

大学生を対象としたカードを用いた信頼区間の学習の可能性

小山達也

青森県立保健大学健康科学部栄養学科

抄 録

〔目的〕 近年、科学的根拠に基づいた栄養学の実践が求められている。科学的根拠を理解するためには推測統計の知識が必要である。信頼区間の理解のために実施した、カードを用いたシミュレーション教育について報告する。

〔方法〕 2019年1月に大学3年生23名を対象に実施した。1～9までの数字が書かれたカード2枚ずつ計18枚を用意し、この18枚のカードを母集団とみなし、母集団から無作為に4枚のカードを抽出させた。引いた4枚のカードが、母集団から無作為抽出された標本に相当することを説明した。カードを引くごとに平均値および95%信頼区間を計算した。この手順について、カードを引く学生と引かれる学生を順次交代し100回実施した。100回の95%信頼区間をグラフ表示し、100個の95%信頼区間のうち、95回は母平均を含み、5回は母平均を含まないことを視覚的に理解できるようにした。

〔結果〕 今回の試行では、95回が母平均を含み、5回が母平均を含まない結果となった。各回の結果が表示されるようにグラフを作成することで、1回ごとの95%信頼区間は母平均を含むか含まないかのいずれかであり、100回を通して考えたとき95回は母平均を含んでいることが視覚的に理解できるようにした。

〔結論〕 本シミュレーション教育は、学生の95%信頼区間の理解が高まる教材となる可能性がある。

《キーワード》 統計教育, 推定, 信頼区間

I. 緒 言

近年、科学的根拠に基づいた栄養学の実践が求められている。科学的根拠を理解するためには、統計学の知識が必要となってくる¹⁾。統計学は記述統計と推測統計に大きく2つに分けられる。わが国では、記述統計については高等学校までに学習する機会はあるが、推測統計については高等学校までに学習する機会が少ない²⁾。しかし科学的根拠を理解するためには記述統計だけでなく推測統計についても理解する必要がある¹⁾。

推測統計はさらに推定と検定に分けることができる。推定は点推定と区間推定があり、95%信頼区間は区間推定の1つである。母平均を推定する際に、標本から得られた平均値により1つの値で推定することが点推定であり、ある区間をもって推定することが区間推定となる。標本平均を m 、母平均を μ 、母分散を σ^2 、抽出したサンプルサイズを n としたとき、母分散が既知のときの母平均 μ の95%信頼区間は次の式で求めることができる。

$$m - 1.96 \times \sqrt{\sigma^2/n} \leq \mu \leq m + 1.96 \times \sqrt{\sigma^2/n}$$

何度も実験を繰り返せば、95%の確率で母平均がこの範囲に含まれる。しかし、実際の調査においては何度も繰り返すことは少なく、頻度的な理解は難しい。95%信頼区間を理解する方法には、理論的に学習する方法やシミュレーションを用いて学習する方法がある。統計学的推測を学習する際には、シミュレーションを用いて学習する方が理論的に学習するよりも、事前に関連する知識を学習することが少なく³⁾、統計の理解度が高まることが報告されている^{4,5)}。

そこで、青森県立保健大学の栄養学科3年生を対象に、カードを用いて95%信頼区間のシミュレーション教育を実施したので報告する。本シミュレーションの目的は、母集団平均の95%信頼区間を構成する作業を100回繰り返したとき、100回のうち95回前後は母集団平均を捉えているということを視覚的に観察し、95%という意味を頻度的に理解させることである。

II. 方 法

2019年1月に、栄養学科3年生の希望者23名を対象として統計講座を実施した。受講生には個人情報扱わず、講義の概要を研究報告することを口頭で説明した。受講生は、本講義以前に人間総合科学科

連絡先 小山達也 (E-mail: t_koyama@auhw.ac.jp)
 青森県立保健大学 健康科学部 栄養学科
 〒030-8505 青森市浜館間瀬58-1
 Tel: 017-765-4214 Fax: 017-765-2117
 (2019年6月24日受付: 2019年10月30日受理)

目や公衆衛生学，食事摂取基準論，栄養疫学，公衆栄養学等の専門科目で統計学について，95%信頼区間も含め基礎的な事柄を学習している。一方，統計講座は，卒業研究を進めていく中で必要となる統計学の理解を深めることを目的とした。統計講座の全体の概要を表1に示す。各回の講義は筆者が180分実施した。講義内容についてはパワーポイントを用いて説明を行い，適宜ホワイトボードに補足事項を書く形で進めた。また，筆者が各回90分程度かけて資料準備を行った。

第3回目の推定で，95%信頼区間を取り上げた。95%信頼区間の学習の前段階として，母集団と標本との関係，点推定と区間推定の関係，母数と標本統計量を説明した。その後，95%信頼区間を理解するためのシミュレーションを行った。このシミュレーション教育は書籍⁶⁾に記述されていることを再現したものである。

