Educational Measurement
11(43), 131-152.
Jem.atu.ac.ir
DOI:

ATU PRESS

Identifying the most optimum Logistic distribution-based model for item nonresponse imputation

Alireza Khoshgooyanfard[1] PhD student, measurement and assessment Dep, Allameh Tabatabaei University, Tehran, Iran.

Mohamadreza Falsafi Nejad Associat professor, measurement and assessment Dep, Allameh Tabatabaei University, Tehran, Iran.

Noorali Farrokhi Associat professor, measurement and assessment Dep, Allameh Tabatabaei University Tehran, Iran.

Abstract

Nonresponse is an inevitable challenge to large-scale studies and can result in wasting money, time and human resource involved in data collection and can also prevent the studies from obtaining their objects especially scores distribution. Imputation methods have thus been invented to estimate item nonresponses in order to make inference from a completed data set. Using a simulation study on a real data set in the form of a multivariate experimental design, this paper evaluates the accuracy of three models including cumulative logit model, graded response model and explanatory item response model. The results show that the imputed values of all three models are acceptable under random nonresponse mechanism although the imputed values of the explanatory item response model are always more accurate than those of the other models. If nonrandom nonresponses are occurred, explanatory item response model has acceptable imputed values only at 5% nonresponse rate and the other models are not accurate at all. The results also show that it is more accurate to impute individual item nonresponses and then compute the total score instead of directly imputing the total score.

Keywords: bias; error; large-scale study; nonresponse rate; ordinal item response model.

[1] Corresponding Author: Email Address

How to Cite: xxxxxxx

Original Reserch / Review / ...
Received: 2017/06/17
Accepted: 2020/01/12
eISSN: 2476-6240 ISSN: 2252-004X

مقاله پژوهشی/مروری/سخن سردبیر

تاریخ دریافت: ۹۶/۰۳/۲۷؛ تاریخ پذیرش: ۹۸/۱۰/۲۲

eISSN: 2476-6240 ISSN: 2252-004X

اندازه گیری تربیتی
دوره ۱۱، شماره ۴۳، بهار، ۱۴۰۰، ۱۳۱-۱۵۲
atu.ac.ir
DOI:

شناسایی بهینه‌ترین مدل مبتنی بر توزیع لجستیک برای جایگزینی پرسش‌های بی‌پاسخ*

علیرضا خوشگویان فرد — دانشجوی دکتری، سنجش و اندازه گیری، دانشگاه علامه طباطبائی، تهران، ایران.

محمد رضا فلسفی نژاد — دانشیار گروه سنجش و اندازه گیری، دانشگاه علامه طباطبائی، تهران، ایران.

نورعلی فرخی — دانشیار گروه سنجش و اندازه گیری، دانشگاه علامه طباطبائی، تهران، ایران.

چکیده

بی‌پاسخی چالشی اجتناب‌ناپذیر در برابر مطالعات بزرگ‌مقیاس است و می‌تواند از یک‌سو موجب اتلاف هزینه، زمان و نیروی انسانی درگیر در گردآوری داده‌ها شود و از سوی دیگر، مطالعه را از رسیدن به اهدافش به‌ویژه پارامترهای توزیع نمرات باز دارد. از این‌رو، روش‌های جایگزینی برای برآورد پاسخ پرسش‌های بی‌پاسخ ابداع شده‌اند تا امکان استنباط از یک مجموعه داده کامل‌شده را فراهم آورند. این مقاله از طریق شبیه‌سازی روی یک مجموعه داده واقعی بر اساس یک الگوی طرح آزمایش چندمتغیره، دقت سه مدل جایگزینی را شامل لُجیت‌های تراکمی، پاسخ مدرج و پرسش‌پاسخ تبیینی مورد ارزیابی قرار می‌دهد. یافته‌ها نشان می‌دهند که تحت بی‌پاسخی تصادفی، جایگزینی‌های هر سه مدل در حدی مطلوب قرار دارند هر چند جایگزینی‌های مدل پرسش‌پاسخ تبیینی همواره دقیق‌تر از دو مدل دیگر است. اگر بی‌پاسخی غیرتصادفی باشد، تنها در نرخ بی‌پاسخی ۵ درصد به نتایج مطلوبی برای مدل پرسش‌پاسخ تبیینی دست خواهیم یافت و دقت دو مدل دیگر قابل پذیرش نیست. همچنین، یافته‌ها حاکی از آن هستند که جایگزینی پرسش‌های دارای بی‌پاسخی و محاسبه نمره کل فرد از ترکیب پاسخ‌های جایگزین‌شده و پاسخ‌های واقعی کم‌خطاتر از جایگزینی مستقیم نمره کل او است.

* این مقاله برگرفته از رساله دکترا با عنوان «شناسایی رویکرد بهینهٔ جایگزینی بی‌پاسخی در سنجش سازه‌های روان‌شناختی» است.

کلید واژه‌ها: خطا؛ سوگیری، مدل پرسش‌پاسخ ترتیبی، مطالعه بزرگ‌مقیاس، نرخ بی‌پاسخی، پارامترهای توزیع نمرات.

مقدمه

هزینه، زمان و نیروی انسانی زیادی صرف اجرای مطالعات بزرگ‌مقیاسی[۱] می‌شود که به اندازه‌گیری صفتی در جامعه‌های آماری بزرگ (هم به لحاظ جمعیت و هم جغرافیا) دست می‌زنند. بخش قابل‌توجهی از این منابع به گردآوری داده‌ها یعنی تماس با افراد نمونه در منزل یا محل کار و مصاحبه با آنان برای پاسخگویی به پرسش‌های وسیله اندازه‌گیری اختصاص دارد. از این‌رو، هرگونه ضعف در ساخت وسیله اندازه‌گیری یا فرایند اجرای آن، ضررهای جبران‌ناپذیری را بر مطالعه بزرگ‌مقیاس تحمیل می‌کند؛ ضررهایی که هم شامل اتلاف منابع و هم مانع رسیدن مطالعه به اهداف پژوهشی خود می‌شود.

اغلب، هدف از اجرای مطالعه بزرگ‌مقیاس، هنجاریابی و شناسایی پارامترهای توزیع صفت مورد اندازه‌گیری در جامعه آماری تحت مطالعه برای برنامه‌ریزی و مدیریت مبتنی بر شواهد است؛ بنابراین، دریافت پاسخ پرسش‌های وسیله اندازه‌گیری از تک‌تک افراد نمونه، نقشی مؤثر در موفقیت چنین مطالعاتی ایفا می‌کند. با این حال، مطالعات بزرگ‌مقیاس از سوی عوامل مختلفی نظیر ضعف وسیله اندازه‌گیری، مصاحبه‌گران کم‌مهارت یا ویژگی‌های فرهنگی، اجتماعی، شخصیتی و شناختی افراد نمونه تهدید می‌شود (تورانو، ریپس و رازینسکی[۲] ۲۰۰۰؛ گراوز و همکاران[۳] ۲۰۱۱؛ هاکس[۴] و همکاران ۲۰۰۲؛ فالر[۵] ۲۰۱۳)، زیرا این عوامل به‌تنهایی یا در تعامل با یکدیگر می‌توانند به عدم تمایل فرد نمونه به مشارکت در مطالعه، ایجاد محافظه‌کاری و خودسانسوری در پاسخگویی یا عدم درک پرسش‌ها یا مقوله‌های پاسخ شود به نحوی که چالش‌های جدی در برابر کیفیت داده‌ها و رسیدن به اهداف این مطالعات ایجاد می‌گردد. یکی از این چالش‌های جدی، بی‌پاسخ ماندن یک یا چند پرسش وسیله اندازه‌گیری از سوی افراد از نمونه است (بادنر[۶] ۲۰۰۶)؛ وسیله‌هایی در این پژوهش مورد توجه است که اولاً، پاسخ به پرسش‌های آن به صورت درست یا غلط ارزیابی نشوند (وسیله اندازه‌گیری از نوع امتحان نباشد) و ثانیاً، فرد نمونه واجد شرایط پاسخگویی به پرسش یا پرسش‌های بی‌پاسخ باشد و

---

۱. large-scale
۲. Tourangeau, Rips and Rasinski
۳. Groves
۴. Hox
۵. Fowler
۶. Bodner

مثلاً به دلیل پرش، برخی پرسش‌ها بی‌پاسخ نمانده باشند (مثلاً فرد مجرد است و بخشی از پرسش‌ها به متأهل‌ها اختصاص دارد).

