

データサイエンスと帰納的思考

大分工業高等専門学校一般科 樋口勇夫¹

1 はじめに

数理・データサイエンス・AI教育が全国的に推進されている。データサイエンスは数学、情報科学、統計学などの複合学問といわれ、これらの分野を横断的に指導することが望ましい。AI時代におけるデータサイエンス教育の方法として、筆者は統計学、情報学、数学を3本の柱として教育することを工学教育(2022)[1]で提言した。この中で統計学は帰納的推論の学問であり、数学によって客観的な裏付けを与えられるものであることを述べた。

統計学と数学については共通する部分も多いが、統計学が数学の中にすべて含まれているわけではない。しかしながら、一般に統計学は数学の一分野と捉えられ、数学の中での教育が求められることが多い。数学の中で統計学を教えようとするとき、確率論を駆使した裏付けの部分をしっかりと説明するが、その結果として得られるものについての説明が不足することが少なくない。その原因として、数学は演繹的論証の学問であり、教育者が帰納的推論に慣れていないことが考えられる。

帰納的推論の難しいところは、過去の事例を学んでも他の分野への応用ができるかどうかははっきりしないところにある。したがってその分野での専門家の知見を積み重ねることが有効といえる。そして、新たな知見を得るためには臨機応変に対処して仮説を組み立てる必要がある。そのような帰納的思考を学ぶためにはデータから知見を得るような経験が役に立つ。しかし、教育の中、例えば実験によってそのような経験を与えようすると、一部の意欲的な学生はインターネットなどで結果を調べてしまい、自分で試行錯誤して導き出すことをしない。学生にとってはその方がタイムパフォーマンスがいいということであろう。そのような中で帰納的思考を体験させる方法としてカードゲームのような模擬的な問題を使うことを考えた。

本研究では統計学と数学との関係を明らかにし、帰納的思考の養成の必要性を示す。そのうえで帰納的思考に慣れさせるための実例として、カードゲームを用いた教育について紹介する。

2 研究成果

2.1 データ分析と帰納的推論

データ分析ではデータから何らかの仮説を組み立てる。統計学でよく出てくる仮説検定では、帰無仮説から得られるはずの数値の確率を計算し、帰無仮説とデータとの整合性を調べる。帰無仮説から確率計算によって p 値を得る過程は間違いなく演繹的論証であり数学の一分野である。その結果、 p 値が小さければ帰無仮説が棄却される。このように統計

¹E-mail: i-higuchi@oita-ct.ac.jp

学の数理的な部分は数学として教育されるが、では何故仮説検定を使うのか、何故その結果が行動の指針になるのか、といった内容はデータ自体が確率的に生じるものであり、その根源となる確率分布をどう探るのか、といった内容に起因し、ここで帰納的推論を行うということを明らかにしないと、教育が難しい部分である。

ところが多いの学生は帰納的思考でものごとを学ぶ機会が少ない。実際、教育の題材となる多くの問題について「調べればわかる」という現状の中で、データから試行錯誤で真理を探るという行動を選ぶ価値を学生が感じることは難しいだろう。

2.2 帰納的推論を扱うカードゲーム

この問題の解決策の一つとして、カードゲームのようなトイモデルを使うことが挙げられる。隠されたルールを用意し、ルール通りにカードを並べていく中で、そのルールを推測するゲームをする。このようなゲームであれば決まった戦略というものとは存在せず、ネットで調べれば答えがわかるものではないはずである。

そのような目的で作られたカードゲームとして Eleusis Express というものがある。これは 2006 年にミシガン州のグレートバレー州立大学の数学教授である John Golden が開発したカードゲームである。これはルールに従って並べられたカードの列からどのようなルールなのかを見出すゲームである。Golden は Robert Abbott が 1956 年に生み出した Eleusis を改良してこのゲームを作ったと述べている。そして Abbott はこのゲームは帰納的推論を必要とすると述べている。

Eleusis Express は次のようなゲームである ([3] よりルール抜粋)。

概要 プレイヤーのうちの 1 人が、プレイ可能なカードについてのルールをつくる。例えば「最後にプレイされたカードと異なる色のカードを出さなければならない」。他のプレイヤーはルールに合致すると思うカードを出していく。正しいカードを出したプレイヤーは正しいと思うルールを宣言することができる。

人数 3~8 人、おそらく 4~5 人が最適と思われる。

カード ジョーカーを抜いたトランプ 52 枚を 2 組用いる。

目的 プレイヤーの一人がディーラーとなり、秘密のルールを選択する。ディーラー以外のプレイヤーがカードをプレイしていき、手札をすべて出し切るか、秘密のルールを正しく言い当てることを目指す。

準備 ディーラーは自分が選んだ秘密のルールを紙に書き、その後各プレイヤーにカードを 12 枚ずつ配る。ディーラーは残りのカードのうち一番上のカードを表向きにし、テーブル上に置く。

