15(57), Autumn 2024, 7-31 ISSN: 2252-004x
eISSN: 2470-6240 **Article Type:** Reserch Article
Received in revised form June 07, 2023 **Received** December 15, 2022
Accepted July 17, 2023 **Published** Online October 06, 2024

Educational Measurement
https://jem.atu.ac.ir/

ATU PRESS

The Risk of the Metropolis-Hastings Robbins-Monroe Algorithm in Multidimensional Multidimensional Models of Item-Response Theory Considering the Role of Missing Data

Mehdi Molaei Yasavoli *
Corresponding Author, PhD Student in Measurement and Assessment, Allameh Tabataba'i University, Tehran, Iran. E-mail: molaei.atu92@yahoo.com

Ali Delavar
Professor, Department of Measurement and assessment, Allameh Tabataba'i University, Tehran, Iran. E-mail: dr.delavarali@gmail.com

Mohammad Asgari
Associate Professor, Department of Measurement and Assessment, Allameh Tabataba'i University, Tehran, Iran. E-mail: drmasgari423@gmail.com

Jalil Yonesi
Associate Professor, Department of Measurement and Assessment, Allameh Tabataba'i University, Tehran, Iran. E-mail: jalilyounesi@gmail.com

Vahid Rezaei Tabar
Associate Professor, Department of Statistics, Allameh Tabataba'i University, Tehran, Iran. E-mail: vhrezaei@gmail.com

Abstract

The efficiency and accuracy of parameter estimation constitute one of the most vital psychometric matters in behavioral science measurements. In the realm of item-response theory models, the presence of diverse algorithms such as MHRM, together with their application in tests featuring missing data, present considerable obstacles within the field. The primary objective of this study was to explore the hazards associated with the MHRM algorithm in multidimensional models of item-response theory, specifically in scenarios involving polytomous data, by taking into account the nature and extent of missing data. The methodological approach employed in this research endeavor was experimental, utilizing a multi-group post-test design. A study sample was fabricated through simulation studies conducted under diverse conditions for 27 independent variables, with 100 replications executed for each instance. The model utilized in this study was a multidimensional scaled response model, and the parameters scrutinized comprised the slope and threshold values of the questions. To generate and scrutinize the data, R statistical software was utilized. The outcomes of the study unveiled that the MHRM algorithm exhibits a decreased inferred risk as compared to both EM and MCEM algorithms. The results also indicated that a statistically significant difference exists in the risk of slope and threshold parameters across the three distinct mechanisms of missing data. However, no significant variation was detected in relation to the independent variable of missing data. Furthermore, a significant interaction was observed between the type of algorithm and the missing mechanism, thereby revealing the optimal performance of the MHRM algorithm. Consequently, as the algorithm is employed, a decrease in the mean and variance of the MSE slope and threshold parameters occurs in all three loss mechanisms, resulting in their convergence. Consequently, it can be strongly advocated that the application of the MHRM algorithm is indispensable within data exhibiting high rates of lost data and assorted loss mechanisms. Researchers are, therefore, counseled to employ the MHRM algorithm in data analysis involving intricate structures such as high data loss and diverse missing patterns.

Keywords: MHRM Algorithm, risk, multidimensional multidimensional models, item-response theory, missing data

Cite this Article: Molaei Yasavoli, M., Delavar, A., Asgari, M., Yonesi, j., & Rezaei Tabar, V. (2024). The Risk of the Metropolis-Hastings Robbins-Monroe Algorithm in Multidimensional Multidimensional Models of Item-Response Theory Considering the Role of Missing Data. *Educational Measurement*, *15*(57), 7-31. https://doi.org/10.22054/jem.2023.65417.3334

Publisher: Allameh Tabataba'i University Press
DOI: https://doi.org/10.22054/jem.2023.65417.3334

Extended Abstract

Introduction

The problem of insufficient data estimation and analysis methodologies in the presence of missing data, particularly when dealing with a high dimension of data points, is considered one of the most pervasive issues within the field of data analysis. Noteworthy researchers such as Finch 88888,aaaa ş 77717,Koo & Snnnnnnnnm,Cai ()))) , aaaa ooooœ & Muthén (2012), and Lesaffre & Spiessens (2001) have all highlighted this persistent problem. Given that many statistical assessments necessitate the availability of entire data sets (Schouten, Lugtig, & Vink, 2018), missing data presents a widespread and pervasive challenge in scientific research domains, such as cognitive and affective evaluations (Finch, 2008). The prevalence of missing data in the dataset leads to significant drawbacks, including biased parameter estimations, inflated standard errors, information loss, and compromised generalizability of findings (Dong & Peng, 2013; Finch, 2010).

To effectively apply methodologies pertaining to missing data, a thorough and valid assessment of their performance is of utmost importance. In light of the mounting emphasis on missing data methodologies within scientific research, it becomes crucial to identify under what specific conditions a particular technique can and should be employed (Schouten, Lugtig, & Vink, 2018). The issue of missing data becomes increasingly significant when the number of dimensions in psychological measurements escalates. The "curse of dimensionality," or the presence of high data, as proposed by Richard Bellman in 1957, poses one of the most challenging issues in the field of measurement. It represents the exponential increase in problem complexity as the number of variables (or dimensions) expands (Kuo & Sloan, 2005). The notion of the "curse of dimensionality" pertains to a range of phenomena that emerge during data analysis and organization in high-dimensional spaces, across disciplines such as numerical analysis, machine learning, and data mining. Specifically, in the context of test design (especially within the realm of human traits studies), the increase in dimensions is associated with magnified computational complexity and diminished estimation accuracy (Liu & Pierce, 1994; Cai, 2010).

Over the past years, multiple methods have been devised to enable more efficient and effective estimation of item-response theory models with high dimensions. Notably, several methodologies that have been

employed for decades encompass: (1) marginal maximum likelihood based on the maximum expectation algorithm, (2) the Monte Carlo maximum expectation algorithm (MCEM), and (3) full Bayesian estimation through simulations. An additional estimation technique, known as Markovian chain Monte Carlo (MCMC), has been proposed by researchers in the field. According to existing literature, researchers have noted disadvantages concerning the inaccuracies associated with parameter estimation when implementing each of these algorithms, highlighting their limitations in achieving precise results.

The EM algorithm relies on fixed Gaussian-Hermite quadrature points for estimation purposes. However, when the number of dimensions escalates to four or five (or higher), the estimation process encounters an issue (Liu & Pierce, 1994; Naylor & Smith, 1982). When implementing the Monte Carlo maximum expectation algorithm (MCEM), the simulation size must be considerably enlarged, particularly during the latter iterations, to achieve parameter convergence. This is due to the parameter estimates' increased closeness to the maximum likelihood function (Cai, 2010). A notable challenge encountered in the Markovian chain Monte Carlo (MCMC) algorithm pertains to the meticulous specification of a suitable proposed distribution, which may involve trial-and-error experimentation (Bashkov & DeMars, 2017).

In an effort to circumvent the limitations observed in prevalent estimation methodologies within multidimensional item-response theory models, researchers have developed the Metropolis-Hastings-Robins-Monroe (MHRM) algorithm. This approach combines the Metropolis-Hastings sampling (Patz) algorithm with Robbins-Monro sampling (Robbins, Monro, 1951) within the framework of maximum likelihood estimation (Cai, 2010). The Metropolis-Hastings-Robbins-Monroe (MHRM) algorithm serves as a composite technique grounded in the principles of marginal probability, wherein stochastic approximations are employed to integrate samples of the distribution of the latent variables (Cai, 2010).

Contrasting with the Markovian chain Monte Carlo (MCMC) approach, which approximates the overall latent distribution, the Metropolis-Hastings-Robbins-Monroe (MHRM) algorithm concentrates on point estimates and standard errors. This feature enables MHRM to achieve convergence at a markedly accelerated pace compared to MCMC. Notably, within the MHRM algorithm, the role

of the Markov chain transitions is not to facilitate an accurate approximation of the level to be achieved, but rather to merely determine the directional changes throughout the iterative process (Cai, 2010).

In light of the existing research and the current challenges, the purpose of this study is to examine the multi-dimensional item-response theory models using the MHRM algorithm in conjunction with the EM and MCEM algorithms. This investigation aims to shed light on a foundational issue within this domain. The investigation scrutinized the performance of the MHRM algorithm relative to the established conventional methods, namely EM and MCEM algorithms, with respect to both the type and quantity of missing data.

Literature Review

Research on the efficacy of the MHRM algorithm has been relatively minimal in quantity thus far. Yang & Cai (2014) conducted a study iilled “eiii nnnnnn nn aacggddddd effecss uuuuuh iiiii ii ar lll iieeeel latent variable modeling with algommmmmhhhhh oocss nn eaaa ccnng eee efficiency of parameter estimation in the context of multi-level hidden variable models. In this study, the MHRM algorithm was employed with the aim of maximizing marginal probability estimation of background effects. The results indicate that the MHRM algorithm can produce estimates and standard errors efficiently.

