TOPICS

分割表を用いた高校生の左右握力に関する統計解析

東京理科大学 理工学部 情報科学科 助教 生亀（いき） 清貴（きよたか）

はじめに

近年，統計学という学問が大変注目されています。ベイズ統計，ビッグデータ解析，データサイエンティストといった単語が次々にピックアップされており，さまざまな業界で統計学を用いて何か有用な知見を得ようという取り組みが行われています。もはや社会で生活する上で，統計データと係わることは必須になっているとまで言えるでしょう。

これらの統計解析は，パソコンの統計解析用ソフトウェアを用いて行われることがほとんどです。具体的にはMicrosoftのExcelやフリーソフトウェアのR，他にもSASやSPSSなどが挙げられます。パソコンの発達により，私たちは多少のプログラムの知識を身につけるだけで，高度な解析をパソコン上で手軽に行うことができます。しかし解析用ソフトを使うとき，パソコン内でいったいどのような処理が行われているかを，私たちはしっかり把握していると言えるでしょうか。

統計学というのは，数理的な学問です。まず数学があって，それを論理的に積み上げることによって解析手法は生まれます。この事実を忘れて，ただ統計ソフトの使い方を覚えるだけに終始するならば，たしかにデータをパソコンに入力することで表面上は解析が行えているように見えるかもしれませんが，それだけでは不十分です。本当にその解析法を使うことが正しいかどうか，データを正しい型式で入力できているかどうか，あるいは得られた結果を正しく解釈できているかといった疑念が常につきまとうからです。

統計学の理論を十分に知らずに解析ソフトを使うのならば，全く意味のない解析を行ったり，誤った結果を広めたり，世の中に氾濫する統計データに逆に踊らされるという結果になりかねません。したがって，統計学の数理的な面にも触れることで，統計的な思考力を養い，統計リテラシーを身につけることが重要です。

本稿では，筆者が取り組んでいる研究テーマの一つである，分割表の対称性に関する研究を元に，統計学におけるデータ解析がどのような手順で行われるかを述べます。本論で数理的な側面をすべて紹介することはできませんが，おおまかな流れを知ってもらえれば幸いです。

左右握力の関連性

人間の左右握力は，多くの人で右の方が強いと考えられています。しかしこの「強い」という言葉の定義が具体的に何を示しているかまでは，知っている人は少ないのではないでしょうか。同じ年齢層あるいは同じ職種などを対象とする一定の集団（母集団）において，その集団の人の左右の握力にはどのような関連性が存在するかを調べてみましょう。

表１は2011年から2015年にかけて，東京理科大学のオープンキャンパスに参加した男子高校生805名の握力を測定したデータです。表１を眺めてみると，例えば，右握力が(1)強い，左握力も(1)強い，という人が74名いるこ

とが分かります。もともと測定された握力データは，例えば右握力46.2kg，左握力41.0kgというように連続的な値ですが，ここでは文部科学省の体力・運動能力調査結果の全国平均値を元に，右握力，左握力ともに⑴強い，⑵普通，⑶弱い，という３段階に分類しています。このように連続的な値をいくつかの離散的な値に変換しデータを集約することは，統計的な解析を行いやすくするため，またシンプルな解釈を得るためによく行われます。

右握力と左握力が共に⑴強い，共に⑵普通，または共に⑶弱いという男子高校生を合わせると587名にのぼり，実に男子高校生全体の73％を占めることが分かります。多くの男子高校生は左右握力のバランスがとれており，バランスの取れていない人は左右握力の差が大きくなるにつれて少なくなる傾向にあります。したがって，男子高校生の左右握力には強い関連性があることがうかがえます。さらに，表１では，例えば，右握力が⑴強い，左握力が⑵普通という人が89名，一方で右握力が⑵普通，左握力が⑴強い，という人が10名おり，差にして79名，比率にして8.9倍も違っています。どうも右握力と左握力は対称的ではないように思われます。

そもそも一定の集団において左右握力が対称的であるとは，どのような場合を意味するのでしょうか。一言で対称的といっても，統計学では定義がいろいろとあるのですが，例えば，任意に選んできた人の右握力がi，左握力がj（$>i$）である確率（p_{ij}と記す）が，右握力がj，左握力がiである確率（p_{ji}と記す）に等しいことを意味します。すなわち，表の主対角線に関して，対称的な確率がそれぞれ等しいと解釈できます。ここで主対角線とは，やや専門的な用語になりますが，表の左上から右下にかけての対角線を意味します。表１に対しては，右握力が⑴強い，左握力が⑴強いというセルと，右握力が⑵普通，左握力が⑵普通というセル，さらに右握力が⑶弱い，左握力が⑶弱いというセル，計３つのセルを通る直線となります。このような確率構造が成り立つという仮説を対称モデルと呼びます。

