

独立性の検定に関する HTML をベースとした教材の操作口 グの分析

北本卓也 (山口大学)、金子真隆、野田健夫 (東邦大学)

発表の構成

- 独立性の検定に関する HTML ベースの教材 (北本)
- 教材を用いた学習活動における学習者の行動の分析 (金子)

数理統計教材

- データサイエンス教育の重要性が指摘されている。
https://www.mext.go.jp/a_menu/koutou/suuri_datascience_ai/00002.htm
- データサイエンスの重要な手法の1つとして数理統計学があり、統計的推定や検定など様々な応用がある。
- 医療系の学部では、統計的検定が重要 (薬の臨床試験を行った結果、本当に薬が効果があるかどうか)。
- これには仮説検定という手法が用いられるが、わかりにくいという声が学生から上がっている。
- そこで、試行錯誤を行いながら仮説検定の手法を理解するための Web 上の教材を作成した。

独立性の検定

- 薬の効果が本当にあるかどうかのチェックに使うことができる。
- 薬剤群とコントロール群に分けて、それぞれの群での「治った」「治らなかった」場合の数をチェック

| | A | B | C | D |
|---|--------|-----|--------|-----|
| 1 | | 治った | 治らなかった | 合計 |
| 2 | 薬剤群 | 50 | 30 | 80 |
| 3 | コントロール | 50 | 50 | 100 |
| 4 | 合計 | 100 | 80 | 180 |

独立性の検定

下の表の B2, B3, C2, C3 のセルにデータを入力してください。

| | A | B | C | D |
|---|--------|-----|--------|-----|
| 1 | | 治った | 治らなかった | 合計 |
| 2 | 薬剤群 | 50 | 30 | 80 |
| 3 | コントロール | 50 | 50 | 100 |
| 4 | 合計 | 100 | 80 | 180 |

(STEP1) データの入力

期待度数の表

| | A | B | C | D |
|---|--------|-------------|-------------|-----|
| 1 | | 治った | 治らなかった | 合計 |
| 2 | 薬剤群 | 44.44444444 | 35.55555555 | 80 |
| 3 | コントロール | 55.55555555 | 44.44444444 | 100 |
| 4 | 合計 | 100 | 80 | 180 |

(STEP2) 期待度数の表の作成

5

独立性の検定

χ^2 の値を計算してみましょう。

$$\begin{aligned}\chi^2 &= \sum_{i,j} \frac{(a_{ij} - e_{ij})^2}{e_{ij}} \\ &= \frac{(a_{11} - e_{11})^2}{e_{11}} + \frac{(a_{12} - e_{12})^2}{e_{12}} + \frac{(a_{21} - e_{21})^2}{e_{21}} + \frac{(a_{22} - e_{22})^2}{e_{22}} \\ &= \frac{(60 - 55)^2}{55} + \frac{(40 - 45)^2}{45} + \frac{(50 - 55)^2}{55} + \frac{(50 - 45)^2}{45} \\ &= 2.02\end{aligned}$$

(STEP3) χ^2 二乗値の計算

(STEP4) p 値の計算



6

教材の機能

- データを入力すると、独立性の検定を自動的に計算する。
- 計算過程も画面に表示する。
- カイ二乗分布のグラフは Cinderella を用いて表示。
- カイ二乗の値は Numerical Recipes in C のソースコードを Cinderella に移植して計算

| | A | B | C | D |
|---|--------|-------------|-------------|-----|
| 1 | | 治った | 治らなかった | 合計 |
| 2 | 薬剤群 | 44.44444444 | 35.55555555 | 80 |
| 3 | コントロール | 55.55555555 | 44.44444444 | 100 |
| 4 | 合計 | 100 | 80 | 180 |