In a reeeacch ddddy iilled “Crrrrr rrrr rr iiii nnnnnnnn sssss rrr a Multi-lllllllllll ll aeed Reeeeeee eeee ”,” ccccc ccc by ooo & Sheng (2016), the effectiveness of various algorithms in the context of the multi-dimensional graded response model within the domain of item-response theory was put to the test. Within the scope of this study, EM and MHRM estimation methodologies were analyzed under diverse data quality conditions to appraise the abilities of each estimation method and draw meaningful conclusions regarding their relative efficacy.

Bashkov & DeMars (2017) conducted a research study titled “ff ccnnncy Tett ff MMMM MMMmmmmmm iiii ggggggg tttt i-Dimensional Multi-Level eeee ””,” ”””””” h aiddd oo eaaeeeee hhe efficiency of the MHRM algorithm in dichotomous data within the framework of a three-parameter model. The research scrutinized the capabilities of the MHRM algorithm in retrieving item parameters, variance, and hidden covariance, and assessed its efficiency in

estimating abilities at the level of both individual subjects and cluster groups (e.g., schools). The findings revealed that the MHRM algorithm demonstrated exceptional proficiency in estimating item parameters, individual subject data, and group-level parameters within a multi-dimensional multi-level modeling framework.

Methodology

The present research adopted an experimental approach, seeking to compare the phenomenon under examination. In this study, a factorial research design of the three-way type was applied, wherein the independent variables are as follows: type of algorithm, type of missing data, and quantity of missing data.

In this research, in order to compare the risk level of MHRM algorithm, Monte Carlo simulated data has been used. According to the role of missing data, the simulation study in this research generally had four stages:

1. The first step was to simulate a complete multivariate data set, representing the target population.
2. Subsequently, the same data set was subjected to deliberate removal based on the specified type and amount of missing data.
3. In the third stage, missing data were estimated using diverse statistical methods.
4. In the final phase, statistical inferences were drawn based on the obtained parameters.

By comparing the statistical inferences derived from various missing data scenarios, it is possible to assess the performance of missing data and evaluate the power of different statistical techniques.

In the first stage of this study, the multi-dimensional graded response model was employed to account for the specificities of the data (polytomous in nature). The model incorporates two parameters: slope and threshold. In the data generation process, the distribution of subject abilities was based on a normal distribution with a mean of zero and a standard deviation of one.

When generating the data, the slope parameter values were derived from a log-normal distribution, with a mean of 0.3 and standard deviation of 0.2. Additionally, the difficulty parameter was considered based on a normal distribution with a mean of zero and a standard deviation of one. Given the high correlation and multiple collinearity

issues, the average risk of these coefficients was taken into account. Considering the multi-dimensional nature of the response pattern, there were 5 slopes and 4 thresholds, respectively.

The statistical population and sample size in this research were carefully determined based on a multitude of conditions influenced by the algorithm type, missing data type, and amount of missing data. In this section, three types of missing data (MAR, MNAR, and MCAR) with varying percentages of missing data (10%, 50%, and 90%) were intentionally induced in the data. These sets of data were subsequently subjected to parameter estimation utilizing three distinct methods: EM, MCEM, and MHRM algorithms.

This research examined missing data across three levels based on the number of subjects. A 5-dimensional model, widely used in various psychological tests (such as the 5-factor personality model), was utilized, and sample sizes of 1000 data points were adhered to consistently for all conditions.

To address the research questions effectively, a three-way factorial variance analysis was employed, focusing on the risk associated with the parameters (slope and threshold of questions). The data generation and analysis process made use of various statistical software packages, including R mirt, interactions, car, and psych.

Discussion

The results revealed that in at least some comparisons between the levels, all three variables—namely, the algorithm type, the type of missing data, and the amount of missing data—were shown to significantly impact the square root of the average error in parameter estimation. Based on the comparisons between the MHRM algorithm and EM and MCEM algorithms, the results indicated that the MHRM algorithm proved to be significantly more accurate and efficient compared to the other two methods. In other words, the MHRM algorithm demonstrated lower risk in parameter estimation across all assessed scenarios. Interestingly, within the context of slope parameter estimation, the EM algorithm was found to be considerably more fallible and error-prone than both the MCEM and MHRM algorithms.

Examining the impact of different types of missing data patterns on parameter estimation, it was observed that the MCAR mechanism was shown to carry considerably less risk in terms of threshold parameter estimation, in comparison with both the MAR and MNAR methods. It

is noteworthy to highlight the significant influence of the interaction between the type of algorithm and the type of missing data mechanism on the overall performance and accuracy of the parameter estimation process.

Investigating the interplay between the type of algorithm and the type of missing data mechanism, it was noted that when the MHRM algorithm was employed, there was a subtle distinction in performance amongst the three types of missing data patterns. Crucially, the overall effect size of the estimated risk levels decreased significantly, and the error values across all three modalities began to converge. The same concept can be observed with the MCEM algorithm, wherein a notable distinction amongst the three types of missing data mechanisms is seen, with the MNAR pattern exhibiting the highest error magnitude. This phenomenon is particularly apparent in the estimation of the threshold parameter, yielding larger effect sizes. However, in relation to the slope parameter, the effect size is considerably smaller.

Another important aspect scrutinized in this study was the impact of the amount of missing data on the risk level in estimating both the slope and threshold parameters, alongside the interactive effects of this factor with the algorithm type and the method of handling missing data. Regarding the evaluation of different missing data methods, the results indicated that there was no substantial variance in performance when assessing the threshold and slope parameters across various levels of data absence.

Moving forward, in the subsequent evaluations, the results of the interaction between the amount of missing data and two other independent variables revealed significant findings. Specifically, the interaction between the amount of missing data and the type of algorithm was observed, with the values for all three levels being very similar and no significant difference observed when using either the MHRM or MCEM algorithms. Interestingly, when considering the EM algorithm, there is indeed a noticeable difference across the different levels of missing data, with higher amounts exhibiting heightened risk.

Conclusion

Based on the study findings, it appears that the MHRM algorithm, which utilizes a combination of Metropolis-Hastings (MH) sampling and Robbins-Monroe (RM) sampling, effectively enhances the maximum likelihood estimation. The key difference between the

MHRM and MCMC algorithms lies in the fact that the former focuses specifically on point estimates and standard errors, while the latter aims to approximate the entire posterior distribution. This targeted approach of the MHRM algorithm results in much faster convergence compared to both MCMC and EM. Indeed, it is important to note that since the MHRM algorithm takes advantage of Markov chain jumps to determine the direction of change during iterations, the convergence process occurs much faster and with greater accuracy compared to alternative methods. Based on the study findings, it is evident that the MHRM algorithm tends to exhibit lower risk when taking into account intervention factors like the missing data mechanism and type. Consequently, for analyzing complex data that commonly appear in behavioral science, it is highly recommended to use the MHRM algorithm to ensure less error in parameter estimation.

اندازه‌گیری تربیتی
https://jem.atu.ac.ir/

دوره ۱۵، شماره ۵۷، پاییز ۱۴۰۳، ۷-۳۱
eISSN: 2470-6240 ISSN: 2252-004x

نوع مقاله: مقاله پژوهشی
تاریخ دریافت: ۱۴۰۱/۰۹/۲۴؛ تاریخ بازنگری: ۱۴۰۲/۰۳/۱۷
تاریخ پذیرش: ۱۴۰۲/۰۴/۲۶؛ تاریخ انتشار: ۱۴۰۳/۰۷/۱۵

مخاطره الگوریتم متروپلیس هستینگز رووینز مونرو در مدل‌های چندارزشی چند بعدی نظریه سؤال پاسخ با در نظر گرفتن نقش داده‌های گمشده

مهدی مولایی یساولی* *نویسنده مسئول*، دانشجوی دکتری سنجش و اندازه‌گیری (روان‌سنجی)، دانشگاه علامه طباطبائی، تهران، ایران. رایانامه: molaei.atu92@yahoo.com

علی دلاور استاد تمام گروه سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی، تهران، ایران. رایانامه: dr.delavarali@gmail.com

محمد عسگری دانشیار گروه سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی، تهران، ایران. رایانامه: drmasgari423@gmail.com

جلیل یونسی دانشیار گروه سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی، تهران، ایران. رایانامه: jalilyounesi@gmail.com

وحید رضایی تبار دانشیار گروه آمار، دانشگاه علامه طباطبائی، تهران، ایران. رایانامه: vhrezaei@gmail.com

