

母平均の差の検定に関するシミュレーション学習を可能とする HTML をベースとした教材の作成と改良

山口大学・教育学部 北本 卓也

Takuya Kitamoto, Faculty of Education, Yamaguchi University

東邦大学・薬学部 金子 真隆

Masataka Kaneko, Faculty of Pharmaceutical Sciences, Toho University

東邦大学・理学部 野田 健夫

Takeo Noda, Faculty of Science, Toho University

1 はじめに

本研究では、数理統計分野でもとりわけ応用範囲が広い「母平均の差の検定」について、シミュレーションを伴う学習を行わせるための教材について取り上げる。実験授業の対象者が、数理統計の授業を一通り受講した薬学部3年生6人だったため、該当者にとって身近な血圧降下剤の効果の検証を題材とした。周知のごとく、この検定については、2群のデータに対応がある場合とない場合の2通りがあり、背景となるメカニズムがかなり異なる。平常授業の際も、理論的な枠組みを説明した上で、計算例を用いた演習を行っているが、試験の際などに方法を逆転させてしまう事例が稀ではない。このような事態を極力防ぐために、メカニズムの差異を認識できるようなシミュレーション用教材の作成に着手したわけであるが、こうした教材を直ちに大人数の平常授業に投入した場合、筆者の経験ではあまり効果を発揮しない事例が少なくなく、まずは少人数を対象に試行的に使用し、対象者からコンテンツ自体やこれを用いた授業設計の改善に向けたフィードバックを得ることが必要だと考えられる。本稿で紹介するのはこのような試行事例である。

2 HTML をベースとした教材について

2.1 教材を作成したシステムの概要

本教材は HTML 上のコンテンツであり、ブラウザ上で動作する。このため、パソコンのみならず、タブレットやスマートフォンから閲覧することが可能であり、教材としての汎用性が高い。また、本教材は [1] のシステムを用いて構成されている。この [1] のシステムでは、基本的にはブラウザ上からその構成を行えるようになっており、HTML のソースコードを直接編集することなく、教材の作成が可能になっている（ただし、本教材の作成においては下記に述べるように一部でソースコードを直接編集することが必要となった）。

2.2 教材の作成方法

編集モードでは、図1のようなメニューが現れる。

機能	名前	場所	オプション		
login 設定		eps_basis の後▼	local ▼		
radio button 追加	r1	eps_basis の後▼	選択肢: A,B,C		
textarea 追加	t1	eps_basis の後▼	行数: 10	列数: 80	
Cinderella 追加	c1	eps_basis の後▼			
Quill 追加	q1	eps_basis の後▼			
Javascript ボタン 追加	j1	eps_basis の後▼	label: 実行	行数: 10	列数: 80
Chart 追加	g1	eps_basis の後▼	label1: グラフ描画	データ形式: textarea ▼	
表計算 追加	s1	eps_basis の後▼	label1: 表計算	label2: ダウンロード	label3: csv出力
ファイル入出力 追加	gf	eps_basis の後▼			
画面共有 追加	d1	eps_basis の後▼	送信 ▼		
TV電話 追加	t1	eps_basis の後▼	送受信 ▼		
P2Pデータ 追加	d1	eps_basis の後▼	テキストエリア ▼		
削除	test1				

図1 編集モードでのメニュー

このメニューで例えば、「textarea 追加」をクリックすると、画面上に textarea を追加することが可能である。「Cinderella 追加」をクリックすると、画面上に図2のようなボタンが追加される。

☒ (c3): ☒ ファイルを選択 選択されていません

図2 「Cinderella 追加」をクリック

このボタンをクリックし、Cinderella ([2],[3]) で作成した HTML ファイルを選択するとその Cinderella の図が画面上に現れる。また、「JavaScript ボタン追加」をクリックすると、図3のように textarea と「実行」のラベルがついたボタンが画面に追加される。

☒ (j3):

