数学科教育研究法 D

# 区間推定の考えの理解を目指した教材

M18-1704　嶋崎　翼
M18-1706　瀬戸彩菜
M18-1707　武塙 健

【単元名】
数学B 「統計的な推測」

【単元の目標】
・具体的な事象において区間推定の考え方を理解すること。
・不確実な事象の起こりやすさに着目し、実験を通して、主張の妥当性について、批判的に考察すること。

【単元観】
データの活用領域では，中学 1 年生，中学 2 年生で不確定な事象の起こりやすさについて考察することをしてきた．また，中学 3 年生では，標本調査を行い，母集団の傾向を推定し判断することをしてきた．この既習事項をふまえ，数学 I では，「仮説検定の考え」を学習をしていく．

高等学校学習指導要領解説理数編には，以下のように記されている．

「具体的な事象において仮説検定の考え方を理解するとともに，不確実な事象の起こりやすさに着目し，主張の妥当性について，実験などを通して判断したり，批判的に考察したりすること」（文部科学省，2018，p.48）

ここでは，「飛行機の定時運航率の比較問題」を題材とし，エクセルを使った無作為抽出の実験を通して，区間推定の考え方を理解することを課題とする．

【教材観】
第 1 時では，一回の標本調査からの推測は信用性が低く，多数回標本調査を行うことで，真の値に近づくことを実感させる．また，「標本数を増やして，試行回数を減らす」または「標本数を減らして，試行回数を増やす」のどちらが良いかを考えさせる状況も作っている．各班で出た分布から，どちらの航空会社の方が優秀であるかを判断させる．そこで，基準とする点によって判断が異なることを実感させる．

第２時では，各班で出した結論を基に，クラス全体での結論を出す．その際に，平均値や最頻値などの代表値に着目することが考えられる．しかし，分布を見ると，いつもそうなるとは限らないため，判断できないことに気がつく．一点を決めて判断しようとすると，困難であるため，誤差に注目して範囲から判断することを経験させる．その経験から，仮説検定の考えについて理解させる．

【指導計画（２時間扱い)】

| 時間 | ねらい | 学習内容・学習活動 |
|---|---|---|
| １時間目 | １回の標本調査では，信頼性が低いことを実感する．<br>代表値や分布など，基準とする点によって判断が異なることに気づく． | 標本調査を通して，班ごとにどちらの航空会社の方が優秀かを判断させる． |
| ２時間目 | 代表値や分布など，基準とする点によって判断が異なることを理解する．<br>仮説検定の考え方について理解する． | 第１時で班ごとに出した結論から，クラス全体の結論を出す．<br>１点に着目して判断を下すことだけでなく，誤差を考えて，範囲に着目して判断を下す． |

【第１時の指導】

（１） 第１時のねらい

・１回の標本調査では，信頼性が低いことを実感する．

・代表値や分布など，基準とする点によって判断が異なることに気づく．

（２） 第１時の展開

| 時間 | 学習内容・学習活動<br>○質問・発問・指示　・学習活動<br>S 生徒の反応　　T 教師の行動 | 指導上の留意事項<br>・留意点　◇評価規準(評価方法) |
|---|---|---|
| 導入 | （飛行機の写真と架空の航空会社の遅延率を提示する）<br>T：飛行機はよく遅延をするが，実際どれだけ遅延しているのか．JEL 社の全数調査では，定時運航率は 73.8%だった．これをライバル社の ENE 社と比べて，自社の方が優秀だということを宣伝に使いたい．しかし ENE 社の全数調査のデータは提供してもらえず，自社でできる範囲内で調べることしかできない．（予算的にも全数調査ができない．）ではどのようにして ENE 社の定時運航率を調べ，自社と比べてどうかを調べられるだろうか．<br>S：標本調査を行う． | |
| | ○標本調査をもとに他社の遅延率を推測しよう<br>・班ごとに標本調査を行う．<br>S：10 回のうち，遅延が 8 回<br>S：10 回のうち，遅延が 7 回<br>S：10 回のうち，遅延が 6 回… | ・標本調査はデータを無作為抽出して行うことを確認する．<br>◇コンピュータを用いて，無作為に標本を取り出し整理すること（思考力，判断力，表現力） |
| | T：自社の方が遅延率が高いと思う班は手を挙げて．自社の方が遅延率が低いと思う班は手を挙げて．<br>【「高い」が 6 班，「低い」が 4 班と仮定する】 | |
| | ○この結果を受けて，自社の遅延率は高いと判断すべきか．低いと判断すべきか．<br>S1：高い(もしくは低い)という班の方が多いか | ◇実験を通して判断し，批判的に考察すること（思考力，判断力，表現力）<br>・S1 の反応の要因として，次の 3 点が考えられ |

