

そ 0.3 以下であり、統計的に値が小さいことがわかった。

3.3 相関分析のまとめ

今回の分類毎の単回帰分析では、いずれも決定係数が低く、回帰モデルで解答所要時間を推定できる精度はない。また、全ての穴抜きの構造に対して二変数間である程度の相関があるという結果から、条件を課す、または、解答所要時間、正解率のほかに回帰分析の変数を追加することで実用性のある回帰モデルが作成できる可能性がある。

4. 解答所要時間の分析

4.1 分析の目的

解答所要時間と正解率の相関関係はそれほど強くないことが分かった。そこで解答所要時間のみに注目し、どのような傾向があるかを分析する。分析を行うことで穴抜きと問題の解答所要時間を確認し、穴抜きについての所要時間の考察を行う。

4.2 で解答所要時間の分布が既存のどの分布に近いかを評価し、4.3 で穴抜きの模範解答をトークン数で分け、解答が他のトークン数に比べて異なるような穴抜きを調査している。

4.2 解答所要時間の分布

学生が個別の穴抜きに要する解答所要時間の分布を図 5 に示す。データ総数は 1001 である。x 軸は解答所要時間、y 軸は相対出現度を表している。また、所要時間の階級は 5 秒間隔で集計している。ヒストグラム上のオレンジの線はカーネル密度推定を用いて境界面の設定によらない分布を作図している。また、解答の出現頻度が最も高くなるのは 0~10 秒の間であった。

また、分布についてシャピロ-ウィルク検定、コルモゴロフ-スミルノフ検定を用いて正規性の検定を行った。その結果、どちらも正規分布に従わないという結果になった。解答所要時間の分布にトークン数で分布を分け正規性について調べたが、どのトークン数の場合でも正規分布にはならなかった。

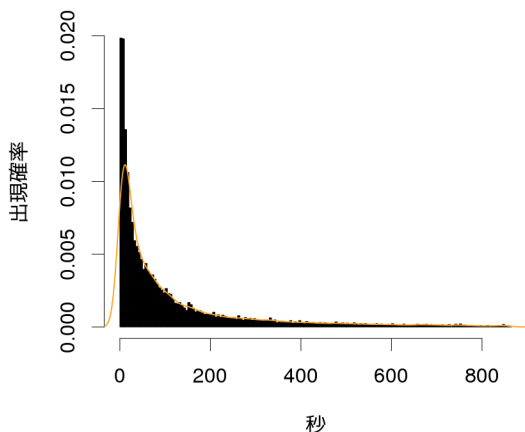


図 4：解答所要時間のヒストグラム

4.3 分布の評価

穴抜きの解答所要時間がどのような分布にしたがっているかを AIC (Akaike's Information Criteria) を用いて評価した。その結果、AIC 値は、正規分布の場合 -2851.648、二項分布の場合 6.009477、ポアソン分布の場合 5055.581 であった。従って、この中で最もモデルの当てはまりが良いものは二項分布である。

4.4 トークン数と解答所要時間の相関分析

穴抜き毎の解答所要時間を対象として分析する。

分類：最小データ数 210

- トークン毎：トークン数 1~12 (ただし、トークン数が 9 と 11 の問題は定義されていない)

分析方法

本研究ではトークン数毎の関係を定量的に検定するために t 検定を採用した。t 検定を用いて総当たりでトークン数毎の有意差を調べている。t 検定の前提条件となる正規性については分布の評価より、解答所要時間の分布は二項分布に近く正規分布に従わないことが分かっている。また、トークン毎に分けた解答所要時間の分布に対してシャピロウィルク検定を用いてすべてのトークン数で正規分布に従わないことを確認している。

トークン数毎の解答所要時間の分散については F 検定を用いてトークン数のすべての組み合わせで等分散であることを確認している。

表 3 ではトークン数毎の解答所要時間の有意差の有無を示す。○は p 値が 5%未満のため有意差が確認できた。一方、×のセルは p 値が 5%以上のため、有意差がないと判断した。

