

Original Article



# Investigating the Person Fit Statistics of Teenagers Participating in the Iranian University Entrance Exam 2022

Elham Pakrad<sup>1</sup> , Keyvan Salehi<sup>1,\*</sup> , Rob Meijer<sup>2</sup> , Balal Izanloo<sup>3</sup> , Ali Moghadamzadeh<sup>1</sup> 

<sup>1</sup> Faculty of Psychology and Education, University of Tehran, Tehran, Iran

<sup>2</sup> Faculty of Behavioral and Social Sciences, University of Groningen, Groningen, Netherlands

<sup>3</sup> Faculty of Psychology and Education, Kharazmi University, Tehran, Iran

## Abstract

### Article History:

Received: 20/05/2023

Revised: 01/06/2023

Accepted: 06/06/2023

ePublished: 22/09/2023

\*Corresponding author: Keyvan Salehi, Faculty of Psychology and Education, University of Tehran, Tehran, Iran.  
Email: keyvansalehi@ut.ac.ir

**Background and Objectives:** Various factors can play a role in producing aberrant response. The presence of aberrant response has a significant effect on the validity of the test. If the test scores do not have adequate validity, the score does not reflect the true latent trait of the test. Individual person fit statistics are indicators for evaluating the compatibility of test scores with a certain measurement model or with a group of examinees. This research was conducted with the aim of investigating how to compute the person fit statistics of the biology course of the 2022 university entrance exam using the PerFit package in the R software.

**Materials and Methods:** In this descriptive study, the answer sheets of 1000 candidates aged 16 to 18 years of the national exam who answered 50 items in the biology course were examined. To compute individual person fit statistics, PerFit package was used in R software. Also, a number of its functions were also examined.

**Results:** The commands for compute individual person fit statistics, and were presented and the values of these statistics were calculated.

**Conclusion:** The PerFit package was introduced as a powerful tool for measuring individual person fit and it was shown that the individual person fit statistics can be easily used for academic achievement tests using this package.

**Keywords:** Iranian university entrance exam; Person fit; Response pattern; Teenagers

**Please cite this article as follows:** Pakrad E, Salehi K, Meijer R, Izanloo B, Moghadamzadeh A. Investigating the Person Fit Statistics of Teenagers Participating in the Iranian University Entrance Exam 2022. *Pajouhan Scientific Journal*. 2023; 21(3): 186-195. DOI: 10.61186/psj.21.3.186



## Extended Abstract

### Background and Objective

Various factors can contribute to unusual test answers, including the inappropriateness of the test for the test-takers, as well as the unusual behavior of the test-takers in answering the items, such as cheating. It is crucial to determine the validity of test scores. If test scores are invalid, they do not represent the true latent trait being measured. Several person-fit statistics have been proposed to detect aberrant patterns of item scores. There are some empirical applications of person-fit statistics in related literature. However, some researchers have found it difficult to implement person-fit analysis in practice. The PerFit package discussed in this article can be useful for a wide range of professionals. This article will begin with a theoretical overview of person-fit patterns. Afterward, the introduction and classification of person-fit statistics and the review of person-fit statistics that are available in the PerFit package are discussed. In the final step, it will be shown how to use the main PerFit functions in practice. For this purpose, this study aimed to investigate the person-fit statistics in the answer sheet of the items of the biology course of the 2022 university entrance exam.

### Materials and Methods

The study utilized a descriptive method. A theoretical overview was conducted on person-fit statistics, which were then introduced and categorized. Afterward, the person-fit statistics available in the PerFit package of R software were introduced and reviewed. Using 1,000 answer sheets from the 2022 university entrance exam, researchers demonstrated how PerFit functions could be applied on a 50-item biology test. In this article, only the person-fit statistics included in the PerFit package were discussed. Most of the person-fit statistics in the PerFit package were group-based. R is a language and environment for statistical computing and graphics and provides a wide range of statistical techniques (e.g., linear and non-linear modeling, classical statistical tests, time series analysis, classification, and clustering) and graphical techniques that are highly extensible. In the following, to method of using the main PerFit functions to estimate person-fit statistics is introduced; subsequently, the software-based results are presented.

### Results

The PerFit package is available on the Comprehensive Archive Network at: <https://CRAN.R-project.org/package=PerFit>. It is recommended to upload  $N \times I$  matrices ( $N$ =test-takers,  $I$ =items) into R, with binary scores of 0 and 1 in dichotomous mode, or scores ranging from 0 to  $m-1$  in polytomous mode ( $m \geq 3$ ), depending on one's

person-fit statistics. The number of scoring options should be the same for all questions. Missing values are allowed. By default, this package includes two dichotomous datasets: InadequacyData ( $N=806$ ,  $I=28$ ) and Intelligence Data ( $N=1,000$ ,  $I=26$ ), and a polytomous dataset: PhysFuncData ( $N=714$ ,  $I=10$ ). The purpose of this paper is to demonstrate how person-fit statistics can be used to high-stak test data, and how to compute these persons fit statistics using the PerFit package. In this article, the biology score dataset ( $N=1,000$ ,  $I=50$ ) was used, which includes the answer sheets of 1,000 candidates for the 2022 university entrance exam who answered 50 biology items. At first, we converted the answer sheet into dichotomous (correct and incorrect). The data were entered into R software as a 1000x50 matrix. We showed how to use the PerFit package commands and functions to compute  $U3$ ,  $I_z^*$  and  $H^t$  statistics.

### Discussion

Assessing the fit of an individual to a research background is an active area of interest that has attracted the attention of many psychological researchers. Various research streams, such as evaluating the performance of individual fit statistics in different conditions, identifying sources of non-fit responses, and searching for guidelines for selecting appropriate individual fit statistics, continue to be favored by researchers. Some researchers have attempted to not only conduct quantitative inspections to discover potential sources of non-fit patterns but also to qualitatively examine the sources and reasons for these non-fits. A comprehensive and thorough examination of the sources of individual non-fit can provide valuable information for assessing the validity of psychological assessment tools susceptible to fake responses and also for providing educational advancement tests. As a result, in this study, several individual fit statistics on real data from the biology national exam were discussed, which are used to detect non-fit score patterns.

### Conclusion

In this article, we briefly reviewed the main concepts of measuring the person fit and provided an example of how to perform such an analysis in R software using the PerFit package. Developments in this field continue, and researchers are expanding and compiling more packages for this software. Measuring the person fit is an active research field that has attracted the attention of many psychometric researchers. Various research streams, such as evaluating the performance of individual person fit statistics in different conditions, identifying sources of aberrant responses, and seeking guidelines for choosing the appropriate individual person fit statistic, are still of interest to researchers.