| | | |
|---|---|---|
| | ら高い(もしくは低い).<br>S：標本調査なので，確実には判断できない<br>S：1 回の標本調査でしか判断していないそれぞれの班の結論は信用性が低い.<br>S：班の数にそれほど差がないということは，ほとんど違いがないのではないか.<br>→S1：このように判断することはできない.<br>S2：判断できない.<br><br>○各班で 1 回標本調査をするだけでは不十分でした．どのようにしたらよいか.<br>S3：それぞれの班で行った標本調査の結果を合わせればよい.<br>S4：それぞれの班で，より多くの母数の標本調査をおこなえばよい.<br>S5：それぞれの班で多数回の標本調査をおこなえばよい.<br><br>○まずは各班で 10 回ずつ標本調査を行い，分布をかいてみよう.<br>・班ごとに 10 回ずつ標本調査を行い，分布をかく. 平均や最頻値など，何に注目するかどうかなどの議論の時間を設ける.<br>・各分布を大きめの紙に書いてもらい，黒板に貼る. そのさい，自分たちの班で判断の基準にしたものを書き込むなど，あとで説明しやすいようにしてもらう. | る.<br>➢ 標本調査と全数調査は同じ結果が出る.<br>➢ 標本調査の結果と真の値の差は，母数の大小に無関係である.<br>➢ 標本調査の結果はあまり幅広くならない.<br><br>これらの点を生徒が指摘することで，この考えが適当でないことを明らかにする.<br><br>◇標本調査の結果を批判的に考察すること（思考力，判断力，表現力）<br>S3・S4 に関しては，のちの議論でも適応できる考え方であるため，まずは S5 の考え方である，複数回の標本調査を行うこととする. |
| 展開 | ○標本調査の結果をどのように見ればよいか.<br>・各班の説明と，それに対する意見を募る.<br>（判断できる：SA　判断できない：SB）<br>平均値<br>SA1：標本調査による結果の平均値が自社の遅延率より低いため，自社の遅延率は低いと判断できる.<br>SB1：標本調査による結果の平均値が自社の遅 | ◇不確実な事象の起こりやすさに着目し，主張の妥当性について，批判的に考察できているか（思考力，判断力，表現力）<br><br>生徒が判断に用いると考えられるのは，これまでに学んだことのある平均値，最頻値と，データの分布を視覚的に表したことによる視覚的な情報（データの多さ，分布の形）である. |

| | | |
|---|---|---|
| | 延率と近いため，判断できない.<br>最頻値<br>SA2：標本調査による結果の最頻値が自社の遅延率より低いため，自社の遅延率は低いと判断できる.<br>SB2：標本調査による結果の最頻値が自社の遅延率と近いため，判断できない.<br><br>データの個数<br>SA3：標本調査による結果のうち，自社の遅延率より低い結果の方が，自社の遅延率より高い結果よりも多くあるため，自社の遅延率は低いと判断できる.<br>SB3：標本調査による結果のうち，自社の遅延率より低い結果と，自社の遅延率より高い結果の個数はそこまで変わらないため，判断できない.<br>分布の形<br>SA4：データの分布を見て，自社の遅延率より高い部分と低い部分を比べると，高い部分の方が多く見えるため，自社の遅延率は低いと判断できる.<br>SB4：データの分布を見て，自社の遅延率より高い部分と低い部分を比べると，あまり変わらないように見えるため，判断できない.<br>SB5：すべてのデータが自社の遅延率より高いわけではないため，判断できない.<br><br>T：今回発表されたいろいろな意見については，次回議論していきます. | |

