

2019PCカンファレンス

【セミナー1】

CIEC会誌

『コンピュータ&エデュケーション』

に採択されるために

－新しい投稿カテゴリーが期待するもの－

●パネリスト

横川博一（神戸大学／CIEC会誌編集委員長）

武沢 護（早稲田大学大学院・高等学院／CIEC会誌編集委員）

村上正行（大阪大学／CIEC会誌編集委員）

寺尾 敦（青山学院大学／CIEC会誌編集委員）

CIEC会誌

『コンピュータ&エデュケーション』 に採択されるために

－新しい投稿カテゴリーが期待するもの－

2019PCC SEMINAR

企画 会誌編集委員会



CIEC会誌『コンピュータ&エデュケーション』 に採択されるために －新しい投稿カテゴリーが期待するもの－

パネリスト

横川 博一（神戸大学）
武沢 譲（早稲田大学大学院・高等学院）
村上 正行（大阪大学）
寺尾 敦（青山学院大学）

概 要

会誌では投稿論文の種別を「研究論文」に加え、
「実践論文」「研究ノート」「実践報告」を
新たに設けました。これらがどんなものを求めているのか、
採択の成否を分けるもの、質を高める統計手法など
について検討します。

最近の会誌：特集・巻頭インタビュー

号 (発行年月)	特集／巻頭インタビュー
Vol.42 (2017.06)	[特集] これからのICT教育環境基盤を探る [巻頭] AI 搭載ロボットで世界基準の外国語教育とグローバル人材育成を推進する／田中良一さん
Vol.43 (2017.12)	[特集] 学習の動機づけとICT利用教育 [巻頭] 子どもたちの可能性のある未来のために－「教育」と「ICT」をつなぐイノベーターとして／川合睦さん
Vol.44 (2018.06)	[特集] ITでアーカイブと教育は変わる [巻頭] 数理・データサイエンス教育のゆくえ／斎藤政彦さん
Vol.45 (2018.12)	[特集] AI 時代における教育と学習について考える [巻頭] 新会長に聞く CIEC－過去・現在・そして未来／若林靖永さん
Vol.46 (2019.06)	[特集] ロボットを活用した教育実践と可能性 [巻頭] 分身ロボット Orihime が語るもの／吉藤健太郎さん

最近(過去 5 年)の投稿数・採択率

号	投稿数 (特集除く)	採択数	不採択	採択率
Vol.38	13	8	5	62%
Vol.39	10	7	3	70%
Vol.40	8	7	1	88%
Vol.41	8	3	5	38%
Vol.42	11	5	6	45%
Vol.43	13	10	3	77%
Vol.44	17	11	6	65%
Vol.45	25	15	10	60%
Vol.46	15	10	5	67%

査読方法

- 論文，特集など，すべてについて2名が査読
 - （1）編集委員以外の理事に査読を依頼することがあり得る。
この場合でも主査は編集委員。
 - （2）特集：2名のうち1名は特集担当編集委員
 - （3）不採択後の再投稿：査読者を変更

投稿カテゴリー

本誌は、とくにコンピュータやネットワークを活用した教育・研究に関する理論的・実践的な原稿を歓迎します。

改正前

論文

図表を入れて刷り上がり 6 ページ以内

活用事例

図表を入れて刷り上がり 4 ページ

論文および事例研究

図表を入れて刷り上がり 6 ページ以内

改正後 **NEW!**

研究論文・実践論文

コンピュータやネットワークを活用した教育・研究に関する理論的・実践的研究。
図表を入れて 6 ページ以内

研究ノート・実践報告

図表を入れて 4 ページ以内

ソフトウェアレビュー

自作のものを含む公開されたソフトウェアについて、その機能、特徴および教育・学習への有用性を会員に広く紹介することを目的としたもの。図表を入れて 2 ページ以内

採択の一般的基準

- 新規性も重要ですが..
- **研究論文**：当該のテーマ，設定された課題について検証済みであり，同じことを検証する必要がない，あるいは同じことをすれば同じ効果が約束されている。
- **実践論文**：対象とする母集団に対して適用した方法の効果が明確である。

採択されるために：ストーリー性

その研究・実践のオモシロさを伝える

Introduction

- ① Background
- ② Issue/gap
- ③ Purpose

- ▶ どんな背景や問題意識があるのか？
 - ▶ この研究・実践にはどんな意味・意義があるのかを伝える。
- ▶ これまでに何が明らかになっていて、何が明らかになっていないのか？
 - ▶ 何を問題にするのかを明確に伝える。
 - ▶ この研究・実践が明らかにしようとすることを明確に伝える。
- ▶ どんな方法でそれを明らかにしようとするのか？
 - ▶ 方法論の妥当性とゴールを伝える。

採択されるために：ストーリー性

その研究・実践のオモシロさを伝える

Body

①Participants

②Method

③Results

- ▶ 誰を対象にしたのか？
 - ▶ この研究・実践の対象を，必要な情報とともに明確に伝える。
- ▶ どんな方法を使ったのか？
 - ▶ 何この研究・実践の方法（手続きや授業展開，効果の検証の方法など）を，正確かつ十分な情報量で伝える。
 - ▶ その方法はこの研究・実践の目的に合っているかどうか伝わるよう記述する。
- ▶ どんな結果が得られたのか？
 - ▶ 目的に合った結果の示し方をする。
 - ▶ そこに「思い込み」はないか？
 - ▶ 「都合のよい」データだけを示していないか？

採択されるために：ストーリー性

その研究・実践のオモシロさを伝える

例

3. 結果と考察

まず、映画を使って、授業外でタスクを行いながら学習したことについて、アンケート調査（5段階評価、5が「大いに思う」）を行った結果を表1に示す。

質問番号	評価平均値	標準偏差
Q1	4.3	0.56
Q2	4.8	0.76
Q3	3.9	1.21
Q4	4.2	0.45
Q5	4.5	0.78

また、タスクに関する自由記述回答を一部抜粋する。

- タスクがあったので、映画を見ようという気が湧きました。
- 映画の英語はむずかしかったけど、タスクで聞き取りのポイントが分かったので、理解しやすかったです。

- ▶ 教育の現場—授業や学生—を研究のフィールドとすることのむずかしさ
- ▶ 授業は「うまくいけば」、それでいい?!
- ▶ ICTを活用した授業の効果をどう検証するか?
- ▶ どんな授業を行ったか？それがポイント！
- ▶ そこに「思い込み」はないか？
 - ▶ 何らかの形で「客観的」なデータを示す
- ▶ 都合のいいデータになっていないか？
 - ▶ 多方面から分析・考察する
 - ▶ 評価値が低い学習者のコメントは？
 - ▶ 結果を一般化する

