

第 18 回 統計教育の方法論ワークショップ・理数系教員授業力向上研修会(東京)
The 18th Japan Conference for Teaching Statistics, JCOTS2021

令和3年2月27日(土)

セッション V: 日本での高等学校統計・データサイエンスの実践 I

推測統計の指導について



広島大学附属中・高等学校
橋本 三嗣



広島大学

背景(社会からの要請)

- ・ AI実装社会の到来でデータサイエンスの重要性が増し、データから意思決定・行動につながる統計的思考・判断・表現力を育成することが急務となっている。
- ・ ICTの活用に関して学校教育への期待も強まり、文部科学省が進めるGIGAスクール構想もその実現に向けて準備が進んでいる。
- ・ 学校内の設備面が充実する中、実際にどのように活用して指導を展開するかという問題が出てきた。

背景(数学の捉え方の変化)

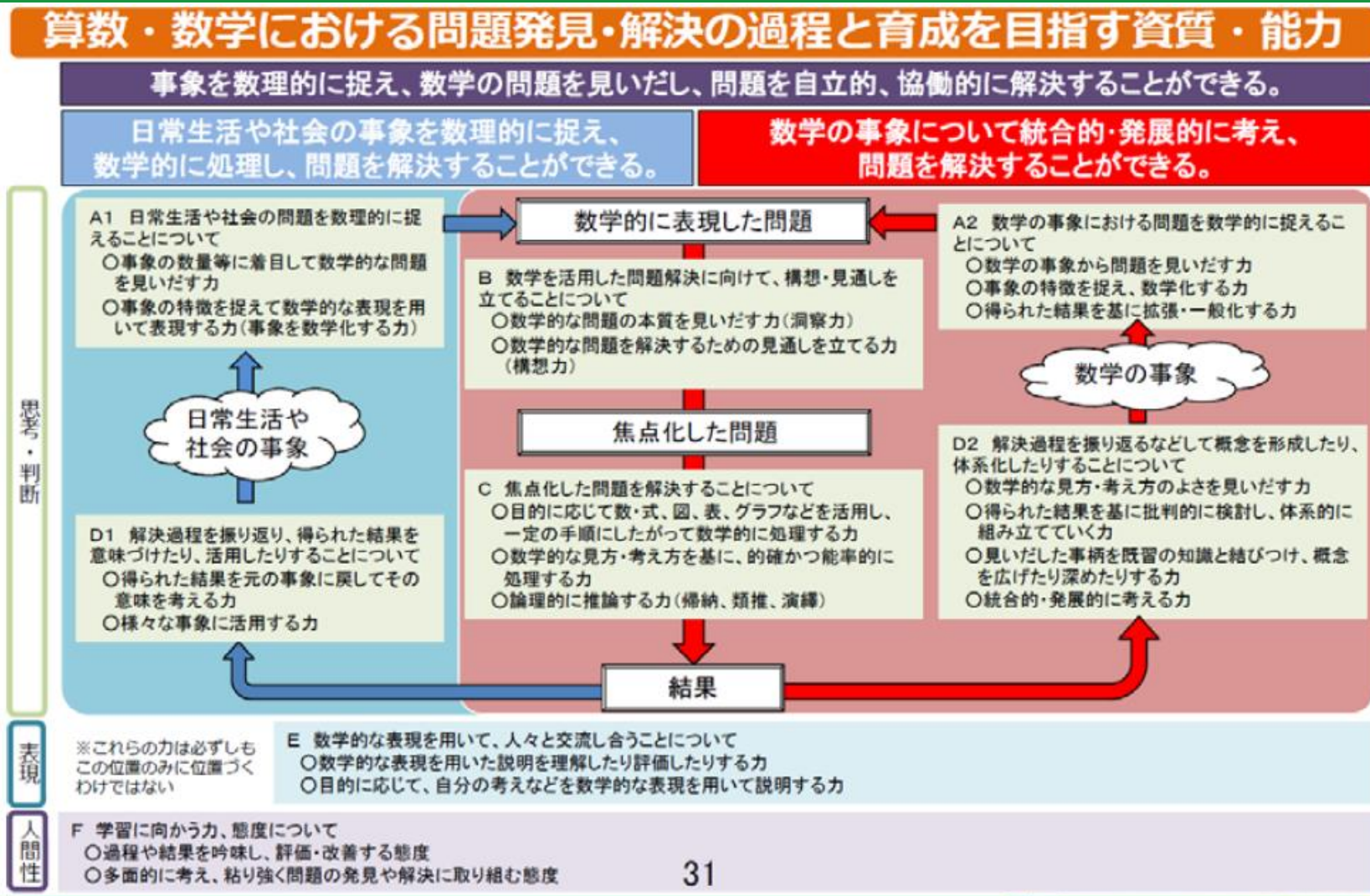
日本学術会議数理科学委員会数理科学分野の参照基準検討分科会(2013)、「報告 大学教育の分野別質保証のための教育課程編成上の参照基準 数理科学分野」

