

論文概要

学 科 東京医療保健大学
医療情報学科
学籍番号 HI012057
氏 名 永田 有希

統計処理ソフトウェア「R」の GUI 化

近年、医療ビッグデータが注目され始め、これらのデータを用いた解析結果により、新しいエビデンスが生まれ、より質の高い医療の提供がなされることが期待されている。医療に限らず他分野でもビッグデータが注目され、その解析結果を経営戦略などに活用することが期待されている。しかし、手軽に使用できる統計解析ソフトウェアが少なく、大半のものが有償で高価なため、日常的に統計解析を行うことは個人レベルでは難しい。また、コマンド入力を必要とするものや、統計解析を行う際に実施する解析手法を指定する必要があるものなど、ある程度の統計解析に関する知識が必要となる統計解析ソフトが多く、初学者にとって使用しやすい統計解析ソフトウェアは少ない。このことから、統計解析を行うことが出来る人が育たず、その結果、人手不足となり、蓄積されたデータを活用できずにいるということが考えられる。

そこで本研究では、コマンドの勉強及び、統計学に関する専門知識を必要とすることなく、誰でも無償で統計解析を行うことが出来る統計解析ソフトウェアの開発を行った。

開発は、R と Java、NetBeans を使用し、画面や入力方法などの仕様において JMP のように初学者が使いやすいものができるよう工夫した。

本ソフトウェアは、CSV ファイルのデータの表示、1 変量の分布、2 変量の関係の画面で解析したい項目を指定し、指定された項目の尺度に適した、統計解析の結果を表示させることができる、初学者にとって使いやすい統計解析ソフトウェアとなっている。

目次

第1章	はじめに	
1.1	研究の背景	1
1.2	目的	1
第2章	ソフトウェア・インターフェース	
2.1	Java	2
2.2	NetBeans	2
2.3	R	2
第3章	研究の方法	
3.1	研究の方法	3
3.2	設定	3
3.2.1	環境変数の設定	3
3.2.2	NetBeans の設定	4
第4章	研究の結果	
4.1	ソフトウェアを使用するための準備	6
4.2	ソフトウェアの使用法	6
4.3	実施できる統計解析	6
4.4	入力画面	6
4.5	1 変量の分布	7
4.5.1	クロス集計	8
4.5.2	サマリー	9
4.5.3	フローチャート図	10
4.6	2 変量の関係	11
4.6.1	カイ二乗検定	12
4.6.2	分散分析	13
4.6.3	ロジスティック回帰分析	13
4.6.4	線形回帰分析	13
4.6.5	フローチャート図	14
第5章	プログラムの説明	
5.1	CSV	15
5.2	1 変量の分布	16
5.3	2 変量の関係	19

第6章	まとめ	
6.1	今後の課題	22
6.2	結論	22
	謝辞	23
	参考文献	23