採択されるために：ストーリー性

その研究・実践のオモシロさを伝える

Conclusion

- ① Discussion
- ② Conclusion
- ③ Further study

- ▶ 得られた結果は何を示唆しているか？
 - ▶ 研究・実践の目的に照らして、結果が示唆することを客観的に伝える。
 - ▶ 今回得られた結果を、先行研究をふまえて、できるだけいろいろな角度から考察する。
- ▶ 研究・実践の目的・方法・結果，ふたたび
 - ▶ 研究・実践の目的に対する解答，結果が示唆すること，残されている課題をまとめて伝える。

採択されるために：ストーリー性に加えて 論文をオモシロクするために

- 教育の改善が主目的であって、コンピュータの利用は手段に過ぎないことを忘れていないでしょうか？
- グラフィックス等を多用することで学生・生徒に評判が良かったことをもって授業効果が上がったと誤解していないでしょうか？
- コンピュータの利用を優先するあまり、授業本来の目的・目標やシラバスで約束したことを逸脱していないでしょうか？
- それまでの授業の問題点は解決されたでしょうか？また、授業は改善されたでしょうか？

抄録(Abstract) の充実を 論文のイントロではなくサマリー

例

The Reading Span Test (RST) can provide insights into the efficient functioning of the Working Memory (WM) and how participants allocate their finite WM resources to the processing and storage functions. **However, the original version of RST does not clarify** the types of processing that exhaust participants' WM resources because various types of processing functions (i.e., orthographical, phonological, syntactic, semantic, and pragmatic processing) are employed during the course of RST.

Background of research

Therefore, **the first purpose of this study is to specify** the type of processing that places a burden on the WM of the Japanese EFL learners and makes the maintenance of previously processed information more difficult. **It further explores** the relationship between RST scores and the English proficiency among the Japanese EFL learners. **We employed the following four types of RSTs:** a standard version (RST 1), an RST with a Japanese-translation verification task (RST 2), an RST with a grammatical judgment task (RST 3), and an RST with a semantic judgment task (RST 4). **The following two types of English proficiency tests were also administered:** a grammar test and a reading comprehension test. Forty-four Japanese EFL learners completed these four RSTs and two English proficiency tests.

Research question

Research method

The results showed that the span scores of the RST 3 were significantly lower than those of the other RSTs, irrespective of the scoring methods. In addition, the span scores of the RST 2 and RST 3 correlated significantly with both grammar and reading test scores, irrespective of the scoring methods. **These results suggest that** syntactic processing places a heavy cognitive load on WM resources, thereby leaving few resources for storage functions while the Japanese EFL learners are performing RSTs. **They also suggest that** their English proficiency largely depends on the degree of automatization of syntactic processing.

Results and discussion

魅力的なタイトルを タイトルは研究論文の最大のサマリー

- 必要な情報を最大限盛り込んで、魅力的なタイトルを！

例

**ワーキングメモリに対する認知的処理負荷の影響
ー日本人英語学習者を対象としたリーディングスパン
テストによる検討ー**

**The effect of processing cognitive load on working
memory: Performance of the Japanese EFL learners
in reading span tests**

- 何についての研究か？（どんな分野の研究か？）
- 誰を対象にした研究か？
- どんな方法や手法を用いた研究か？

有益な論考の投稿を歓迎します

そして、投稿する前に

- 音読してみてください
 - 誤字脱字はありませんか？
 - 意味不明の表現・文章はありませんか？
- 英語のネイティブ・チェックを受けてください
- 共著者みんなが目を通してください
 - 論考の内容は、共有され、コンセンサスが得られたものですか？
 - 複数の目で十分推敲してください。
- ストーリーがありますか？
 - 目的や研究課題と結果・考察に矛盾はないでしょうか？
 - 適切な構成で、適切なことばと文章で、簡潔かつ効果的に記述できていますか？