大学における数理科学を「数学を中心とし、数学から生まれた統計学や応用数理などの分野と、数学教育や数学史など数学と他の学問分野との境界分野を合わせた学問分野」とした上で、「これからの時代の市民にとって、数理科学的な事象の把握・処理の能力は欠かせない。市民が正しい判断を行うためには、データに基づき物事を量的に把握することが必要不可欠であるが、そのような能力の涵養において、数理科学教育(算数・数学教育)が果たす役割は大きい。」と指摘。



初等・中等教育の算数・数学の教科カリキュラムに影響

算数・数学の学習過程のイメージ

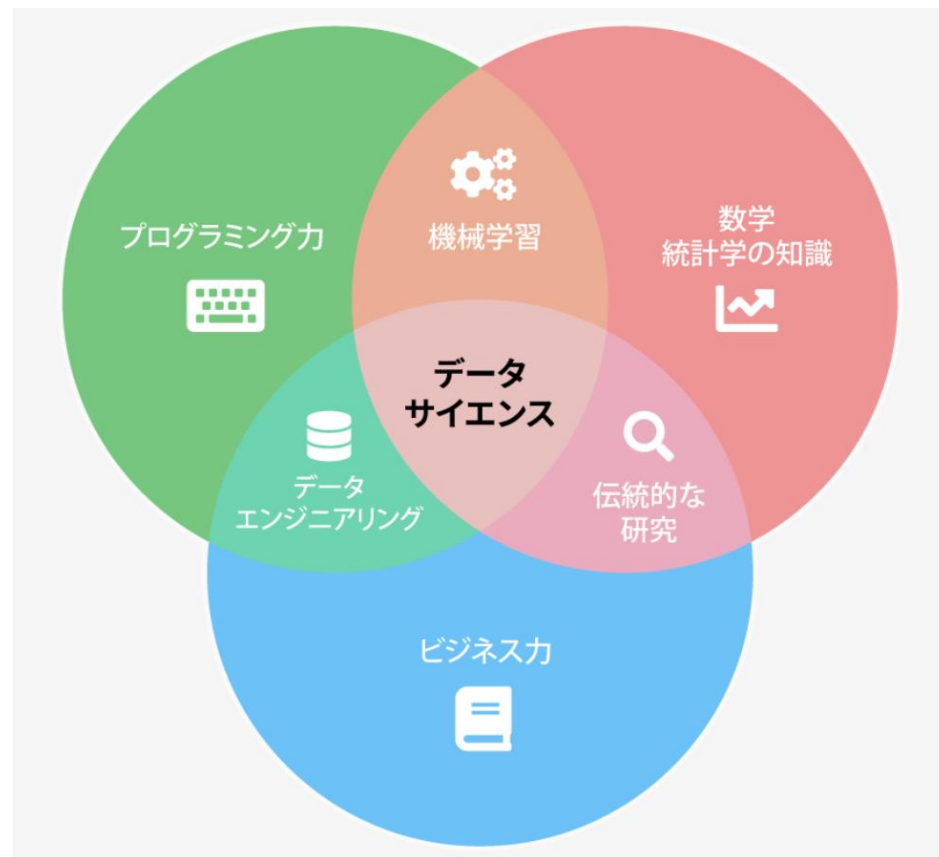


報告する内容

推測統計の指導事例
推測統計(推定と検定)

