



مقدمه‌ای بر ارزش‌یابی پرسش‌های آمارگیری

محیا قاسمی و مهدی جباری نوقابی

دانشگاه فردوسی مشهد، گروه آمار، mahya_ghasemi@mail.um.ac.ir

دانشگاه فردوسی مشهد، گروه آمار، jabbarinm@um.ac.ir

چکیده

امروزه با گسترش تحقیقات در علوم مختلف، بسیاری از پدیده‌ها از طریق ابزارهایی مانند پرسش‌نامه‌ها مورد مطالعه قرار می‌گیرند. به منظور استفاده‌ی کارآمد از داده‌های آمارگیری، در مطالعه‌ی حاضر به بررسی روش‌های ارزیابی پرسشنامه‌های آمارگیری پرداخته شده است. بدین منظور، ارزش‌یابی سوال را مؤلفه‌بندی کرده و انواع مختلف روش‌ها برای این مؤلفه‌ها ارائه داده و مورد بررسی قرار می‌گیرد.

کلمات کلیدی: ارزش‌یابی، اعتبار، پایایی، طرح‌های آمارگیر.

مقدمه

روایی معمولاً به صورت یک تابع همبستگی بین (Y_i, μ_i) تعریف می‌شود. بدین معنا که مقادیر Y_i به مقادیر μ_i وابسته هستند که مقدار روایی به صورت زیر محاسبه می‌شود:

$$validity(Y) = \frac{\sum_{it} (Y_{it} - \bar{Y})(\mu_i - \bar{\mu})}{\sqrt{\sum_{it} (Y_{it} - \bar{Y})^2 \sum_{it} (\mu_i - \bar{\mu})^2}} = corr(Y_i, \mu_i).$$

با توجه به این تعریف روایی به صورت عددی بین صفر و یک بیان می‌شود. هر چه عدد بزرگتر باشد، اعتبار بیشتر است. روایی را می‌توان با استفاده از داده‌های خارج از آمارگیری برآورد کرد. معمول‌ترین اشتباهی که در اصطلاح شناسی خطاهایی که به سوال‌ها مربوط می‌شود، نحوه‌ی ارتباط "روایی" با "اربابی" است. روایی تابعی از همبستگی بین پاسخ و مقدار واقعی است. بنابراین روایی یکی از خواص پاسخ‌دهی تکی به سوال‌ها می‌باشد. برخی از انحرافات مانند کم گزارش‌دهی موجب کاهش همبستگی بین پاسخ‌ها و مقادیر واقعی نمی‌شوند ولی باعث اربابی خواهند شد. اربابی زمانی مطرح می‌شود که بین پاسخ افراد و مقدار واقعی آنها برای همه‌ی تکرارها و افراد، انحرافی سیستماتیک وجود داشته باشد:

$$bias(Y_{it}) = E_i [E_t(Y_{it}) - \mu_i]$$

این یعنی میانگین پاسخ‌ها یک برآورد ارباب از میانگین مقادیر واقعی است:

$$bias(\bar{Y}) = E \left[\frac{\sum_i Y_{it}}{N} \right] - \frac{\sum_i \mu_i}{N}$$

جمله‌ی اول مقدار مورد انتظار پاسخ برای همه‌ی تکرارها است. اربابی به مقادیر μ_i بستگی دارد، اما روایی فقط به تغییر پذیری مقادیر μ_i بستگی دارد. به همین علت، مفهوم اربابی تنها برای اندازه‌گیری متغیرهای عینی حقایق یا حوادث به کار می‌رود. با استفاده از داده‌های هر یک از عناصر جامعه هدف و استفاده از آماره‌های جامعه‌ی که تحت تأثیر خطای پاسخ آمارگیری نبوده‌اند می‌توان اربابی را برآورد کرد.

پایایی و واریانس پاسخ: پایایی به پایداری نتایج آزمون بستگی دارد، یعنی اگر آزمون تکرار شود و نتیجه‌ی آن با نتایج آزمون‌های پیشین نزدیک و یکسان باشد، در این صورت آزمون پایا است. بنابراین پایایی بر حسب مؤلفه‌های یک واریانس، تغییرپذیری ε_{it} برای همه پاسخ‌گویان و تکرارها تعریف می‌شود. که به شکل زیر داریم:

$$per(Y_{it}) = \frac{E_i(\mu_i - \bar{\mu})^2}{E_i(\mu_i - \bar{\mu})^2 + E_{it}(\varepsilon_{it} - \bar{\varepsilon})^2}$$