実行

図3 「JavaScript ボタン追加」をクリック

この「実行」のボタンをクリックすると、textarea 書かれた JavaScript のコードが実行される。

このメニューを用いて教材を作成していったが、本教材には今までにない次の特徴があり、HTML のソースコードの一部を手で入力することが必要になった。

- (i) 本教材ではわかりやすさのため、Cinderella の図の横にその図に対応する計算過程を表示するようにしている (図4を参照)。

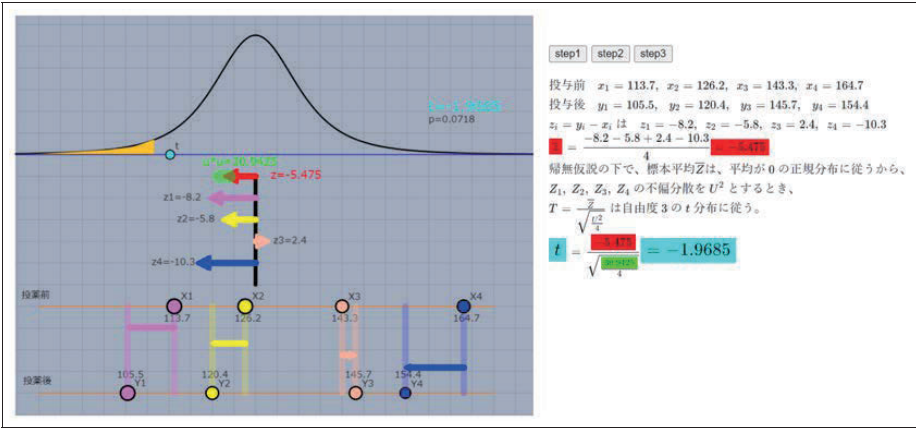


図4 システムの全体図

このため、HTML のレイアウトを操作する命令を直接 HTML ファイルに書き込んで、Cindrella の図と KaTeX ([4]) の位置の調整を行った。

- (ii) 計算の過程を理解しやすくするため、KaTeX を用いて計算過程の数式を画面に表示している (図 4 を参照)．今まではブラウザベースのリッチテキストエディタである Quill ([5]) を用いて間接的に KaTeX を使ってきたが、本教材で使われるような複雑な数式は Quill ではうまく表現することができなかった．そのため、KaTeX の命令を直接、HTML のソースコードに書き込んで教材を作成した。

2.3 教材の特徴

本コンテンツは、Cindrella による図を用いることで、視覚的な操作を通じてデータを扱うことを可能にしている．図の点の配置がデータの値と結びついているため、データの値と計算結果との関係性を視覚的に理解しながらデータの値を操作することが可能である．また図の点（データの値）の変更が行われた場合でも、それに応じた計算過程が表示されるため、計算過程と図のデータとの関連が直感的に理解しやすい．

さらに、この計算過程は一度に全てを表示せず、全体を 3 つのステップに分けて少しずつ表示されるようにしている．図 5 は第 4 節に説明する実験授業を経て改定を繰り返し、最終的にできあがったコンテンツの図解であるが、上半分が対応ありの場合、下半分が対応なしの場合のシミュレーションを行うためのものである．この中にある Step1, Step2, Step3 のボタンをクリックしていくと、それぞれ標本平均、不偏分散、 t 検定量の各実現値が計算・表示され、仮説検定の計算過程が少しずつ追加表示されるようになっている（それに応じて Cindrella の図の方も少しずつ表示部分を変えている）．

動的な図解と数式表現の連携をとることは、この検定の全体像を理解する上で重要である一方で、すべての情報を一度に表示すると、情報量が過多になって学習者に大きな認知負荷 [6] がかかることが想定される．このような段階的な表示は、ユーザーが計算

