Archive of SID

Journal of Educational Measurement & Evaluation Studies
Vol. 8, No. 23, Autumn 2018

فصلنامه مطالعات اندازه‌گیری و ارزشیابی آموزشی
سال هشتم، شماره ۲۳، پاییز ۱۳۹۷، صص ۳۰- ۷

اثرات بی‌پاسخی پرسش بر شاخص‌های روان‌سنجی: شواهدی بر نامناسب بودن راهکار حذف[1]

علیرضا خوشگویان‌فرد*

محمدرضا فلسفی‌نژاد**

نورعلی فرخی***

تاریخ دریافت مقاله: ۹۶/۰۳/۱۶

تاریخ پذیرش مقاله: ۹۶/۱۰/۰۳

چکیده

بی‌پاسخی پرسش چالشی جدی پیش روی اندازه‌گیری‌ها در مطالعات بزرگ‌مقیاس است، مطالعاتی که نقشی اساسی در تصمیم‌گیری‌ها و سیاست‌گذاری مربوط به آموزش و فرهنگ ایفا می‌کنند. به دلیل ایجاد سوگیری و خطا در برآورد پارامترهای توزیع نمره‌ها و شاخص‌های روان‌سنجی، چالش بی‌پاسخی هم بر فرایند ساخت وسیله اندازه‌گیری و هم بر استفاده از آن اثر می‌گذارد. از این رو، هدف این مقاله تأثیر نرخ‌های مختلف بی‌پاسخی پرسش را بر تعدادی از این شاخص‌ها و پارامترها تحت سازوکارهای کاملاً تصادفی و غیرتصادفی با استفاده از یک مجموعه داده واقعی شبیه‌سازی می‌کند. یافته‌ها نشان می‌دهند که بی‌پاسخی کاملاً تصادفی کمتر از ۵ درصد قابل چشم‌پوشی است مشروط بر آنکه افراد بی‌پاسخ متفاوت از افراد دارای پاسخ نباشند. در مقابل، اگر بی‌پاسخی با هر نرخی غیرتصادفی باشد، شاخص‌های روان‌سنجی (آلفای کرونباخ، ضریب پایایی دو نیمه کردن، ضریب روایی) و پارامترهای توزیع نمره‌ها (میانگین، انحراف‌معیار و چارک‌های نمره‌ها) می‌توانند دارای سوگیری و نادقیق باشند. همچنین، بی‌پاسخی غیرتصادفی می‌تواند به فاصله‌های اطمینان عریض و نتایج گمراه‌کننده‌ای منجر شود. از این رو، شواهد حاصل از این مطالعه شبیه‌سازی از حذف افراد بی‌پاسخ برای رفع مشکل بی‌پاسخی حمایت نمی‌کند.

واژه‌های کلیدی: پایایی، توزیع نمره‌ها، خطا، روایی، سوگیری، مطالعه بزرگ‌مقیاس، نرخ بی‌پاسخی.

---

[1]. این مقاله برگرفته از رساله دکتری با عنوان «شناسایی رویکرد بهینهٔ جایگزینی بی‌پاسخی در سنجش سازه‌های روان‌شناختی» است.

* دانشجوی دکتری سنجش و اندازه‌گیری، دانشگاه علامه طباطبایی

**دانشیار گروه سنجش و اندازه‌گیری، دانشگاه علامه طباطبایی (نویسنده مسئول: falsafinejad@yahoo.com)

*** دانشیار گروه سنجش و اندازه‌گیری، دانشگاه علامه طباطبایی

www.SID.ir

Archive of SID

مقدمه

بسیاری از دولت‌ها و نهادهای متولی آموزش و فرهنگ به اندازه‌گیری ارزش‌ها، باورها، نگرش‌ها و صفات روان‌شناختی در قالب مطالعات بزرگ‌مقیاس[1] دست می‌زنند تا از یافته‌های آن برای ارزیابی فعالیت‌های خود، سیاستگذاری‌ها و برنامه‌ریزی‌های کلان استفاده کنند (گراوز و همکاران[2]، ۲۰۱۱). این مطالعات بزرگ مقیاس در سطح شهرها، استان‌ها یا کشور به اجرا در می‌آیند و موفقیت آنها تنها در گرو ساخت وسیله‌های اندازه‌گیری کم‌نقص و نظارت جامع و دقیق بر اجرای آن است تا داده‌هایی معتبر و قابل تعمیم از جامعه تحت مطالعه گردآوری شود و مبنای تصمیم‌گیری‌های حساس قرار گیرد.

مطالعات بزرگ‌مقیاس از سوی عوامل مختلفی تهدید می‌شوند، از قبیل ۱ـ ضعف وسیله اندازه‌گیری به لحاظ موضوع، طراحی پرسش‌ها، حساسیت آنها یا طیف پاسخ؛ ۲ـ رفتار مصاحبه‌گر در برقراری ارتباط، جلب اعتماد و همکاری فرد نمونه و ۳ـ موقعیت فرهنگی، اجتماعی، شخصیتی و شناختی فرد نمونه (تورانو، ریپس و رازینسکی[3]، ۲۰۰۰؛ گراوز و کوپر[4]، ۲۰۱۲؛ هاکس و همکاران[5] ۲۰۰۲؛ فالر[6]، ۲۰۱۳). بنابراین، با وجود تلاش‌های فراوانی که صرف ساخت و اجرای مطلوب وسیله‌های اندازه‌گیری در مطالعات بزرگ‌مقیاس می‌شود، مشکلات اجتناب‌ناپذیری در اندازه‌گیری رخ می‌دهد که بی‌پاسخی پرسش از جدی‌ترین آنها به شمار می‌آید به نحوی که در بیشتر مطالعات بزرگ‌مقیاس، یک یا چند پرسش وسیله اندازه‌گیری از سوی تعدادی از افراد نمونه بی‌پاسخ باقی می‌ماند (بادنر[7]، ۲۰۰۶).

اساسی‌ترین چالشی که بی‌پاسخی پرسش به همراه دارد اندازه‌گیری ناقص صفت و ناتوانی در محاسبه نمره کل از پاسخ‌های پرسش‌های وسیله اندازه‌گیری است. برخلاف بی‌پاسخی پرسش‌های امتحانی که معادل تعلق گرفتن نمره صفر به آن

---

1. large scale
2. Groves et al
3. Tourangeau, Rips & Rasinski
4. Groves & Couper
5. Hox et al
6. Fowler
7. Bodner

www.SID.ir

Archive of SID

پرسش قلمداد می‌شود، بی‌پاسخی پرسش را نمی‌توان در اندازه‌گیری صفات روان‌شناختی، ارزش‌ها، باورها یا نگرش‌ها معادل نمره صفر به شمار آورد. بنابراین، نمره‌گذاری برای فردی که به برخی از پرسش‌ها پاسخ نداده است، امکان‌پذیر نیست. از این رو، برخی پژوهشگران به حذف افراد دارای پرسش‌های بی‌پاسخ اقدام می‌کنند تا با چالش محاسبه نمره کل رو به رو نشوند (دیلمن و همکاران[1] ، ۲۰۰۲).

حذف افراد دارای بی‌پاسخی پرسش، راهکاری ساده و عملی به نظر می‌رسد زیرا پس از حذف این افراد، تنها کسانی در نمونه باقی می‌مانند که به همه پرسش‌های وسیله اندازه‌گیری پاسخ داده‌اند و به راحتی می‌توان نمره کل این افراد را محاسبه کرد و انواع شاخص‌های روان‌سنجی و روش‌های آماری را برای مجموعه داده کامل به کار برد. با وجود این، راهکار حذف به کاهش اندازه نمونهٔ بهینه و مواجهه با چالش‌های جدی‌تری در برآورد شاخص‌های روان‌سنجی و پارامترهای هنجاریابی منجر می‌شود.

به طور کلی، بی‌پاسخی می‌تواند به سه شکل رخ دهد: کاملاً تصادفی، تصادفی و غیرتصادفی (بی‌هاسکاران و اسمیت[2]، ۲۰۱٤؛ سیمن و همکاران[3] ، ۲۰۱۳؛ تانگ، لیتل و راناتان[4]، ۲۰۰۳). بی‌پاسخی کاملاً تصادفی[5] (MCAR)، کمترین آسیب را برای برآورد شاخص‌های روان‌سنجی و پارامترهای هنجاریابی به همراه دارد. در این وضعیت، بی‌پاسخ ماندن یک پرسش از سوی فرد هیچ ارتباطی با مقدار آن پرسش، پرسش‌های دیگری وسیله اندازه‌گیری (و به طور کلی صفت مورد اندازه‌گیری) یا متغیرهای کمکی ندارد. در حالی که برای بی‌پاسخی تصادفی[6]، احتمال بی‌پاسخ ماندن پرسش به مقدار آن پرسش برای افراد بی‌پاسخ بستگی ندارد بلکه ناشی از مقدار متغیر برای پرسش‌های دیگر وسیله اندازه‌گیری یا متغیرهای کمکی است. مشکل‌زاترین بی‌پاسخی، نوع غیرتصادفی است که به مقدار پاسخ پرسش دارای بی‌پاسخی بستگی دارد به نحوی که تغییر مقدار این پاسخ بر بی‌پاسخ گذاشتن آن پرسش اثر می‌گذارد.

