

円グラフと帯グラフの量的把握の比較検討

隅谷 孝洋^{1,a)} 坂下 潤一郎²

概要：比率を表す二つのグラフ表現である円グラフと帯グラフについて、その量的把握、すなわち表している比率の読み取りや、大小関係の判定について800名の実験を行った。結果として、比率の読み取り精度については円グラフのほうが帯グラフよりもやや高いが、処理時間は帯グラフのほうが短い。また、大小関係の判定については、帯グラフのほうが精度が高く処理時間も短かった。また、円グラフでは角度で認識しているということが通常言われているが、日常見慣れているアナログ時計の読み取りとも深く関係しているかもしれないということが示唆された。

キーワード：円グラフ、帯グラフ、クラウドソーシング、認知実験、グラフ表現、データ可視化、データ科学教育

Comparative study on quantitative perception of pie charts and band charts

TAKAHIRO SUMIYA^{1,a)} JUN'ICHIRO SAKASHITA²

Abstract: We conducted an experiment with 800 participants to understand the quantitative perception of two graphical representations that express ratios: pie charts and band chart. This involved interpreting the ratios they represent and discerning their relative magnitudes. The results revealed that the accuracy of interpreting ratios was slightly higher in pie charts than in band charts. However, the processing time was shorter for band charts. Furthermore, in terms of understanding relative sizes, band charts showed higher accuracy and shorter processing times. Although it is commonly said that we recognize pie charts through angles, our findings suggest a deep connection with how we're used to reading analog clocks in our daily lives.

Keywords: pie chart, band chart, crowdsourcing, cognitive experiment, graphical representation, data visualization, data science education

1. はじめに

データ科学に対する興味が高まり、全国的な教育が期待されている。本稿でテーマとしている「グラフ表現」は、統計学的な基礎知識と並んでデータ科学の基礎になるものと考えられる。グラフ表現の中でも円グラフは頻繁に用いられ、日常生活でも非常に多く目にするものになっている。

一方、Rで円グラフを描く関数 `pie` のヘルプを開くと、

「円グラフは、情報を表示する方法としては最悪だ」と書かれている。理由は、人間は（長さなどの）線形な尺度であれば比較がしやすいが、面積の比較は難しいから^{*1}、ということだ。棒グラフやドットプロットを使った方が良いとも書いてある。

確かに、形が異なるものの面積を比較するのはかなり難しい。しかし、棒グラフが長さで数量を表すと同時に面積でも表しているのと同様に、円グラフは本質的には角度で大きさを表しているのであり、人はこれを読むときに面積ではなく、角度で認識している。Rのヘルプの元になっている Cleveland も、1995年では円グラフは角度での認識で

^{*1} The eye is good at judging linear measures and bad at judging relative areas.

¹ 広島大学 情報メディア教育研究センター
Hiroshima University, Information Media Center

² 広島大学大学院 先進理工系科学研究科
Hiroshima University, Graduate School of Advanced Science and Engineering

^{a)} sumiya@hiroshima-u.ac.jp

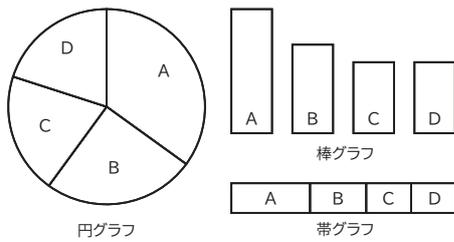


図 1 円グラフと代替

あると述べている。

いずれにせよ円グラフは very bad way であり、限定的な利用にとどめるべきとされている ([1], [2], [3])。しかしそれにもかかわらず、冒頭に述べたように円グラフは巷に溢れている。3D 円グラフをはじめとする誤用もたっぷりある [4]。そんなにも読み取りにくいものをなぜ使うのか。

ところで、棒グラフやドットプロットは数量の大小を比較する点では円グラフにまさるが、比率を示すことには向かない。図 1 のように、円グラフでは 1 番目の A のセグメントが全体の 1/3 ほどで、B までを加えると半数を超えることが容易に読み取れるが、棒グラフではそうではない。棒グラフは割合を表示する方法としては不適切なのだ。

では割合はどう表すべきかという点、帯グラフを用いるのが良いとされている。帯グラフは角度ではなく長さを使って表示するので。しかし、このグラフは、経年変化など複数の割合を比較する用途には用いられるが、その他の場合にはやはり円グラフが利用されることが多い。