の手順を少しずつ理解できるようにし、その認知負荷が低減されるように配慮したものである。

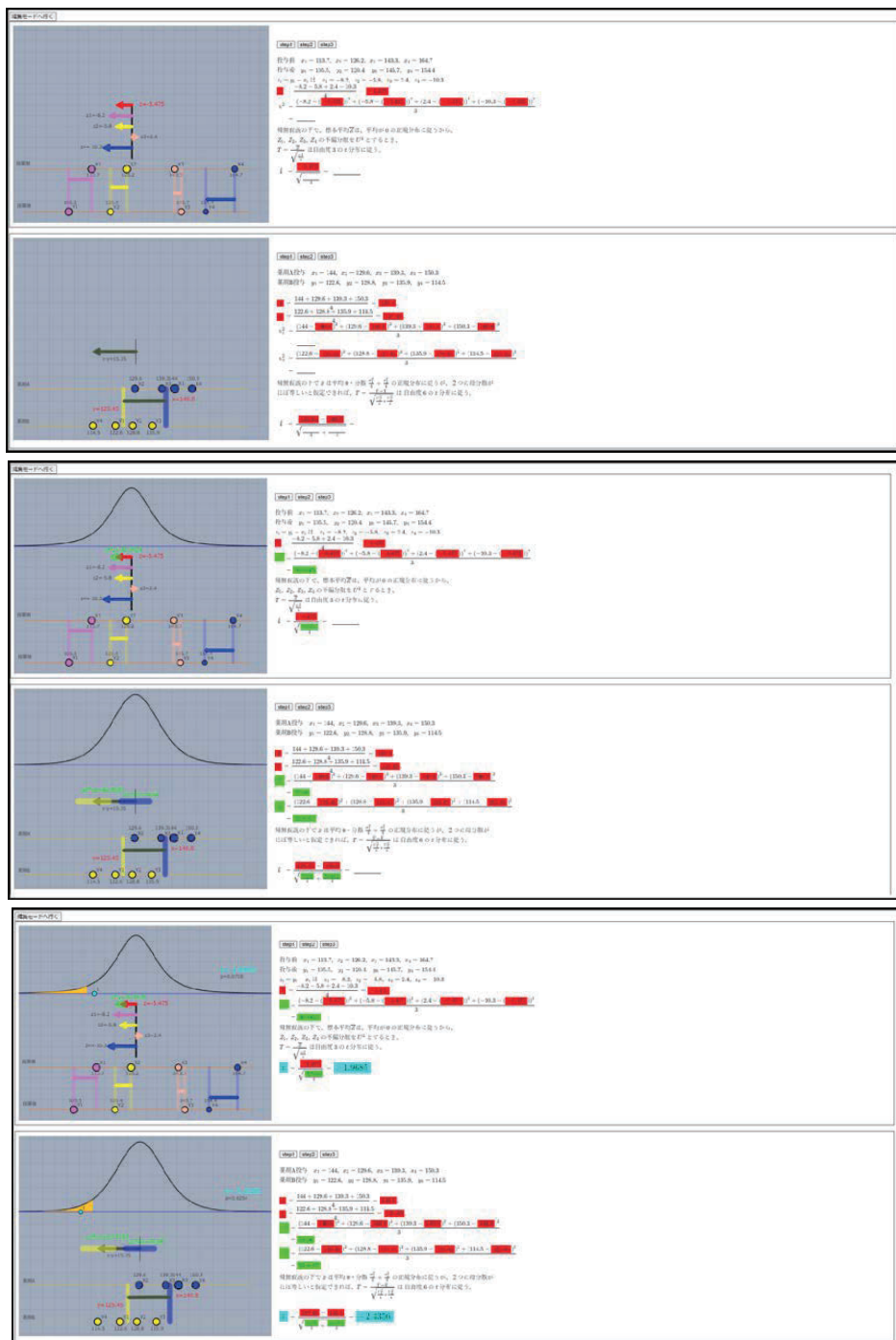


図5 完成したコンテンツ

また、図6に示す追加コンテンツでは「対応がある場合」と「対応のない場合」の2つの Cinderella の図を1つの教材上で表示しているが、この2つの図のデータの相互コピーが可能である。これにより、同じデータが与えられたときの「対応がある場合」と「対応がない場合」のそれぞれの検定結果を確認できる。それが、「対応がある場合」と「対応がない場合」の2つケースでの仮説検定のメカニズムの違いの理解へと繋がると期待できる。

