



実験授業の実施状況



1. 調査の背景



COGNITIVE LOAD THEORY

(1) Human cognitive architecture

(Sweller, Morrienboer, Paas, 1998)

(2) Measurement of cognitive load 探索的活動での操作のゆらぎ？

subjective measure, secondary task measure,
physiological measure, ...

(Zheng 2018)

(3) Instructional effects generated by CLT

Goal free effect

Worked example effect

Split-Attention effect

Modality effect

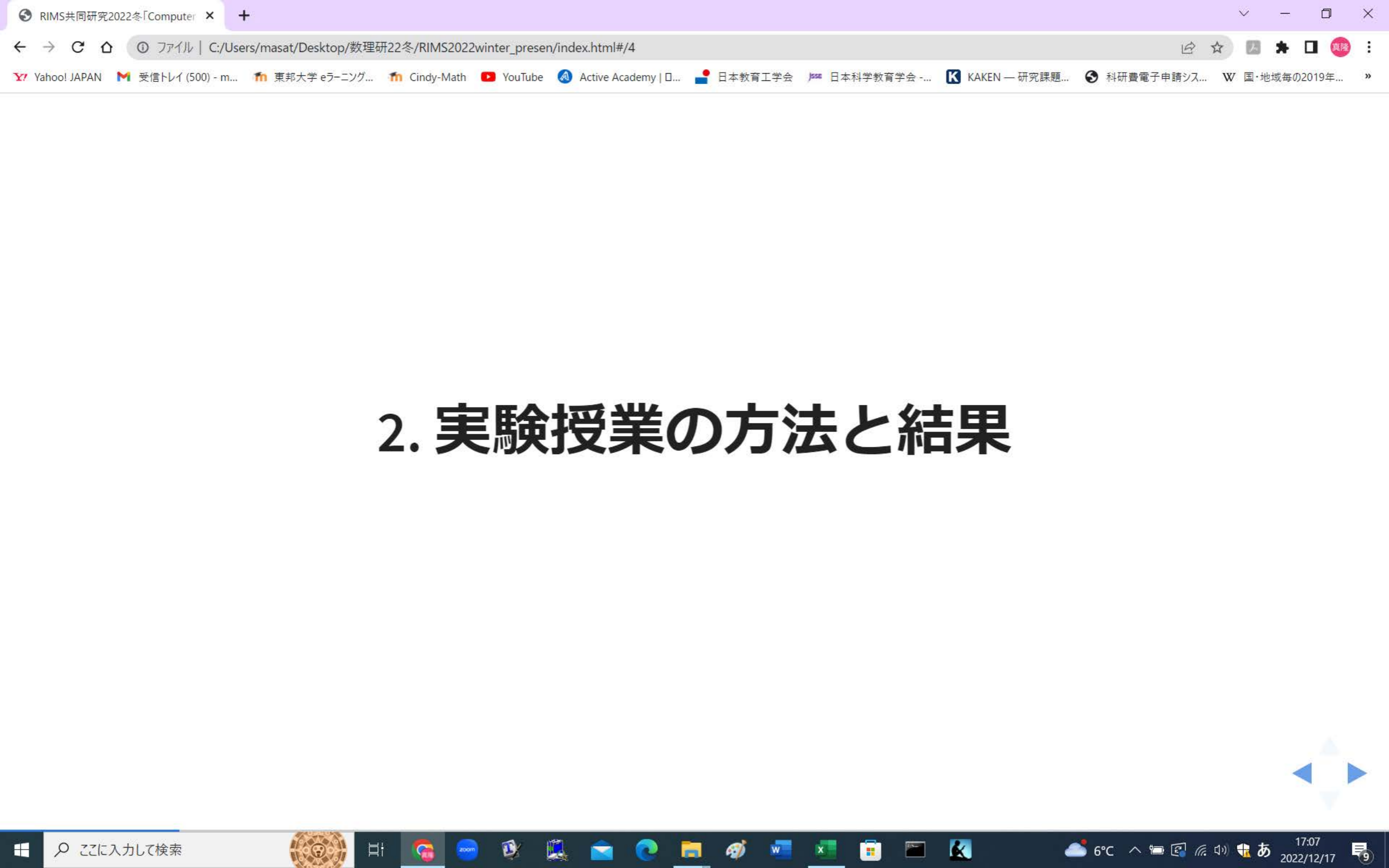
Expertise reversal effect

Element interactivity effect

....

(Sweller, Ayres, Kalyuga, 2011)





2. 実験授業の方法と結果

対象者

- 数理統計の基礎（検定の概念など）を履修中の大学2年生
- 女子80名 男子29名 計109名
- 通常授業の中で、MoodleをLMSとして使用
- Moodleではスライド教材掲載・多肢選択式の小テストを利用
- 数式入力の経験はなし
- CindJSによる動的コンテンツの提示は毎回実施
- 「帰無仮説」「対立仮説」「検定統計量」「有意水準」「p値」といった検定の基本概念については、数値例とともに、母平均や母分散の検定など複数種類の検定について提示・計算練習済



実施フロー

教科書の読み合わせ

84

3.1 いろいろな検定

次の表は、ある都市において1年間観測した毎日の交通平均混雑指数(人)の観測数(観測度数)である。この観測度数のポアソン分布へのあてはまり方を検定しよう。

混雑指数	0	1	2	3	4	5以上	計
観測数	72	100	91	60	22	20	365
期待度数	40	100	98	59	26	13	365

データから求めた1日平均混雑指数は1.90であった。表の期待度数は、母数λをこの値で置き換えたポアソン分布(3.1)によって計算したものである。未知母数がない場合、検定統計量(5.1)の自由度はn-1であるが、あてはめる理論分布に含まれる未知母数をデータから推定した場合、推定した母数の数だけ自由度が小さくなることが知られている。今の場合、推定された未知母数は1つであるから、(5.1)の自由度は(n-1)-1=64-2=62=6となる。有意水準を5%とすると、検定値は

$$\chi^2 \geq \chi^2_{0.05}(62) = 78.48$$

χ^2 の観測値を計算すると

$$\chi^2 = \frac{(72-40)^2}{40} + \frac{(100-100)^2}{100} + \frac{(91-98)^2}{98} + \frac{(60-59)^2}{59} + \frac{(22-26)^2}{26} + \frac{(20-13)^2}{13} = 5.645$$

この値は棄却域に入らないから、観測度数がポアソン分布に従っていないとはいえない。

5.1.2 独立性の検定

母集団が2つの属性からなる場合、たとえば「学歴と支持政党」や「健康状態と喫煙習慣」など、それらの属性の間に互いに関連があるかないか、すなわち独立であるかどうかを判定する独立性の検定について説明しよう。

検定値は、以下のような表にまとめられる。これを分割表またはクロス表という。分割表の一番上の行には、属性の分類を書く。これを表頭という。また、一番左の列にも属性の分類を書く。これを表側という。

学歴	関	不関	計
関	70	0	70
不関	0	80	80
計	70	80	150

属性の間に関連のある場合、分割表の表側がわかれば表頭の表側は簡単に予想できる。

学歴	関	不関	計
関	10	60	70
不関	20	60	80
計	30	120	150

属性の間に関連がない時、分割表で関をとりた人と不関をとりた人、それぞれの歴史の成順の内部が完全に比例している。

85

5.1 適合度と独立性の検定

表の作り方から、2つの属性の関連が検出できれば対角線に沿って大きな値が出てくると考えられる。その極端な場合が前ページの左側の表である。一方、右側の表は関連がない場合の例である。

