

## بی‌پاسخی اقلام اطلاعاتی در آمارگیری‌ها و روش‌های مقابله با آن

آرمان بیداربخت‌نیا

مرکز آمار ایران

**چکیده.** در این مقاله، به رده‌بندی و بررسی الگوهای مختلف بی‌پاسخی در اقلام اطلاعاتی و منابع آن‌ها پرداخته می‌شود و سپس روش‌های جلوگیری از بروز این نوع بی‌پاسخی مورد بررسی قرار می‌گیرد. همچنین در مورد چگونگی تأثیر دانش گردآوری داده‌ها بر روش‌های آماری مقابله با بی‌پاسخی در اقلام اطلاعاتی بحث می‌شود.

### ۱- مقدمه

یکی از مشخصه‌های مهم کیفیت داده‌های حاصل از آمارگیری‌ها، میزان بی‌پاسخی در اقلام اطلاعاتی است. این خطا زمانی رخ می‌دهد که یک واحد اطلاع‌گیری (مثلاً، یک فرد) اطلاعات تولید می‌کند، اما به دلایلی، در مورد یک یا چند قلم اطلاعاتی یا پرسش خاص، اطلاعات مناسبی دریافت نمی‌شود. پیشگیری کامل از این خطا البته امکان‌پذیر نیست، اما می‌توان آن را به‌طور قابل ملاحظه‌ای کاهش داد. این کار نه تنها پژوهشگران را با اطلاعات بیش‌تری برای تحلیل مواجه می‌سازد و از ارزیابی در برآوردها جلوگیری می‌کند، بلکه به آنان کمک می‌کند تا با به‌کارگیری متغیرهای کمکی، جانمایی و تعدیل بهتری انجام دهند. تحلیلگران تا چندی پیش، این خطا را با حذف بی‌پاسخ‌ها و اجرای تحلیل فقط با مقادیر با پاسخ رفع می‌کردند، که از دست دادن اطلاعات از این طریق، باعث عدم کارایی برآوردها و کاهش توان آماری آزمون‌ها می‌شد.

---

واژگان کلیدی: بی‌پاسخی؛ تعدیل؛ جانمایی.

به‌منظور برخورد با این نوع خطا، شناخت، پیشگیری، و روش‌های مقابله با آن (حذف، جانمایی، ...) مانند حلقه‌های زنجیر به هم پیوسته‌اند، که در این مقاله وجوه مختلف آن‌ها مورد بررسی قرار می‌گیرد، به این ترتیب که نخست به رده‌بندی و بررسی الگوهای مختلف بی‌پاسخی در اقلام اطلاعاتی و منابع آن‌ها پرداخته می‌شود و سپس روش‌های جلوگیری از بروز آن‌ها بررسی می‌گردد. در انتها چگونگی تأثیر دانش گردآوری داده‌ها بر روش‌های آماری مقابله با بی‌پاسخی در اقلام اطلاعاتی مورد بحث قرار می‌گیرد.

## ۲- انواع بی‌پاسخی در اقلام اطلاعاتی

### ۲-۱- بی‌پاسخی کاملاً تصادفی

هرگاه عدم دریافت اطلاعات برای یک پرسش، بی‌ارتباط با مقادیر به‌دست‌آمده از دیگر پرسش‌ها، و نیز پاسخ واقعی خود پرسش باشد، بی‌پاسخی کاملاً تصادفی رخ داده است. مثلاً، از قلم‌افشادگی پرسش به‌طور تصادفی از جانب مصاحبه‌گر یا پاسخگو، از این نوع است. در این حالت، متغیرهای (پرسش‌های) بی‌پاسخ، یک نمونه‌ی تصادفی از کل متغیرها هستند و هیچ‌گونه تفاوت سیستماتیکی بین مقادیر بی‌پاسخ و باپاسخ وجود ندارد؛ لذا در نتایج حاصل، آریبی وجود نخواهد داشت و فقط توان آماری کاهش می‌یابد. این نوع بی‌پاسخی را قابل چشم‌پوشی گویند. در این حالت، تحلیل آماری را فقط با مقادیر باپاسخ می‌توان انجام داد.