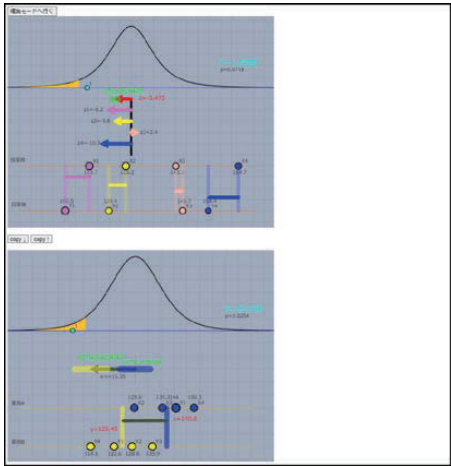


図6 2種の検定の比較に用いたコンテンツ

Cinderella 同士は互いにデータをやり取りする機能を持っていないので、このデータの複製は JavaScript のプログラムを用いて行われている。すなわち、Cinderella 上の点の配置のデータは一旦、JavaScript のシステムにコピーされ、その後その値がもう一方の Cinderella の方へコピーされる。この機能は上で説明した「JavaScript ボタン追加」で作成されたテキストボックス内に JavaScript のプログラムを書くことで実現されている。

また、Cinderella はそれ自身で HTML ファイルを作成する機能を持っているが、その機能では1つの HTML 上に1つの Cinderella の図しか置くことができず、2つ以上の Cinderella の図を持つ本教材を作成することは不可能である。それに対して、本システムでは上で説明したように「Cinderella 追加」のボタンをクリックすることで1つの HTML ファイルに任意数の Cinderella の図を配置することができる。

3 教材を用いた基本的な授業設計

母平均の差の検定において、対応がある場合とない場合とで計算過程上最も大きく異なる点は、対応がない場合は2群のそれぞれについてまず標本平均の実現値を計算するのに対し、対応がある場合は互いに紐づいたデータどうして差を取った上で標本平均の実現値を計算することである。授業の設計上、対応がある場合には紐づいたデータの差が正規分布に従うこと、対応がない場合には2群のデータがそれぞれ正規分布に従うこ

とがそれぞれ背景となるという、理論的な差異を強調することが出発点となる。作成したコンテンツにおいても、対応ありの場合にはいったん紐づいたデータどうしの差をとってその標本平均（赤い矢印で表示される）を計算していること、対応なしの場合には2群の標本平均（黄色と青色の縦線で表示される）を計算した上でその差を計算していることが、図上で強調されている。

初学者にとっては決して易しい課題ではないため、以上の説明を教科書を用いて行った後、図7にある簡単な計算課題に取り組んでもらうことにしている。課題では、データがまったく同じであるにもかかわらず、対応ありの状況となしの状況で検定結果が異なるように工夫してある。

1. 血圧を下げる治療に用いる新薬の臨床試験において、5名の被験者の投与前と投与6週間後の収縮期血圧（単位 mmHg）を測定し、以下のデータを得た。投与前に比べて投与後の血圧が下がっているといえるか、片側検定により有意水準5%で検定したい。

投与前	184	156	136	178	181
投与後	169	152	125	177	177

投与前後の標本をそれぞれ X_i , Y_i とおき、検定統計量 $T = \frac{\bar{X} - \bar{Y}}{\sqrt{U^2/n}}$ の実現値 t を求めると、 $\bar{X} - \bar{Y}$ の実現値が $\boxed{ア} \cdot \boxed{イ}$, U^2 の実現値が $\boxed{ウ} \cdot \boxed{エ} \cdot \boxed{オ}$ であるので

$$t = \frac{\boxed{ア} \cdot \boxed{イ}}{\sqrt{\boxed{カ} \cdot \boxed{キ} \cdot \boxed{ク}}} = \boxed{ケ} \cdot \boxed{コ} \cdot \boxed{サ}$$

T は自由度 $\boxed{シ}$ の t 分布に従うから、棄却域が $\boxed{ス} \cdot \boxed{セ} \cdot \boxed{ソ} \cdot \boxed{タ} \leq t$ となることより、差が $\boxed{チ}$ と結論される。 $\boxed{チ}$ については、「ない」であれば1を、「ある」であれば2をマークせよ。