上の分割表では、2つの属性の属性はそれぞれ関と不関の2種類に分類され、観測度数は分割された1つのセルに配置されている。このような分割表を2×2分割表という。一般に、表側の属性がr種類に、表頭の属性がc種類に分類され、観測度数がr×c個のセルに配置されている場合、r×c分割表という。

右の表は、ある大手企業の従業員から無作為に抽出した195人について、喫煙習慣と健康状態を同時に調べた結果をまとめた2×2分割表である。

健康状態	関	不関	計
喫煙習慣	45	25	70
なし	90	35	125
計	135	60	195

このデータを用いて、喫煙習慣と健康状態の間に関連があるかどうかを有意水準5%で検定する方法を考えよう。まず、表のように分割をおく。

帰無仮説 H_0 : 喫煙習慣と健康状態の間に関連がない
対立仮説 H_1 : 喫煙習慣と健康状態の間に関連がある

期待度数を確率と見なせば、195人の従業員から無作為に1人選ぶとき、喫煙習慣のある従業員である確率は $\frac{70}{195}$ 、健康状態が良い従業員である確率は $\frac{135}{195}$ 、だから、喫煙習慣と健康状態の間に関連がなければ、テーブルの右上のセルより、喫煙習慣があり健康状態が良い従業員である確率は $\frac{70}{195} \times \frac{135}{195}$ したがって、195人のうち喫煙習慣があり健康状態が良い従業員の数は $195 \times \frac{70}{195} \times \frac{135}{195} = 48.5$ であると期待される。

同時に、それぞれの期待度数を計算すると、右の表のようになる。

健康状態	関	不関	計
喫煙習慣	48.5	21.5	70
なし	86.5	38.5	125
計	135	60	195

適合度の検定と同様に、統計量

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(\text{観測度数} - \text{期待度数})^2}{\text{期待度数}} \quad (5.2)$$

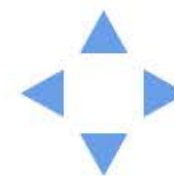
を算出する。この統計量は、標準の大きさが大きいとき、近似的に自由度1の χ^2 分布に従うことが知られている。期待度数と観測度数のズレが大きいたま χ^2 の値は大きくなり、また、 χ^2 の値が大きいたま H_0 は棄却される。 χ^2 分布表より、 $\chi^2_{0.05}(1) = 3.84$ だから、有意水準5%のときの棄却域は



実施フロー

教科書の読み合わせ

- 具体例の数値を見ながら、相対する概念としての「独立性」と「相関性」の言語的なイメージを確立
- 独立性が成り立つ場合、クロス集計表の値が「期待度数」に相当することから、期待度数の計算の仕方を説明
- 所与のクロス集計表の傾向が、「独立性」と「相関性」のどちらに近いかわかるためには、集計表の数値と期待度数との偏差を測ることが自然である点を説明
- 本検定に関しては、「独立性」の成立が帰無仮説、「相関性」の成立が対立仮説となることを説明



実施フロー

スライドでのサマリー

復習 独立性の検定

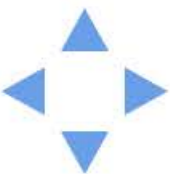
互いに独立な2つの属性をもつ母集団の $l \times m$ 分割表で

$$\chi^2 = \sum_{i,j} \frac{(a_{ij} - e_{ij})^2}{e_{ij}}$$

は、標本の大きさ S が大きいとき自由度 $(l - 1)(m - 1)$ の χ^2 分布に従う。

	B_1	...	B_m	計
A_1	a_{11}	...	a_{1m}	T_1
...
A_l	a_{l1}	...	a_{lm}	T_l
計	S_1	...	S_m	S

e_{ij} は期待度数



RIMS共同研究2022冬「Computer」 x +

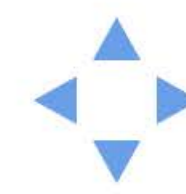
ファイル | C:/Users/masat/Desktop/数理研22冬/RIMS2022winter_presen/index.html#/6/3

Yahoo! JAPAN 受信トレイ (500) - m... 東邦大学 eラーニング... Cindy-Math YouTube Active Academy | 日本教育工学会 JSEE 日本科学教育学会 KAKEN 研究課題... 科研費電子申請シス... 国・地域毎の2019年...

実施フロー

スライドでのサマリー

- 帰無仮説と対立仮説の設定から、集計表の数値と期待度数の偏差をもとにした χ^2 検定量を選択することが自然であることを説明
- 設定した統計量が自由度1の χ^2 分布に従うことを説明
- 通常の場合と比べ、自由度が1の場合は χ^2 分布の密度関数のグラフが変則的な形になることを復習
- 有意水準を5%とすることにより、棄却域は χ^2 分布の密度関数のグラフにおける右側確率5%の領域に設定されることを説明
- 「正の相関」と「負の相関」については特にふれていない



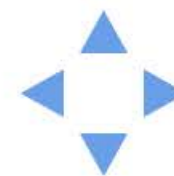
ここに入力して検索

17:08 2022/12/17 6°C

実施フロー

コンテンツの提示と操作内容の説明

- 属性をひとつを「薬剤投与群とコントロール群」、もうひとつを「病変の治癒の可否」とする設定を説明
- コントロール群の数値を60と40に固定
- 薬剤投与群の総数を100とする制約条件のもと、治癒群の数値を変更した際に χ^2 値や p 値がどのように変動するか、 χ^2 分布の密度関数のグラフと対照しながら10分程度確認してもらう
- 操作前に事後課題についても説明
コントロール群の数値を 60/40 から少し動かした上で、 p 値を0.9, 0.1, 0.05, 0.01 の各値に最も近づけるような、治癒群の数値を予想してもらいたい



実施フロー

思考状況の確認

- 少人数を対象とした予備実験の際のような口頭での確認は困難
- 操作時の入力値のログをタイムスタンプ付きで集積
探索時の入力値のゆらぎに認知負荷の高さが現れていないか
操作に要する時間の遷移からもimplicationが得られないか
- ポストテスト
操作時の治癒群の数値と p 値とを一覧表として提示
コントロール群の数値を 62/38 に変更し、指定した p 値である 0.9, 0.1, 0.05, 0.01 の各値に最も近づけるような、治癒群の数値を予想させた
コントロール群に合わせて薬剤投与群の入力値も 2 増やせば、期待度数も 2 増えて偏差は変化せず、 χ^2 値も大きく変化しないことを判断できるか



実施結果のポイント

ポストテストの解答状況

- 完全正解かそれに近い (p 値が 0.01 の場合のみずれたもの) と判断されるもの 13例
- 部分正解と判断されるもの (全体にフィットは良いが、 p 値が 0.9 の場合に62よりも小さい値を予想してしまっている) 8例
- 不正解と判断されるもの 88例

このうち

ヒントの表の数値をそのまま読んだと判断されるもの 35例

全体に60よりも小さい数値が入力されているもの 2例

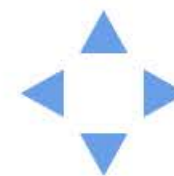
(ヒントの表が効果を発揮したと考えられる)