高校2年の数学Bの授業で、
問題解決的な活動をどのように展開するか。

実践と生徒の反応
(2017年9月実施)



<https://www.data-artist.com/contents/data-science.html>(2021.02.12)



指導事例「正規分布を利用した区間推定」

(1)授業の概要

高校2年生を対象にして、
学校設定科目「GS社会と統計」の時間に正規分布を用いた区間推定を指導した後の1時間で実施

課題「標本の平均から古語辞典の見出しの総数について推測しよう！」

→古語辞典の見出しの集合を母集団として、母平均の区間推定を行い、母集団の中の見出しの総数について推測

時間	生徒の学習活動	授業者の指導・支援
導入 5分	1 問題解決の見通しを立てる。	- 古語辞典を示し、「見出しの総数を母平均の区間推定を利用して求める」という問題を提示する。 - 学習課題を、標本調査の計画を立てて実施すること、収集したデータから区間推定を行い、その構成要素の総数について推測することに設定する。
展開 10分 15分	課題1 標本調査の計画を立てて、古語辞典の見出しの総数を調べよう。	
	2 乱数さいを利用したデータ収集の計画を立てる。	無作為抽出の方法を確認し、乱数さいを利用してどのように調べるページを決めるかを話し合い、ルールを決める。
	3 グループでデータ収集を行う。	標本の大きさ n の値が大きい方が望ましいが、15分間で調査できるページのデータを収集する。そのためグループにより調べるページ数が異なる。
10分 5分 まとめ 5分	課題2 収集したデータの標本平均から母平均を区間推定し、見出しの総数について推測しよう。	
	4 収集したデータの標本平均から母平均の95%の信頼区間を求める。	- ノートPCの表計算ソフトを利用させる。 - 各グループで求めた母平均の信頼区間を黒板に書かせる。
	5 求めた母平均の95%の信頼区間から見出しの総数について推測する。	古語辞典のページ数が1423であることから、母平均の95%の信頼区間の各辺を1423倍して値の範囲を求める。
	6 記載されている値と比較し、その違いについて考察する。	見出し約23000と記載されていることを伝え、推測した結果と比較する。
	7 本時の学習を振り返る。	学習を振り返り、標本調査の方法や結果をまとめる。



指導事例「正規分布を利用した区間推定」

母平均の推定

標本の大きさ n が大きいとき、母平均 m に対する信頼度 95 % の信頼区間は

$$\bar{X} - 1.96 \cdot \frac{\sigma}{\sqrt{n}} \leq m \leq \bar{X} + 1.96 \cdot \frac{\sigma}{\sqrt{n}}$$

1. データを収集する (15分)

乱数さいでページを決める (), 古語辞典の見出しの数を数える ()
 プリントに記入する (), エクセルに入力する ()

0671 : 14	0544 : 14	0966 : 16
0821 : 15	0938 : 9	0832 : 11
0319 : 17	0120 : 9	0227 : 14
1368 : 16	0944 : 15	0724 : 6
0798 : 2	1308 : 17	1417 : 20
1073 : 13	0926 : 8	1334 : 19
0341 : 7	1246 : 19	0675 : 17
0546 : 10	0594 : 11	0781 : 22
1322 : 19	1320 : 12	0497 : 18
0246 : 9	0643 : 14	0412 : 15
0991 : 17	1314 : 20	1101 : 20
1231 : 20	0673 : 22	1034 : 15
0914 : 19	0062 : 10	1361 : 17
0196 : 16	1164 : 20	0422 : 10
0458 : 15	1017 : 15	0661 : 17
0309 : 24	0460 : 11	
0378 : 18	0128 : 16	
0514 : 14	1261 : 11	
0480 : 17	0964 : 10	
0864 : 11	0970 : 7	
1375 : 17	0800 : 13	
	0293 : 12	

2. 母平均 m に対する信頼度 95 % の信頼区間を求める (10分)

