

연구노트

통계패키지 사용성 평가를 통한 한국형 통계패키지 개발 가능성 모색*

이재열** · 신인철*** · 이영조**** · 김청택***** · 박원호***** · 진윤아***** · 김혜진*****

사회과학연구에서 통계분석을 적용한 양적 접근이 활성화됨에 따라, 많은 연구자들이 다양한 범용통계패키지를 이용하고 있다. 다만 일부 통계패키지를 제외하고는 대부분의 연구자는 외국산 유료 소프트웨어를 이용하고 있을 뿐만 아니라, 학습을 위해 많은 시간과 비용을 지불하고 있다. 이 연구는 한국형 통계패키지 개발을 목표로 추진된 융·복합 연구에서 진행된 통계패키지 사용성 평가결과를 중심으로 기존 범용통계패키지의 특성을 탐색해보았다. 그리고 웹기반 통계패키지를 새롭게 개발하여 기존 범용통계패키지와 비교를 통해 개선사항을 확인하였다. 이와 더불어 평가자들에게 사용 의향성을 확인하여 초급자용 통계패키지를 개발하고자 할 경우 학습용이성과 효율성, 그리고 이해용이성이 매우 중요한 방향성임을 확인하였다. 이러한 결과는 앞으로 통계패키지를 개선하거나 새롭게 개발할 때 참고할 기준이 될 것으로 기대한다.

주제어: 통계패키지, 사용성 평가, 질적 비교연구방법론

* 이 논문을 심사하여 주신 익명의 심사자 세 분께 감사드립니다. 이 논문은 2017년도 융복합 연구과제 지원사업의 지원을 받아 수행된 연구이다.

** 서울대학교 사회학과 교수(jyyee@snu.ac.kr), 주저자.

*** 서울시립대학교 도시사회학과 조교수(incshin@uos.ac.kr), 교신저자.

**** 서울대학교 통계학과 교수(youngjo@snu.ac.kr).

***** 서울대학교 심리학과 교수(ctkim@snu.ac.kr).

***** 서울대학교 정치외교학부 교수(wpark@snu.ac.kr).

***** 서울대학교 아시아연구소(yoona03@snu.ac.kr).

***** 서울대학교 아시아연구소(kiyo3@snu.ac.kr).

I. 서론

최근 사회과학연구에서 통계학적 방법론을 적용한 양적 접근이 그 어느 때보다도 활성화되고 있다. 이러한 추세 속에서 글로벌 IT기업인 IBM은 2009년 대형 통계소프트웨어사인 SPSS를 인수하는 등 세계적으로 통계소프트웨어의 개발·활용에 대한 관심이 높아지고 있다. 그러나 현재 국내에서 범용적으로 활용되고 있는 통계 패키지인 SAS, SPSS, STATA 등은 모두 외국산 제품들이기 때문에 구입과 임대에 많은 외화가 낭비될 뿐만 아니라, 언어 등과 같이 국내 이용자들에게 친화적이지 않은 문제점을 갖고 있다. 뿐만 아니라 통계패키지의 설치와 운용 과정에서 컴퓨터의 메모리를 과도하게 차지하여 대용량 자료 처리에 한계를 가지고 있다. 또한 범용통계프로그램의 특성상 다양한 분석방법들이 포함되어 있어 실제 이용하지 않은 분석방법에 대한 비용까지 지불해야 하는 비경제성도 문제이다. 이러한 이유 때문에 최근 빅데이터의 활용 가능성의 증대와 함께 무료 통계소프트웨어인 R에 대한 관심이 커지고 있는 상황이다. 하지만 R의 경우 오픈소스 프로그램이 가지고 있는 메모리 관리 문제나 보안성의 문제뿐만 아니라 GUI 기능이 약하고 숙련에 있어 다소 많은 노력이 요구된다는 한계가 있어 통계에 익숙하지 않거나 소프트웨어 활용 능력이 낮은 사람들에게 접근성이 좋지 않은 단점을 가지고 있다.

이에 외국산 통계패키지를 대체할 수 있는 국산 통계패키지를 개발하려는 노력들이 전개되어 왔다. 대표적으로 이정진 외(1995)는 생존분석, 베이지안 분석, 보험 통계 분야 등에 적용할 수 있는 통계패키지 설계·개발을 시도한 바 있다. 하지만 MS-DOS 시스템에 기반하였기 때문에 당시에 상용화된 SPSS의 GUI(Graphical User Interface)에 비취볼 때 사용자 친화적이지 못해 범용적으로 활용되지 못한 바 있다. 이공희(2000)는 다양한 통계분석기법이 포함된 소위 패키지 형태의 소프트웨어보다는 X-12-ARIMA에 바탕을 둔 BOK-X-12-ARIMA 프로그램을 개발하여 계절변동조정에 대한 전문적인 지식이 없는 통계작성(이용)자가 계절변동조정통계를 정교히 작성할 수 있는 소프트웨어 개발을 시도한 바 있으나, 여전히 시계열분석에 많이 활용되는 SAS나 MATLAB과 비교했을 때 상대적 우위성이 결여되어 사용자들에게 주목을 받지 못하였다.

이와는 달리, 유의미한 성과를 이룬 국내 연구도 있는데, 대표적으로 서울대 통

계학과에서는 빅데이터분석이 가능한 통계패키지인 가칭 SRC-STAT를 개발한 바 있다. 이 통계패키지는 평균, 분산 등 기초통계부터 의학분야에서 사용되는 생존자료, 사회과학 분야에서 활용되는 다변량 자료 및 다수준 모형뿐만 아니라, 금융분야에서 활용되는 시계열자료 등 다양한 자료를 분석할 수 있는 범용 프로그램이다. 특히 다양한 다단계 일반화모형 등에 기반한 독창적인 통계기법과 계산 알고리즘을 구현하고, 데이터 분석결과에서 직관적으로 데이터의 의미를 알 수 있도록 일목요연하게 시각화하는 등 기존 기술과 차별화했다. 그러나 GUI 환경 등 여전히 사용자의 편의성을 높이기 위한 추가적인 후속 연구가 필요한 상태였다.