2. あってはならないことだが、前問のデータを取った際、資料の整理が不十分だったため、数値がどの被験者のものか確定できない状況であることが判明したという。このとき、投与前に比べて投与後の血圧が下がったといえるか、片側検定により有意水準5%で判断すると、検定統計量 $T = \frac{\bar{X} - \bar{Y}}{\sqrt{U_1^2/n_1 + U_2^2/n_2}}$ の自由度が $\boxed{ツ}$ で、棄却域が $\boxed{テ} \cdot \boxed{ト} \cdot \boxed{ナ} \leq t$ であるのに対し、 T の実現値が $\boxed{ニ} \cdot \boxed{ヌ}$ であることより、 $\boxed{ネ}$ と結論される。 $\boxed{ネ}$ については、「ない」であれば1を、「ある」であれば2をマークせよ。

図7 平常授業時の計算課題

対応なしの場合に厄介なのは、2群の母分散が大きく異なる場合、自由度や検定統計量の形が複雑になってしまうこと（Welchの t 検定）である。このため、2群の母分散が大きく異ならない状況を想定し、より簡単な計算を適用できる場合のみ取り扱うこととしている（Studentの t 検定）。コンテンツ上では、2群のデータの不偏分散が黄色と青色の帯で図示されるようになっており、観察しているデータが上記の前提を満たしているか、逐次確認できるようになっている。

当然のことながら、コンテンツを用いてシミュレーションを行わせるのは計算課題が完了した後になる。まず、コンテンツを提示しながら、対応ありの場合は薬剤を投与した前後での最高血圧、対応なしの場合は2種類の薬剤を別個の被験者に投与した際の最高血圧をそれぞれ表すという設定を説明した後、実際にデータ点を動かしてみ、標本平均や不偏分散が計算・可視化される経過を見せることになる。同時に、検定統計量の確率分布である t 分布について、対応ありの場合は自由度が3であるのに対し、対応なしの場合は自由度が6となる点についても強調することになる。ここで最も大きな問題は、学生自身にコンテンツを操作させる前の段階で、どの程度のシミュレーションを提示するかという点である。さらに、学生に操作させる初期段階での制約をどこまでかけるか

という点も問題となった．実際，事前に与える情報の量を抑え，完全に学生の自由に操作させると，探索の柔軟性は高まるが，逆に探索の方向性を定めるのに苦勞するという心配も出てくる．次節に示す実験授業では，この点の設計が最大のポイントとなった．

シミュレーションのしめくりの課題として，「相互コピー」のコンテンツを用い，同じデータを用いているにもかかわらず，対応ありの場合と対応なしの場合とで検定結果が異なるケースを一つ予想するように指示することにした．さらに，この課題は，学生自身による自由な探索を行わせる手前の段階で提示しておき，探索のひとつの目標を与えることとした．コンテンツ自体は相互のコピーが可能な状態だが，Student の t 検定にかかる制約上，基本的には対応なしの場合のコンテンツでデータ点を並べ，2 群の不偏分散が大きく異なることを確認した上で，該当のデータを対応ありのコンテンツにコピーする流れになる．対応なしの場合の 2 群の標本平均の実現値の差が大きく，検定結果が有意となるものの，データの順番を入れ替えると対応ありの場合での検定結果が逆転するケース（図9 左）が最も把握しやすいが，逆に，対応ありの場合で紐づいたデータの差が小さいにもかかわらず差のばらつきが小さいために検定結果が有意になるものの，差の小ささが影響して対応なしの場合の検定結果が有意にならないケース（図9 右）も存在する．こうした感覚は，動的コンテンツを操作してみることなしに獲得するのは極めて難しいと考えられる．