【第２時の指導】

（１） 第２時のねらい

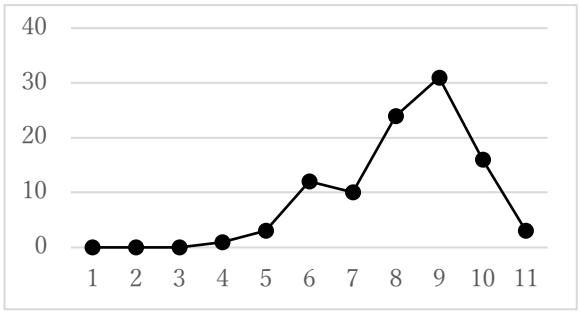
・代表値や分布など，基準とする点によって判断が異なることを理解する．

・仮説検定の考え方について理解する．

（２） 第２時の展開

| 時間 | 学習内容・学習活動<br>○質問・発問・指示　・学習活動<br>S 生徒の反応　　T 教師の行動 | 指導上の留意事項<br>・留意点　◇評価規準(評価方法) |
|---|---|---|
| 導入 | ○標本調査の結果をどのように見ればよいか。<br>（前回の議論の内容を提示しながら，足りない部分や，どのように決めていけばいいかについて議論する．）<br>（判断できる：SA　判断できない：SB）<br>平均値<br>SA1：標本調査による結果の平均値が自社の遅延率より低いため、自社の遅延率は低いと判断できる。<br>SB1：標本調査による結果の平均値が自社の遅延率と近いため、判断できない。<br>最頻値<br>SA2：標本調査による結果の最頻値が自社の遅延率より低いため、自社の遅延率は低いと判断できる。<br>SB2：標本調査による結果の最頻値が自社の遅延率と近いため、判断できない。<br>【1･2 に関する議論】<br>S：平均値(最頻値)より、例えば 0.001 でも低ければ、自社の遅延率の方が高いと判断してよいのか。ダメなら、いくつからいいと判断できるのか。<br><br>データの個数<br>SA3：標本調査による結果のうち、自社の遅延率より低い結果の方が、自社の遅延率より高い結果よりも多くあるため、自社の遅延率は低いと判断できる。<br>SB3：標本調査による結果のうち、自社の遅延率より低い結果と、自社の遅延率より高い結果の個数はそこまで変わらないため、判断できない。 | 生徒が判断に用いると考えられるのは、これまでに学んだことのある平均値、最頻値と、データの分布を視覚的に表したことによる視覚的な情報（データの多さ、分布の形）である。<br>◇標本調査の結果を批判的に考察すること（思考力、判断力、表現力）<br><br>◇不確実な事象の起こりやすさに着目し、主張の妥当性について、批判的に考察できているか（思考力、判断力、表現力）<br><br>感覚的な判断に対しては、その判断をより明確に、他者に伝わるように表現するよう求める。 |