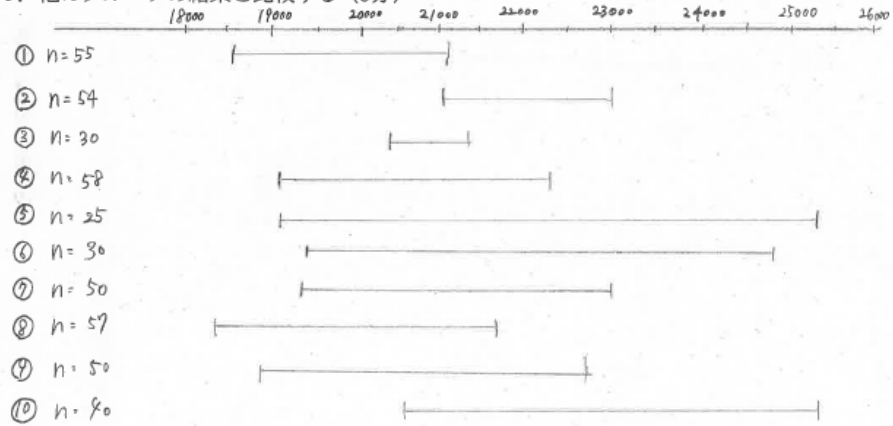
$$14.6 - 1.96 \cdot \frac{4.4}{\sqrt{7.6}} \leq m \leq 14.6 + 1.96 \cdot \frac{4.4}{\sqrt{7.6}}$$

$$13.4268 \leq m \leq 15.71$$

3. 古語辞典の見出しの総数を推測する (5分)

$$19106 \leq 1423m \leq 22357$$

4. 他のグループの結果と比較する (5分)



5. 授業の振り返りを行う (5分)

標本の大きさが大きいと、範囲が狭まることなどが分かった。

指導事例「正規分布を利用した区間推定」

(2)生徒の発言やつぶやき 〔調べるページの決め方〕

生徒A 「この辞典は13～1435ページまで見出しがあり、古語辞典が全部で10冊しかないことも考えるとグループで分担して全ページの見出しを数えるのは無理だな。」

生徒B 「そうそう。時間内に終わらないよ。」

生徒C 「乱数さいを利用するみたいだけど、千の桁まで決めるにはさいころが4ついるね。」

授業者 「どうやってページを決めますか？」

生徒C 「4つの乱数さいで桁数を決めておいて、千の位は偶数を0, 奇数を1にすればいい。辞典にないページは無視ということで。」

生徒B 「そうしよう。」

授業者 「乱数さいを使って同じページに決まったらどうしますか？」

生徒(複数) 「2回目以降は無視しよう。」





指導事例「正規分布を利用した区間推定」

(2)生徒の発言やつぶやき 〔見出しの総数の求め方〕

授業者「見出しの総数について推測するにはどんな計算をすればよいですか？」
生徒D「古語辞典は全部で1423ページあるから、 m の不等式の各辺を1423倍すればいい。」

授業者「母平均 m の95%の信頼区間は、“1ページあたりの見出しの数の範囲”を表しているため、1423倍すると、見出しの総数について推測できます。計算してみましょう。」