---

1. Dillman et al
2. Bhaskaran & Smeeth
3. Seaman et al
4. Tang, Little & Raghunathan
5. missing completely at random (MCAR)
6. missing at random (MAR)

www.SID.ir

Archive of SID

بی‌پاسخی‌های کاملاً تصادفی یا تصادفی تنها به افزایش خطای برآورد منجر می‌شوند در حالی که بی‌پاسخی غیرتصادفی دلالت بر تفاوت بین افراد پاسخگو و افراد دارای بی‌پاسخی پرسش دارد و نمونه را با سوگیری شدید رو به رو می‌سازد. به تبع آن، برآوردهای حاصل از نمونهٔ کاهش‌یافته هم با افزایش خطا (واریانس) و هم با بیش‌برآوردی[۱] یا کم‌برآوردی[۲] همراه می‌شوند. بنابراین، می‌توان دریافت که راهکار حذف افراد دارای بی‌پاسخی پرسش، معرف بودن نمونه را مخدوش می‌کند در حالی که انتخاب نمونه‌ای معرف از جامعه آماری، اصلی‌ترین شرط استنباط‌های آماری معتبر در مطالعات بزرگ‌مقیاس است، مانند برآورد ضرایب روایی و پایایی، هنجاریابی نمره، فاصله‌های اطمینان یا آزمون‌های آماری.

به‌رغم مشکلاتی که حداقل به صورت نظری قابل پذیرش است، بررسی‌هایی نظیر راث[۳] (۱۹۹٤) آزار[٤] (۲۰۰۲) یا پاف و اِندرز[٥] (۲۰۰٤) نشان می‌دهند چالش بی‌پاسخی در بسیاری از پژوهش‌ها نادیده گرفته می‌شود و برخی پژوهشگران ترجیح می‌دهند افراد دارای بی‌پاسخی را بدون توجه به تأثیرات آن از نمونه حذف کنند؛ حتی هیچ اشاره‌ای در تعدادی از پژوهش‌ها به نرخ بی‌پاسخی و حذف شدن افراد بی‌پاسخ نمی‌شود و تنها از طریق درجه آزادی آزمون‌ها می‌توان به این امر پی برد.

نادیده گرفتن مشکلات بی‌پاسخی و در پیش گرفتن راهکار حذف افراد دارای بی‌پاسخی پرسش باعث شده است تا برخی پژوهشگران به بررسی «تجربی» اثرات بی‌پاسخی بر برآورد شاخص‌ها و پارامترهای مختلف اقدام کنند. برای مثال، در مقاله‌های کائو، تسیاتیس و دیویدیان[٦] (۲۰۰۹) بی‌پاسخی در برآورد پارامتر میانگین توزیع، در مقاله‌های مک‌دونالد، ترستون و نلسون[۷] (۲۰۰۰) و اِندرز (۲۰۰۳) بی‌پاسخی در برآورد شاخص آلفای کرونباخ و در مقاله گراهام و کافمن[۸] (۲۰۱۲) بی‌پاسخی در مدل معادلات ساختاری بررسی شده است. اغلب این بررسی‌ها به برآورد پارامترهای

---

1. overestimate
2. underestimate
3. Roth
4. Azar
5. Peugh & Enders
6. Cao, Tsiatis and Davidian
7. Mcdonald, Thurston & Nelson
8. Graham & Coffman

www.SID.ir

Archive of SID

پیچیده مانند ضرایب مدل‌های آماری و استنباط‌هایی از این دست اختصاص دارند و کمتر به شاخص‌های روان‌سنجی (به استثنای آلفای کرونباخ) یا پارامترهای هنجاریابی نظیر چندک‌های توزیع نمره‌ها توجه داشته‌اند.

در این مقاله با نگاهی فراگیرتر، سه شاخص آلفای کرونباخ، ضریب دو نیمه کردن و ضریب روایی در کنار پنج پارامتر هنجاریابی میانگین، انحراف معیار و چارک‌های توزیع نمره‌ها بررسی می‌شوند. این بررسی، برخلاف بسیاری از مقاله‌ها، تحت نرخ‌های مختلف بی‌پاسخی و دو وضعیت بی‌پاسخی کاملاً تصادفی و بی‌پاسخی غیرتصادفی صورت می‌گیرد زیرا انتظار می‌رود تأثیرات بی‌پاسخی غیرتصادفی به مراتب جدی‌تر از بی‌پاسخی کاملاً تصادفی باشد. بر این اساس، دو پرسش زیر پاسخ داده می‌شوند:

۱ـ آیا همه شاخص‌های روان‌سنجی و پارامترهای هنجاریابی از هر نرخ بی‌پاسخی دچار سوگیری و کاهش دقت برآورد می‌شوند یا نرخی از بی‌پاسخی وجود دارد که بتوان برای برخی شاخص‌ها نادیده گرفت و افراد دارای بی‌پاسخی را از نمونه حذف و استنباط را بر اساس سایر افراد نمونه صورت داد؟

۲ـ میزان اثرگذاری دو نوع بی‌پاسخی تصادفی (بدون سوگیری) و بی‌پاسخی غیرتصادفی (دارای سوگیری) بر سوگیری و خطای برآورد شاخص‌های روان‌سنجی و پارامترهای هنجاریابی چگونه است؟

روش پژوهش

تأثیر بی‌پاسخی تصادفی و غیرتصادفی تحت نرخ‌های مختلف بی‌پاسخی از طریق شبیه‌سازی با انتخاب نمونه‌های بوت اِستریپی مطابق رویه مقاله کرومری و هینز[1] (۱۹۹٤) و در قالب الگوی یک طرح آزمایش دو عاملی[2] با استفاده از نرم‌افزار SAS/IML بررسی می‌شود. اساس این روش، تولید مجموعه داده‌های متعددی است که وضعیت‌های تحت مطالعه (نرخ بی‌پاسخی و نوع بی‌پاسخی) را دارا باشند. در این مقاله، تأثیر «نرخ‌های مختلف بی‌پاسخی» و «بی‌پاسخی تصادفی و دو نوع بی‌پاسخی

[1]. Kromrey & Hines
[2]. two-factor experimental design

www.SID.ir

Archive of SID

غیرتصادفی» بر شاخص‌های روان سنجی مورد ارزیابی قرار می‌گیرد. از این‌رو، نرخ‌های بی‌پاسخی ۲/۵، ۵، ۷/۵، ۱۰، ۱۵، ۲۰، ۲۵ و ۳۰ درصد به دفعات در وضعیت‌های تصادفی و غیرتصادفی در یک مجموعه داده «واقعی» ایجاد می‌شوند که به بررسی ۱۸ وضعیت مختلف بی‌پاسخی منجر خواهد شد. با استفاده از یک مجموعه داده پایه مربوط به یک نمونه واقعی، در این پژوهش می‌توان نتایج حاصل از یک نمونه کامل (بدون بی‌پاسخی) را با یک نمونه کاهش‌یافته در اثر بی‌پاسخی مقایسه کرد.

در حالت تصادفی، این نرخ‌ها با حذف «تصادفی» افراد از نمونه واقعی صورت می‌گیرد. برای مثال، ۲/۵ درصد از نمونه به طور تصادفی، حذف و شاخص‌های روان‌سنجی برای افراد باقی‌مانده محاسبه می‌شود. عمل حذف ۲/۵ درصد از نمونه، ۱۰هزار بار جداگانه روی این مجموعه داده اعمال می‌شود تا توزیع هر یک از شاخص‌های روان‌سنجی تحت بی‌پاسخی آشکار شود. در حالت غیرتصادفی، یک بار عمل حذف در بین افرادی که بالاترین نمره‌های کل را دارند و بار دیگر در میان افرادی که پایین‌ترین نمره‌های کل را دارند به‌طور تصادفی صورت می‌گیرد. بنابراین، بی‌پاسخی در میان افراد خاصی (دارندگان نمره‌های بالا یا پایین) ایجاد می‌شود تا با ایجاد تفاوت بین گروه پاسخگویان و گروه دارای بی‌پاسخی در نمونه سوگیری رخ دهد.