1～9までの数字が書かれたカードを2枚ずつ計18枚を用意し，この18枚のカードを母集団として，母集団から無作為に4枚のカードを抽出させた。引いた4枚のカードが，母集団から無作為抽出された

標本に相当することを確認させた。カードは1～9までの離散変数であり，各2枚であるため個々のカードが引かれる確率は一様分布に従う。しかし，引いたカードの平均値は中心極限定理により正規分布に従うことになる。本講義では母分散は既知（各回で引かれたカードの分散）として，得られたカードの平均値を m ，分散を σ^2 としたとき，母平均の95%信頼区間の下限は $m - 1.96 \times \sigma$ ，上限は $m + 1.96 \times \sigma$ となる。4枚のカードの数字を予め数式を入力したエクセルシート[®]に入力することで，平均値および95%信頼区間が計算されるようにした（図1）。この18枚のカードから4枚のカードを引く操作について，カードを引く学生と引かれる学生を順次交代し，100回行った。95%信頼区間を構成する作業を100回実施した後，100回の95%信頼区間をグラフから，100個の95%信頼区間のうち，95回は母平均を含み，5回は母平均を含まないことを視覚的に理解できるようにした。

エクセルについては教員用のパソコンだけ準備し，教員のエクセルの入力画面をスクリーンに投影する形で実施し，学生は特に準備物はなく実施した。

その後，信頼区間の幅は標本の大きさに依存することを，シミュレーションで用いた例から抽出するカードの枚数を変更することと関連付けて理解させ，信頼区間の幅は標本のばらつきにも依存することを100回の操作の中の例で理解させた。最後に，信頼水準についても説明を行った。

表1. 統計講座の概略

回	内容	キーワード
1	データ尺度	質的変数，量的変数，名義尺度，順序尺度
2	基本統計量	平均値，中央値，最頻値，範囲，四分位範囲，分散，標準偏差，変動係数，標準化得点
3	推定と検定	母集団，標本，標準誤差，信頼区間，帰無仮説，有意確率，有意水準
4	仮説検定	t検定，一元配置分散分析，多重比較，カイ二乗検定，相関分析，回帰分析，相関係数，線型回帰，ロジスティック回帰
5	実例紹介	論文抄読

III. 結果

100回カードを引くことで得られた平均値の95%信頼区間の結果を図2に示す。今回の試行では95回が母平均を含み，5回が母平均を含まない結果となった。各回の結果が表示されるよう事前にグラフの設

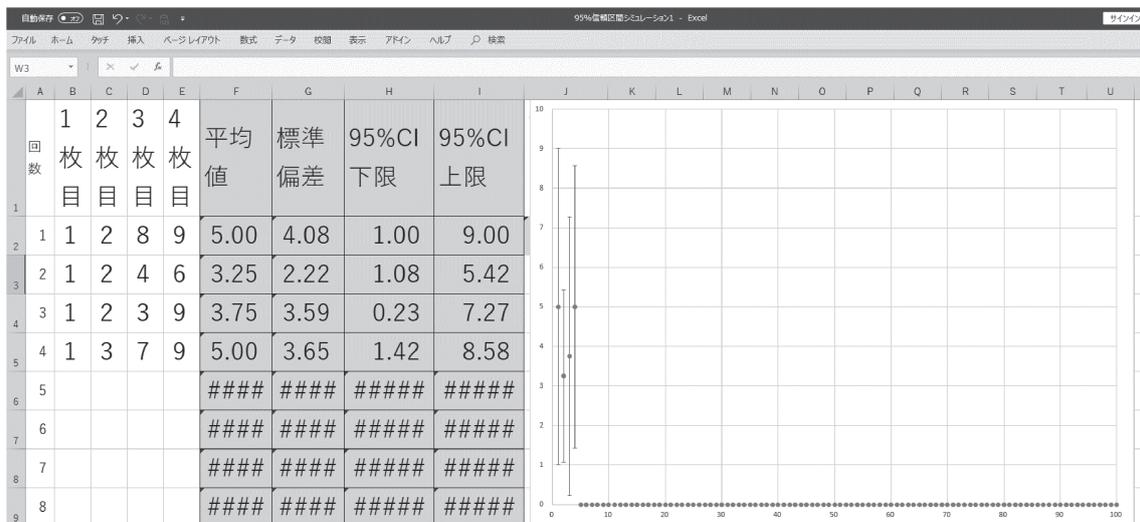


図1. 95%信頼区間を算出するエクセルシート

平均値，標準偏差，95% CI 下限，95% CI 上限には予め計算式を入力しておき，1～4枚目までに数字を入力することで計算結果が表示されるようにした。併せて散布図のグラフを用いて，右のグラフも表示されるようにした。