بی‌پاسخی پرسش هم ساخت وسیله اندازه‌گیری و هم اجرای موفقیت‌آمیز آن را در مطالعات بزرگ‌مقیاس با مشکل رو به رو می‌سازد. اصلی‌ترین مشکل، به ناتوانی در محاسبه نمره کل به دلیل اندازه‌گیری ناقص صفت برای فرد دارای پرسش‌های بی‌پاسخ باز می‌گردد. از این‌رو، پژوهشگران و بسیاری از نرم‌افزارهای آماری مانند SPSS، Minitab و Stata به‌طور پیش‌فرض، افراد دارای بی‌پاسخی پرسش را از نمونه حذف می‌کنند تا تنها کسانی در نمونه باقی بمانند که به همه پرسش‌ها پاسخ داده‌اند (راث[۱] ۱۹۹۴؛ آزار[۲] ۲۰۰۲). این راهکار هم هزینه و زمان صرف‌شده برای گردآوری داده از افراد دارای بی‌پاسخی پرسش را به هدر می‌دهد و هم باعث بروز مشکلات جدی‌تری برای برآورد شاخص‌های روان‌سنجی و پارامترهای توزیع نمرات می‌شود (وندر هایدن[۳] ۲۰۰۶).

در واقع، حذف افراد دارای بی‌پاسخی پرسش از نمونه، به کاهش اندازه نمونه یعنی افزایش خطای نمونه‌گیری و در صورت وجود تفاوت بین افراد پاسخگو با افراد بی‌پاسخ، به سوگیری نمونه یعنی عدم انعکاس بخشی از جامعه آماری در نمونه منجر می‌شود. در واقع، چنین نمونه‌ای معرّف جامعه آماری خود نیست و می‌تواند موجب افزایش خطای برآوردهای ضرایب پایایی و روایی یا پارامترهای مربوط به توزیع نمرات شود و حتی این برآوردها را با سوگیری قابل‌توجهی رو به رو می‌سازد. از این‌رو، انجمن روان‌شناسی آمریکا در امور علمی در سال ۱۹۹۹ به ذکر نرخ بی‌پاسخی در مطالعات و دوری از حذف آن‌ها توصیه می‌کند (ویلکینسون[۴] ۱۹۹۹) و از آن زمان، گنجاندن روش‌های جایگزینی[۵] قوی و پیچیده در نرم‌افزارهای آماری و روان‌سنجی بیش از پیش صورت می‌گیرد (پاف و اندرز[۶] ۲۰۰۴). روش‌هایی که برآوردی از پاسخ پرسش بی‌پاسخ به دست می‌دهند تا جایگزین بی‌پاسخی شود (اندرز[۷] ۲۰۱۰).

---

۱. Roth
۲. Azar
۳. Vander Heijden
۴. Wilkinson
۵. imputation
۶. Peugh and Enders
۷. Enders

روش‌های جایگزینی، از روش‌های ساده‌ای مانند جایگزین کردن میانگین پاسخ‌های داده‌شده به‌جای پاسخ‌های داده‌نشده تا روش‌های پیچیده جایگزینی مبتنی بر مدل را در بر می‌گیرند. جایگزین‌های مبتنی بر مدل، برآوردهایی هستند که از برازش یک مدل بر داده‌ها به دست می‌آیند (کریشناموتی، مالیک و متیو[۱] ۲۰۰۹). جایگزینی رگرسیونی یکی از شناخته‌شده‌ترین این روش‌ها است که نمره کل غیرقابل محاسبه فرد دارای پرسش‌های بی‌پاسخ را از طریق یک مدل رگرسیونی برآورد می‌کند؛ نمره کل به‌عنوان متغیر وابسته مدل و متغیرهایی مانند سن، جنس یا تحصیلات به‌عنوان متغیرهای پیش‌بین آن به کار می‌روند. برای این منظور، ضرایب مدل رگرسیونی بر اساس افرادی که دارای بی‌پاسخی در نمره کل نیستند، تعیین و سپس، نمره کل افراد دارای بی‌پاسخی با قرار دادن متغیرهای پیش‌بین مربوط به آن‌ها در مدل، برآورد می‌شود (ژانگ[۲] ۲۰۱۶).

جایگزینی رگرسیونی، نمره کل فرد دارای بی‌پاسخی پرسش را به‌طور مستقیم برآورد می‌کند بی‌آن‌که اهمیتی دهد فرد به چند پرسش از وسیله اندازه‌گیری پاسخ و به چند پرسش پاسخ نداده است، در نهایت نمره کل او مورد جایگزینی قرار می‌گیرد و پرسش‌های دارای پاسخ او که اطلاعاتی «واقعی» از فرد به شمار می‌آیند، بدون استفاده باقی می‌مانند. راهکاری که برای مقابله با بی‌پاسخی پرسش در این مقاله دنبال می‌شود جایگزینی پرسش‌های بی‌پاسخ و استفاده همزمان از پرسش‌های دارای پاسخ و پرسش‌های جایگزین‌شده برای محاسبه نمره کل است؛ به این ترتیب، اگر فردی به برخی پرسش‌ها پاسخ دهد و به برخی پاسخ ندهد، ابتدا پاسخ‌های او به پرسش‌های بی‌پاسخ برآورد (جایگزین) می‌شوند و سپس، پاسخ‌های برآوردشده (جایگزین‌شده) با پاسخ‌های پرسش‌های دارای پاسخ ترکیب می‌شوند تا در محاسبه نمره کل، پاسخ‌های واقعی فرد از پرسش‌های دارای جواب نیز نقش داشته باشند. این امکان وجود دارد که نمره کل حاصل دقیق‌تر از نمره کلی باشد که به‌طور مستقیم برآورد (جایگزین) می‌شود زیرا در برآورد مستقیم نمره کل هیچ استفاده‌ای از پرسش‌هایی که فرد به آن‌ها پاسخ داده است، نمی‌شود و همه پرسش‌ها هم دارای پاسخ و هم بی‌پاسخ نادیده گرفته می‌شوند.