| | | |
|---|---|---|
| | 分布の形<br>SA4：データの分布を見て、自社の遅延率より高い部分と低い部分を比べると、高い部分の方が多く見えるため、自社の遅延率は低いと判断できる。<br>SB4：データの分布を見て、自社の遅延率より高い部分と低い部分を比べると、あまり変わらないように見えるため、判断できない。<br>【3･4 に関する議論】<br>S：個数が多いか少ないか、あまり変わらないかの基準は何か。<br><br>SB5：すべてのデータが自社の遅延率より高いわけではないため、判断できない。<br>【5 に関する議論】<br>S：外れ値が 1 つでもあれば認めないのか。たとえば、10000 回の標本調査の結果が自社の遅延率より低くても、10001 回目の標本調査の結果が自社の遅延率よりも高くなったら、それだけで判断できなくなるのか。<br><br>○それぞれの議論の焦点を明確にする。<br>【1･2 に関する議論より】<br>近さを判断する基準が必要になる。<br>【3･4 に関する議論より】<br>データの個数や分布の形を判断する基準が必要になる。<br>【5 に関する議論より】<br>1 つ 1 つのデータを見て判断するのではなく、あくまで全体の傾向から判断するべきである。 | |
| 展開 | T：班によって、色々な根拠から判断できる、判断できないと考えることができた。しかし、各班の結論はバラバラ。クラスで 1 つの結論を出したいが、どうするべきか。<br>S：すべての班の結果を合わせる。<br>T：どのようにして合わせようか。<br>S6：各班が標本調査で集めたデータをまとめて、母数 | |

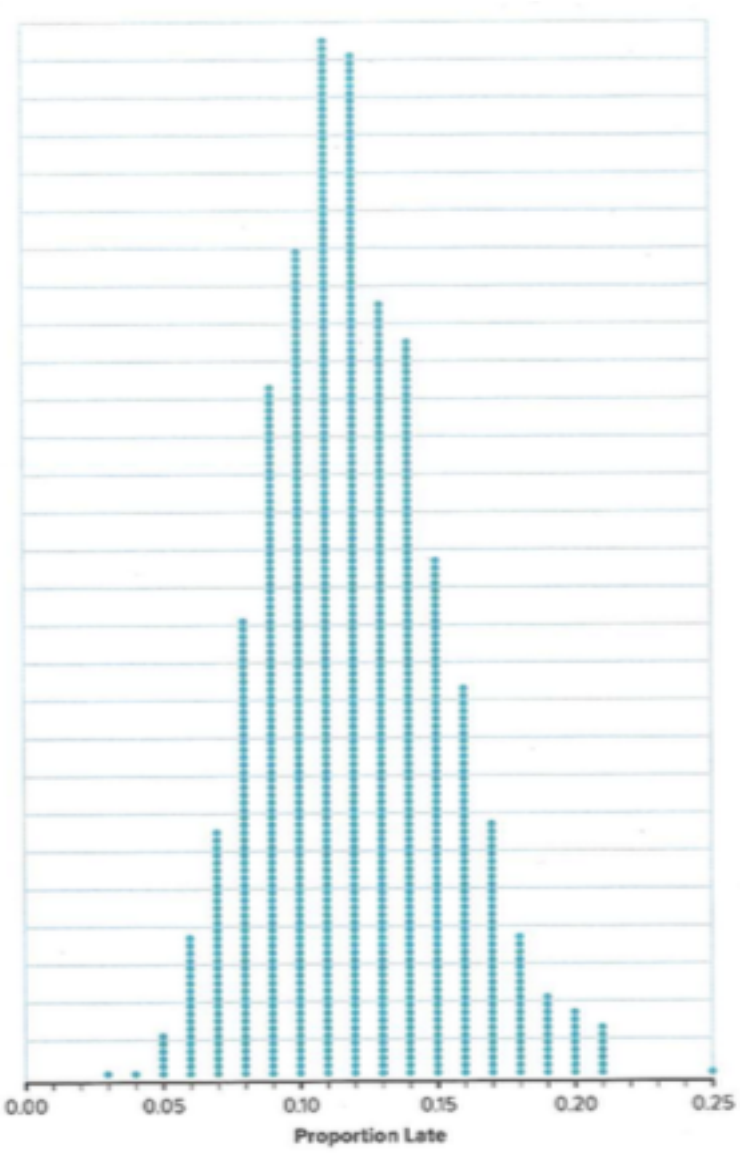
| | | |
|---|---|---|
| | 100 の標本調査として捉え、その結果から判断する。<br>  <br><br><br>S7：各班の標本平均の平均を求め、その結果から判断する。<br>→72.5%<br><br>S8：各班の標本平均を集め、それから判断する。<br>→小さい方から順に次の 10 個<br>7,7.1,7.1,7.1,7.2,7.2,7.3,7.4,7.5,7.6<br><br>【S6・S7 に関する議論】<br>S：S6 の考え方だと、結局近さがわからないから判断できない。<br>S：でも、各班のデータを集めた結果だから、この平均値は信用していいんじゃないかな。<br>S：S6 の平均値と S7 の値は一緒になる。<br>S：計算順序が違うだけで、結局同じものになっている。<br>S：実際標本調査で 72.5 以外の数値が出ているものもある.<br>S：もう一回標本調査を行うと数値が変わってしまう.<br><br>【S8 に関する議論】<br>S：これも結局平均をとったら S7 と同じ。<br>S：最小値 7・最大値 7.6 ってことは、高いか低いか判断できないんじゃないかな。<br>S：それだと、また外れ値があったら問題になるんじゃないのかな。<br>S：1 つ 1 つのデータじゃなくて、標本平均だから、外 | S6・S7 は点推定の考え方、 |