円グラフの方が帯グラフよりも「見栄えが良い」のはそののだろうか、実際の読み取り性能にはどの程度の差があるのだろうか。この研究では、この両者の比較を実験的におこなった。

2. 先行研究

認知心理学やデータ可視化の研究においては、さまざまな調査がある。Cleveland と McGill [5] は、R の Help の文言の元となった研究を行なっている。彼らが position-angle と呼ぶ対応で、棒グラフと円グラフを比較、position-length と呼ぶ対応で、棒グラフと積み上げ棒グラフ（帯グラフと類似）を比較している。

Spence ら [6] は、円グラフ、帯グラフ、棒グラフにおいて、セグメント同士もしくはセグメントを合算したものの同士で大小関係を認識させる実験を行なっている。円グラフはセグメントを合算させた場合に読み取り精度が良かったと報告している。

Skau [7] らは円グラフを構成する要素を細かく分割し、角度だけや弧だけなどの場合において、認知がどのように変わるかを調査している。

なお、本研究は 2020 年度に広島大学情報科学部に提出された著者の一人の卒業研究 [8] を発展させたものである。

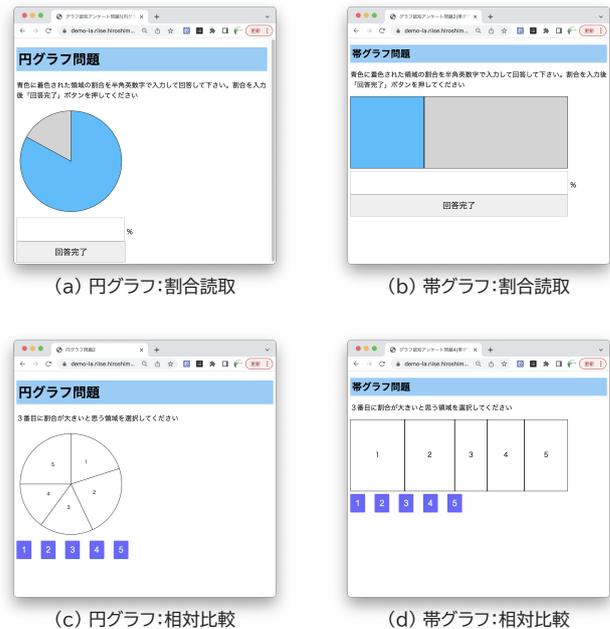


図 2 実験用 Web ページ

卒業研究においては、Google Form を用いて 29 名の学生に対して円グラフと帯グラフについての認知実験をおこなった。その際の主な仮説は以下のようなものだった。

まず円グラフでは、読み取る扇形の領域の中心角が 90° 、 180° 、 270° (25%, 50%, 75%) となる時は直角か直線となり、加えて時計の $3 \cdot 6 \cdot 9$ 時部分にあたり日常的に見慣れていることも踏まえて、比較的認知がしやすく、誤差が少ないのではないかと考える。一方帯グラフに関しては、箱型の図形を 2 等分するのは比較的容易であるため、対応する 50% 付近では誤差が少ないと考えられるが、ほかの割合、例えば 25% では円グラフは中心角が 90° なため認知しやすいが、帯グラフでは箱を 4 分割しなければならず、個人差が生じやすく誤差が大きくなると考える。[8]

いくつかの問題があり、この実証には至らなかったが後述するように本研究ではこれらの点も確認ができています。

3. 実験

3.1 調査項目

円グラフと帯グラフの読み取り性能について、以下の 2 点を調べる。

- (1) 割合読取：あるセグメントの割合を読み取る
 - (2) 相対比較：セグメント間の大小関係の判定をする
- まず (1) を調べるため、図 2-(a), (b) に示すようなグラフを提示し、割合を整数値でキーボード入力するようにした。そして、(2) を調べるため、図 2-(c), (d) に示すようなグラフを提示し「3 番目に大きい箇所」を選択するようにし



図3 一人分の実験データ

た*2。

卒業研究 [8] では、(2) の実験では円グラフと帯グラフともほとんどが正解で、差が出なかった。同じように読み取れた場合は、時間がかからない方が認知的な負荷が少ないと言えるのではないかと考え、今回の実験では (1) も (2) も一つ一つの解答にかかる時間を測るようにした。