実施結果のポイント

操作ログの状況

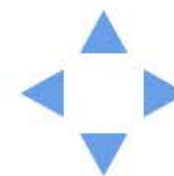
- 数値を全角で入力したり、数字の0をアルファベットのoで入力したりといった誤入力が目立った
- 指定された p 値の近辺以外に該当する入力値や、60未満の入力値を試す場面も少なからず見られる
- 1回の数値入力にとどまり、その後の探索が進んでいない事例が複数（4例）ある。これはデータ送信上のトラブルである可能性は小さく、周囲の友人の操作を観察していた可能性が高い
- 正解群と不正解群の間で、操作時間に関する有意差は見られず、全体として、数式表現・棒グラフ等の要素単体での理解でなく、これらの要素を統合した理解が目指されたと推察



実施結果のポイント

ポストテストの解答解釈

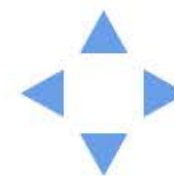
- 完全正解かそれに近いもの
徹底して入力値を2大きくできていることから、 χ^2 統計量の数式表現に十分に注意が払われたと判断される
- 部分正解と判断されるもの
 χ^2 統計量の数式表現に一定程度注意が払われたものの、期待度数の動きについて理解が十分でなかったと判断される
- 不正解と判断されるもの
入力値の増減と χ^2 値の大小との連関については観察できたものの、数式表現までは注意が向かなかったと判断される



実施結果のポイント

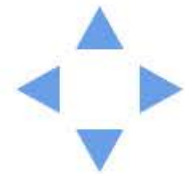
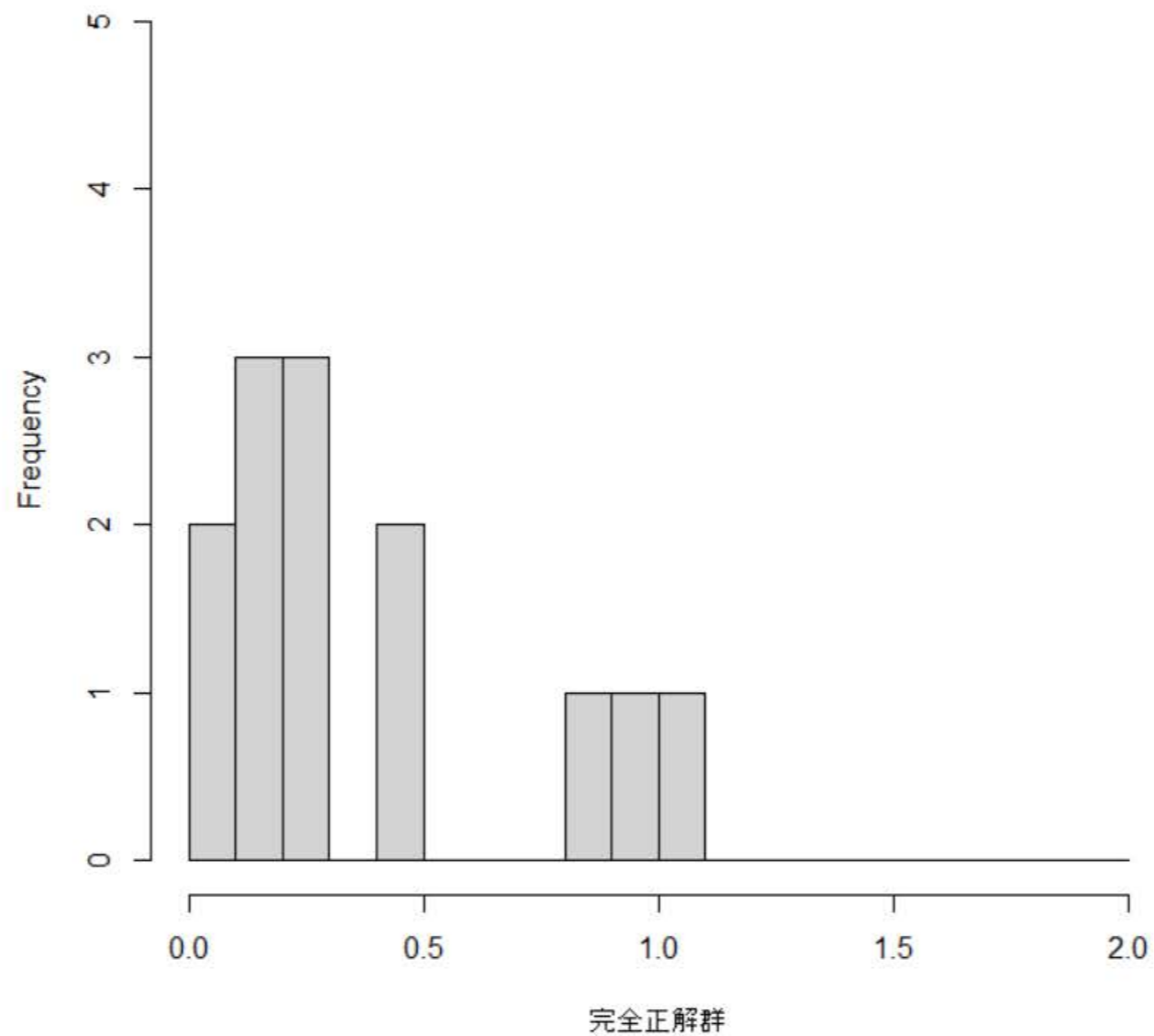
操作ログに関する群間の比較

- 不正解群の中だけでも操作プロセスは多岐にわたり、理解状況を操作ログのみで特徴づけるには限界があるか？
- ただ、正解群にはほとんど見られないのに不正解群には一定程度見られる操作パターンが存在 => at risk のシグナル？
正解ゾーン（60から80）を飛び越えて上下の領域を往復
試しに、入力値の総変動量の時間平均 ($/s$) のヒストグラムを作成してみると・・・