生徒D「私のグループ3は $21629.6 \leq 1423m \leq 24760.2$ となります。幅が3130.6もあります。」

授業者「幅を小さくするにはどのようにすればよいですか？」

生徒E「調べるページの数を増やすとよいです。」

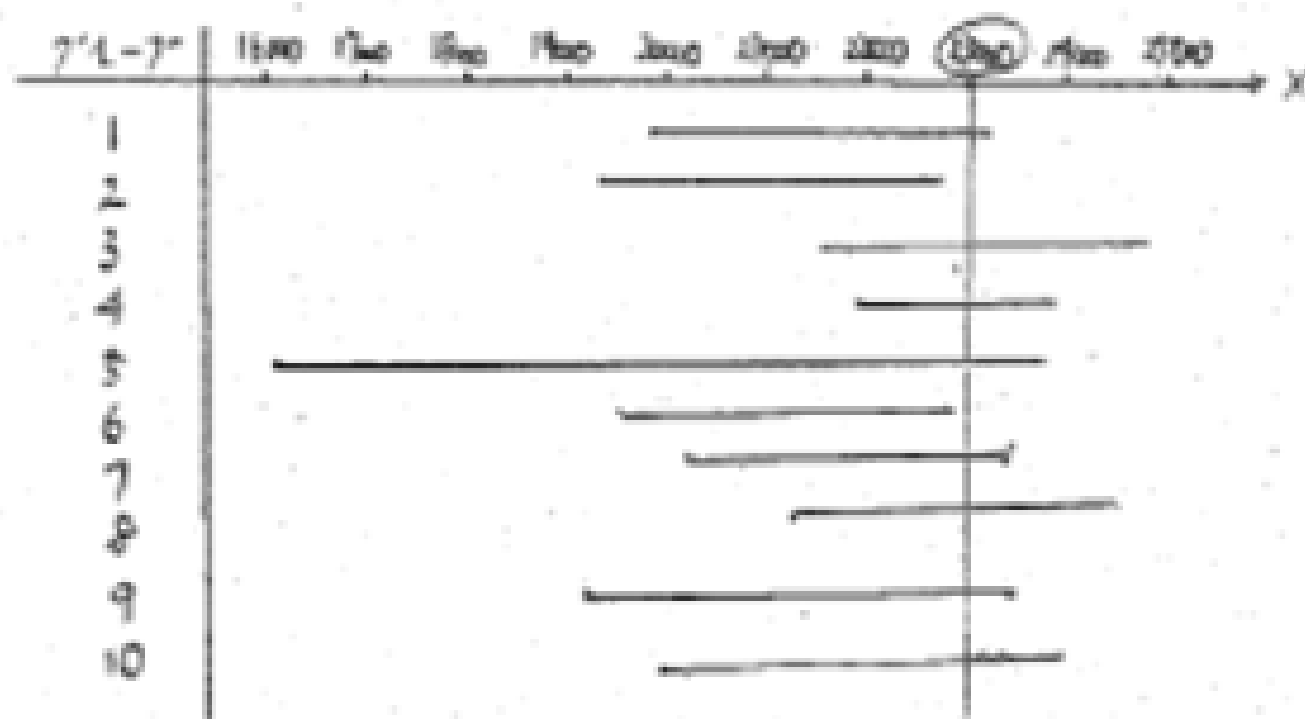
授業者「そうですね。母平均 m の信頼区間から、区間の幅は $2 \times 1.96 \cdot s / \sqrt{n}$ となります。 n の値を大きくすると、 s の値の変動は一般には小さいので、区間の幅は小さくなります。例えば、データ数が最も少ないグループ5は区間の幅が一番長いですね。」

指導事例「正規分布を利用した区間推定」

各グループの求めた母平均 μ の信頼区間

グループ	母平均 μ の信頼区間 (95%)	グループ	母平均 μ の信頼区間 (95%)
1 (50)	$14.0 \leq \mu \leq 16.4$	6 (61)	$13.8 \leq \mu \leq 16.0$
2 (44)	$13.6 \leq \mu \leq 16.1$	7 (58)	$14.2 \leq \mu \leq 16.4$
3 (48)	$15.2 \leq \mu \leq 17.4$	8 (47)	$15.0 \leq \mu \leq 17.2$
4 (45)	$15.4 \leq \mu \leq 16.7$	9 (63)	$13.5 \leq \mu \leq 16.2$
5 (40)	$11.4 \leq \mu \leq 16.7$	10 (55)	$14.0 \leq \mu \leq 16.6$

(グループの () はデータ数を表している, 小数第2位を四捨五入)





指導事例「正規分布を利用した区間推定」

(2)生徒の発言やつぶやき 〔見出しの数が約23000と発表した後〕

授業者「実はこの古語辞典の見出しは約23000だそうです。」

生徒D「すごい。23000は私のグループ3で計算した不等式の範囲に入っています。」

生徒E「私のグループ8も入っています。」

授業者「皆さんのグループで23000は不等式の範囲に入っていますか？」

生徒E「僕たちのグループ2は、 $19352.8 \leq 1423m \leq 22910.3$ となり、23000は不等式の範囲に入っていません。しかし、約23000なので僕たちのグループの不等式は全くの外れではないと思います。」

