

理科教育に関する調査や 学習評価における統計的处理について

柳本 高秀

筆者は、今年度当センターの第2回研修会（コンピューターを活用した効果的な評価の測定について）において、講師を担当した。本稿では、その研修会から、理科教育に関する調査や学習評価における統計的な処理を、有意差に着目して実施した内容について、その一部を報告する。

〔キーワード〕 理科教育に関する調査 学習評価 統計的处理 有意差

はじめに

学力・学習状況調査等、理科教育に関する調査が数多く存在するが、その調査結果には、平均値、中央値、標準偏差、分散等の様々な基礎的統計量が記載されている。また、SSHでの取り組みに代表される、研究開発における評価においても、実践した内容の検証・評価に、基礎的な統計処理を用いた分析がなされている。

このように、様々な教育効果の検証を定量的な分析で行う必要が叫ばれる中、教育現場においては、各種調査等に用いられてる基礎的な統計について、その内容が十分に理解されているとは言いがたい現状がある。

そこで、本稿では、評価に関する分析の観点を広げることを目的として行った研修会の内容について、その一部を報告する。

1 研修会の目的

昨年11月28日、情報処理教育センターを会場に、所員30名を対象に研修会を実施した。研修会の目的は、以下の通りである。

本日の研修会の目的

PCを活用してアンケート調査などの結果を統計的に処理する方法を、差の検定を中心に実習し、評価に関する分析の観点を広げる。

2 研修会のメニュー

研修会のメニューは、以下の通りであった。

本日のメニュー

1. はじめに、本日の研修会の目的
2. 統計の基礎用語と本日扱う内容
3. アンケート調査の評価・分析
4. 学習効果測定の評価・分析
5. まとめ

3 研修会の具体的内容

①統計用語

研修会の冒頭、基礎的な統計用語とその説明が記載されたプリントを配布した。

使用した用語については、表1に示した通りであった。

表1 研修会で使用した基礎統計用語

- ・統計的仮説検定
- ・代表値（平均値、中央値など）
- ・正規分布
- ・標準偏差
- ・カイ二乗検定
- ・帰無仮説
- ・対立仮説
- ・P値
- ・有意水準
- ・自由度
- ・分散
- ・t分布
- ・t検定
- ・対応のあるデータ
- ・対応のないデータ
- ・F検定
- ・等分散性の検定

②学テの分析を例とした基礎的統計量の意味

表2 学力テストの成績表

	児童数	平均正答数	平均正答率(%)	中央値	標準偏差
北海道(公立)	42,951	10.8 / 15	71.8	11.0	3.1
全国(公立)	1,080,663	10.9 / 15	72.9	12.0	3.1

表2の成績表を基に、中央値と標準偏差の意味について説明した。中央値とは、データを大きさの順に並べた時の真ん中に位置する値であり、北海道(公立)では、42,951人の真ん中、つまり、21,476番目の児童は、中央値の11問を正答したことがわかる。

また、標準偏差に関しては、「平均値±1標準偏差」の範囲には、全データの約68%が含まれることから、北海道(公立)では、「10.8±3.1」の値から、平均正答数で7.7から14.1の児童が約68%を占めていることがわかる。

③カウント数の差の検定

(1) 直接確率計算

一般的に、カウント数の有意差を分析するためには、直接確率計算とカイ二乗検定という2つの方法がある。そのうち直接確率計算は、主に、表3のような1×2表の分析に用いられる。

表3 試案Aに対する試案(人)

賛成	反対
2	7

表3の結果は、ある試案Aに対して、無作為に9人を選び、賛成と反対を尋ねた結果である。「賛成者と反対者の人数に実質的な差があるかどうか」を直接確率計算によって検定すると、

表4 直接確率計算の結果

直接確率計算

賛成 2 反対 7 合計 9 (N=9) この9人は無作為
「このような賛成者と反対者の人数に差があるといえるか」

賛成(r)	反対	組み合わせ	確率(p)	
0	9	1	0.002	0.2%
1	8	9	0.018	1.8%
2	7	36	0.070	7.0%
3	6	84	0.164	
4	5	126	0.246	
5	4	126	0.246	
6	3	84	0.164	
7	2	36	0.070	7.0%
8	1	9	0.018	1.8%
9	0	1	0.002	0.2%
		512	1.000	

直接確率検定では、まず、 ${}_N C_r$ (N個からr個を取り出す組み合わせの数)を計算する。例えば、表3では、9人の中から賛成者2人を選んでいるから、 ${}_9 C_2 = 9! / 2! \cdot (9-2)! = 36$ で、36通りということになる。

表4から分かるように、賛成者が2名の時と、賛成者が7名の時は、出現確率が同じとなる。また、賛成・反対の数が、(0・9), (1・8), (8・1), (9・0)の出現確率も賛成・反対(2・7), (7・2)より低いので、合わせて計算すると、 $p = 0.18 ((0.002 + 0.018 + 0.070) \times 2)$ となる。