جامعه آماری شامل تمام موقعیت‌های بی‌پاسخی در سنجش سازه‌های روان‌شناختی است اعم از شرایط اجرا، آزمودنی‌ها، زمان و پرسش‌ها؛ بنابراین با یک ابرجامعه[1] سروکار داریم که در عمل، تنها یکی از این موقعیت‌ها یعنی زیرجامعه‌ای[2] از آن برای این پژوهش در دسترس خواهد بود. با وجود این، شبیه‌سازی به‌عنوان روش تولید داده‌ها کمک خواهد کرد تا از این زیرجامعه، وضعیت‌های‌های بالقوه متعددی ایجاد شود و بتوان به انبوهی از الگوهای متنوع بی‌پاسخی دست یافت.

داده‌هایی که عملیات شبیه‌سازی روی آنها صورت می‌گیرد از اجرای پرسشنامه‌ای شامل ۳۰ پرسش به دست آمده است که «امید به آینده» را از ۱۳۸۶۷ نفر ساکن در

---

1. super population
2. subpopulation

www.SID.ir

Archive of SID

مراکز استان‌های کشور با نمونه‌گیری خوشه‌ای سه مرحله‌ای[1] اندازه گرفته است (هر پرسش دارای طیف پاسخ ترتیبی «اصلاً»، «کم»، «متوسط» و «زیاد» است که از ۱ تا ٤ امتیازدهی می‌شوند). در کنار این ۳۰ پرسش، ۱۰ پرسش دیگر نیز وجود دارد که «رضایت از وضعیت کنونی زندگی» را اندازه می‌گیرند. انتظار می‌رود رابطه «امید به آینده» و «رضایت از وضعیت کنونی زندگی» قوی و مثبت باشد. این همبستگی به عنوان ضریب روایی وابسته به ملاک در نظر گرفته می‌شود.

شاخص‌های مورد بررسی در این مقاله عبارتند از ضریب آلفای کرونباخ، ضریب پایایی با روش دو نیمه کردن، ضریب روایی (همبستگی نمره «امید به آینده» و نمره «رضایت کنونی»)، میانگین توزیع نمره‌های کل افراد نمونه، چارک‌های توزیع نمره‌های کل و انحراف معیار توزیع نمره‌های کل. مقدار این شاخص‌های هشتگانه برای وسیله اندازه‌گیری «امید به آینده» بر اساس مجموعه داده کامل (۱۳۸٦۷ نفر) به شرح جدول (۱) است.

جدول (۱) شاخص‌های روان‌سنجی و توزیع نمره‌ها وسیله اندازه‌گیری «امید به آینده» بر اساس داده‌های کامل

آلفای کرونباخ	دو نیمه‌کردن	ضریب روایی	میانگین	چارک اول	چارک دوم	چارک سوم	انحراف معیار
۰/۷۸۸	۰/۸۱۱	۰/۵۹٦	۵۸/٤	۵۱	۵۷	٦٤	۹/٦

تأثیر بی‌پاسخی نیز بر شاخص‌های مورد بررسی در جدول (۱)، در قالب دو ویژگی سوگیری و خطا (انحراف‌معیار) بررسی می‌شود. برای این منظور، مراحل سه‌گانه زیر اجرا می‌شوند:

۱ـ ایجاد نرخ بی‌پاسخی مورد نظر در نمونه؛ برای مثال، برای نرخ بی‌پاسخی ۲/۵ درصد، باید ۳٤۷ نفر از نمونه ۱۳۸٤۷ نفری اصلی حذف شود.

۲ـ محاسبه شاخص مورد مطالعه بر اساس نمونه باقی‌مانده؛ در اینجا شاخص‌های هشتگانه جدول (۱) برای مجموعه داده کاهش‌یافته محاسبه می‌شوند.

[1]. three-stage cluster sampling

www.SID.ir

Archive of SID

۳ـ تکرار مراحل ۱ و ۲ به تعداد ۱۰هزار مرتبه برای دستیابی به توزیع تک‌تک شاخص‌های مورد مطالعه و به دست آوردن تقریبی از سوگیری و خطای برآورد شاخص‌ها.

یافته‌ها

یافته‌ها در چهار بخش ارائه می‌شوند. نخست، تأثیر بی‌پاسخی تصادفی و غیرتصادفی بر سوگیری و سپس، تأثیر آن بر افزایش خطای برآورد شاخص‌ها بررسی می‌شود. پس از آن، میانگین مربع خطای برآورد شاخص‌ها به‌عنوان معیاری ترکیبی از سوگیری و خطا مورد توجه قرار می‌گیرد. در پایان، تأثیر بی‌پاسخی بر عریض شدن فاصله اطمینان میانگین به‌عنوان یکی از پرکاربردترین روش‌های آماری بررسی خواهد شد.

تأثیر بی‌پاسخی بر سوگیری

شکل (۱)، سوگیری هشت شاخص را به ازای نرخ‌های بی‌پاسخی ۲/۵، ۵، ۷/۵، ۱۰، ۱۵، ۲۰، ۲۵ و ۳۰ درصد تحت بی‌پاسخی کاملاً تصادفی، بی‌پاسخی غیرتصادفی در میان افراد با نمره‌های بالا و بی‌پاسخی غیرتصادفی در میان افراد با نمره‌های پایین نشان می‌دهد. همان‌طور که انتظار داریم، بی‌پاسخی کاملاً تصادفی موجب سوگیری در هیچ‌یک از شاخص‌ها نمی‌شود (نوسان اندکی در آلفای کرونباخ، یا انحراف معیار نمره‌ها مشاهده می‌شود که قابل چشم‌پوشی است). در مقابل، شاخص‌های آلفای کرونباخ، ضریب پایایی دو نیمه کردن، ضریب روایی و انحراف معیار تحت بی‌پاسخی غیرتصادفی دچار کم‌برآوردی می‌شوند به‌نحوی که با افزایش نرخ بی‌پاسخی بر شدت سوگیری افزوده می‌شود.

شاخص‌های میانگین نمره‌ها و چارک‌ها تحت بی‌پاسخی غیرتصادفی نمره‌های بالایی دچار سوگیری منفی (کم‌برآوردی) و تحت بی‌پاسخی نمره‌های پایینی دچار سوگیری مثبت (بیش‌برآوردی) شده‌اند؛ زیرا بی‌پاسخی در افراد دارای نمره‌های بالایی باعث می‌شود تا داده‌های بزرگی که میانگین یا چارک‌ها را به سمت خود متمایل می‌سازند، از داده‌ها حذف شوند و میانگین و چارک‌ها کاهش یابند. عکس این موضوع با وقوع بی‌پاسخی در افراد دارای نمره‌های پایینی برای شاخص‌های مورد بررسی رخ داده است. انحراف معیار نمره‌ها از چنین قاعده‌ای تبعیت نمی‌کند زیرا با