表１　男子高校生の左右握力

右握力	左握力 ⑴強い	⑵普通	⑶弱い	計
⑴強い	74	89	3	166
⑵普通	10	444	93	547
⑶弱い	0	23	69	92
計	84	556	165	805

ここで一点重要なことがあります。今私たちが知りたいのは，805名の測定データに対して左右握力が対称的であるかどうかではなく，男子高校生全体（母集団）に対して，左右握力が対称的かどうかです。しかし男子高校生全体の確率構造は未知であり，その具体的な値を知ることは不可能です。そこで，得られた805名の測定データを男子高校生全体の標本と捉え，男子高校生全体の左右握力が対称的かどうかを推測することを考えます。これは統計学における推測統計と呼ばれる解析手法です。詳細は省略しますが，表１の男子高校生の左右握力データにおいては，右握力と左握力は対称的ではないと結論づけられます。

さて，男子高校生の左右握力に関しては対称性が成り立ちませんが，ではどのような構造ならば成り立つでしょうか。次節では，表１のデータに対して成り立ちそうな仮説（モデル）を考えてみたいと思います。

モデル提案

表１のデータをよく観察してみると，表の左下よりも右上に多くの観測度数が集中していることが分かります。これは多くの人間が右利きであり，右の握力の方が強いであろうという直感とも一致します。このとき，表の右上から左下にかけての対角線，ここでは逆対角線と呼ぶことにしますが，この逆対角線

$H_{AS}: p_{11}=p_{33},\ \ p_{12}=p_{23},\ \ p_{21}=p_{32}$

右握力	左握力 (1)強い	(2)普通	(3)弱い	計
(1)強い	p_{11}	p_{12}	p_{13}	$p_{1\bullet}$
(2)普通	p_{21}	p_{22}	p_{23}	$p_{2\bullet}$
(3)弱い	p_{31}	p_{32}	p_{33}	$p_{3\bullet}$
計	$p_{\bullet 1}$	$p_{\bullet 2}$	$p_{\bullet 3}$	1

図１　逆対角対称モデル

$H_{AGS}: p_{11}+p_{12}+p_{21}=p_{23}+p_{32}+p_{33}$

右握力	左握力 (1)強い	(2)普通	(3)弱い	計
(1)強い	p_{11}	p_{12}	p_{13}	$p_{1\bullet}$
(2)普通	p_{21}	p_{22}	p_{23}	$p_{2\bullet}$
(3)弱い	p_{31}	p_{32}	p_{33}	$p_{3\bullet}$
計	$p_{\bullet 1}$	$p_{\bullet 2}$	$p_{\bullet 3}$	1

図２　逆対角グローバル対称モデル

に関しては対称性が成り立つかもしれません。

ここで未知の確率 $\{p_{ij}\}$ についての３つの仮説（モデル）を考えます。

１つ目は，逆対角線に関しての対称性を表す仮説です。この仮説の数式的な定義を**図１**に詳しく記載しました。図の見方としては，同じ色の丸で囲まれた部分の確率が等しいことを示しています。例えば，青い丸で囲んだ２つの確率については，任意に選んできた人の左右握力が共に(1)強いである確率（図１におけるp_{11}）が，左右握力が共に(3)弱いである確率（図１におけるp_{33}）と等しいことを意味します。赤丸（p_{12}とp_{23}）と緑丸（p_{21}とp_{32}）についても，同様の構造が成り立ちます。この仮説をここでは逆対角対称モデルと呼ぶことにします。

２つ目は，逆対角線に関して左上の確率の総和と，右下の確率の総和が等しいという仮説です。この仮説の数学的な定義を**図２**に記載しました。紫色の図形で囲われたそれぞれの部分の確率の和が等しい，すなわち，p_{11}とp_{12}とp_{21}の総和とp_{23}とp_{32}とp_{33}の総和が等しいという仮説です。この仮説をここでは逆対角グローバル対称モデルと呼ぶことにします。

３つ目の仮説に入る前に，一つ新たな定義を説明します。今，左右どちらかの握力だけに注目して，例えば，任意に選んできた人の右の握力がiである確率を$p_{i\bullet}$で定義することにします。つまり$p_{1\bullet}$はp_{11}とp_{12}とp_{13}の総和で表されることになります。左の握力についても同様に，左の握力がiである確率を$p_{\bullet i}$で定義することにします。３つ目の仮説は，この周辺確率について述べたものです。この仮説の数式的な定義を**図３**に記載しました。図１と同様に，同じ色の丸で囲まれた部分の確率が等しいことを示しています。具体的には，右握力が(1)強い確率と，左握力が(3)弱い確率が等しく，他にも右握力が(2)普通の確率と左握力が(2)普通の確率，最後に右握力が(3)弱い確率と左握力が(1)強い確率が，それぞれ等しいという構造です。この仮説をここでは逆対角周辺同等モデルと呼ぶことにします。