3.2 実験用 Web ページ

今回の実験にあたり、実験用の Web ページを PHP と Javascript を用いて自作した。アクセスすると以下のような動作を行い、被験者が回答したデータをサーバに保存する。

- (1) 実験前の簡単なアンケートを提示
- (2) 2セグメントの円グラフを提示し、最初の部分（着色して表示）の割合を整数値で入力。これを10回繰り返す
- (3) 2セグメントの帯グラフを提示し、最初の部分（着色して表示）の割合を整数値で入力。これを10回繰り返す
- (4) 5セグメントの円グラフを提示し、3番目に大きいところを指定。これを5回繰り返す
- (5) 5セグメントの帯グラフを提示し、3番目に大きいところを指定。これを5回繰り返す
- (6) 実験後の簡単なアンケートを提示
- (7) 完了コードを提示

ここで、(2) と (3) では、10組の乱数を生成し、円グラフと帯グラフでは同じ組の乱数がそれぞれ異なる順で出題される。(4) と (5) では、あらかじめ作成した5組の数値の組み合わせから出題した。これも (2), (3) と同様、円グラ

*2 3番目を選択するようにしているのが重要で、卒論 [8] (p.23) では5つのうち「4番目に大きい箇所」と指定したところ、著者の意図していた「小さい方から4番目」を選んだ被験者よりも「大きい方から4番目」を選んだ被験者の方が相当多かった。なので、どちらに解釈しても同じ結果となる「3番目」とした

フと帯グラフで同じ組み合わせが異なる順序で出題されるようにした。

また、出題の順序が結果に影響することがないように、(2) と (3) の提示順序はアクセスする毎に入れ替わるようにした。(4) と (5) も同様である。

実験前と実験後のアンケート内容は以下の通りである。

(1) 実験前のアンケート

- 性別
- 年代
- 割合を表すのによく利用するのは、円グラフか帯グラフか（例を示す）
- グラフの領域が示す割合を読み取りやすいと思うのは、円グラフか帯グラフか（例を示す）

(2) 実験後のアンケート

- 割合を読み取りやすかったのは、円グラフか帯グラフか
- 順序を読み取りやすかったのは、円グラフか帯グラフか

まとめると、図3のように6件のアンケート回答を含むメタデータと、一人当たり30件の回答データとが得られることになる。

3.3 実験の実施

できるだけ多様なひとびとからデータを取得したいため、本実験ではクラウドソーシングを利用した。具体的には、「クラウドワークス*3」を利用した。

クラウドワークスは、登録した「クラウドワーカー」に仕事を依頼するためのサービスである。今回のような調査目的には、「タスク形式（特定の一人と契約するのではなく、多くの人から作業を集める）」として作業内容を提示し、一件あたりの作業単価と、集めたい件数、締め切り日時などを指定することになる（図4）。アンケート調査サービスではないので、男女比や年代の指定はできない。

あるワーカーがタスクをこなして完了させた場合、依頼者は承認作業をおこない、承認されたものに対して報酬が



図4 CrowdWorksでの「クラウドワーカー」募集画面

*3 CrowdWorks <https://crowdworks.jp/>

表 1 被験者の男女構成

男性	女性	その他
362	435	2

表 2 被験者の年代構成

30 歳未満	30-39 歳	40-49 歳	50-59 歳	60 歳以上
107	293	240	119	40

支払われるという仕組みになっている。

承認作業のため、白木ら [9] の図 2 にあるように、実験用 Web ページ (3.2) に完了コードを表示する機能を持たせ、被験者にはクラウドワークスのアンケート機能で完了コードを提出してもらった。提出された完了コードとサーバに記録されたものを突き合わせれば、ある実験結果を出したクラウドワーカーが特定できる。

実験結果の信頼性を担保するため、「努力の最小限化」をしたものを検出する方法がいくつも提案されている ([9], [10])。今回は回答が連続して同一など不誠実回答の基準を立ててチェックしたが、記録された実験結果で不承認とすべきものはなかった。一件のみ、実験結果の記録なしに正しくない完了コードを提出していたものを不承認とした。