投稿のお誘い

『コンピュータ&エデュケーション』への 投稿をお待ちしています

- 投稿は随時受け付けています
- 47号締切は8月31日です

新投稿システムが導入されます

- 49号（2020年3月1日受付開始）から



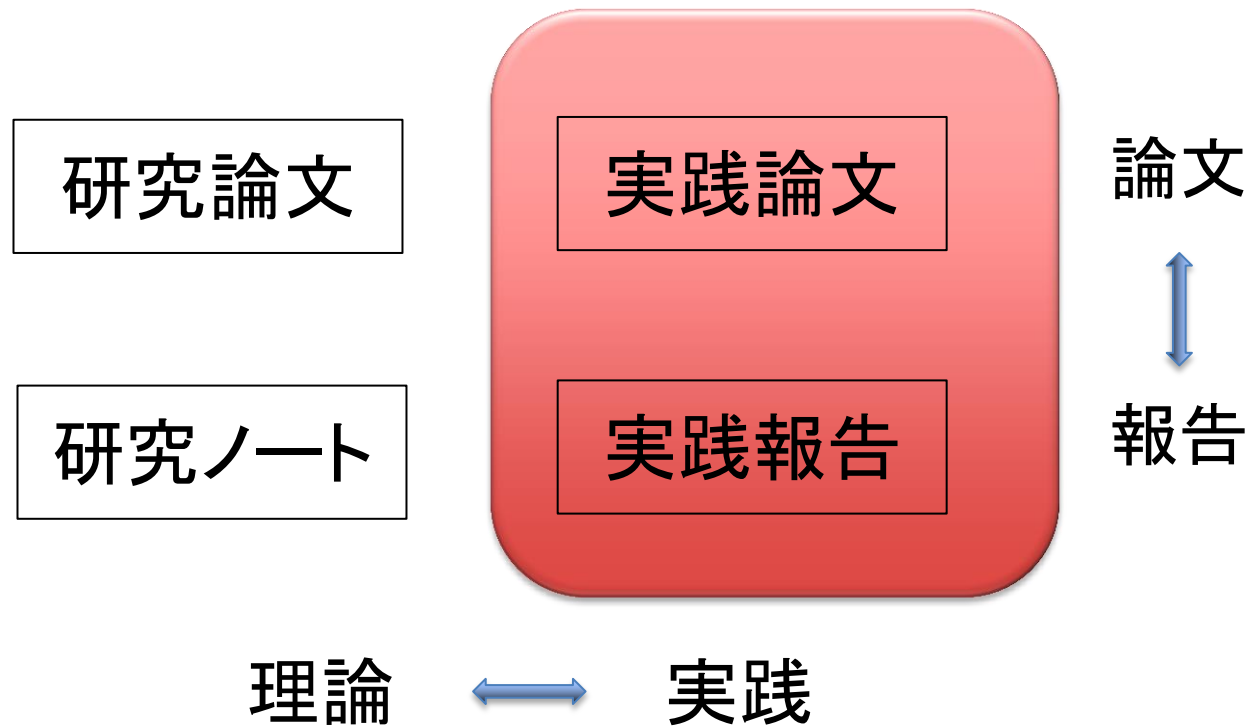
CIEC会誌「コンピュータ&エデュケーション」に 採択されるために

-実践論文・実践報告に期待するもの-

CIEC会誌編集委員
武 沢 護



新しい投稿カテゴリー





実践論文・実践報告の内容と記述

○研究実践に関する記述について

- ・問題意識 → みんなが困っている
- ・新規性 → やってみたいくなる
- ・再現性(普遍性) → うちの学校でもできる

○研究実践方法の手順の記述について

- ・実践手順の正確な記述
- ・被験者のプロフィール(実験群、統制群)情報
- ・データ収集の正確性、信頼性・・・サンプリング(方法、量など)

○データ分析の正確性、妥当性、信頼性について

○論文内における術語の正確な記述について

論文内での定義もしくは先行研究での援用

実践論文 実践報告 ある事例を例に

事例の概要:

- 研究テーマ: 情報関係の授業プログラムの開発
- 研究目的: 授業プログラムの妥当性と有効性を検証
- 研究方法
 - 事前事後アンケート: 各1回
 - 検証授業: 1回
 - 対象者: 2校・3学年 合計80名程度
- 測定方法: 事前、事後アンケート
 - 事後アンケートには自由記述を含む
- 分析方法: 事前・事後アンケート分析
 - テキストマイニング・計量テキスト分析を含む



実践論文 実践報告 手順の記述について

○研究実践方法の手順の記述

査読前

原論文：
対象者：2校・3学年
人数：合計80名程度
実践の手順の記述なし
被験者のプロフィールなし



査読後

追加：
本文)
・授業回数、アンケート実施の詳細など
注)
・被験者へのアンケート結果
（国のアンケート結果と比較）
・関連する情報モラルの教育の実施の有無

実践手順の記述の追加要請

- ・被験者のプロフィール
- ・当該のテーマに関する学習履歴等の記述

→ 正確な記述の必要性



実践論文 実践報告 アンケート分析について

○データ分析の正確性、妥当性、信頼性

手法:事後アンケート(%):

項 目	とても	まあまあ	あまり	まったく
授業のテーマが身近になった。	45.8	42.2	10.8	1.2
グループディスカッションで考えが深まった。	33.8	53.0	9.6	2.4
気をつけることがわかった。	77.1	22.9	0	0

考察:

有意義な話し合いができたとうかがえる。

教材や授業の妥当性や有効性が理解される。

→ 事後アンケートのみからの結論
妥当性・信頼性に欠ける

実践論文 実践報告 アンケート分析について

○データ分析の正確性、妥当性、信頼性

手法:事後アンケート(4点法) t-検定

質問項目	事前 AV	事前 SD	事後 AV	事後 SD	T-Value
知識1	3.28	0.8	3.5	0.6	2.11
知識2	3.28	0.82	3.82	0.38	5.71
態度	2.84	0.96	3.39	0.62	5.43

考察:

いずれの項目も実践後の平均値は上がっており、有意差も認められた。

→ t-検定を利用
妥当性？

実践論文 実践報告 アンケート分析(自由記述)について

○データ分析の正確性、妥当性、信頼性

List				
#	抽出語	品詞/活用	頻度	
1	学習	サ変名詞	6	
2	思う	動詞	6	
3	機械	名詞	4	
4	Azure	未知語	3	
5	感じる	動詞	3	
6	機能	サ変名詞	3	
7	使いこなせる	動詞	3	
8	使う	動詞	3	
9	非常	形容動詞	3	
10	便利	形容動詞	3	
11	活用	サ変名詞	2	
12	慣れる	動詞	2	
13	興味	名詞	2	
14	自分	名詞	2	
15	操作	サ変名詞	2	
16	複雑	形容動詞	2	
17	AI	未知語	1	
18	R	未知語	1	
19	プログラミング	サ変名詞	1	
20	メリット	名詞	1	
21	意味	サ変名詞	1	
22	演算	サ変名詞	1	
23	確実	形容動詞	1	
24	学ぶ	動詞	1	
25	学べる	動詞	1	
26	楽しい	形容詞	1	
27	関数	名詞	1	
28	機会	名詞	1	
29	記録	サ変名詞	1	

出現頻度

手法:

テキストマイニング(計量テキスト分析):

例文: サンプル文

考察: 名詞・動詞の係り受け頻度上位を列挙

→ 結果から考察に工夫が必要

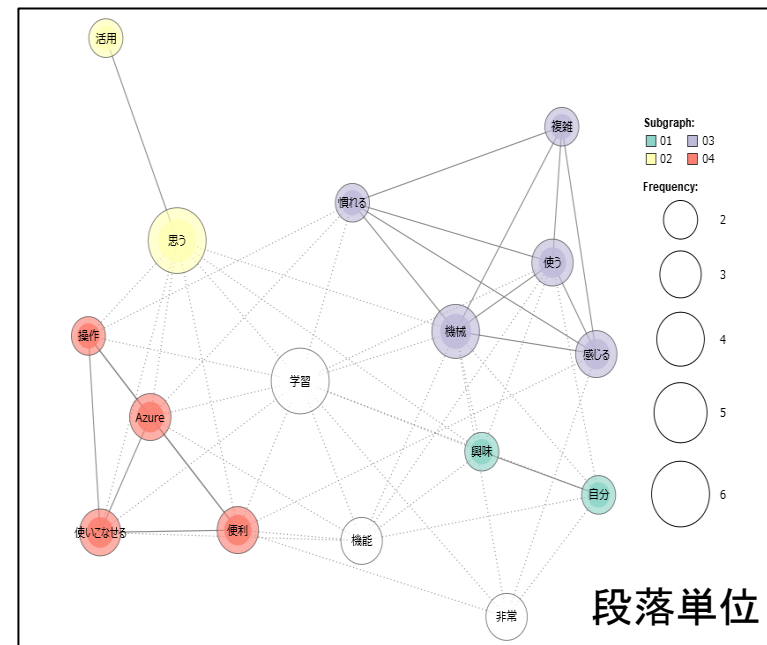
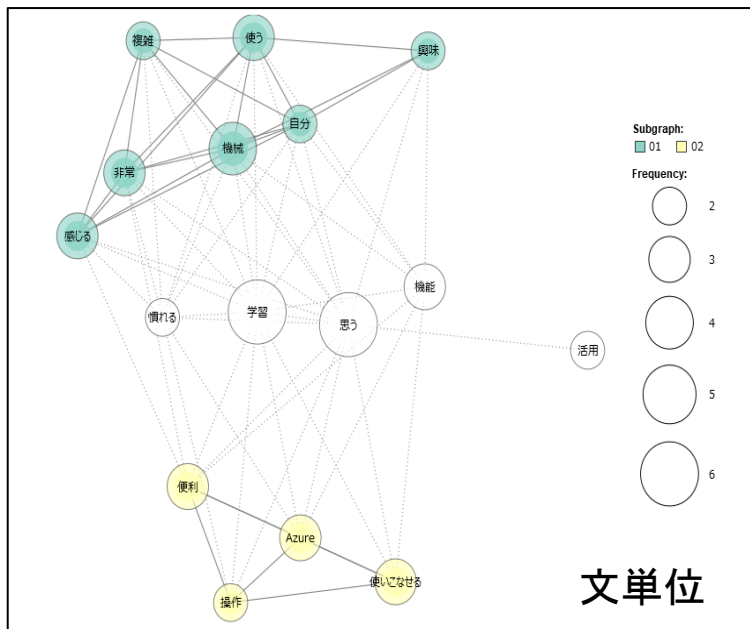
→ 実践事例の場合、少数の意見にも傾聴の価値はある!