## بررسی آماره‌های برازش فرد در نوجوانان شرکت کننده در آزمون سراسری سال ۱۴۰۱

الهام پاک راد<sup>۱</sup>، کیوان صالحی<sup>۱\*</sup>، راب میجر<sup>۲</sup>، بلال ایزانلو<sup>۳</sup>، علی مقدم زاده<sup>۱</sup>

<sup>۱</sup> دانشکده‌ی روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران

<sup>۲</sup> گروه روان‌سنجی و آمار، دانشکده‌ی علوم رفتاری و اجتماعی، دانشگاه خرونینگن، خرونینگن، هلند

<sup>۳</sup> دانشکده‌ی روان‌شناسی و علوم تربیتی، دانشگاه خوارزمی، تهران، ایران

### چکیده

**سابقه و هدف:** عوامل مختلفی می‌توانند در تولید پاسخ‌های نامتعارف نقش داشته باشند. وجود پاسخ‌های نامتعارف بر روی آزمون تأثیر قابل ملاحظه‌ای دارد. اگر نمرات آزمون رویی مناسب نداشته باشند، نمره نشان‌دهنده ویژگی پنهان واقعی آزمون نیست. آماره‌های برازش فرد شاخص‌هایی برای ارزیابی تطابق نمره‌ی آزمون با یک مدل اندازه‌گیری معین یا با گروهی از آزمون شوندگان هستند. این پژوهش با هدف بررسی نحوه محاسبه آماره‌های برازش فرد در پاسخنامه درس زیست شناسی آزمون سراسری سال ۱۴۰۱ با استفاده از بسته PerFit در نرم‌افزار R انجام شد.

**مواد و روش‌ها:** در مطالعه توصیفی حاضر پاسخنامه ۱۰۰۰ داوطلب ۱۶ تا ۱۸ ساله آزمون سراسری رشته تجربی که به ۵۰ سوال درس زیست‌شناسی پاسخ داده بودند، بررسی شد. برای محاسبه آماره‌های برازش فرد از بسته PerFit در نرم‌افزار R استفاده شد. همچنین تعدادی از توابع محاسباتی و گرافیکی آن نیز بررسی شد.

**یافته‌ها:** دستورات محاسبه آماره‌های برازش فرد  $U3$  و  $l_z^*$  و  $H^t$  ارائه و مقادیر این آماره‌ها محاسبه شد.

**نتیجه‌گیری:** بسته‌ی PerFit به عنوان ابزاری قدرتمند برای سنجش برازش فرد معرفی شد و نشان داده شد که آماره‌های برازش فرد را با استفاده از این بسته می‌توان به راحتی برای آزمون‌های پیشرفت تحصیلی به کار برد.

**واژگان کلیدی:** برازش فرد؛ الگوی پاسخ؛ نوجوانان؛ آزمون سراسری

تاریخ دریافت مقاله: ۱۴۰۲/۰۲/۳۰

تاریخ داوری مقاله: ۱۴۰۲/۰۳/۱۱

تاریخ پذیرش مقاله: ۱۴۰۲/۰۳/۱۶

تاریخ انتشار مقاله: ۱۴۰۲/۰۶/۳۱

تمامی حقوق نشر برای دانشگاه علوم پزشکی همدان محفوظ است.

\* نویسنده مسئول: کیوان صالحی، دانشکده‌ی روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران.  
ایمیل: keyvansalehi@ut.ac.ir

**استناد:** پاک راد، الهام؛ صالحی، کیوان؛ میجر، راب؛ ایزانلو، بلال؛ مقدم زاده، علی. بررسی آماره‌های برازش فرد در نوجوانان شرکت کننده در آزمون سراسری سال ۱۴۰۱. مجله علمی پژوهان، تابستان ۱۴۰۲؛ ۳۱(۳): ۱۹۵-۱۸۶.

### مقدمه

پیش‌بینی کرد. اگر الگوی پاسخ مشاهده شده و نمره‌ی آزمون برای یک آزمون دهنده با الگو و نمره‌ی مورد انتظار یا نمره‌ی سایر آزمون-دهندگان مطابقت نداشته باشد، الگوی پاسخ نامتعارف شکل می‌گیرد [۲]. عوامل متعدد و متنوعی، شامل نامناسب بودن آزمون برای آزمون دهندگان، رفتارهای نامتعارف آزمون دهندگان در پاسخ به سوالات نظیر تقلب و بسیاری از موارد دیگر وجود دارند که می‌توانند پاسخ‌های غیرعادی ایجاد کنند [۳].

تقلب یا پیش آگاهی از سؤال ممکن است نمرات آزمون را افزایش

پاسخ به سؤال‌های یک آزمون و نمره‌ی کسب شده می‌تواند از عوامل مختلفی، از جمله مناسب بودن آزمون برای آزمون دهنده از نظر سطح توانایی و نوع دانش و مهارتی که آزمون برای سنجش آن در نظر گرفته شده است، تأثیر پذیرد [۱]. هنگامی که یک آزمون برای آزمون دهنده‌ی با سطح مشخصی از توانایی مناسب است، الگوی پاسخ سؤال یا نمره‌ی آزمون را می‌توان با درجه‌ی خاصی از قطعیت بر اساس مدل اندازه‌گیری زیربنایی استفاده شده یا الگوهای پاسخ و نمرات دیگر آزمون دهندگان با سطح توانایی مشابه در نمونه یا جامعه

ناپارامتریک نامیده می‌شوند [۹].

پاسخ‌های نامتعارف می‌تواند به کسب نمرات کاذب بالا یا کاذب پایین برای شرکت کنندگان در آزمون منجر شود. وجود انحراف در نتایج آزمون می‌تواند نمرات آزمون را نادرست و تفسیرها و کاربردهای پیشنهادی آن‌ها را کم ارزش کند [۱۰]. میجر به پنج عامل شامل تقلب (Cheating)، خوابیدن (sleeping)، حدس زدن (guessing)، پاسخ‌گویی خلاقانه (Extremely Creative Examinees) و پاسخ‌گویی تصادفی اشاره می‌کند که می‌توانند به نامتعارفی پاسخ‌های آزمون‌دهنده به سؤال‌های یک آزمون منجر شوند و نمرات کاذب بالا یا پایین ایجاد کنند [۱].