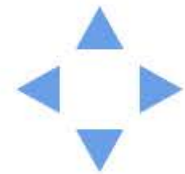
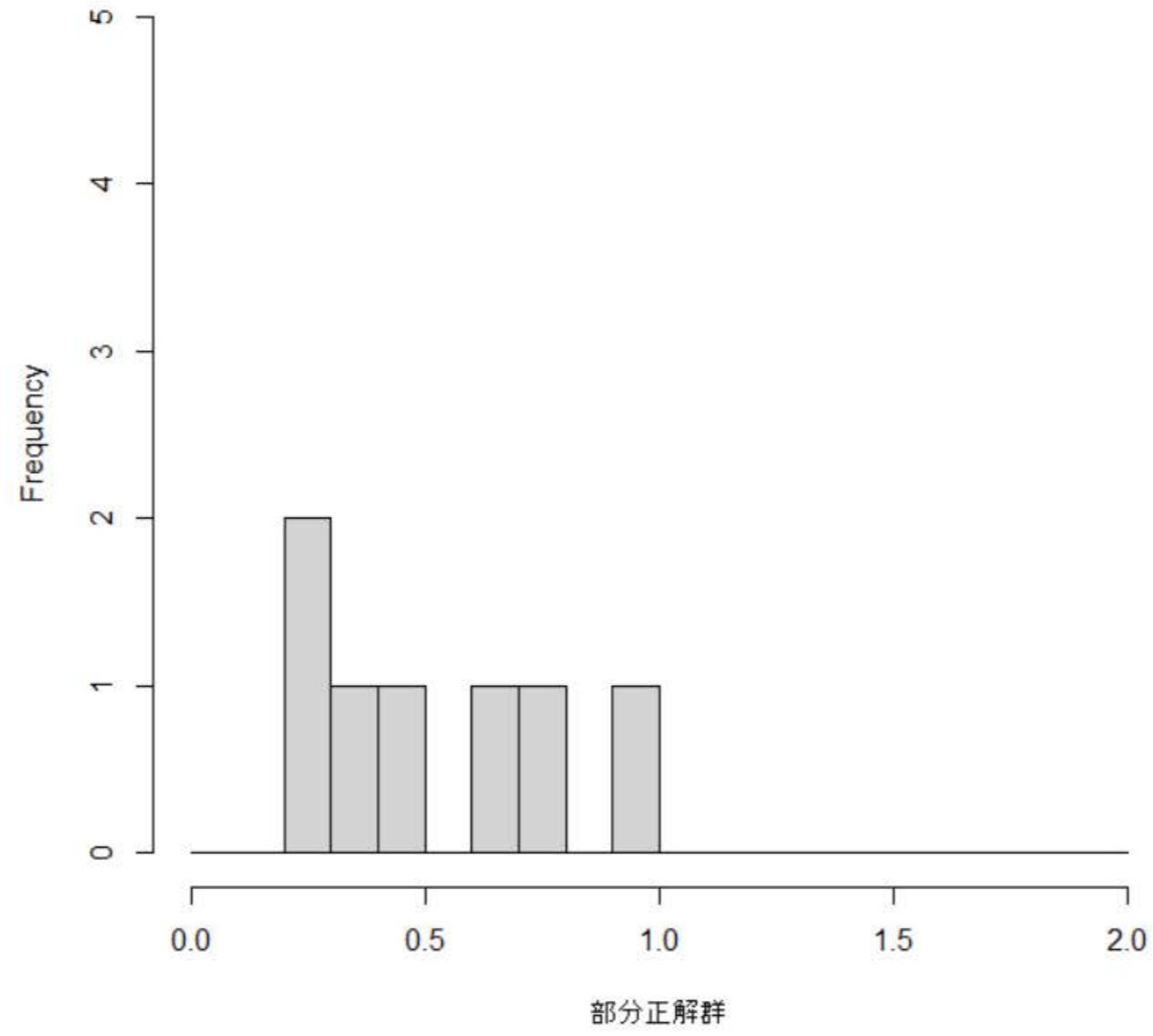
実施結果のポイント

Histogram of 完全正解群



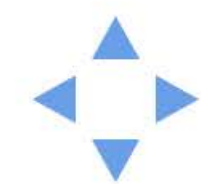
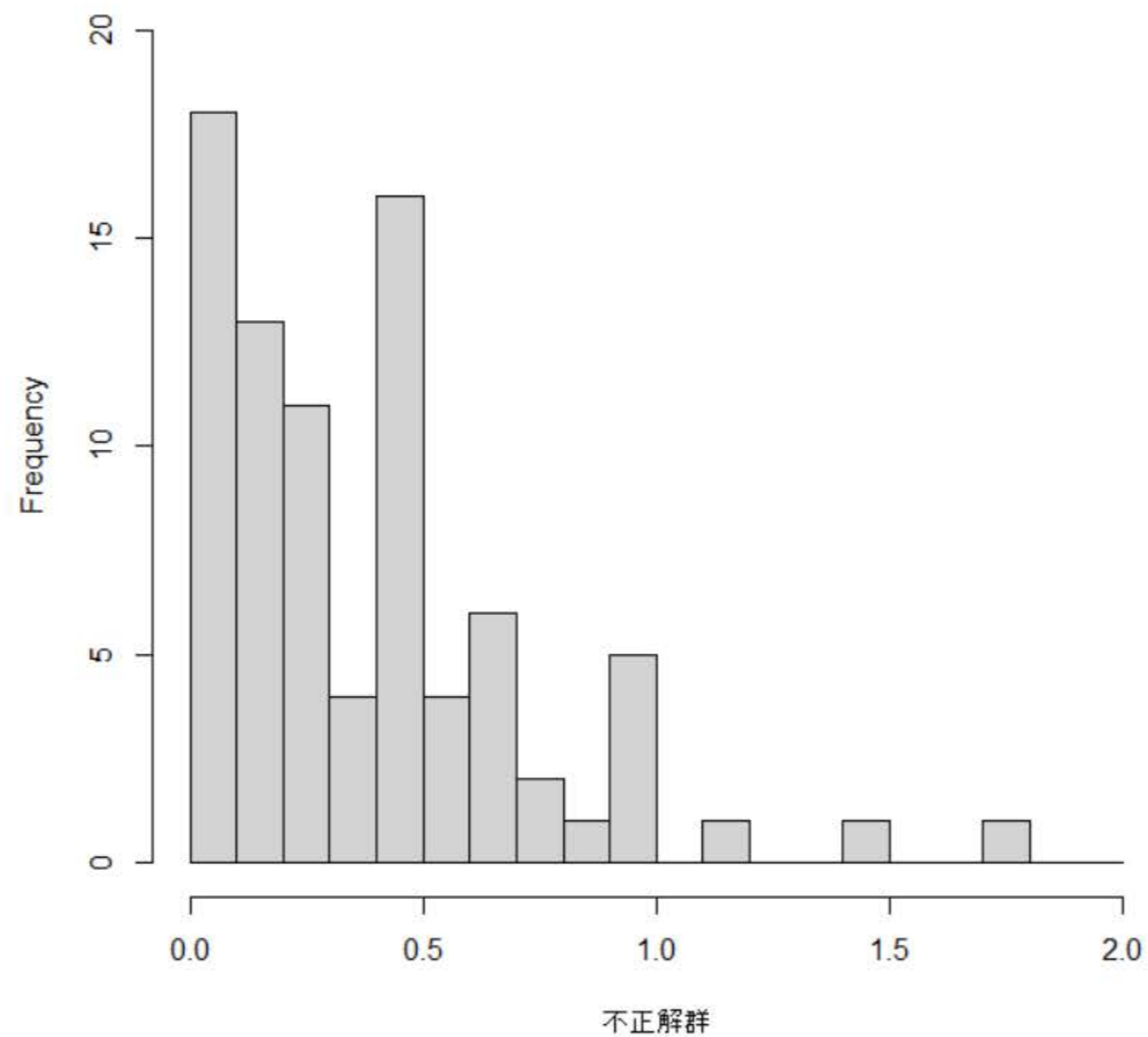
実施結果のポイント

