

# 新学習指導要領における統計分野の内容について —新数学 I に追加された内容を中心に—

まつい しんや  
松井 真也

本稿では平成30年3月に告示された学習指導要領で改訂された統計領域の内容について、数学 I を中心に解説します。数学 I におけるデータの分析の主な変更点は次のとおりです。以下、学習指導要領を「指導要領」、同解説を「解説」と書きます。

- ・四分位範囲、箱ひげ図が中学校に移行
- ・「仮説検定の考え方」の追加
- ・「質的データ」と「分割表」の扱いが解説に記載
- ・用語「外れ値」が指導要領に、その目安が解説に記載

このうち、仮説検定と分割表について以下で解説します。

## §1. 「仮説検定の考え方」について

そもそも仮説検定とはなにか。端的に言えば

仮説検定とは、確率を用いた「背理法」

です。仮説検定と背理法はよく似た概念なので、比較すると理解が深まると考えられます。

### 〈背理法の流れ〉

命題 P が真であることを示したいとき

- ① P が偽であると仮定する。
- ② 仮定に「矛盾がある」ことを示す。

### 〈仮説検定の流れ〉

ある確率  $p$  に伴って事象 A が起きている時、この確率  $p$  が「 $p = \alpha$  ではない」と主張したいとき

- ①  $p = \alpha$  であると仮定する。
- ②  $p = \alpha$  であるとき、「事象 A が起こる確率が極端に小さい」ことを示す。

背理法における「矛盾」が、仮説検定における「事象 A が起こる確率が極端に小さい」ことに対応します。ここが理解できれば、仮説検定はそれほど難し

いものではありません。

具体例を1つ示します。

「コインを10回投げたら9回表が出た。このコインの表が出る確率は  $\frac{1}{2}$  であると言えるか。」

示したいことは「 $\frac{1}{2}$  ではない」です。これを**対立仮説**と言います。「ではない」ことを示すことに注意してください。仮説検定を行う時、まずはこの対立仮説を否定して「 $\frac{1}{2}$  である」と仮定します。これを**帰無仮説**と言います。帰無仮説のもとで、表が9回以上出る確率を求めます。なぜ9回以上なのかは後で述べます。数学 A で学ぶ反復試行の確率により求める確率は

$${}_{10}C_9 \left(\frac{1}{2}\right) \left(\frac{1}{2}\right)^9 + \left(\frac{1}{2}\right)^{10} = \frac{11}{1024} \doteq 0.0107\cdots$$

と計算できます。この0.0107……を「極端に小さい」と判断すれば、表が出る確率は「 $\frac{1}{2}$  ではない」と主張できるのです。この「極端に小さい」と判断する基準の確率を**有意水準**と言います。一般的にこの有意水準は5%、または1%とすることが多く、このケースでは5%有意です。ただし、有意水準を1%とすると、先ほどの確率は1%より大きいので「極端に小さい」とは言えなくなります。つまり「表が出る確率は  $\frac{1}{2}$  ではないとは言いきれない」

ことになります。「表が出る確率は  $\frac{1}{2}$  である」と断言するのは間違いです。あくまでも「ではないとは言えない」です。ここも仮説検定の大切なポイントですので押さえてください。数学 I で扱う「仮説検

定の考え」では、この程度のことまで押さえておけば十分です。

〈「仮説検定の考え」で押さえないこと〉

- ・仮説検定は、確率を用いた「背理法」である。
- ・対立仮説と帰無仮説を立て、帰無仮説のもとで「現実起こったことが起きる確率」を計算し、求めた確率が「極端に小さい」ことを示す。
- ・極端に小さいと判断して対立仮説が正しいと判断する。

ちなみに、9回以上の確率を求めた理由は2つあります。1つは、表が10回出ているということは、少なくとも9回は表が出ている、つまり10回表ということは9回表が出て、さらに1回表が出たと考えるということです。もう1つはコインを投げる回数が増えた場合、「ちょうどの確率」はほぼ0になるということです。例えば10回投げて「ちょうど」9回表が出る確率はそれなりにありますが、10000回投げて表が「ちょうど」9000回の確率はほぼ0になってしまい、検定を行うことができません。表が9000回以上出る確率ならばそれなりの値になり、検定を行うことができます。

## §2. 「仮説検定」について

数学Iのレベルでは§1.で述べた程度のことを理解すれば十分です。ここからはもう少し詳しく仮説検定について説明します。

仮説検定は昭和53年告示の学習指導要領までは高校の教科書でも扱われていた内容でした。当時の指導要領では確率に関する内容が「確率・統計」という科目にまとめられており、系統立てられた指導がなされていました。確率・統計の学習内容は多くの教科書で

場合の数→確率→確率分布→区間推定→仮説検定の順でした。平成元年告示の指導要領で仮説検定が学習内容から外れ、実際は期待値までを学習させるケースがほとんどになりました。その期待値も平成21年告示の指導要領で数学Bに移行し、ほとんど生徒が期待値を学ばないことになりました。今回の改定で期待値は数学Aに戻ってきます。条件付き確率も残りますので数学Aの学習内容は単純に増えることになります。その上30年近く高校では教えられてこなかった仮説検定を、しかも必修の数学Iで指導することになります。確率は数学A、期待値を