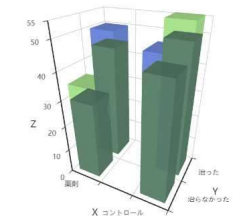
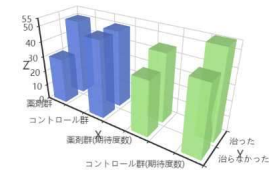
$$\begin{aligned}\chi^2 &= \sum_{i,j} \frac{(a_{ij} - e_{ij})^2}{e_{ij}} \\ &= \frac{(a_{11} - e_{11})^2}{e_{11}} + \frac{(a_{12} - e_{12})^2}{e_{12}} + \frac{(a_{21} - e_{21})^2}{e_{21}} + \frac{(a_{22} - e_{22})^2}{e_{22}} \\ &= \frac{(60 - 55)^2}{55} + \frac{(40 - 45)^2}{45} + \frac{(50 - 55)^2}{55} + \frac{(50 - 45)^2}{45} \\ &= 2.02\end{aligned}$$



7

教材の機能

- 入力したデータと期待度数の表を 3 D グラフで表示。
- その 3 D グラフを重ねて表示することでデータの差が重要であることを示す。



8

画面の構成

χ^2 の値を計算してみましょう。

$$\begin{aligned}\chi^2 &= \sum_{ij} \frac{(a_{ij} - e_{ij})^2}{e_{ij}} \\ &= \frac{(a_{11} - e_{11})^2}{e_{11}} + \frac{(a_{12} - e_{12})^2}{e_{12}} + \frac{(a_{21} - e_{21})^2}{e_{21}} + \frac{(a_{22} - e_{22})^2}{e_{22}} \\ &= \frac{(60 - 55)^2}{55} + \frac{(40 - 45)^2}{45} + \frac{(50 - 55)^2}{55} + \frac{(50 - 45)^2}{45} \\ &= 2.02\end{aligned}$$

下の表の B2, B3, C2, C3 のセルにデータを入力してください。

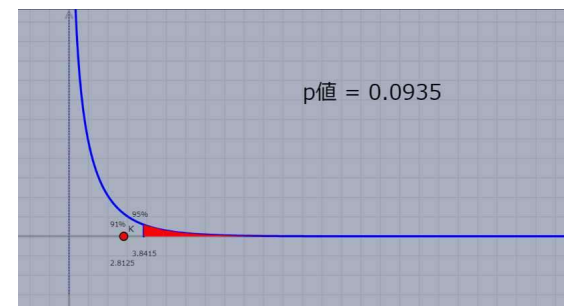
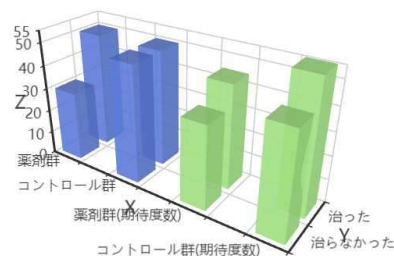
| | A | B | C | D |
|---|--------|-----|--------|-----|
| 1 | | 治った | 治らなかった | 合計 |
| 2 | 薬剤群 | 50 | 30 | 80 |
| 3 | コントロール | 50 | 50 | 100 |
| 4 | 合計 | 100 | 80 | 180 |

KaTeX

Jspreadsheet

9

画面の構成



Apache Echarts

Cinderella, CindyJS

10

KaTeX

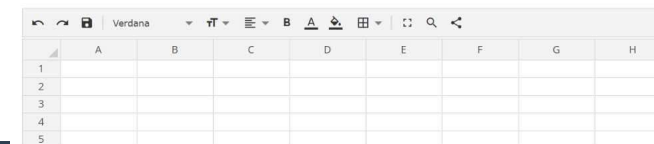
- ブラウザ上で TeX のソースをレンダリングするエンジン (JavaScript で書かれている)
- MathJax も同様にソフトウェアであるが、KaTeXの方が高速 (ただし、機能は MathJaxの方が上)
- Webで検索したところ、LATEX.js というものもあり (見たところ良さそうだが使ったことはない)。