Histogram of 部分正解群



実施結果のポイント

Histogram of 不正解群



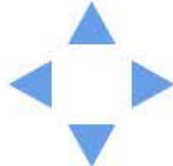
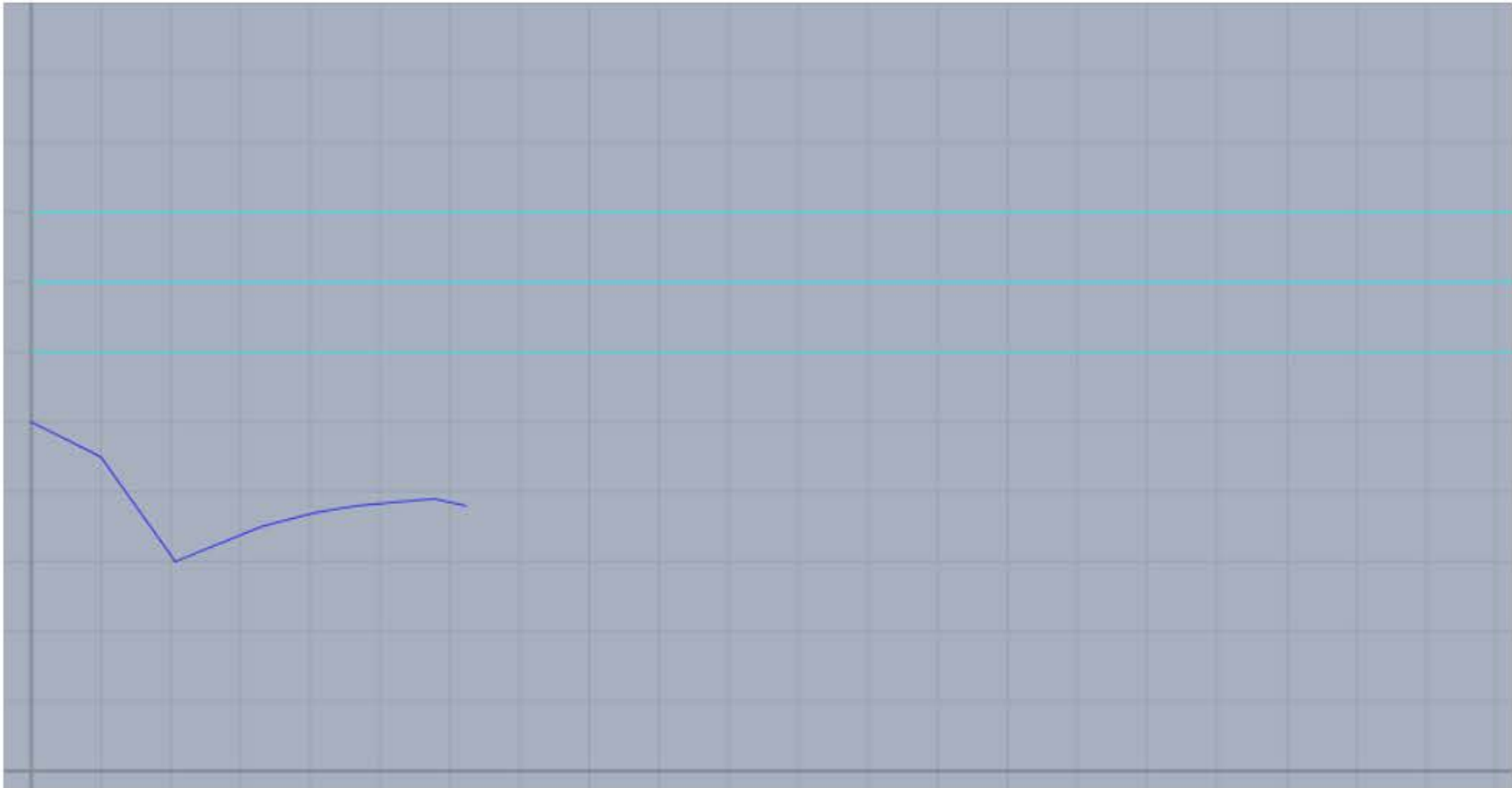
RIMS共同研究2022冬「Computer」

ファイル | C:/Users/masat/Desktop/数理研22冬/RIMS2022winter_presen/index.html#/7/7

Yahoo! JAPAN 受信トレイ (500) - m... 東邦大学 eラーニング... Cindy-Math YouTube Active Academy | 日本教育工学会 日本科学教育学会 KAKEN 研究課題... 科研費電子申請シス... 国・地域毎の2019年...

実施結果のポイント

不正解群で特に心配なパターン



ここに入力して検索

17:11 2022/12/17


RIMS共同研究2022冬「Computer」

ファイル | C:/Users/masat/Desktop/数理研22冬/RIMS2022winter_presen/index.html#/7/8

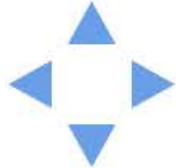
Yahoo! JAPAN 受信トレイ (500) - m... 東邦大学 eラーニング... Cindy-Math YouTube Active Academy | 日本教育工学会 日本科学教育学会 KAKEN 研究課題... 科研費電子申請シス... 国・地域毎の2019年...

実施結果のポイント

不正解群で特に心配なパターン



The graph displays a blue line on a light blue grid. The line starts at a point on the left, slopes downwards to a local minimum, then remains relatively flat with a very slight upward trend towards the right, ending with a small, sharp peak before dropping slightly.



ここに入力して検索

17:11 2022/12/17


RIMS共同研究2022冬「Computer」

ファイル | C:/Users/masat/Desktop/数理研22冬/RIMS2022winter_presen/index.html#/7/9

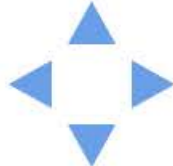
Yahoo! JAPAN 受信トレイ (500) - m... 東邦大学 eラーニング... Cindy-Math YouTube Active Academy | 日本教育工学会 日本科学教育学会 KAKEN 研究課題... 科研費電子申請シ... 国・地域毎の2019年...

実施結果のポイント

不正解群で特に心配なパターン



The graph displays a blue line on a light blue grid. The line starts at a baseline, has a small dip, then rises to a sharp peak, followed by a sharp dip below the baseline, and finally returns to the baseline. This pattern is characteristic of a specific data distribution or error profile.