이에 서울대학교에서는 사회과학대학, 자연과학대학, 자유전공학부 등 다양한 계열의 학과가 참여한 융·복합 연구를 통해 사회과학, 의학 그리고 일반 사용자들의 친화성을 높여 SRC-STAT를 통계패키지의 한국의 대표적인 통계패키지로 성장시키기 위한 기초연구를 수행하였다. 사실 전 세계적으로 융합과 학제간 연구가 빠르게 증가하고 있고(Leahy & Moody 2014), 인문사회과학적 상상력과 사고력이 모든 분야에 중요해지면서 인문학에 대한 대중적인 관심이 고조되고 있지만, 실제 대학과 학계에서 융합은 인문사회과학 전공자에게 상당한 도전과 위기감을 심어주고 있다. 이에 비해 통계방법론에 대한 협동연구와 관련 패키지 개발을 위한 학제간 융합은, 여러 학문분야에서 경험연구를 수행하는 연구자들이 각기 다른 이론, 방법 및 절차, 연구자료들을 적용할지라도 STAT, SPSS, R과 같은 범용통계패키지를 공통적으로 사용하기 때문에 가능한 분야이다.

범용통계패키지는 통계적 전문지식을 갖춘 통계학자와 소프트웨어 개발기술을 가진 컴퓨터공학자들만으로는 개발될 수 없다. 개발된 통계패키지를 사용하는 것은 통계적 지식과 소프트웨어에 대한 전문성을 갖고 있지 않은 일반 사용자들이기 때문에 범용성을 높이기 위해서는 무엇보다도 분석자료의 입력 및 변환, 통계분석방법의 결정과 변수의 설정, 통계분석의 실행, 분석결과 도출 및 EXCEL이나 아래 한글 등으로의 외부 반출(export) 등이 용이해야 한다. 이를 위해서는 통계패키지를 활용하여 연구를 수행하는 자연과학, 사회과학, 의학 등 다양한 학문분야의 사용자들로부터의 환류(feedback)가 뒤따라야 하고, 통계프로그램 개발자는 이를 수용하여 지속적으로 시스템을 개선할 필요가 있다.

이처럼 한국형 통계패키지 개발사업은 단기간에 완료될 수 없으며 중장기적인 비전에 따라 단계적으로 추진할 필요가 있다. 다만 본 연구는 한국형 통계패키지 개발이라는 중장기 사업의 기초연구로서의 성격을 갖고 있으며, 그 첫 단계인 사용성

평가의 결과가 주된 연구의 내용이다. 이 연구에서는 3회에 걸쳐 진행되었던 사용성 평가 중, 기존 범용통계패키지와 신규통계패키지를 대상으로 다양한 척도를 활용하여 평가한 1차와 3차 평가의 결과를 중심으로 하여 이용자들이 각기 통계분석에 활용하고 있는 기존 범용통계패키지의 사용성에 있어 어떠한 장·단점을 얘기하고 있는지, 그리고 사용자 편의성을 고려하여 새롭게 개발한 웹기반 통계패키지에 대해서는 어떻게 평가하는지 확인하였다. 그리고 신규통계패키지에 대한 사용 의향성에 미치는 조건들을 탐색해 봄으로써 한국형 통계패키지 개발을 위한 방향성을 확인하는 데 그 목적을 두고 있다.

II. 통계패키지의 사용성 평가의 기준과 평가 과정

1. 사용성 평가의 개념과 기준

이 연구에서 가장 핵심적인 내용은 새롭게 개발한 통계패키지의 사용성에 대한 평가이다. 이러한 사용성 평가는 소프트웨어의 품질과 밀접한 관련이 있다. 전기전자 기술자협회(Institute of Electrical and Electronics Engineering 1990) 표준 610.12에 따르면 소프트웨어의 품질은 시스템, 구성요소, 그리고 과정이 사용자의 요구와 기대에 부합하는 정도를 의미한다(Weheba, Attar, & Salha 2017).

아직까지 통계패키지의 품질과 관련된 명확한 기준이 부재한 상태이기 때문에 일반적인 소프트웨어에 대한 품질을 사용자 관점에서 평가하기 위한 국제표준인 ISO/IEC(International Organization for Standardization/International Electrotechnical Commission) 9126에서 제시한 기준을 살펴보면 다음과 같다. 품질평가의 기준 중 내·외부 품질은 크게 기능성, 신뢰성, 사용성, 효율성, 유지·보수성, 이식성 등 6개 품질특성과 각각의 품질특성별로 총 27개의 부특성으로 평가된다. 그리고 사용품질과 관련해서는 효과성, 생산성, 안정성, 만족도 등으로 나누어 평가된다(이하용·양해술 2013). 여기서 HCI(Human-Computer Interaction) 연구에서 핵심 개념인 사용성(usability)은 어떤 도구나 인간이 만든 물건, 서비스를 어떤 특정 목적을 달성하기 위해 사용할 때에 얼마나 사용하기 쉬운가를 말한다(주예균·김희현 2014). 특수한 환경에서 특수한 목적을 수행하는 사용자가 자신의 목적에 맞게 쉽고 효율적으

로 이용할 수 있을 때 우수한 소프트웨어일 것이다. 이에 Shackel(1991)은 사용성은 소프트웨어의 객관적 특성뿐만 아니라 사용자가 구동하는 과정에서 인식하는 주관적 측면도 매우 중요함을 강조한다.

이러한 특성을 반영하여 ISO/IEC 9126에서 제시하는 사용성은 이해성, 학습성, 운영성, 선호도, 준수성 등과 같은 부특성으로 나뉘는데, 연구자에 따라 사용성의 부특성을 추가하거나 변형하여 활용하고 있다. 대표적으로 닐슨(Nielsen 1994)이 인터랙션 설계에서 제시한 사용성을 평가하기 위한 10가지의 원칙으로 효율성, 유효성, 학습용이성, 기억용이성, 이해용이성, 피드백, 지침, 유연성, 보편성, 일관성이 있다. 다만 앞서 지적한 것처럼 현재 통계패키지의 사용성 평가를 위한 평가기준이 명확히 제시되지 않고 있기에 본 연구에서도 이러한 10가지의 원칙을 기준으로 <표 1>과 같이 변형하여 측정하였다.

이와 더불어 이 연구에서 개발하고자 했던 통계패키지는 R 프로그램을 기반으로 하되 GUI(Graphic User Interface)를 강화하여 사용자 편의성을 높이고자 시도하였다. 편정민·신동은(2007)의 연구에서는 사용자 커뮤니케이션 중심의 디자인 환경으로 변화하는 추세에 맞춰 모바일 GUI의 사용성을 다양한 각도로 분석할 수 있는 체크리스트를 제시한 바 있다. 이 연구에서 개발을 시도한 통계패키지는 원자료를 저장할 수 있는 어떠한 기기를 이용하더라도 웹을 통해 접속하여 해당 원자료를 업로드만 하면 누구나 쉽게 분석할 수 있는 통계패키지를 지향하고 있기 때문에 GUI의 요소가 매우 중요했다. 이에 본 연구에서는 편정민·신동은(2007)의 연구에서 개발한 GUI 체크리스트를 바탕으로 통계패키지에 맞도록 변형하여 측정하였다. 이와 함께 최근 통계패키지 사용자들이 정보를 효과적으로 전달하기 위해 그래프를 활용하는 경우가 많아지고 있어 통계패키지를 통해 산출되는 그래픽 기능에 대한 평가를 위한 척도도 이용하였다.