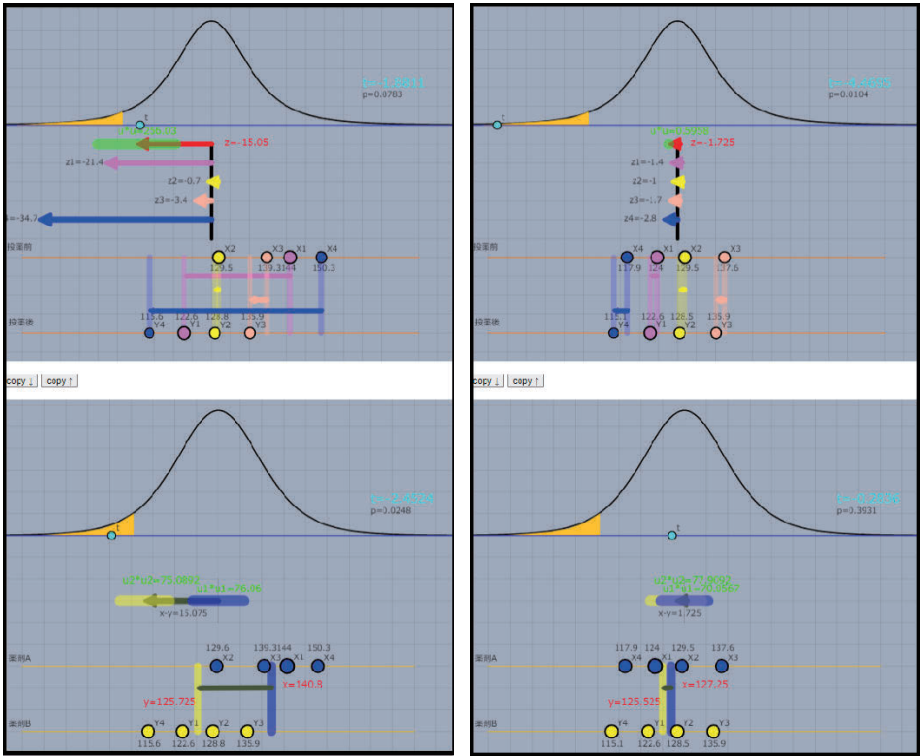


図8 検定結果が逆転する事例

4 実験授業実施とそれをふまえたコンテンツ・授業設計の改定

我々のこれまでの研究でも、数理統計を題材とした学習を行わせる場合、概念の言語的側面、数式表現、モデルの可視化などといった複数の modality にわたる情報を扱う必要から、学習者に高い認知負荷 [6] が掛かることがわかっている。本研究でも、実験授業における学習者の状況を観察し、こうした認知負荷を低減することがコンテンツ・授業設計改定の目的となった。ここでは、いわゆる *documentational approach* [7] の立場から、そうした改定の状況を報告する。

合計5回にわたる実験授業の実施にあたっては、該当者が数理統計の授業を受講してから数か月程度経過していたこともあり、まずは基本的な事項を教科書で振り返った後に、図7にある平常授業時の計算課題を用いて対応がある場合とない場合のそれぞれについて計算の流れを振り返るところから開始した。

問題文中における検定統計量の記載が簡略であることにもよると考えられるが、教科書の振り返りの際、対応ありの場合には投与前後のデータの引き算を先行させることを特に強調した最後の2回の授業の対象者以外は、投与前・投与後それぞれのデータの平均の計算を先行させ、結果として不偏分散の計算の際に混乱を生じた。このことは、正規分布に従う確率変数の線形結合が、再び正規分布に従う確率変数を与えるというメカニズムが、学習者にとっていかにとらえにくいものであるかを鮮明に示している。

最初の2回の実験授業における対象者のコンテンツ操作については、特に目標は定めず、いろいろとデータを取り替えて検定を繰り返す中で検定のメカニズムを観察してもらうことを指示し、いわば *goal free* の状況での試行となった。結果的に、なかなか探索目標が定まらず、標本平均の差の実現値が広がっているのに t 値が有意な方向に動かないケースが存在し、その背景に不偏分散の増大があるということ、また、同様のメカニズムが対応なしの場合にも働きうること、といった点に関する認識を持つまでに相当の時間を要した。図9は、2回の実験授業で使用され、数式表現をフルに表示させた状態のコンテンツであるが、左側に示されるモデルの可視化と、計算過程を示す右側の数式表現の部分の連関についてなかなか注意が向かなかったことが原因の一つであることが推察された。このため、標本平均や不偏分散の実現値として可視化されている部分と、数式表現でこれに該当する部分との対応をいかにして効果的に表出するかという点について、さらなる工夫が求められることとなった。その一方で、簡易化のために数式表現を除外することは避けた方が良いという指摘や、対応ありの場合となしの場合の計算のメカニズムの違いについて、認識しやすくする効果はあるだろうとの指摘もあった。また、さらに具体的な改善点として、可視化の中に「投与前」「投与後」といった状況説明の字句を補った方が良いとか、統計量の実現値は対応する図形要素の近くに表示させた方が良いといったアドバイスもあった。最終課題では、対応なしの場合のデータの順番を入れ替えることによる検定結果への影響について認識するのにも、かなりの時間を要した。この状況からは、特に大人数の平常授業で実施する場合、ある程度教員側で誘導する必要があることが示唆された。