ここに入力して検索

17:11 2022/12/17

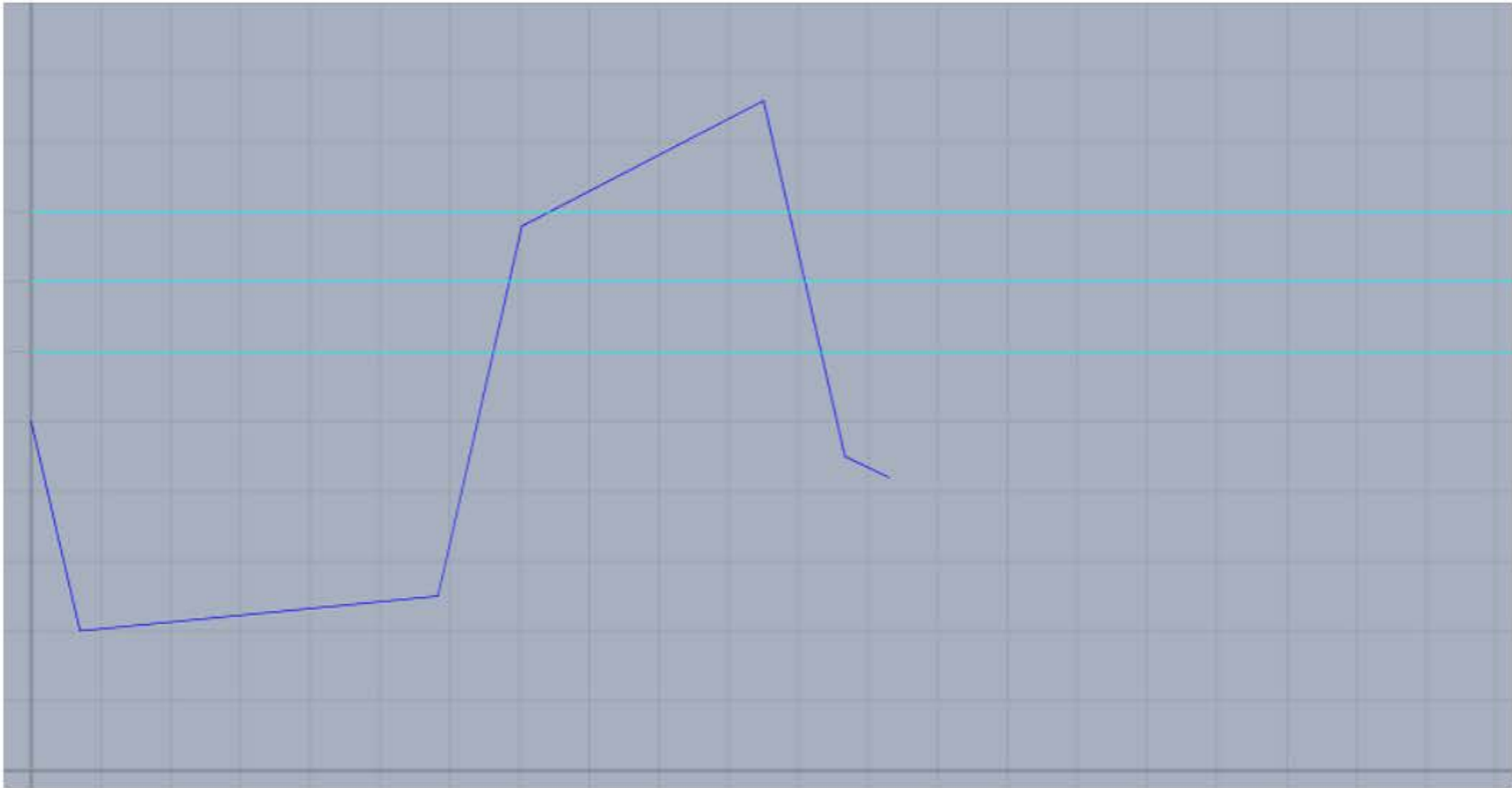
RIMS共同研究2022冬「Computer」

ファイル | C:/Users/masat/Desktop/数理研22冬/RIMS2022winter_presen/index.html#/7/10

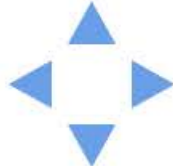
Yahoo! JAPAN 受信トレイ (500) - m... 東邦大学 eラーニング... Cindy-Math YouTube Active Academy | 日本教育工学会 日本科学教育学会 KAKEN 研究課題... 科研費電子申請シス... 国・地域毎の2019年...

実施結果のポイント

不正解群で特に心配なパターン



The graph displays a blue line on a light blue grid. The line starts at a low point on the left, rises to a higher point, reaches a peak, and then drops sharply to a low point on the right. This pattern likely represents a distribution of incorrect answers.



ここに入力して検索

17:12 2022/12/17

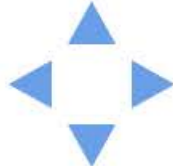

RIMS共同研究2022冬「Computer」

ファイル | C:/Users/masat/Desktop/数理研22冬/RIMS2022winter_presen/index.html#/7/11

Yahoo! JAPAN 受信トレイ (500) - m... 東邦大学 eラーニング... Cindy-Math YouTube Active Academy | 日本教育工学会 日本科学教育学会 KAKEN 研究課題... 科研費電子申請シ... 国・地域毎の2019年...

実施結果のポイント

不正解群で特に心配なパターン



ここに入力して検索

7°C 17:12 2022/12/17

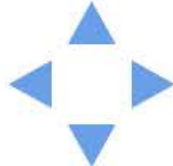
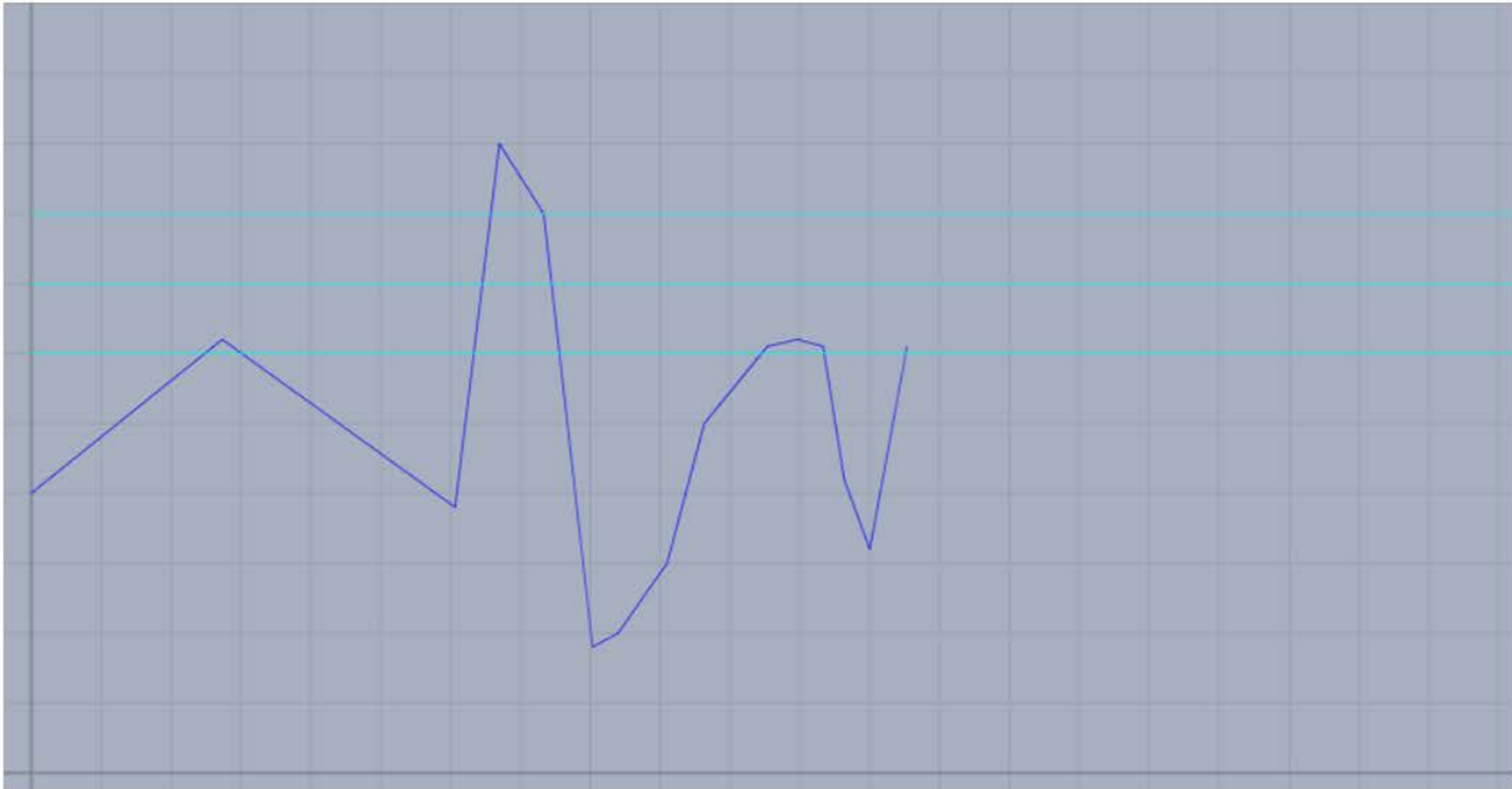
RIMS共同研究2022冬「Computer」

ファイル | C:/Users/masat/Desktop/数理研22冬/RIMS2022winter_presen/index.html#/7/12

Yahoo! JAPAN 受信トレイ (500) - m... 東邦大学 eラーニング... Cindy-Math YouTube Active Academy | 日本教育工学会 日本科学教育学会 KAKEN 研究課題... 科研費電子申請シ... 国・地域毎の2019年...