<https://latex.js.org/>

11

Jspreadsheet

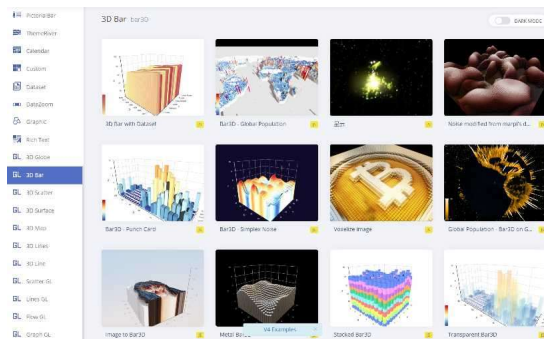
- ブラウザ上で Excel のような表計算を実現する JavaScript のライブラリ
- MIT License で高機能の Pro 版がある
- <https://www.wasp.co.jp/blog/169> によると x-spreadsheet という似た機能を持つものもある (こちらは CANVAS を使う)
- ツールバーを設定することも可能



12

Apache Echarts

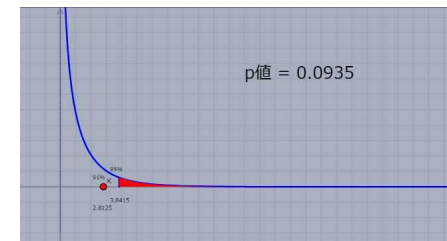
- ブラウザ上で 3D を含む様々なグラフを描画するための Javascript のライブラリー
- Apache License 2.0
- マウス操作で 3D グラフの視点の変更、拡大・縮小が可能



13

Cinderella, CindyJS

- Cinderella は単体で動作する動的幾何ソフトウェア。Cinderella 上で CindyScript というプログラミング言語でプログラムが動く。
- Cinderella で作成した図はホームページ (html) に変換する事が可能。
- CindyJS は Cinderella をブラウザ上で動かすためのライブラリー。



14

教材の特徴

- 独立性の検定の教材の例
<https://bellcurve.jp/statistics/blog/14038.html>
- 上の例と比較すると学習者が自分でデータを入力して試行錯誤することが可能 ⇒ 学習内容の深い理解につながると期待

15

教材の変遷

下の表の a_{ij} がデータが入る所です。

| | B_1 | ... | B_m | \sum |
|--------|----------|-----|----------|--------|
| A_1 | a_{11} | ... | a_{1m} | T_1 |
| ... | ... | ... | ... | ... |
| A_l | a_{l1} | ... | a_{lm} | T_l |
| \sum | S_1 | ... | S_m | S |

| | A | B | C | D |
|---|--------|-----|--------|-----|
| 1 | | 治った | 治らなかった | 合計 |
| 2 | 薬剤群 | 70 | 30 | 100 |
| 3 | コントロール | 40 | 40 | 80 |
| 4 | 合計 | 110 | 70 | 180 |

右の表の B2, B3, C2, C3 のセルにデータを入力してください。
D2,D3,B4,C4 にはそれぞれの行、列の合計が自動的に入ります。

下の期待度数の表を計算します。

| | B_1 | ... | B_m | \sum |
|--------|----------|-----|----------|--------|
| A_1 | e_{11} | ... | e_{1m} | T_1 |
| ... | ... | ... | ... | ... |
| A_l | e_{l1} | ... | e_{lm} | T_l |
| \sum | S_1 | ... | S_m | S |

| | A | B | C | D |
|---|--------|-------------|-------------|-----|
| 1 | | 治った | 治らなかった | 合計 |
| 2 | 薬剤群 | 61.11111111 | 38.88888888 | 100 |
| 3 | コントロール | 48.88888888 | 31.11111111 | 80 |
| 4 | 合計 | 110 | 70 | 180 |

ここで、 e_{ij} は $T_i \times \frac{S_j}{S}$ で計算されます。

16

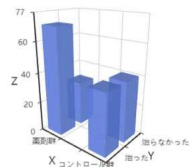
教材の変遷

下の表の a_{ij} がデータが入る所です。

| | | | | |
|-------|----------|-----|----------|-------|
| | B_1 | ... | B_m | 計 |
| A_1 | a_{11} | ... | a_{1m} | T_1 |
| ... | ... | ... | ... | ... |
| A_i | a_{i1} | ... | a_{im} | T_i |
| ... | ... | ... | ... | ... |
| 計 | S_1 | ... | S_m | S |