図9 最初の2回の授業で使用されたコンテンツ

図10は、最初の2回での対象者からのフィードバックを受け、改定を施した上で3回目の実験で使用したコンテンツである。図中に見られる通り、説明字句の補足のほかに、可視化の中で統計量の実現値を色付けし、数式表現の方でも対応する値を同じ色で上塗りする形をとっている。



図10 第3回授業で使用されたコンテンツ

さらに、探索開始に先だって、標本平均の差の実現値が広がっているのに t 値が有意な方向に動かないケースを明示的に提示し、類似のケースがあった場合、その背景を数式表現にも注目しながら考察するように促した。対象者へのヒアリングの結果から、これによって可視化と数式表現との連関への意識が高まったことが推察され、操作の方向性が安定化された上、対応なしの場合でデータの順番を入れ替えることにより、対応ありの場合の検定結果を容易に変更しうる点を自力で認識することが可能となった。その一方で、可視化の中で、統計量の実現値の数値と、対応する図形要素とで色付けが異なっており、ややわかりづらいという指摘があった。

図 11 は、以上のフィードバックを受け、改定を施した上で第 4 回の実験で使用したコンテンツである。図中に見られる通り、標本平均に関わる部分は赤色、不偏分散に関わる部分は黄緑色、 t 値に関わる部分は水色に統一されている。



図 11 第 4 回授業で使用されたコンテンツ

第 3 回と同様の提示を行ったほか、操作途中で不偏分散の計算メカニズムを可視化をもとに考えるように促し、これをもとに標本平均の差と t 値との「不整合な」動きの原因を考えさせた。結果として、操作プロセスが不偏分散の動きの観察にフォーカスされ、第 3 回以前と比べて操作時間がかなり短縮され、対応ありの場合でデータの順番が検定結果に及ぼす重要性も迅速に認識できる結果となった。その一方で、対応ありの場合のみ不偏分散の計算過程が欠落しており、これを補った上で、さらにその計算過程に現れる標本平均も赤色で表示した方が良いとの指摘があった。また、この実験におけるように、操作状況に基づいて適時の教員による介入が行われないと、同様のスムーズな進行を得るのが難しい可能性についても言及があった。さらに踏み込んで、どのように動かして良いかわからない学生が出てくる可能性を想定し、初期状態をある程度限定し、動かすデータを指定して観察させるという方法についても提案があった。

図 5 は、以上のフィードバックを受け、改定を施した上で最終回の実験で使用したコンテンツである。図中に見られる通り、数式表現を追加したほか、対応ありの場合となしの場合を比較するコンテンツについても、配色を統一している。第 4 回の対象者からのフィードバックに基づき、対応ありのコンテンツ操作開始時点で

$$X1 = 110, Y1 = 100, X2 = 120, Y2 = 110$$

$$X3 = 130, Y3 = 130, X4 = 140, Y4 = 145$$

という指定をした上で、 $Y3$ の値をまず動かしてみるようにとの指示を与えた。結果的に検定は有意とはならず、 $Y3$ の値を減少させているにもかかわらず、途中から t 値が増加に転じるという現象を短時間のうちに観察することになる。実験の際は、これに引き続く数分のうちに、こうした逆転の原因が不偏分散の増大であることを認識した上で、

減少量が小さいにもかかわらず分散が小さいために結果が有意となる例を見出すに至った．必然的に，対応なしの場合の同様な事例や，データの順番を入れ替えることによって対応ありの場合の検定結果が反転する事例についても，ごく短時間のうちに見出すことが出来る結果となった．