### ۲-۲- بی‌پاسخی تصادفی

وقتی که عدم دریافت اطلاعات برای یک پرسش، با مقادیر به‌دست‌آمده از دیگر پرسش‌ها مرتبط بوده، اما ارتباطی با پاسخ واقعی آن پرسش نداشته باشد، بی‌پاسخی تصادفی رخ داده است. به‌عنوان مثال، یک شخص مسن به‌دلیل ضعف حافظه به‌سختی می‌تواند پرسش مربوط به یک حادثه را که در گذشته رخ داده است، پاسخ دهد. لذا پاسخ به پرسش مربوط به آن حادثه ارتباط زیادی با متغیر سن دارد، اما به خود آن حادثه ارتباطی ندارد و متغیر سن می‌تواند به‌عنوان یک متغیر کمکی در روش‌های آماری مناسب به کار

گرفته شود. بی‌پاسخ‌ها در این حالت، نمونه‌ای تصادفی از کل مقادیر درون رده‌هایی هستند که متغیرهای مشاهده‌شده‌ی دیگر ایجاد کرده‌اند. این نوع بی‌پاسخی را نیز به شرطی که مدل آماری مناسبی برای تحلیل داده‌ها به کار گرفته شود، قابل چشم‌پوشی گویند.

### ۳-۲- بی‌پاسخی ناتصادفی

هرگاه عدم دریافت اطلاعات برای یک پرسش، مرتبط با پاسخ واقعی آن پرسش باشد، گوییم بی‌پاسخی ناتصادفی رخ داده است. به‌عنوان مثال، اگر پاسخگو متوجه شود که پاسخ واقعی او امر نامطلوبی در جامعه تلقی می‌شود (مثل مصرف الکل)، ممکن است با جمله‌هایی چون «نمی‌دانم» یا به‌طور کلی با عدم پاسخ‌گویی مواجه شویم. در این حالت ممکن است اریبی بسیار زیادی ایجاد شود. این نوع بی‌پاسخی را بر خلاف دو حالت قبلی، بی‌پاسخی غیر قابل چشم‌پوشی گویند و راه‌حل ساده‌ای برای مقابله با آن‌ها در دست نیست. به هر حال در تحلیل باید بی‌پاسخی را در مدل منظور کرد تا از اریبی کاسته شود.

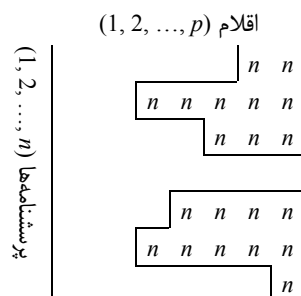
### ۳- الگوهای بی‌پاسخی در ارقام اطلاعاتی

همان‌گونه که اشاره شد، آگاهی از مکانیسم بی‌پاسخی، اعم از کاملاً تصادفی، تصادفی، یا ناتصادفی، می‌تواند محقق را در کاهش بی‌پاسخی به‌طور قابل ملاحظه‌ای یاری دهد. مطالعه‌ی الگوهای بی‌پاسخی در ارقام اطلاعاتی، گام مهمی در راه شناخت این مکانیسم است.

#### ۳-۱- الگوی بی‌پاسخی یکنوا (جزئی)

در این الگو، از پرسش خاصی به بعد، همه‌ی پرسش‌ها بی‌پاسخ هستند (شکل ۱). این الگو ارتباط مستقیمی با زمان دارد، زیرا پس از یک زمان مشخص است که با بی‌پاسخی مواجه می‌شویم. دو مثال عمده از این الگو، مرگ و میر در آمارگیری‌های پانلی<sup>۱</sup> و توقف در حین مصاحبه می‌باشند، که آمارشناسان زیادی (از جمله فان‌دیل [۵] و انگل و راینه‌که [۲]) در این زمینه کار کرده‌اند.

