

日本赤十字九州国際看護大学/Japanese Red Cross Kyushu International College of Nursing

分数と集計

メタデータ	言語: Japanese 出版者: 公開日: 2020-08-04 キーワード (Ja): キーワード (En): fraction, tabulation, cross tabulation 作成者: 守山, 正樹 メールアドレス: 所属:
URL	https://jrckicn.repo.nii.ac.jp/records/715

This work is licensed under a Creative Commons
Attribution-NonCommercial-ShareAlike 3.0
International License.



第2章 分数と統計

皆さん、こんにちは。今回は統計学の2回目、テーマは分数です。皆さんはすでに小学校の算数の時間に分数を学んでいます。その一方、統計学でも分数は基本です。まず分数的な発想についてお話しします。

1 分数的な発想

・一人の人に起こる出来事

まず前回の復習、1人の人における出来事の起こり方は、それが起こって「いない」か「いる」か、0か1かで、離散型確率変数（離散量）表されます。今日は、皆さん自身に起きる可能性のある出来事として「新型コロナウイルス COVID-19 への感染」を考えてみます。既に、新型コロナウイルスに感染した経験があると考える人は1を、経験がない人と考える人は0を思い浮かべてください。

・集団(例:このクラス 100人)

皆さんが所属するクラスの総数を100名とします。この100名に過去1カ月のウイルス感染について質問した結果、例えば7人から「感染あり」の答えが得られたとします。この感染の発生を数値で表したらどうなるでしょうか。多くの皆さんは、この問題をそれほど難しいとは思わず「100分の7です」などと分数で書き表すと思います。しかしこのように分数で書き表すことは、昔から常識だったわけではありません。

・古代から分数はあった

数の相対的な大きさを表すために分数を使う考え方は古代エジプトやギリシャの時代からあり、5世紀にはさらに進んだ形がインドでも現れたとされます。古代の分数の考え方は、目に見える物体やお金をいくつかに分けるような状況で使われていました。分数の考え方で集団を捉えていたわけではありません。

・[ものを分配する分数]と[集団の出来事の起こり方を表す分数]との違い

ものを分配する場合の分数は、幾何学的に、視覚的に表現できます。例えば目の前に1個のリングがあり、それを6人で分けるとしたら6分の1、この6分の1は視覚的に容易に捉えられます。

一方、このクラスで新型コロナウイルスへの感染がどのくらい起きたのかを分数で表すとしたら、合計(Σ , シグマ)を求める計算を二回行わなければなりません。分母には、クラス全員の合計(人数)が必要です。分子には、出来事が起こった人の合計(感染者数)が必要です。このように、分数の計算を行うためには、そこにいる人々を集団として捉え、その人数を数える発想が必要ですが、このような発想は15世紀までは未成熟でした。

・人間集団に分数を使い始めたのは17世紀

歴史的にみると、感染症の流行が分数の考え方を進歩させました。コレラやペストなど感染症の大規模な流行の現状を把握するため、死亡に関連して分数の考え方が導入され、死亡率の計算が実用化したのは17世紀と言われています。その代表はイギリス人ジョン・グラウント(1620-74)です。グラウントは当時たびたびペストが流行していたロンドンにおいて、各教区の教会から入手した出生と死亡に関する情報を分析し、1662年には「死亡調書の自然的および政治的観察」と

題する革新的な本を出版しています。グラントは、計算で利用した数値が、ロンドンで発生した全出生と全死亡の一部分でしかないことを認識した上で、そこからロンドン全体の状況を推測することも行いました。この考え方「標本のデータから、より大きな母集団の推測・推定を行う」は現在では推測統計（統計的推測）と呼ばれています。

2 分数と集計表

1) 初めての統計量、分数

では推測統計の考え方をを用いて、実際に分数を計算してみます。次のデータは今年の統計学の全受講者（100 名、操作的母集団）を対象として行った調査結果の一部分です。100 名から 6 名を無作為に選び、得られた標本（a さんから f さんまで 6 名の事例を含む）について、「朝食の摂取」を離散量（離散型確率変数）として示します。（朝食なし 0；朝食あり 1）

事例	朝食
a	あり
b	なし
c	あり
d	なし
e	なし
f	あり

「朝食あり」の割合を分数で示すにはどうしたらよいでしょうか。まず分母、全体の人数は、a さん b さん c さん... と数えて f さんまでで合計 6 名です。次は分子、「あり」の人数の合計は 3 名、よって「朝食あり」を表す分数は $3/6$ となります。この 6 名のような小さな集団のことを標本、サンプルと呼びます。さて、推測統計の目的は、標本から母集団の様子を推測することです。今回の標本は 6 名と少数ですが、ここで得られた $3/6$ から、母集団（今年の全受講者 100 名）の様子をどこまで推測できるのでしょうか。「推測」とは何らかの根拠をもとに予想をすることを意味します。絶対に正しい結果を導くのではなく、利用できるデータから予想を積み重ねることが大切です。 $3/6$ は一つの根拠と位置付けられます。

2) クロス集計表と分数

さて上述の例では一つの離散量（朝食あり・なし）を集計し、分数を計算しました。実際の調査では、離散量がもう一つ増え、二つの離散量が組み合わせられた場合が出て来ます。この場合はどうしたらよいでしょうか。組み合わせる（クロスする）場合の集計では、まず表（クロス集計表）による整理が大切です。二つの離散量のカテゴリーごとに分割して集計することから、分割表とも呼ばれます。次に例を示します。