最終回の実験のみ，学生2名による協調学習の形をとったことで，かなり多くのフィードバックが寄せられたが，平常授業の際も，学生個々による操作ではなく，協調学習の形式にした方が理解を得やすいのではないかという提案もあった．もっとも，平常授業の際に学生間の会話が少なからず発生している実態に鑑みると，実質的にはそれほどの差がないのではないかと期待される．

5 結論と今後の研究の方向性

いみじくも対象となった学生自身の言葉にある通り，本研究で用いたコンテンツに記載される数式表現の情報は一般的な学生にとって相当程度負荷の高い情報を含むものであり，探索の範囲を限定してその一部をフォーカスするなど，負荷を軽減する方策が必要だと推察される．対象者が多くはないものの，本研究における5回の実験授業を通じて，大人数の平常授業でもある程度実行可能な授業計画が得られたものと考えられる．今後，実際にこの形で平常授業を実施し，従前のような学習者による計算方法の誤選択がどの程度まで減らせるか，検証したいと考えている．

教材作成システムの面でも改善が必要と思われる点が2点出てきた．1つ目はKaTeXに関わることである．今回作成した教材では，計算過程を示すためにKaTeXを用いている．これまで数式表示が必要な場合はQuillを使っていたため，KaTeXを直接使う機能を備えていなかった．今回の教材は数式が長く複雑になるため，Quillでは不十分で直接KaTeXを使う必要があったことから，上述のようにHTMLファイルの一部を手で直接書く必要があった．こうした状況を踏まえ，KaTeXを直接使うためのボタンをシステムに装備する予定である．2つ目は教材におけるWebの各要素の配置に関わることである．本システムはHTMLの各要素をWeb画面に配置していく形で教材を作成する．Web要素はある程度自由に配置可能であるが，完全に自由に配置できるわけではない．今回の教材では，Cinderellaの横に計算式を表示するためKaTeXを用いたが，Web要素の横に別のWeb要素を配置する機能を本システムは持っていなかった．そこでHTMLのレイアウトを操作する命令をHTMLファイルに手で書き込んだが，今後はレイアウトを操作するための仕組みをシステムに導入する必要がある．

また，第4回実験の対象者の指摘にもある通り，探索に自由度を持たせれば持たせるほど，教員が介入して方向づけする必要性が生じる可能性が上昇する．そのためには，当然，教員が学習者個々の探索状況を逐次把握できていることが前提となる．本研究のコンテンツに実装されている操作ログの取得・送信機能はそうした目的に適う可能性を秘めるが，統計量を決定する要因が少なくないだけに，得られた操作ログから学習者の思考状況をどのように推定するかという点で，かなりの困難を伴うことが予想される．この課題を克服するためには，最終回の実験におけるような協調学習の機会を積み重ねることで，学習者間の対話から思考状況を推定し，これと操作パターンとを対照するほ

かないと考えられ，やはり今後の研究課題である．

謝辞

本研究は JSPS 科研費 21K02752，22K02951 の助成を受けている．

参考文献

- [1] T. Kitamoto, M. Kaneko, S. Takato : "E-learning system with Computer Algebra based on JavaScript programming language", Proc. of ATCM 2018, Yogyakarta, 2018, 123-133.
- [2] Cinderella Official Home Page : URL <https://www.cinderella.de/tiki-index.php> (2023 年 12 月 7 日閲覧)
- [3] CindyJS Official Home Page : URL <https://cindyjs.org/> (2023 年 12 月 7 日閲覧)
- [4] KaTeX Official Home Page : URL <https://katex.org/> (2023 年 12 月 7 日閲覧)
- [5] Quill Official Home Page : URL <https://quilljs.com/> (2023 年 12 月 7 日閲覧)
- [6] J. Sweller, P. Ayres, S. Kalyuga : Cognitive Load Theory, Springer, 2011.
- [7] G. Gueudet, B. Pepin, L. Trouche : From Text to Lived Resources: Mathematics Curriculum Materials and Teacher Development, Springer, 2013.