<표 1> 통계패키지 사용성 평가를 위한 척도

구분	속성	측정 문항
통계 패키지의 전반적 사용성 통계	효율성	통계분석을 위한 최소한의 단계를 거치도록 구성되어 있다
	유효성	통계분석 시 발생한 오류처리를 쉽게 할 수 있도록 지원한다
	학습 용이성	원하는 통계분석방법을 쉽게 찾을 수 있도록 설계되어 있다

패키지의 전반적 사용성	기억용이성	이전에 수행한 방법과 동일한 과정을 거쳐 분석해야 할 때 쉽게 할 수 있도록 설계되어 있다
	이해용이성	해당 통계분석방법과 방법별 옵션이 어떤 것인지를 쉽게 이해할 수 있도록 구성되어 있다
	피드백	통계분석의 검정절차에 따라 메뉴나 옵션들이 구성되어 있어 조작성이 용이하다
	지침	사용자 설명서가 친절하고 유용하게 구성되어 있다
	유연성	사용자 개인 기호에 맞게 맞춤형으로 인터페이스를 변경할 수 있다
	보편성	다양한 수요자들의 요구에 맞게 통계방법이 포함되어 있다
	일관성	이 통계패키지를 사용하면 안정적이고 일관된 분석결과를 얻을 수 있다
GUI 사용성		통계방법들이 통일성 있게 잘 짜여 있다
	레이아웃	통계정보 제공의 순서가 효율적으로 배치되어 사용이 편리하다
		메뉴와 버튼이 눈에 잘 띈다
		레이아웃이 간결하고 복잡하지 않다
		타이틀만으로도 정보의 성격을 예상할 수 있다
	텍스트	텍스트의 이해를 돕기 위해 다른 디자인요소가 효율적으로 사용되고 있다
그래픽 기능 사용성		텍스트의 색상이 정보를 인지하는 데 적합하다
		통일된 스타일의 사용으로 내용의 이해가 빠르다
		원하는 그래프를 구현하기 쉽게 GUI나 명령어가 간결하다.
		다양한 그래프를 구현할 수 있도록 많은 그래프 기능들이 포함되어 있다.
	산출된 그래프의 디자인이 세련되고 깔끔하다.	
	그림판 등의 다른 프로그램을 사용하여 편집 없이 바로 논문에 수록 가능한 결과를 얻을 수 있다	

2. 사용성 평가의 과정

이 연구에서는 앞서 살펴본 척도를 중심으로 현재 사용자가 통계분석에 사용하고 있는 통계프로그램과 이 연구에서 새롭게 개발한 통계패키지를 각각 평가하도록 하였다. 사용성 평가는 총 3차례에 걸쳐 실시되었는데, 첫 번째 사용성 평가는 2019년 6월에 서울대학교 한국사회과학자료원에서 개최한 중급통계 무료공개강좌에서 새롭게 개발한 통계패키지를 시연하고 수강자 59명을 대상으로 실시하였다. 다만, 불성실 응답자 10명을 제외하고 본 연구에서는 49명의 결과만을 이용한다. 두 번째 사용성 평가는 같은 기관에서 실시한 중급통계 수강생을 대상으로 1시간 동안 통계패키지 시연 후 37명에게서 사용성에 대한 의견을 수렴하였다. 다만 두 번째 사용성 평가는 앞서 제시한 척도를 활용하지 않고 전반적인 개선의견을 수렴하기 위해 실시하였기 때문에 이 연구의 결과분석에서는 제외한다. 마지막 사용성 평가는 그동안 사용성 평가결과에서 도출된 개선의견을 바탕으로 통계패키지를 지속적으로 보완·개선한 이후에 진행되었다. 2020년 4월에 서울대학교 한국사회과학자료원에서 중급통계를 수강한 수강생을 대상으로 신규 통계프로그램에 대한 온라인 동영상 시청한 후 사용성을 평가하도록 하였으며, 총 20명이 참여하였다. 이러한 과정을 통해 실시한 사용성 평가에 참여한 평가자가 적은 것은 사실이며, 이 연구가 가지고 있는 한계이기도 하다. 다만, 통계패키지는 한글이나 MS 소프트웨어와는 달리 목적성이 특수하고 대상이 매우 한정적이기 때문에 Statgraphics이라는 통계패키지를 개발하고 사용성을 평가한 Weheba, Attar, & Salha(2017)의 연구에서도 총 73명의 대학원생을 대상으로 실시하는 등 평가자 모집에 한계가 있음은 주지할 필요가 있다.

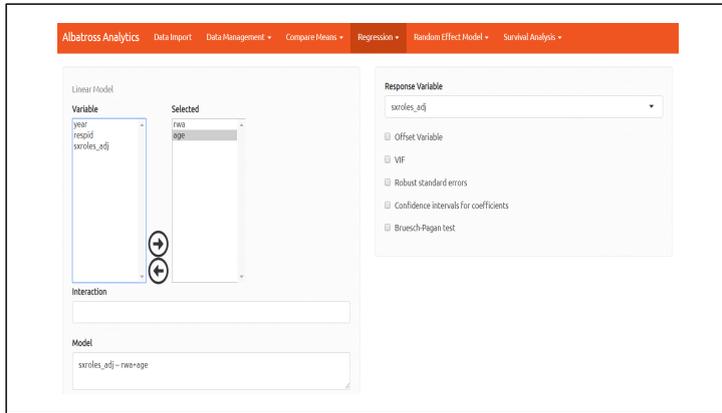
이러한 사용성 평가 과정에서 한번의 큰 변화가 있었다. 1차와 2차 사용성 평가에서는 이 연구가 시작되기 이전에 초기 개발단계였던 SRC-STAT에 대해 사회과학분야 전문가들이 운용 전반에 대해 테스트를 해보고 그 결과를 반영하여 개선한 것이었다. 다만 초기 단계에 있던 통계프로그램에서는 통상적인 통계패키지처럼 사용자가 개인용 컴퓨터에 프로그램을 설치하여 운용하는 방식이었다면, 1차와 2차 사용성 평가에서는 R-Shiny를 토대로 interactive web app로 제작된 웹기반 통계패키지로 전환하였다는 점에서 한층 진일보한 성과라 하겠다.