実践論文 実践報告 アンケート分析(自由記述)について

○データ分析の正確性、妥当性、信頼性

手法: 共起ネットワーク

自然言語処理において任意の文や文書にある文字列と別の文字列が同時に出現するパターンをグラフ化したもの



考察: カテゴリー分類、語間の関連性などに言及

→ 結果から考察に工夫が必要

実践論文 実践報告 アンケート分析(自由記述)について

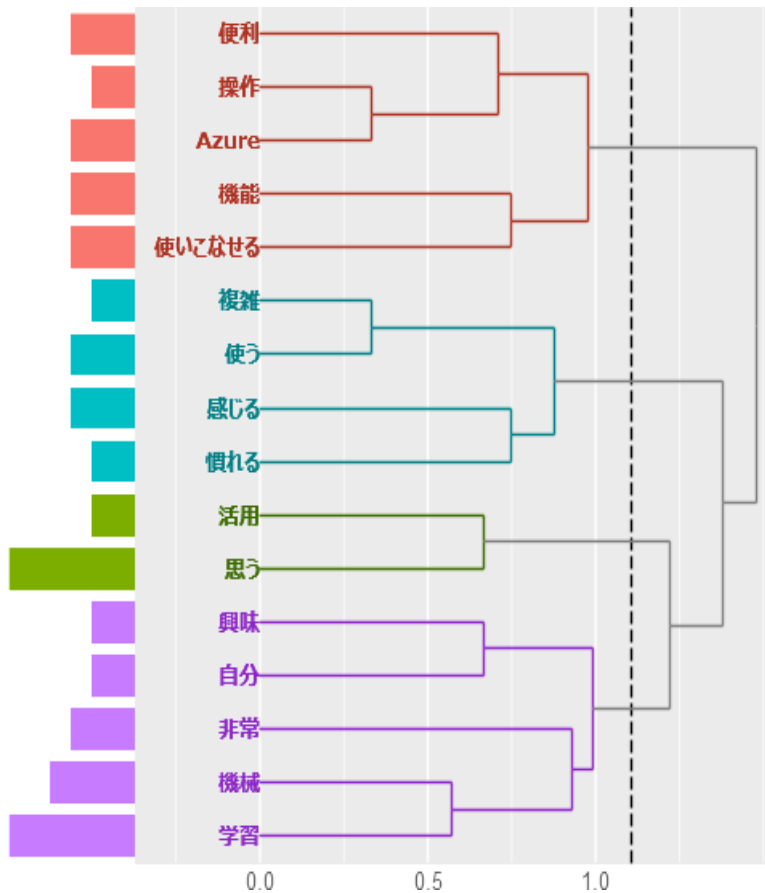
○データ分析の正確性、妥当性、信頼性

手法:階層的クラスター分析

ある集団について各個体の属性あるいは形質データにもとづき似たものどうし(データ間の距離)をいくつかの群にまとめて類型を作り出す分類手法で, 統計的多変量解析法の一つ。

考察:

カテゴリー分類、語間の関連性などに言及



→ 結果からどのように考察するか



会誌をよりよくするために(実践論文・実践報告)

ー教育分野でのデータ分析のむずかしさー

- 優れた実践事例の紹介
 - みんなが困っている
 - やってみたいくなる
 - うちの学校でもできる
- 実践後の分析手法の妥当性と信頼性
 - 教育実験環境の設定と整備(被験者のプロフィールなど)
 - 実証実験の頻度
 - 被験者の標本数
 - 統制群と実験群設定の妥当性

CIEC会誌
『コンピュータ&エデュケーション』
に採択されるために
—新しい投稿カテゴリーが期待するもの—

「研究論文」「研究ノート」で扱う課題とは？

大阪大学 全学教育推進機構
教育学習支援部 教授 村上正行

masayuki@murakami-lab.org

Twitter ID: @munyon74

<https://www.facebook.com/masayuki.murakami.14>



投稿カテゴリー（横川先生スライドより）

本誌は、とくにコンピュータやネットワークを活用した教育・研究に関する理論的・実践的な原稿を歓迎します。

改正前

論文

図表を入れて刷り上がり 6 ページ以内

活用事例

図表を入れて刷り上がり 4 ページ

論文および事例研究

図表を入れて刷り上がり 6 ページ以内

改正後 **NEW!**

研究論文・実践論文

コンピュータやネットワークを活用した教育・研究に関する理論的・実践的研究。
図表を入れて 6 ページ以内

研究ノート・実践報告

図表を入れて 4 ページ以内

ソフトウェアレビュー

自作のものを含む公開されたソフトウェアについて、その機能、特徴および教育・学習への有用性を会員に広く紹介することを目的としたもの。図表を入れて 2 ページ以内

「研究論文」「研究ノート」に 求められるものとは？

“コンピュータやネットワークを活用した
教育・研究に関する理論的研究”

日本教育工学会 論文種別

- 1. 論文 (Paper)
教育工学に関係のある独創的な研究結果，新規な方法・結果等で，信頼性が認められ，学問や教育工学の発展に役立つ内容を，順序立てて明瞭に記述したもの。会員の参考となる新しいデータをまとめたもの。
- 2. 教育システム開発論文 (Paper on Educational System Development)
教育の改善等を目的にして教育システムを開発し，既存の要素技術の組み合わせ方や方法，あるいは，開発したシステムや要素技術自体に新規性があるもの。開発したシステムの教育的な有効性を実証していなくても，教育システム開発論文として，この分野に関心を持つ会員や読者にとって価値と有効性があるもの。
- 3. 教育実践研究論文 (Paper on Educational Practice Research)
教育実践に貢献できる問題提起と意義があり，学問としての教育実践学を構築していなくても，研究手法や道具の開発，要因の分析，実践の改善や学習環境づくり，教師の教育実践力について，新たな点があるもの。教育実践研究論文として，この分野に関心を持つ会員や読者にとって価値と有効性があるもの。

日本教育工学会 論文種別

- 4. 資料 (Technical Information)

教育工学に関する研究資料, 実践報告, 調査報告, 教育システム開発, 教育コンテンツ開発など, 会員に参考になる有効な結果が得られているもの.

