

統計処理ソフトを用いた演習

大阪公立大学工業高等専門学校

梶真理香・鬼頭秀行・梶崎亮・稗田吉成・松野高典・室谷文祥¹

1 はじめに

「大阪公立大学工業高等専門学校」(以下公立大高専)は、2022年度に学校名称を大阪府立大学工業高等専門学校(以下府大高専)から変更し、新しい教育組織・内容(コース再編・カリキュラム改革等)を展開している。IoTやAI等の科学技術の進歩や社会の変化と要請に応えられるよう、コンピュータの活用を求める科目が増加し、授業の中でも積極的に使用するために、1年次は授業で使用するノートパソコンを希望者全員に貸与している。貸与した端末には、自由にソフトウェアをインストールできるよう、管理者権限を与えている。2年次からは学校が指定する条件に合った端末を各家庭で購入して使うCYOD(Choose Your Own Device)を導入している。

2021年8月に府大高専の「総合的な工学教育システムを背景にしたデータ思考力の涵養」プログラムが文部科学省による「数理・データサイエンス・AI教育プログラム(リテラシーレベル)」([1])に認定された。2022年3月には「DXマインドの気付き・動機付けから始める専門技術者の育成」事業が文部科学省による「デジタルと専門分野の掛け合わせによる産業DXをけん引する高度専門人材育成事業」([2])に採択された。公立大高専では本事業の補助金(以下DX補助金)により、デジタルと専門分野の教育を掛け合わせた実験・実習カリキュラムを実施するために、その取組の基盤となる教育設備等の整備を行い、デジタル化が進む産業分野をけん引する高度専門人材を育成することが可能になった。

公立大高専では1年次の数学科目として「基礎数学A(2単位,前期)」「基礎数学B(2単位,後期)」「基礎数学C(2単位,通年)」,4年次の数学科目として「確率統計(2単位,前期)」が設定されている。また,1年次の専門基礎科目として「情報1(2単位,後期)」が設定されており,Excelの初歩的な操作を学ぶ。各科目の内容は,国立高専機構が策定したモデルコアカリキュラム([3])の内容と特に違いはないため,詳細についてはシラバス([4])で確認してもらうこととして,ここでは説明を省く。

上述の背景を踏まえ,数学の1年次科目において,これまでの板書による通常の授業と合わせてICTを活用した数学教育を行うことにした。本稿では「DXマインドの気付き・動機付けから始める専門技術者の育成」事業に関連し,統計の導入教育として1年次科目「基礎数学B」において実施した,統計処理ソフトを用いた演習について報告する。

なお,公立大高専では,eラーニングシステムにおける学習管理システム(LMS, Learning Management System)としてMoodle(Modular Object-Oriented Dynamic Learning Environment)を利用している。数学の1年次科目におけるMoodleの学力補充指導への活用事例については日本高専学会誌([5])で紹介しているので,そちらを参照していただきたい。

¹E-mail: muroya@omu.ac.jp

2 統計処理ソフトを用いた演習

公立大高専では、1年次の数学において確率統計の内容に繋がる「場合の数」を学年末試験の直前に扱う。また、授業最終週を学年末試験後に設定し、試験返却に充てている。以上を踏まえ、授業最終週の数学授業において、統計処理ソフトを用いた演習を行うことにした。1年次後期科目「基礎数学 B」では週 2 回の授業を実施していることから、「基礎数学 B」の最終週を統計処理ソフトを用いた演習に充てることにした。

基礎数学 B の最終週 2 回のうちの 1 回目で使用使用する統計処理ソフトは、「情報 1」の授業で初歩的な操作を学んだ Excel とし、2 回目の授業で使用使用するソフトは、5 年次の卒業研究等で学生が将来的に使用する可能性がある「R」「SPSS」とした。学生に身近な題材を扱うべく、データには匿名の成績表を用いた。また、統計の導入教育という観点から、各回とも統計処理ソフトの使用方法を修得し、学生自らデータを分析できることを目標とした。演習の際は、ノートパソコンを持参し忘れた学生や充電切れとなった学生を対象に、DX 補助金で購入した端末を貸与し、全員が授業に参加できるよう配慮した。