R-Shiny 패키지는 RStudio에서 무료로 제공하는 웹 어플리케이션 제작 프레임워크로 기존의 웹 개발 언어를 사용하지 않고 R 스크립트만으로 대화식 웹 어플리케이션을 만들 수 있다는 장점을 가지고 있다(나중화·황은지 2017). 이러한 특징 때문에 R에서 이용할 수 있는 무수히 많은 분석방법 중에 자신의 필요에 맞는 일부 분석방법만을 취사선택하여 개인형 통계패키지를 만드는 사용자들이 점차 확대되고 있다. 특히 R-Shiny 기반으로 통계패키지를 구축할 경우 기존 통계패키지처럼 별도의 프로그램 설치 과정을 거치지 않고 해당 URL에 접속만 하면 누구나 쉽게 이용할 수 있어 설치과정에서 발생하는 오류로부터 자유롭다는 장점을 가지고 있다. 이 연구를 진행하는 과정에서도 이를 고려하여 웹기반 통계패키지로 전환한 것이다. 다만 1~2차 사용성 평가에서는 테스트용으로 개발된 것이기에 서버 용량의 제한으로 다수의 접속에 한계가 있어 3차 사용성 평가에서는 대형 서버로 이전하고 통계패키지 전반을 개선한 이후에 실시하였다.

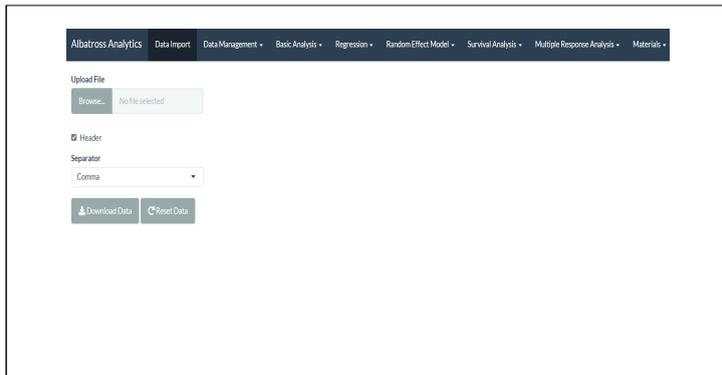
하지만 R-Shiny를 토대로 한 웹 기반 소프트웨어가 기존의 범용패키지와 비교가 가능할 것인지에 대해서는 비판의 여지가 분명하다. 만약 현재 이 연구에서 이용한 신규통계패키지가 R-Shiny에서 제공하는 분석기법들을 단순 조합한 것이라면 이러한 비판은 적절할 것이다. 그러나 이 연구에서 개발한 통계패키지는 크게 두 가지의 특성을 종합한 것이라는 측면에서 차별성을 갖는다. 먼저 자료 불러오기, 자료의 변환 및 변수의 계산, 그리고 기초통계 및 중급통계에 해당되는 기술통계량, 회귀분석, 로지스틱 분석 등은 기존의 R-Shiny에서 제공하는 것을 활용하였다. 하지만 이 연구에서 개발한 신규통계패키지는 서울대학교 통계학과 이영조 교수팀에서 오랜 개발과정을 통해 국내·외적으로 검증받은 통계기법인 HGLM 모형도 포함하고 있다. HGLM 모형은 선형 독립변수의 확률효과를 추정할 수 있도록 GLM 모형을 확장한 것으로 이번 연구에서 개발된 신규통계패키지를 통해 누구나 쉽게 이용할 수 있다. 이러한 후자의 특성이 R-Shiny 기반 웹용 패키지나 다른 범용통계패키지와 차별성이라 하겠다.¹⁾

1) 이 연구에서 이용한 신규통계패키지를 확인하고자 할 경우 다음의 링크주소를 입력하면 웹을 통해 접속 가능하다(<http://cheoling.snu.ac.kr:3838/DHGLM/>).

<1차 및 2차 사용성 평가 통계패키지 첫 화면>



<3차 사용성 평가 통계패키지 첫 화면>



<그림 1> 각 차수별 사용성 평가에 활용된 신규통계패키지

III. 통계패키지의 사용성 평가의 결과

1. 통계패키지에 대한 전반적인 사용성 평가결과

먼저, 사용성 평가의 참여자들이 현재 주로 사용하고 있는 범용통계패키지의 사용성을 평가하도록 한 1차 결과를 살펴보기로 한다. 또한 동일한 척도를 이용하여

새롭게 개발된 통계패키지에 대해서도 평가하였다(<그림 2> 참조). 물론 사용성 평가결과는 평가자가 기존에 사용하고 있는 통계패키지에 대한 충성도와 숙련도에 따라 차이가 있겠으나 제한된 표본 수로 인해 세부화할 경우 유의미한 분석이 어려워 결과의 제시에서는 생략하였다.

평가참여자들이 현재 각자 주로 이용하고 있는 통계패키지를 10개의 영역에 대해 평가한 결과를 살펴보면 각 범용통계패키지별로 가지고 있는 특성이 잘 드러난다. SPSS 이용자의 경우에는 10개의 영역 중 학습 용이성(5.2점), 일관성(5.1점)을 높게 평가하고 있었으나, 유연성(3.5점)과 유효성(3.8점)은 낮게 나타났다. SPSS 통계패키지의 경우에는 현재 명령어를 통해 프로그램을 운용할 수 있으나 대부분의 사용자들이 GUI 방식을 주로 활용하고 있기에 통계를 처음 접하는 입문자들이 많이 활용하고 있다. 현재 SPSS에서 각 통계방법별로 설정된 디폴트(default)를 변경하지 않고 일반적인 절차에 따를 경우 쉽게 분석결과를 도출할 수 있도록 설계되어 있다. 이러한 특성 때문에 실제 사용자들 역시 SPSS를 이용할 경우 보다 쉽게 학습할 수 있고 일관성 있는 결과를 얻을 수 있다고 평가하는 것으로 판단된다. 하지만 다른 통계패키지에 비해 사용자가 인터페이스를 변경하거나 분석과정에서 오류가 발생할 경우 직관적으로 처리하는 지원 체계는 미흡한 것으로 평가되고 있다.