除く確率分布は数学Bの内容ですから、これだけでも従前のような指導はできないことがわかります。

先ほどの「仮説検定の考え方」で押さえないことで「帰無仮説のもとで現実起こったことが起きる確率を計算」と書きました。そもそも、数学Aを学んでいなければこの確率を計算で求めることはできません。解説には「不確実な事象の起こりやすさに着目し、実験などを通して、問題の結論について判断したり、その妥当性について批判的に考察したりできるようにする。」とあるように、数学Iにおける「仮説検定の考え方」では「実験結果に基づく統計的確率を用いて」検定を行うこととなります。したがって、数学Iにおいては検定を行うためには「実験結果」が準備されている必要があります。例えば解説にある「コインを30回投げるという試行」を何回か(実際には何千回という単位)繰り返した結果が示され、その結果をもとに「出現回数/試行回数」という単純な計算で統計的確率を求めて検定することになります。このように、「帰無仮説のもとで起こる確率」を直接求めて検定する方法を**直接法**、もしくは**正確検定**と言います。また、求めた確率の値を**p値**と言います。直接法による検定が高等学校で扱われるのは今回が初めてであろうと思われます。では、以前はどのように検定していたかという点、まず<sup>(※)</sup>確率分布から区間推定によって信頼区間を求め、実際に起こった結果がこの信頼区間の外側にあるかどうかで判断します。この信頼区間の外側のことを**棄却域**と言います。非常に端折って書きましたので何のことかわからないかもしれませんが、この(※)の信頼区間を求めるところまでは、現行の指導要領でも数学Bで扱っている内容です。教科書にも詳しく書かれていますので、お手元の数学Bの教科書にある確率分布の部分をご覧ください。確率分布や区間推定についての詳細は教科書をお読みいただくことにして、ここで押さえておきたいことは、この方法では「帰無仮説のもとで起こる確率」を求めていないという点です。このように棄却域を定めて検定する方法を**間接法**と言います。間接法の手順を以下に示します。

〈仮説検定の手順(間接法)〉

- ・対立仮説と帰無仮説を立てる。
- ・有意水準を定め、帰無仮説のもとでの確率分布を仮定して信頼区間を定める。

- ・実際に起こった事象が信頼区間の外側にあれば対立仮説が正しいと判断する。
- ・信頼区間の内側にあれば帰無仮説が正しいと言うわけではない。

このように、次期指導要領では従前の仮説検定の方法とは異なる方法を扱うことになります。実は今回の指導要領にはもう一つ大きな特徴があります。それは対立仮説の立て方です。対立仮説の立て方には2種類の考え方があります。例えばAかBかを回答する二者択一のアンケートをとった場合、

- ① AはBより多いと言えるか。
- ② AとBに差はあると言えるか。

と言う2種類の仮説の立て方があります。①は実際にAが多いという結果があって、それが統計的にも多いと言っていいかどうかという問題です。②はそもそも今回たまたまAが多かっただけで、実際はBが多いという可能性も考えるという問題です。従来は②の考え方が多く、信頼区間も閉区間で定めることが多かったので棄却域は信頼区間の両側に存在します。このような検定を**両側検定**と言います。一方①の場合、実際Aが多いという結果が出ているわけですからBが大きい可能性は考えません。とすると棄却域は片側だけに存在することになります。これを**片側検定**と言います。

数学Iでは基本的に片側検定しか扱いません。直感的にも片側検定の方が理解しやすく、「考え方」を指導するには直接法、片側検定で十分だからです。数学Bにおいても、解説のp.111にある仮説検定の方法に「(3) 帰無仮説  $H_0$  が真であると仮定した場合に事象Eが起こる確率  $p$  を求める。」とあるように、基本的には直接法をもとに指導することになります。ただし、手計算で直接確率を求めることはかなり困難なことから、入試においては間接法を中心に出题されると思われます。直接法、間接法両方の視点、片側検定、両側検定の両方の考え方を指導する必要があると個人的には考えています。

### §3. 質的データと量的データについて

ここからは**分割表**について解説をします。分割表、もしくはクロス集計表とは、2つ以上の変数の間の関係を記録し、分析するための表です。Wikipedia ([1]) には、以下のような2×2分割表の例があります。

(表1)

	右利き	左利き	計
男性	43	9	52
女性	44	4	48
計	87	13	100

上の表のように、分割表は一般に質的データの関係を示すのに用いられます。**質的データ**と**量的データ**の扱いは小学校で既に学習しており、解説には、「小学校算数科では、文字情報として得られる「質的データ」と数値情報として得られる「量的データ」の双方を、中学校では、主に後者を取り扱っている。「数学I」では、「質的データ」と「量的データ」の双方や、複数の「質的データ」や「量的データ」が紐付けられた複数の種類のデータを取り扱う。」と記載があります。