تحقق راهکار جایگزینی مستلزم برآورد پاسخ در خُردترین سطح وسیله اندازه‌گیری یعنی پرسش است که اغلب، نوعی متغیر مقوله‌ای (اسمی یا ترتیبی) با محدودیت‌های

---

۱. Krishnamoorthy, Mallick and Mathew
۲. Zhang

محاسباتی به شمار می‌آید. به دلیل این محدودیت‌ها، نمی‌توان پاسخ پرسش را به‌طور مستقیم از طریق یک مدل رگرسیونی، برآورد و جایگزین کرد زیرا علاوه بر محدودیت‌های محاسباتی چنین متغیرهایی، هیچ تضمینی وجود ندارد که پاسخ حاصل از مدل در دامنه پاسخ‌ها قرار گیرد (برای مثال، مدل ممکن است برای پرسشی با چهار مقوله پاسخ که به صورت ۱، ۲، ۳ و ۴ نمره‌دهی می‌شود، مقدار ۵ یا ۳/۵ را به‌عنوان جایگزین به دست دهد که هیچ‌یک منطبق بر مقادیر ۱، ۲، ۳ و ۴ نیستند). از این‌رو، احتمال انتخاب مقوله‌های مختلفِ پاسخ از طریق مدل‌های مربوط به «متغیرهای مقوله‌ای» محاسبه و محتمل‌ترین مقوله به‌عنوان مقدار جایگزین برگزیده می‌شود.

روش‌های «آماری» مختلفی برای جایگزینی پرسش‌هایی با طیف پاسخ مقوله‌ای قابل استفاده هستند از جمله تحلیل تشخیصی[۱] خوشه‌بندی[۲] یا استفاده از مدل رگرسیون لجستیک. با وجود این، به نظر می‌رسد مدل‌های حوزه «روان‌سنجی» نیز قابلیت جایگزینی را دارا باشند زیرا بین احتمال انتخاب مقوله‌های پاسخ و پارامتر توانایی فرد ارتباط برقرار می‌کنند و به این ترتیب، می‌توان بر اساس مقدار توانایی فرد، محتمل‌ترین گزینه‌ای انتخابی او را شناسایی کرد. از این‌رو، در این مقاله، علاوه بر مدل آماری شناخته‌شده لُجیت‌های تراکمی[۳](همان مدل رگرسیون لجستیک برای متغیرهای ترتیبی)، از مدل پاسخ مدرج سامی‌جیما[۴] و مدل پرسش‌پاسخ تبیینی[۵] نیز که به حوزه روان‌سنجی تعلق دارند، برای اولین‌بار برای تعیین احتمال انتخاب مقوله‌های مختلف پاسخ برای پرسشی با مقوله‌های پاسخ ترتیبی مانند طیف لیکرت استفاده و کاربردی تازه از خانواده مدل‌های پرسش‌پاسخ ارائه می‌شود.

مدل لُجیت‌های تراکمی نوعی مدل رگرسیون لجستیک برای متغیرهای ترتیبی است که احتمال انتخاب یک مقوله یا مقوله‌های پیش از آن (احتمال تراکمی) را با استفاده از متغیرهای پیش‌بین مدل‌سازی می‌کند (اُکانل[۶] ۲۰۰۷). مدل پاسخ مدرج، تعمیمی از مدل پرسش‌پاسخ دو پارامتری برای پرسش‌هایی با مقوله‌های پاسخ ترتیبی است. این مدل نیز احتمال تراکمی انتخاب یک مقوله را مدل‌سازی می‌کند با این تفاوت که از مقدار توانایی

۱. discriminant analysis
۲. cluster analysis
۳. cumulative logits model
۴. Samejima’s graded response model
۵. explanatory item response model
۶. O’Connel

فرد به‌عنوان پیش‌بین استفاده می‌شود[۱](سامی‌جیما، ۲۰۰۸). مدل پرسش و پاسخ تبیینی، شکل عمومی هر دو مدل رگرسیون و پرسش‌پاسخ است به نحوی که همزمان از توانایی و متغیرهای دیگر به‌عنوان پیش‌بین بهره می‌برد. در واقع، مدل‌های سه‌گانه به خانواده مدل‌های خطی تعمیم‌یافته تعلق دارند و همگی از تابع پیوند[۲]لُجیت (لگاریتم نسبت دو احتمال) استفاده می‌کنند (دو بوئک و ویلسون[۳] ۲۰۰۴) و به همین دلیل در این مقاله مدل‌های مبتنی بر توزیع لجستیک نامیده می‌شوند. با وجود تعلق این مدل‌ها به یک خانواده از مدل‌ها، انتظار نمی‌رود همگی آن‌ها از دقت یکسانی در جایگزینی برخوردار باشند، زیرا حداقل نوع استفاده از پیش‌بین‌ها در آن‌ها متفاوت است.

از آنچه گذشت می‌توان دریافت که دقت جایگزینی می‌تواند اولاً به نوع مدل انتخابی برای جایگزینی و ثانیاً، به رویکردِ جایگزینی یعنی رویکرد خُرد (جایگزینی پرسش‌ها و سپس محاسبه نمره کل) یا جایگزینی مستقیم نمره کل بستگی داشته باشد. علاوه بر تنوع مدل‌ها و این دو رویکرد، دو عامل دیگر را نیز می‌توان در دقت جایگزینی نقش‌آفرین قلمداد کرد: نرخ بی‌پاسخی و نوع آن. نرخ بی‌پاسخی نشان‌دهنده نسبت موارد بی‌پاسخ در مجموعه داده‌ها است و نوع بی‌پاسخی می‌تواند تصادفی (سیمن[۴] ۲۰۱۳) یا غیرتصادفی یعنی دارای سوگیری (تانگ[۵] ۲۰۰۳) باشد به این معنا که قشر خاصی از جامعه (مثلاً زنان بیشتر از مردان، افراد کم‌سواد بیشتر از افراد دارای تحصیلات یا افراد محافظه‌کار بیشتر از افراد غیرمحافظه‌کار)، تمایل بیشتری به بی‌پاسخ گذاشتن پرسش داشته باشند. در حالت تصادفی، پاسخ دادن یا ندادن به پرسش امری تصادفی تلقی می‌شود به نحوی که پس از حذف موارد بی‌پاسخ نیز نمونه باقی‌مانده همچنان نمونه‌ای تصادفی از جامعه آماری تحت مطالعه است در حالی که در بی‌پاسخی غیرتصادفی، افراد بی‌پاسخ ویژگی‌های خاصی دارند و بین آنان و افراد پاسخگو تفاوت جدی وجود دارد، لذا بی‌پاسخی منجر به سوگیری در نمونه و استنباط‌های حاصل از آن می‌شود.

دور از انتظار نیست که دو عامل پیش‌گفته دارای تأثیرات قابل‌توجهی بر دقت جایگزینی باشند زیرا از یک‌سو، نرخ بی‌پاسخیِ کم به معنای دسترسی به پاسخ افراد

---

۱. سامی‌جیما از یک تابع کلی در معرفی مدل استفاده می‌کند ولی در اینجا، تابع لجستیک به کار رفته است.