이에 비해 STATA 통계패키지의 경우에는 10개의 영역에서 고르게 긍정적인 평가를 받는 것으로 나타났다. 이 중에서 일관성(5.9점)을 가장 높게 평가하고 있었으며, 학습용이성이나 기억용이성(각각 5.5점) 측면에서도 높은 점수를 받았다. 앞서 살펴본 SPSS와 비교했을 때 10개의 영역에서 상대적으로 높은 사용성을 인정받고 있는 것으로 나타났다. 세 번째로 R 통계패키지의 사용성 평가결과를 살펴보면 다양한 수요자들의 요구에 맞게 통계방법을 포함할 수 있다는 보편성에서 가장 높은 점수(5.8점)를 받았다. R의 경우에는 다양한 분야의 사용자들이 개발한 통계방법을 무료로 웹상에 게시하고 이용하고자 하는 사람이면 누구나 쉽게 다운로드받아 이용할 수 있도록 하고 있는데, 이러한 특성을 반영한 평가결과라 하겠다. R은 이러한 보편성이 강점이지만 이것이 또 다른 한계로 작용할 수 있는데, 사용자가 개발한 통계분석방법을 활용하는 과정에서 오류가 발생했을 경우 빠른 피드백을 받아 문제를 해결하기가 쉽지 않다. 이러한 특징 때문에 10개의 영역 중 피드백에 대한 평가가 3.2점으로 가장 낮았다. 반면, 이 연구에서 개발한 신규통계패키지는 R을 기반으로 하되 사용자로 하여금 명령어 대신에 GUI를 통해 운용할 수 있도록 설계되었기

때문에 R과 SPSS 통계패키지의 중간적인 형태라 할 수 있다. 이러한 특징 때문에 SPSS에 대한 사용성 평가처럼 학습용이성(5.2점), 효율성(5.1점), 이해용이성(5.1점), 기억용이성(5.0점) 등에서는 높은 평가를 받았지만, SPSS처럼 일반적으로 많이 활용되는 통계방법들만이 포함되어 있기 때문에 보편성(4.1점)은 낮은 평가를 받았다. 그리고 이 신규통계패키지는 R을 기반으로 하고 있기는 하지만 현재 포함되어 있는 분석방법 이외에 사용자의 편의에 따라 추가적으로 확장시킬 수 없기 때문에 R과 달리 유연성(3.9점) 역시 낮은 평가를 받았다. 반면에 R에서 상대적으로 낮은 평가를 받았던 피드백(5.0점)의 경우에는 높은 점수를 받았다.



<그림 2> 주이용 통계패키지에 대한 사용성 평가결과(1차 평가, 7점 만점)

새롭게 개발된 통계패키지만을 대상으로 했을 때 1차 평가와 3차 평가의 결과가 어떻게 달라졌는지를 비교한 결과는 <표 2>와 같다. 물론 이 결과는 사용성 평가의 대상자가 1차 평가와 3차 평가에서 같지 않고, 평가자 규모 역시 다르기 때문에 절대적인 비교는 어렵지만 상대적인 변화를 확인할 수는 있다.

<표 2> 신규통계패키지에 대한 사용성 평가결과의 비교(7점 만점)

속성	1차 평가		3차 평가	
	평균	표준편차	평균	표준편차
효율성	5.3	(0.88)	5.4	(1.47)
유효성	4.5	(1.07)	3.6	(1.43)
학습용이성	5.2	(1.17)	5.9	(1.27)
기억용이성	5.0	(1.19)	4.8	(1.68)
이해용이성	5.1	(1.23)	5.4	(1.35)
피드백	5.0	(1.12)	5.6	(1.54)
지침	4.1	(1.08)	4.8	(1.29)
유연성	3.9	(1.17)	3.7	(1.26)
보편성	4.1	(1.30)	4.7	(1.75)
일관성	4.9	(1.05)	5.0	(1.38)
표본 수(명)	49		20	

주: 표에 제시된 결과는 평균(표준편차)임.

앞서 논의한 바와 같이, 2차 사용성 평가 이후부터는 웹기반 통계패키지로 변경되었고 그 과정에서 1차와 2차 사용성 평가에서 제시되었던 평가의견을 반영하여 개선한 시스템을 대상으로 3차 사용성 평가가 이루어졌다. 그 결과 학습용이성 측면에서 긍정적인 평가가 크게 높아졌으며, 이해용이성 측면 역시 개선된 것으로 나타났다. 또한, 웹기반 통계패키지 개발 과정에서 분석방법별 예제를 사용자들이 확인할 수 있도록 하였는데 이로 인해 피드백에 대한 평가 역시 향상된 것으로 나타났다. 다만, 신규통계패키지 역시 R에 기반을 하고 있기 때문에 웹기반 통계패키지라 하더라도 오류처리에 즉각적으로 대응하는 문제(유효성)나 기존의 R이나 STATA처럼 새로운 분석방법을 사용자가 다운로드받아 활용하는 것(보편성)에서의 문제는 지속적으로 개선될 필요가 있는 것으로 나타났다.

2. 통계패키지의 GUI와 그래픽 기능에 대한 사용성 평가결과

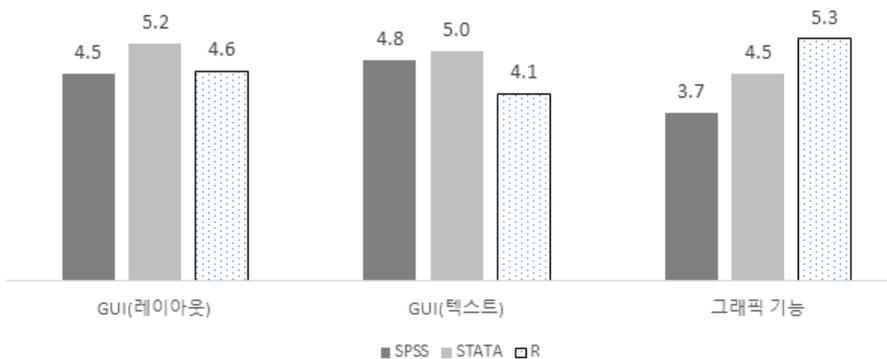
통계학의 입문자들이 통계패키지를 이용하는 방법에서 어려워하는 것이 소위 신택스(syntax)라고 하는 명령어를 이용하는 경우이다. 현재 이용자가 많은 범용통계 패키지 중에서 명령어 기반으로 구동되는 R과는 달리 SPSS와 STATA는 명령어뿐

만 아니라 GUI 방식 모두를 활용한 자료의 처리와 통계분석이 가능하다. 그럼에도 불구하고, SPSS 이용자들은 GUI 방식을 주로 이용하고 STATA 이용자들은 명령어 방식에 친숙하다. 반면 이 연구를 추진하는 과정에서 개발된 신규통계패키지는 R에 기반하고 있지만 사용자는 주로 GUI 방식으로 운용할 수 있도록 설계되어 있다.