تقلب به رفتارهایی اطلاق می‌شود که در آن، آزمون‌دهنده به‌طور نامشروع برای سؤال‌هایی که قادر به پاسخ‌گویی صحیح به آن‌ها نیست، از طریق پیش‌آگاهی از سؤال‌ها یا کپی کردن پاسخ‌های سایر آزمون‌دهندگان یا پاسخ‌های ارائه‌شده توسط معلمان، پاسخ‌های صحیح را دریافت می‌کند. در الگوی خوابیدن، آزمون‌شونده در شروع آزمون مشکل دارد و پس از سازگاری با آزمون، پاسخ برخی از سؤال‌ها آسان‌تر را که معمولاً در ابتدای آزمون می‌آیند، بررسی نمی‌کند؛ در نتیجه، نسبت پاسخ‌های صحیح در سؤال‌ها آسان به پاسخ‌های صحیح در سؤال‌ها با دشواری متوسط و زیاد، کمتر است. حدس‌زدن زمانی اتفاق می‌افتد که آزمون‌دهنده پاسخ‌های صحیح برخی از سؤال‌های آزمون را که پاسخ صحیح آن‌ها را نمی‌داند، حدس بزند. پاسخ‌گویی خلاقانه و پیچیده می‌افتد که آزمون‌دهنده با توانایی بالا به‌دلیل تفسیرهای خلاقانه و پیچیده از سؤال‌ها، پاسخ‌های نادرستی به برخی از سؤال‌های آسان بدهد [۱۱].

### آماره‌های برازش فرد

آماره‌ی برازش فرد میزان معقول بودن پاسخ‌های آزمون‌دهنده به مجموعه‌ای از سؤال‌های آزمون را اندازه می‌گیرد. همان‌طور که در [۱۱] نشان داده شده است، تعداد زیادی آماره به‌منظور شناسایی آزمودنی‌ها با پاسخ نامتعارف ابداع شده است. اکثر آن‌ها در جدول ۱ ارائه شده‌اند. جدول شامل مجموعاً ۳۶ آماره‌ی برازش فرد است که به‌عنوان ناپارامتریک یا پارامتریک طبقه‌بندی می‌شوند؛ البته در این مقاله، فقط به بررسی آماره‌های برازش فردی که در بسته‌ی PerFit موجود هستند، پرداخته شده است. بیشتر آماره‌های برازش فرد موجود در بسته‌ی PerFit مبتنی بر گروه هستند. استثناها آماره‌های لگاریتم درست‌نمایی هستند که اسنیدرز [۱۲] معرفی کرده و در اسگو [۱۳] آن‌ها را بهبود بخشیده است. در ادامه، بر آماره‌های برازش فرد و توابع مرتبط که در PerFit دسترس هستند، تمرکز خواهیم کرد.

### آماره‌های پارامتریک؛ برازش فرد مبتنی بر مدل

در این حالت مفروضات مدل بررسی می‌شوند: ۱. تک‌بعدی بودن؛ همه‌ی سؤال‌ها در آزمون عمدتاً یک ویژگی نهفته را اندازه‌گیری می‌کنند؛ ۲. استقلال موضعی؛ پاسخ به سؤال‌های مختلف مشروط به توانایی از نظر آماری مستقل هستند؛ ۳. یکنواختی توابع پاسخ سؤال؛ هر تابع پاسخ سؤال به‌طور یکنواخت غیرکاهشی است [۹]. مدل‌های پارامتریک اغلب برای توصیف داده‌ها استفاده می‌شوند.

دهد؛ در حالی که بی‌توجهی یا حدس‌زدن ممکن است نمرات آزمون را کاهش دهد. تشخیص معتبر بودن نمرات آزمون بسیار مهم است. اگر نمرات آزمون نامعتبر باشند، نشان‌دهنده‌ی ویژگی پنهان واقعی در حال اندازه‌گیری نیستند [۴]. چندین آماره‌ی برازش فرد برای تشخیص الگوهای نامتعارف نمرات سؤال‌ها پیشنهاد شده است. برخی از کاربردهای تجربی تحلیل‌های برازش فرد (Person fit) در متون مرتبط وجود دارند؛ به‌عنوان مثال [۲] و [۵] و [۶] به‌ترتیب برای کاربرد در آموزش ابتدایی، آزمون‌های سرنوشت ساز و آزمون‌های شناختی دست به تحلیل آماره‌های برازش فرد زده‌اند. با این حال، برخی از محققان و متخصصان اجرای تحلیل برازش فرد را در عمل دشوار می‌دانند. این ممکن است با پیچیدگی ریاضی برخی از آماره‌های برازش فرد و فقدان نرم‌افزاری که به متخصصان در انجام این نوع تحلیل کمک کند، توضیح داده شود [۷]. بسته‌ی PerFit [۸] که در این مقاله درباره‌ی آن بحث شده است، می‌تواند برای طیف وسیعی از متخصصان مفید باشد.

این مقاله با مروری نظری بر الگوهای برازش فرد آغاز خواهد شد؛ سپس، به معرفی و دسته‌بندی آماره‌های برازش فرد و بررسی آماره‌های برازش فردی که در بسته‌ی PerFit موجود است، پرداخته می‌شود. در مرحله‌ی نهایی نشان داده خواهد شد که چگونه می‌توان از عملگرهای اصلی PerFit در عمل استفاده کرد. بدین منظور، در این مطالعه به بررسی الگوهای برازش فرد در پاسخ‌نامه‌ی سؤال‌های درس زیست‌شناسی آزمون سراسری سال ۱۴۰۱ پرداخته شده است.

### مواد و روش‌ها

مطالعه‌ی حاضر با روش تحلیلی توصیفی انجام شد. بدین منظور، در ابتدا مروری نظری بر انواع آماره‌های برازش فرد انجام گرفت و این آماره‌ها معرفی و دسته‌بندی شدند. سپس، به معرفی و بررسی آماره‌های برازش فردی که در بسته‌ی PerFit از نرم‌افزار R موجود است، پرداخته شد. در نهایت، با استفاده از ۱۰۰۰ پاسخ‌نامه‌ی درس ۵۰سوالی زیست‌شناسی داوطلبان آزمون سراسری رشته‌ی تجربی سال ۱۴۰۱ نشان داده شد که چگونه می‌توان از عملگرهای اصلی PerFit در عمل استفاده کرد.

الگوهای پاسخ نمرات آزمون‌دهندگان ممکن است اطلاعات بیشتری را نسبت به نمره‌ی کل یک آزمون فراهم کند [۹]. یکی از رویکردهای استفاده‌شده برای مطالعه‌ی الگوهای پاسخ سؤال، تحلیل برازش فرد است که به‌طور کلی، شامل مقایسه‌ی الگوی پاسخ مشاهده‌شده‌ی آزمون‌دهنده با الگوی پاسخ مورد انتظار وی است [۲]. به‌طور کلی برای تعیین الگوی پاسخ مورد انتظار دو رویکرد وجود دارد: ۱. الگوی پاسخ مورد انتظار با استفاده از یک مدل نظری یا ریاضی، مانند مدل نظریه‌ی پاسخ سؤال تولید می‌شود که مشخص‌کننده‌ی رابطه‌ی بین فرد و سؤال‌های آزمون است؛ ۲. رویکرد دوم الگوی پاسخ مورد انتظار بر اساس الگوهای پاسخ مشاهده‌شده از همه‌ی آزمون‌دهندگان موجود در نمونه است. آماره‌های برازش فرد مرتبط با رویکرد اول و دوم، به‌ترتیب، شاخص‌های پارامتریک و

بر مدل و مبتنی بر گروه را ارائه می دهد. به طور خاص، PerFit شامل  $H^t$  است که در چندین مطالعه شبیه سازی نشان داده اند که از اکثر آماره های برازش فرد بهتر عمل می کند [۲]،  $l_z^*$  به عنوان نسخه های به روز شده از آماره ی رایج  $l_z$  و چندین آماره ی برازش فرد مناسب برای داده های چندارزشی، در این بسته موجود است [۵].