授業者「見出しの総数について、不等式の範囲を図に表して比較してみましょう。」

指導事例「正規分布を利用した区間推定」

(2)生徒の発言やつぶやき 〔図をかいた後〕

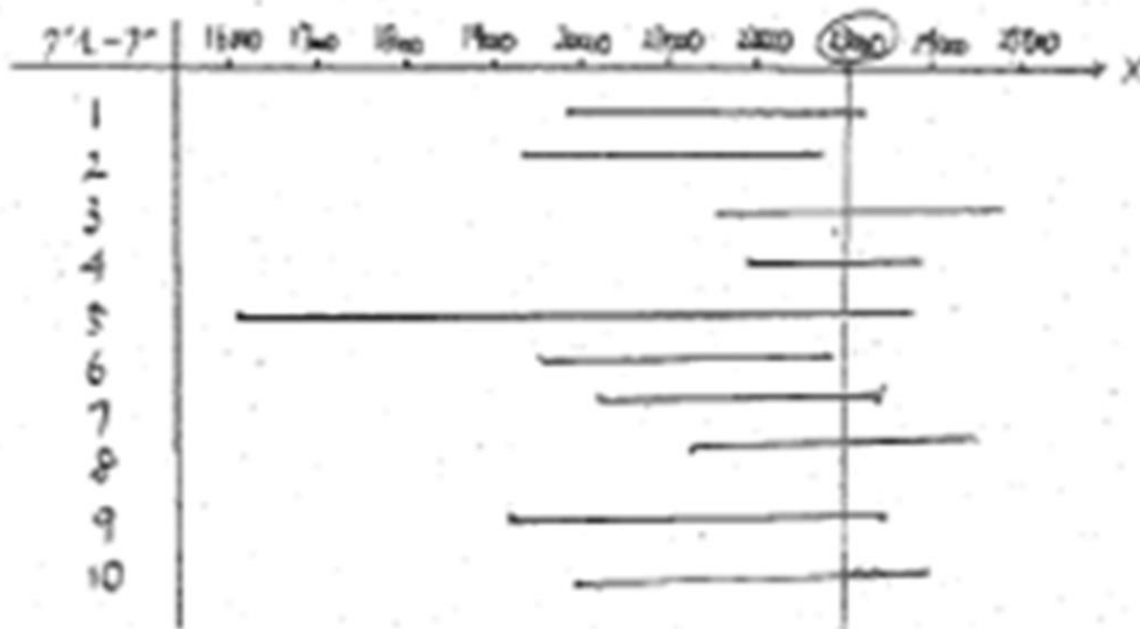
生徒F「グループ2と6で23000が不等式の範囲の右側に外れています。」

生徒E「(図から)標本が変わると不等式の範囲が変化することがわかります。」

また全てのグループで不等式の範囲は23000より左側によっています。」

生徒F「これらの結果から見出しの総数は約23000といっても、実際は23000より少ないのかな。全てグループで22000は不等式の範囲に入っているしね。」

授業者「区間推定することで、見出しの総数について不等式の範囲で考え、その結果の状況の比較から統計的には23000より少ない可能性があるという話になりました。」



指導事例「正規分布を利用した区間推定」

(3)生徒は活動の意義をどう感じたか **〔授業後インタビュー〕**

- 予想した値の範囲に実際の値が入っているのがすごいと思った。また、入っていないグループの結果から何故かを考えるのが楽しかった。
- 予想から、公表されている実際の値は本当かと思うようになった。確かめる方法を知ることができたのがよかった。
- n の値を大きくすると、信頼区間の幅が狭くなること、標本の標準偏差の平均が一定の値に近づくことが分かった。
- 見出し約23000語収録と記載された古語辞典の「約」という表現が絶妙だと思った。安易に信じてはいけないと思った。
- 各グループの結果を表や図に整理すると、分かりやすくなると思った。

数学の単元として指導するために

- 統計の科目ならば、週に1回(50分)で35回
→ 単元として扱うと週に2回($50分 \times 2$)で35回
- 数学の授業として、どのような数学的活動を行うのかのプランが必要
- 式やグラフに関して、時には要点のみを示すことにして、その活用を扱うことも大切
- 様々な授業の適切な場面でICTを活用