今３つの仮説を考えましたが，この仮説には包含関係があり，例えば，確率構造に対して逆対角対称モデルが成り立てば，逆対角グローバル対称モデルと逆対角周辺同等モデルがどちらも成り立つという関係があります。次節ではその点も含めて，実際に男子高校生の握力を解析してみましょう。

データ解析

前節で導入した３つのモデルに対して，「検定」という解析手法を用いて実際にモデルの当てはまりを見てみます。ここからいくつか専門的な用語が出てきますが，それらの用語をすべて理解せずとも，話の流れだけを追ってもらえれば大丈夫です。

まず，逆対角対称モデルに対する尤度比（ゆうどひ）カイ二乗統計量は5.53と計算されます。この値を自由度３のカイ二乗分布の上側５％点である7.81と比較して，大きければ有意水準５％で帰無仮説（この例では，男子高校生の左右握

$$H_{AMH} : p_{1\bullet} = p_{\bullet 3},\quad p_{2\bullet} = p_{\bullet 2},\quad p_{3\bullet} + p_{\bullet 1}$$

	左握力			
右握力	(1)強い	(2)普通	(3)弱い	計
(1)強い	p_{11}	p_{12}	p_{13}	$p_{1\bullet}$
(2)普通	p_{21}	p_{22}	p_{23}	$p_{2\bullet}$
(3)弱い	p_{31}	p_{32}	p_{33}	$p_{3\bullet}$
計	$p_{\bullet 1}$	$p_{\bullet 2}$	$p_{\bullet 3}$	1

図３　逆対角周辺同等モデル

力に逆対角対称モデルが成り立つ）を棄却します。帰無仮説を棄却するとき，帰無仮説の下で，すなわち，逆対角対称モデルが成り立つと仮定したとき，表１のデータが得られる確率は５％より低いと推測します。つまり，帰無仮説が正しいと仮定したときに，表１のデータが出現する確率は５％より下なのだから，帰無仮説は間違っているとみなすのが妥当だろうと考えるのです。今回は５％点より小さいので，帰無仮説は間違っているとはいえないと考え，帰無仮説を採択します。少し複雑な言い回しになってしまいましたが，母集団において，逆対角対称モデルの確率構造が成り立っていると考えます。

他の２つのモデルも同様に当てはまりを調べてみましょう。逆対角グローバル対称モデルに対する尤度比カイ二乗統計量は0.40と計算されます。この値も自由度１のカイ二乗分布の上側５％点である3.84と比較して小さいので，逆対角グローバル対称モデルが成り立っていると考えます。逆対角周辺同等モデルに対する尤度比カイ二乗統計量は0.64であり，自由度２のカイ二乗分布の上側５％点である5.99と比較して小さいので，逆対角周辺同等モデルが成り立っていると考えます。

３つのモデルはどれも当てはまりが良いという結果が得られました。ここでもう少し踏み込んで，この中でどのモデルがもっとも当てはまりが良いか，すなわち解釈を得るのが適切であるかを調べてみます。逆対角対称モデルと逆対角グローバル対称モデルには包含関係があるので，まずこの２つの仮説を比較してみましょう。こちらも詳細は省略しますが，逆対角対称モデルの方がより当てはまりが良いことが分かります。さらに逆対角対称モデルと逆対角周辺同等モデルの比較をしてみると，逆対角周辺同等モデルの方がより当てはまりが良いことが分かります。

以上により，男子高校生の左右握力については，逆対角周辺同等モデルがもっとも当てはまりが良いという結果が得られました。つまり，男子高校生に左右握力に関しては，右握力が(1)強い確率と左握力が(3)弱い確率が同等であり，また右握力が(3)弱い確率と左握力が(1)強い確率も同等であることが分かりました。

終わりに

本稿は統計学や数学にあまり詳しくない人に対しても，統計学の数理的な面を知ってほしいという趣旨で執筆しました。難解な用語は避け，なるべく丁寧に書いたつもりですが，それでもまだ難しく思える箇所があったかもしれません。これは筆者の力不足ですが，これを読んで統計学，特に数理の面に興味をもったならば，ぜひ統計学という学問を勉強してみて下さい。

さまざまな分野で注目されている統計学ですが，その話題性に振り回されるだけでなく，統計学のセンスをしっかりと身につけることができれば，そこから有用な解釈と結論を得ることができるはずです。