### استفاده از بسته ی PerFit

آر (R) زبان و محیطی برای محاسبات آماری و گرافیک است. پروژه ی R را از سال ۱۹۹۵ در گروه آمار دانشکده ی اوکلند (Auckland)، روبرت جنتمن (Robert Gentleman) و راس ایهاکا (Ross Ihak) (نام این برنامه از حرف اول نام این دو نفر گرفته شده است) شروع کردند و به سرعت مخاطبان زیادی یافت. R طیف گسترده ای از فنون آماری (نظیر مدل سازی خطی و غیرخطی، آزمون های آماری کلاسیک، تحلیل سری های زمانی، طبقه بندی و خوشه بندی) و فنون گرافیکی را که بسیار توسعه پذیر است، ارائه می کند. در حال حاضر، این زبان را تیمی بین المللی نگهداری می کند و داوطلبانه توسعه می دهد. R به عنوان نرم افزار رایگان تحت شرایط مجوز عمومی جنو (بنیاد نرم افزارهای رایگان) در قالب کد منبع باز است. این برنامه بر طیف گسترده ای از پلتفرم های یونیکس (UNIX) و سیستم های مشابه از جمله فری بی اس دی (FreeBSD) و لینوکس (Linu)، ویندوز و مک او اس (MacOS) کامپایل و اجرا می شود.

در ادامه، چگونگی استفاده از عملگرهای اصلی PerFit به منظور برآورد آماره های برازش فرد معرفی می شود و بعد از آن، نتایج مبتنی بر نرم افزار می آید و تفسیر می شود.

### یافته ها

#### محاسبه ی آماره های برازش فرد

بسته ی PerFit در شبکه ی جامع بایگانی (CRAN) به آدرس زیر در دسترس است:

<https://CRAN.R-project.org/package=PerFit>

و می توان از آنجا با استفاده از دستور زیر برنامه را نصب کرد:

```
R> install.packages("PerFit")
```

آماره های برازش فرد موجود در بسته ی PerFit در جدول ۱ فهرست شده اند. بسته به آماره های برازش فرد، ماتریس های  $N \times I$  با نمره های ۰ و ۱ (در حالت دوارزشی) یا نمره های  $0, 1, \dots, m-1$  (در حالت چندارزشی) باید در R بارگذاری شوند. تعداد گزینه های نمره باید برای همه ی سؤال ها یکسان باشد. مقادیر گم شده مجاز هستند. این بسته به طور پیش فرض شامل دو مجموعه ی داده ی دوارزشی (InadequacyData ( $N = 806, I = 28$ )) و IntelligenceData ( $N = 1000, I = 26$ ))، و یک مجموعه ی داده ی چندارزشی (PhysFuncData ( $N = 714, I = 10$ )) است (8). در این مقاله، از مجموعه ی داده ی ZistData ( $N = 1000, I = 50$ ) که شامل پاسخ نامی ۱۰۰۰ داوطلب آزمون سراسری سال

برای این کار، توابع ریاضی شکل بسته برای توابع پاسخ سؤال معمولاً بر اساس تابع لجستیک انتخاب می شوند. این توابع هم بر حسب پارامترهای فرد و هم پارامترهای سؤال تعریف می شوند [۱۴]. بسته های نرم افزاری متعددی وجود دارد که امکان برازش مدل های پارامتریک را با داده ها فراهم می کند. در R برخی از بسته های مفید عبارتند از: eRm [۱۵]، jirtoys [۱۶]، ltm [۱۷]، mcIRT [۱۸] و TAM [۱۹].

#### آماره های ناپارامتریک؛ برازش فرد مبتنی بر گروه

آماره های برازش فرد مبتنی بر گروه یا ناپارامتری، صرفاً به برازش مدل های ریاضی شکل بسته برای داده ها محدود نمی شوند، اگرچه فرض های تک بعدی بودن، استقلال موضعی و یکنواختی هم باید بررسی شوند [۲۰]. به طور خاص، الگوی نمره ی سؤال در صورتی که تا حد زیادی از معمول ترین رفتار پاسخ در نمونه منحرف شود، به عنوان نامتعرف علامت گذاری می شود. آماره های برازش فرد مبتنی بر گروه معمولاً با استفاده از دشواری های مرحله سؤال در حالت چندارزشی که به سادگی برابر با نسبت نمره های صحیح در حالت دوارزشی هستند، محاسبه می شوند [۵].

#### مقایسه ی عملکرد آماره های برازش فرد

با توجه به انبوهی از آماره های برازش فرد موجود در دسترس متخصصان، مهم است که بدانیم کدام آماره از سایر آماره ها بهتر است؛ یعنی کدام آماره بالاترین توان را برای تشخیص رفتار پاسخ نامتعرف دارد و در عین حال، نرخ های مثبت کاذب را تحت کنترل نگه می دارد. ارزیابی عملکرد استراتژی های آماری رقیب به طور سنتی، با استفاده از مطالعات شبیه سازی و در موارد نادر، با استفاده از داده های تجربی انجام می شود [۹].

یافته ها ممکن است با توجه به طراحی مطالعه ی شبیه سازی و اینکه کدام آماره ی برازش فرد در مطالعه استفاده شده است، متفاوت باشند [۹]؛ با این حال، یکی از نتیجه گیری های کلی تجربی این است که آماره های برازش فرد ساده ی مبتنی بر گروه مانند  $H^t$  و  $U3$  نسبت به اکثر آماره های برازش فرد مبتنی بر مدل در انواع مختلف مجموعه داده ها عملکرد بدتری ندارند و در برخی از موارد حتی بهتر عمل می کنند [۳] و [۵]. این خبر خوبی برای متخصصان است، به این معنا که: ۱. آماره های برازش فرد مبتنی بر گروه معمولاً ساده تر محاسبه می شوند؛ ۲. مدل های اندازه گیری ناپارامتری نسبت به مدل های IRT لجستیک پارامتریک محدودتر به داده های تجربی هستند [۲۰].