در آمارگیری‌های پانلی، بی‌پاسخی حاصل از ترک پانل توسط پاسخگو نمی‌تواند کاملاً تصادفی باشد. این بی‌پاسخی اگر به موضوع مورد مطالعه بستگی نداشته باشد، یک بی‌پاسخی تصادفی است و در غیر این صورت، ناتصادفی خواهد بود. توقف در حین مصاحبه عموماً در مصاحبه‌های تلفنی اتفاق می‌افتد.



شکل ۱- الگوی بی‌پاسخی یکنوا

### ۳-۲- الگوی بی‌پاسخی آزاد

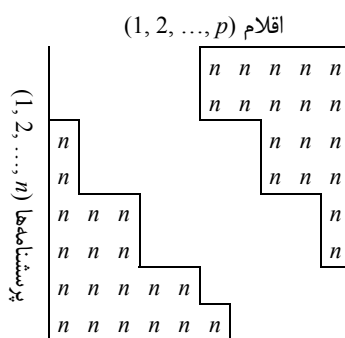
اگر الگوی بی‌پاسخی در ماتریس داده‌ها از فرم یکنوا پیروی نکند، الگوی آزاد نامیده می‌شود. الگوهای بی‌پاسخی آزاد به دو گروه باقاعده (شکل ۲) و بی‌قاعده (شکل ۳) تقسیم می‌شوند.

#### ۳-۲-۱- الگوی باقاعده: دو دلیل عمده برای این نوع بی‌پاسخی وجود دارد:

آ) برخی از پرسش‌ها مختص همه‌ی پاسخگویان نیستند، به این نوع بی‌پاسخی، بی‌پاسخی منطقی نیز گفته می‌شود.

ب) از طرح یکسانی برای ثبت اطلاعات مربوط به زیرمجموعه‌های متفاوتی از پرسش‌ها یا ایجاد محرک‌هایی برای افراد متفاوت استفاده می‌شود. بر خلاف الگوی بی‌پاسخی منطقی، در این‌جا همه‌ی پرسش‌ها مربوط به تمام پاسخگویان هستند. با این حال، طرح همه‌ی پرسش‌ها مستلزم تکمیل پرسش‌نامه‌ای حجیم و صرف زمان زیادی است؛ لذا

زیرمجموعه‌های خاصی از پرسش‌ها از گروه‌های متفاوتی از افراد پرسیده می‌شوند. طرح‌هایی کاربردی در این زمینه اجرا شده، که مزایای آن‌ها مشهود است؛ از جمله این که نیاز به منابع و وقت کمتری است، پاسخگویان برای تکمیل پرسش‌نامه تحت فشار کمتری قرار دارند، و گردآوری داده‌ها کارآتر است.



**۲-۲-۳- الگوی بی‌قاعده:** در مورد تعدادی از افراد یا همه‌ی آنان بی‌پاسخی برای برخی از پرسش‌ها در فواصل نامنظم اتفاق می‌افتد. این نوع بی‌پاسخی در ارقام اطلاعاتی ممکن است به دلایل عمده‌ای رخ دهد:

آ) اطلاعات مربوط به پرسش خاصی برای یک پاسخگو دریافت نشود؛ مثلاً، یک پرسش به‌طور تصادفی از قلم بیفتد، فرد پاسخگو پاسخ را ندادند، یا در مورد پرسش خاصی حاضر به پاسخ‌گویی نشود. مسئله‌سازترین حالت، زمانی است که موفق به دریافت هیچ اطلاعاتی از سوی پاسخگو نشویم؛ زیرا در این حالت ممکن است مکانیسم‌های متفاوتی از بی‌پاسخی را در پیش رو داشته باشیم.

در حالتی که پرسش به‌طور تصادفی از قلم می‌افتد، با بی‌پاسخی کاملاً تصادفی مواجه هستیم، که در این حالت، مکانیسم بی‌پاسخی قابل چشم‌پوشی است و می‌توان روش‌های ساده‌ی آماری مانند روش حذف listwise را به کار برد.

