日本赤十字九州国際看護大学/Japanese Red

Cross Kyushu International College of

Nursing

my標本からクラス全体のデータへ

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): sample, population, sampling, random number, estimation, point estimation, interval estimation, standard error 作成者: 守山, 正樹 メールアドレス: 所属:
URL	https://jrckicn.repo.nii.ac.jp/records/722

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

第 10 回　ｍｙ標本からクラス全体のデータへ

さて前回の授業では、12 人だけのデータからも統計学のレポートが書けるという話をしました。しかし皆さんが利用可能なのは､my 標本だけではありません｡my 標本を抽出した母集団(操作的母集団)、皆さんの先輩 150 人分のデータについても、利用できます。今回は my 標本から、その母集団（操作的母集団）に目を向けて、統計的な用語の説明を中心に行います。

１　my 標本から母集団へ

・標本

まず皆さんが前回救出した小さな集団選んだ 12 人のデータセット、my 標本、これを標本と言います

・操作的な母集団

実際に標本抽出を行うことができる母集団のことです。2017 年あるいは 2018 年に、この統計学を履修した皆さんの先輩が、この操作的な母集団です。

・概念的な母集団

教科書によると、調べたい対象全体を表す理想的な母集団と定義されています。皆さんの先輩が操作的母集団だとすると、その背景にある、より大きな、理想的な集団といえば、「他の大学の看護学生も含めた、日本全体の看護学生の集団」といえます。

２　標本抽出と乱数

母集団から標本を選ぶ際、適当に選ぶわけにはいきません。私たち人間が適当にこの選ぶ作業を行うと主観が入る可能性があります。主観を取り除き、無作為にランダムに選ぶ方法が求められます。全ての標本に番号が振られた母集団の名簿をもとに、乱数表などを使ってランダムな順番で番号を選ぶ方法が単純無作為抽出法です。今回の 12 人は皆さん自身が乱数表を使って選んだのではなく、私が乱数を使って、すでに皆さんのために順番をランダムに並び替えていたデータを利用しました。乱数表は、数字を無作為に並べている表のことです。昔の統計学の教科書には必ず乱数表が付いていました。最近はエクセルや、またスマートフォンでも iPhone であれば、簡単に乱数を発生させることができます｡iPhone の計算機は横にすると関数電卓になります｡Rand キーを押すごとに、１未満の乱数が発生します。

インターネット上にも乱数を発生できるサイトがあります。

https://keisan.casio.jp/exec/system/1425449868

他方、印刷された乱数表は見かけることが少なくなりました。

３　推定

推定とは、標本での計算から得た標本の特徴（標本の統計量）から、母集団の特徴（母集団の統

計量）を推し量ることを、統計学的な推定といいます。

・抽出と推定の関連

抽出と推定は統計学ではセットにして使うことが多い考え方です。教科書の始めの方、5 頁から 8 頁の記述を読み直してください。母集団を大きなスープの鍋、標本をそこから取ったスプーン一杯として説明されています。どのようにして大きなスープ鍋からスプーン一杯を取るかの方法が「標本抽出」、スプーン一杯の味からスープ鍋全体の味付けを予測するのが「推定」という関係が示されています。

・点推定

具体例で説明します。皆さんが本当に知りたいのは操作的母集団。皆さんの先輩におけるアルバイトの時間だとします。しかし皆さんの先輩は人数が多すぎて、全員のデータから計算するのは困難だとします。でも my 標本 12 名であれば、すぐに計算でき、アルバイト時間の平均がある値（たとぇば 70 分）になったとします。この値（70 分）をもって、皆さんの先輩のアルバイト時間を 70 分と推測したとき、その値を求める行為を「点推定」といいます。推定した結果を、一点の値として表わすのが、点推定です。教科書の 91 ページを見ると「適切な標本抽出法を用いて、標本平均は、母集団の平均の点推定値として最も優れているはずである」とあります。最も優れていると言っても、皆さんはそれぞれに my 標本を持っているわけですから、点推定値も my 標本によって、同じではありません。my 標本ごとに、やや異なる点推定値が得られることになります。

・標準誤差

皆さんの標本から得た平均の点推定値は、操作的母集団の平均に対して、どのくらいの誤差を持っているでしょうか。この誤差を標準誤差といいます。標準誤差の計算式は以下のとおりです。

$$\text{標準誤差} = \frac{\text{標準偏差}}{\sqrt{n}}$$

・区間推定

点推定によって平均を推定したとき、その推定値は、推定の誤差を考慮しなければならないために確率的な表現が必要です。つまり点推定のように特定の点を指定するのではなく、真の平均値は値 A と値 B の間におそらく含まれている、という考え方です。これを区間推定といいます。区間推定、特に 95%信頼区間は以下の式で表わされます。

$$\text{95\%信頼区間} = \text{標本平均} \pm 1.96 \times \text{標準誤差}$$

4　操作的母集団をどう分析するか？

さてここまでは、12 名からなる標本（my 標本）を中心に分析を考えてきました。この my 標本でレポートを書くことも出来ます。

しかし、皆さんは my 標本だけでなく、すでに操作的母集団全体のデータも利用することができます。12 人ではなく、150 人のデータ全てを使ってレポートを書くことも可能です。