چکیده

کارایی و خطای برآورد پارامترها، در اندازه‌گیری‌های علوم رفتاری یکی از مهم‌ترین موضوعات روان‌سنجی است. وجود الگوریتم‌های گوناگون مانند MHRM و کاربرد آن‌ها در آزمون‌های دارای داده گمشده، یکی از چالش‌های موجود در حوزه مدل‌های نظریه سؤال پاسخ است. هدف این پژوهش بررسی مخاطره الگوریتم MHRM در مدل‌های چندبعدی نظریه سؤال پاسخ در داده‌های چند ارزشی با در نظر گرفتن مکانیسم و میزان داده گمشده متفاوت، بود. روش پژوهش مورداستفاده آزمایشی و با استفاده از طرح پس‌آزمون چند گروهی بود. نمونه موردمطالعه بر اساس مطالعات شبیه‌سازی تحت شرایط مختلف متغیرهای مستقل (نوع الگوریتم، نوع داده گمشده و میزان داده گمشده) در ۲۷ حالت با ۱۰۰ تکرار برای هرکدام، ایجاد شد. مدل مورداستفاده مدل پاسخ مدرج چندبعدی و پارامترهای موردبررسی شیب و آستانه سؤالات بود. جهت بررسی مخاطره هر یک از پارامترها در حالت‌های مختلف آزمایشی شاخص میانگین توان دوم خطاها (MSE) مورداستفاده قرار گرفت. جهت تولید و تحلیل داده‌ها از نرم‌افزار آماری R استفاده شد. نتایج پژوهش نشان داد الگوریتم MHRM در قیاس با الگوریتم‌های EM و MCEM دارای مخاطره برآورد کمتری است. همچنین نتایج نشان داد که در میزان مخاطره پارامترهای شیب و آستانه، بین سه مکانیسم متفاوت داده‌های گمشده تفاوت معنی‌داری وجود دارد ولیکن در رابطه با متغیر مستقل میزان داده‌های گمشده، تفاوت معنی‌داری مشاهده نشد. همچنین بین نوع الگوریتم و مکانیسم گمشدگی نیز تعامل معنی‌داری وجود داشت که حکایت از عملکرد مطلوب الگوریتم MHRM داشت. درنتیجه زمانی که از این الگوریتم استفاده می‌شود، میانگین و واریانس MSE پارامترهای شیب و آستانه در هر سه مکانیسم گمشدگی، هم‌زمان که کاهش می‌یابند، به یکدیگر نزدیک نیز می‌شوند. پس می‌توان گفت کاربرد الگوریتم MHRM در داده‌های با میزان داده گمشده بالا و انواع گمشدگی، ضروری است؛ بنابراین، به پژوهشگران توصیه می‌شود که از الگوریتم MHRM در تحلیل داده‌های با ساختار پیچیده از قبیل میزان داده گمشده بالا و انواع مکانیسم گمشدگی بهره گیرند.

کلیدواژه‌ها: الگوریتم MHRM، مخاطره، مدل‌های چند ارزشی چندبعدی نظریه سؤال پاسخ، داده‌های گمشده

استناد به این مقاله: مولایی یساولی، مهدی، دلاور، علی، عسگری، محمد، یونسی، جلیل، و رضایی تبار، وحید. (۱۴۰۳). مخاطره الگوریتم متروپلیس هستینگز رووینز مونرو در مدل‌های چندارزشی چند بعدی نظریه سؤال پاسخ با در نظر گرفتن نقش داده‌های گمشده. *فصلنامه اندازه‌گیری تربیتی*، ۱۵(۵۷)، ۷-۳۱.
https://doi.org/10.22054/jem.2023.65417.3334

ناشر: دانشگاه علامه طباطبائی

مقدمه

علوم رفتاری با وجود انبوهی از متغیرهای پیچیده، یکی از حوزه‌های مطالعاتی جذاب در سال‌های گذشته بوده است که البته همراه با چالش‌ها و دشواری‌های فراوانی بوده است. یکی از مهم‌ترین مشکلات در حوزه تحلیل داده‌ها، ناکارآمدی روش‌های برآورد و تحلیل داده‌ها در صورت موجود بودن داده‌های گمشده (ccccch, 8888; Aşşşş , 7717) و وجود تعداد ابعاد بالا است (Kuo & Sloan, 2005; Cai, 2005; Asparouhov & Muthén, 2012 & Lesaffre & Spiessens, 2001). اکثر داده‌های پژوهشی در روان‌شناسی، جامعه‌شناسی و تعلیم و تربیت از موضوعات انسانی جمع‌آوری می‌شود و داده‌های ازدست‌رفته یکی از مهم‌ترین مشکلات پژوهشگرانی است که در این زمینه‌ها کار می‌کنند (Akşşş , 7717). ازآنجاکه اکثر تحلیل‌های آماری به داده‌های کامل نیاز دارند (Schouten et al., 2018)، داده‌های ازدست‌رفته یک مشکل رایج و فراگیر در پژوهش‌های علمی از قبیل ارزیابی‌های شناختی و عاطفی (Finch, 2008) است. داده‌های ازدست‌رفته معمولاً در شرایطی اتفاق می‌افتد که شرکت‌کنندگان به دلیل عدم آگاهی، تردید و عدم انگیزه و غیره به برخی از سؤالات در فرآیند جمع‌آوری داده‌ها پاسخ نمی‌دهند (Enders, 2010). وجود مشاهدات گمشده در داده‌ها منجر به مشکلاتی مانند برآورد پارامترهای مغرضانه، تورم خطاهای استاندارد، از دست دادن اطلاعات و تعمیم‌پذیری ضعیف نتایج می‌شود (Dong et al., 2013). در زمان تعیین یک روش مناسب برای مدیریت مشاهدات گمشده، محققان باید به میزان و مکانیسم داده‌های ازدست‌رفته دقت کنند (Enders, 2010). به‌طورکلی سه مکانیسم مهم در رابطه با گمشدگی داده‌ها ارائه شده است: ۱) گم‌شدن تصادفی[1]: در این حالت گم‌شدگی در متغیر دارای مقادیر گمشده به متغیری که برای همه افراد شرکت‌کننده در مطالعه به‌طور کامل مشاهده شده است، دارد ولی به خود متغیر دارای مقادیر گمشده بستگی ندارد. ۲) گم‌شدن تصادفی کامل[2]: در این حالت گم‌شدگی در متغیر دارای مقادیر گمشده، به هیچ‌یک از متغیرهای دیگر و خود متغیر دارای مقادیر گمشده بستگی ندارد؛ بنابراین در این حالت تحلیل بر اساس داده‌های مشاهده شده، برآوردهای نااریبی از پارامترها (و نیز خطای معیار آن‌ها) نتیجه خواهد داد. ۳) گم‌شدن غیر تصادفی[3]: در این حالت گم‌شدگی در

1. Missing at random (MAR)
2. Missing completely at random (MCAR)
3. Missing not at random (MNAR)

متغیر دارای مقادیر گمشده به متغیری که برای همه افراد شرکت‌کننده در مطالعه به‌طور کامل مشاهده شده است، و به خود متغیر دارای مقادیر گمشده بستگی دارد (Van Buuren & Groothuis-Oudshoorn, 2011). برای استفاده صحیح از روش‌های مقابله با داده‌های ازدست‌رفته، ارزیابی کامل و معتبر از عملکرد این روش‌ها حیاتی است. ازآنجایی‌که پژوهش‌های علمی به‌طور فزاینده‌ای بر روش‌شناسی داده‌های ازدست‌رفته تکیه می‌کند، مهم است که بدانیم تحت چه شرایطی می‌توان و باید از یک تکنیک خاص استفاده کرد (Schouten et al., 2018). علاوه بر این پژوهشگران نشان داده‌اند که اگر داده‌های گمشده نادیده گرفته شوند، ممکن است مشکلاتی را در تخمین پارامترهای دشواری آیتم در زمینه نظریه سؤال-پاسخ (IRT) ایجاد کند (Finch, 2008). مسئله داده‌های گمشده، زمانی که تعداد ابعاد موردسنجش در اندازه‌گیری روان‌شناختی افزایش می‌یابد، اهمیت بیشتری پیدا می‌کند. نفرین ابعاد[1] یا وجود تعداد داده‌های بالا، یکی از مهم‌ترین چالش‌ها در حوزه اندازه‌گیری است که توسط Richard Bellman (1957) مطرح شد تا رشد فوق‌العاده سریع مشکلات را با افزایش تعداد متغیرها (ابعاد) توصیف کند (Kuo & Sloan, 2005). نفرین ابعاد به پدیده‌های مختلفی اشاره دارد که هنگام تحلیل و سازمان‌دهی داده‌ها در فضاهای با ابعاد بالا در حوزه‌هایی مانند تحلیل عددی[2]، یادگیری ماشین[3] و داده‌کاوی[4] به وجود می‌آیند. درواقع می‌توان گفت افزایش ابعاد در یک آزمون (به‌ویژه در حوزه مطالعات صفات انسانی) بر پیچیدگی‌های محاسباتی و افزایش خطا در دقت برآورد می‌افزاید (Liu & Pierce, 1994; Cai, 2010).

یکی از رویکردهایی که توجه ویژه‌ای به مشکلات موجود از قبل نفرین ابعاد و داده‌های گمشده، در تحلیل داده‌های رفتاری داشته است و از مدل‌ها و روش‌های محاسباتی توانمندی بهره گرفته، نظریه سؤال پاسخ (Hambleton & Swaminathan, 1985; van der Linden & Hambleton, 1997) است. نظریه سؤال پاسخ به‌عنوان یک نظریه روان‌سنجی جدید با برخورداری از مدل‌های متنوع و پیشرفته و کاربرد الگوریتم‌های متنوع در محاسبات ریاضیاتی، به‌طور گسترده در آزمون‌های آموزشی و اندازه‌گیری، ارزیابی روانی و پیامدهای بهداشتی مورداستفاده قرار گرفته است (Bartolucci et al., 2015; Gibbons et al., 2016).