#### نرم افزار R

برخی از آماره های برازش فرد در R در دسترس هستند؛ مانند  $l_z$  (بسته های آی آر توپز (irtoys)، ال تی ام (ltm)، ام آی آر تی (mirt) [۲۱]، اس آی آر تی (sirt) [۲۲]، آماره های اینفیت (infilt) و اوتفیت (outfit) مبتنی بر مدل راش [۲۳]، بسته های ای آر ام (eRm)، ام آی آر تی و اس آی آر تی). PerFit مجموعه ی گسترده تری از آماره های برازش فرد مبتنی

انتساب توصیه نمی شوند؛ زیرا ممکن است به نتایج اریب برازش فرد منجر شوند. برای نسبت‌های بزرگ مقادیر گم‌شده، بهتر است آماره های برازش فرد محاسبه نشود [۲۶].

خروجی یک تابع آماره‌ی برازش فرد یک شیء (object) از کلاس "PerFit" است که از لیستی با ۱۲ عنصر تشکیل شده است:

۱. PFScores : لیستی به طول N از مقادیر آماره‌های برازش فرد؛
  ۲. PFStatistic: آماره‌های برازش فرد به کار برده شده؛
  ۳. PerfVects: پیامی که نشان می‌دهد آیا بردارهای پاسخ کامل (همه ۰ یا همه ۱) از تحلیل حذف شده‌اند؛
  ۴. ID.all0s : نشانگرهای ردیفی بردارهای حذف شده پاسخ همه-۰ از تحلیل (در صورت وجود)؛
  ۵. ID.all1s : نشانگرهای ردیفی بردارهای حذف شده پاسخ همه-۱ از تحلیل (در صورت وجود)؛
  ۶. matrix: ماتریس داده پس از تعیین مقادیر گم‌شده (در صورت وجود، در غیر این صورت ماتریس داده‌ی اصلی است)؛
  ۷. Ncat: تعداد دسته‌های پاسخ؛
  ۸. IRT.PModel: مدل پارامتری آی‌آرتی استفاده شده؛
  ۹. IP: ماتریس پارامترهای سؤال؛
  ۱۰. Ability.PModel: روشی که برای برآورد توانایی‌ها استفاده می‌شود؛
  ۱۱. Ability: بردار پارامترهای توانایی N؛
  ۱۲. NAs.method: روشی که برای مقابله با مقادیر گم‌شده استفاده می‌شود.
- به‌غیر از آماره‌هایی که در حال حاضر در بسته گنجانده شده است،

۱۴۰۱ است که به ۵۰ سؤال زیست شناسی پاسخ داده‌اند، استفاده شد.

به‌طور معمول، برای استفاده از تابع برازش فرد فقط می‌توان ماتریس داده (ماتریس آرگومان) را ارائه داد. به‌عنوان مثال، برای محاسبه‌ی مقادیر  $H^t$  برای مجموعه‌ی داده‌ی ZistData، کد زیر اجرا شد:

```
R> library("PerFit")
R> data("ZistData ", package = "PerFit")
R> Ht("ZistData")
```

تابع  $H^t$  آرگومان‌های بیشتری دارد؛ اما همه‌ی آن‌ها پیش‌فرض‌های از پیش تعریف شده دارند. به‌طور کلی، آرگومان‌های توابع برازش فرد دوارزشی به قرار زیر هستند:

```
(matrix,
  NA.method = "Pairwise", Save.MatImp = FALSE,
  IP = NULL, IRT.PModel = "2PL", Ability = NULL,
  Ability.PModel = "ML",
  mu = 0, sigma = 1)
```

و آرگومان‌های توابع برازش فرد چندارزشی به قرار زیر هستند:

```
(matrix, Ncat,
  NA.method = "Pairwise", Save.MatImp = FALSE,
  IP = NULL, IRT.PModel = "GRM", Ability = NULL,
  Ability.PModel = "EAP")
```

کاربر ممکن است انتخاب کند که ماتریس داده‌های وارد شده را در یک فایل ذخیره کند (Save.MatImp = TRUE). وقتی نسبت مقادیر گم‌شده اندک است و از مکانیسم MCAR پیروی می‌کند، رویه‌های مقادیر گم‌شده‌ی موجود در PerFit احتمالاً برای اکثر اهداف عملی مناسب هستند؛ اما در صورتی که به طرح‌های دقیق‌تری برای مقابله با مقادیر گم‌شده نیاز باشد، باید به بسته‌های تخصصی R مانند آملیا (Amelia) [۲۴] و می (mi) [۲۵] مراجعه شود. با این حال، زمانی که نسبت گم‌شده زیاد است، هر دو استراتژی حذف دوتایی و

جدول ۱: دسته‌بندی آماره‌های پارامتریک و ناپارامتریک برازش فرد

آماره های برازش فرد	توابع PerFit	نوع سؤال	رویکرد
$r_{pbis}$	$r_{pbis}$	دوارزشی	ناپارامتری
$C$	C.Sato	دوارزشی	ناپارامتری
$G, G_n$	G.Gnormed	دوارزشی	ناپارامتری
$A, E, D$	A.KB, E.KB, E.KB	دوارزشی	ناپارامتری
$U3, ZU3$	U3, ZU3	دوارزشی	ناپارامتری
$C^*$	Cstar	دوارزشی	ناپارامتری
$NCI$	NCI	دوارزشی	ناپارامتری
$H^T$	HT	دوارزشی	ناپارامتری
$G^P$	Gpoly	دوارزشی	ناپارامتری
$G_n^P$	Gnormed.poly	چندارزشی	ناپارامتری
$U3^P$	U3poly	چندارزشی	ناپارامتری
$l_z$	lz	دوارزشی	پارامتری
$l_z^P$	lzpoly	چندارزشی	پارامتری
$l_z^*$	lzstar	دوارزشی	پارامتری

```
R> set.seed(123)
R> cutoff (Ht.out, Blvl = 0.10)
$Cutoff
[1] 0.0072
$Cutoff.SE
[1] 0.0058
$Prop.flagged
[1] 0.152
$Tail
[1] "lower"
$Cutoff.CI
  2.5%   97.5%
-0.0062 0.0156
```