実施結果のポイント

部分正解群でも...



ここに入力して検索

7°C 17:12 2022/12/17


RIMS共同研究2022冬「Computer」

ファイル | C:/Users/masat/Desktop/数理研22冬/RIMS2022winter_presen/index.html#/7/13


Yahoo! JAPAN 受信トレイ (500) - m... 東邦大学 eラーニング... Cindy-Math YouTube Active Academy | 日本教育工学会 日本科学教育学会 KAKEN 研究課題... 科研費電子申請シス... 国・地域毎の2019年...

実施結果のポイント

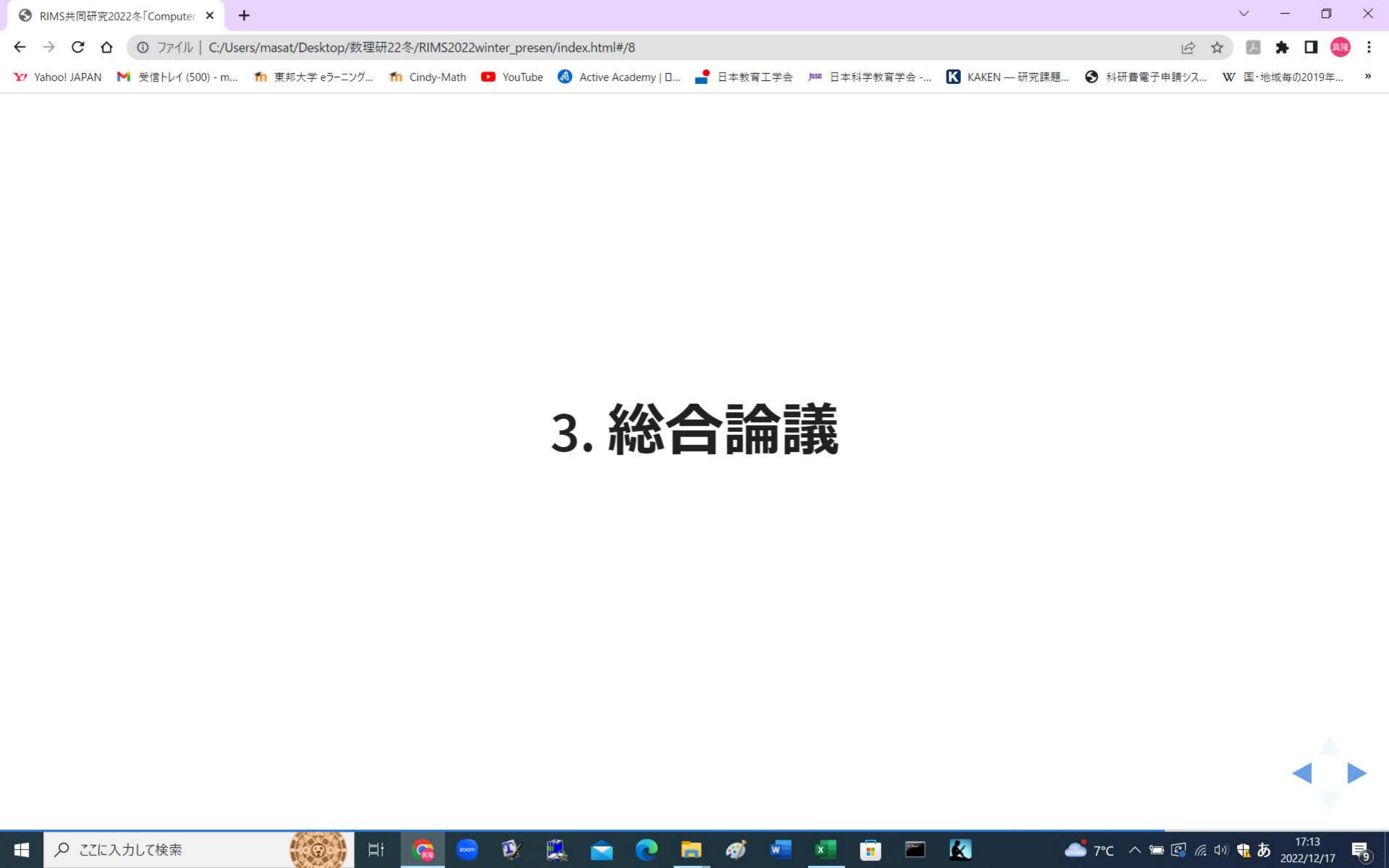
部分正解群でも...



The graph displays a blue line on a grid. The line starts at a low value on the left, rises to a moderate peak, dips slightly, and then rises sharply to a higher peak on the right before dropping to zero. The grid has vertical lines every 1 unit and horizontal lines every 2 units.