www.SID.ir

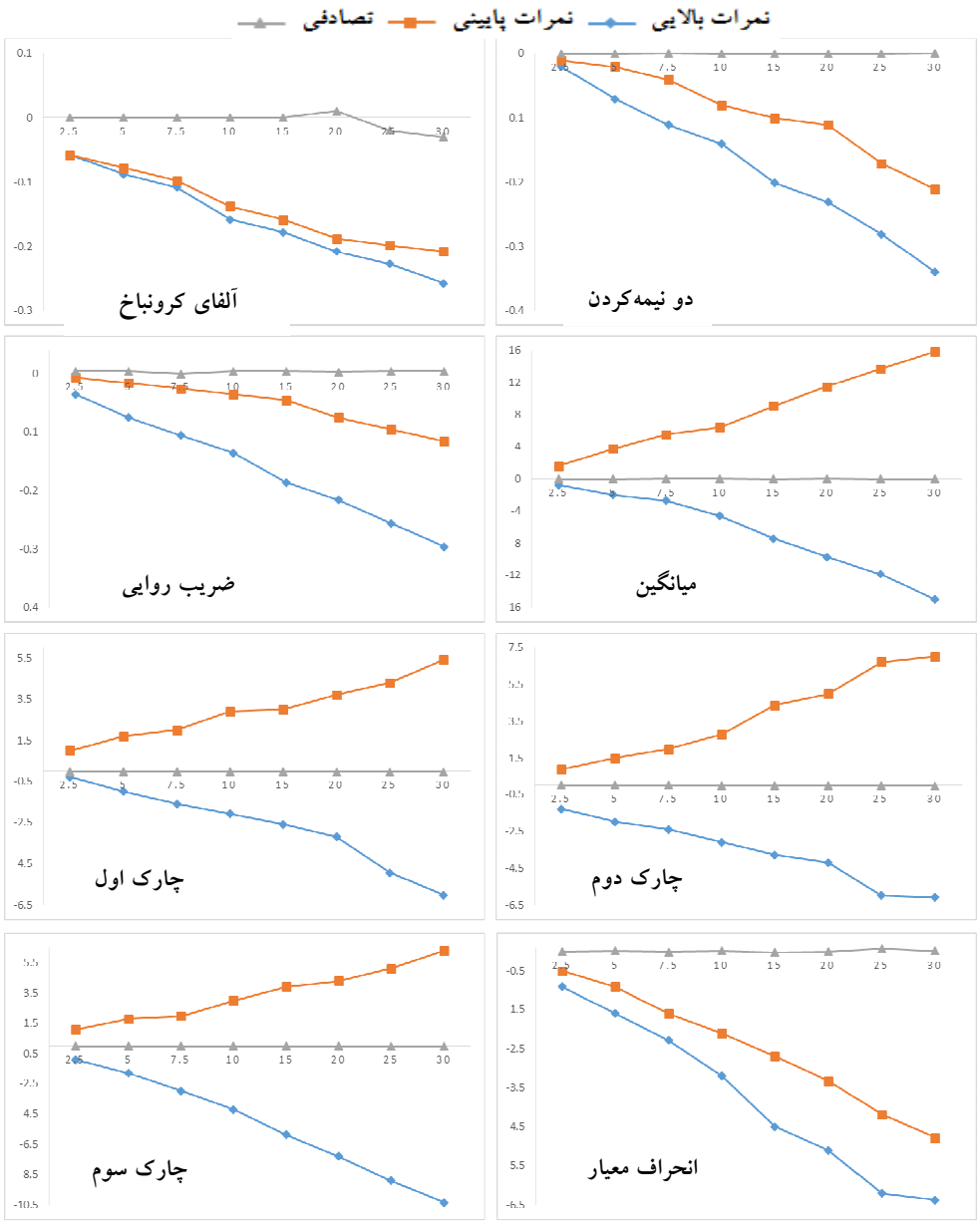
Archive of SID

حذف نمره‌های بالایی یا پایینی از تنوع نمره‌ها کاسته می‌شود و در هر حال، پراکندگی آنها کاهش می‌یابد.

به نظر می‌رسد سوگیری‌های جدی در شاخص‌ها پس از نرخ بی‌پاسخی ۵ درصد رخ می‌دهد به‌طوری که تا نرخ بی‌پاسخی ۵ درصد، حداکثر ۰/۱ کم‌برآوردی در آلفای کرونباخ، ضریب پایایی دو نیمه کردن و ضریب روایی دیده می‌شود ولی کم‌برآوردی برای نرخ‌های بی‌پاسخی بیشتر به ویژه بالای ۲۰ درصد به بیش از ۰/۲ نیز می‌رسد. بنابراین، آلفای ۰/۷۸۸ مربوط به داده‌های کامل، تحت بی‌پاسخی غیرتصادفی ۳۰ درصد در نمره‌های بالایی، ۰/۵۸۸ برآورد می‌شود یا ضریب پایایی دونیمه کردن ۰/۸۱۱، کمتر از ۰/۵ برآورد می‌شود. ضریب روایی نیز با مشکل کم‌برآوردی جدی روبه‌رو است به‌طوری که ضریب روایی ۰/۵۹۶ تحت همان نرخ بی‌پاسخی، ۰/۳ برآورد می‌شود.

www.SID.ir

Archive of SID



شکل (۱) سوگیری برآورد شاخص‌های روان‌سنجی به ازای نرخ‌های مختلف بی‌پاسخی تحت بی‌پاسخی کاملا تصادفی و غیرتصادفی

www.SID.ir

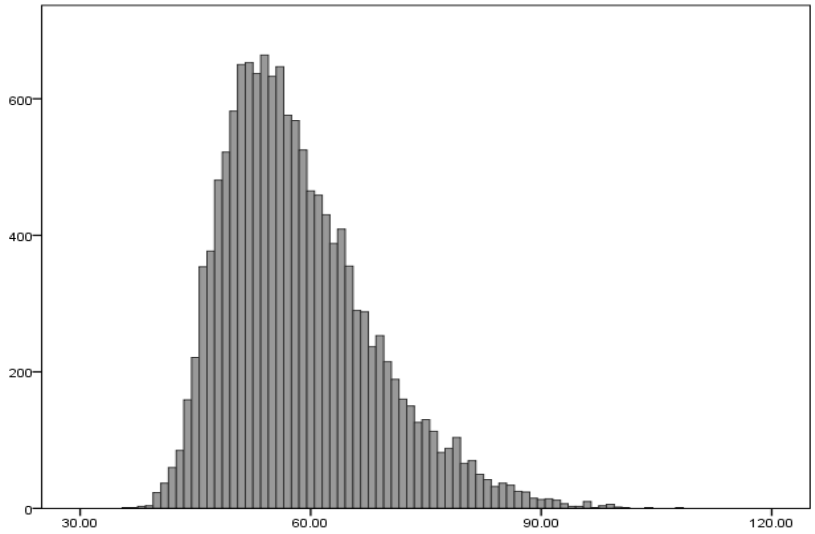
Archive of SID

چارک‌ها کمتر از میانگین تحت تأثیر بی‌پاسخی هستند به نحوی که میانگین تا ۱٦ واحد دچار سوگیری بیش‌برآرودی یا کم‌برآرودی شده است ولی چارک‌ها حداکثر تا ٦/٥ واحد (به جز چارک سوم که در بی‌پاسخی ۳۰ درصد نمره‌های بالایی، دچار ۱۰ واحد کم‌برآوردی شده است). برای مثال، ممکن است میانگین نمره‌ها در بی‌پاسخی ۲۰ درصد به جای ٥۸/٤، ۷۰/٤ یا ٤۸/٤ برآورد شود. دلیل تأثیر بیشتر بی‌پاسخی بر میانگین به وابستگی محاسبه میانگین به «مقدار» همه نمره‌ها و وابستگی محاسبه چارک‌ها به «تعداد» نمره‌ها بازمی‌گردد. بنابراین، حذف موارد بی‌پاسخ به طور مستقیم باعث تغییر مقدار میانگین نمره‌ها می‌شود در حالی که چنین تأثیری روی چارک‌ها ندارد.

بی‌پاسخی غیرتصادفی نمره‌های بالایی باعث سوگیری بیشتری نسبت به نمره‌های پایینی در همه شاخص‌های هشتگانه شده است. این امر می‌تواند با توزیع نمره‌های بالایی در داده‌های کامل مرتبط باشد؛ شکل (۲)، توزیع نمره‌ها را برای داده‌های کامل نشان می‌دهد که توزیعی چوله به راست است یعنی تعدادی از افراد با نمره‌های بسیار بالا وجود دارند که می‌توانند میانگین، واریانس یا چارک‌های توزیع نمره‌ها را به سمت خود متمایل کنند. به عبارت دیگر، نمره‌های بالایی، واریانس بیشتری نسبت به نمره‌های پایینی دارند و رخ دادن بی‌پاسخی در میان آنها نیز تأثیرات شدیدتری نسبت به بی‌پاسخی در نمره‌های پایینی به همراه خواهد داشت.