**تابع *flagged.resp*:**

```
flagged.resp(x, cutoff.obj = NULL, scores = TRUE, ord =
TRUE, ModelFit =
"NonParametric", Nreps = 1000, IP = x$IP, IRT.PModel
= x$IRT.PModel,
Ability = x$Ability, Ability.PModel = x$Ability.PModel,
mu = 0, sigma = 1,
Blvl = 0.05, Breps = 1000, CIlvl = 0.95, UDIvl = NA)
این تابع مشخص می کند که کدام آزمون شوندگان در نمونه
بر اساس یک نقطه‌ی برش پرچم‌گذاری آماره‌ی برازش فرد توسط یک
شده‌اند.
```

```
R> U3.out <- U3(ZistData)
R> set.seed(72)
R> U3.cutoff <- cutoff(U3.out, Blvl = .01)
R> flagged.resp(U3.out, cutoff.obj = U3.cutoff, ord =
TRUE)
```

```
$Scores
  FlaggedID It3 It21 It15 It23 ... It47 It13 It28 It27
PFscores
[1,] 91 0 0 0 0 0 0 0 0 0 0.7310
[2,] 130 0 0 0 0 0 0 1 0 0 0.7080
[3,] 345 0 0 0 0 0 0 0 0 0 0.7822
[4,] 412 0 0 0 0 0 0 0 0 0 0.7837
[5,] 421 0 0 0 0 0 0 0 0 0 0.7873
[6,] 478 0 0 0 0 0 1 0 0 0 0.7049
[7,] 507 0 0 0 0 0 0 0 0 0 0.7488
[8,] 549 0 0 0 0 0 0 0 0 0 0.7366
[9,] 589 0 0 0 0 0 0 0 0 0 0.7120
[10,] 607 0 0 0 0 0 ... 0 0 0 1 0.7374
[11,] 649 0 0 0 0 0 ... 0 0 0 0 0.7031
[12,] 683 0 0 0 0 0 ... 0 0 0 1 0.7023
[13,] 72 0 0 0 0 0 ... 0 0 0 0 0.7261
[14,] 769 0 0 0 0 0 ... 0 0 0 0 0.7873
[15,] 778 0 0 0 0 0 ... 0 0 0 0 0.7057
[16,] 797 0 0 0 0 0 ... 0 0 0 0 0.7388
[17,] 822 0 0 0 1 ... 0 0 0 0 0.7044
[18,] 864 0 0 0 0 0 ... 0 0 0 0 0.8643
[19,] 897 0 0 0 0 0 ... 0 0 0 0 0.8760
[20,] 931 0 0 0 0 0 ... 0 0 0 0 0.7313
[21,] 956 0 0 0 0 0 ... 0 0 0 1 0.7222
```

**تابع *PRFplot*:**

```
PRFplot (matrix, respID, h = .09, N.FPts = 101, VarBands
= FALSE,
  VarBands.area = FALSE, alpha = .05, Xlabel = NA,
Xcex = 1.5, Ylabel = NA,
  Ycex = 1.5, title = NA, Tcex = 1.5, NA.method =
"Pairwise", Save.MatImp =
  FALSE, IP = NULL, IRT.PModel = "2PL", Ability =
NULL, Ability.PModel =
```

تابع بیشتری نیز در دسترس هستند، به این نام‌ها:

```
.the print ,flagged.resp ,cutoff ,PerFit.SE ,PerFit.PFS
plot methods in PerFit ,summary
```

**تابع *PerFit.PFS*:**

```
PerFit.PFS(matrix, method = NULL, simplified = TRUE,
NA.method = "Pairwise",
  Save.MatImp = FALSE, IP = NULL, IRT.PModel =
NULL, Ability = NULL,
  Ability.PModel = NULL, mu = 0, sigma = 1)
```

تابع *PerFit.PFS* امکان محاسبه‌ی بیش از یک آماره‌ی برازش فرد را به‌طور هم‌زمان فراهم می کند.

```
R> PFS.out <- PerFit.PFS(ZistData, method = c("U3",
"Izstar", "Ht"), simplified = TRUE)
R> PFS.out
```

	U3	Izstar	Ht
1	0.2962	0.0578	0.1237
2	0.1859	1.5706	0.1793
3	0.0848	1.8427	0.2426
4	0.3654	0.3611	0.0649
5	0.3104	-0.5771	0.1093

تابع *PerFit.SE* خطاهای استاندارد چک‌نایف را برای نمرات آماره‌ی برازش فرد در شیء  $x$  محاسبه می کند. این فرایند شامل محاسبه‌ی آماره‌های برازش فرد برای هر مجموعه‌ی داده با حذف  $i$ امین سؤال ( $i = 1, \dots, I$ ) و به‌دنبال آن، محاسبه‌ی انحراف استاندارد چک‌نایف مقادیر  $I$  آماره‌های برازش فرد، برای هر آزمون شونده در مجموعه‌ی داده است.

```
R> Ht.out <- Ht (ZistData)
R> Ht.SE <- PerFit.SE (Ht.out)
R> Ht.S
```

```
PFscores PFscores.SE
[1,] 0.1237 0.0591
[2,] 0.1793 0.0598
[3,] 0.2426 0.0720
[4,] 0.0649 0.0459
[5,] 0.1093 0.0591
```

**تابع *cutoff*:**

```
cutoff(x, ModelFit = "NonParametric", Nreps = 1000, IP
= x$IP, IRT.PModel =
  x$IRT.PModel, Ability = x$Ability, Ability.PModel =
x$Ability.PModel,
  mu = 0, sigma = 1, Blvl = 0.05, Breps = 1000, CIlvl = 0.95,
UDIvl = NA)
```

این تابع اجازه می دهد تا مقدار یک نقطه‌ی برش را در یک سطح از پیش تعیین شده برآورد کنیم. مقادیر آماره‌ی برازش فرد در نمونه در حد یا بیشتر از مقدار نقطه‌ی برش به پرچم‌گذاری بردارهای نمرات سؤال مربوط به‌عنوان نامتعارف منجر می شود. نقطه‌ی برش مقادیر  $H^t$  محاسبه‌شده از داده‌های *ZistData* بر اساس نرخ خطای اسمی ۱۰ درصد به شرح زیر است:

```
= 1000, CIlvl =
  0.95, UDIvl = NA, Type = "Density", Both.scale =
TRUE, Cutoff = TRUE,
  Cutoff.int = TRUE, Flagged.ticks = TRUE, Xlabel = NA,
Xcex = 1.5, title =
  NA, Tcex = 1.5, col.area = "lightpink", col.hist =
"lightblue", col.int =
  "darkgreen", col.ticks = "red", ...)
```