앞서 전반적인 통계패키지의 사용성에 대해 평가한 결과를 살펴보면, 이후에는 GUI를 중심으로 사용자들이 주로 활용하고 있는 통계패키지의 GUI에 대한 사용성 평가결과를 비교해 보도록 한다. 이와 함께 최근 연구자들이 가시성이 높으면서 효과적이고 간명하게 정보를 전달하기 위한 목적으로 그래프를 많이 사용하고 있기에 각 통계패키지의 그래픽 기능성에 대한 사용성 평가결과도 살펴보고자 한다.

먼저 현재 가장 많이 활용되고 있는 SPSS, STATA, R의 GUI와 그래픽 기능에 대한 사용성 평가를 살펴보면 <그림 3>과 같다. GUI 측면에서 보면 STATA가 SPSS나 R에 비해 상대적으로 높은 평가를 받고 있는 것으로 나타났다. 특히, R은 명령어 기반이기 때문에 RStudio를 사용하지 않는 이상 GUI가 매우 빈약하고, 따라서 텍스트와 관련하여 R이 가장 낮은 평가를 받았다.

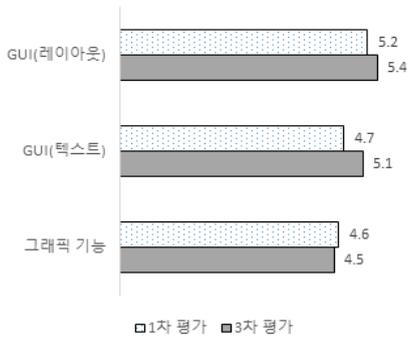
이와는 대조적으로 그래픽과 관련해서는 R이 가장 우수한 평가를 받았으며, 그 다음으로 STATA, SPSS 순이었다. 최근 빅데이터 분석에 대한 연구자들의 관심이 증대되면서 R이 제공하는 다양한 그래픽 기능들의 유용성도 주목받고 있는데 이러한 결과는 이를 뒷받침하는 결과라 하겠다. 다만 여기서 제시된 결과는 제한된 유의표집(Purposive or Judgemental Sampling)을 통해 모집된 평가자들을 대상으로 한 것이기 때문에 차이의 통계적 유의미성을 판단하기에는 한계가 있다.



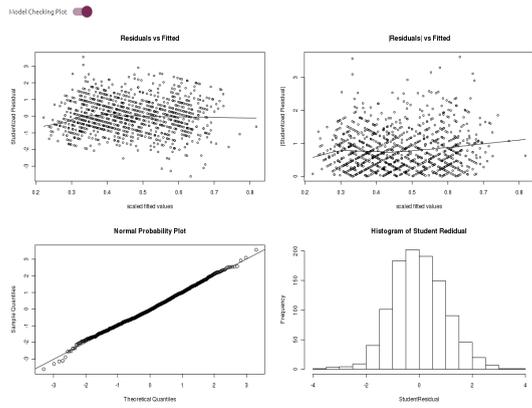
<그림 3> 주이용 통계패키지에 대한 사용성 평가결과(1차 평가, 7점 만점)

다음으로 새롭게 개발된 통계패키지의 GUI와 그래픽 기능에 대해 1차와 3차에 실시한 사용성 평가결과를 비교해보면 <그림 4>와 같다. 전반적으로 신규통계프로그램의 GUI에 대한 사용성 평가는 향상된 것으로 나타났다. 이것은 웹기반 통계패키지로 전환하는 과정에서 1차와 2차 사용성 평가결과를 토대로 사용자 편의성을 위해 GUI를 지속적으로 개선한 결과라 하겠다. 그럼에도 불구하고 그래픽 기능의 사용성에 있어서는 큰 변화가 없는 것으로 나타났다. <그림 4>의 우측에 제시된 것처럼, 각 분석방법별로 많이 활용되는 그래프를 한번의 선택으로 바로 도출될 수 있도록 구축하였음에도 불구하고, default로 설정된 그래프 이외에 추가적인 그래프를 포함시킬 수 없다는 측면이 고려된 것으로 판단된다.

<1차와 3차 사용성 평가결과>



<신규통계패키지의 그래프 사례>



<그림 4> 신규통계패키지에 대한 사용성 평가결과(7점 만점)

IV. 신규통계패키지 사용 의향성 분석

1. 분석방법과 주요 변수

앞서 살펴본 범용통계패키지와 신규통계패키지에 대한 사용성 평가 이외에 이 연구에서는 새롭게 구축된 신규통계패키지의 사용 의향에 대해서도 조사하였다. 이

에 어떤 사람들이 기존에 사용하던 통계패키지 대신에 신규통계패키지를 이용하고자 하는지를 살펴보고자 한다. 각 평가에서 동일한 척도와 관련 문항이 포함되어 있기에 1차 사용성 평가와 3차 사용성 평가에 응답하였던 전체 응답자 69명을 대상으로 분석하였다.

다만 평가자 69명을 대상으로 분석할 경우 표본의 수가 작은 것도 문제이지만, 이들이 자발적 참여로 모집된 것이기에 표본의 대표성을 담보하기 어렵다. 그렇기 때문에 기존의 확률통계분석방법을 이용하여 검증하는 데는 한계가 있다. 따라서 본 연구에서는 질적 비교 분석(Qualitative Comparative Analysis, 이하 QCA)을 이용하였다. QCA는 사례 중심적 접근인 질적 연구방법과 변수 중심적인 양적 연구방법의 장점을 적절히 결합시킨 연구방법으로, 변수의 수와 사례의 수를 기준으로 한 이차원 행렬에서 적은 사례 수를 이용하는 비교방법(comparative method) 중의 하나이다(Ragin 1987). 무엇보다도 QCA는 소위 진리표(truth tables)라는 것을 이용함으로써 인과적 복잡성을 가시적으로 표현할 수 있을 뿐 아니라 이를 통해 여러 조건들의 인과적 관계를 분석할 수 있으며, 논리적 최소화(logical minimization)의 원리를 통해 같은 결과를 공유하는 사례들 내에서의 공통성과 차이점을 규명할 수 있다는 장점을 가지고 있다(신인철 2015).