表 3：トークン数毎の解答所要時間の有意差検定

| | | | | | | | | | |
|----|---|---|---|---|---|---|---|---|---|
| 1 | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| 2 | × | × | × | ○ | × | × | × | × | ○ |
| 3 | ○ | × | ○ | ○ | × | ○ | ○ | ○ | ○ |
| 4 | × | ○ | × | ○ | × | × | × | × | ○ |
| 5 | ○ | ○ | × | × | × | × | × | × | × |
| 6 | × | ○ | × | × | × | × | × | × | ○ |
| 7 | ○ | ○ | × | × | × | × | × | × | ○ |
| 8 | ○ | ○ | × | × | × | × | × | × | × |
| 10 | ○ | ○ | × | × | × | × | × | × | × |
| 12 | ○ | ○ | × | × | × | × | × | × | × |

有意差検定の結果からおおむねトークン数の差が広がると解答所要時間に差が出ることが分かった。しかし、トークン数が離れた場合でも有意差がないと判定された事例がある。これは正解トークン数が多い穴抜きでも、慣用表現 (イディオム) などの影響により、解答所要時間が短い場合があるためだと考えられる。

5. 穴抜きの分類

5.1 分析の目的

pgtracer では学生の学習ログデータを収集しているが、教員がすべての学習ログデータを閲覧し、分析するには手間がかかる。そこで各ユーザーが持つ学習ログデータの分類を行う。分類をすることで解答に特徴のある穴抜きを抽出することを目的とする。

対象データおよび評価尺度

- 穴抜き (78 個)
- 平均解答所要時間(初回時・最高得点時)
- 平均正解率(初回時・最高得点時)
- 平均解答回数(初回時・最高得点時)

分析方法

問題毎の分類でも学生の分類と同様に、ワード法を用いたクラスタ分析が最も均等にクラスタの形成が行われた。図5では4つのクラスタが形成される様子を樹形図で示す。

問題毎の分類分析結果

クラスタの作成を定量的に評価するためにコーフェン相関係数を調べた結果、距離行列とコーフェン相関係数の歪みが少ない方法はユークリッド距離を用いた重心法だったが、その方法を用いるとクラスタ毎の要素数が不均等だったが、そのため本研究では樹形図により、グループ化の状況を視覚的に判断し、各クラスタの要素数が均等に近く、最もクラスタの形成が上手くいっていたキャンベラ距離を用いたワード法を採用した。

枠線は各クラスタの分類を示している。また、クラスタ番号を1~4とすると、データ数はクラスタ1が16個、クラスタ2が20個、クラスタ3が28個、クラスタ4が20個だった。

樹形図で把握した各クラスタについて変数である解答所要時間、正解率の間にどのような違いがあるのかを箱ひげ図を用いて調べた(図6、図7)。クラスタ間の有意差検定にデータ数が少ないこと解答所要時間の分布が正規分布に従わないことがわかっているので、ウィルコクソン順位和検定を採用した。

