

(5) こうした場合，特殊なルールを用いて切り捨てるか切り上げるかを決定する方法も慣習的に用いられてきた。たとえば第 m 位の前の数字が偶数であれば切り上げ値に，奇数であれば切り捨て値に丸めるというルールが知られている（Diaconis & Freedman, 1979，安田・原，1982）。しかしコンピュータによる計算・作表が主流になった現代では，このルールはあまり採用されていない。たとえば一般に普及している表計算ソフト「Excel」の丸め機能は「5」で終わる小数を，本稿と同様に一律に切り上げている。

⑹ たとえば％の小数点以下を四捨五入して整数部分のみの％に丸めるとき，カテゴリー i の丸め後の％を 1％ 修正すると，相対誤差 I は $(q_i+1)/q_i$ または $q_i/(q_i-1)$ となる。これらの値はどちらも q_i が大きくなるほど小さくなり，1 に近づく。したがって q_i の最も大きいカテゴリー，すなわち度数が最大のカテゴリーで調整を行えば誤差が最も小さくなるとわかる。

⑺ 四者五入の原理は切り捨て値と切り上げ値の「算術平均」を境界線として，切り捨てるか切り上げるかを判断する方法であり，本文の例のように 1％ と 2％ の間の値をそのどちらかに丸める場合，1 と 2 の算術平均である「1.5」が，その境界線となる。これに対して相対誤差を最小化する場合には，算術平均の 1.5 ではなく，「幾何平均」である $\sqrt{(1\times 2)}=1.4142\cdots$ を境界線として，切り捨てるか切り上げるかを判断することになる。よって本文に記したように，1.45 は切り上げにより 2 に丸められる。相対誤差を最小化したい場合，切り捨て値を d，切り上げ値を u，切り捨て・切り上げの境界線を x とすると，$u/x=x/d$ が成り立ち，これを変形すると $x^2=du$ ゆえに $x=\sqrt{(du)}$ なので，切り捨て値と切り上げ値の幾何平均を境界線とする丸めの方法を採るべきだとわかる。また相対誤差の定義の仕方によっては，上記の他にも特殊な丸めの方法がありうる。たとえば相対誤差を $\max(1,p/q)-\min(1,p/q)$ などと定義すれば，d と u の「調和平均」すなわち $2du/(d+u)$ が切り捨て・切り上げの境界線になる。$d=1, u=2$ の場合，1.333…が境界線になるから，たとえば 1.4％ は 1％ でなく 2％ に丸められる。

⑻ 表 7 は 3 時点の例だが，2 時点の相対度数分布表でも同様の事態が生じうる。たとえば時点 1 の真の％がカテゴリー abc の順に 33.45％，33.0％，33.55％，時点 2 の真の％が同じく 33.4％，33.35％，33.25％ だったとしよう。小数点以下を四捨五入すると，時点 1 は順に 33％，33％，34％ で合計 100％ になるが，時点 2 は全カテゴリーが 33％ で合計 99％ になる。時点 2 の a を 34％ にする調整を行えば単調性が満たされなくなる。a の真の％は時点 1 から 2 にかけて下降しているのに調整後の％は上昇しているからである。一方，時点 2 の b または c を 34％ にする調整を行うと，今度は整合性が満たされない。時点 2 において a の真の％は b および c を上回っているのに，調整後の％は b または c を下回ることになるからである。

⑼ 手計算の時代に本稿が主張する表記法，すなわち丸めた％をそのまま表示し，その合計が 100 にならなくても合計欄には一律に「100」と記すやり方を採れば，ミスプリや誤記を一目で発見するチャンスを減らし，検算の手間をかえって増やすことになるかもしれない。

⑽ ただし公表された相対度数分布表の中には，大雑把に丸めた比率のみを示し，各カテゴリーの度数を示さないものも存在するから，複雑な繰り返し計算に耐える精度の比率を計算することは常に可能ではない。しかしその種の相対度数分布表を 2 次利用して，繰り返し計算を伴う分析をすること自体がそもそも間違いである。こうした場合は簡易表示された分布表のデータソースに遡り，各カテゴリーの正確な度数を確認し，そこからできるかぎり正確な相対度数分布表を作成して分析するべきであろう。

⑾ 合衆国の下院における議員定数の人口比例配分の問題（Balinski & Young, 1982＝1987）の場合，配分する資源の最小単位（1 議席）の増減が各州の利害に大きく影響するので，各州の合衆国全体に占める人口構成比をいくら正確に求めても問題は解消されない。ただしこの問題は比率の計算・表示そのものではなく，その比率から求めた配分議席数を整数に丸めるところで生じているから，本稿の射程を越えた問題と言ってよい。これに対して金銭の人口比例配分の問題であれば，金銭の最小単位（1 円，1 ドルなど）の増減が，諸集団の利害に大きく影響するとは考えられないので，構成比をできるかぎり正確に計算し，その結果に基づいて資源の最小単位ぎりぎりまで正確な分配をすれば，大きな問題は生じないだろう。

〔参考文献〕

Balinski, M. L., 1996, How should data be rounded?, *Institute of Mathematical Statistics Lecture Notes－Monograph Series,* 28： 33–44.

Balinski, M. L., and Young, H. P., 1982, *Fair Representation,* Yale University.（訳書：M. L. バリンスキー＆H. P. ヤング著，越山康（監訳）・一森哲男（訳），1987，『公正な代表制　－ワン・マン－ワン・ヴォートの実現を目指して－』，千倉書房．）

Bohrnstedt, G. W., and Knoke D., 1988, *Statistics for Social Data Analysis 2nd. Ed.,* Peacock.（訳書：ボーンシュテット＆ノーキ著，海野道郎・中村隆（監訳），1990，『社会統計学－学生版－』，ハーベスト社）．

Diaconis, P., and Freedman, D., 1979, On rounding percentages, *Journal of the American Statistical Association,* 74： 359–364.

一森哲男，2006，「ばらつきを考慮した議員定数配分方法について」，『日本応用数理学会論文誌』，16：265–276.

岩永雅也・大塚雄作・高橋一男，1996，『社会調査の基礎』，放送大学教育振興会．

Mosteller, F., Youtz, C., and Zahn, D., 1967, The distribution of sums of rounded percentages, *Demography,* 4： 850–858.

日本統計学会（編），2012，『資料の活用　日本統計学会公式認定統計検定 4 級対応』，東京図書．

西平重喜，1985，『統計調査法　改訂版』，培風館．

佐伯胖・松原望（編），2000，『実践としての統計学』，東京大学出版会．

篠原清夫・清水強志・榎本環・大矢根淳（編），2010，『社会調査の基礎　－社会調査士 A・B・C・D 科目対応－』，弘文堂．

辻新六・有馬昌宏，1987，『アンケート調査の方法　－実践ノウハウとパソコン支援－』，朝倉書店．

安田三郎，1969，『社会統計学』，丸善．

安田三郎・原純輔，1982，『社会調査ハンドブック　第 3 版』，有斐閣．

〔付記〕

佛教大学社会学部現代社会学科のゼミ生，松下雅さんに貴重なヒントを頂いたことを感謝いたします。もちろん，本稿に関する一切の責任は筆者にあります。

（やまぐち　よう　現代社会学科）

2014 年 10 月 31 日受理