本稿では、授業で使用した図表を引用するとともに、5 章に配布資料の抜粋を掲載しているので、参照いただきたい。

2.1 Excel を用いた演習

最終週 2 回のうちの 1 回目では、表 1 の通り、数学と英語の 10 名分の成績をサンプルデータとして用意し、中学校で学んだ平均値、中央値、最頻値、度数、箱ひげ図、および、「情報 1」の授業で学んだ分散、標準偏差、ヒストグラム、相関係数、回帰直線の復習を行い、Excel の関数、およびグラフ・図表の作成機能を用いて、サンプルデータの分析方法を紹介した。次に、偏差値の定義を紹介し、図 1 の通り、サンプルデータの A は数学と英語が同じ点数であるが、偏差値が違うことを確認した。その後、演習時間を取り、各学生がサンプルデータの分析を行った。

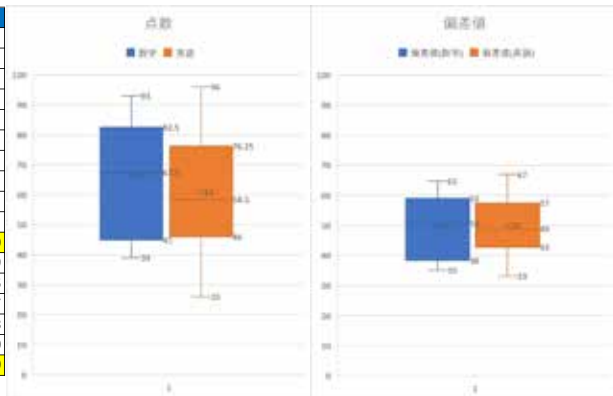
表 1: サンプルデータ

名前	A	B	C	D	E	F	G	H	I	J
数学	71	90	77	80	39	93	45	45	58	64
英語	71	96	51	66	48	89	26	51	72	40

続いて、教員が 2022 年度「基礎数学 A」および「基礎数学 C」の前期中間成績を用いて、統計基本量の導出と、箱ひげ図、度数分布表、ヒストグラムを作成する方法を解説した。その後、各学生が「基礎数学 A」および「基礎数学 C」の前期末成績を用いて、図 2 の結果を導出すべく、演習時間を取った。演習中は Excel の相対参照、絶対参照の使い分けに躓く学生も見られたため、適宜フォローを行った。

また、授業資料作成時に河合塾の Web サイト [6] で公開されていた「2022 年度第 3 回全統共通テスト模試（マーク式）」の成績統計資料データを用いて、標本数が多いデータの分析方法を紹介した。

名前	数学	英語	偏差値(数学)	偏差値(英語)
A	71	71	53	55
B	90	96	63	67
C	77	51	56	45
D	80	66	58	52
E	39	48	35	44
F	93	89	65	64
G	45	26	38	33
H	45	51	38	45
I	58	72	46	55
J	64	40	49	40
平均値 μ	66.2	61	50	50
中央値	67.5	58.5	51	49
最頻値	45	51	38	45
最高点	93	96	65	67
最低点	39	26	35	33
分散	332.6	429	100	100
標準偏差 σ	18.2	20.7	10	10



$$y=50+10*(x-\mu)/\sigma$$

図 1: サンプルデータを用いた分析結果

階級名	度数 (A)	度数 (C)	相対度数 (A)	相対度数 (C)
100	100	7	0.044	0.025
95	95-99	22	0.138	0.063
90	90-94	16	0.101	0.138
85	85-89	23	0.145	0.107
80	80-84	17	0.107	0.089
75	75-79	10	0.063	0.107
70	70-74	14	0.088	0.064
65	65-69	9	0.057	0.107
60	60-64	8	0.050	0.063
55	55-59	6	0.038	0.057
50	50-54	13	0.082	0.069
45	45-49	2	0.013	0.050
40	40-44	6	0.038	0.006
35	35-39	4	0.025	0.019
30	30-34	1	0.006	0.019
25	25-29	1	0.006	0.006
20	20-24	0	0.000	0.000
15	15-19	0	0.000	0.000
10	10-14	0	0.000	0.000
5	5-9	0	0.000	0.000
0	0-4	0	0.000	0.000
合計	150	150		
平均点	76.4	73.2		
標準偏差	18.1	17.2		