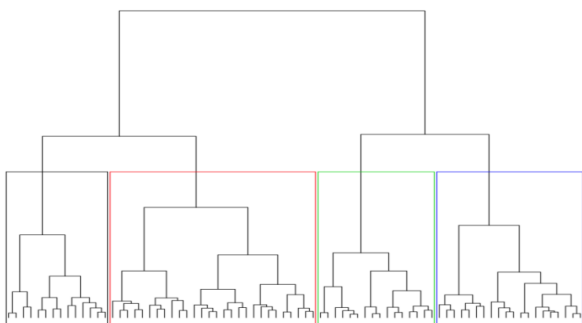


図5：ワード法でのクラスタ形成

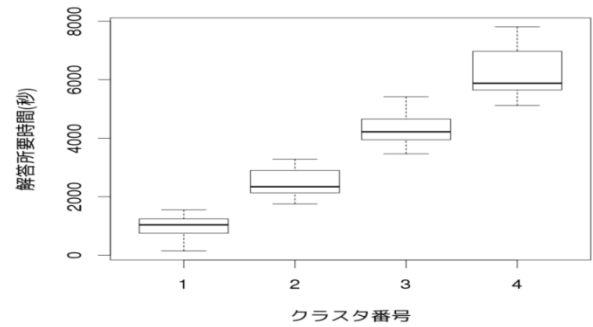


図6：クラスタ毎の解答所要時間分布

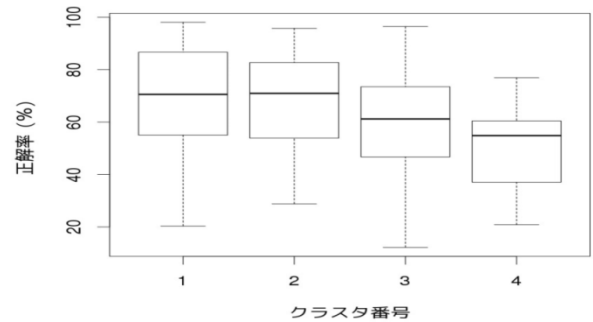


図7：クラスタ毎の正解率分布

クラスタ1では解答所要時間が短い穴抜きを分類できた。また、各クラスタの初回解答所要時間の違いについてウィルコクソンの順位和検定を有意水準5%で行った結果、すべてのクラスタの組み合わせで有意差がみられた。

正解率の箱ひげ図ではクラスタ2から3,4になるにつれて正解率が低下する傾向が見られた。また、各クラスタの正解率についても同様にウィルコクソンの順位和検定で検定した結果、クラスタ1と2の間には有意差がなく、クラスタ1と3、クラスタ2と3の間では有意差がみられた。また、初回解答回数についての有意差検定では正解率と同様のクラスタの間で有意差がみられた。

分析結果の考察

図5で得られた分類を元にそれぞれのクラスタについて穴抜きを観察し、クラスタの傾向を調べた。表4ではクラスタ毎に具体的な穴抜きを3つ挙げ、初回解答時の解答データを示し、プログラム中の穴抜きを表現するため、穴抜きの前後を抜き出し、対象の穴抜きを四角で括弧で表現している。また、具体例を挙げるにあたって初回解答所要時間が平均に近いものをサンプリングした。穴抜きの傾向を調べる方法としてサンプリングした穴抜きの最も重複する誤答を調べた。

クラスタ1

表7の穴番号1では穴抜きの正解が「right=n-1;」に対し、解答に「;」をつけ忘れていた間違いが最も多く、誤答全体で21回の内4回を記録していた。また、穴番号2,3についても正解に近いような解答をしていることが分かった。このことからクラスタ1では比較的難易度の低い穴抜きが所

| クラス番号 | 穴番号 | 問題 | プログラムの穴抜き |
|-------|-----|---------|---|
| 1 | 1 | 二分探索 | //4.探索間を配列の全域とする left=0; right = n - 1; |
| | 2 | 二分探索 | if(x < array [mid]){ right = mid - 1; } |
| | 3 | 二分探索 | cout<<"探索位置"<< right <<endl; |
| 2 | 4 | ハッシュ関数 | if(storeResult== -1){\ |
| | 5 | ヒープソート | for(i==0;i<N;i++){ insert (array [i]);} |
| | 6 | 括弧の対応検査 | if(paren == rightParenTable [i]){\ |
| 3 | 7 | ハッシュ関数 | //質問データを検索し、結果を表示する searchResult = search (x); |
| | 8 | ヒープソート | //1.データ x をヒープの最後尾に追加する m=n; heap [n] = x ; |
| | 9 | 括弧の対応検査 | while(*p!= NULL &&isParen (*p ==0){ |
| 4 | 10 | ハッシュ関数 | while(hashTable [i] != 0){\ |
| | 11 | ヒープソート | while(heap[(m - 1) / 2] < x){ |
| | 12 | ヒープソート | heap[n-1]=0; n -- ; yPos=0; |

表4：クラス毎の穴抜きの具体例

属していると考察する。

クラス2

穴番号4では穴抜きの正解が「-1」に対し、「0」と解答したもののがもっとも多く、誤答全体で16回の内、10回を記録した。また、穴抜き番号5では最も多かった誤答が「heap[i]=insert(array[i]);」となっており、問題は理解しているが文の前後関係などの見落としがあるなどの傾向がみられた。このことからクラス2ではクラス1ほど単純ではないが、比較的學生が理解できている穴抜きが所属していると考察する。

クラス3

穴番号7については最も多い誤答が「search(x);」であった。これは誤答全体で28回の内、8回あるため、関数の戻り値をsearchResultに格納する必要があることに気付いていない、複雑なトレース表を正確に追えていないなどが考えられる。