- 5. ショートレター

教育工学に関する, 実践的な研究成果, 新しい試み やその結果等, 教育システム開発等を記述したもの. 新規性 (オリジナリティ) は必ずしも高くなくても, 実践研究の方法と成果を明確に記述したもの. 4ページを厳守.

教育システム情報学会 論文カテゴリ

- 1-a) 一般論文

研究, 開発, 検討の結果をまとめたものであり, 新規性, 信頼性が高いもの. あるいは, 多くの研究を独自の視点でまとめ, 将来の研究分野の方向性を示したものであり, 新規性, 信頼性が高いもの. 教育実践を伴う評価に基づく高い有用性の提示は要求されないが, 研究の位置づけが関連研究との比較検討により明確になっていること.

- 1-b) 実践論文

情報システム・機器を利用した教育実践の結果をまとめたもので, その仕組みや条件が明確に記述され, 汎用性の高い知見や方法が客観的な形式で導出されており, 有用性, 信頼性が高いもの. あるいは, 情報システム・機器利用教育に関わるデータを包括的にまとめたもので, 有用性, 信頼性が高いもの. 高い新規性は要求されないが, 研究の位置づけが関連研究との比較検討により明確になっていること.

教育システム情報学会 論文カテゴリ

- 1-c) ショートノート

研究の動機, 方法, 結果などの技術的な点が明確に記述されており, 新規性, 信頼性があるもの. 一般論文のような研究の位置づけ, 他の関連研究との比較検討などに関する記述は要求されない.

- 1-d) 実践速報

情報システム・機器を利用した教育実践の結果をまとめたもので, その仕組みや条件が明確に記述されており, 有用性, 信頼性があるもの. 実践論文のような研究の位置づけ, 他の関連研究との比較検討などに関する記述は要求されない.

「研究論文」「研究ノート」に 求められるものは？

“コンピュータやネットワークを活用した
教育・研究に関する理論的研究”



新規性・信頼性が求められる

教育システム情報学会における 新規性・信頼性・有用性

- 新規性

主要な内容が公知・既発表または既知のことから容易に導き得るものでないこと

- 信頼性

論旨が通っており、また前提・結論などを信頼し得る何らかの根拠を示していること

- 有用性

内容が、教育・産業・学術の発展に何らかの意味で役立つものであること

特徴

- 新規性
 - 手法そのものが新しい
 - 従来研究に対しての研究の位置づけが必要
- 信頼性
 - 他人が同じ研究を行っても、同じ結果が得られる
 - そのための条件(範囲)を明確にしている
 - データを十分集めている
 - 一貫性がある(論理的な飛躍がない)
- 有用性
 - 対象となる教育実践の有効性を示している
 - 他人にとって、参考になる知見が示されている

「研究論文」「研究ノート」で扱う 研究テーマ・課題とは？

さまざまな研究が考えられるが、過去の論文賞
12編を対象として、大きく下記に分類した

- 開発・実験系（6編）
 - － 教材
 - － 教育方法
- システム開発系（3編）
 - － 教育支援システム
- 調査系（3編）
 - － 実態調査

開発・実験系

- 新しい教育方法や教材などの開発と、有効性の検証
 - 安藤 公彦・柴田 千尋・稲葉 竹俊(2018)「深層学習技術を用いた自動コーディングによる協調学習のプロセスの分析」Vol.43, p.79-p.84)
 - 白井 詩沙香・福井 哲夫(2014)「数式自動採点システム STACK における数式入力方法の改善」Vol. 37, pp.85-90
 - 廣瀬 英雄・月原 由紀・鈴木 敬一(2009)「項目反応理論による評価を加味した数学テストとe-learningシステムへの実装の試み」Vol.24, p.70～p.76

開発・実験系

- 新しい教育方法や教材などの開発と、有効性の検証
 - 佐藤 和彦・倉重 健太郎・岡田 吉史・佐賀 聡人 (2012)「学生のやる気を引き出す「見える」ソフトウェア開発演習の実現と評価」Vol.31, p.94-p.99
 - 松波 紀幸・大熊 雅士・長南 良子・福島 健介・牧野 豊・島田 文江・宮本 裕之・生田 茂(2008)「情報活用能力育成のためのカリキュラム開発とその検証-情報収集,判断・選択に重点をおいて-」Vol.22, p.70～p.75
 - 水野 邦太郎(2007)「本と人・人と人との絆を結ぶ互恵的な読書環境の創出」Vol.19, p.75～p.84

システム開発系

- 教育支援に関するシステムの開発と評価
 - 澤口 隆・巽 靖昭(2016)「バックグラウンド稼働クリッカー(bgClicker)の開発」Vol.38,2015.6
 - 伊藤 貴一・熊坂 賢次・諏訪 正樹・花房 真理子「自己探求する柔らかい構造化ツール『Hipparu-McS』の実装と評価」Vol.29, p.80～p.85
 - 八木 龍平、國藤 進「効果的学習技能が埋め込まれた Web テキスト読解支援システム -読解方略 SQ3R 法の適用を促進する Web インターフェースの開発と評価-」Vol.18,p.90～p.98

調査系

- 教育・学習に関する調査と分析
 - 片平 昌幸・中村 彰(2011)「新入生のICT素養と学習効果の統計学的評価」Vol.29, p.86～p.91
 - 妹尾 堅一郎(2004)「学習コミュニティを支えるメディア環境:「社会調査」の6年間に見る学習環境の変容」Vol.13, p.123～p.131
 - 妹尾 堅一郎(2004)「学習コミュニティにおける「支援者」の役割:慶應SFC「社会調査方」におけるコラボレーションの実際」Vol.14, p.62～p.70