1. Curse of dimensionality
2. Numerical Analysis
3. Machine Learning
4. Data Mining

در سال‌های گذشته چندین روش برای ایجاد امکان برآورد مدل‌های نظریه سؤال پاسخ با ابعاد بالا و کارآیی بیشتر در فرآیند تخمین، توسعه یافته‌اند. برای چند دهه، مدل‌های نظریه سؤال پاسخ با استفاده از: (۱) بیشینه احتمال حاشیه‌ای[1] بر اساس الگوریتم بیشینه انتظار[2]، (۲) الگوریتم بیشینه انتظار مونت‌کارلو (MCEM) و (۳) برآورد بیزین کامل[3] با استفاده از شبیه‌سازی مونت کارلوی زنجیره مارکوفی[4] (MCMC) تخمین زده می‌شود.

طبق ادبیات پژوهش، پژوهشگران برای هر یک از این الگوریتم‌ها، معایبی ذکر کرده‌اند که حکایت از ناتوانی آن‌ها در برآورد دقیق پارامترها دارد. الگوریتم EM از نقاط کوادراتور گاوسی-هرمیت ثابت، برای برآورد استفاده می‌شود ولی زمانی که تعداد ابعاد به چهار یا پنج (یا بیشتر) افزایش یابد، فرایند تخمین با مشکل روبه‌رو می‌شود؛ زیرا تعداد کل نقاط کوادره با قدرتی برابر با تعداد ابعاد افزایش می‌یابد و ارزیابی انتگرال‌ها را بسیار دشوار می‌کند (Liu & Pierce, 1994; Naylor & Smith, 1982). الگوریتم MCEM نیز برای رسیدن به همگرایی نقطه‌ای برآورد پارامترها، اندازه شبیه‌سازی باید بسیار افزایش یابد، به‌ویژه در چند تکرار آخر، زیرا تخمین‌های پارامتر به حداکثر تابع احتمال نزدیک‌تر می‌شوند (Cai, 2010). علاوه بر این، زمان همگرایی MCEM به این دلیل افزایش می‌یابد که برای هر تکرار مرحله E، نمونه‌گیری مجموعه جدیدی از قرعه‌کشی‌های تصادفی را ایجاد می‌کند. در الگوریتم MCMC نیز، یک چالش بزرگ، مشخصات توزیع پیشنهادی مناسب است، که ممکن است نیاز به تجربه با آزمون‌وخطا باشد. علاوه بر این، ازآنجاکه MCMC هنوز هم کل سطح پاسخ را در فضای چندبعدی تقریب می‌زند، برای مشکلات چند متغیره ممکن است هنوز به زمان محاسبات گسترده‌ای نیاز داشته باشد، که استفاده از آن را در عمل منع می‌کند. علاوه بر این، ارزیابی همگرایی در MCMC اعمال‌شده در مدل‌های پیچیده با بسیاری از سؤالات یا بسیاری از ابعاد صفت نهفته، می‌تواند دست‌وپا گیر باشد و نیاز به قضاوت انسان دارد (Bashkov & DeMars, 2017).

در پاسخ به محدودیت‌های روش‌های محبوب برآورد در مدل‌های چندبعدی نظریه سؤال پاسخ که در بالا توضیح داده شد، Cai (2010) یک الگوریتم متروپلیس-هستینگز روبینز-

1. Marginal Maximum Likelihood (MML)
2. Expectation Maximization (EM)
3. Fully Bayesian estimation
4. Markov Chain Monte Carlo estimation

مونرو[1] (MHRM) معرفی کرد که نمونه‌گیری متروپولیس- هستینگز (Patz & Junker, 1999) را با نمونه‌گیری روبینز-مونرو (Robbins & Monro, 1951) ترکیب می‌کند تا بیشینه درستنمایی را تسهیل کند. الگوریتم متروپلیس- هستینگز روبینز-مونرو (MHRM) یک الگوریتم ترکیبی مبتنی بر احتمال حاشیه‌ای است که در آن نمونه‌های توزیع شرطی عوامل از طریق تقریب تصادفی ترکیب می‌شوند (Cai, 2010). MHRM برخلاف MCMC که کل توزیع پسین را تقریب می‌زند، بر برآورد نقطه‌ای و خطاهای استاندارد تمرکز دارد. این اجازه می‌دهد تا MHRM خیلی سریع‌تر از MCMC به همگرایی برسد. به‌طور خاص، در MHRM، جهش‌ها در زنجیره مارکوف صرفاً به‌منظور تعیین جهت تغییر در طول تکرارها است، نه فراهم آوردن تقریب دقیق سطحی که ممکن است، هدف باشد (Cai, 2010).

پیشینه پژوهش

در رابطه با توانمندی الگوریتم MHRM پژوهش‌های محدودی صورت گرفته است. Yang & Cai (2014) در پژوهشی با عنوان «برآورد اثرات زمینه‌ای از طریق مدل‌سازی متغیر پنهان چند سطحی غیرخطی با الگوریتم» با هدف مطالعه بهبود بهره‌وری برآورد در به دست آوردن بیشینه برآورد احتمال حاشیه‌ای اثرات زمینه‌ای در چارچوب مدل متغیر پنهان چند سطحی غیرخطی با کاربرد الگوریتم MHRM است. نتایج نشان می‌دهد که الگوریتم MHRM می‌تواند تخمین‌ها و خطاهای استاندارد را به‌طور کارآمد تولید کند. Kuo and Sheng (2016) نیز در پژوهشی با عنوان «مقایسه روش‌های تخمین برای یک مدل پاسخ درجه‌بندی چندبعدی» به بررسی توانمندی الگوریتم‌های متفاوت در مدل چندبعدی پاسخ مدرج نظریه سؤال پاسخ پرداختند. در این پژوهش روش‌های برآورد EM و MHRM تحت شرایط متفاوت کیفیت داده مورد مقایسه قرار گرفتند و توانمندی هر یک از روش‌های برآورد مورد ارزیابی قرار گرفت. Bashkov and DeMars (2017) در پژوهشی با عنوان «آزمون کارایی الگوریتم متروپلیس هستینگز روبینز مرنرو در برآورد مدل‌های چندبعدی چند سطحی» کارایی الگوریتم متروپلیس هستینگز روبینز مرنرو را در داده‌های دو ارزشی با استفاده از مدل سه پارامتری مورد ارزیابی قرار دادند و کارایی MHRM در بازیابی پارامترهای سؤال، واریانس و کوواریانس پنهان، و همچنین برآورد توانایی درون خوشه‌ها و خوشه‌ها

1. Metropolis-Hastings Robbins-Monro

(به‌عنوان‌مثال، مدارس) در یک مطالعه شبیه‌سازی، بررسی شد که نتایج نشان داد، الگوریتم MHRM عملکرد خوبی در برآورد سؤال، آزمودنی و پارامترهای سطح گروه مدل داشت.

در کل مطالعات نشان داده‌اند که استفاده از روش‌های برآورد ذکرشده باعث می‌شود فرآیند تخمین برای مدل‌های چندبعدی ازنظر محاسباتی فشرده[1] و غالباً لجباز[2] باشد که تعداد زیادی از ابعاد در آن دخیل باشد (Cai, 2010; Chalmers, 2012). درنتیجه الگوریتم MHRM برای غلبه بر مشکلات برآورد بیشینه احتمال و سایر الگوریتم‌های مرسوم و ارائه تخمین‌های مفید برای داده‌هایی با تعداد متفاوتی از سؤال، ابعاد مختلف و یا داده‌های گمشده ارائه شد (Kuo & Sheng, 2016). درنتیجه با عطف به پژوهش‌های انجام‌شده و نظر به چالش‌های موجود، در این پژوهش از جهت تحلیل مدل‌های چندبعدی نظریه سؤال پاسخ از MHRM در کنار الگوریتم‌های EM و MCEM استفاده شد و به بررسی این موضوع اساسی پرداخته شد که آیا الگوریتم MHRM در قیاس با الگوریتم‌های مرسوم دیگر یعنی الگوریتم‌های EM و MCEM با در نظر گرفتن نوع و میزان داده‌های گمشده عملکرد بهتری دارد؟

روش

پژوهش حاضر ازنظر فلسفی اثبات گرایانه، هدف کاربردی، روش کمی، از نوع آزمایشی و به دنبال مقایسه پدیده موردمطالعه بود. هدف اصلی از کاربرد روش آزمایشی در این پژوهش، تعیین شرایطی که در آن پدیده معینی مبتنی بر مدل‌های نظریه سؤال پاسخ در مطالعات شبیه‌سازی اتفاق می‌افتد. طرح پژوهشی مورداستفاده در این پژوهش عاملی و از نوع سه راهه است که متغیرهای مستقل عبارت‌اند از نوع الگوریتم، نوع داده‌های گمشده و میزان داده‌های گمشده. در این پژوهش به‌منظور مقایسه میزان ریسک الگوریتم MHRM از داده‌های شبیه‌سازی‌شده مونت‌کارلو استفاده شده است. دلیل استفاده از داده‌های شبیه‌سازی‌شده مونت‌کارلو این بود که اساساً دستیابی به داده‌های واقعی آزمون‌های متفاوت برای مطالعات روش‌شناسی مقایسه الگوریتم‌ها در شرایط آزمایشی متفاوت بسیار دشوار است؛ زیرا ایجاد شرایط واقعی آزمون‌ها برای تعداد زیادی از آزمودنی‌ها با تکرارهای بسیار زیاد و در شرایط یکسان کار بسیار دشواری است و فقط در تخیل می‌تواند، صورت بگیرد