分割表を扱うにあたり、まずは質的データと量的データの違いを簡単におさえておく必要があります。

- ・質的データ…右利き、左利きのように分類や種類を区別するためのデータ。順序の無い名義尺度と順序のある順序尺度の2種類に分けられる。
  - ・量的データ…気温や身長など、数値として意味のあるデータ。気温(°C)のように比には意味がない間隔尺度と身長のように比にも意味がある比例尺度の2種類に分けられる。
- また、身長のように連続な値を取りうる連続データと、人数や個数のように離散的な値を取る離散データの2種類にも分けられる。

中学校では量的データを中心に扱われていますが、解説では「質的データ間の関係を探る際には、データを漏れないように分割した二次元表に整理したり、一旦整理したデータを、さらに異なる性質で漏れないように分類したりする」とあるので、高等学校の授業において分割表の読み方を扱うことになると思われます。

### §4. 分割表を指導するにあたっての注意点

分割表を読むにあたって、是非おさえておきたいポイントを述べます。解説には以下のような「合否が判定される試験において、ある本を使って学習をしたかどうかを尋ねた結果例」が挙げられています。

(表2)

	合	否
使用有	65%	35%
使用無	49%	51%

先程の例と違い、ある本の使用の有無と合否の関心の比率しか示されていません。このように百分率で示した表は厳密には**百分率分割表**と言い、分割表とは区別するのが普通です。百分率分割表だと元のデータの大きさがわからないだけでなく、実際の現象を見誤ることがあります。実際、解説には度数は次の表3の通りであり「度数を見ることで、本の使用者自体が少なかったこともわかる。」と述べられています。

(表3)

		合	否
使用有	1学年	3	4
	2学年	8	2
使用無	1学年	12	38
	2学年	32	10

上の分割表をもとに使用の有無と合否の関係を表す分割表を作り直すと以下ようになります。

(表4)

	合	否
使用有	11	6
使用無	44	48

先程の百分率分割表と合わせて、以下のような表を作るのが一番正確でわかりやすいです。

(表5)

	合	否
使用有	65% (11)	35% (6)
使用無	49% (44)	51% (48)

(括弧内は度数を表す)

分割表を読む際は、比率にとらわれると実際の現象を見誤ることがあることに注意するとともに、データの大きさを意識することの重要性を指導する必要があります。データの大きさが重要であることは、仮説検定を指導する上でも重要なポイントです。

先程の表3は実は単純な分割表ではなく、「使用の有無と合否」という分割表と「学年と合否」「使用の有無と学年」という3つの分割表を合わせたものになっています。このように複数の分割表をまとめたものを**多重分割表**(多重クロス表)と言います。数学Iの段階ではこのような多重分割表から表4のように分割表を作ることや、分割表の度数からある程度のデータの傾向を読み取ることができれば十分であると考えます。

## §5. 補足

表2の分割表は明らかに「本の使用の有無が合否

に影響を与える」という因果関係を前提に作られています。分割表だけから因果関係を読み取ることには原則できません。因果関係を読み取るとは統計学の分野ではなく、データの背景にある理論に依存する、つまり解釈によるということには気をつけてください。また、表4から表2は行合計が100%になるように作られています。これは「百分率分割表を作る際、因果関係を前提とする場合は原因側が100%になるようにまとめる」という原則があるためであることを補足しておきます。

分割表で興味があるのは「2つの変数間に関係が有るかどうか」です。例えば先程の表2の例だと「本の使用の有無と合否に関係が有るか無いか」という点です。大学の統計学の範囲ですが「独立性の検定」というものがあり、 $\chi^2$ 統計量というものを計算して近似的に検定することができます。詳しい説明は以下の文献に任せますが、 $\chi^2$ 統計量の計算自体は2×2分割表であれば電卓で計算できる程度なので、紹介していただければ生徒の興味関心を引き出せるのではないかと思います。

### 《参考文献》

指導要領及び解説以外の主な参考文献をあげます。

- [1] 分割表-Wikipedia  
<https://ja.wikipedia.org/wiki/分割表>  
 統計学の基礎的な部分を理解するには、高専で使われている確率・統計の教科書が良いと思います。
  - [2] 新確率統計 高遠節夫 他著 大日本図書
  - [3] 新版確率統計 岡本和夫 他著 実教出版  
 もう少し専門的なものとして、
  - [4] 独習 統計学 24 講
  - [5] 独習 統計学 応用編 24 講  
 いずれも 鶴田陽和 著 朝倉書店
  - [6] 統計検定 2級対応 統計学基礎  
 日本統計学会編 東京図書
  - [7] 統計学入門 東京大学教養学部統計学教室編  
 東京大学出版会
- などがあります。また、分割表について詳しく書かれた専門書には、上記の〔5〕の他に
- [8] 分割表の統計解析 -二元表から多元表まで-  
 宮川雅巳・青木敏 著 朝倉書店
- があります。

(岐阜県立各務原高等学校)