承認作業に関連するデータチェックについては、事前に想定していなかった落とし穴があった。クラウドワークスのシステムの仕様上、提出された完了コードは「承認後」でないと一括ダウンロードができないのだ。20 件ずつを画面で確認することができるだけだった。当初は一括ダウンロードして、提出された完了コードと実験用サーバで発行済みの完了コードを突き合わせる予定だったが、それができなくなった。そのため、画面上での目視確認できる範囲内で完了コードの確認作業をおこなった。承認後一括ダウンロードした完了コードを確認したところ、不正なものは 1 件もなかった。

2023 年 5 月 23 日 7:00 に最大 800 件と設定して作業依頼を公開^{*4}したところ、5 時間弱で設定件数が達成された。後日内容を確認したところ前述の不承認が 1 件出たが、不承認作業後すぐに追加の作業が提出され、800 件の実験結果を得ることができた。

実験用 Web ページでは各ページの回答を提出しないと先に進めないようにしていたため、完了コードが発行されているものについては、すべての回答を提出していることになる。しかし、なんらかの不具合で完了コードが発行されているにもかかわらず実験データに欠損があったものが 1 件だけ存在した。これは被験者の責任範囲外の問題であると考えられるため承認したが、データ分析からは除外することとした。

よって、今回得られた実験結果は 799 件である。

4. 実験結果

799 件の被験者の男女比、年齢構成は表 1, 2 のとおりである。

表 3 割合読取の結果 (全てのデータ)

	N	誤差	平方誤差	回答時間
円グラフ	7990	-0.66 (5.8)	34.3 (290.6)	12.7 (14.8)
帯グラフ	7990	0.21 (4.5)	20.1 (196.7)	9.11 (9.85)

表 4 割合読取の結果 (| 誤差 | < 15)

	N	誤差	平方誤差	回答時間
円グラフ	7801	-0.22 (3.4)	11.5 (25.4)	12.8 (14.8)
帯グラフ	7947	0.26 (3.5)	12.5 (22.0)	9.10 (9.86)

4.1 割合読取

図 5 に、割合読取の全てのデータをプロットした。横軸が実際の割合、縦軸が被験者が読み取った割合である。左が円グラフの結果で右が帯グラフの結果になっている。“reading actual numbers from a pie chart is next to impossible[11]”とも言われているが、「意外と読めている」というのが第一印象だった。平方誤差 (実際の値と被験者が読み取った値の差の 2 乗) と、反応時間の平均を表 3 に示す。括弧内の数は標準偏差である。

読み取り誤差は円グラフのほうが大きく、かかった時間も円グラフのほうが長い。

図 5 の円グラフの方をみると、右上がりの対角線の下に結構な数の点からできている線が見えるだろう。これらは $y = \frac{60}{100}x$ の直線に乗っている。円グラフで示された割合をパーセントで読み取るタスクなのだが、一周を時計の分表示の 60 とみなして読み取っているデータと思われる。75%を 45「分」という具合に回答してしまっているのだ。これを便宜的に「時計読み」と呼ぼう。

左下りの対角線に薄く乗っているのは、支持された方とは逆のセグメントを読んでいるものである。95%なのを 5%と読んでしまう。これを便宜的に「裏読み」と呼ぼう。「裏読み」がいくらかあることは卒業研究時にわかっていたため、今回はそれを避けるように説明書きや出題の際の表示に注意を払った。

「時計読み」に関しては想定していなかったのだが、結果的に 150 から 200 件 (2%前後) がこれに相当すると思われた。正答になっている割合と読み取り誤差の関係などについても検証したいため、逆に読むものも含め、こういうものについては外れ値として取り扱いたい。

表 3 より、円グラフと帯グラフの読取誤差の標準偏差は 5.8 と 4.5 であるので、読み取り誤差の絶対値が 15 (概ね標準偏差の 3 倍) 以上のものを外れ値として扱うこととした。外れ値を除いた場合の散布図を図 6 に示す。「時計読み」は少し残っているがほぼ排除されていると思われる。