QCA 분석에 이용된 변수는 크게 결과변수와 조건변수로 나뉜다. 당연히 결과변수는 신규통계패키지의 이용 의향으로, 의향이 있는 경우 1의 값을 부여하여 더미변수화하였다. 다음으로 조건변수는 앞서 사용성 평가에 대한 결과에서 신규통계패키지가 상대적으로 우위에 있던 속성인 학습용이성, 이해용이성, 그리고 효율성을 포함하였다. 이들 변수들은 7점 척도로 측정되었는데 표준화과정을 통해 0~1의 값을 갖도록 변환하여 이용하였다. 이 외에 평가자가 현재 사용하고 있는 통계패키지의 이용기간과 통계패키지에 대한 친숙도를 고려하였다. 이용기간은 평가자에 따라 매우 다양하였는데, 초급자 수준을 약 6개월 미만으로 설정하였고 6개월 미만인 경우는 1, 그렇지 않을 경우에는 0의 값을 부여하여 더미화하였다. 그리고 현재 사용하고 있는 통계패키지에 대한 친숙도는 7점 척도로 측정되었으며 앞서 살펴본 다른 조건변수와 마찬가지로 표준화과정을 통해 0~1의 값을 갖도록 변환하여 이용하였다.

2. 사용 의향성 분석결과

QCA 방법을 이용하여 신규통계패키지의 사용 의향성을 분석한 결과는 <표 3>

에 제시된 바와 같다. 이 결과는 앞서 살펴본 조건변수들의 인과적 관계를 분석하고 논리적 최소화 원리를 적용하여 공통적인 조건결합을 도출한 결과이다. 다만 본격적인 해석에 앞서 인과 조건을 명시하기 위해서는 집합이론적 일치성과 포괄성을 확인할 필요가 있다. 여기서 일치성은 주어진 조건변수들의 조합 사례들이 논의되는 결과변수의 부분집합이 되는 정도를 의미한다(이승윤 2014). QCA를 적용한 기존 연구에서는 절대적인 기준은 아니지만 일치성의 기준으로 0.7, 0.75, 0.8 등이 이용되는데, 이 연구에서는 대표적인 QCA 전문가인 Schneider & Wagemann(2010)의 연구에 따라 0.75를 기준으로 하였다. 다음으로 고려해야 할 것은 포괄성 지수로 인과적 조합의 실증적 타당성을 나타내고 전통적인 회귀분석에서 상관관계의 강도와 유사한 개념이다. 이 연구에서는 이승윤(2014)이 사례로 적용했던 기준에 따라 0.3보다 작은 조건변수들의 조합은 제외하고자 한다. 그렇기 때문에 아래의 <표 3>에서 D(기존 통계패키지 이용 기간)은 결과의 해석에서 제외하였다.

<표 3> 신규통계패키지 사용 의향성에 대한 QCA 분석결과

조건의 조합	포괄성	일치성
s*E	0.36	0.75
s*U	0.40	0.79
s*L	0.39	0.78
L*E	0.50	0.76
D	0.04	1.00

주: 1) S: 기존 통계패키지 친숙도 / E: 신규통계패키지 효율성

U: 신규통계패키지 이해용이성 / L: 신규통계패키지 학습용이성

D: 기존 통계패키지 이용 기간

2) 여기서 소문자는 해당 조건의 값이 작음을, 대문자는 값이 큼을 의미함.

일치성과 포괄성 기준에 부합하는 조건들의 결합을 중심으로 신규통계패키지의 사용 의향성을 해석해보면 다음과 같다. 먼저, s*E, s*U, 그리고 s*L의 조건들은 공통적으로 기존 통계패키지에 대한 친숙도가 적은 경우(s)로, 이러한 평가자들은 신규통계패키지의 효율성(E), 이해용이성(U), 학습용이성(L) 때문에 신규통계패키지를 사용할 의향을 갖고 있는 것으로 나타났다. 즉, 기존에 사용하고 있는 통계패키지에 대한 친숙도가 적고 새롭게 개발된 신규통계패키지의 효율성을 높게 평가하는 사람들일수록 신규통계패키지에 대한 사용 의향성을 더 보였다. 이러한 경향은 신

규통계패키지에 대한 이해용이성이나 학습용이성에서도 유사하게 나타났다. 즉, 기존에 통계패키지를 사용하고 있다고 하더라도 해당 통계패키지에 대한 친숙도 또는 능숙도가 적은 초심자들에게 신규통계패키지가 가지고 있는 이러한 용이성과 효율성이 매력적으로 다가갈 수 있음을 보여준다.

이전의 결과에서는 친숙도가 매우 중요한 조건변수로 작용한 데 반해 L*E의 결과는 기존 통계패키지의 친숙도나 신규통계패키지의 이해용이성과 상관없이 학습용이성과 효율성만으로 사용 의향성을 설명하는 부분이다. 즉, 신규통계패키지가 학습용이성(L)과 효율성(E)을 동시에 가지고 있다고 생각하는 평가자일수록 기존의 통계패키지 대신에 이용할 의사를 보인 것으로 나타났다.

V. 결론

이 연구는 한국형 통계패키지 개발을 목표로 수행되었던 서울대학교 융·복합연구에서 설계·개발한 통계패키지에 대한 사용성 평가에 대한 탐색적 분석을 통해 통계패키지 이용자들이 자신이 주로 사용하는 통계패키지에 대해 어떻게 평가하고 있는지를 살펴보고, 신규통계패키지의 사용 의향성을 분석하는 내용으로 구성되었다.

먼저 기존 범용통계패키지에 대한 사용성 평가의 결과를 비교해보면, 사용자들은 각각의 통계패키지가 가지고 있는 특성을 고려하여 자신의 수준에서 합리적으로 선택하고 있음을 알 수 있었다. 예를 들어, SPSS 이용자들은 학습용이성이 가장 우선시 되었다면, STATA 이용자들은 일관성을, R 이용자들은 보편성을 우선시 하는 등 통계패키지마다 각기 다른 특성을 가지고 있음을 알 수 있었다. SPSS 이용자들은 유연성에, R 이용자들은 피드백 부분에서 아쉬움을 나타낸 반면 STATA 이용자들은 다른 통계패키지에 비해 전반적으로 사용성에 있어 만족도가 높은 것으로 나타났다. 이러한 결과들은 향후 통계패키지를 개선할 때 참고할 수 있는 사항이면서 통계방법을 배우고자 하는 입문자들이 통계패키지마다 가지고 있는 특성을 고려하여 자신의 목적에 맞는 소프트웨어를 선택함에 있어서도 도움이 될 것으로 기대한다.