اگر واریانس انحراف پاسخ $E_{it}(\varepsilon_{it} - \bar{\varepsilon})^2$ کم باشد، این نسبت پایایی به یک نزدیک می‌شود. یعنی این که اندازه مربوط دارای پایایی زیاد است و اگر تغییرپذیری پاسخ‌ها در تکرارها زیاد شود آنگاه پایایی به مقدار صفر نزدیک می‌شود. معمولاً در آمارگیری به جای پایایی از عبارت "واریانس پاسخ" که در نقطه مقابل پایایی است استفاده می‌کنند. اگر یک سوال برای جامعه دارای پایایی زیاد باشد داری واریانس پاسخ کم است. دو روش برای ارزیابی ثبات پایایی مورد استفاده قرار می‌گیرد:

الف) مصاحبه‌های مکرر با پاسخ‌گویان یکسان: این روش به نام مطالعه مصاحبه مجدد نیز شناخته شده است. ب) استفاده از شکل‌های چندگانه یک ساختار: در این روش سوال‌های مربوط به ساختار مورد نظر به چند حالت دیگر طراحی شده و از پاسخ‌گویان پرسیده می‌شود. آلفای کرونباخ معیاری برای سنجش پایایی چنین شاخص‌های چند سوالی است که به طور گسترده مورد استفاده قرار می‌گیرد. پایایی شاخص تولید شده به تعداد سوال‌ها، k و متوسط همبستگی درونی (\bar{r}) آن‌ها بستگی دارد:

$$\alpha = \frac{k\bar{r}}{1 + (k-1)\bar{r}}$$

مقدار زیاد آن نمایانگر این است که پاسخ‌های ارائه شده به یک سوال، پاسخ دیگر سوال‌ها را تحت تأثیر قرار داده است تا همبستگی مثبت زیادی را القا کند. مقدار کم نشان دهنده‌ی این است که سوال‌ها ساختار یکسانی را اندازه‌گیری نمی‌کنند.

بحث و نتیجه‌گیری

با توجه به آنچه که تا به حال گفته شده است با به کارگیری روش‌های ارزش‌یابی می‌توان قدرت مطالعات را بالا برد. توجه محققان و پژوهشگران به ارائه‌ی نتایج آماری می‌تواند نتیجه ارزش‌یابی پرسش‌ها باشد، چرا که با کمک این روش‌ها می‌توان به پرسشنامه‌ای استاندارد دست یافت که باعث ایجاد نتایج آماری دقیق و صحیحی از پژوهش می‌شود، که به کمک این نتایج می‌توان تصمیم‌گیری درستی انجام داد.

مراجع

- [۱] گراوزم، فولرج، (۲۰۰۴) ترجمه‌ی صالحی، م. روش‌شناسی آمارگیری، پژوهشکده‌ی آمار، تهران
- [۲] حیدری چروده، م. راهنمای سنجش روایی و پایایی در پژوهش‌های فرهنگ‌های و اجتماعی، انتشارات جهاد دانشگاهی، مشهد
- [3] Graeser, A., Bommareddy, S. (1996). Integrating Questionnaire Design with a cognitive computational model of human question answering, *schwars,N, and sud mab,S(esd)*, 175-343.
- [4] Lord, F., Norich, M. (1968). *Statistical theories of mental test scores*, Addison-wesley, Oxford, England

برای انجام هر مطالعه‌ای در علوم مختلف، اولین کاری که بعد از مشخص شدن هدف انجام می‌شود این است که طرح آمارگیری مشخص شود. برای این که بتوان از این آمارگیری نتیجه‌ی مطلوبی به دست آورد، باید طرح‌های آمارگیری را مورد ارزش‌یابی قرار داد. متأسفانه فرایند سنجش و ارزش‌یابی سوال‌های آمارگیری، شامل ارزیابی مجموعی یا نهایی است که در پایان دوره‌ی آمارگیری برای رتبه‌بندی و تصمیم‌گیری صورت می‌گیرد. تردیدی نیست که نتایج حاصل از این تصمیم‌گیری‌ها دقیق نبوده و ممکن است عواقب ناگواری به همراه داشته باشد. در مقاله‌ی حاضر به چند روش برای ارزش‌یابی سوال‌های آمارگیری پرداخته شده است.