۲. link function

۳. De Boeck and Wilson

۴. Seaman

۵. Tang

بیشتری از نمونه است که برازش مدل‌های جایگزینی را عینی‌تر و واقعی‌تر می‌سازد و از سوی دیگر، بی‌پاسخی غیرتصادفی به معنای در دسترس نبودن پاسخ‌های گروهی از افراد است که پاسخ‌هایی متفاوت از سایر افراد نمونه دارند. پس درمجموع می‌توان مدل‌های جایگزینی را به لحاظ نوع متغیرهای پیش‌بین (متغیر کمکی، توانایی یا هر دو)، نرخ بی‌پاسخی (درصد افراد دارای بی‌پاسخی) و نوع بی‌پاسخی (تصادفی یا غیرتصادفی) را در دقت جایگزینی دخیل دانست. از این‌رو، این مقاله در پی پاسخگویی به پرسش‌های زیر است:

۱ـ دقت کدام‌یک از سه نوع مدل مبتنی بر توزیع لجستیک بیشتر است؟

۲ـ چه تغییری در دقت مدل‌های سه‌گانه تحت بی‌پاسخی کاملاً تصادفی و غیرتصادفی ایجاد می‌شود؟

۳ـ چه تغییری در دقت مدل‌های سه‌گانه تحت نرخ‌های مختلف بی‌پاسخی رخ می‌دهد؟

۴ـ دقت جایگزینی مستقیم نمره کل در مقایسه با جایگزینی پرسش‌ها چگونه است؟

دقت مدل‌های سه‌گانه از طریق شبیه‌سازی و در قالب یک الگوی طرح آزمایش سه متغیره دوعاملی (MANOVA) مطابق جدول ۱ و با استفاده از نرم‌افزار SAS/IML ارزیابی می‌شود. شبیه‌سازی با ایجاد موقعیت‌های متعدد در مجموعه داده‌ها، شرایط بررسی تأثیرات این موقعیت‌ها را فراهم می‌سازد. برای این منظور، نرخ‌های مختلف بی‌پاسخی (۵، ۱۵ و ۲۵ درصد) روی یک مجموعه داده واقعی در دو وضعیت بی‌پاسخی کاملاً تصادفی و غیرتصادفی اعمال می‌شوند به نحوی که ابتدا تعدادی از پاسخ‌های واقعی افراد تبدیل به موارد بی‌پاسخ و سپس، از طریق هر سه مدل جایگزین می‌شوند. برای مثال، بی‌پاسخی ۵ درصد، با انتخاب تصادفی ۵ درصد از افراد نمونه (برای وضعیت تصادفی) یا انتخاب غیرتصادفی آنان (برای وضعیت غیرتصادفی) و سپس، حذف حداقل یکی از پاسخ‌های هر یک از آن‌ها به‌تصادف صورت می‌گیرد. به این ترتیب، می‌توان دقت مدل‌ها را از مقایسه مقدار جایگزین‌شده با پاسخ واقعی افراد مشخص کرد. عملیات انتخاب افراد و تبدیل پاسخ‌ها به بی‌پاسخی، ۱۰۰۰ بار برای هر یک از موقعیت‌های آزمایش تکرار می‌شود تا ارزیابی دقیق‌تری از دقت هر یک از مدل‌ها به دست آید.

جدول ۱. الگوی آزمایش سه متغیره دوعاملی برای ارزیابی مدل‌های سه‌گانه (متغیرهای پاسخ: دقت مدل لُجیت‌های تراکمی CLM، مدل پاسخ مدرج GRM و مدل پرسش‌پاسخ تبیینی EIRM)

نرخ بی‌پاسخی / نوع بی‌پاسخی	۵ درصد	۱۰ درصد	۱۵ درصد
کاملاً تصادفی	CLM GRM EIRM	CLM GRM EIRM	CLM GRM EIRM
غیرتصادفی	CLM GRM EIRM	CLM GRM EIRM	CLM GRM EIRM

برای ارزیابی دقت برآوردها تحت دو وضعیت بی‌پاسخی تصادفی و غیرتصادفی، اعمال نرخ‌های مختلف بی‌پاسخی یک‌بار به شکل کاملاً تصادفی (انتخاب صرفاً «تصادفی» افراد و تبدیل حداقل یکی از پاسخ‌های آنان به حالت بی‌پاسخ) و یک‌بار با اعمال بی‌پاسخی در میان افراد دارای نمرات بالا یا پایین صورت می‌گیرد. در حالت اخیر، بی‌پاسخی موجب سوگیری در نمونه می‌شود زیرا بی‌پاسخی به صفت مورد اندازه‌گیری بستگی دارد و پاسخ‌های قشر خاصی از جامعه یعنی گروه دارندگان نمرات بالا یا پایین به شکل معرّفی در نمونه منعکس نمی‌شود؛ بنابراین، ۶ موقعیت بی‌پاسخی مطابق جدول ۱ ایجاد و در هر موقعیت، ۳ مدل جایگزینی استفاده می‌شود.

برای مقایسه دقت جایگزینی براساس رویکرد جایگزینی مستقیم نمره کل و جایگزینی پرسش‌ها (جایگزینی پرسش‌ها و سپس محاسبه نمره کل)، بهترین مدل مبتنی بر توزیع لجستیک که از ارزیابی مدل‌های سه‌گانه به دست آمده است با مدل رگرسیونی برای جایگزینی مستقیم نمره کل مورد بررسی قرار می‌گیرد. برای این منظور، اختلاف میانگین توزیع نمرات حاصل از جایگزینی با میانگین توزیع نمرات حاصل از مجموعه داده کامل در قالب یک الگوی طرح آزمایش دومتغیره دوعاملی مطابق جدول ۲ مورد مقایسه قرار می‌گیرد؛ بنابراین، رویکردی برای جایگزینی مناسب قلمداد می‌شود که اختلاف کمتری با میانگین نمرات حاصل از مجموعه داده کامل داشته باشد.

جدول ۲. الگوی آزمایش دومتغیره دوعاملی برای مقایسه دو رویکرد مستقیم و خُرد (متغیرهای پاسخ: دقت مدل رگرسیونی RM و مدل پرسش‌پاسخ تبیینی EIRM)

نرخ بی‌پاسخی / نوع بی‌پاسخی	۵ درصد	۱۰ درصد	۱۵ درصد
کاملاً تصادفی	RM	RM	RM

EIRM	EIRM	EIRM	
RM EIRM	RM EIRM	RM EIRM	غیرتصادفی

جامعه آماری این تحقیق شامل تمام موقعیت‌های بی‌پاسخی در سنجش سازه‌های روان‌شناختی است اعم از شرایط اجرا، آزمودنی‌ها، زمان و پرسش‌ها؛ بنابراین با یک ابرجامعه[۱]سروکار داریم که در عمل، تنها یکی از این موقعیت‌ها یعنی زیرجامعه‌ای[۲]از آن برای این پژوهش در دسترس خواهد بود. با وجود این، شبیه‌سازی به‌عنوان روش گردآوری داده‌ها کمک خواهد کرد تا از این زیرجامعه وضعیت‌های بالقوه متعددی ایجاد و شرایط یک مطالعه آزمایشی فراهم شود.