まとめ

- 「研究論文」「研究ノート」を書く上で必要な点は？
 - 新規性は、どの点にあるか？
 - 信頼性は、十分にあるか？
 - 得られる成果は、教育・学習に役立つものか？
 - 開発・実践系
 - システム開発系
 - 調査系

2019 PC CONFERENCE セミナー 1

2019年8月8日（木）

CIEC会誌『コンピュータ & エデュケーション』に採択されるために
ー新しい投稿カテゴリーが期待するものー

統計的手法の適切な利用 ー投稿論文の質を高めるためにー

寺尾 敦

青山学院大学

atsushi [at] si.aoyama.ac.jp

お話しする内容

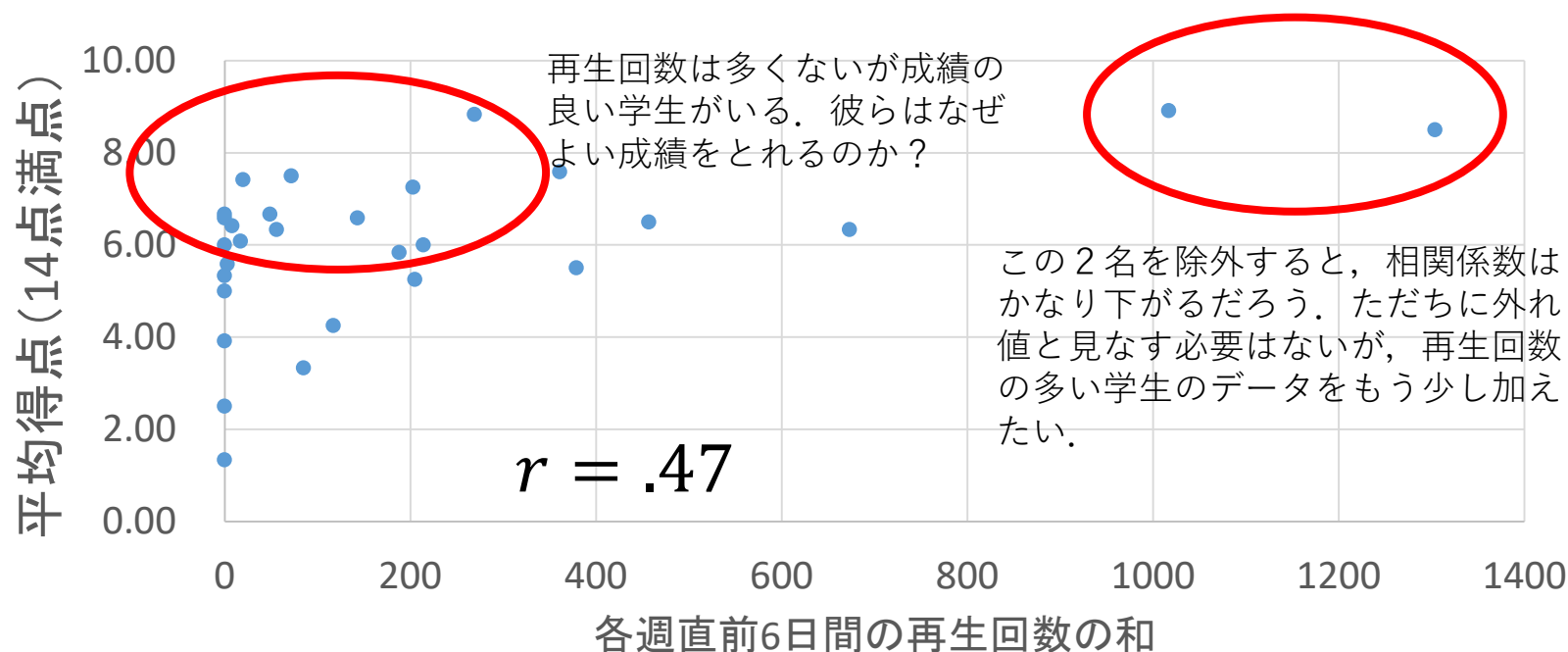
- データを図示する
- 母集団を意識する
 - 結果を一般化できる範囲は？「研究論文」か「実践論文」か？
- 検定の前提条件を意識する
 - 上位群・下位群という設定は適切なのか？
 - 2群の母分散の違いには（ときに）重要な示唆がある
- 統計的仮説検定に加えて（の代わりに）
 - 信頼区間を報告する
 - 効果量を報告する

- 『コンピュータ & エデュケーション』に掲載されたいくつかの論文について批判的なコメントをするところがあるが、それら論文の価値を損なうものではない.

データを図示する

- さまざまな計算の前に、まずはデータをよく観察する.
- 基本的な道具の例
 - ヒストグラム
 - 箱ひげ図（あるいは、ヴァイオリンプロット）
 - 散布図
 - 折れ線グラフ
- これらの図は論文中でもしばしば役立つ

例：英語音声教材の再生数とテスト成績との散布図（針持，2016）



母集団を意識する

- 『コンピュータ & エデュケーション』の投稿カテゴリには「研究」と「実践」がある.
- どちらのカテゴリの投稿するか判断材料のひとつとして、得られた結論を一般化できる範囲を考える. これは用いたデータの母集団である.
 - 一般化できる（と著者が考える）範囲が比較的広ければ「研究」、そうでなければ「実践」として投稿する.
 - 例：特定の中学校だけでデータをとっていても、日本の中学生全体（これが想定する母集団）に結論が一般化できると考えるなら「研究」で投稿する. 階層線形モデルで、中学の違いによる変量効果が存在するかどうかを考えてもよい.