^{*4} <https://crowdworks.jp/public/jobs/9155449>

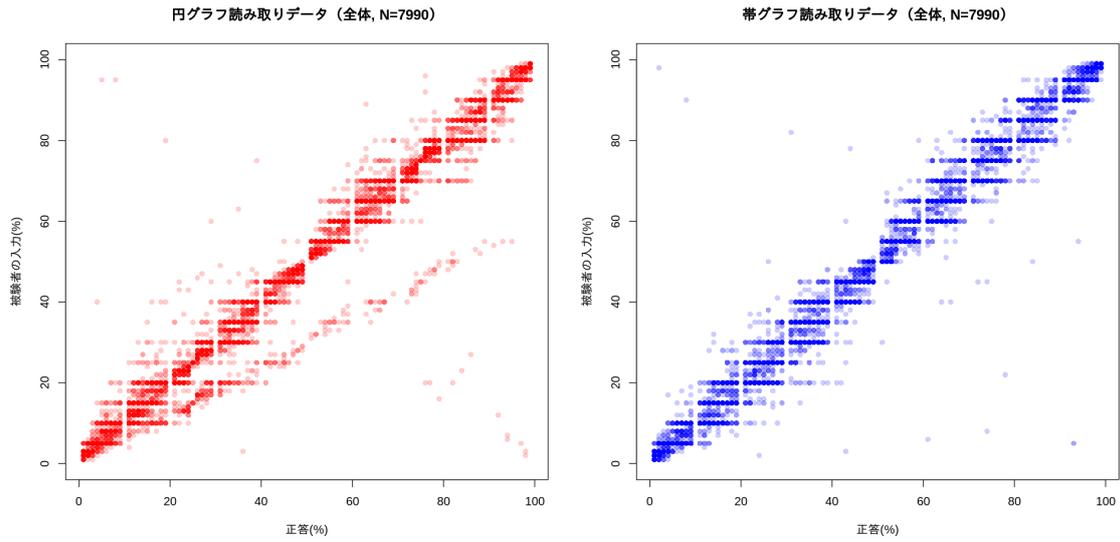


図 5 割合読取の結果 (全ての結果)

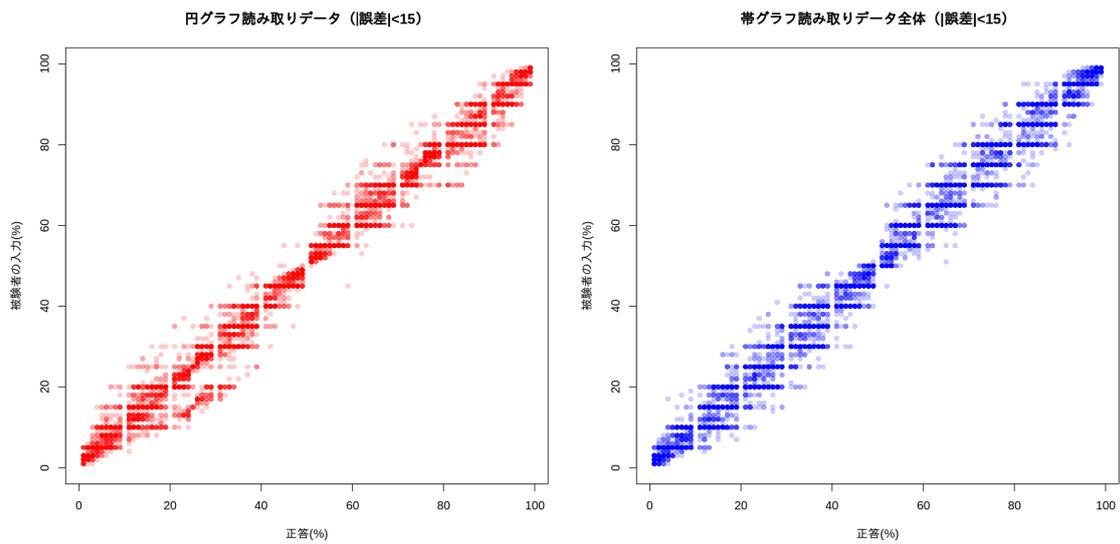


図 6 割合読取の結果 ($|\text{誤差}| < 15$)