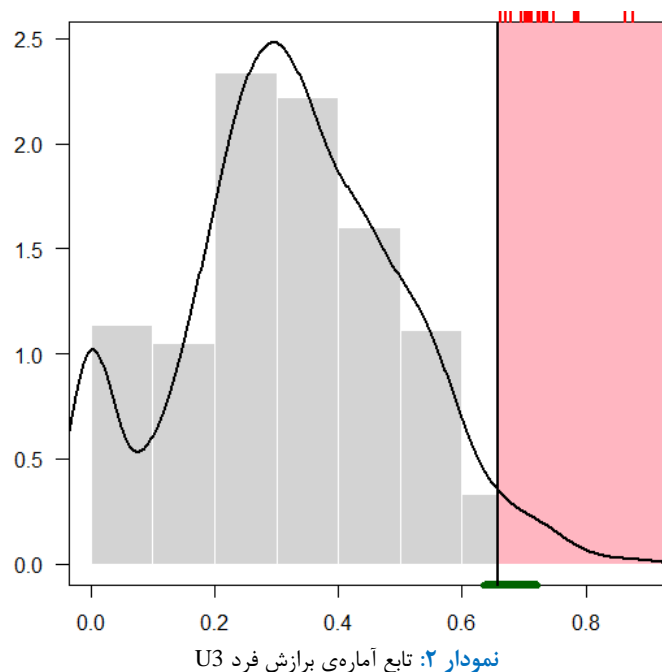
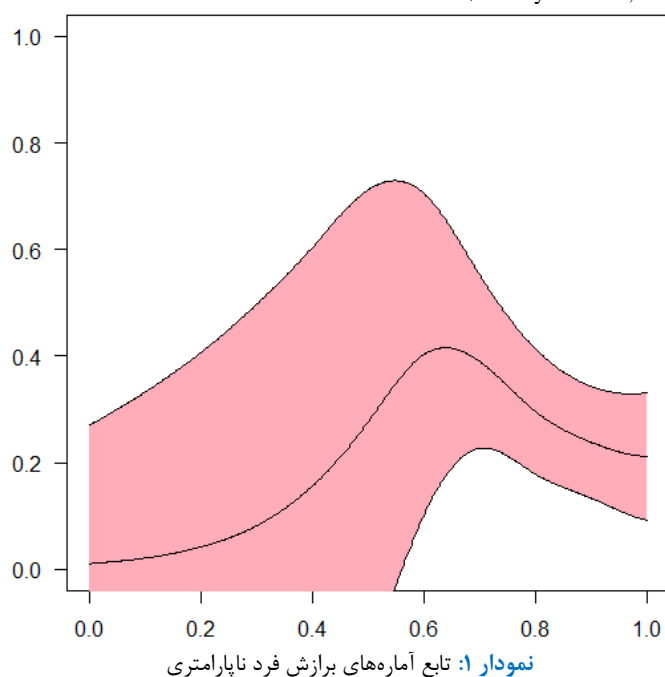
نمودار ۲ توزیع تجربی مقادیر آماره های برازش فرد مشخص شده توسط شیء x کلاس PerFit را نشان می دهد. هیستوگرام، چگالی یا ترکیبی از هر دو نمایشگر امکان پذیر است (آرگومان Type با گزینه های "Histogram"، "Density" (default) و "Both"). (نمودار ۲).  
`plot(U3poly(data, Ncat = 2), Blvl = 0.01)`

```
"ML", mu = 0, sigma = 1, message = TRUE)
تابع PRFplot آماره های برازش فرد ناپارامتری را برای همهی
آزمون دهندگان شناسایی شده در بردار respID ترسیم می کند. این
تابع فقط برای داده های دوارزشی اعمال می شود. محور طول ها مقادیر
دشواری و محور عرض ها مقادیر احتمال پاسخ صحیح را نشان می دهد
(نمودار ۱).
```

```
PRFplot(ZistData, 1000, VarBands = TRUE,
VarBands.area = TRUE, message = FALSE)
```

روش plot برای PerFit :

```
plot(x, cutoff.obj = NULL, ModelFit = "NonParametric",
Nreps = 1000, IP =
  x$IP, IRT.PModel = x$IRT.PModel, Ability = x$Ability,
Ability.PModel =
  x$Ability.PModel, mu = 0, sigma = 1, Blvl = 0.05, Breps
```



## بحث

آزمون‌های انطباقی رایانه‌ای خواهند بود. آماره‌های برازش فرد اولین گام برای ردیابی افرادی است که رفتار پاسخ‌گویی آن‌ها یا بخشی از آن، نتیجه‌ی ویژگی‌های دیگری غیر از توانایی پنهانی است که آزمون قصد اندازه‌گیری آن را دارد. تحلیل الگوهای نمره‌ی سؤال ممکن است اطلاعات بیشتری را درباره‌ی آزمون شوندگان نسبت به تحلیل نمرات آزمون نشان دهد. در این بررسی، درباره‌ی چندین آماره‌ی برازش فرد بر روی داده‌های واقعی درس زیست شناسی آزمون سراسری بحث شد که می‌توان از آن‌ها برای تشخیص الگوهای نمره‌ی نامتعارف استفاده کرد.

## تشکر و قدردانی

مقاله‌ی حاضر حاصل بخشی از رساله‌ی دکتری نویسنده‌ی اول در دانشگاه تهران است. نویسندگان بدین وسیله نهایت سپاس و قدردانی خود را از سازمان سنجش آموزش کشور ابراز می‌دارند.

## تضاد منافع

هیچ‌گونه تعارض منافع بین نویسندگان وجود ندارد.

## سهم نویسندگان

تمامی نویسندگان امور مربوط به مقاله و پیگیری‌های آن را بر عهده داشته‌اند.

## ملاحظات اخلاقی

ملاحظات اخلاقی مربوط به مقاله به طور کلی رعایت شده است.

## حمایت مالی

جهت انجام این پژوهش هیچ‌گونه حمایت مالی دریافت نشده است.

در این مقاله، به‌طور خلاصه مفاهیم اصلی سنجش برازش فرد را بررسی کردیم و یک مثال درباره‌ی نحوه‌ی انجام چنین تحلیلی در نرم‌افزار R با استفاده از بسته‌ی PerFit ارائه کردیم. تحولات در این زمینه ادامه دارد و پژوهشگران در حال گسترش و تدوین بسته‌های بیشتری برای این نرم‌افزار هستند. سنجش برازش فرد زمینه‌ی تحقیقاتی فعالی است که توجه بسیاری از محققان روان‌سنجی را به خود جلب کرده است. جریان‌های تحقیقاتی متنوعی، مانند ارزیابی عملکرد آماره‌های برازش فرد در شرایط مختلف، شناسایی منابع پاسخ نامتعارف [۲۷] و [۵] جست‌وجوی دستورالعملی برای انتخاب آماره‌ی برازش فرد مناسب [۲۸] همچنان مورد علاقه‌ی پژوهشگران است. برخی از محققان سعی کردند علاوه بر بازرسی‌های کمی برای کشف منابع بالقوه‌ی الگوهای پاسخ نامتعارف، به بررسی کیفی منابع و دلایل این عدم برازش‌ها بپردازند. بررسی همه‌جانبه و دقیق منابع ایجادکننده‌ی عدم برازش فرد می‌تواند اطلاعات ارزشمندی را برای بررسی اعتبار نمرات ابزارهای روان‌شناختی که مستعد پاسخ‌های جعلی هستند و همچنین، آزمون‌های پیشرفت تحصیلی ارائه دهد.