www.SID.ir

Archive of SID



شکل (۲) توزیع نمرات امید به آینده برای داده‌های کامل

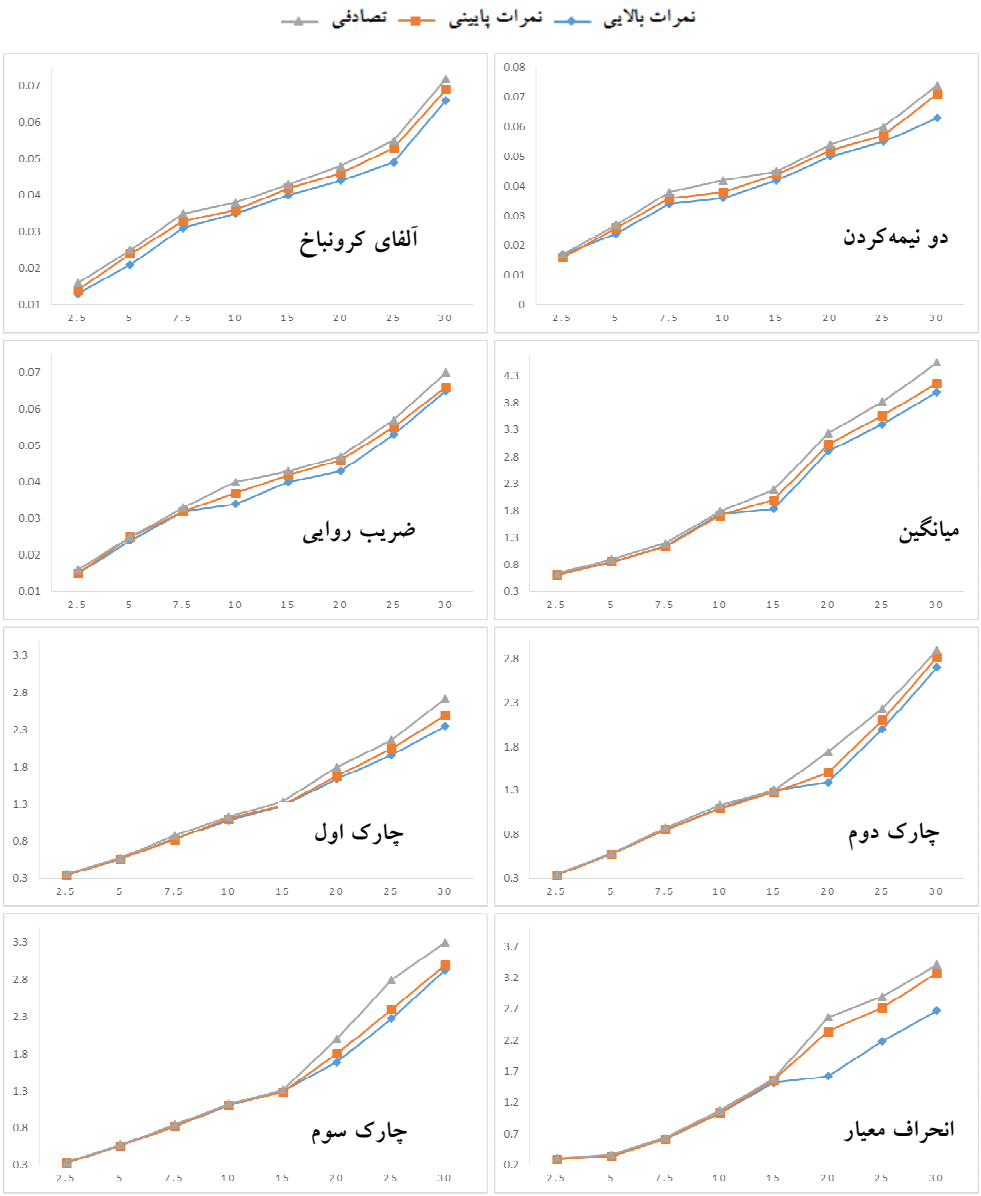
تأثیر سوگیری بر خطای برآورد شاخص‌ها

شکل (۳)، انحراف معیار برآورد شاخص‌ها و به تعبیری نوسان و خطای آنها را به ازای نرخ‌های بی‌پاسخی ۲/۵، ۵، ۷/۵، ۱۰، ۱۵، ۲۰، ۲۵ و ۳۰ درصد تحت بی‌پاسخی تصادفی و غیرتصادفی نشان می‌دهد. کاملاً آشکار است که با افزایش نرخ بی‌پاسخی (تصادفی یا غیرتصادفی)، خطای برآورد شاخص‌های هشتگانه نیز افزایش می‌یابد یعنی برآوردهای شاخص‌های روان‌سنجی در نرخ‌های بالای بی‌پاسخی بی‌ثبات‌تر و غیرقابل اطمینان‌تر می‌شوند.

در شاخص‌های مربوط به پایایی و روایی، همواره خطای برآوردها برای بی‌پاسخی تصادفی بیشتر از دو نوع بی‌پاسخی غیرتصادفی (در نمره‌های بالایی و نمره‌های پایینی) است در حالی که در شاخص‌های هنجاریابی (میانگین، انحراف‌معیار و چارک‌های توزیع نمره‌ها)، تفاوت بین خطای برآوردها برای بی‌پاسخی تصادفی و دو نوع غیرتصادفی آن، پس از نرخ بی‌پاسخی ۱۵ درصد آشکارتر می‌شود. در این حالت نیز خطای برآورد شاخص‌ها در بی‌پاسخی تصادفی بیشتر از دو نوع بی‌پاسخی غیرتصادفی است.

www.SID.ir

Archive of SID



شکل (۳) خطای برآوردهای شاخص‌ها روان‌سنجی به ازای نرخ‌های مختلف بی‌پاسخی تحت بی‌پاسخی کاملاً تصادفی و غیرتصادفی

www.SID.ir

Archive of SID

یافته‌های شکل (۳) نشان می‌دهد که خطای برآورد شاخص‌ها در بی‌پاسخی تصادفی بیشتر از خطای برآورد شاخص‌ها در بی‌پاسخی غیرتصادفی است. این تفاوت می‌تواند ناشی از این واقعیت باشد که در بی‌پاسخی کاملاً تصادفی، حذف «تصادفی» افراد باعث می‌شود تا تنوع داده‌ها بیشتر حفظ شود (به‌طور تصادفی هم از افرادی با نمره‌های بالا، هم پایین و هم متوسط تعدادی حذف می‌شوند)، ولی در بی‌پاسخی غیرتصادفی افراد، حذف تنها از میان افرادی با نمره‌های بالا یا پایین که سهم بیشتری در ایجاد پراکندگی در نمره‌ها دارند، صورت می‌گیرد، بنابراین با کاهش واریانس نسبت به بی‌پاسخی تصادفی رو به رو هستیم. به عبارت دیگر، واریانس نمره‌ها در نمونهٔ کاهش‌یافته از بی‌پاسخی تصادفی بیشتر از نمونهٔ کاهش‌یافته از بی‌پاسخی غیرتصادفی است. مانند تفاوتی که بین سوگیری برآورد شاخص‌ها در بی‌پاسخی غیرتصادفی نمره‌های بالایی با نمره‌های پایینی وجود داشت، خطای برآورد شاخص‌ها در بی‌پاسخی نمره‌های بالایی نیز کمتر از بی‌پاسخی نمره‌های پایینی است. این امر نیز به توزیع نمره‌های مجموعه داده کامل در شکل (۱) بازمی‌گردد؛ تنوع نمره‌های بالایی بیشتر از تنوع نمره‌های پایینی است، بنابراین حذف افرادی با نمره‌های بالا، بیشتر به همسانی داده‌ها (کاهش واریانس) منجر می‌شود تا حذف افرادی با نمره‌های پایین.

تأثیر بی‌پاسخی بر میانگین مربع خطا

تاکنون، تأثیر بی‌پاسخی جداگانه بر خطا (انحراف‌معیار) و سوگیری برآوردهای شاخص‌های روان‌سنجی بررسی شده است. ترکیب این دو کمّیت نیز به‌طور هم‌زمان در قالب میانگین مربع خطا[1] (مجموع مربع سوگیری و واریانس) قابل بررسی است. جدول (۲) شامل میانگین مربع خطای شاخص‌های هشتگانه به ازای نرخ‌های بی‌پاسخی ۲/۵، ۵، ۷/۵، ۱۰، ۱۵، ۲۰، ۲۵ و ۳۰ درصد تحت بی‌پاسخی تصادفی و دو نوع غیرتصادفی است.

نتایج جدول (۲)، کم و بیش با یافته‌های شکل‌های (۱) و (۳) هماهنگ است. همواره با افزایش نرخ بی‌پاسخی، میانگین مربع خطای برآوردهای همه شاخص‌ها افزایش می‌یابد به نحوی که این افزایش در بی‌پاسخی غیرتصادفی به شکل آشکاری

[1]. mean square error

www.SID.ir

Archive of SID

بیشتر از بی‌پاسخی تصادفی است. همچنین، در نرخ‌های بی‌پاسخی بالا، مقدار میانگین مربع خطا به شدت متورم می‌شود که نشان‌دهنده خطا و سوگیری (برای بی‌پاسخی غیرتصادفی) قابل توجهی در برآوردهای شاخص‌های هشتگانه است.