クラス4

穴抜き番号10では誤答の解答は多いものの重複する誤答が少ない傾向がみられる。また、穴番号11では誤答全体で38回の内、「m」と解答したものが11回見られ、重複する誤答の数も多くみられた。このことから、クラス4では學生が理解できなかった穴抜きや、間違いやすい傾向にある穴抜き、穴抜き自体は単純だが、トレース表で数値を辿るのに時間が掛かる穴抜きが所属していると考察する。

5.2 穴抜きの分類分析のまとめ

穴抜きの分類では各クラスでどのような違いがあるかを観察した。その結果、各クラス間の解答所要時間においてウィクソクソソ順位と検定より、全ての組み合わせ

で有意差が見られた。また、クラス1から4になるにつれ解答所要時間が大きくなっており、正解率は徐々に小さくなっていった。また、この分類は初回解答時と最高得点時の学習データを用いたが、初回時の学習データからもほぼ同様のクラスが形成されたため、最高得点時の学習データは穴抜きの分類にはあまり寄与していないと考えられる。

また、クラス毎に重複する誤答について調べ、各クラスの傾向を考察した。クラス1,2では比較的単純な穴抜きが分類され、クラス3,4でアルゴリズムやトレース表といった複数の要素の理解ができなければ解けない穴抜きが分類されていると考えられる。また、クラス4では學生が理解できなかった穴抜きや決まった誤答に誘導されている穴抜き、穴抜き自体は単純だがトレース表を辿るのに時間が掛かる穴抜きの3つが分類されていると考察した。

6. 學生の分類

6.1 分析の目的

学習ログデータを學生の解答毎に分類する。この分類を行うことで解答に有意差のある學生を分類する。また、分類を調べることにより、同じ分類に所属する學生の得意箇所、不得意箇所を把握することを目指している。

対象データ

- 學生 (72人)
- 問題にかけた平均所要時間
- 問題への解答回数
- 問題の平均点

分析方法

學生毎の学習ログデータにクラス分析を行った。クラ

スタ分析ではユークリッド距離を用いた最遠距離法、ユークリッド距離を用いた重心法、キャンベラ距離を用いたウォード法の3つの場合で分析を行った。

学生の分類分析結果

樹形図による視覚化

図8では3つの大きなクラスタが形成されている様子を示している。クラスタの作成を定量的に評価するためにコーフェン相関係数を調べた結果、距離行列とコーフェン相関係数の歪みが少ない方法はユークリッド距離を用いた重心法だったが、その方法を用いると、小さなクラスタがいくつかできてしまい、うまく分類できていなかった。そのため本研究では樹形図により視覚的に判断し、最もクラスタの形成が上手くいっていたキャンベラ距離を用いたウォード法を採用した。また、クラスタ番号を1~3とすると、それぞれのデータ数は24、25、23であった。

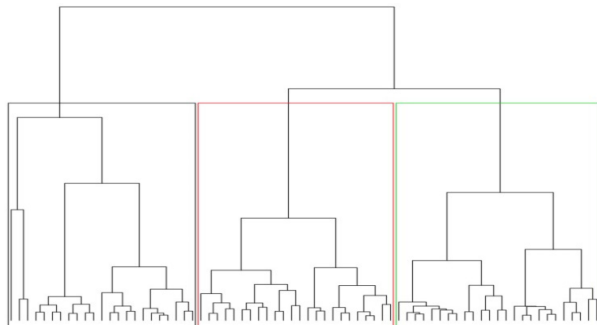


図8：ウォード法でのクラスタ形成

クラスタの詳細

図9では樹形図で示したクラスタの平均所要時間、解答回数、平均点の詳細を箱ひげ図を用いて視覚化を行った。この図では各クラスタ間の平均所要時間の差は少なくとも200秒以上あることが確認できる。