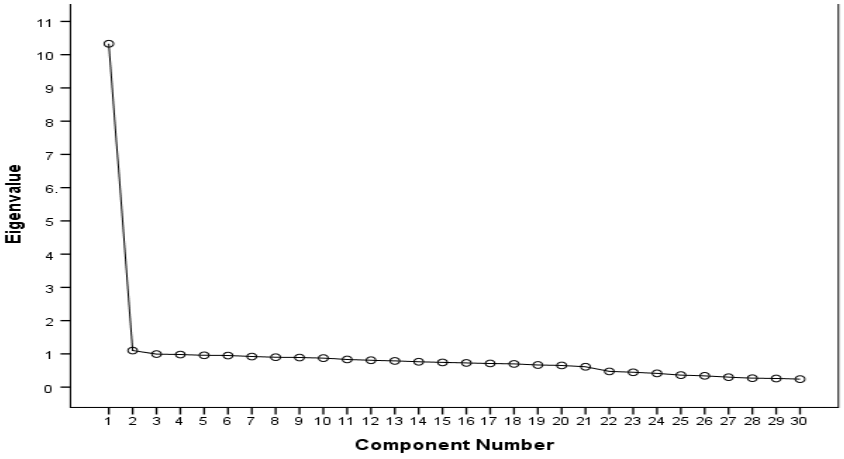
داده‌هایی که بی‌پاسخی در آن‌ها اعمال می‌شود به اجرای پرسشنامه‌ای شامل ۳۰ پرسش تعلق دارند که «امید به آینده» را از ۱۳۸۶۷ نفر ساکن در مراکز استان‌های کشور با نمونه‌گیری خوشه‌ای سه مرحله‌ای اندازه گرفته است. هر پرسش دارای طیف پاسخ ترتیبی «اصلاً»، «کم»، «متوسط» و «زیاد» است که از ۱ تا ۴ امتیازدهی می‌شوند، لذا از مدل‌های سه‌گانه پیش‌گفته استفاده می‌شود که مربوط به پرسش‌هایی با مقوله‌های پاسخ ترتیبی هستند. متغیر پیش‌بین مدل‌ها و مدل رگرسیون را نمره حاصل از ۱۰ پرسش دیگر تشکیل می‌دهد که «رضایت از وضعیت کنونی زندگی» را اندازه می‌گیرد زیرا همبستگی این نمره با نمره امید به آیند برابر با ۰/۵۹۶ است؛ بنابراین، متغیر پیش‌بین یکسانی در همه مدل‌ها به کار رفته است تا اختلاف دقت آن‌ها با متغیر پیش‌بین ارتباطی نداشته باشد.

یافته‌ها در دو قسمت ارائه می‌شوند: یافته‌های مرتبط با مقایسه مدل‌های سه‌گانه (پرسش‌های ۱ تا ۳) و یافته‌های مرتبط با مقایسه دو رویکرد جایگزینی نمره کل و جایگزینی خُرد (پرسش ۴). پیش از آن، تک‌بعدی بودن وسیله اندازه‌گیری «امید به آینده» به‌عنوان یکی از مفروضات مدل‌های پرسش‌پاسخ مورد بررسی قرار می‌گیرد. شکل ۱، نمودار اِسکری[۳] مربوط به ۳۰ پرسش سنجش «امید به آینده» را نمایش می‌دهد که آشکارا دلالت بر وجود یک مؤلفه جامع دارد که نسبت واریانس بیشتری را در مقایسه با مؤلفه‌های دیگر به خود اختصاص می‌دهد.

---

۱. super population
۲. subpopulation
۳. scree plot

شکل ۱. نمودار اسکری برای تشخیص تک‌بعدی بودن وسیله اندازه‌گیری

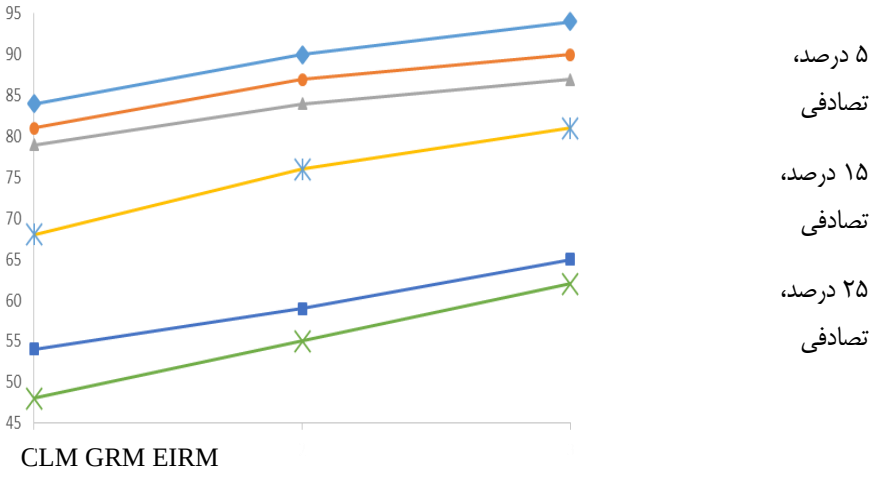


مقایسه مدل‌های سه‌گانه: تحلیل واریانس سه‌متغیره روی دقت جایگزینی مدل‌های سه‌گانه مطابق الگوی جدول ۱ نشان می‌دهد که «نرخ بی‌پاسخی» و «نوع بی‌پاسخی» دارای اثر متقابل روی دقت جایگزینی مدل‌های سه‌گانه هستند ($Wilk's\ Lambda = ۰/۲۳۹,\ p = ۰/۰۰۰$). شکل ۲ دقت این مدل‌ها را تحت سه نرخ بی‌پاسخی ۵، ۱۵ و ۲۵ درصد در دو وضعیت بی‌پاسخی کاملاً تصادفی و غیرتصادفی نشان می‌دهد. دقت هر سه مدل تحت بی‌پاسخی تصادفی همواره بیشتر از بی‌پاسخی غیرتصادفی است.

آزمون تجزیه واریانس یک‌متغیره برای مقایسه دقت هر یک از مدل‌ها در موقعیت‌های شش‌گانه (و همچنین آزمون‌های تعقیبی متناظر با آن‌ها که در این مقاله گزارش نشده‌اند)، اختلاف معناداری را در سطح ۱ درصد بین دقت جایگزینی مدل لُجیت‌های تراکمی ($F = ۲۷۷۶۸,\ p = ۰/۰۰۰$)، دقت جایگزینی مدل پاسخ مدرج سامی‌جیما ($F = ۲۶۵۳۰,\ p = ۰/۰۰۰$) و دقت جایگزینی مدل پرسش‌پاسخ تبیینی ($F = ۲۰۵۲۸,\ p = ۰/۰۰۰$) نشان می‌دهد. با توجه به شکل ۲، مدل پرسش‌پاسخ تبیینی تحت بی‌پاسخی تصادفیِ ۵ درصد تا ۹۴ درصد به درستی پاسخ را جایگزین می‌کند که رقم قابل‌توجهی است. همان‌طور که انتظار داریم، با افزایش نرخ بی‌پاسخی، دقت این مدل کاهش می‌یابد به نحوی که در نرخ‌های ۱۵ و ۲۵ درصد، جایگزینی دقیق به ترتیب به ۹۰ و

۸۷ درصد کاهش می‌یابد ولی همچنان رقم مطلوبی به شمار می‌آید. به‌طور کلی، کم‌دقت‌ترین وضعیت (۷۹ درصد) به‌جایگزینی مدل لُجیت‌های تراکمی برای نرخ بی‌پاسخیِ تصادفی ۲۵ درصد تعلق دارد.