جز در نرخ‌های پایین بی‌پاسخی (۲/۵ و ۵ درصد)، میانگین مربع خطای برآورد ضریب پایایی دونیمه کردن هم برای بی‌پاسخی تصادفی و هم غیرتصادفی بیشتر از آلفای کرونباخ است. به علاوه، میانگین توزیع نمره‌ها نیز دارای میانگین مربع خطایی به مراتب بیشتر از چارک‌های توریع نمره‌هاست. با اینکه در بررسی شکل‌های (۱) و (۳)، نرخ بی‌پاسخی غیرتصادفی ۵ درصد یا کمتر به‌ظاهر قابل چشم‌پوشی بود، شاهد میانگین مربع خطای بزرگی برای میانگین نمره‌ها و در صورت محافظه‌کاری، میانگین مربع خطای برخی چارک‌ها هستیم. از این‌رو، تحت نرخ بی‌پاسخی غیرتصادفی ۵ درصد نیز احتیاط باید کرد.

www.SID.ir

Archive of SID

جدول (۲) مقدار میانگین مربع خطا (MSE) برای برآورد شاخص‌های هشتگانه

نوع و نرخ بی‌پاسخی		آلفای کرونباخ	پایایی دو نیمه‌کردن	ضریب روایی	میانگین نمرات	چارک اول نمرات	چارک دوم نمرات	چارک سوم نمرات	انحراف‌معیار نمرات
نمرات بالایی	۲/۵	۰/۰۰۳۵	۰/۰۰۰۷	۰/۰۰۱۵	۰/۹۸۷۹	۰/۲۰۷۶	۰/۸۰۱۶	۰/۹۱۶۳	۰/۸۹۳۵
	۵	۰/۰۰۸۲	۰/۰۰۵۶	۰/۰۰۶۴	۴/۷۱۰۶	۱/۳۱۷۰	۴/۲۹۶۵	۳/۵۹۸۷	۲/۶۸۰۴
	۷/۵	۰/۰۱۲۶	۰/۰۱۳۵	۰/۰۱۲۳	۹/۵۸۵۰	۳/۲۶۲۲	۶/۴۸۵۹	۵/۶۱۷۴	۵/۶۲۱۱
	۱۰	۰/۰۲۶۲	۰/۰۲۱۲	۰/۰۱۹۷	۲۴/۲۵۸۹	۵/۵۳۸۸	۱۰/۷۴۹۳	۱۹/۰۳۳۸	۱۱/۲۸۴۵
	۱۵	۰/۰۳۳۳	۰/۰۴۲۲	۰/۰۳۶۳	۶۰/۰۰۱۰	۸/۳۵۴۲	۱۵/۹۸۸۸	۳۶/۳۶۳۹	۲۲/۵۶۰۴
	۲۰	۰/۰۴۵۴	۰/۰۵۶۲	۰/۰۴۸۵	۱۰۳/۶۸۱۴	۱۲/۸۶۵۷	۱۹/۵۸۳۲	۵۶/۴۲۸۴	۲۸/۷۳۹۷
	۲۵	۰/۰۵۴۴	۰/۰۸۲۰	۰/۰۶۸۵	۱۵۴/۲۶۰۲	۲۸/۳۴۸۰	۴۰/۷۳۷۸	۸۴/۳۹۰۲	۴۳/۳۴۳۷
	۳۰	۰/۰۷۰۹	۰/۱۲۰۳	۰/۰۹۲۰	۲۴۴/۰۵۹۳	۴۱/۷۶۷۶	۴۵/۱۰۸۳	۱۱۶/۱۳۳۵	۴۷/۸۶۵۴
نمرات پایینی	۲/۵	۰/۰۰۳۶	۰/۰۰۰۴	۰/۰۰۰۳	۲/۹۳۷۰	۱/۱۳۸۴	۰/۹۲۱۶	۱/۳۱۸۲	۰/۳۳۵۸
	۵	۰/۰۰۶۷	۰/۰۰۱۱	۰/۰۰۰۹	۱۴/۴۱۷۶	۳/۱۹۹۱	۲/۶۱۴۲	۳/۵۵۰۲	۰/۹۲۷۰
	۷/۵	۰/۰۱۰۷	۰/۰۱۱۰	۰/۰۰۱۷	۳۱/۵۲۴۶	۴/۷۲۲۴	۴/۷۲۵۹	۴/۶۳۲۵	۲/۹۳۷۰
	۱۰	۰/۰۲۰۳	۰/۰۲۹۱	۰/۰۰۲۷	۴۴/۰۲۲۵	۹/۷۴۰۸	۹/۰۳۰۳	۱۰/۲۲۹۹	۵/۵۲۵۴
	۱۵	۰/۰۲۶۷	۰/۰۳۰۱	۰/۰۰۳۹	۸۴/۵۶۰۹	۱۰/۶۵۱۲	۲۰/۹۰۷۹	۱۷/۰۱۷۶	۷/۹۳۹۲
	۲۰	۰/۰۳۷۵	۰/۰۳۹۵	۰/۰۰۷۹	۱۴۱/۸۹۷۴	۱۶/۶۶۷۵	۲۷/۱۲۲۸	۲۱/۲۹۵۶	۱۶/۵۸۳۲
	۲۵	۰/۰۴۲۰	۰/۰۴۷۲	۰/۰۱۲۲	۲۰۰/۶۷۳۳	۲۲/۴۶۴۹	۵۰/۲۳۰۲	۳۱/۲۱۰۷	۲۴/۹۵۹۹
	۳۰	۰/۰۴۸۰	۰/۰۴۹۶	۰/۰۱۷۸	۲۶۹/۱۵۴۲	۳۵/۱۹۱۶	۵۷/۲۲۷۲	۴۸/۴۳۳۸	۳۳/۵۹۳۷
تصادفی	۲/۵	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۳	۰/۳۹۷۰	۰/۱۳۱۱	۰/۱۲۱۱	۰/۱۱۶۳	۰/۰۹۶۷
	۵	۰/۰۰۰۶	۰/۰۰۰۷	۰/۰۰۰۶	۰/۸۱۰۰	۰/۳۳۱۸	۰/۳۳۸۷	۰/۳۲۴۹	۰/۱۳۴۱
	۷/۵	۰/۰۰۱۲	۰/۰۰۱۴	۰/۰۰۱۱	۱/۴۳۱۳	۰/۷۶۳۹	۰/۷۶۵۶	۰/۷۱۷۴	۰/۴۰۳۳
	۱۰	۰/۰۰۱۴	۰/۰۰۱۸	۰/۰۰۱۶	۳/۱۹۴۳	۱/۲۸۱۴	۱/۲۸۳۸	۱/۲۵۶۶	۱/۱۵۵۷
	۱۵	۰/۰۰۱۸	۰/۰۰۲۰	۰/۰۰۱۹	۳/۳۴۹۰	۱/۷۸۴۹	۱/۶۸۰۰	۱/۷۲۶۶	۲/۴۹۰۵
	۲۰	۰/۰۰۲۴	۰/۰۰۲۶	۰/۰۰۲۲	۱۰/۴۶۵۳	۲/۸۵۶۵	۳/۰۲۰۷	۲/۹۰۰۳	۶/۵۹۹۸
	۲۵	۰/۰۰۳۴	۰/۰۰۳۶	۰/۰۰۳۳	۱۴/۶۳۱۰	۵/۷۰۰۶	۴/۹۸۷۲	۵/۱۸۹۷	۸/۴۱۴۹
	۳۰	۰/۰۰۶۱	۰/۰۰۶۵	۰/۰۰۴۹	۲۰/۶۵۷۱	۷/۴۲۶۵	۸/۳۷۶۸	۹/۷۴۱۵	۱۱/۶۲۱۴