| | | |
|---|---|---|
| | れ値はそれほど影響しないんじゃないかな。<br>S：これまでは 1 つの値で判断してたけど、最小値と最大値を決めて判断しちゃいけないのかな。<br><br>T：S6・S7 の考え方と S8 の考え方は、どのような点が同じで、どのような点が異なっていましたか。<br>S：どちらも、データ全体を用いて判断しようとしている。<br>S：S6・S7 の考え方は、すべてのデータを 1 つの値として捉えている。S8 の考え方は、すべてのデータからある範囲を求めている。<br>S：S6 は判断基準が明確なので比べやすくてよいが，真の値を明確に評価することはできず，少し心配.<br>S：S7 は値の範囲で考えられるので真の値からずれたことは言わないが，比べにくい.<br>S：範囲があると，調査の誤差の範囲まで考えることができるので，範囲内の時は言い切れないが，範囲を超えたときに自信を持って大小を言い切れる. | |
| まとめ | T：標本調査の結果をほかの値を比べる時，標本調査の結果と真の値の間にはずれが生じた．それを表現するためには，標本平均のような値以外に，範囲を使うことで，標本調査の結果のずれまで考慮した比較が行える．このような考え方に信頼区間というものがあり，今回のように標本調査で真の値を表現する時に使うほか，研究の調査などでも用いられることがある. | |

【原本の問題】

６ヶ月間で、アラスカ航空のフライトのうち１２％が時間通りに到着しなかった。ジャスミンはこの割合を知らないで、それを推定しようとするために１００の出発便の無作為標本を取った。

a. n の値は？p の値は？
b. ジャスミンの１００便の無作為標本では、２０便が遅れた。遅れている全てのアラスカ航空便の割合のジャスミンの点推定$\hat{p}$は何か？彼女の標本誤差は何か？
c. 以下おおよその標本分布は、１００の出発フライトのそれぞれ１０００の異なる無作為標本からの p の値を示している。プロットから、サイズ１００の無作為標本が、ジャスミンの得たサンプルと同じくらい高いか、さらに高い割合を推定しなさい。