وقتی پاسخگو با وجود تمایل خود، قادر به پاسخ‌گویی نیست، ممکن است حالت بی‌پاسخی تصادفی رخ دهد (مثلاً، بی‌پاسخی به‌دلیل ضعف حافظه بر اثر پیری)، که در

این مورد می‌توان روش‌های جانمایی را به کار گرفت. در عین حال در برخی موارد ممکن است با حالت بی‌پاسخی ناتصادفی روبه‌رو شویم، که بر خلاف دو حالت قبل، مکانیسم بی‌پاسخی آن غیر قابل چشم‌پوشی است و مدل‌های متناسب با بی‌پاسخی باید در تحلیل منظور گردد.

ب) اطلاعات دریافت‌شده از جانب پاسخگو قابل استفاده نباشد. مثلاً، پاسخ به پرسش خاصی ناممکن، خارج از محدوده، غیر قابل کدگذاری، یا ناخوانا باشد. این مشکل عموماً ارتباط مستقیمی با طرح پرسش یا پاسخ واقعی آن دارد – که به‌عنوان بی‌پاسخی ناتصادفی از آن یاد شد – و اطلاع از آن می‌تواند منجر به شناخت مکانیسم بی‌پاسخی و کاهش آریبی حاصل از آن گردد. مثلاً، با استفاده از پیش‌آزمون می‌توان این مورد را کنترل کرد.

پ) اطلاعات مفید از بین رفته باشد. مثلاً در وارد کردن داده‌ها اشتباه شده باشد. این نوع بی‌پاسخی، کاملاً تصادفی و لذا قابل چشم‌پوشی است.

تفاوت‌های مهمی بین این سه عامل وجود دارد. دو عامل نخست در مرحله‌ی گردآوری داده‌ها اتفاق می‌افتند و اغلب در نتیجه‌ی مشکلاتی است که در فرایند پرسش و پاسخ به وجود می‌آید. اما عامل سوم بر اثر اشتباهاتی است که در مرحله‌ی پردازش داده‌ها (وارد کردن، ویرایش، کدگذاری) رخ می‌دهد.

$$\begin{array}{c} \text{اقلام } (1, 2, \dots, p) \\ \left| \begin{array}{cccc} n & & & \\ & n & & \\ & & & n \\ & & n & \\ n & n & & \\ & n & n & \\ & & & n \\ n & & n & n \end{array} \right| \\ \text{پرسشنامه‌ها } (1, 2, \dots, n) \end{array}$$

شکل ۳- الگوی بی‌پاسخی آزاد بی‌قاعده

#### ۴- کاهش بی‌پاسخی در ارقام اطلاعاتی

شناخت عوامل و خاستگاه‌های بی‌پاسخی در ارقام اطلاعاتی، بهترین راه برای جلوگیری از بروز این نوع خطا می‌باشد. در این بخش، به بررسی این خاستگاه‌ها می‌پردازیم.

##### ۴-۱- روش گردآوری اطلاعات

سه روش عمده که در مصاحبه‌ها برای گردآوری اطلاعات به کار می‌روند، عبارت‌اند از: استفاده از پرسش‌نامه‌های خویش‌اجراگرانه (self-administered)، مصاحبه‌ی حضوری، مصاحبه‌ی تلفنی. ممکن است در یک آمارگیری (مانند آمارگیری پانلی)، ترکیبی از روش‌های مختلف به کار گرفته شود.

دلی‌یو [۳] نشان داده است که مصاحبه‌های حضوری و تلفنی نسبت به پرسش‌نامه‌های خویش‌اجراگرانه (مانند پرسش‌نامه‌ی آمارگیری‌های پستی) با بی‌پاسخی‌های کمتری همراه هستند. البته در مورد پرسش‌های حساس، نتیجه کاملاً بر عکس می‌باشد. همچنین بررسی‌های مقایسه‌ای نشان می‌دهد، مصاحبه‌هایی که در آن‌ها از رایانه برای ثبت اطلاعات استفاده می‌شود، در مقایسه با مصاحبه‌هایی که از کاغذ و قلم استفاده می‌کنند، از درصد خطا و بی‌پاسخی کمتری برخوردارند.