1. Intensive
2. Intractable

اما استفاده از مطالعات مونت‌کارلو این تخیل را در محیطی شبیه‌سازی‌شده و تحت کنترل به واقعیت تبدیل می‌کند. این روش سریع‌تر، ارزان‌تر و آسان‌تر از جمع‌آوری اطلاعات از آزمودنی‌های زنده است و به محقق اجازه می‌دهد، مدل‌ها و روش‌های جدید روان‌سنجی را به‌سرعت و با هزینه بسیار کمی موردپژوهش قرار دهند. مطالعه شبیه‌سازی در این پژوهش با توجه به نقش داده‌های گمشده، به‌طورکلی دارای چهار مرحله بود: (۱) یک مجموعه داده چند متغیره و کامل شبیه‌سازی‌شده و جمعیت موردنظر در نظر گرفته شد. (۲) سپس این مجموعه داده بر اساس نوع و میزان، دارای گمشدگی شد. (۳) داده‌های دارای گمشدگی در حالت‌های مختلف با استفاده از روش‌های آماری تخمین زده شد. (۴) استنباط‌های آماری برای مجموعه پارامترهای حاصل‌شده انجام شد. مقایسه این استنباط‌ها نشانه‌ای از عملکرد داده‌های گمشده و توانمندی روش‌های آماری را نشان می‌دهد.

در مرحله اول ابتدا مدل مورداستفاده در این پژوهش به دلیل نوع داده‌ها (چند ارزشی)، مدل پاسخ مدرج چندبعدی استفاده که دو پارامتر شیب و آستانه را در خود جای دارد. در تولید داده‌ها، شکل توزیع توانایی آزمودنی‌ها بر اساس توزیع نرمال با میانگین صفر و انحراف معیار یک در نظر گرفته شد. تعیین شکل توزیع توانایی آزمودنی‌ها معمولاً از توزیع یک جامعه مشخص به‌صورت تصادفی انتخاب می‌شود که اکثر اوقات توضیح استاندارد است (Han & Hambleton, 2014). به لحاظ نظری هر یک از اشکال مختلف می‌توانند به‌عنوان توزیع پیشین فرض شوند، اما اغلب به نظر می‌رسد توزیع طبیعی انتخاب درست‌تری باشد (Embretson & Reise, 2013). پارامترهای سؤالات (شیب و آستانه) نیز بر اساس پیشنهاد Bulut & SÜNBÜL (2017) عمل شد. مقادیر پارامتر شیب بر اساس توزیع لوگ نرمال با میانگین ۰/۳ و انحراف معیار ۰/۲ و پارامتر دشواری بر اساس توزیع نرمال با میانگین صفر و انحراف معیار یک در نظر گرفته شد. لازم به ذکر است که به دلیل اینکه با توجه به تعداد ابعاد و الگوی پاسخ به ترتیب ۵ شیب و ۴ آستانه وجود داشت که به دلیل وجود همبستگی بالا و هم خطی چندگانه، میانگین مخاطره این ضرایب در نظر گرفته شد.

جامعه آماری و حجم نمونه در این پژوهش آزمون‌هایی هستند که با توجه به شرایط مختلفی که با توجه به متغیرهای نوع الگوریتم، نوع داده‌های گمشده و میزان داده‌های گمشده ایجاد می‌شود، تعیین شدند. در این بخش داده‌ها با نوع داده گمشده MAR، MNAR و MCAR و میزان داده گمشده ۱۰ درصد، ۵۰ درصد و ۹۰ درصد تولید شد که در معرض

برآورد پارامترها با ۳ روش الگوریتم EM، MCEM و MHRM قرار گرفتند. در زمینه درصد داده‌های گمشده طبق نظر Van Buuren and Groothuis-Oudshoorn (2011) عمل شد که دو روش برای در نظر گرفتن درصد داده‌های گمشده با اهداف شبیه‌سازی معرفی شده است. در روش اول بر اساس تعداد آزمودنی‌ها و در روش دوم تعداد خانه‌های موجود در ماتریس سؤالات و آزمودنی‌ها عمل می‌شود. در این پژوهش از روش اول استفاده شد که در سه سطح و بر اساس حجم آزمودنی‌ها، داده‌های گمشده منظور شد. لازم به ذکر است مدل ۵ بعدی (به‌عنوان مدل پرکاربرد در بسیاری از آزمون‌های روان‌شناختی از قبیل مدل‌های ۵ عاملی شخصیت) و حجم نمونه ۱۰۰۰ داده برای همه شرایط در نظر گرفته شد. در ادامه حالت‌های مختلف تولید داده در جدول ۱ ارائه شده است.

جدول ۱. حالت‌های مختلف تولید داده‌های شبیه‌سازی‌شده

نوع الگوریتم	نوع داده گمشده	درصد گمشدگی	تکرار	نوع الگوریتم	نوع داده گمشده	درصد گمشدگی	تکرار	نوع الگوریتم	نوع داده گمشده	درصد گمشدگی	تکرار
EM	MAR	۱۰	۱۰۰	MHRM	MAR	۱۰	۱۰۰	MCEM	MAR	۱۰	۱۰۰
		۵۰	۱۰۰			۵۰	۱۰۰			۵۰	۱۰۰
		۹۰	۱۰۰			۹۰	۱۰۰			۹۰	۱۰۰
	MNAR	۱۰	۱۰۰		MNAR	۱۰	۱۰۰		MNAR	۱۰	۱۰۰
		۵۰	۱۰۰			۵۰	۱۰۰			۵۰	۱۰۰
		۹۰	۱۰۰			۹۰	۱۰۰			۹۰	۱۰۰
	MCAR	۱۰	۱۰۰		MCAR	۱۰	۱۰۰		MCAR	۱۰	۱۰۰
		۵۰	۱۰۰			۵۰	۱۰۰			۵۰	۱۰۰
		۹۰	۱۰۰			۹۰	۱۰۰			۹۰	۱۰۰

برای پاسخگویی به سؤالات پژوهش با توجه به طرح‌های تحقیق (عاملی) از تحلیل واریانس عاملی سه راهه استفاده شد. به‌طور خاص، مخاطره[1] پارامترها (تمیز و آستانه سؤال‌ها) ارزیابی شد. جهت تولید و تحلیل داده‌ها از نرم‌افزارهای آماری R بسته‌های mirt، interactions، car و psych استفاده شد.

1. Risk

یافته‌ها

ابتدا میانگین و انحراف استاندارد متغیرهای وابسته پژوهش (میانگین توان دوم خطاهای شاخص‌های شیب و آستانه) به تفکیک متغیرهای مستقل در جدول ۲ ارائه شده است.

جدول ۲. اطلاعات توصیفی میانگین توان دوم خطاها (MSE)

الگوریتم	نوع داده گمشده	میزان داده گمشده	شیب (a)		دشواری (b)	
			M	SD	M	SD
EM	MAR	۱۰ درصد	۰/۰۱۱۱	۰/۰۰۳۰	۰/٤۱۱۳	۰/۱۸۸۵
		۵۰ درصد	۰/۰۱۰٦	۰/۰۰۲٦	۰/٤۱۲٦	۰/۲۰۱۱
		۹۰ درصد	۰/۰۱۰٦	۰/۰۰۳۱	۰/۳۲۹۹	۰/۱۹۵٤
	MNAR	۱۰ درصد	۰/۰۱۰۱	۰/۰۰۲٦	۰/۳٦۹۰	۰/۲۱٤۷
		۵۰ درصد	۰/۰۱۰۱	۰/۰۰۲۹	۰/۳۹۱٤	۰/۲۱۵۷
		۹۰ درصد	۰/۰۰۹۲	۰/۰۰۲٤	۰/۳٦۵۱	۰/۲۰۱۳
	MCAR	۱۰ درصد	۰/۰۱۱٤	۰/۰۰۲۸	۰/۲۹۳۹	۰/۱٦۹۰
		۵۰ درصد	۰/۰۰۹۸	۰/۰۰۲۷	۰/۳۷۷۸	۰/۱۹۷۰
		۹۰ درصد	۰/۰۰۹۵	۰/۰۰۲٤	۰/۳۷۱۵	۰/۲۱٦۲
MCEM	MAR	۱۰ درصد	۰/۰۰٤۹	۰/۰۰۱۰	۰/۳۷۲۷	۰/۱۹۲۵
		۵۰ درصد	۰/۰۰۵۱	۰/۰۰۱۱	۰/۳۲۰۱	۰/۱۵۷۷
		۹۰ درصد	۰/۰۰۵٤	۰/۰۰۱٤	۰/۳۰۷۹	۰/۱۸۸۰
	MNAR	۱۰ درصد	۰/۰۰٤٦	۰/۰۰۰۹	۰/۳٦۹۲	۰/۱۹۰۱
		۵۰ درصد	۰/۰۰٤۹	۰/۰۰۱۰	۰/٤۳۲۵	۰/۲۱٤٤
		۹۰ درصد	۰/۰۰۵٤	۰/۰۰۱٤	۰/۳۹۳٤	۰/۲۳۹۹
	MCAR	۱۰ درصد	۰/۰۰٤۹	۰/۰۰۱۰	۰/۳۱٦۸	۰/۱۸۹۰
		۵۰ درصد	۰/۰۰٤۸	۰/۰۰۱۰	۰/۳۲۳۱	۰/۱۸۹۲
		۹۰ درصد	۰/۰۰۵۳	۰/۰۰۱۳	۰/۳٦۹۹	۰/۲۱۲۲
MHRM	MAR	۱۰ درصد	۰/۰۰٤٦	۰/۰۰۱۱	۰/۲۳۵٤	۰/۱٦۳۵
		۵۰ درصد	۰/۰۰٤۲	۰/۰۰۱۱	۰/۲۲۹۰	۰/۱۸۲۳
		۹۰ درصد	۰/۰۰٤۹	۰/۰۰۱٤	۰/۲۳۳۵	۰/۱٦٦٤
	MNAR	۱۰ درصد	۰/۰۰٤۱	۰/۰۰۱۱	۰/۲۷۵۱	۰/۲۱۹٤
		۵۰ درصد	۰/۰۰٤٦	۰/۰۰۱۲	۰/۳۳۵۳	۰/۲۱٤٤
		۹۰ درصد	۰/۰۰٤۹	۰/۰۰۱٤	۰/۲۲۱۵	۰/۱۷۱۹
	MCAR	۱۰ درصد	۰/۰۰۵۲	۰/۰۰۱۳	۰/۲۵۸۱	۰/۱۸۳۷
		۵۰ درصد	۰/۰۰٤۵	۰/۰۰۱۲	۰/۲۳۹۱	۰/۱۷۸۰
		۹۰ درصد	۰/۰۰٤۸	۰/۰۰۱۳	۰/۲۲۸۲	۰/۱٦٤۲