d. プロットから、0.05 以下の標本誤差を持つ標本分布の割合を推定しなさい。

e. この状況で$2\sqrt{\frac{p(1-p)}{n}}$の値を計算しなさい。この文脈でこの値が何を意味するかを説明しなさい。

【解答】

a. n=100、p=0.12

b. $\hat{p}$=0.20

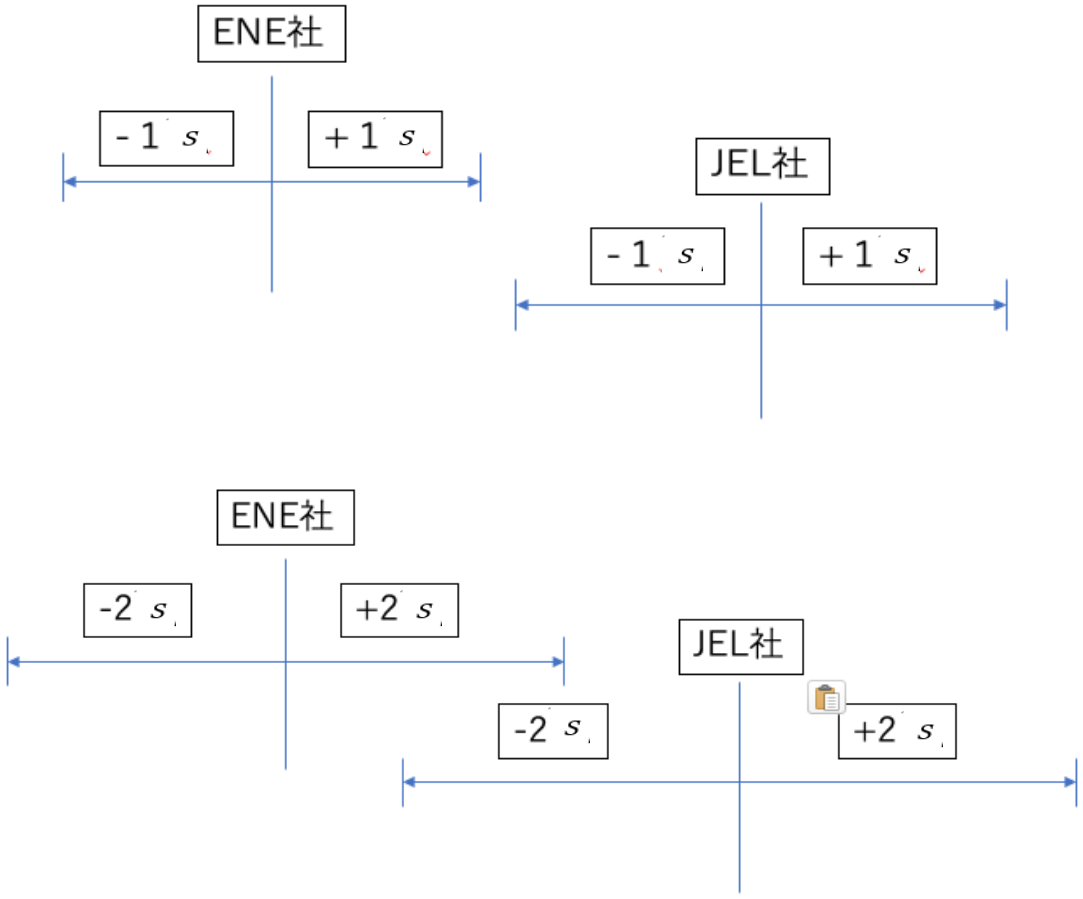
標本誤差=|0.20-0.12|=0.08

c. $\frac{17}{1000}$=0.017

d. $\frac{927}{1000}$=0.927

e. $2\sqrt{\frac{p(1-p)}{n}}=2\sqrt{\frac{0.12(1-0.12)}{100}}$=0.06499=0.065

0.135≤p≤0.265

この範囲に、p=0.12 が入っていないため、$\hat{p}$は遅すぎると判断することができる。

信頼区間の指導の前に、仮説検定の考え方では、標準偏差（s）と信頼区間の関係を明らかにする必要がある。標準偏差は、(平均値)±１sには、約６８.３％が分布しており、(平均値)±２sには、約９５.５％が分布しているという特徴を持っている。例えば、下のように、分布している区間が重なっている場合と重なっていない場合とでは、どのように判断が変わってくるかを理解させることも大切である。信頼区間の前に、このような群間比較をすることも意義がある。