プレイ プレイヤーは手番において、カードを 1 枚テーブル上に出す。ディーラーはそのカードが正しいか誤っているか宣言する。正しかった場合、そのカードをメインの列における最後のカードの右隣に配置する。誤っていた場合、そのカードを直前のカードの下側に配置する。誤ったカードを出したプレイヤーはカードを 1 枚引き手札に加える。正しいカードを出したときはカードを引かないため、手札の枚数が 1 枚減ることになる。

パス プレイヤーは正しくプレイできるカードを持っていないと宣言することもできる。このプレイヤーは手札を（全員に）開示し、ディーラーはこのプレイヤーが正しいか誤っているか回答する。

- この宣言が誤っていた場合は、ディーラーは開示された手札から正しくプレイできた筈のカード 1 枚を選んで場に並べ、プレイヤーはカードを 1 枚引き手札に加える。
- この宣言が正しかった場合は、プレイヤーは開示された手札を全て破棄しカードを新たに引き直す。このとき破棄した枚数より 1 枚少ない枚数のカードを引く。

カードのプレイと同様に、正しければ 1 枚手札が減り、間違っていた場合は手札枚数は変わらない、ということになる。

宣言 プレイヤーが正しくカードを出す、あるいは正しくパス宣言した場合、プレイヤーはルールを宣言する権利が与えられる。宣言したルールが正しかった場合ゲームを終了する。間違っている場合はそのままゲームを続行する。

その他 ここでは割愛するが、得点計算のルールがあり、それを採用するのであればディーラーを各自が 1 回ずつ行って合計点で競う。1 回のゲームで終わるのであれば、カードを出し切るか、正しくルールを宣言したプレイヤーの勝ちである。

原作の Eleusis の作者である Abbott は [3] で以下のように述べている。

- このゲームは帰納的推論を扱う数少ないゲームの一つ
- ほとんどのゲームではプレイヤーにルールと状況が提示され、それに基づいた演繹的推論が求められる
- Eleusis ではプレイヤーは状況を観察し、支配するルールを発見する
- そのためには仮説を立て、プレイで検証していくことになる
- この過程は科学者が自然法則を解明するときの過程と似ている

2.3 カードゲームの拡張

2022 年に筆者はこのゲームのための新しいカードセット、E.E.E. (Eleusis Express Expansion) を作成した。



図 1: E.E.E.

E.E.E. は以下のようなカードセットである。

色 赤黒の 2 色

形 ○▽□の 3 種類, それぞれジャンケンのグー・チョキ・パーに対応する

数字 0~9, 10(T), 11(A) の 12 種類

枚数 同じカードが 2 枚ずつ, 計 144 枚

大きな特徴としては色, 形, 数字の 3 要素が独立に交じる構造になっていることが挙げられる。また, 形の 3 種類にジャンケンの 3 すくみの要素を加えてみた。これにより, 「○, □, ▽の順を繰り返す」というルールを「前の手に勝つように出す」と言い換えることができる。3 つの形によるサイクルを解釈できるようになるため, 「奇数の後は前の手に勝つ, 偶数の後は前の手に負けるようにカードを出す」, 「黒のカードを出すときは前の手に勝つように, 赤のカードを出すときは前の手に負けるようにする」などといったルールを作るときに直感的に記述することができる。背面にこの 3 すくみ関係を書くとゲームの中で使いやすいと考えられるので, 次を作るときは検討したい。

3 実践の報告

大分高専では数理・データサイエンス・AI 教育プログラムのために, 2022 年度より本科 1 年生の一般科目として情報 I・II を開設した [2]。情報 I は情報リテラシーの基礎, 情報 II は AI・データサイエンス概論となっている。情報 II の 1 コマを利用し, E.E.E. を体験してもらった。これ以外に専攻科の科目や, 囲碁将棋部の活動の中でも体験してもらっている。

この授業では 90 分のうちの最初 20 分程度で説明を行い, 6 人または 7 人のグループに分かれて体験してもらった。1 ゲームは 20 分ほどで, それぞれ 2 ゲームまたは 3 ゲームプ

レイしていた。盤面に出たカードは以下の進行例で示すような形で記録をとらせた。ゲームごとに、どこでどのように当たりをつけたか、それをもとにどのような選択をしたのか、などの思考過程も振り返えるよう促した。

授業ではルールを考えるのに時間をかけてしまうことも予想されたので、ルールを書いたカードを作り、その中から選んでもらうこともできるようにした。3分の1くらいはこのカードを利用していた。

多くの学生から「あまり他のゲームでは経験しない考え方だけど、新鮮で面白かった」という感想が出た。一部の学生は Eleusis Express について検索していたが、攻略法は特に見つからなかったようだ。

ゲームが長引くと、目的が自分のカードを出し切ることより、ルールを解明することに移っていく傾向がみられた。その結果ルールを解明するために協力する様子も多くみられた。