www.SID.ir

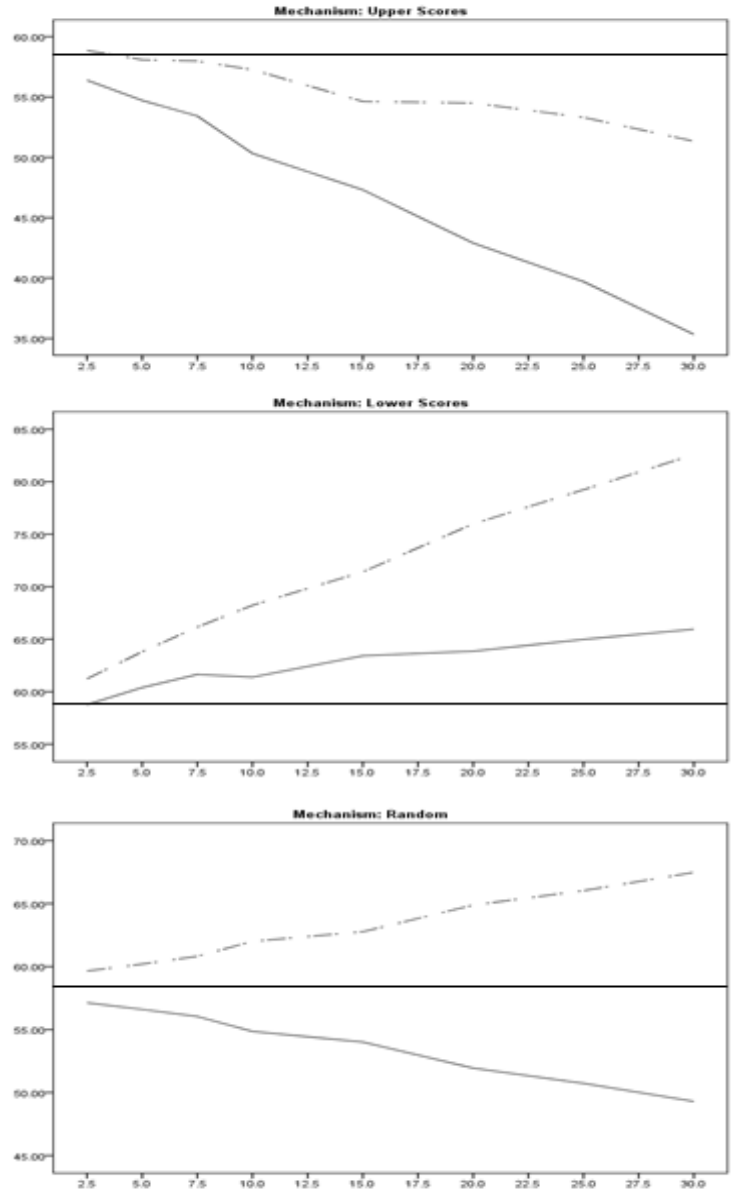
Archive of SID

تأثیر بی‌پاسخی بر پهنای فاصله اطمینان

علاوه بر تأثیراتی که بی‌پاسخی بر سوگیری و خطای شاخص‌های روان‌سنجی دارد، نرخ‌های بالای بی‌پاسخی و اُفت نمونه ناشی از حذف افراد بی‌پاسخ، کاهش توان‌های آزمون‌های آماری و عریض‌شدن طول فاصله‌های اطمینان را به همراه دارد. در شکل (٤) حد بالا و حد پایین فاصله‌های اطمینان برای میانگین به ازای نرخ‌های بی‌پاسخی ۲/۵، ۵، ۷/۵، ۱۰، ۱۵، ۲۰، ۲۵ و ۳۰ درصد ترسیم است (شکل بالا برای بی‌پاسخی در نمره‌های بالایی، شکل وسط برای بی‌پاسخی در نمره‌های میانی و شکل پایین برای بی‌پاسخی تصادفی است). خط افقی در این نمودارها، مقدار میانگین یعنی ۵۸/٤ را برای داده‌های کامل از جدول (۱) نشان می‌دهد.

www.SID.ir

Archive of SID



شکل (٤) فاصله اطمینان۹۵ درصدی برای میانگین توزیع نمره‌ها به ازای نرخ‌های مختلف بی‌پاسخی

www.SID.ir

Archive of SID

دو نکته در شکل (٤) به‌طور آشکار قابل توجه است؛ نخست اینکه با افزایش نرخ بی‌پاسخی، حد بالا و پایین فاصله اطمینان‌ها از هم دور می‌شوند و فاصله‌های عریض‌تری ایجاد می‌شود. نکته دیگر اینکه به استثنای بی‌پاسخی تصادفی که به ازای همه نرخ‌های بی‌پاسخی، میانگین داده‌های کامل در میان حد بالا و پایین فاصله قرار می‌گیرد (خط افقی در میان حد بالا و حد پایین قرار دارد)، فاصله‌های اطمینان مربوط به بی‌پاسخی نمره‌های بالایی و پایینی، به‌سختی حتی برای بی‌پاسخی ۲/۵ درصد هم شامل میانگین داده‌های کامل می‌شود.

برای مثال، در نرخ بی‌پاسخی تصادفی ۲۰ درصد، فاصله به صورت ٥١/٩٦ تا ٦٤/٨٧ به دست می‌آید که شامل مقدار ٥٨/٤ یعنی میانگین نمره‌های کامل است؛ در حالی که برای همین نرخ بی‌پاسخی ولی در نمره‌های بالایی به فاصله ٥٣/٦ تا ٤٣/٨ خواهیم رسید که شامل مقدار ٥٨/٤ نیست. در واقع، هم سوگیری و هم خطا هم‌زمان در این فاصله اطمینان‌ها منعکس شده‌اند به‌نحوی که در بی‌پاسخی غیرتصادفی، نتایج گمراه‌کننده‌ای را درباره میانگین توزیع نمره‌ها به دست می‌دهند.

نتیجه‌گیری

در کنار روش‌هایی مانند جایگزینی[1] پرسش‌های بی‌پاسخ یا وزن‌دهی برای اصلاح بی‌پاسخی (لیتل و روبین[2]، ۲۰۱٤)، ساده‌ترین و اولین شیوه برخورد با افراد دارای بی‌پاسخی که پژوهشگران در پیش می‌گیرند و اغلب نرم‌افزارها نیز به طور پیش‌فرض اعمال می‌کنند، حذف آنها از مجموعه داده‌هاست. این حذف به دو صورت کامل و دوبه‌دو صورت می‌گیرد. در تحلیل موارد کامل[3]، فردی که حداقل دارای یک پرسش بی‌پاسخ است به طور کامل از مجموعه داده‌ها حذف می‌شود و مجموعه داده‌ای شامل آن دسته از افراد نمونه باقی می‌ماند که به همه پرسش‌های وسیله اندازه‌گیری پاسخ داده‌اند[4] (کی‌نال و همکاران[5]، ۲۰۱۰). در کنار سادگی این روش و به‌رغم مشکلات

---

1. imputation
2. Little & Rubin
3. complete case analysis
٤. این معادل انتخاب گزینه Exclude cases listwise در نرم‌افزار SPSS است.
5. Knol et al

www.SID.ir

Archive of SID

جدی آن که در این مقاله بررسی شد، دو مزیت عمده باعث در پیش گرفتن تحلیل موارد کامل شده است. نخست اینکه، روش‌های آماری معمول بدون هیچ نگرانی درباره موارد بی‌پاسخ برای این مجموعه داده قابل استفاده هستند. مزیت دیگر اینکه تحلیل انفرادی پرسش‌ها و مقایسه نتایج آنها با یکدیگر به دلیل اجرای محاسبات تک‌تک پرسش‌ها روی مجموعه داده واحد امکان‌پذیر است.

مزیت دوم، هم در نمره‌گذاری و هم ساخت وسیله اندازه‌گیری جدید از اهمیت ویژه‌ای برخوردار است. برای مثال، در مرحله تحلیل پرسش‌های وسیله اندازه‌گیری، پارامتر دشواری تک‌تک پرسش‌ها محاسبه و با یکدیگر مقایسه می‌شوند. اگر ۳۰۰ نفر از یک نمونه ۱۰۰۰ تایی به یک یا چند پرسش پاسخ نداده باشند، با حذف کامل این افراد از نمونه، ۷۰۰ نفر باقی می‌مانند که به همه پرسش‌ها پاسخ داده‌اند. به این ترتیب، پارامتر دشواری تک‌تک پرسش‌ها بر اساس پاسخ‌های این ۷۰۰ نفر محاسبه می‌شود. اکنون اگر افراد بی‌پاسخ حذف نشوند، این امکان وجود دارد که پرسشی را ۹۰۰ نفر، پرسش دیگری را ۸۵۰ نفر و پرسش دیگری را ۹۵۰ نفر پاسخ داده باشند، پس دشواری پرسش اول بر اساس ۹۰۰ نفری محاسبه می‌شود که به آن پاسخ داده‌اند و به همین ترتیب، دشواری پرسش‌های دوم و سوم بر اساس ۸۵۰ و ۹۵۰ نفر پاسخگوی متناظرشان. بدین معنا که مقایسه دشواری این سه پرسش بر مبنای نمونه‌های مختلفی صورت می‌گیرد که در مرحله مقایسه دشواری پرسش‌ها به منظور انتخاب پرسش‌های مناسب برای ساخت وسیله اندازه‌گیری منطقی نیست.

به‌رغم دو مزیت تحلیل موارد کامل، این روش به حذف همه اطلاعات افراد دارای بی‌پاسخی پرسش حتی پاسخ‌های داده شده آنان منجر می‌شود تا نمونه‌ای با پاسخ‌های کامل برای همه به دست آید؛ کافی است فردی تنها به یک پرسش پاسخ نداده باشد، تا به‌طور کامل از مجموعه داده‌ها کنار گذاشته شود در حالی که اگر تحلیل انفرادی پرسش‌ها مورد توجه نباشد، مثلاً به دنبال همبستگی دوبه‌دوی پرسش‌ها باشیم، می‌توان حذف را تنها محدود به افرادی کرد که به یکی از این دو پرسش پاسخ نداده‌اند که به آن، تحلیل موارد موجود[1] گفته می‌شود (راناتان[2]، ۲۰۰٤). در این شیوه

---

1. available case analysis
2. Raghunathan

www.SID.ir

Archive of SID

اُفت اندازه نمونه کمتر است زیرا حذف افراد نمونه تنها ناشی از دو پرسش و نه همه پرسش‌ها است[۱].