사실 이 연구는 한국형 통계패키지 개발이라는 큰 목표를 이루기 위한 기초적인 단계의 연구성과이다. 이미 통계분석 방법론을 자신의 연구에 활용하고 있는 연구

자들은 적지 않은 기간 동안 사용해온 통계패키지가 있는데 약간의 불편함 때문에 새로운 통계패키지를 새롭게 학습한다는 것은 많은 비용을 수반하는 일이다. 그렇기 때문에 신규통계패키지를 개발한다고 하더라도 저변 확대에 대한 의문이 제기되는 것도 사실이다. 하지만 R을 제외하고 대부분의 범용통계패키지들이 적지 않은 비용을 지불해야 하는 유료 소프트웨어이고, 통계방법을 배우고자 하는 입문자나 간단한 탐색적 분석만을 도출하고자 하는 사용자의 경우에 당장의 필요가 없는 통계패키지 학습에 많은 노력과 시간을 들여야 하는 것도 사실이다. 이 연구 과정에서 구축된 신규통계패키지는 향후 지속적인 사용성 평가를 거쳐 한층 더 업그레이드하고 웹 서버를 안정화해야 하는 등 더 많은 과제가 남아 있다. 하지만 사용 의향성 평가결과를 토대로 보면, 통계방법론 입문자나 통계패키지 학습에 어려움을 겪는 이용자들에게 매우 유용할 수 있다는 가능성도 확인하였다.

이러한 유의미한 성과를 도출했음에도 불구하고, 이 연구에서 도출된 결과는 여전히 몇 가지 한계를 갖고 있다. 무엇보다도 사용성 평가의 대상자를 실험설계에 준하는 엄밀한 과정을 통해 충분한 표본을 모집하지 못한 것이 가장 큰 한계라 하겠다. 엄밀한 설계와 충분한 표본이 확보되었다면 다양한 추론통계방법도 가능했을 것이다. 또한 3차에 걸쳐 진행된 사용성 평가의 대상자가 각각 다르고 평가가 진행되는 과정에서 시스템이 개선되었고 이로 인해 평가자들의 평가결과가 어떻게 변화되는지를 확인할 수 없었다는 것이다. 또한 각 차수별 평가를 진행하면서 평가자들에게 어떠한 내용이 개선되었는지를 충분히 공지하고 이전 버전의 품질도 함께 경험할 수 있도록 했다면 좀 더 유의미한 결과를 도출했을 것이다. 다만 앞서 언급한 바와 같이 이 연구는 한국형 통계패키지를 개발하여 누구나 이용 가능할 수 있도록 공개하기 전까지 지속적으로 사용성 평가를 진행할 예정이기에 향후 평가에서는 이러한 한계를 명확히 인식하고 개선하여 유의미한 결과를 도출하는 데 반영해나갈 예정이다.

참고문헌

- 나중화·황은지. 2017. “R의 Shiny를 이용한 시각화 분석 활용 사례.” 《한국데이터정보과학회지》 28(6): 1279-1290.

- 신인철. 2015. “‘학업중단 퍼즐’을 이해하기 위한 세 접근.” 성균관대학교 박사학위논문.
- 이승희. 2000. “한국형 계절변동조정 프로그램 BOK-X-12-ARIMA.” 《응용통계연구》 13(2): 225-236.
- 이승윤. 2014. 《퍼지셋: 질적 비교연구방법론의 이론과 적용》. 고려대학교출판부.
- 이정진·강근석·이윤오·김지현·이창수·김성철. 1995. “전문가용 한국형 통계패키지 개발연구 I: 생존분석, 베이지안분석, 보험통계를 중심으로” 《한국통계학회 논문집》 2(2): 434-444.
- 이하용·양해술. 2013. “UI 중심의 소프트웨어 사용성 평가 방법.” *The Journal of Digital Policy & Management* 11(5): 105-117.
- 주예균·김희현. 2014. “모바일 쇼핑 애플리케이션 인터페이스 디자인 사용성 평가.” 《디지털디자인학연구》 14(3): 63-72.
- 편정민·신동은. 2007. “사용성 평가를 위한 모바일폰 GUI 체크리스트 개발.” 《한국디자인포럼》 15: 603-615.
- Institute of Electrical and Electronics Engineering. 1990. *Standard 610.12, IEEE Standard Glossary of Software Engineering Terminology*. New York: Institute of Electrical and Electronics Engineering.
- Leahey, E. and J. Moody. 2014. “Sociological Innovation through Subfield Integration.” *Social Currents* 1(3): 228-256.
- Nielsen, J. 1994. *Enhancing the Explanatory Power of Usability Heuristics*. Proc. ACM CHI'94 Conf. (Boston, MA, April 24-28), 152-158.
- Schneideraand, C.A. and C. Wagemann. 2010. “Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets.” *Comparative Sociology* 9: 397-418.
- Weheba, G., M. Attar, and M. Salha. 2017. “A Usability Assessment of a Statistical Analysis Software Package.” *Journal of Management and Engineering Integration* 10(2): 81-89.

<접수 2021.05.11; 수정 2021.05.11; 게재확정 2021.05.18 >

Possibility of a Korean Statistical Package Development through the Statistical Package Usability Evaluation *

Jaeyeol Yee

(Seoul National University)

Inchol Shin

(University of Seoul)

Youngjo Lee

(Seoul National University)

Cheongtag Kim

(Seoul National University)

Won-Ho Park

(Seoul National University)

Yoon Chin

(Seoul National University Asia Center)

Hyejin Kim

(Seoul National University Asia Center)

In social science research, a quantitative approach applying statistical analysis is being invigorated, and many researchers are using various general-purpose statistical packages. However, with the exception of some statistical packages, most of them are charged foreign software. And users are paying a lot of time and money for learning. Under aimed at developing a Korean statistical package, this study explored the characteristics of the general-purpose statistical package and newly developed web-based statistical package, based on the results of the usability evaluation of the statistical package. By evaluating the use intention of newly developed web-based statistical package we found that learning ease, efficiency, and ease of understanding should be considered when developing software for beginners. These results are expected to serve as a reference for other developers to improve or develop new statistical packages in the future.

Key words: statistical package, usability evaluation, qualitative comparative analysis

* This work was supported by the Seoul National University Research Grant in 2017.