次にいくつかのゲーム進行例を示す。以下の進行例では、黒のカードを●▼■、赤のカードを○▽□を用いて表す。

3.1 ゲーム進行例 1

EEtable_sample1 by KETCindy

▽	○	▼	●	■	●	□	▽	■	□
T	3	5	1	7	0	3	1	8	6
■ A		▽ 6	■ 1	▽ 8	▽ 0			■ A	● 5
		□ 7	■ 0						▼ 4
		■ 9							▽ 6
		● A							

図 2: 進行例 1

ほとんどの場合、失敗したカードをみることでルールが見えてくる。図 2 の例では▼ 5 の後に失敗が 4 つ続いている。これらは条件を満たさないカードであるから、失敗した 4 つの共通点を考えることがルールの解明につながる。ここでは単純な色の違いや形の違いではなさそうに見える。この段階で数字 (6,7,9,11) に着目すると、小さめの数だったら？とを考えていけば、数字の大小が関係しているようだとわかる。

この例でのルールは「数字の増減が交互に起こる」というものであった。数字の大小の場合、それが絶対的か相対的なものかについても検討する必要がある。

3.2 ゲーム進行例 2

図3の例でも、○0の後の4つの失敗カードがヒントになる。状況的には進行例1と同様の数字の大小に関するルールが候補に挙がるだろう。しかしこの状況は偶然で、実際は「黒の後は素数、赤の後は素数でない数」というルールであった。5,2,2,3の共通点が素数だと気づけるかどうかは鍵になるといえる。

基本的な戦略としては、同じ状況で失敗したカードの共通点を考え、それにあてはまらないカードを出していく。しかしながら他のルールで通用した考え方が必ずしも次に通用するとは限らないので注意は必要である。

EEETable_sample3 by KETCindy

▽	○	●	○	■	■	▼	□	□	▽
7	0	8	5	T	7	7	3	6	8
●7	■5	○8				▽T	▽3		□3
	●2	□4				○6			
	▽2								
	○3								

図 3: 進行例 2

3.3 ゲーム進行例 3

3つ目の例として少し複雑なルールの例を挙げる。(図4) このゲームのルールとしてはあまりお勧めできないくらいの難しさと思っていたが、実際は途中で判定ミスがあり、その修正がヒントとなって正解にたどり着いたようである。

- 赤と黒は交互に
- 数字が前のカードと同じかそれより大きいカード、ただし最大のカードのあとは何でもよい
- 同じ形が3枚続いた場合、b.における大小が入れ替わる
- のルールはいわゆる「大富豪」の「革命」ということで、このルールを採用した学生は「革命ルール」と呼んでいた。

この例ではルールミスがあった場所は修正してある。

「だんだん数字が大きくなる」というのはわりとすぐわかり、途中で順が逆転したことも意外と早く気づいていた。判定ミスがなくても正解にたどり着いていた可能性は十分あったと考えられる。

EEtable_sample5_1 by KETCindy

□ 5	■ T	○ A	▼ 4	○ 5	● 6	○ 7	■ 4	□ 3	● 1	▽ 0	■ 7
▽3	□2	□T	■A			□9		▼5	○T		○8
●0	○2		■1			▽8			▽A		
○0	●1										
▼3	○0										

EEtable_sample5_2 by KETCindy

□ 2	▼ 1	□ 1	▼ 0	○ A	● A	▽ A	▼ 7	□ 0	■ 4	□ 4	
▼4		▼2					▽9				
●5		●A									

図 4: 進行例 3

4 まとめと今後の課題

本研究ではデータサイエンス教育において、帰納的思考を養成するための工夫について述べた。データサイエンス教育には帰納的推論からのアプローチも必要と思われるが、そのためには帰納的推論を体験する機会を与えることが有効と考えられる。実際に授業で導入してみたところ、多くの学生にとって帰納的思考は新鮮な体験となったようで、概ね好評であった。

今後の課題はいくつかあるが、まずは、オンラインの環境を整え、気軽に体験できるようにしたい。カードを使うゲームであるが、GMがプレイヤーのカード内容を把握しても構わない。そのため、メールで将棋やチェスの対戦ができるように、データのやり取りをメールの平文で行っても問題が無い。計算機上でプレイできる環境ができれば、ほぼそのままオンラインでの実施が可能な環境になると考えられる。計算機による判定ができるようになればここに確率的な揺らぎを組み込むことも考えられる。計算機を使わずにこのような揺らぎを組み込むことも検討中であるが、ディーラーの処理をどう簡略化するかが大きな課題となっている。また、試行錯誤の過程を学べるチュートリアルを何らかの形で作成することによってより気軽に体験できるようにしたい。

参考文献

- [1] 樋口勇夫, 統計教育を重視したデータサイエンス教育, 工学教育 **70**(4) (2022), 4-8.
- [2] 樋口勇夫, 「数理・データサイエンス・AI 教育プログラム認定制度」と大分高専の取り組み, 高専教育フォーラム (4) (2022), 4-5.
- [3] Robert Abbott (竹田原裕介 訳), ロバート・アボット 新テーブルゲーム作品集, 合同会社ニューゲームズオーダー (2022).