برای مثال، اگر محاسبه همبستگی دو پرسش مورد توجه باشد، لازم است واریانس هر یک از دو پرسش و کوواریانس آنها محاسبه شود که برای این منظور، هم واریانس‌ها و هم کوواریانس بر اساس همه افرادی که به هر دو پرسش پاسخ داده‌اند، محاسبه می‌شود تا در نهایت، همبستگی دو پرسش به دست آید. این امکان وجود دارد که واریانس هر پرسش در میان همه افرادی محاسبه شود که به آن پرسش پاسخ داده‌اند و نه افرادی که به هر دو پرسش پاسخ داده‌اند، ولی همبستگی حاصل از به‌کارگیری این واریانس‌ها می‌تواند کمتر از ۱- یا بیشتر از ۱ شود. از این‌رو، ضروری است محاسبات تنها در میان افرادی صورت پذیرد که به هر دو پرسش پاسخ داده‌اند.

این مقاله نشان داد که روش حذف برای رویارویی با مسئله بی‌پاسخی پرسش، روش قابل دفاعی نیست. با وجود این، در شرایط بسیار محدود و نادر، این روش می‌تواند تا اندازه‌ای راهگشا باشد. با توجه به یافته‌ها، نسبت افراد دارای بی‌پاسخی در «کل» نمونه نباید بیش از ۵ درصد باشد به نحوی که با اُفت چشمگیر اندازه نمونه و به دنبال آن افزایش خطای نمونه‌گیری روبه‌رو نباشیم. خطای نمونه‌گیری باعث افزایش خطای اندازه‌گیری، خطای برآورد پارامترهای مدل‌های IRT و شاخص‌های روان‌سنجی می‌شود. نباید از نظر دور داشت که رقم ۵ درصد مربوط به کل نمونه است نه صرفاً تک‌تک پرسش‌ها. به این ترتیب، در یک وسیله ۳۰ پرسشی که برای هر پرسش تنها ۱ درصد بی‌پاسخی رخ داده است، متوسط درصد افرادی که به همه پرسش‌ها پاسخ داده‌اند $۰/۹۹^{۳۰} = ۷٤$ درصد خواهد بود یعنی ۲٦ درصد از افراد نمونه حذف می‌شوند. بنابراین هر چه تعداد پرسش‌های بیشتری دارای نسبت پاسخگویی کم باشند، اُفت نمونه با حذف موارد بی‌پاسخ به‌شدت افزایش می‌یابد.

یافته‌ها به خوبی سوگیری تحت بی‌پاسخی غیرتصادفی در نمره‌های بالایی یا پایینی را حتی تحت نرخ بی‌پاسخی ۵ درصد نشان دادند. بنابراین، افراد حذف‌شده نباید تفاوت جدی با افراد دارای پاسخ داشته باشند تا حذف آنان موجب سوگیری نتایج نشود. تشخیص بی‌پاسخی تصادفی از غیرتصادفی دارای راهکار دقیقی نیست

۱. این معادل انتخاب گزینه Exclude cases pairwise در نرم‌افزار SPSS است.

www.SID.ir

Archive of SID

ولی از طریق مقایسه ویژگی‌های جمعیت‌شناختی افراد بی‌پاسخ با افراد پاسخگو یا مقایسه نحوه پاسخگویی به پرسش‌هایی که هر دو گروه به آنها پاسخ داده‌اند، تا حدی می‌توان به این تشخیص دست یافت.

در شرایطی که علاوه بر مطالعه تمام نمونه، بررسی زیرگروه‌ها نیز مورد توجه باشد، دور از انتظار است که روش حذف کارامد باشد زیرا با حذف افراد بی‌پاسخ از نمونه، تعداد نمونه مربوط به زیرگروه‌ها می‌تواند با کاهش بیشتری در مقایسه با کل نمونه روبه‌رو شود. برای مثال، اگر مطالعه بزرگ‌مقیاس در پی اندازه‌گیری صفتی در یک شهر باشد و پس از آن، بررسی وضعیت این صفت در گروه‌های سنی مختلف مورد توجه قرار گیرد، حذف افراد دارای بی‌پاسخی، هم کاهش کل اندازه نمونه را به همراه خواهد داشت و هم موجب کاهش اندازه نمونه همه یا تعدادی از زیرگروه‌های سنی خواهد شد که هر یک به تنهایی دارای اندازه نمونه کوچکی نسبت به نمونه کلی هستند.

در مجموع، حذف افراد دارای بی‌پاسخی از نمونه چه با رویکرد تحلیل موارد کامل باشد چه با رویکرد تحلیل موارد موجود، راهکاری نیست که به طور کامل به حل مشکل بی‌پاسخی بیانجامد (وندر هایدن[1]، ۲۰۰۶). در ساخت وسیله اندازه‌گیری، هم تحلیل انفرادی پرسش‌ها و هم تحلیل دوبه‌دوی آنها مورد نیاز است و در به کارگیری یک وسیله اندازه‌گیری معتبر، نمره‌گذاری و هنجاریابی از پاسخ‌های کل پرسش‌های وسیله اندازه‌گیری استفاده می‌کند. بنابراین، بعید به نظر می‌رسد با راهکار حذف بتوان حتی تحت بی‌پاسخی کاملاً تصادفی نیز از مشکلات اُفت نمونه دور ماند مگر اینکه نرخ بی‌پاسخی بسیار ناچیز باشد. از این‌رو، ضروری است نرخ بی‌پاسخی پرسش‌ها و شیوه برخورد با آنها در مطالعات بزرگ‌مقیاس به روشنی، بیان و از شیوه‌های کارامد و دقیق‌تر برای مقابله با این چالش کمک گرفته شود.

[1]. van der Heijden

www.SID.ir

Archive of SID

منابع

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

Bhaskaran, K.; & Smeeth, L. (2014). What is the difference between missing completely at random and missing at random? *International Journal of Epidemiology*, *43* (4), 1336-1339.

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

Cao, W.; Tsiatis, A. A.; & Davidian, M. (2009). Improving efficiency and robustness of the doubly robust estimator for a population mean with incomplete data. *Biometrika*, 96, 723-734.

Dillman, D. A.; Eltinge, J. L.; Groves, R. M.; & Little, R. J. *(2002). Survey nonresponse in design, data collection, and analysis. Survey nonresponse*, *3*-26.

Enders, C. K. (2003). Using the expectation maximization algorithm to estimate coefficient alpha for scales with item-level missing data. *Psychological methods*, *8* (3), 322.

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

Graham, J. W.; & Coffman, D. L. (2012). Structural equation modeling with missing data. *Handbook of structural equation modeling*, 277-295.

Groves, R. M.; & Couper, M. P. (2012). *Nonresponse in household interview surveys*. John Wiley & Sons.

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. D.; Couper, M. P.; Groves, R. M.; De Heer, W.; Kuusela, V.; …& Belak, E. (2002). The influence of interviewers' attitude and behaviour on household survey nonresponse: An international comparison. In: R. M. Groves, D. A. Dillman, J. L. Eltinge, & R.J.A. Little (Eds). *Survey nonresponse*. New York: Wiley, pp. 103-120

Knol, M. J.; Janssen, K. J.; Donders, A. R. T.; Egberts, A. C.; Heerdink, E. R.; Grobbee, D. E.; ... & Geerlings, M. I. (2010). Unpredictable bias when using the missing indicator method or complete case analysis for missing confounder values: an empirical example. *Journal of Clinical Epidemiology*, *63* (7), 728-736.

www.SID.ir

Archive of SID

Kromrey, J. D.; & Hines, C. V. (1994). Nonrandomly missing data in multiple regression: An empirical comparison of common missing-data treatments. *Educational & Psychological Measurement*, *54* (3), 573-593.

Little, R. J.; & Rubin, D. B. (2014). *Statistical analysis with missing data*. John Wiley & Sons.

Mcdonald, R. A.; Thurston, P. W.; & Nelson, M. R. (2000). A Monte Carlo study of missing item methods. *Organizational Research Methods*, *3* (1), 71-92.

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. E. (2004). What do we do with missing data? Some options for analysis of incomplete data. *Annu. Rev. Public Health*, 25, 99-117.

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

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

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.

www.SID.ir