しかし、標本から操作的母集団へと視点を移して分析するためには、それなりの方法を用いる必要があります。

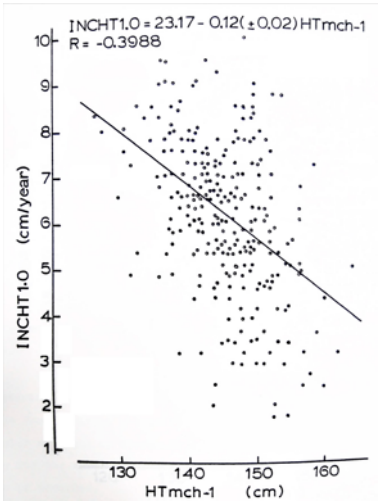
my 標本は 12 名と数が限られたデータでしたので、じっくりと一人ひとりのデータを観察し、質的な考察も可能でした。しかし 150 人以上ものデータだと、じっくりデータをひとり一人見る方法は時間がかかりすぎます。もっと早く、データの全体像を把握する必要があります。この目的で、もっともよく使われるのがエクセルです。皆さんは昨年、情報学でエクセルを使っているので、基本は理解しているはずです。

エクセルを使う場合、特に便利なのは、ピポットグラフの方法です。これを上手に使うと、簡単に度数分布表や棒グラフやクロス集計表を作ることが出来ます。

さて、皆さん全員がエクセルを使える環境にあるのであれば、エクセルの話をするのですが、どうでしょうか。現実には、パソコンやエクセルが利用できず、スマートフォンやタブレットで学んでいる皆さんも多いと思います。

そこで、次回からのオンデマンド授業でも、スマートフォン中心に話をしたいと思います。スマートフォンだけで、適切なサイトを使えば、様々な統計計算が可能です。しかし、もしエクセルが利用可能であるなら、昨年学んだ知識を活用して、ぜひエクセルを用いた計算にもチャレンジしてください。

５　手書きして考えることの大切さ



それから、ネットやコンピューターに頼るだけでなく、自分でグラフ用紙に手書きで点をプロットして考えることも大切です。

画面に示すのは、私が 40 年前、ネットやパソコンが無い時代に、手で書いた散布図です。１枚の図を描くのに何日もかかりましたが、図を描きながら、いろいろなことを考えました。現在は、エクセルで一瞬にして図が描けてしまうため、あのときほどは、深く図を見て考えることをしなくなったかもしれません。それから、回帰と相関の授業のときに紹介したゴルトンの図も思い出してください。ゴルトンの時代は、コンピューターはもちろん、機械的な計算機も存在しない時代です。あのとき、ゴルトンは多くのスイートピーの種の重さを測り、手書きでグラフを書きました。そして回帰という現象を発見したのです。

そういえば、皆さんの先輩にも、レポートにグラフを手書きした人がいました。資料の中に入れていますので、参考にしてください。

演習問題

1. 標本、操作的母集団、概念的な母集団について動画とは別な例を考え、50 字以内で書いてください。

2. 自分でいくつか乱数を発生させた上で、どのような乱数が現れたか、それを見て何を感じたか、何に用いたいかなど、80 字以内で書いてください。あなたのスマホで乱数が発生できない場合は、以下のサイトを利用してください。

https://keisan.casio.jp/exec/system/1425449868

3. あなたの my 標本 12 名のデータには、身長、体重、睡眠時間など、いくつかの連続量が含まれています。どれか一つ連続量を選んで区間推定を行い、得られた 95%信頼区間を以下に記してください。

4. あなたはレポートを書く際に、どのような手段で計算を行う予定ですか？またどのような計算手段に関心がありますか？もし将来、今回のレポートを書く際にも、本格的な統計パッケージを利用したいなら、JASP がお勧めです。

https://jasp-stats.org/

しかし JASP を使うには、あなたの自由になるパソコンが必要で、また基本操作は英語で行う必要があります。いったん使用方法を習得したら、エクセルよりもずっと容易に高度な統計計算ができます。（JASP に関心のある人が多ければ、次回以降の授業で基本を説明します）

以下、複数回答が可能です；ワークシートで手計算したい／電卓で計算したい／js-STAR などネット上の計算サイトを利用したい／パソコンでエクセルを使いたい／パソコンで JASP など本格的統計パッケージを利用したい。