(計算式の例) F2のセル=average (B2 : E2)，G2のセル=stdev (B2 : E2)，H2のセル=F2-1.96*G2/2，I2のセル=F2+1.96*G2/2

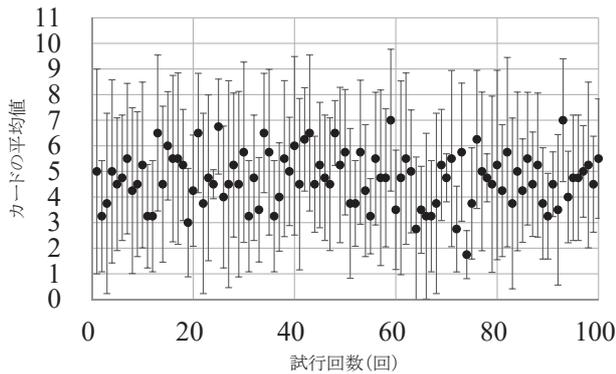


図2. 95%信頼区間のシミュレーション結果
平均値を点で、95%信頼区間をひげとして各回の信頼区間を图示した。

定を行えば、1回ごとの95%信頼区間は母平均を含むか含まないかのいずれかであり、100回を通して考えたとき95回は母平均を含んでいることがわかる。また、今回は試行回数が少ないうちは、95%の確率で母平均を含むとは必ずしもいえないが、100回の試行では95%の確率で母平均が含まれていた。

学生はすぐに演習目的と作業内容を理解して取り組んだ。当初は95%信頼区間の性質上、20~30回目くらいに母平均を含まないカードの組合せが起こると予想していたが、実際には48回目に初めて母平均を含まない信頼区間が得られた。結果的には100回中95回が母平均を含み、5回が母平均を含まないという理論通りの結果が得られた。そのため、試行の途中で解説を行うよりも、すべての試行が終了してから説明した方が、説明しやすく、試行回数は100回程度必要であると考えられる。

IV. 考察

数字の書かれたカードとエクセルを用いることで、比較的簡便に95%信頼区間を体験的に理解させる講義を行った。数字の書かれたカードを引く、引かれる操作を学生自身が行うことで、学生主体の講義を行える。今回の講義では、数字の書かれたカードを引く操作を100回実施したが、95%信頼区間が母平均を含む確率が、100回程度実施することではじめて95%程度になった。

95%信頼区間の幅はサンプルサイズによっても変化する。そのため、本講義ではカードを引く枚数を4枚にした。しかし、カードを引く枚数を8枚にするなどサンプルサイズを変えて同様の手順を行うことで、信頼区間の幅が変化する様子が観察され、95%信頼区間の幅はサンプルサイズにも影響することが学習できた可能性がある。ただし、サンプルサイズにより信頼区間の幅は変化するが、母平均を含む確率は変化しないので、カードを引く回数を変更する必要はないと考えられる。

95%信頼区間の上限値、下限値により、臨床的に重要であるかどうかを冷静に判断する手段が得られる。近年、統計学的検定によりp値だけを示すので

はなく、統計学的推定を示すことが求められている⁷⁾。また2017年3月に実施された第31回管理栄養士国家試験にも出題されるなど、管理栄養士も95%信頼区間の意味の理解が望まれている。しかし、大学生や大学院生だけでなく⁸⁾、研究者の多くが95%信頼区間を適切に理解していないことが報告されている^{9,10)}。平成30年度管理栄養士専門分野人材育成事業「教育養成領域での人材育成」報告書において、栄養学研究の学修目標の1つとして、「情報リテラシーを獲得し、統計学の基本的知識を説明できる」という項目はあるが、栄養学研究でどのような統計学の基本的知識を教えているのかについての現状は報告されていない¹¹⁾。

信頼区間という用語の定義の理解、標本平均と母平均の区別と両者の関係についての理解、信頼水準という概念の理解、標本の大きさや標本のばらつきが信頼区間の幅に影響を与えることへの理解、将来に繰り返されたときに信頼区間に基づいて何が推測されるか理解すること、信頼区間を正確に解釈する方法を理解することが、信頼区間について考える際に重要となってくる⁸⁾。一方で、信頼区間について誤解して理解していることも報告されている¹²⁾。