一般的な統計的仮説検定での有意水準は5% ($p = 0.05$)であるので、「賛成者と反対者の人数は、5%の水準で有意ではない」ことがいえる。

(2) カイ2乗検定

有意差検定には、図1のように、条件間の比較を行うものもある。

1) 有意差のある例

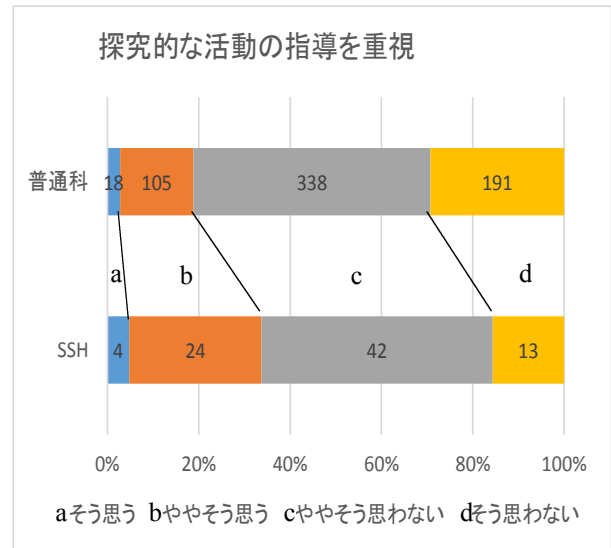


図1 普通科(SSH未指定)とSSH指定校での探究的な活動の重視の割合(教員向け)

図1の結果は、「探究的な活動の指導を重視しているか」との問いに関する結果であり、これらの結果において、普通科(SSH未指定)とSSH指定校の教員間で、肯定的意見と否定的意見に有意差があるのかを検定した。

表 5 図 1 をもとにした検定結果

	そう思う	ややそう思う	ややそう思わない	そう思わない	
普通科	18	105	338	191	
SSH	4	24	42	13	
実測値					
普通科		肯定的 123	否定的 529	合計 652	合計の割合 0.89
SSH		28	55	83	0.11
		151	584	735	1.00
期待値					
		肯定的 133.9	否定的 518.1	合計	
		17.1	65.9		
期待値の133.9は、151×0.89で計算 (同様に、17.1は、151×0.11)					
$\chi^2(1) = 9.973$, $p < .01$ 1%水準で有意					

肯定的意見として、「そう思う」、「ややそう思う」の人数を、否定的意見として「ややそう思わない」、「そう思わない」の人数を合算した結果は、表5の実測値に記載されている。

$$\chi^2 = \frac{(\text{実測値} - \text{期待値})^2}{\text{期待値}} + \frac{(\text{実測値} - \text{期待値})^2}{\text{期待値}} + \dots$$

図 2 カイ二乗検定におけるカイ二乗値

表5における期待値133.9とは、肯定的意見の合計151人に全体の割合0.89をかけたものである。また、カイ二乗値が大きいということは、期待値からのズレが大きいということ＝「つまり差がある」ということである。

表 6 χ^2 値分布

df	χ^2 値			df	χ^2 値		
1	2.706	3.841	6.635	8	13.362	15.507	20.090
2	4.605	5.991	9.210	9	14.684	16.919	21.666
3	6.251	7.815	11.345	10	15.987	18.307	23.209
4	7.779	9.488	13.277	12	18.549	21.026	26.217
5	9.236	11.071	15.086	14	21.064	23.685	29.141
6	10.645	12.592	16.812	15	22.307	24.996	30.578
7	12.017	14.067	18.475	16	23.542	26.296	32.000
出現確率	.10	.05	.01	出現確率	.10	.05	.01
有意水準	有意傾向	5%	1%	有意水準	有意傾向	5%	1%

検定の結果、カイ二乗値が9.973となった。この値は、表6の自由度(df)が1，出現確率0.01の6.635よりも大きいことから、SSH指定校と未指定校の間では、肯定的意見と否定的意見に、1%水準で有意な差があるといえる。