شکل ۲. دقت مدل لُجیت‌های تراکمی (CLM)، مدل پاسخ مدرج (GRM) و مدل پرسش‌پاسخ تبیینی (EIRM) تحت نرخ‌های بی‌پاسخی ۵، ۱۵ و ۲۵ درصد در دو وضعیت بی‌پاسخی تصادفی و غیرتصادفی



جدول ۳ نتایج آزمون آماری برای مقایسه دقت جایگزینی مدل‌های سه‌گانه را در تک‌تک موقعیت‌های شش‌گانه بی‌پاسخی ارائه می‌دهد. همان‌طور که مشاهده می‌شود تمام آزمون‌ها در سطح ۱ درصد معنادار هستند. با مراجعه به شکل ۲، می‌توان دریافت که همواره دقت مدل پرسش‌پاسخ تبیینی بیشتر از دو مدل دیگر و دقت مدل پاسخ مدرج سامی‌جیما بیشتر از مدل لُجیت‌های تراکمی است. ترتیب دقت مدل‌های سه‌گانه برای بی‌پاسخی غیرتصادفی همانند بی‌پاسخی تصادفی است با این تفاوت که مقدار دقت همواره در حد مطلوبی قرار ندارد. جز دقت مدل پرسش‌پاسخ تبیینی در نرخ بی‌پاسخیِ ۵ درصد که به ۸۱ درصد می‌رسد، دقت مدل‌ها در وضعیت‌های دیگر کمتر از ۸۰ درصد است. مدل لُجیت‌های تراکمی از کم‌ترین دقت برخوردار است به نحوی که در نرخ‌های بی‌پاسخی غیرتصادفی ۵، ۱۵ و ۲۵ درصد به ترتیب دارای دقت ۶۶، ۵۴ و ۴۸ درصد است. به‌طور

کلی، رفتار مدل‌های سه‌گانه در نرخ بی‌پاسخی غیرتصادفی ۵ درصد به حالت تصادفی نزدیک‌تر از نرخ‌های دیگر حالت غیرتصادفی است.

در هر دو حالتِ بی‌پاسخی تصادفی و غیرتصادفی، همواره دقت مدل پرسش‌پاسخ تبیینی بیشتر از هر دو مدل پاسخ مدرج و لُجیت‌های تراکمی است. همچنین، دقت مدل پاسخ مدرج همواره از دقت مدل لُجیت‌های تراکمی بیشتر است. دقت مدل پاسخ مدرج به مدل پرسش‌پاسخ تبیینی نزدیکتر از مدل لُجیت‌های تراکمی است. به نظر می‌رسد استفاده از توانایی به‌عنوان پیش‌بین مؤثرتر از استفادهٔ صرف از متغیر کمکی دیگری مانند «رضایت از وضعیت کنونی زندگی» باشد.

جدول ۳. آزمون مقایسه دقت جایگزینی مدل‌های سه‌گانه در موقعیت‌های مختلف بی‌پاسخی

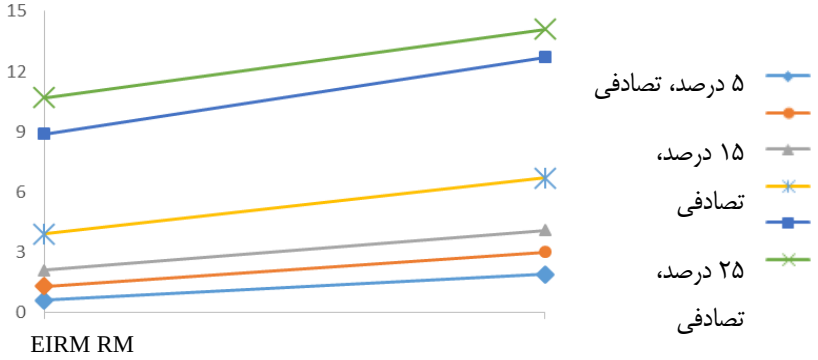
نوع نرخ	تصادفی	غیرتصادفی
۵ درصد	Wilk's Lambda = ۰/۰۰۰۳۵, p = ۰/۰۰۰	Wilk's Lambda = ۰/۰۰۱, p = ۰/۰۰۰
۱۵ درصد	Wilk's Lambda = ۰/۰۰۰۳۶, p = ۰/۰۰۰	Wilk's Lambda = ۰/۰۰۰۷, p = ۰/۰۰۰
۲۵ درصد	Wilk's Lambda = ۰/۰۰۰٤, p = ۰/۰۰۰	Wilk's Lambda = ۰/۰۰۰۸, p = ۰/۰۰۰

مقایسه جایگزینی نمره کل و رویکرد خُرد: با توجه به یافته‌های قسمت قبل، مدل پرسش‌پاسخ تبیینی نسبت به دو مدل دیگر از دقت بیشتری برخوردار بود. پس در این قسمت، نتیجهٔ جایگزینی پرسش‌های بی‌پاسخ از طریق این مدل و محاسبه نمره کل از ترکیب پاسخ‌های جایگزین‌شده و پاسخ‌های واقعی با جایگزینی مستقیم نمره کل از طریق یک مدل رگرسیون خطی مورد مقایسه قرار می‌گیرند. برای این منظور، قدرمطلق اختلاف میانگین نمرات کل حاصل از دو رویکرد جایگزینی با مقدار حاصل از نمرات نمونه کامل مقایسه می‌شوند.

تحلیل واریانس دومتغیره روی دقت جایگزینی‌های دو مدل نشان می‌دهد که «نرخ بی‌پاسخی» و «نوع بی‌پاسخی» دارای اثر متقابل روی قدرمطلق اختلاف میانگین نمرات مربوط به این دو رویکرد هستند (Wilk's Lambda = ۰/۰۰۰۲۱, p = ۰/۰۰۰). شکل ۳،

قدرمطلق خطای دو رویکرد جایگزینی را برای سه نرخ بی‌پاسخی ۵، ۱۵ و ۲۵ درصد در دو وضعیت بی‌پاسخی تصادفی و غیرتصادفی نشان می‌دهد. به سرعت می‌توان به چند نکته اساسی در این شکل دست یافت. نخست اینکه، بر اساس آزمون‌های آماری جدول ۴ و شکل ۳، همواره میانگین حاصل از رویکرد خُرد (EIRM) به میانگین داده‌های کامل نزدیک‌تر از رویکرد جایگزینی مستقیم (RM) است. دیگر اینکه، خطای هر دو رویکرد در بی‌پاسخی تصادفی کمتر از بی‌پاسخی غیرتصادفی است. در نهایت، با افزایش نرخ بی‌پاسخی هم بر خطای رویکرد مستقیم ($F=۱۷۲۹۱۲۵۵, p=۰/۰۰۰$) و هم بر خطای رویکرد خُرد ($F=۳۴۳۶۱۷۱۲, p=۰/۰۰۰$) افزوده می‌شود.

شکل ۳. مقایسه خطای برآورد میانگین نمرات از جایگزینی مستقیم نمره کل با مدل رگرسیونی (RM) و جایگزینی پرسش‌ها با مدل پرسش‌پاسخ تبیینی (EIRM)



این نمودار نشان می‌دهد که خطای هر دو رویکرد در نرخ بی‌پاسخی تصادفیِ ۵ درصد قابل‌قبول است به نحوی که قدرمطلق خطای به کارگیری مدل پرسش‌پاسخ تبیینی حداکثر به ۰/۶ و مدل رگرسیونی به ۱/۹ می‌رسد. به علاوه، قدرمطلق خطا در نرخ بی‌پاسخی تصادفیِ ۱۵ درصد به ترتیب به ۱/۳ و ۳ و در نرخ بی‌پاسخی تصادفیِ ۲۵ درصد به ترتیب به ۲/۱ و ۴/۱ می‌رسد. پذیرش این مقدار از خطا در مقایسه با خطای حذف افراد دارای بی‌پاسخ از نمونه به مراتب قابل‌تحمل‌تر و منطقی‌تر است. در مقایسه با حالت تصادفی، وضعیت برای نرخ‌های مختلف بی‌پاسخی غیرتصادفی مطلوب به نظر نمی‌رسد. جز در نرخ بی‌پاسخی غیرتصادفیِ ۵ درصد که خطای مدل پرسش‌پاسخ تبیینی به ۳/۹ می‌رسد، خطای سایر وضعیت‌ها بیشتر از ۶/۷ است هرچند این مقدار خطا نیز بهتر از خطای حذف افراد دارای بی‌پاسخ از نمونه است.