در ادامه جهت پاسخگویی به سؤال پژوهش مبنی بر اینکه آیا الگوریتم MHRM در قیاس با الگوریتم‌های مرسوم دیگر یعنی الگوریتم‌های EM و MCEM با در نظر گرفتن تعداد ابعاد مختلف و تعداد متفاوت سؤالات عملکرد بهتری دارد؟ از روش تحلیل واریانس چندمتغیره عاملی[1] استفاده شد. در این نوع تحلیل باید مفروضه‌هایی رعایت گردند تا بتوان به نتایج به‌دست‌آمده اطمینان کرد. مفروضه اول وجود ساختار همبستگی بین متغیرهای وابسته بود که نتایج آزمون کرویت بارتلت نشان داد بین متغیرهای وابسته پژوهش همبستگی لازم جهت تحلیل‌های چندمتغیری وجود دارد (P=۰/۰۰۱ و Square=۲۱۰۲۱/۴۰-). یکی دیگر از این مفروضه‌ها، بررسی همسانی ماتریس‌های واریانس-کوواریانس است که بدین منظور از آزمون باکس[2] استفاده شده است. میزان معناداری آزمون باکس از ۰,۰۵ کوچک‌تر است (P=۰/۰۰۱ و F=۱۲/۹۸ و Box`s M=۱۰۲۰/۵۹)، لذا نتیجه گرفته می‌شود که ماتریس واریانس-کوواریانس‌ها همگن نیست و لذا برقراری این مفروضه نقض شده است. در ادامه جهت بررسی مفروضه نرمال‌بودن داده‌ها از آزمون نرمال بودن چندمتغیره شاپیرو ویلک[3] استفاده شد که مقدار به‌دست‌آمده (P=۰/۰۱ و Mvw=۰/۹۱) نشان از عدم نرمال بودن داده‌ها دارد. با توجه به رد برخی از مفروضه‌ها و البته به دلیل وجود تعداد داده‌های یکسان در هر یک از سطوح طرح عاملی، در بررسی نتایج تحلیل کوواریانس چندمتغیری از آزمون مقاوم اثر پیلایی[4] که نسبت به نقض مفروضه‌ها پایدارتر است، استفاده شد. در جدول ۳ آماره چندمتغیری برای هر یک از متغیرهای مستقل و تعامل بین آن‌ها ارائه شده است.

جدول ۳. نتایج تحلیل کوواریانس چندمتغیری

متغیرها	اثر پیلایی	F	df1	df2	P	η2
الگوریتم	۰/٦۹	۷۱٤/۸٦	٤	٥۳٤٦	۰/۰۰۱	۰/۳٤۸
نوع داده گمشده	۰/۰۲	۱۰/۸۷	٤	٥۳٤٦	۰/۰۰۱	۰/۰۰۸
میزان داده گمشده	۰/۰۰٥	۲/۲٤	٤	٥۳٤٦	۰/۰۸	۰/۰۰۲
الگوریتم*نوع داده گمشده	۰/۰۱	٤/٤۲	۸	٥۳٤٦	۰/۰۰۱	۰/۰۰۷
الگوریتم*میزان داده گمشده	۰/۰۳	۹/٥۱	۸	٥۳٤٦	۰/۰۰۱	۰/۰۱٥
نوع داده گمشده * میزان داده گمشده	۰/۰۱	٤/۸۳	۸	٥۳٤٦	۰/۰۰۱	۰/۰۰۷
الگوریتم*نوع گمشدگی*میزان گمشدگی	۰/۰۱	۲/٤۰	۱٦	٥۳٤٦	۰/۰۰۱	۰/۰۱۲

1. Factorial Multivariate Analysis of variance
2. Box's Test of Equality of Covariance Matrices
3. Shapiro-Wilk test for Multivariate Normality
4. Pillais Trace

همان‌گونه که در جدول ۳ مشاهده می‌شود، آماره چندمتغیری مربوطه یعنی اثر پیلایی در سطح خطای ۰/۰۱۶ (بر اساس آلفای بونفرونی) برای هر سه متغیر مستقل و حالت‌های مختلف تعامل بین آن‌ها معنی‌دار است (P=۰/۰۰۱). بدین ترتیب ترکیب خطی متغیرهای وابسته (شاخص MSE ضرایب شیب و آستانه) حداقل یکی از متغیرهای مستقل (نوع الگوریتم، نوع داده گمشده و میزان داده گمشده) تأثیر پذیرفته است. همچنین لازم به ذکر است تمامی تعامل‌های بین متغیرهای مستقل با یکدیگر معنی‌دار است. با توجه به اینکه آزمون چندمتغیری مذکور معنادار بوده و ترکیب متغیرهای وابسته از حداقل یکی از متغیرهای مستقل اثر پذیرفته است، لذا بعدازآن به پیگیری وضعیت اثرگذاری هر یک از متغیرهای مستقل بر هر یک از متغیرهای وابسته و نیز تعامل بین آن‌ها به‌طور مجزا پرداخته شده است. جهت بررسی این موضوع از آزمون تجزیه‌وتحلیل واریانس عاملی تک متغیره استفاده شد که نتایج آن در جدول ۴ ارائه شده است.

جدول ٤. نتایج تجزیه‌وتحلیل واریانس تک‌متغیره جهت مقایسه میانگین توان دوم خطاها (MSE)

متغیر وابسته	متغیر مستقل	df	MS	F	P	η^2
شیب	الگوریتم	۲	۰/۰۰۹	۲۵۰۲/۲۸	۰/۰۰۱	۰/٦٥
	نوع داده گمشده	۲	۰/۰۰۱	۹/۹۷	۰/۰۰۱	۰/۰۰۷
	میزان داده گمشده	۲	۰/۰۰۱	۱/۸۹	۰/۱٥	۰/۰۰۱
	الگوریتم*نوع داده گمشده	٤	۰/۰۰۱	٥/٦٤	۰/۰۰۱	۰/۰۰۸
	الگوریتم*میزان داده گمشده	٤	۰/۰۰۱	۱۸/۰۱	۰/۰۰۱	۰/۰۳
	نوع داده گمشده * میزان داده گمشده	٤	۰/۰۰۱	٤/٤۳	۰/۰۰۱	۰/۰۳
	الگوریتم*نوع گمشدگی*میزان گمشدگی	۸	۰/۰۰۱	۱/۹۸	۰/۰٤	۰/۰۰٦
	خطا	۲٦۷۳	۰/۰۰۱	–	–	–
آستانه	الگوریتم	۲	۳/۷۸	۱۰۰/۲۸	۰/۰۰۱	۰/۰۷
	نوع داده گمشده	۲	۰/٤٤	۱۱/۸۹	۰/۰۰۱	۰/۰۰۹
	میزان داده گمشده	۲	۰/۱۷	۲/٦۱	۰/۰٦	۰/۰۰۳
	الگوریتم*نوع داده گمشده	٤	۰/۱۲	۳/۲٤	۰/۰۱	۰/۰۰٥
	الگوریتم*میزان داده گمشده	٤	۰/۰٥	۱/۲٥	۰/۲۸	۰/۰۰۲
	نوع داده گمشده * میزان داده گمشده	٤	۰/۲۰	٥/۲۲	۰/۰۰۱	۰/۰۰۸
	الگوریتم*نوع گمشدگی*میزان گمشدگی	۸	۰/۱۰	۲/۸٥	۰/۰۰٤	۰/۰۰۸
	خطا	۲٦۷۳	۰/۰۳	–	–	–

با توجه به نتایج جدول ۳، F مشاهده شده در سطح خطای ۰/۰۵ تفاوت معناداری را بین میانگین گروه‌های مورد بر اساس متغیرهای مستقل (نوع الگوریتم و نوع داده‌های گمشده)، در هر دو متغیر وابسته یعنی مخاطره شیب و آستانه نشان داد. ولیکن در رابطه با متغیر مستقل میزان داده‌های گمشده، بنا بر شواهد به‌دست‌آمده می‌توان گفت بر تغییرات مخاطره هر دو پارامتر شیب و آستانه نقش معنی‌داری ندارد. همچنین تعامل بین نوع الگوریتم و میزان داده‌های گمشده در رابطه با تغییرات MSE آستانه معنی‌دار نبود. این نتایج نشان می‌دهد تفاوت در سطوح هر یک از متغیرهای مستقل (به‌ویژه نوع الگوریتم و نوع داده‌های گمشده)، در هر یک از متغیرهای وابسته، میزان میانگین توان دوم خطای متفاوتی را به دنبال دارد. در ادامه با استفاده از آزمون تعقیبی بونفرونی به بررسی تفاوت نمرات متغیرهای وابسته در بین گروه‌های هر یک از متغیرهای مستقل به‌صورت زوجی پرداخته شد که نتایج در جدول ۵ ارائه شده است.