検定の前提条件を意識する

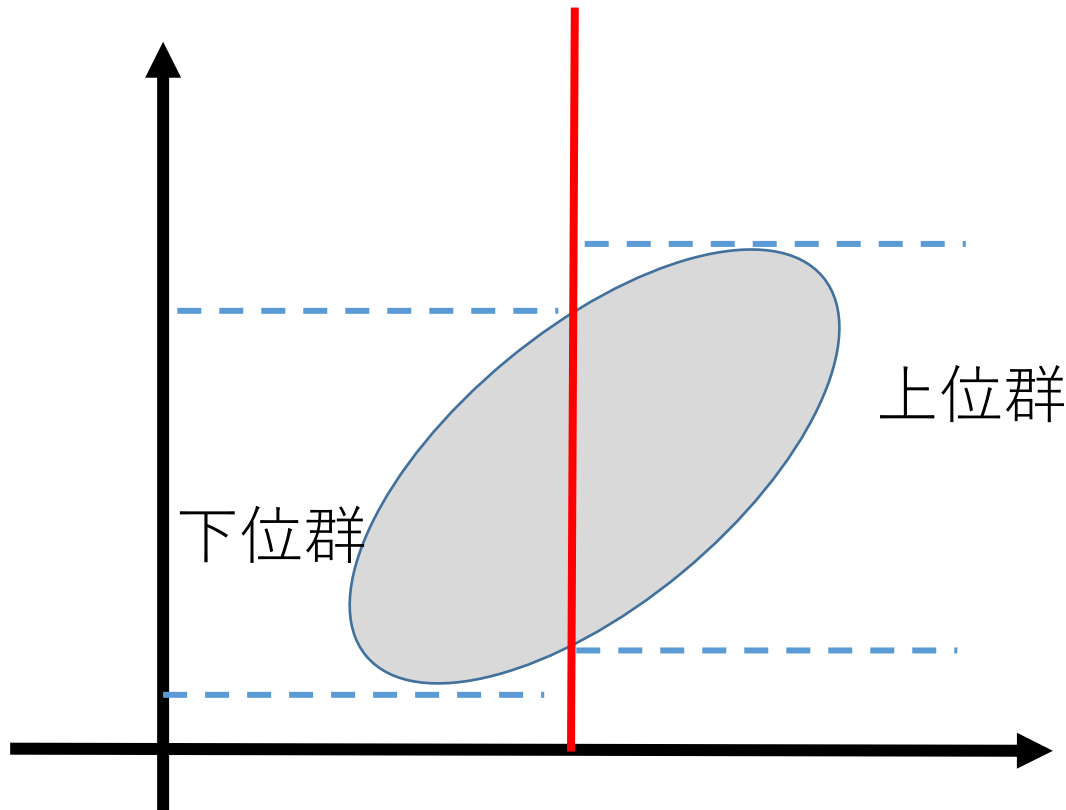
- これは母集団を意識することと深く関係している.
- パラメトリックな検定では母集団分布についての仮定がなされる. これは検定を行う前提条件となる.
- 2つの母集団平均の差の検定（いわゆる t 検定）の前提条件
 - 母集団が正規分布
 - 無作為抽出された標本 (i.i.d)
 - 母集団分散が等しい
- 3つの前提条件のうち、等分散性だけが検定でチェックされることが多い. 他の2つ成立しているかどうかは頭で考える.

2つの母集団平均の差の検定

- 等分散性の検定の結果が有意（つまり，等分散とは言えない）ならばウェルチの t 検定が使われることが多い。
 - 2段階の検定を行うこの手続きが不適切だという意見がある。
 - しかし，ユーザとしては当面この手続きに従っておけばよいだろう。
- 母集団分布として正規分布が仮定できなければ，マン・ホイットニーの U 検定を用いることが多い。
 - 標本の大きさが小さいからこの分析を用いるという記述（奥原ほか，2019）はやや言葉足らず。
- 2つの母数団分散が異なる場合には， U 検定は不適切となることに注意する。暫定的な結論だが，ウェルチの t 検定の方がよい（名取，2014）。

上位群・下位群の差の検定は適切？

- いわゆる t 検定の前提を考えると，もともとひとつだった標本を上位群・下位群にわけてこの検定を実行することには問題がある.
 - 上位群母集団，下位群母集団を想定し，それぞれが正規分布すると仮定することには無理がある.
 - そのような仮定が可能だとしても，群に分ける境界は標本に依存しており，境界付近では誤判別（母集団では上位群の学生なのに下位群の標本に含まれる）が生じている.
 - せっかくの量的情報が2カテゴリーの順序尺度情報に下落している.
 - 独立変数，従属変数が共に量的変数で，散布図から線形な関係が想定できるならば，相関係数を計算すればよい.



2 群の背後にそれぞれ正規母集団を考えるのは無理がある.

異なる母分散の意味を積極的に吟味

- 等分散性の検定はいわゆる t 検定の前提が満たされているかどうかを確認するために行われる.
- しかし、この検定にはもっと積極的な意味を求めてもよい.
- 教育研究では、2 群の母集団に差があることは、何らかの教育的工夫の効果を意味していることが多い.
- 等分散性の仮定は、教育的工夫によって母集団平均は変化してもよいが、分散は変化しないということを述べている.
- しかし、何らかの教育的工夫によって分散が変化することは十分にあり得ることだろう. その意味を積極的に考えてよい.

信頼区間や効果量を報告する

- 統計的仮説検定の結果は標本の大きさに依存する.
- 非常に小さな p 値が得られても、母数（母平均など）の値のわずかな違いには実質的な意味があまりないかもしれない.
- 仮説検定だけに頼るのではなく、信頼区間、効果量、実質的意味を含めた議論が必要.
 - 「帰無仮説検定は統計的分析の始まりに過ぎず、効果量やCIなどを併せて記載することが結果を適切に報告するために必要である」（APA *Publication Manual*, 6th ed., p.33. 訳は大久保・岡田, 2012）

例：PBLでの達成度の自己評価（佐藤ほか，2019）

達成度項目	1 回目平均	3 回目平均	差の95%信頼区間	
			下限	上限
協調作業	2.065	2.500	0.255	0.615
情報共有	1.987	2.554	0.350	0.802
作業分担	1.489	2.228	0.538	0.940
作業計画立案	1.533	2.272	0.527	0.952
...	...	...	...	...
VRソフトウェア開発	0.793	1.761	0.787	1.148

標本が大きい（ $N=92$ ）ため，仮説検定を行うといずれも $p < .001$ となる．達成度（5段階）の変化について信頼区間を求め，0.5 から 1.0 程度の変化を「ひとつ上のレベルという目標をおおむね達成」と評価した．

例：2タイプのロボットでの関心・意欲の変化（小林ほか，2018）

項目	ロボットの型	事前平均(SD)	事後平均(SD)	効果量 d	有意性検定
関心	Receiver	4.08 (1.17)	4.80 (0.39)	0.90	$p < .01$
	Giver	4.11 (0.99)	4.58 (0.61)	0.59	$p < .01$
学習意欲	Receiver	4.33 (0.89)	4.67 (0.89)	0.39	$p < .05$
	Giver	4.21 (0.79)	4.47 (0.70)	0.36	
活動意欲	Receiver	3.83 (1.27)	4.67 (0.89)	0.80	$p < .01$
	Giver	4.32 (0.82)	4.63 (0.60)	0.45	$p < .05$

論文では、「関心に関しては、・・・Receiver 型の方が効果量が高かった」のように、2つの型のロボットで効果量を比較している。それぞれの型のロボットで学習を行った子どもの人数が異なり、確率 p ではこうした比較ができないことに注意。