اختلاف خطای دو رویکرد در بی‌پاسخی غیرتصادفی همواره بیشتر از بی‌پاسخی تصادفی است به نحوی که حداقل و حداکثر اختلاف برای بی‌پاسخی غیرتصادفی به ترتیب برابر با ۱/۳ و ۲ نمره و برای بی‌پاسخی غیرتصادفی به ترتیب برابر با ۲/۸ و ۴/۱ نمره است. همچنین، خطای هر دو رویکرد در نرخ بی‌پاسخی غیرتصادفیِ ۵ درصد به خطای بی‌پاسخی تصادفی نزدیک‌تر از بی‌پاسخی غیرتصادفی است.

جدول ٤. آزمون مقایسه دقت جایگزینی نمره کل توسط دو رویکرد خُرد و مستقیم

نوع نرخ	تصادفی	غیرتصادفی
۵ درصد	Wilk's Lambda = ۰/۰۰۲۲, p = ۰/۰۰۰	Wilk's Lambda = ۰/۰۰۰۰۱, p = ۰/۰۰۰
۱۵ درصد	Wilk's Lambda = ۰/۰۰۰۰۰۹, p = ۰/۰۰۰	Wilk's Lambda = ۰/۰۰۰۰۰۰۳, p = ۰/۰۰۰
۲۵ درصد	Wilk's Lambda = ۰/۰۰۰۰۵, p = ۰/۰۰۰	Wilk's Lambda = ۰/۰۰۰۰۲, p = ۰/۰۰۰

بحث و نتیجه‌گیری

مسئله بی‌پاسخی دارای پیشینه طولانی در حوزه آمار است به نحوی که چندین کتاب تخصصی در این زمینه نگارش یافته است (رافاناتان[1] ۲۰۱۶؛ مولنبرگز[2] و همکاران، ۲۰۱۴) ولی این مسئله در حوزه روان‌سنجی دارای پیامدهای ویژه‌ای است که آن را از حوزه آمار متمایز می‌سازد. به لحاظ آماری، اُفت نمونه حاصل از حذف افراد دارای بی‌پاسخ به افزایش خطای برآوردها و کاهش توان آزمون‌های آماری منجر می‌شود و در صورت وجود تفاوت عمده بین افراد بی‌پاسخ و پاسخگو با سوگیری نمونه و استنباط‌های آماری و روان‌سنجی حاصل از آن نیز رو به رو خواهیم بود.

هنگامی که این مسئله در مطالعات بزرگ‌مقیاس روان‌شناختی رخ می‌دهد هم با پیامدهای آماری آن درگیر خواهیم شد و هم با پیامدهای خاصی که به لحاظ روان‌سنجی از اهمیت ویژه‌ای برخوردارند مانند مشکل نمره‌گذاری و هنجاریابی نمرات (اندرز، ۲۰۱۰). اغلب وسیله‌های اندازه‌گیری متشکل از تعدادی پرسش هستند که با هم صفتی را اندازه می‌گیرند؛ آنچه اهمیت دارد نمره‌ای است که از کل پرسش‌های وسیله اندازه‌گیری به دست می‌آید و پرسش‌ها تنها نشانگرهایی[3] برای آن صفت به شمار می‌آیند. این بدان معنا است که تک‌تک پرسش‌ها به صورت انفرادی دارای اهمیت نیستند به نحوی که به‌جای هر یک از این پرسش‌ها، پرسش دیگری نیز می‌توانست طراحی شود که همان نقش نشانگری را دارا باشد.

---

۱. Raghunathan
۲. Molenberghs
۳. indicator

این ویژگی وسیله‌های اندازه‌گیری روان‌شناختی، آن‌ها را متمایز از پرسشنامه‌هایی می‌سازد که پرسش‌های جداگانه‌ای با اهداف مختلف را شامل می‌شوند. در این پرسشنامه‌ها، مجموعه پرسش‌ها صفتی را اندازه نمی‌گیرند بلکه هر پرسش یک قلم اطلاعاتی را از فرد نمونه به دست می‌دهد که به‌تنهایی مورد توجه است؛ بنابراین، بی‌پاسخ ماندن یک پرسش از پرسشنامه خللی به پرسش‌های دیگر وارد نمی‌کند و آن‌ها همچنان می‌توانند کارکرد خود را در پرسشنامه دارا باشند. در مقابل، وسیله‌های اندازه‌گیری روان‌شناختی متشکل از پرسش‌های به هم پیوسته‌ای هستند که بی‌پاسخ ماندن حتی یک پرسش نیز فرایند نمره‌گذاری را مختل می‌سازد. در واقع، بی‌پاسخی باعث می‌شود تا محاسبه نمره کل برای فرد دارای بی‌پاسخی ناممکن شود و در عمل، چنین فردی از نمونه کنار گذاشته می‌شود و این آغاز بروز مشکلات آماری حاصل از بی‌پاسخی است.

از این‌رو، بی‌پاسخی چالشی در مقابل اندازه‌گیری در مطالعات بزرگ‌مقیاس روان‌شناختی است و معمولاً با حذف افراد دارای پرسش بی‌پاسخ نادیده گرفته می‌شود، در حالی که این حذف می‌تواند به افزایش خطای نمونه‌گیری یا سوگیری نتایج منجر شود؛ بنابراین، در صورت بروز بی‌پاسخی هم ساخت وسیله اندازه‌گیری و هم به کارگیری آن با مشکل جدی رو به رو خواهد بود (اشلومر، بومن، کارد[1] ۲۰۱۰). به این ترتیب، روش‌های جایگزینی ابداع شده‌اند تا به‌جای حذف افراد از نمونه و تحمیل خطای ناشی از اُفت و سوگیری نمونه بر استنباط‌ها، پاسخ‌های محتمل فرد دارای بی‌پاسخی را برآورد و جایگزین پرسش‌های بی‌پاسخ او کنند تا مجموعه داده‌ای ایجاد شود که دارای خانه‌های خالی نیست. با وجود این، نباید یک مجموعه داده «کامل‌شده» از طریق جایگزینی را همانند یک مجموعه داده «کامل» قلمداد کرد که از ابتدا دارای موارد بی‌پاسخی نیست. از این‌رو، این مقاله ضمن معرفی کاربردی تازه از خانواده مدل‌های پرسش‌پاسخ، دقت سه مدل در جایگزینی پرسش‌های بی‌پاسخ و میزان قابل‌اعتماد بودن برآوردهای آن‌ها از پاسخ‌ها را مورد ارزیابی قرار داد.

در این مقاله از سه مدل مبتنی بر توزیع لجستیک برای برآورد احتمال انتخاب مقوله‌های پاسخ استفاده شد. هر یک از این مدل‌ها دارای ظرفیت‌ها و محدودیت‌هایی در جایگزینی موارد بی‌پاسخ هستند. موفقیت مدل لُجیت‌های تراکمی به متغیرهای پیش‌بین مورد استفاده بستگی دارد؛ متغیرهایی که ارتباطی نسبتاً قوی با پرسش تحت جایگزینی

---

[1] Schlomer, Bauman and Card

داشته باشند تا بتوانند پیش‌بینی محتمل‌تر و دقیق‌تری را به دست دهند در حالی که پژوهشگر در دسترسی به چنین متغیرهایی محدودیت دارد زیرا وسیله‌های اندازه‌گیری مورد استفاده در مطالعات بزرگ‌مقیاس با هدف خاصی ساخته می‌شوند و پرسش‌ها و متغیرهای مرتبط با آن هدف را اندازه می‌گیرند.