学生はシミュレーションを行った方が95%信頼区間についての理解度は深まったように見えた。より視覚的に分かりやすくするため、95%信頼区間が母平均を含まない場合は色を変えるなどの工夫が必要である。また、シミュレーションを行うのに30分程度要するため、もう少し短時間でを行い、同じ講義時間内で95%信頼区間の具体例を紹介した方が理解の定着は良かったと考える。本講義では受講生の理解状況を把握できておらず、本シミュレーションの目的が達成されたかは不明である。そのため今後の課題としては、シミュレーションを用いることで信頼区間が理解できるようになったのかを評価すること、学生の理解を助ける適切な教材としてどのようなシミュレーションが望ましいか検討することがあげられる。また、本論文では統計学的推定について考察したが、推測統計学には統計学的検定もある。統計学的検定ではp値の理解が大切になってくるが、誤解も多い¹¹⁾。統計学的検定を理解できる教材の開発も行っていきたい。

V. 結論

青森県立保健大学栄養学科3年生を対象に、カードを用いて95%信頼区間のシミュレーション教育を実施した。本講義では、学生の理解度を評価しなかったが、本シミュレーションは信頼区間についての誤解が軽減され、理解度の高まる教材となる可能性がある。

謝辞

統計講座の受講生に感謝申し上げます。

利益相反

開示すべき COI 状態はない。

文 献

- 1) 佐々木敏：わかりやすいEBNと栄養疫学, pp. 85-108 (2010) 同文書院, 東京
- 2) 渡辺美智子：知識基盤社会における統計教育の新しい枠組み～科学的探究・問題解決・意思決定に至る統計思考力～. 日本統計学会誌. 2013; 42(2): 253-271.
- 3) Malone CJ., Gabrosek J., Curtiss P., et al.: Resequencing topics in an introductory applied statistics course. *The American Statistician*. 2010; 64(1): 52-58.
- 4) Cobb GW.: The introductory statistics course: A Ptolemaic curriculum? *Technology Innovations in Statistics Education*. 2007; 1: 1.
- 5) Nathan T., Jake C., Karen F., et al.: Assessing the association between precourse metrics of student preparation and student performance in introductory statistics: results from early data on simulation-based inference vs. nonsimulation-based inference. *J Statistics Education*. 2018; 26(2): 103-109.
- 6) 林拓也：社会統計学入門 (改訂版), pp. 1-244 (2018) 放送大学教育振興会, 東京
- 7) Cumming G., Fidler F., Kalinowski P., et al.: The statistical recommendations of the American Psychological Association Publication Manual: Effect sizes, confidence intervals, and meta-analysis. *Australian Journal of Psychology*. 2012; 64(3): 138-146.
- 8) Noelle MC., Anna NB., Martha WA.: Conceptual knowledge of confidence intervals in psychology undergraduate and graduate students. *Statistics Education Research Journal*. 2019; 18(1): 46-62.
- 9) Coulson M., Healey M., Fidler F., et al.: Confidence intervals permit, but don't guarantee, better inference than statistical significance testing. *Frontiers in Psychology*. 2010; 1: 26.
- 10) Sotos AEC., Vanhoof S., Van den N., et al.: Students' misconceptions of statistical inference: a review of the empirical evidence from research on statistics education. *Educational Research Review*. 2007; 2(2), 98-113.
- 11) 栄養改善学会 (2019) 「平成30年度管理栄養士専門分野別人材育成事業「教育養成領域での人材育成」報告書」(http://jsnd.jp/img/H30_houkoku_all4.pdf, 2019年9月15日)
- 12) Wasserstein RL., Lazar NA.: The ASA's statement on p-values: context, process, and purpose. *The American Statistician*. 2016; 70(2): 129-133.

Lesson on confidence interval for university students using cards

Tatsuya Koyama

Aomori University of Health and Welfare

..... (Received June 24, 2019; Accepted October 30, 2019)

ABSTRACT

[Objective] Recently, practice of evidence-based nutrition has been required. Knowledge on inferential statistics is necessary to understand the scientific evidence. This paper reports on the simulation education using cards to understand the confidence interval.

[Methods] Lectures on statistics were held in January 2019 for 23 third-year undergraduate students. A total of 18 cards with numbers from 1 to 9 were prepared, and these 18 cards were regarded as a “population”, from which four cards were randomly extracted. The undergraduate students received the explanation that the four cards drawn was considered as a “sample”, randomly selected from the “population”. Mean and 95% confidence intervals were calculated for each trial. Student conducted these samplings 100 times in turns. A graph of 100 trials of estimated 95% confidence intervals was shown for the students to understand visually.

[Results] Out of 100 trials, 95 trials of estimated 95% confidence intervals included the population mean, whereas 5 trials of estimated 95% confidence intervals did not include it. Showing 100 trials of estimated 95% confidence intervals as a graph would help the students visually understand the above results.

[Conclusions] This simulation education could possibly be a learning material that enhances the undergraduate students to understand 95% confidence interval.

Aomori J. Health Welfare, 1; 29–33: 2019

Key words: statistical education, estimation confidence interval