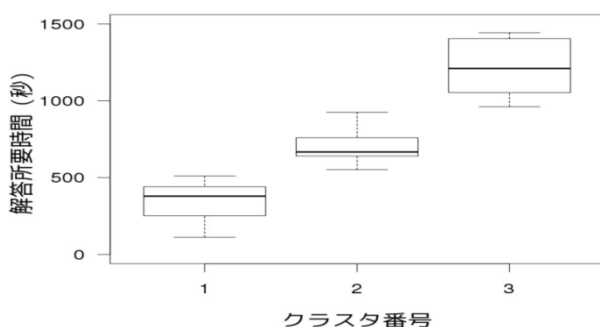


図9：クラスタ毎の解答所要時間の分布

クラスタ毎の解答所要時間、解答回数、平均点について違いを定量的に評価するためにウィルコクソンの順位和検定を行った。また、中央値はクラスタ1が397.5、クラスタ2が667、クラスタ3が1210.455であった。表5では各クラスタの解答所要時間についてウィルコクソンの順位和検定を用いて定量的に評価した結果を示す。検定結果は有意水準5%未満である。

表5：ウィルコクソンの順位和検定結果（解答所要時間）

| クラスタ番号 | 2 | 3 |
|--------|---|---|
| 1 | ○ | ○ |
| 2 | - | ○ |

解答所要時間のみがそれぞれのクラスタ間で有意差があるという結果になった。

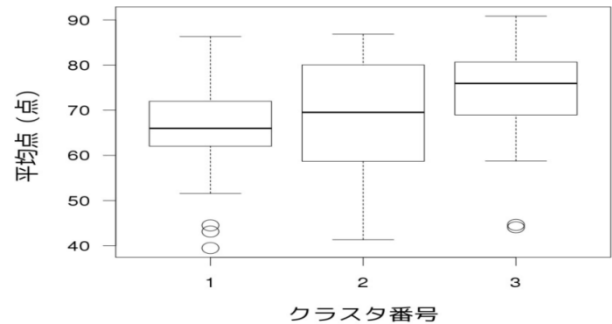


図10：クラスタ毎の平均点の分布

各クラスタの平均点の分布は図10に示すようになっており、中央値はクラスタ1で66.0、クラスタ2で69.5、クラスタ3で76.0であった。

表6：ウィルコクソンの順位和検定結果（平均点）

| クラスタ番号 | 2 | 3 |
|--------|---|---|
| 1 | × | ○ |
| 2 | - | × |

ウィルコクソンの順位和検定の結果（表6）ではクラスタ1とクラスタ3の間でのみ有意差がみられた。そのほか二つでは有意差は見られなかった。

各クラスタの解答回数の分布は図のようになっており、中央値は、クラスタ1で12.5、クラスタ2で15、クラスタ3で17であった。また、ウィルコクソンの順位和検定の結果、クラスタ間の差は見られなかった。

6.2 分析のまとめ

学生の分類について解答所要時間、平均点、解答回数の三要素でクラスタ分析を行った。その結果、各クラスタ間で有意差がみられる要素は解答所要時間のみであった。従って、今回の学習ログデータを学生という観点で分類する場合、解答所要時間に基づいて分類することが最も適当であると考えられる。

7. 学生の解答過程の分析

6節で分類した各クラスタに属する学生の解答過程について観察する。学生の分類では解答所要時間でクラスタ間の有意差がみられた。そこで解答所要時間を代表値とし、実際に異なる学生で同じ問題の解答過程を調べ、解答所要

| クラス 番号 | 二分探索 | | ハッシュ関数 | | ヒープソート | | 括弧の対応検査 | |
|-----------|-------|-------------------|--------|-------------------|--------|-------------------|---------|-------------------|
| | 平均点 | 解答所要 時間 (秒) | 平均点 | 解答所要 時間 (秒) | 平均点 | 解答所要 時間 (秒) | 平均点 | 解答所要 時間 (秒) |
| 1 | 64.22 | 335.15 | 57.13 | 683.87 | 41.38 | 514.67 | 38.14 | 543.05 |
| 2 | 73.56 | 441.44 | 71.43 | 1283.57 | 49.57 | 1099.25 | 59.56 | 1248.89 |
| 3 | 75.98 | 630.31 | 75.02 | 2292.70 | 63.13 | 2232.42 | 65.70 | 1832.82 |

時間の差が生じる理由について考察する。

なお、pgtracer に登録されている7つの問題のうち、数表1、数表2、時刻変換の3つは学生の能力を測る意図はなく、pgtracer の操作修得のために作られた基本問題であるため、具体例としては挙げていない。