Windows taskbar: 7°C, 17:12, 2022/12/17



3. 総合論議



Diagnostics

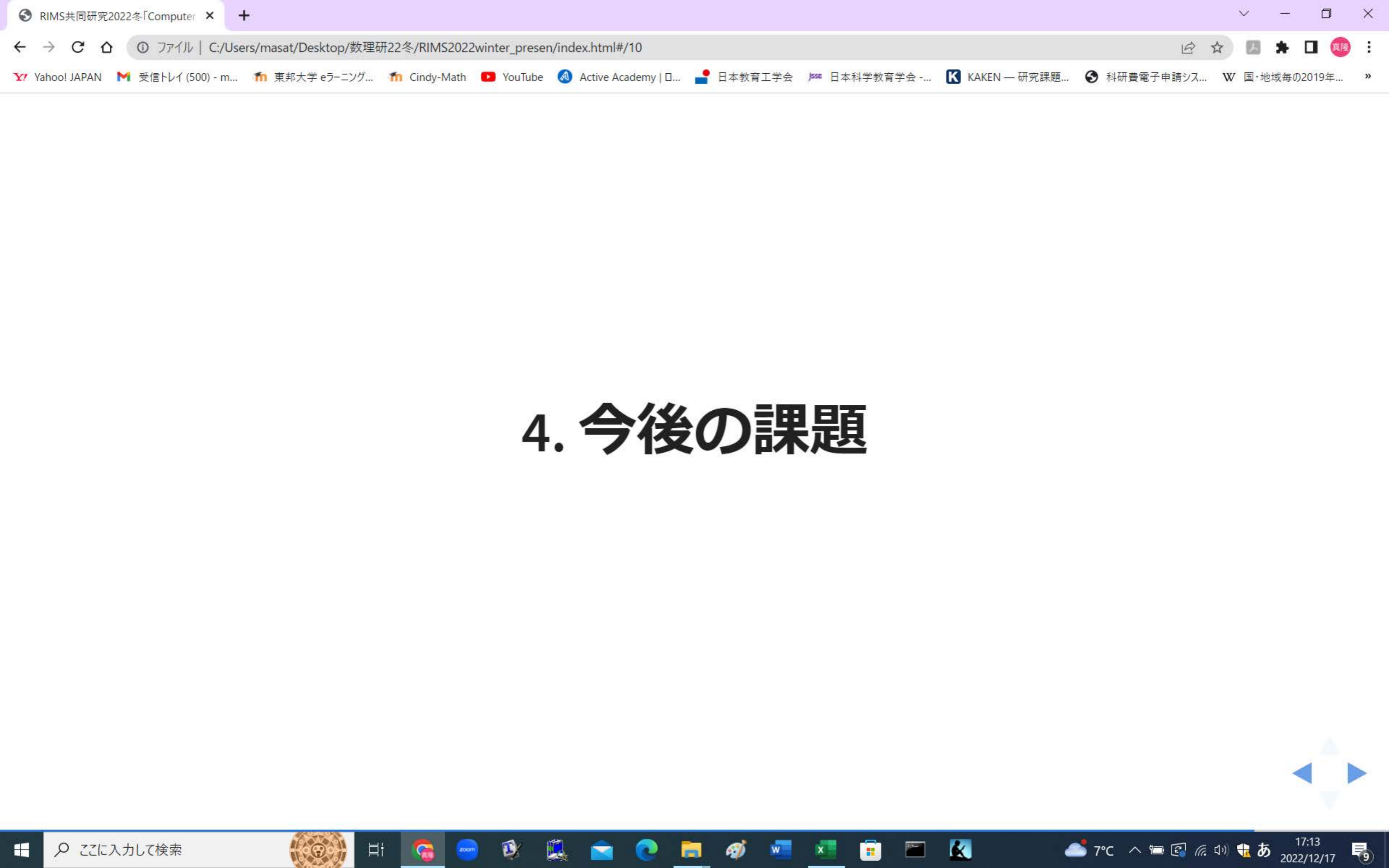
- 本来的に、様々なケースが試行されることが望ましいことを前提としつつも、やはり教育的介入が求められると想定されるケースとしては・・・

単位時間あたりの入力値の変動が大きく、入力値の変更と検定統計量の増減との対応が正確に認識されているか疑わしい

入力値の変動自体は小さいものの、指定された p 値が実現されるのとは別のレンジが探索されており、入力値の変更と p 値の変動の関連性が正確に認識できているか疑わしい

- 探索の自由度が大きいと、データ分析・implicationの抽出はそれだけ難しくなる。Diagnosticsと探索の自由度のどちらをとるか？





4. 今後の課題



今後の課題

- 探索設定の変更

コントロール群と比べ、薬剤投与群で治癒例の入力値が小さくなるのはナンセンスであることを予め教授し、入力値のレンジに制限を加える

それでも幅広い探索が行われる可能性がある一方、事前介入に束縛される事例も想定され、操作ログの評価が難しくなる可能性がある

- タスク設計の変更

小数値の入力事例が見られたので、整数値に限定させるか？



今後の課題

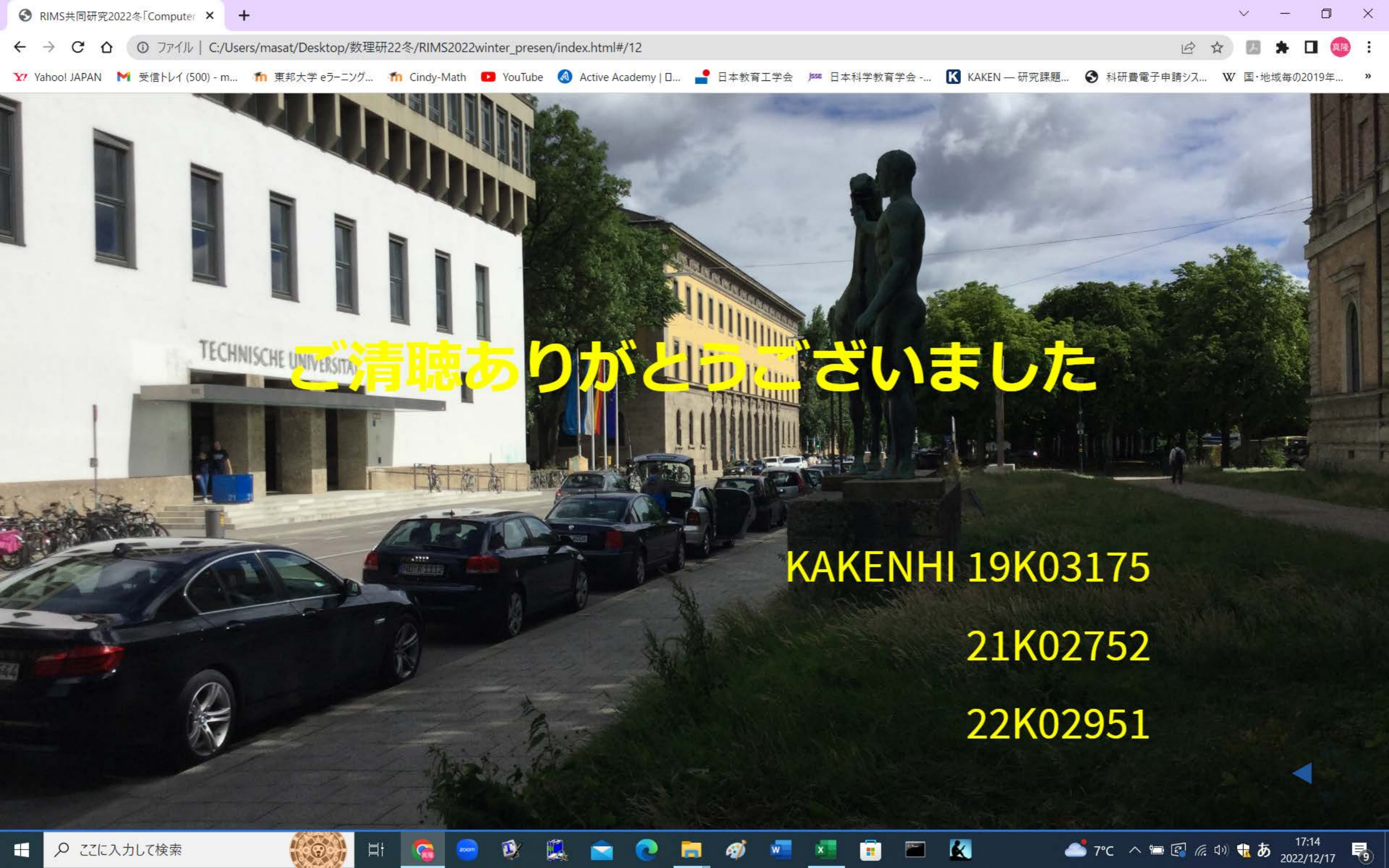
- Delayed post test の検討

検定統計量の意味に関する理解を試すために、コントロール群の数值をさらに少し変えて、指定された p 値を実現するような薬剤投与群の数值を予想してもらう

- 探索プロセスを少し制限することによって、ログデータの傾向がどうなるか？ 特に、Diagnostics のためのシグナルを抽出できないか？

ログデータから post test スコアをある程度予想？





ご清聴ありがとうございました

KAKENHI 19K03175

21K02752

22K02951