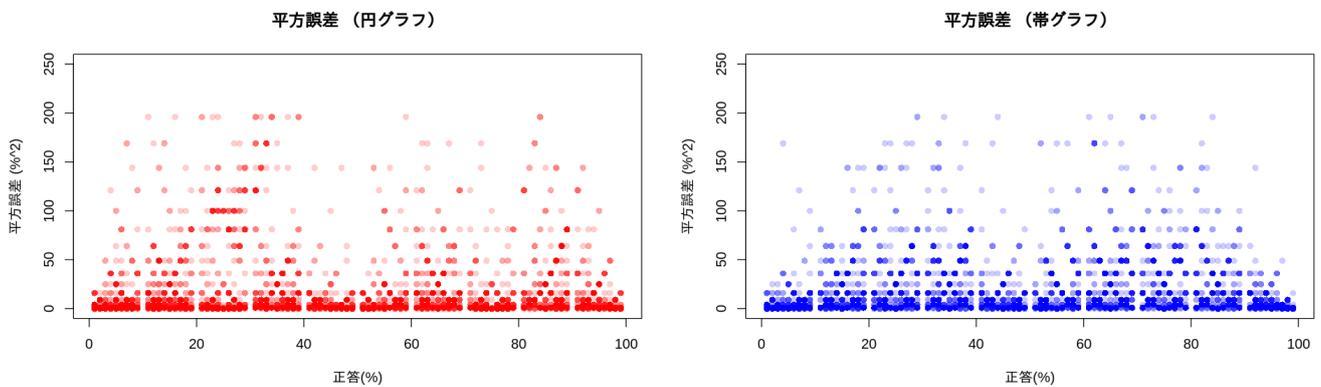
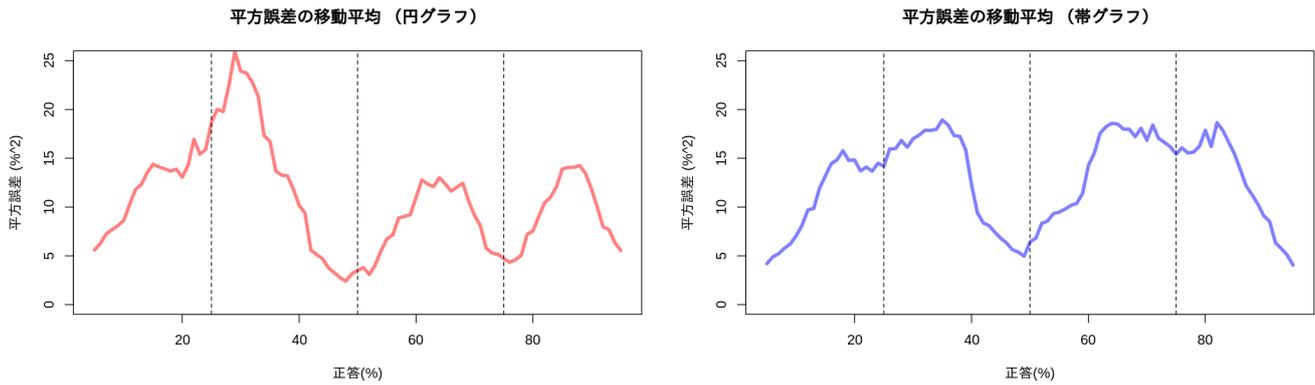
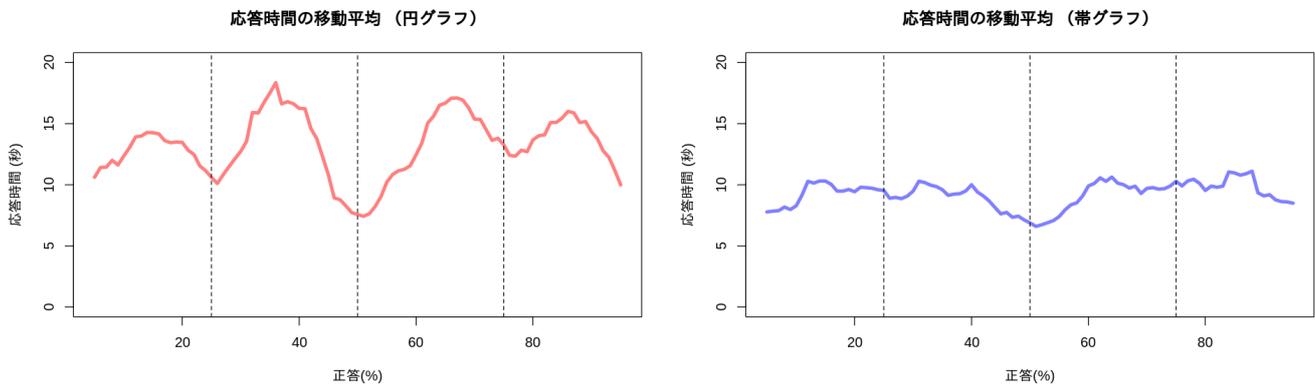