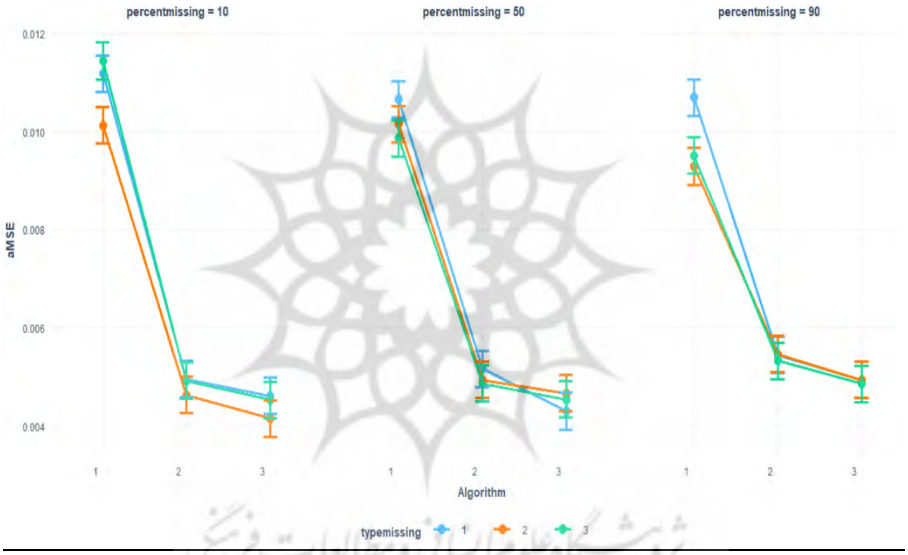
جدول ۵: نتایج آزمون بونفرونی برای مقایسه‌ی MSE گروه‌های متغیرهای مستقل

متغیر وابسته	متغیر مستقل	گروه	تفاوت میانگین	خطای برآورد	P	حد پایین	حد بالا
شیب	نوع الگوریتم	EM و MCEM	۰/۰۰۵	۰/۰۰۱	۰/۰۰۱	۰/۰۰۵	۰/۰۰۵
		EM و MHRM	۰/۰۰۶	۰/۰۰۱	۰/۰۰۱	۰/۰۰۵	۰/۰۰۶
		MCEM و MHRM	۰/۰۰۱	۰/۰۰۱	۰/۰۰۱	۰/۰۰۱	۰/۰۰۱
	نوع داده گمشده	MAR و MNAR	۰/۰۰۱	۰/۰۰۱	۰/۰۰۱	۰/۰۰۱	۰/۰۰۱
		MAR و MCAR	۰/۰۰۱	۰/۰۰۱	۰/۰۲۸	۰/۰۰۱	۰/۰۰۱
		MCAR و MNAR	۰/۰۰۱	۰/۰۰۱	۰/۱۹۹	۰/۰۰۱	۰/۰۰۱
آستانه	نوع الگوریتم	EM و MCEM	۰/۰۱	۰/۰۰۹	۰/۳۹	−۰/۰۰۸	۰/۰۳۶
		EM و MHRM	۰/۱۱	۰/۰۰۹	۰/۰۰۱	۰/۰۹۷	۰/۱۴۱
		MCEM و MHRM	۰/۱۰	۰/۰۰۹	۰/۰۰۹	۰/۰۸۳	۰/۱۲۷
	نوع داده گمشده	MAR و MNAR	−۰/۰۳	۰/۰۰۹	۰/۰۰۱	−۰/۰۵	−۰/۰۱
		MAR و MCAR	۰/۰۰۹	۰/۰۰۹	۰/۹۷	۰/۰۱	۰/۰۳
		MNAR و MCAR	۰/۰۴	۰/۰۰۹	۰/۰۰۱	۰/۰۲	۰/۰۶

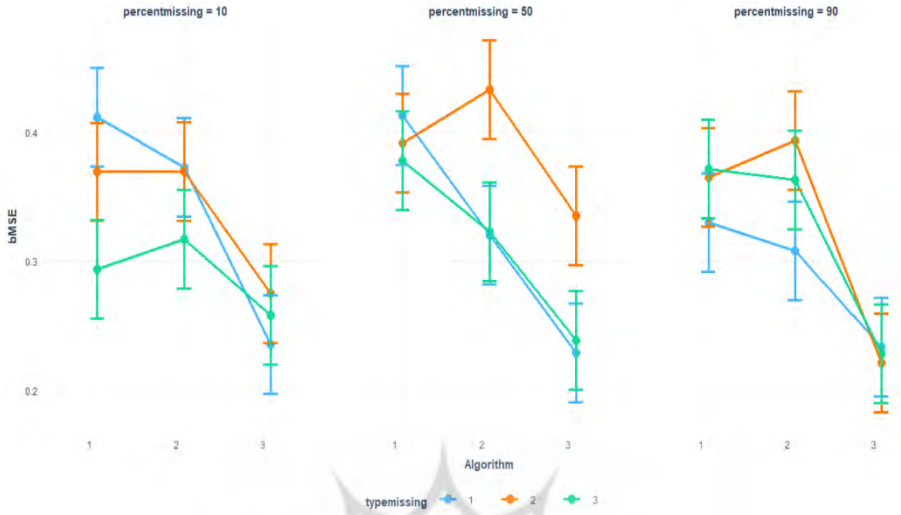
نتایج جدول ۶ نشان می‌دهد که در مقایسه زوجی سطوح مختلف متغیرهای مستقل (نوع الگوریتم و نوع داده‌های گمشده) در متغیر وابسته شیب، نوع الگوریتم در تمامی مقایسه‌ها در سطح اطمینان ۹۵ درصد تفاوت معنی‌داری وجود دارد ولیکن در رابطه با متغیر مستقل نوع داده‌های گمشده بین دو روش MNAR و MCAR تفاوت معنی‌داری مشاهده نشد. در

بررسی نقش متغیرهای مستقل بر ضریب آستانه نیز مشخص شد الگوریتم MHRM نسبت به دو الگوریتم EM و MCEM دارای مخاطره کمتری است ولیکن بین دو الگوریتم EM و MCEM تفاوت معنی‌داری مشاهده نشد. نتایج بررسی میزان مخاطره ضریب آستانه نشان داد در نوع داده‌های گمشده بین دو روش MAR و MCAR و در میزان داده‌های گمشده بین دو مقدار ۱۰ و ۵۰ درصد تفاوت معنی‌داری در سطح خطای ۰/۰۵ وجود ندارد. در ادامه نتایج پژوهش به‌صورت نمودار در شکل‌های ۱ و ۲ به تفکیک متغیرهای وابسته پژوهش ارائه شده است.

شکل ۱. نمودار مقایسه میانگین توان دوم خطاها (پارامتر شیب)



شکل ۲. نمودار مقایسه میانگین توان دوم خطاها (پارامتر آستانه)



بحث و نتیجه‌گیری

هدف پژوهش حاضر بررسی مخاطره الگوریتم متروپلیس هستینگز روبینز مونرو در داده‌های چندبعدی نظریه سؤال پاسخ بود. همچنین در این راستا یکی دیگر از چالش‌های موجود در داده‌های علوم رفتاری یعنی وجود داده‌های گمشده نیز مورد ارزیابی قرار گرفت. نتایج نشان داد هر سه متغیر نوع الگوریتم، نوع و میزان داده‌های گمشده حداقل در یکی از مقایسه بین سطوح، در میزان میانگین توان دوم خطای برآورد پارامترها اثرگذار هستند. در بررسی مقایسه الگوریتم MHRM با الگوریتم‌های EM و MCEM نتایج به‌دست‌آمده نشان از برتری معنی‌دار الگوریتم MHRM داشت. بدین معنی که این الگوریتم در مقایسه با دو الگوریتم EM و MCEM دارای مخاطره کمتری بود. همچنین نکته مهم دیگر این بود که در برآورد پارامتر شیب، الگوریتم EM نسبت به دو الگوریتم MCEM و MHRM دارای مخاطره بسیار بیشتری است.