ارزش‌یابی به معنای جمع‌آوری، تحلیل و تفسیر داده‌ها، جهت بررسی میزان دسترسی به هدف‌های سازمانی و هدف‌های عملکردی می‌باشد. اطلاعات مورد نیاز ارزش‌یابی به کمک اندازه‌گیری و تحلیل جمع‌آوری می‌شوند و هرچه این اطلاعات جامع‌تر و کامل‌تر باشد، ارزش‌یابی درست‌تری صورت می‌پذیرد. ارزش‌یابی سوال دارای دو مؤلفه است، اول اینکه مواردی مانند "سوال‌ها تا چه اندازه خوب فهمیده شده‌اند یا اینکه پاسخ به سوال‌ها تا چه حدی مشکل بوده" بر کیفیت اندازه‌گیری تأثیر می‌گذارد و هراندازه فهم سوال آسان باشد خطای اندازه‌گیری کمتری به وجود می‌آورد. دوم، ارزش‌یابی تطابق پاسخ‌ها با موارد اندازه‌گیری است که همان برآورد مستقیم خطای اندازه‌گیری می‌باشد که این مؤلفه از دو جنبه روایی یا اربابی پاسخ و پایایی یا واریانس پاسخ مورد بررسی قرار می‌دهیم.

سه استاندارد برای سوال‌های آمارگیری وجود دارد که عبارت‌اند از:

الف) استاندارد محتوایی: یعنی آیا سوالی که برای مطالعه‌ی مورد نظر طراحی شده مفهوم را می‌رساند.

ب) استاندارد شناختی: یعنی آیا سوال از طرف پاسخ‌دهنده به طور کامل درک شده است.

ج) استاندارد استفاده‌پذیری: یعنی آیا پرسش‌نامه به صورتی که مدنظر هست توسط پرسش‌گران و پاسخ‌گویان تکمیل می‌گردد.

ارزش‌یابی توصیفی (پیش‌نویسی) پرسش‌ها: ارزش‌یابی توصیفی الگویی است که تلاش می‌کند زمینه‌ی طراحی پرسش‌نامه‌هایی با بهترین کارایی و کمترین خطا را فراهم سازد. گراسر در روش بازنگری کارشناسی، پرسش‌نامه از دو دیدگاه، یکی کارشناسان موضوعی و دیگری کارشناسان طراحی پرسش‌نامه مورد بررسی قرار میدهد. در این روش کارشناسان موضوعی به بازنگری جمله‌بندی سوال‌ها، ساختار سوال‌ها، گزینه‌های پاسخ‌ها، ترتیب سوال‌ها، دستورالعمل پرسش‌گران برای پرسش‌نامه و قواعد راهبردی پرسش‌نامه می‌پردازند.

گفتمان‌های گروه نظارت بر کیفیت و بهره‌وری: به کارگیری گروه‌های نظارت بر کیفیت و بهره‌وری یکی از روش‌های متداول پیش‌برد پرسش‌نامه است. دلیل به کارگیری این گروه این است که اولاً گروه نظارت بر کیفیت و بهره‌وری روش مناسبی است تا در مورد میزان اطلاع پاسخ‌گویان از موضوعات آمارگیر و بخش‌های مهم سوال از نظر آن‌ها اطلاعات کسب کرده و از آن در گروه‌بندی سوال‌ها و رسته‌بندی پاسخ‌ها استفاده کرد. دوماً، این گروه روشی مناسب برای شناسایی کلماتی که پاسخ‌گویان درباره‌ی موضوع به کار می‌برند و نیز چگونگی فهم این کلمات است.

مصاحبه‌های شناختی: در روش مصاحبه‌های شناختی پرسشگران سوال‌های پیش‌نویس را در مصاحبه‌های تکی اجرا می‌کنند و مورد بررسی قرار می‌دهند تا دریابند پاسخ‌گویان تا چه حدی سوال‌ها را فهمیده و پاسخ‌های خود را چگونه تنظیم می‌کنند.

پیش‌آزمون‌های میدانی: پیش‌آزمون‌ها تمرینی برای گردآوری داده‌ها در مقیاس کوچک هستند، در آنها تعداد اندکی مصاحبه با استفاده از روش‌های نمونه‌گیری میدانی مشابه به آمارگیری اصلی به وسیله‌ی پرسش‌گران اصلی انجام می‌شود. هدف این آزمون ارزش‌یابی ابزار آمارگیری، گردآوری داده‌ها و نیز روش‌های انتخاب پاسخ‌گویان است.