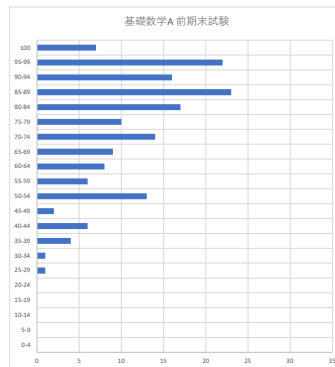


図 2: 2022 年度「基礎数学 A」および「基礎数学 C」前期末試験の分析結果

2.2 R および SPSS を用いた演習

最終週 2 回のうちの 2 回目の事前課題として、各自の端末に統計分析ソフト「R」および R の統合開発環境である「R Studio」をインストールするよう指示した。授業開始の段階で、概ね全員がソフトのインストールを完了していたが、事前課題をこなせなかった学生には、DX 補助金で購入し、これらのソフトをインストールした端末を貸与することで、全員が授業に参加できるよう配慮した。

授業では R Studio の基本操作を説明し、2.1 節で扱ったデータについて統計基本量の導出と箱ひげ図の描画、図 3 の通りヒストグラムの描画を行った。サンプルデータを入力する際に使用する `matrix()` 関数は、縦横に並べた数を入力する際に使用する関数であること、縦横に並べた数は行列と呼ばれ、2 年次通年科目「ベクトル・行列」で学ぶことを紹介した。その後、演習の時間を取り、各学生が R を用いてデータ分析を行った。また、Excel, R の他に、DX 補助金で購入した統計分析ソフト「SPSS」を用いてデータ分析する方法を紹介した。

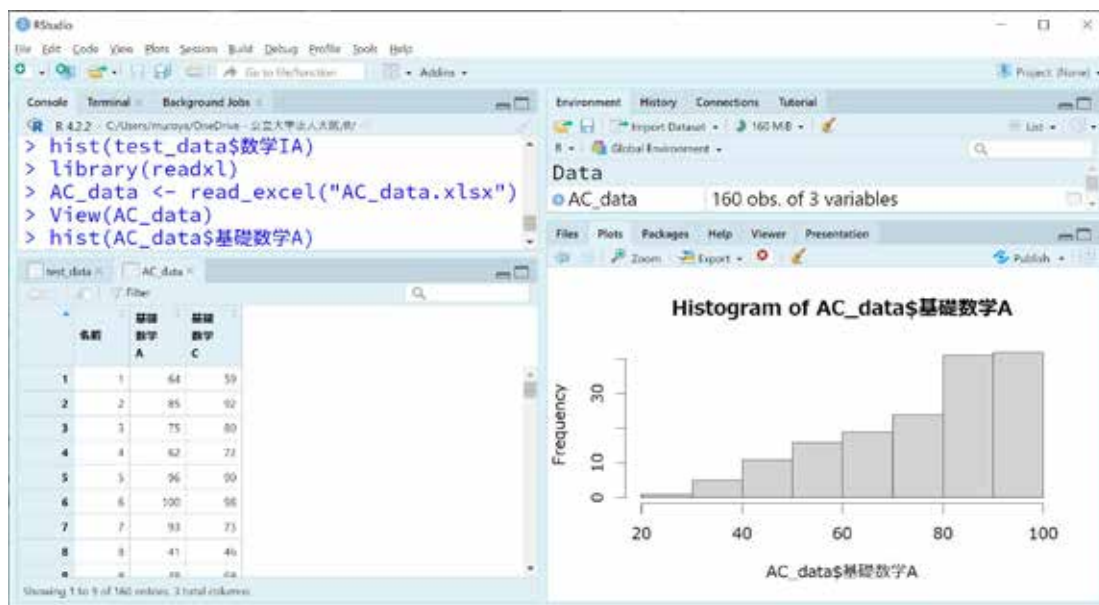


図 3: R Studio による 2022 年度「基礎数学 A」前期末試験のヒストグラム

3 授業アンケートの結果

2022 年度の「基礎数学 B」の最終授業後に授業評価アンケートを実施した。12 項目ある設問の中で、データ分析とコンピュータに関する設問は表 2 の 4 つである。各設問の選択肢は「A (良く当てはまる)」～「E (全く当てはまらない)」の 5 つとした。表 3 はアンケートの各設問に対する 5 つの選択肢の回答数を集計し、選択肢の A～E をそれぞれ 5～1 として各設問の回答の平均を記載したものである。