右の表の B2, B3, C2, C3 のセルにデータを入力してください。
D2,D3,B4,C4 にはそれぞれの行、列の合計が自動的に入ります。

| | | | | |
|---|--------|-----|--------|-----|
| | A | B | C | D |
| 1 | | 治った | 治らなかった | 合計 |
| 2 | 薬剤群 | 70 | 30 | 100 |
| 3 | コントロール | 40 | 40 | 80 |
| 4 | 合計 | 110 | 70 | 180 |

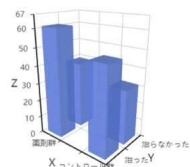


下の期待度数の表を計算します。

| | | | | |
|-------|----------|-----|----------|-------|
| | B_1 | ... | B_m | 計 |
| A_1 | e_{11} | ... | e_{1m} | T_1 |
| ... | ... | ... | ... | ... |
| A_i | e_{i1} | ... | e_{im} | T_i |
| ... | ... | ... | ... | ... |
| 計 | S_1 | ... | S_m | S |

ここで、 e_{ij} は $T_i \times \frac{S_j}{S}$ で計算されます。

| | | | | |
|---|--------|-------------|-------------|-----|
| | A | B | C | D |
| 1 | | 治った | 治らなかった | 合計 |
| 2 | 薬剤群 | 61.11111111 | 38.88888888 | 100 |
| 3 | コントロール | 48.88888888 | 31.11111111 | 80 |
| 4 | 合計 | 110 | 70 | 180 |



17

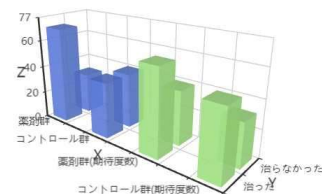
教材の変遷

下の表の B2, B3, C2, C3 のセルにデータを入力してください。

| | | | | |
|---|--------|-----|--------|-----|
| | A | B | C | D |
| 1 | | 治った | 治らなかった | 合計 |
| 2 | 薬剤群 | 70 | 30 | 100 |
| 3 | コントロール | 40 | 40 | 80 |
| 4 | 合計 | 110 | 70 | 180 |

期待度数の表

| | | | | |
|---|--------|-------------|-------------|-----|
| | A | B | C | D |
| 1 | | 治った | 治らなかった | 合計 |
| 2 | 薬剤群 | 61.11111111 | 38.88888888 | 100 |
| 3 | コントロール | 48.88888888 | 31.11111111 | 80 |
| 4 | 合計 | 110 | 70 | 180 |



18

教材の変遷

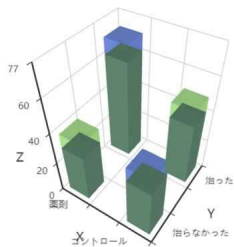
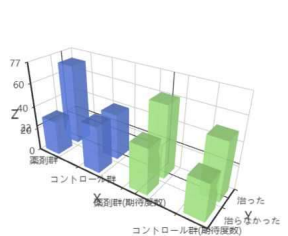
下の表の B2, B3, C2, C3 のセルにデータを入力してください。

| | | | | |
|---|--------|-----|--------|-----|
| | A | B | C | D |
| 1 | | 治った | 治らなかった | 合計 |
| 2 | 薬剤群 | 70 | 30 | 100 |
| 3 | コントロール | 40 | 40 | 80 |
| 4 | 合計 | 110 | 70 | 180 |

期待度数の表

| | | | | |
|---|--------|-------------|-------------|-----|
| | A | B | C | D |
| 1 | | 治った | 治らなかった | 合計 |
| 2 | 薬剤群 | 61.11111111 | 38.88888888 | 100 |
| 3 | コントロール | 48.88888888 | 31.11111111 | 80 |
| 4 | 合計 | 110 | 70 | 180 |

グラフ結合



19

今後の方向性

- 学習者の行動の分析のために操作のログを取っている。
- 終了後にログを「送信」をクリックさせてサーバーへ送信している。
- 定期的にログを自動的にサーバーへ送信するようにシステム変更を検討中。
- サーバーのデータを元にダッシュボードを作成
- 学習者の行動を見てリアルタイムに対応をする、または行動を見てアラートを発する、などが出来ないか。

20

今後の方向性