	疲労感あり	疲労感なし	
朝食あり	n11	n12	
朝食なし	n21	n22	

クロス集計表では、横の並びを行、縦の並びを列、行と列の交差するそれぞれの部分をセルと言います。図に示すのは 4 つのセル（n11、n12、n21、n22）がある 2×2 のクロス集計表です。クロス集計表は、どういう条件の人が何人いるかを整理するのに役立ちます。

ではこのクロス集計表を用いて次のデータを集計してみましょう。ここに示すデータはやはり母集団(昨年の全受講者 100 名)から選んだものですが、今回は 6 名ではなく 10 名を抽出しています。また一つ目の離散量、朝食有無に加え、二つ目の離散量として、疲労感ありなしを示します。A さん B さんとデータを見ながら集計表に正の字を書いていきます。

No	朝食	疲労感
a	あり	あり
b	なし	あり
c	あり	なし
d	なし	あり
e	なし	あり
f	あり	なし
g	なし	なし
h	なし	なし
i	あり	なし
j	なし	あり

	疲労感あり	疲労感なし	計
朝食あり	1	3	4
朝食なし	4	2	6
	5	5	10

最後にこれらの値から自分で分数を工夫してそれがいくつになるかを考えてください。

3 分数と母集団と推測統計

今日は事例 10 名の小さな集団（標本）についてクロス集計を行い分数を得ました。何かを知りたい時それに合わせて分数を考え、その分数が計算できるように調査しデータを集めることは統計学でよく行われます。その際あなたが特徴を調べたいと思う集団、母集団は何かを意識することはとても大切です。ジョン・グラウンツの場合、実際の計算で用いたデータは各教区の教会から得た標本でした。しかしグラウンツが母集団として調べたいと意識していたのは、ロンドン全体でした。

皆さんの場合はどうでしょうか。皆さんが分析した A さんから J さんまでの 10 名のデータは、昨年の受講者全員（100 名）の名簿から、乱数表（数値がランダムに並んでいる表）を用いて、ランダム（無作為）に抽出した標本でした。標本の分析から得た結果を元に、標本の背後にある母集団の様子を推測することを、推測統計といいます。

操作的母集団とは実際に標本抽出を行うことができる母集団です。では操作的母集団以外に、さらにその元に、調べたい対象全体をあらわす理想的な母集団が存在するのでしょうか。・・・そう考えてみると、実は昨年の受講者 100 名の背後には、同じ県内の他の看護大学の学生も考えることができます。さらに広げていくと、この県だけでなく、隣の県、・・・さらに日本全国まで考えると、さらに多くの学生が視野に入ってきます。そのように対象全体を捉えた時、それを理想的な母集団と呼びます。

4 悉皆調査と分数

さて、最後に悉皆調査、全数調査という言葉に触れておきます。分数を用いて、日本全体の状況を考えることができるのでしょうか。

たとえば日本全体の人口を分母に、日本全体の死亡数を分子にとると死亡率が、また日本全体の

出生数を分子にとると出生率が計算できます。国全体の出生率や死亡率はとても重要な値ですので、標本から推測するだけでなく、国民全員についての調査から計算することも行われます。このように、対象をもれなく調べる調査を全数調査（悉皆調査）といいます。出生や死亡などの全数調査についての考え方は、皆さんが秋に学ぶ科目、保健統計学の中で詳しくお話しします。

さて、今回は平均値と偏差、標準偏差についてです。これらも中学や高校の時間に学習したテーマですが、皆さん忘れかけているのではないのでしょうか。ぜひ復習して、数字が与えられた時に、自分で計算できるようにしておいてください。電卓を使っても構いません。しかし、みなさんは簡単な計算は手で出来るような訓練を、小学校から高校にかけて行って来ています。手計算は大切な基本能力です。せっかく身につけた能力は、忘れないで、活用しましょう。その上でさらに電卓、パソコン、統計パッケージなど、より高度な計算方法にも親しんで行ってください。

演習問題

1. 最近の生活で気になる分数は何ですか。分数の具体例を挙げてください。
2. 昨日、ある保健所で1さんから6さんまで6名が、新型コロナウイルス COVID-19 を心配して PCR 検査を受けた結果を以下に示します。PCR 陽性者の割合を分数で表してください。

人	PCR 検査
1さん	陰性
2さん	陰性
3さん	陽性
4さん	陰性
5さん	陽性
6さん	陰性

3. 新型コロナウイルスによる外出自粛の影響調査を行い、二つの離散量（運動ありなし、食欲ありなし）について、10名から以下の結果を得ました。クロス集計表を作成し、4つのセルがどうなるか、それぞれにどのような数値が入るかを報告してください。

No	運動	食欲
1	なし	あり
2	なし	なし
3	あり	あり
4	なし	なし
5	あり	あり
6	あり	なし
7	あり	あり
8	あり	なし
9	あり	あり
10	なし	あり

4. 前の設問で作成したクロス集計表から「運動なし、食欲あり」の割合を分数で表してください。またこの分数から考えられることを述べてください。