表 2: 授業評価アンケート設問

Q1	Excel を用いて、基本統計量を求めることができましたか。
Q2	Excel を用いて、度数分布表とヒストグラムを作成できるようになりましたか。
Q3	R や SPSS を用いて、ヒストグラムを作成できるようになりましたか。
Q4	ノートパソコンを用いた授業を、今後の数学の授業で増やしてほしいですか。

表 3: 授業評価アンケート結果

	A	B	C	D	E	平均
Q1	53	54	25	4	0	4.15
Q2	53	53	19	10	1	4.08
Q3	47	53	24	10	2	3.98
Q4	48	51	24	9	4	3.96

アンケートでは自由記述欄として「優れている点」「改善してほしい点」「感想・その他」

を用意した。以下、自由記述欄の中からデータ分析とコンピュータに関連するものをいくつか紹介する。自由記述欄の校正は、かな表記と句読点の統一に留めている。

1. 「優れている点」

- データ分析の授業で R をしたこと。
- ノート PC を使用した授業はとてもわかりやすい説明をされていてわかりやすかった。
- R について学ぶ時間があったこと。
- 他教科に比べて PC 等を使用した授業が多く、わかりやすかった。

2. 「改善してほしい点」

- データ分析は PC を用いたものだけでなく、手計算で求めるものもやってほしい。

3. 「感想・その他」

- 情報処理の授業では、情報 1 でも学びましたが、改めて Excel を扱うことができてよかったです。

4 まとめと今後の検討事項

表 3 の通り、各設問とも、アンケートに回答した 136 名の学生のうち、7 割から 7 割 5 分程度が「A (良くあてはまる)」「B (ややあてはまる)」と回答している。このことから、本演習を通じて、多くの学生が統計処理ソフトを用いたデータ分析に前向きに取り組むことができたと考えられる。また、多くの学生が、Excel の GUI(Graphical User Interface) のみに慣れ、R や R Studio の CUI(Character-based User Interface) に不慣れであることを想定したが、Q2「Excel を用いて、度数分布表とヒストグラムを作成できるようになりましたか。」と Q3「R や SPSS を用いて、ヒストグラムを作成できるようになりましたか。」のアンケート結果を見る限りにおいては、大きな差異は見られなかった。公立大高専の 1 年次ではコンピュータの活用が進んでおり、自由記述欄では 2 名の学生が「R」を扱ったことを「優れている」と回答した。最終週 2 回のうちの 2 回目の授業では、R に限らず得意なプログラミング言語を使用するよう促したところ、Python を用いてデータ分析を行った学生がいた。

その一方、4 名が Q4「ノートパソコンを用いた授業を、今後の数学の授業で増やしてほしいですか。」で「E(まったくあてはまらない)」と回答したこと、演習中にタイピングや PC の基本操作に不安を覚える学生が散見されたことから、コンピュータが苦手な学生へのフォローも必要であると考えられる。

今後の検討事項として、本演習に係る習熟度の評価が挙げられる。今回は課題提出がなく、客観的な習熟度の評価を行うことができなかったが、今後は課題を出しオンラインで提出させることを検討する。また、2 年次以降の数学科目においても、LMS や CYOD 端末を活用した授業を拡充すべく、数学担当教員一同で力を合わせて取り組んでいきたい。

5 授業資料

以下は、統計処理ソフトを用いた演習で使用した配布資料の抜粋である。

【2】統計とは 5min

・「統計」とは「統(す)べて計る」と書くように、まず、対象とする集団や現象を大量に観察・測定することから始まります。観察・測定には、調査や実験、記録データの収集が含まれます。集められたデータを知りたい目的に沿って整理し、平均や指数などを求めて指標としてとらえたり、グラフでまとめて表現(統計グラフ)したものが統計です。【なるほど統計学園より引用】

- ・中学校ではデータの活用として、平均値、中央値、最頻値、度数、箱ひげ図などを学びました。
- ・「情報1」の授業では、分散、標準偏差、ヒストグラム、相関係数、回帰直線などを学んでいます。
- ・本校のカリキュラムでは4年生の「確率統計」で統計量と標本分布、統計的推定、統計的検定などについて学びます。