2) 有意差のない例

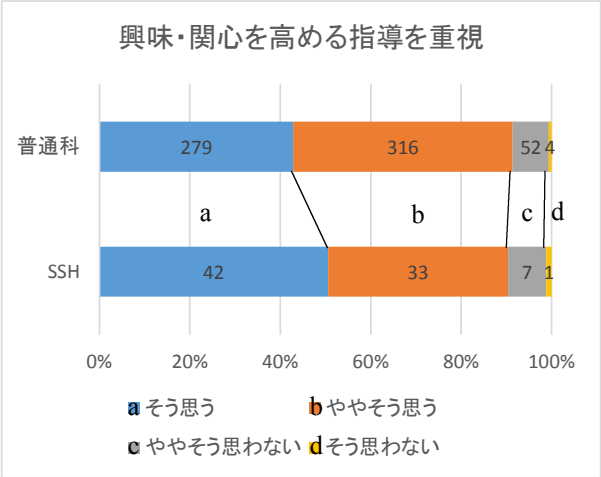


図 3 普通科（SSH未指定）とSSH指定校での興味・関心を高める指導の重視の割合（教員向け）

「興味・関心を高める指導を重視しているか」との問いに対して、普通科（SSH未指定）とSSH指定校の教員間で、肯定的意見と否定的意見に有意差があるのかを検定した。

表 7 図 3 における肯定的意見と否定的意見

	肯定的意見	否定的意見
普通科	595	56
SSH	75	8

前述と同様の方法を使って、検定の結果、カイ二乗値が0.099となった。

この値は、表6の自由度(df)1，出現確率0.10の2.706よりも小さいことから、SSH指定校と未指定校の間では、肯定的意見と否定的意見に、有意な差がないといえる。

④ 平均値の差の検定（t 検定）

ここからは、2つの群の平均値に有意差があるかどうかについて、述べていくことにする。

(1) 対応のない t 検定

表8のようなA, Bの2クラスで、10点満点の小テストを実施した際に、この2クラスの平均点に有意差があるのかを例に演習を行った。両クラスは、別々の生徒で構成されているので、対応のない t 検定を行うことになる。

表8 A, B両クラスにおける小テストの結果
(10点満点)

																平均値	SD
Aクラス	9	10	7	8	9	8	9	6	9	8	7	8	7	10	9	8.27	1.12
Bクラス	5	6	9	6	9	8	6	5	7	5	7	6	6	7	7	6.60	1.25

データの分析には、Excelを使用した。Excel 2013には、「分析ツール」の機能が強化された。

<Excel2013における分析ツールの読み込み方>

- i) ファイル→オプション
- ii) アドイン→管理→Excelアドイン→設定
- iii) 分析ツール→ok

表9 対応のない t 検定をした際のExcel画面

t-検定: 等分散を仮定した2標本による検定

	変数1	変数2
平均	8.26667	6.6
分散	1.352381	1.685714
観測数	15	15
仮説平均	0	
自由度	28	
t	3.703341	
P(T<=t) 片	0.000463	
t境界値 1	1.701131	
P(T<=t) 両	0.000925	
t境界値 2	2.048407	

の2値を比較

t境界値よりt値が大きいため、帰無仮説が棄却 t(28)=3.70, p<.05
AクラスとBクラスの平均点には5%水準で有意な差がある

表8のデータをExcelに入力し、分析ツールを活用することで、このような有意差に関する結果まで表示される(表9)。

(2) 対応のある t 検定

表10 生徒a～oのプレ、ポストテストの結果

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	平均値	SD
プレテスト	9	10	7	8	9	8	9	6	9	8	7	8	7	10	9	8.27	1.12
ポストテスト	5	6	9	6	9	8	6	5	7	5	7	6	6	7	7	6.60	1.25

対応のある t 検定とは、表10のような、クラ

ス内の a～o の15名の生徒における、プレテストとポストテストにおける平均点の有意差について検定するものである。2つのテストを同一の生徒が受けているので、対応のある t 検定となる。

なお、これらの t 検定の検定結果の詳細、ならびにそのほかの研修会の内容については、紙面の都合上、省略する。

4 研修会のまとめ

研修会の最後に、以下の4点について、まとめを行った。

まとめ

1. 統計のデータには、質的データ・量的データがある。
2. 基本統計量の意味を知ると、統計データの意味が見えてくる。
3. 有意差検定を行うと、その差の水準(例5%)がわかる。
4. 様々な報告書(SSHなど)における統計データを検証できる。

* 評価を数値であらわさねばならないとき、今後益々増え、教育的効果等を数値で示していくことが求められる。

5 研修会の評価

研修会の終了後、受講者から以下のような意見をいただいた。

- ・ 徐々にアカデミックな内容を学ぶことができた。調査等で使われている統計的な言葉の本当の意味がわかった。
- ・ 紹介してもらったフリーソフト (Js-STAR), Excel統計, 書籍などを早速、活用してきたいと思う。
- ・ 学校現場でも、様々な調査の分析・解釈は必要だと日頃感じていた。今日の資料をぜひ活用していきたい。

数多くの受講者から、研修会の内容について、今後の指導助言等にも、ぜひ積極的に活用していきたいという評価をいただいた。

主要参考文献

- 1) 福田周;『心理・教育統計法特講』,放送大学,(2013)
- 2) 田中敏;『ユーザーのための教育・心理統計と実験計画法』,教育出版,(2008)
- 3) 平井明代;『教育・心理系研究のためのデータ分析～理論と実践から学ぶSPSS活用法』,東京図書,(2012)

(やなぎもと たかひで 地学研究班)