در این مقاله از نمره حاصل از ۱۰ پرسش برای پیش‌بینی استفاده شد که صفتی مرتبط با صفت اصلی مورد مطالعه را اندازه می‌گرفتند، در حالی که فرصت افزودن پرسش‌های بیشتر به دلیل افزایش طول پرسشنامه در هر مطالعه بزرگ‌مقیاسی در اختیار نیست. اغلب، پژوهشگران می‌توانند از متغیرهایی نظیر سن، تحصیلات و مانند آن برای پیش‌بینی در مدل لُجیت‌های تراکمی استفاده کنند. در مقابل، مدل پاسخ مدرج سامی‌جیما تنها بر اساس مقدار توانایی فرد دست به پیش‌بینی می‌زند؛ استفاده از توانایی به‌عنوان پیش‌بین دو مزیت عمده دارد: نخست اینکه به وجود متغیرهای پیش‌بین دیگر نیازی نیست و دیگر آنکه متغیری قابل‌اعتماد برای پیش‌بینی احتمال انتخاب مقوله‌های پاسخ به شمار می‌آید زیرا احتمال پاسخ به پرسش، تابعی از توانایی فرد است. از همین رو، این مدل دقیق‌تر از مدل لُجیت‌های تراکمی عمل کرد. با وجود این، نباید ظرفیت متغیرهای پیش‌بین دیگر را در کنار توانایی نادیده گرفت. مدل پرسش‌پاسخ تبیینی امکان استفاده از این ظرفیت را فراهم کرد و برتری خود را نسبت به مدل پاسخ مدرج نشان داد.

یافته‌ها حاکی از آن بودند که مدل‌های سه‌گانه به‌ویژه مدل پرسش‌پاسخ تبیینی و پس از آن، مدل پاسخ مدرج سامی‌جیما تحت بی‌پاسخی تصادفی از دقت قابل‌قبولی برای جایگزینی برخوردار هستند. با وجود این، تنها مدل پرسش‌پاسخ تبیینی تحت بی‌پاسخی غیرتصادفی، به دقتی بیشتر از ۸۰ درصد آن هم در نرخ بی‌پاسخی ۵ درصد دست یافت. به نظر می‌رسد. این امر می‌تواند ناشی از این واقعیت باشد که وجود سوگیری در نمونه باعث برازش ضعیف‌تر مدل‌ها در مقایسه با وضعیت تصادفی می‌شود. به‌عبارت‌دیگر، تحت بی‌پاسخی تصادفی، افرادی مشابه افراد بی‌پاسخ نیز در میان پاسخگویان وجود دارند که به برازش بهتر مدل بر داده‌ها کمک می‌کنند و پیش‌بینی موارد بی‌پاسخ را دقیق‌تر می‌سازند؛ بنابراین، اگر شواهدی دال بر غیرتصادفی بودن بی‌پاسخی در دست باشد، برازش مدل‌های جداگانه در زیرگروه‌های نمونه به‌جای برازش یک مدل بر کل داده‌ها، بتواند به افزایش

دقت پیش‌بینی آن‌ها کمک کند. به این ترتیب، ممکن است تا حدی از تأثیرات جدی بی‌پاسخی غیرتصادفی بر برازش مدل‌های جایگزینی کاسته شود.

این مقاله نشان داد که رویکرد خُرد بر رویکرد جایگزینی مستقیم نمره کل افراد برتری دارد؛ بنابراین، در مطالعات بزرگ‌مقیاسی که به دنبال شناسایی پارامترهای توزیع نمرات هستند، جایگزینی پرسش‌های دارای بی‌پاسخی از طریق یک مدل مناسب و سپس، محاسبه نمره کل از ترکیب پاسخ‌های جایگزین‌شده و پاسخ‌های واقعی فرد به نتایج کم‌خطاتری نسبت به‌جایگزینی مستقیم نمره کل او می‌انجامد. با وجود این، باید با احتیاط از روش‌های جایگزینی در نرخ‌های بی‌پاسخی غیرتصادفی بالای ۵ درصد استفاده کرد تا منجر به داده‌سازی غیرواقعی و نتایج گمراه‌کننده نشود. راهکار برازش مدل در زیرگروه‌های نمونه می‌تواند در مصون ماندن از این خطا نیز راهگشا باشد.

ORCID

Alireza Khoshgooyanfard http://orcid.org/0000-0002-6639-9339

Mohamadreza Falsafi Nejad http://orcid.org/0000-0002-7105-1644

Noorali Farrokhi http://orcid.org/0000-0003-0191-2717

منابع

Azar, B. (2002). Finding a solution for missing data. *Monitor on Psychology*, 33(7), 70.

Bodner, T. E. (2006). Missing data: Prevalence and reporting practices. *Psychological Reports*, 99, 675–680.

De Boeck, P., & Wilson, M. (2004). *Explanatory item response models*. Springer New York.

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

Fowler Jr, F. J. (2013). *Survey research methods*. Sage publications.

Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey methodology* (Vol. 561). John Wiley & Sons.

Hox, J., de Leeuw, E., Couper, M. P., Groves, R. M., De Heer, W., Kuusela, V., & Martin, J. (2002). The influence of interviewers' attitude and behavior on household survey.

Krishnamoorthy, K., Mallick, A., & Mathew, T. (2009). Model-based imputation approach for data analysis in the presence of non-detects. *Annals of Occupational Hygiene*, 53(3), 249-263.

Molenberghs, G., Fitzmaurice, G., Kenward, M. G., Tsiatis, A., & Verbeke, G. (Eds.). (2014). *Handbook of missing data methodology*. CRC Press.

O'Connell, A. A. (2006). *Logistic regression models for ordinal response variables* (No. 146). Sage.

Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of educational research*, 74(4), 525-556.

Raghunathan, T. (2016). *Missing Data Analysis in Practice*. CHAPMAN & HALL/CRC.

Roth, P. L. (1994). Missing data: A conceptual review for applied psychologists. *Personnel psychology*, 47(3), 537-560.

Samejima, F. (2008). Graded response model based on the logistic positive exponent family of models for dichotomous responses. *Psychometrika*, 73(4), 561-578.

Seaman, S., Galati, J., Jackson, D., & Carlin, J. (2013). What Is Meant by" Missing at Random"?. *Statistical Science*, 257-268.

Schlomer, G. L., Bauman, S., & Card, N. A. (2010). Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology*, 57(1), 1-10.

Tang, G., Little, R. J., & Raghunathan, T. E. (2003). Analysis of multivariate missing data with nonignorable nonresponse. *Biometrika*, 747-764.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.

van der Heijden, G. J., Donders, A. R. T., Stijnen, T., & Moons, K. G. (2006). Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: a clinical example. *Journal of clinical epidemiology*, 59(10), 1102-1109.

Wilkinson, L. (1999). Task Force on Statistical Inference, American Psychological Association, Science Directorate (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*(8), 594-604.

Zhang, Z. (2016). Missing data imputation: focusing on single imputation. *Annals of translational medicine*, *4*(1), 1-8.