【4】エクセルを用いてデータの分析をしてみよう

・情報1 講義ノート 08, 11, 12 を参考

・sample ファイル(説明) 10min

・偏差値の説明 5min

データの値を X 、平均値を μ 、標準偏差を σ としたとき、偏差値 $= 50 + \frac{10(X - \mu)}{\sigma}$ で定義される。

偏差値とは、元のデータを、平均を 50、標準偏差を 10 のデータに変換したときに示す値である。

- ・sample ファイル 15min (5min 休憩)
- ・基礎数学 A・C の前期中間試験の結果(説明) 10min
- ・相関係数の説明 5min
- ・基礎数学 A・C の前期末試験の結果 15min
- ・共通テスト模試(河合塾実施)の結果(説明) 5min
- ・共通テスト模試(河合塾実施)の結果 15min (途中で1回目授業終わり)

【5】R を用いてデータの分析をしてみよう

・R は R と RStudio をインストールして実行する

・どうしてもインストールできなかった場合は、教員からインストール済みの PC を借りて実行する

・【4】の内容を R を用いて分析する

・sample ファイルの数学について分析する 5min

```
> summary(c(71,90,77,80,39,93,45,45,58,64))
> 数学 <- c(71,90,77,80,39,93,45,45,58,64)
> summary(数学)
> hist(数学)
> boxplot(数学)
```

・sample ファイルの英語について分析する 5min

・sample ファイルの数学と英語について分析する 5min

```
> 数学英語 v1 <- matrix(c(71,90,77,80,39,93,45,45,58,64,71,96,51,66,48,89,26,51,72,40),10,2)
> 数学英語 v1
> summary(数学英語 v1)
> hist(数学英語 v1[,1])
> hist(数学英語 v1[,2])
> boxplot(数学英語 v1)
```

・csv ファイル(sample.data.csv)を読み込んで分析する 10min

```
> 数学英語 v2 <- read.csv("sample.data.csv") 注:読み込むファイルは作業ディレクトリに入れておく
> 数学英語 v2
> summary(数学英語 v2)
> hist(数学英語 v2[,2])
> hist(数学英語 v2[,3])
> boxplot(数学英語 v2[,c(2,3)])
```

・偏差値についてどうするか? 25min (5min 休憩)

・基礎数学 A・C の前期末試験の結果 15min

・共通テスト模試(河合塾実施)の結果 15min

【6】Python や SPSS を用いてデータの分析をしてみよう 10min (【5】が早く終わった人対象)

・Python は Google Colaboratory を用いて実行する(情報1 講義ノート 10 を参考)

・SPSS は教員からインストール済みの PC を借りて実行する

・【4】【5】の内容を Python や SPSS を用いて分析する

【7】まとめ 5min

・今回は全てのデータが揃っている例についての統計に触れた

・4年生の確率統計の授業では一部のデータから全体についてを考える統計を学ぶ

参考文献

- [1] 文部科学省, 数理・データサイエンス・AI教育プログラム認定制度 (リテラシーレベル),
https://www.mext.go.jp/a_menu/koutou/suuri_datascience_ai/00002.htm
(2023年11月14日)
- [2] 文部科学省, デジタルと専門分野の掛け合わせによる産業DXをけん引する高度専門人材育成事業,
https://www.mext.go.jp/a_menu/koutou/sankangaku/1413155_00013.htm
(2023年11月14日)
- [3] 国立高専機構, モデルコアカリキュラム,
https://www.kosen-k.go.jp/about/profile/main_super_kosen.html
(2023年11月14日)
- [4] 大阪公立大学工業高等専門学校, 2023年度本科シラバス1年生 (2023年),
<https://www.ct.omu.ac.jp/studies/classes/syllabus-reg/>
(2023年11月14日)
- [5] 梶真理香・鬼頭秀行・檜崎亮・稗田吉成・松野高典・室谷文祥, 大阪公立大高専におけるICTを活用した数学教育, 日本高専学会誌, **27** (4)(2022), 21–24.
- [6] 河合塾, 成績統計資料データ,
<https://www.kawai-juku.ac.jp/zento/statistics/>
(2023年11月14日)