در بررسی نقش نوع داده‌های گمشده در برآورد پارامترها هم مشخص شد، که به مکانیسم داده گمشده MCAR نسبت به دو روش MAR و MNAR به‌ویژه در رابطه با پارامتر آستانه دارای مخاطره پایین‌تری است. نکته مهمی که در این رابطه می‌توان بیان کرد معنی‌دار بودن تعامل بین نوع الگوریتم و نوع مکانیسم داده‌های گمشده است. در بررسی تعامل بین نوع الگوریتم و مکانیسم داده گمشده مشخص شد که وقتی از الگوریتم MHRM استفاده

می‌شود تفاوت اندکی بین سه نوع مکانیسم داده گمشده وجود دارد و به‌طورکلی اندازه اثر میزان مخاطره به‌طور قابل‌ملاحظه‌ای کاهش می‌یابد و مقادیر خطاها در هر سه حالت به یکدیگر نزدیک می‌شود. در این زمینه وقتی از الگوریتم MCEM استفاده می‌شود، تفاوت معنی‌داری بین سه نوع مکانیسم گمشدگی به وجود می‌آید که البته بیشترین مقدار خطا درمکرد مکانیسم MNAR اتفاق می‌افتد. این شرایط در برآورد پارامتر آستانه با اندازه‌های اثر بالاتر، به‌خوبی قابل‌مشاهده است ولیکن در رابطه با پارامتر شیب اندازه اثر بسیار پایینی مشاهده می‌شود.

چالش بعدی که در این پژوهش مورد ارزیابی قرار گرفت، میزان داده‌های گمشده و نقش آن در مخاطره برآورد پارامترهای شیب و آستانه و البته تعامل آن با دو متغیر مستقل دیگر یعنی نوع الگوریتم و مکانیسم داده‌های گمشده بود. در رابطه با نوع داده گمشده نتایج نشان داد که در بررسی هر دو پارامتر آستانه و شیب تفاوت معنی‌داری بین سطوح مختلف آن، وجود ندارد. ولیکن در بررسی‌های بعدی و تعامل میزان داده‌های گمشده با دو متغیر مستقل دیگر نتایج معنی‌داری مشاهده شد. در این میان می‌توان به تعامل بین میزان داده‌های گمشده با نوع الگوریتم اشاره کرد که می‌توان گفت با در نظر گرفتن الگوریتم MHRM و MCEM، مقادیر هر سه سطح داده‌های گمشده بسیار به نزدیک به هم می‌شود و تفاوت معنی‌داری وجود ندارد ولیکن در رابطه با الگوریتم EM تفاوت معنی‌داری بین سه سطح داده گمشده وجود دارد و در مقادیر بالاتر داده گمشده مخاطره بالاتری مشاهده می‌شود.

بنا بر نتایج به‌دست‌آمده به نظر می‌رسد الگوریتم MHRM با استفاده از ترکیب نمونه‌گیری متروپولیس- هستینگز با نمونه‌گیری روبینز-مونرو بیشینه درستنمایی را به‌خوبی تسهیل می‌کند. ازآنجایی‌که الگوریتم MHRM برخلاف MCMC که کل توزیع پسین را تقریب می‌زند، بر برآورد نقطه‌ای و خطاهای استاندارد تمرکز دارد و همین امر اجازه می‌دهد الگوریتم MHRM خیلی سریع‌تر از MCMC و EM به همگرایی برسد. درنتیجه می‌توان گفت ازآنجایی‌که در MHRM، جهش‌ها در زنجیره مارکوف به‌منظور تعیین جهت تغییر در طول تکرارها است، همگرایی سریع‌تر و دقیق‌تر اتفاق می‌افتد. در پایان با توجه به یافته‌های پژوهش می‌توان نتیجه‌گیری کرد که الگوریتم MHRM با در نظر گرفتن عوامل مداخله‌گری چون مکانیسم و نوع داده‌های گمشده، دارای مخاطره‌ی کمتری است و لذا پیشنهاد می‌شود

در تحلیل داده‌های پیچیده که در علوم رفتاری به فراوانی یافت می‌شود، از الگوریتم MHRM استفاده گردد تا برآوردها با خطای کمتری همراه باشند.

تعارض منافع

لازم به ذکر است بین نویسندگان هیچ‌گونه تعارض منافعی وجود ندارد.

سپاسگزاری

مقاله حاضر برگرفته از رساله دکتری رشته سنجش و اندازه‌گیری دانشگاه علامه طباطبایی است. از تمامی اساتید گران‌قدر و دانشجویان دکتری وردی سال ۹۶ گروه سنجش و اندازه‌گیری دانشکده روان‌شناسی دانشگاه علامه طباطبایی تهران که در انجام این پژوهش یاری‌رسان گروه پژوهش بودند، تشکر و قدردانی می‌گردد.

References

Jackson, L. M. (2019). *The psychology of prejudice: From attitudes to social action* (2nd ed.). American Psychological Association

Asparouhov, T., & Muthén, B. (2012). General random effect latent variable modeling: Random subjects, items, contexts, and parameters. In *annual meeting of the National Council on Measurement in Education*, Vancouver, British Columbia.

kk bşş, .. (2017). Examinooon of hhe Effccss of ii ffrrnnt ssss ing Data Techniques on Item Parameters Obtained by CTT and IRT. *International Online Journal of Educational Sciences, 9*(3).

Bartolucci, F., Bacci, S., & Gnaldi, M. (2015). *Statistical analysis of questionnaires: A unified approach based on R and Stata* (Vol. 34). CRC Press.

Bashkov, B. M., & DeMars, C. E. (2017). Examining the performance of the Metropolis–Hastings Robbins–Monro algorithm in the estimation of multilevel multidimensional IRT models. *Applied psychological measurement*, *41*(5), 323-337.

Bulut, O., & SÜNBÜL, Ö. (2017). Monte Carlo Simulation Studies in Item Response Theory with the R Programming Language R Programlama Dili ile Madde Tppki uu rmnnnda oo nee Crr oo mmnnasyon Çşşşşmrrrr ı. *Journal of Measurement and Evaluation in Education and Psychology, 8*(3), 266-287.

Cai, L. (2010). High-dimensional exploratory item factor analysis by a Metropolis–Hastings Robbins–Monro algorithm. *Psychometrika*, *75*(1), 33-57.

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software, 48*(6

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahway.

Dong, Y., & Peng, C. Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus, 2*(1), 1-17.

Enders, C. K. (2010). Applied missing data analysis. Guilford Press.

Embretson, S. E., & Reise, S. P. (2013). *Item response theory*. Psychology Press.

Finch, H. (2008). Estimation of item response theory parameters in the presence of missing data. *Journal of Educational Measurement, 45*(3), 225-245.

Finch, W. H. (2010). Imputation methods for missing categorical questionnaire data: A comparison of approaches. *Journal of Data Science, 8*(3), 361-378.

Gibbons, R. D., Weiss, D. J., Frank, E., & Kupfer, D. (2016). Computerized adaptive diagnosis and testing of mental health disorders. *Annual Review of Clinical Psychology, 12*, 83-104.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1985). Principles and applications of item response theory.

Han, K. T., & Hambleton, R. K. (2014). User's Manual for WinGen 3: Windows Software that Generates IRT Model Parameters and Item Responses (Center for Educational Assessment Report No. 642). *Amherst*, MA: University of Massachusetts.

Liu, Q., & Pierce, D. A. (1994). A note on Gauss—Hermite quadrature. *Biometrika, 81*(3), 624-629.

Kuo, F. Y., & Sloan, I. H. (2005). Lifting the curse of dimensionality. *Notices of the AMS, 52*(11), 1320-1328.

Kuo, T. C., & Sheng, Y. (2016). A comparison of estimation methods for a multi-unidimensional graded response IRT model. *Frontiers in psychology, 7*, 880, 1-29.

Lesaffre, E., & Spiessens, B. (2001). On the effect of the number of quadrature points in a logistic random effects model: an example. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 50*(3), 325-335.

Linden, W. J. & van der, & Hambleton, RK (1997). *Handbook of modern item response theory*, 9-39.

Naylor, J. C., & Smith, A. F. (1982). Applications of a method for the efficient computation of posterior distributions. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 31*(3), 214-225.

Patz, R. J., & Junker, B. W. (1999). Applications and extensions of MCMC in IRT: Multiple item types, missing data, and rated responses. *Journal of educational and behavioral statistics, 24*(4), 342-366.

Robbins, H., & Monro, S. (1951). A stochastic approximation method. *The annals of mathematical statistics*, 400-407.

Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*(3), 581-592.

Schouten, R. M., Lugtig, P., & Vink, G. (2018). Generating missing values for simulation purposes: a multivariate amputation procedure. *Journal of Statistical Computation and Simulation, 88*(15), 2909-2930.

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of statistical software, 45*, 1-67.

Yang, J. S., & Cai, L. (2014). Estimation of contextual effects through nonlinear multilevel latent variable modeling with a Metropolis–Hastings Robbins–Monro algorithm. *Journal of Educational and Behavioral Statistics, 39*(6), 550-582.