図 7 平方誤差プロット ($|\text{誤差}| < 15$)

図 8 平方誤差の移動平均 ($|\text{誤差}| < 15$)図 9 応答時間の移動平均 ($|\text{誤差}| < 15$)

この場合の誤差などの平均、標準偏差は、表 4 の通りである。円グラフのほうが「時計読み」や「裏読み」がより多く発生するということを記憶に留めつつ、これ以降の記述については、このデータを用いることにする。

表 4 では、読取誤差は円グラフのほうが僅かながら帯グラフよりも少ない結果になっている。統計的有意差もある(ウィルコクソンの順位和検定, $p < 2.2 \times 10^{-16}$)。一方、回答時間は帯グラフのほうが有意に短い(同, $p < 2.2 \times 10^{-16}$)。

表示されている割合(正答)と読み取り誤差の関係を調べるため、横軸に正答をとり、縦軸に読み取り誤差の 2 乗(平方誤差)をプロットしたものが図 7 である。これでは傾向がわかりにくいため、正答を時間に見立てて移動平均をとったのが図 8 である。1 ずつずらしながら、 ± 2 以内で平方誤差の平均をとってプロットした。これを見ると、円グラフが正答 50%と 75%で平方誤差が半分以下になっているのがわかる。25%でもこのような減少があると予想していたが、排除しきれしていない「時計読み」の影響でそこは見えなくなっている。一方帯グラフの方では、25%と 75%でやや減少しているが、非常に限定的である。

また、同様に応答時間にたいして移動平均を求めたのが図 9 である。円グラフでは 25%付近での平方誤差の著しい増加に関わらず、応答時間は想定通り減少しているのが

読み取れる。また、全体的に帯グラフの方が円グラフよりも応答時間が短いことがわかる。

4.2 相対比較

円グラフまたは帯グラフの「3 番目に大きいセグメントを回答する」問題は、セグメント同士の相対的な大小関係を判定する必要がある。離れたところにあるものの比較も含めて正確に判定できなければ正解することはできないはずだ。しかし、このタスクについては、卒業研究の時点で、多くの人がほぼ正解を出すことができることがわかっていた。そのため、今回はページ遷移してから回答するまでの時間をミリ秒単位で測定するようにした。応答時間も測定したい意図は、作業開始の段階で被験者には伝えてある。

このタスクで得られた実験データは、799 人のそれぞれに対して円グラフ 5 件合計 3995 件と帯グラフ 5 件合計 3995 件となる。それぞれのうち正解は 3276 件 (82.0%) と 3622 件 (90.7%) だった。個人ごとに処理時間の平均をとり、それを横に円グラフ、縦に帯グラフの散布図として表示すると図 10 のようになる。対角線を引くと一見して帯グラフの方が処理時間が短いものが多いのがわかる。ウィルコクソンの符合順位和検定を行うと $p < 2.2 \times 10^{-16}$ で正解率、処理時間とも有意差が認められた。

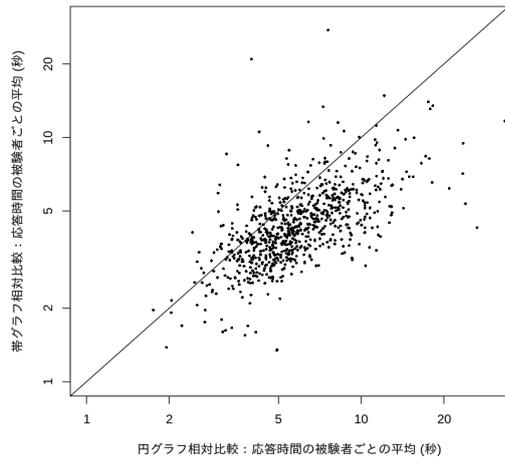


図 10 相対比較に要した応答時間

5. 考察

本研究では円グラフと帯グラフについて、割合の読み取りと相対的な大小判定においてそれぞれの優位性、特質を測るためにクラウドソーシングを活用した実験をおこなった。

その結果、割合読み取りに対しては、円グラフの方が帯グラフよりも精度がやや高いということが示された。これは直感に反するかもしれない。ひとつ理由として考えられるのは、通常帯グラフには値を読み取るための目盛りがついているが、本実験ではそれがまったくない状態で読み取らなくてはならなかったことがある。円グラフでも目盛りがないことには変わらないが、これは通常目盛りなしで目にする。それでも読み取れるのは何故かという、角度を直接読み取れる訳ではなく、「アナログ時計の読み取り」が慣れた処理として存在するからではないかと推察される。時計の文字盤が極めて見慣れたものとして存在し、円グラフの読み取りの時にも「見えない目盛り」のように働いているのではないかという仮説である。