クラスターの具体的な考察

学生の解答過程を考察するにあたり、同じ問題について、クラスターにより異なる学生の振る舞いを分析する。

表7には各クラスターにおける問題別の平均点(100点満点)と解答所要時間の平均値を示す。クラスター1では、4つの問題について、解答所要時間には大きな違いはない。一方、クラスター2および3では問題毎の解答所要時間にはより大きな違いがある。特にクラスター番号3の解答所要時間は、他のクラスターと比較して長い、平均点もより高いことが、どの問題についても言える。

我々は二分探索問題、ハッシュ関数、ヒープソートについてクラスターから代表的な学生を選び解答過程を調べた。その結果、どの問題についても同様の傾向が見られた。

クラスター1の中でサンプリングした学生は、冒頭の穴抜きを埋めるために時間が掛かるが、他の穴抜きは比較的短時間で解いている学生、特定の穴抜きに時間をかけているが、それ以外では短時間で解答している学生が所属しており、基本的に順調に解答を行った学生が所属していた。しかし、表7を参照すると、どの問題に対しても他のクラスターよりも所要時間が短く、正解率も低くなっている。そのため、クラスター1には解ける問題は解答するが、理解できない問題には考える時間を費やさない学生が多く含まれていると考えられる。また、クラスター1では解答を行っていない(問題を見ただけの)学生、一部のみ解答を行っている学生など、特に解答所要時間が少ない学生が多く所属している。このクラスターに属する学生に対しては、重点的なケアが必要と考えられる。

クラスター2の中でサンプリングした学生は、ある程度の時間をかけて解答しているが、正解率は幅広く分布しており、一般的な学生が多く含まれていると考えられる。

一方、クラスター3の中でサンプリングした学生は問題初めの穴抜きを含め、特定の穴抜きに時間をかけている箇所が多かった。また、二分探索問題では探索の範囲指定部分で解答に時間をかけている学生が多くみられた。クラスター

3では理解できない問題に対しても積極的に取り組む学生が所属していると考えられる。特にクラスター3の学生には解答が正解になるまで複数回答していた事例が見受けられ、このことが解答所要時間と正解率が他のクラスターよりも高い原因だと考えられる。

8. おわりに

本論文ではプログラミング教育支援ツール pgtracer を用いて収集した穴抜き問題に対する学習ログデータを用い、Rを用いたクラスター分析と、各クラスターに関する考察を行った。穴抜きの解答所要時間と正解率は、必ずしも相関関係にないことが分かった。また、解答所要時間はトークン数に依存する傾向があることが検定結果より分かった。

学習データの分類では穴抜きと学生の2種類のデータに対してクラスター分析を行った。穴抜きの分類では全てのクラスターの解答所要時間と特定のクラスターの正解率について有意差のある穴抜きを分類できた。また、穴抜きの傾向を調べ、各クラスターに所属する穴抜きの特徴を考察した。

また、学生の分類では解答所要時間に有意差のあるものを検出し、分類した学生の解答過程を分析した。

今後の研究課題としては、本研究で抽出した各クラスターの学生に対して、学習態度やプログラミング能力を考慮した穴埋め問題を作成する方法に関する検討が挙げられる。

謝辞：本研究は JSPS 科研費 17K01036 の支援を受けています。

参考文献

- [1] 掛下哲郎, 柳田峻, 太田康介, “穴埋め問題を用いたプログラミング教育支援ツール pgtracer の開発と評価”, 情報処理学会論文誌: 教育とコンピュータ, Vol. 2, No. 2, pp. 20-36, 2016.
- [2] 柳田峻, “穴埋め問題を用いたプログラミング教育支援ツール pgtracer”, 佐賀大学工学系研究科知能情報システム学専攻修士論文, 2015年3月.
- [3] Tetsuro Kakeshita, Kosuke Ohta, Student Log Analysis Functions for Web-Based Programming Education Support Tool pgtracer, 17th International Conference on Information Integration and Web-based Applications & Services (iiWAS2015), pp. 120-128, December 2015.
- [4] 太田康介, “プログラミング教育支援ツール pgtracer におけるログデータ分析機能の開発と学生の理解度分析”, 佐賀大学大学院工学研究科知能情報システム学専攻修士論文, 2016年3月
- [5] 金明哲, “Rによるデータサイエンス(データ解析の基礎から最新の手法まで)”, 森北出版, 2017.