クロス表での効果量

- クロス表では連関係数を報告する.
- 名義尺度の場合
 - クラメールの V
 - ピアソンの C
- 順序尺度の場合
 - グッドマン=クラスカルの γ (ガンマ)
 - ケンドールの順位相関係数 τ_b
 - スチュアートの順位相関係数 τ_c
 - ソマーズの d

例：タイピング習熟度とレポート作成方法の関連（長澤，2019）

	キーボードを				合計 (人)
	ほとんど 見ず	たまに 見ながら	頻繁に 見ながら	ほぼずっと見な がら	
すべてPC	142 (17.2%)	322 (39.1%)	272 (33.0%)	88 (10.7%)	824
スマホ入力PC編集	4 (7.0%)	21 (36.8%)	21 (36.8%)	11 (19.3%)	57
スマホ下書PC入力編集	4 (18.2%)	9 (40.9%)	7 (31.8%)	2 (9.1%)	22
手書き下書 PC入力編集	8 (7.5%)	46 (43.4%)	35 (33.0%)	17 (16.0%)	106
すべてスマホ	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (100.0%)	1

$$\chi^2(12) = 22.2188, p = 0.0351, \underline{V = .08}$$

効果量はかなり小さい

表と統計量の修正

レポート 作成方法	タイピング習熟度			
	ほとんど 見ず	たまに 見ながら	頻繁に 見ながら	ほぼずっと 見ながら
すべてPC	142 (89.9%)	322 (80.9%)	272 (81.2%)	88 (73.9%)
スマホ入力PC編集	4 (2.5%)	21 (5.3%)	21 (6.3%)	11 (9.2%)
スマホ下書PC入力編集	4 (2.5%)	9 (2.3%)	7 (2.1%)	2 (1.7%)
手書き下書PC入力編集	8 (5.1%)	46 (11.6%)	35 (10.4%)	17 (14.3%)
すべてスマホ	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.8%)
合計 (人)	158 (100.0%)	398 (100.1%)	335 (100.0%)	119 (99.9%)

$$\chi^2(12) = 22.22,$$

$$p = .035,$$

$$V = .086$$

「すべてスマホ」
の1人を除外して、
4行4列の表とす
ると、

$$\chi^2(9) = 14.76,$$

$$p = .098,$$

$$V = .070$$

論文の分析をRで再現

```
nagasawa <- matrix(c(142, 4, 4, 8, 0,  
                     322, 21, 9, 46, 0,  
                     272, 21, 7, 35, 0,  
                     88, 11, 2, 17, 1),  
                   nrow = 5)
```

```
nagasawa
```

```
(result <- chisq.test(nagasawa))
```

```
result$expected
```

```
(V1 <- sqrt(result$statistic / (sum(nagasawa) * 3)))
```

```
> (result <- chisq.test(nagasawa))
```

Pearson's Chi-squared test

data: nagasawa

X-squared = 22.219, df = 12, p-value = 0.03514

警告メッセージ:

chisq.test(nagasawa) で: カイ自乗近似は不正確かもしれません

```

> result$expected
      [,1]      [,2]      [,3]      [,4]
[1,] 128.9029703 324.7049505 273.3069307 97.0851485
[2,]  8.9168317 22.4613861 18.9059406  6.7158416
[3,]  3.4415842  8.6693069  7.2970297  2.5920792
[4,] 16.5821782 41.7702970 35.1584158 12.4891089
[5,]  0.1564356  0.3940594  0.3316832  0.1178218
>
> (V1 <- sqrt(result$statistic / (sum(nagasawa) * 3)))
X-squared
0.08563259

```


分析を修正

```
nagasawa2 <- matrix(c(142, 4, 4, 8,  
                      322, 21, 9, 46,  
                      272, 21, 7, 35,  
                      88, 11, 2, 17),  
                    nrow = 4)
```

```
nagasawa2
```

```
(result2 <- chisq.test(nagasawa2))
```

```
result2$expected
```

```
(V2 <- sqrt(result2$statistic / (sum(nagasawa2) * 3)))
```

```
> (result2 <- chisq.test(nagasawa2))
```

Pearson's Chi-squared test

data: nagasawa2

X-squared = 14.755, df = 9, p-value = 0.0979

警告メッセージ:

chisq.test(nagasawa2) で: カイ自乗近似は不正確かもしれません

```

> result2$expected
      [,1]    [,2]    [,3]    [,4]
[1,] 129.030723 325.026759 273.577800 96.364718
[2,]  8.925669  22.483647  18.924678  6.666006
[3,]  3.444995   8.677899   7.304262  2.572844
[4,] 16.598612 41.811695 35.193261 12.396432
>
> (V2 <- sqrt(result2$statistic / (sum(nagasawa2) * 3)))
X-squared
0.06981649

```

お話しした内容

- データを図示する
- 母集団を意識する
 - 結果を一般化できる範囲から論文誌の投稿カテゴリを判断する.
- 検定の前提条件を意識する
 - 上位群・下位群という設定は適切ではない
 - 2群の母分散の違いには（ときに）重要な示唆がある
- 統計的仮説検定に加えて（の代わりに）
 - 信頼区間を報告する
 - 効果量を報告する

文献

- 針持和郎, 「デジタルテキストによる英語リスニングテキストに埋め込まれたデータリスニングの効果」, コンピュータ&エデュケーション, 41, 2016, 22-26.
- 奥原俊ほか, 「グループ学習における音声情報を用いた求人用語学習支援システムの試作」, コンピュータ&エデュケーション, 46, 2019, 46-51.
- 名取真人, 「マン・ホイットニーのU検定と不等分散時における代表値の検定方法」, 霊長類研究, 30, 2014, 173-185.
- 大久保街亜・岡田謙介, 『伝えるための心理統計』, 勁草書房, 2012.
- 佐藤和彦・服部峻・佐賀聡人, 「主体的学習を促す実践的ソフトウェア開発演習の改善とその効果の検証」, コンピュータ&エデュケーション, 46, 2019, 64-69.
- 小林溪太ほか, 「コミュニケーション型ロボットを活用した教育方法の体系的整理と学習効果の検討 — Giver型とReceiver型の比較を通して —」, コンピュータ&エデュケーション, 44, 2018, 79-84.
- 長澤直子, 「日本語入力から見る”PCが使えない大学生問題”」 コンピュータ & エデュケーション, 46, 2019, 58-63.

投稿のお誘い

『コンピュータ&エデュケーション』への 投稿をお待ちしています

- 投稿は随時受け付けています
- 47号締切は8月31日です

新投稿システムが導入されます

- 49号（2020年3月1日受付開始）から