آزمایش‌های تصادفی شده: در روش آزمایش‌های تصادفی شده، طراحان آمارگیری به طور آزمایشی روش‌های مختلف گردآوری داده و همچنین سوال‌هایی با جملات متفاوت را مورد مقایسه قرار می‌دهند. این آزمایش می‌تواند مستقل یا بخشی از پیش‌آزمون‌ها باشد. این گونه آزمایش‌ها شواهدی از تأثیر جنبه‌های متفاوت سوال روی پاسخ‌ها هستند که با توجه به آن می‌توان نسخه‌ی که داده‌های بهتری دارد را انتخاب کرد.

برخی از روش‌های استانداردسازی پرسش‌نامه

در مؤلفه‌ی دوم ارزش‌یابی، ابتدا به توضیح اصطلاح‌های "روانسنجی" و "آماره‌های نمونه‌گیری" می‌پردازیم که از دو سوی متفاوت سوال‌های آمارگیری را تحت نظر قرار می‌دهند. روانسنجی روی پاسخ سوال‌های ارائه شده توسط هر پاسخ‌گو متمرکز است، و در آن از اصطلاح‌های "اعتبار" و "پایایی" استفاده می‌شود. آماره‌های نمونه‌گیری به آماره‌هایی که به خلاصه کردن همه پاسخ‌های تکی می‌پردازد متمرکز می‌شود و از اصطلاح‌های "اربابی" و "واریانس" استفاده می‌کند.

ابزاری جمع‌آوری داده‌ها نخست باید دارای روایی بوده و سپس پایایی داشته‌باشد.

روایی و اربابی پاسخ: روایی به این سوال پاسخ می‌دهد که ابزار اندازه‌گیری تا چه حد خصیصه مورد نظر را می‌سنجد. بدون آگاهی از روایی ابزار اندازه‌گیری نمی‌توان به داده‌های حاصل از آن اطمینان داشت. لورد و نوویک برای اولین بار تعریفی برای روایی ارائه داده‌اند که بر پایه‌ی یک مدل مفهومی ساده از فرایند اندازه‌گیری است که اگر μ_i را مقدار واقعی ساختار برای پاسخ‌گوی i ، Y_{it} را پاسخ به اندازه‌گیری توسط پاسخ‌گوی i در تکرار t ام و ε_{it} را خطای مرتبط با پاسخ Y_{it} در تکرار t ام تعریف شود، بنابراین پاسخی که پاسخ‌گو برای سوال پرسیده شده می‌دهد مدلی به صورت زیر خواهد داشت:

$$Y_{it} = \mu_i + \varepsilon_{it}$$



The 12th Seminar on
Probability and Stochastic
Processes
Semnan University
31 August-1 September 2019

دوازدهمین سمینار احتمال و فرایندهای تصادفی
دانشگاه سمنان ۹-۱۰ شهریور ۱۳۹۸



دوازدهمین سمینار احتمال و فرایندهای تصادفی

چکیده مقالات

۹ و ۱۰ شهریور ماه ۱۳۹۸

دانشگاه سمنان

دانشکده ریاضی، آمار و علوم کامپیوتر



The 12th Seminar on
Probability and Stochastic
Processes
Semnan University
31 August-1 September 2019

دوازدهمین سمینار احتمال و فرایندهای تصادفی

دانشگاه سمنان ۹-۱۰ شهریور ۱۳۹۸



مقدمه‌ای بر ارزش‌یابی پرسش‌های آمارگیری

محیا قاسمی^۱، مهدی جباری نوقابی^۲

^۱گروه آمار، دانشگاه فردوسی مشهد

^۲گروه آمار، دانشگاه فردوسی مشهد

چکیده: امروزه با گسترش تحقیقات در علوم مختلف، بسیاری از پدیده‌ها از طریق ابزارهایی مانند پرسش‌نامه‌ها مورد مطالعه قرار می‌گیرند. به منظور استفاده‌ی کارآمد از داده‌های آمارگیری، در مطالعه‌ی حاضر به بررسی روش‌های ارزیابی پرسشنامه‌های آمارگیری پرداخته شده است. بدین منظور، ارزش‌یابی سوال مؤلفه‌بندی شده و انواع مختلف روش‌ها برای این مؤلفه‌ها ارائه شده و مورد بررسی قرار داده می‌شود.

واژه‌های کلیدی: ارزش‌یابی، اعتبار، پایایی، طرح‌های آمارگیر.

کد موضوع‌بندی ریاضی (۲۰۱۰): 62Dxx, 62xx.

^۲محیا قاسمی: ghasemi@mail.um.ac.ir