## نتیجه‌گیری

PerFit بسته‌ای در نرم‌افزار R است که آماره‌های پرکاربرد برازش فرد موجود را ارائه می‌دهد. برخی از روش‌هایی که احتمالاً در نسخه‌های آینده‌ی بسته قرار می‌گیرند، شامل طیف گسترده‌تری از آماره‌های برازش فرد مبتنی بر مدل و روش‌های مناسب برای داده‌های

## REFERENCES

- Meijer RR. Person-fit research: an introduction. *Appl Meas Educ*. 1996;**9**(1):3-8. DOI: [10.1207/s15324818ame0901\\_2](https://doi.org/10.1207/s15324818ame0901_2)
- Meijer RR, Egberink IJL, Emons WHM, Sijtsma K. Detection and validation of unscalable item score patterns using item response theory: an illustration with harter's self-perception profile for children. *J Pers Assess*. 2008;**90**(3):227-38. DOI: [10.1080/00223890701884921](https://doi.org/10.1080/00223890701884921)
- Karabatsos G. Comparing the aberrant response detection performance of thirty-six person-fit statistics. *Appl Meas Educ*. 2003;**16**(4):277-98. DOI: [10.1207/S15324818AME1604\\_2](https://doi.org/10.1207/S15324818AME1604_2)
- Olson JF, Fremer J. TILSA test security guidebook preventing, detecting, and investigating test security irregularities. 2013.
- Meijer RR, Tendeiro JN. The use of person-fit scores in high-stakes educational testing: how to use them and what they tell us. law school admission council. USA; 2014.
- Conijn JM. Detecting and Explaining Person Misfit in Non-Cognitive Measurement. Netherlands: Tilburg University; 2013.
- Meijer RR, Niessen AS, Tendeiro JN. A practical guide to check the consistency of item response patterns in clinical research through person-fit statistics: examples and a computer program. *Assessment*. 2016;**23**(1):52-62. DOI: [10.1177/1073191115577800](https://doi.org/10.1177/1073191115577800)
- Tendeiro JN. PerFit: Person Fit. 2021. [available from: <https://cran.r-project.org/web/packages/PerFit/index.html>]
- Rupp AA. A systematic review of the methodology for person fit research in Item Response Theory: Lessons about generalizability of inferences from the design of simulation studies. *Psychol Test Assess Model*. 2013;**55**(1):3-8.
- Santos KCP, de la Torre J, Von Davier M. Adjusting Person Fit Index for Skewness in Cognitive Diagnosis Modeling. *J Classif*. 2020;**37**:399-420. DOI: [10.1007/s00357-019-09325-5](https://doi.org/10.1007/s00357-019-09325-5)
- Meijer RR, Sijtsma K. Methodology review: evaluating person fit. *Appl Psychol Meas*. 2001;**25**(2):107-35. DOI: [10.1177/01466210122031957](https://doi.org/10.1177/01466210122031957)
- Drasgow F, Levine MV, Williams EA. Appropriateness measurement with polychotomous item response models and standardized indices. *Br J Math Stat Psychol*. 1985;**38**(1):67-86. DOI: [10.1111/j.2044-8317.1985.tb00817.x](https://doi.org/10.1111/j.2044-8317.1985.tb00817.x)
- Snijders TA. Asymptotic null distribution of person fit statistics with estimated person parameter. *Psychometrika*. 2001;**66**:331-42. DOI: [10.1007/BF02294437](https://doi.org/10.1007/BF02294437)
- Embretson SE, Reise SP. Item Response Theory. Psychology Press; 2013.
- Mair P, Hatzinger R. Extended rasch modeling: the **erm** package for the application of irt models in R. *J Stat Softw*. 2007;**20**:1-20. DOI: [10.18637/jss.v020.i09](https://doi.org/10.18637/jss.v020.i09)
- Partchev I, Maris G, Hattori T. irttoys: A collection of functions related to item response theory (IRT). 2022. [available from: <https://cran.r-project.org/web/packages/irttoys/index.html>]
- Rizopoulos D. Irtm: An R Package for Latent Variable Modeling and Item Response Theory Analyses. *J Stat Softw*. 2007;**17**:1-25. DOI: [10.18637/jss.v017.i05](https://doi.org/10.18637/jss.v017.i05)
- Reif M) 2014. (mclRT: IRT Models for Multiple Choice Items .R package version 0.41, <https://CRAN.R-project.org/package=mclRT>
- Robitzsch A, Kiefer T, Wu M. TAM: Test Analysis Modules. 2023. [available from: <https://cran.r-project.org/web/packages/TAM/index.html>]
- Sijtsma K, Molenaar IW. Introduction to nonparametric item response theory. sage; 2002.
- Chalmers RP. mirt: A multidimensional item response theory package for the r environment. *J Stat Softw*. 2012;**48**:1-29. DOI: [10.18637/jss.v048.i06](https://doi.org/10.18637/jss.v048.i06)
- Robitzsch A. sirt: Supplementary item response theory

- models. 2023. [Available from: <https://cran.r-project.org/web/packages/sirt/index.html>]
23. Wright BD, Masters GN. Computation of OUTFIT and INFIT Statistics. *Rasch Meas Trans.* 1990;**3**(4):84–5.
  24. Honaker J, King G, Blackwell M. Amelia II: A program for missing data. *J Stat Softw.* 2011;**45**:1–47. DOI: [10.18637/jss.v045.i07](https://doi.org/10.18637/jss.v045.i07)
  25. Su YS, Gelman A, Hill J, Yajima M. Multiple imputation with diagnostics ( *mi* ) in *r*: opening windows into the black box. *J Stat Softw.* 2011;**45**:1-31. DOI: [10.18637/jss.v045.i02](https://doi.org/10.18637/jss.v045.i02)
  26. Nering ML, Meijer RR. A comparison of the person response function and the lz person-fit statistic. *Appl Psychol Meas.* 1998;**22**(1):53-69. DOI: [10.1177/01466216980221004](https://doi.org/10.1177/01466216980221004)
  27. Cui Y, Mousavi A. Explore the usefulness of person-fit analysis on large-scale assessment. *Int J Test.* 2015;**15**(1):23-49. DOI: [10.1080/15305058.2014.977444](https://doi.org/10.1080/15305058.2014.977444)
  28. Tendeiro JN, Meijer RR, Schakel L, Maij-de Meij AM. Using cumulative sum statistics to detect inconsistencies in unproctored internet testing. *Educ Psychol Meas.* 2013;**73**(1):143-61. DOI: [10.1177/0013164412444787](https://doi.org/10.1177/0013164412444787)