パーセントで読み取ってくださいと作業を依頼しているのに、「時計読み」をしてしまう人が少なからずいたこと、さらに読み取りの処理時間が帯グラフよりも円グラフの方が有意に長いことは、その仮説を支持していると考えられないだろうか。無意識に分で読み、そのあとパーセントに「換算処理」をしているので時間がかかる。

また、相対比較のタスクでは、精度・処理時間とも帯グラフが優っていた。数値に変換することのない単純な大小判定だと、R の Help でいう「直線的な測度」の優位性がそのまま効いてくるのかもしれない。

2 節で示した、25%、50%、75%仮説は、極めて常識的と思われるが卒業研究時にはデータの的には確認することができなかった。本研究のデータでは、応答時間も含めて考察

すれば概ねデータの的に支持されたと考えて良い。

これらのことを踏まえて、データ表現に取り組む、また教育すると、よりよい結果が得られるのではないだろうか。

6. おわりに

本稿では、今回の実験データに対して円グラフと帯グラフの比較を主な分析対象として考えた。実験データには、回答者の属性や簡易アンケートなど、本稿で扱うことができなかったものが多く含まれている。これらについては、また別の機会を設け、あらためて分析に取り組んでいきたい。

また、前節「考察」において、円グラフとアナログ時計の関係についての見解を述べたが、今後これの裏付けとなるような実験もおこなっていきたいと考えている。

謝辞

広島大学大学院人間社会科学研究所平川真先生には、クラウドソーシングを活用した認知実験について、多くのご教示をいただきました。ここに感謝の意を表明いたします。ありがとうございました。

参考文献

- [1] Cleveland, W. S.: The Elements of Graphing Data, Wadsworth: Monterey, CA, USA. (1985)
- [2] Few, S., & Edge, P.: Save the pies for dessert, Visual business intelligence newsletter, 1-14. (2007)
- [3] 永田ゆかり：データ可視化のデザイン, SB クリエイティブ株式会社 (2020).
- [4] けんけん：あなたの知らない「詐欺グラフ」の世界 (随時更新中), 入手先 <<https://note.com/kenxxxken/n/nce7762dcec30>> (参照 2023-06-17).
- [5] Cleveland, W. S., & McGill, R.: Graphical perception: Theory, experimentation, and application to the development of graphical methods, Journal of the American statistical association, 79(387), 531-554. (1984)
- [6] Spence, I., & Lewandowsky, S: Displaying proportions and percentages, Applied Cognitive Psychology, Vol.5, No. 1, pp.61-77. (1991)
- [7] Skau, D., & Kosara, R.: Arcs, angles, or areas: Individual data encodings in pie and donut charts, In Computer Graphics Forum, Vol. 35, No. 3, pp. 121-130, (2016)
- [8] 坂下 潤一郎：グラフ表現における認知バイアスの検証 -円グラフ・帯グラフの読み取りに関するアンケート調査 -, 令和 3 年度広島大学情報科学部卒業論文, 入手先 <<https://demo-la.riise.hiroshima-u.ac.jp/archive/sakasita2021.pdf>> (参照 2023-06-17) (2021).
- [9] 白木 優馬, 五十嵐 祐：クラウドソーシングを利用したアンケートデータ収集のノウハウと課題, デジタルプラクティス, Vol.9, No.4, (2018).
- [10] Oppenheimer, D. M., Meyvis, T., & Davidenko, N: Instructional manipulation checks: Detecting satisficing to increase statistical power. Journal of experimental social psychology, Vol.45, No.4, pp.867-872. (2009)
- [11] Robert Kosara: Understanding Pie Charts, 入手先 <<https://eagereyes.org/techniques/pie-charts>> (参照 2023-06-17)