##### ۴-۲- پرسش‌نامه

حتی بهترین مصاحبه‌گر نیز ممکن است دچار اشتباه شود. یک پرسش‌نامه‌ی مناسب می‌تواند از اشتباهات مصاحبه‌گر و همچنین پاسخگو تا حدودی جلوگیری کند. ویژگی‌های یک پرسش‌نامه‌ی خوب را به شرح زیر می‌توان برشمرد:

- دارای ویرایش مطلوب باشد. برای این منظور، لازم است اصول مربوط به طرح گرافیکی رعایت شود، مصاحبه‌گر و پاسخگو از پرسشی به پرسش دیگر راهنمایی شوند و پاسخگو به یک درک شهودی برسند.
- پرسش‌ها قابل فهم باشند.
- دسته‌بندی گزینه‌های پاسخ باید با موضوع پرسش، سازگار باشد. تعداد آن‌ها نیز

خیلی مهم است. زیاد بودن گزینه‌ها بهتر از ناقص بودن است؛ زیرا پاسخگو با محدودیت در انتخاب پاسخ مناسب، دچار مشکل می‌شود و مصاحبه‌گر را با پاسخ « نمی‌دانم » یا انصراف از پاسخ‌گویی مواجه می‌سازد. تجربه نشان داده است که از چهار تا هفت گزینه برای دسته‌بندی پاسخ، مناسب‌ترین تعداد است. البته همه‌ی گزینه‌ها باید برای پاسخگو کاملاً معنادار باشند. (به‌عنوان مثال، گزینه‌ی « بی‌طرف » فقط برای پرسش‌های دارای مقیاس دوقطبی معنادار است.) باید توجه داشت که تعداد گزینه‌های پاسخ به نوع گردآوری داده‌ها نیز بستگی دارد. مثلاً، در مصاحبه‌های تلفنی، حد اکثر تعداد پنج گزینه برای پاسخ می‌تواند کارایی داشته باشد.

در صورت عدم رعایت نکات فوق، نتایج حاصل، داده‌ها دارای بی‌پاسخی ناتصادفی خواهند بود. لذا در صورت امکان، جلوگیری از بروز این نوع بی‌پاسخی به وسیله‌ی اصلاح پرسش‌نامه، بسیار حائز اهمیت است.

یکی از سؤال‌های مهم و بحث‌برانگیز در کاهش بی‌پاسخی در ارقام اطلاعاتی این است که: آیا گزینه‌ی « نمی‌دانم » در گزینه‌های پاسخ لحاظ شود یا خیر؟ از دیدگاه عملی نباید امکان گریز از پاسخ‌گویی با استفاده از گزینه‌ی « نمی‌دانم » یا « نظری ندارم » را برای پاسخگو فراهم کرد، مگر در موارد لازم. مثلاً، در آمارگیری‌هایی که در خصوص تصمیم‌گیری مردم در مورد یک رأی‌گیری سؤال می‌شود، گزینه‌ی « نمی‌دانم » نه‌تنها بی‌پاسخی نیست، بلکه حاوی اطلاعات می‌باشد و در چنین مواردی باید بر حسب نیاز، امکان این انتخاب را برای پاسخگو فراهم کرد.

راه‌هایی برای تصمیم‌گیری در این موارد پیشنهاد شده است. مثلاً، در مصاحبه‌های حضوری گزینه‌ی « نمی‌دانم » ضروری است، اما مصاحبه‌گر نباید در حین خواندن پاسخ‌ها آن را ذکر کند. در پرسش‌نامه‌های خویش‌اگرانه نیز باید از آوردن این قبیل گزینه‌های پاسخ در پرسش‌نامه خودداری کرد، مگر این‌که واقعاً ضروری باشد. اگر تکمیل پرسش‌نامه به‌کمک رایانه صورت می‌گیرد، می‌توان به پاسخگو اجازه داد که برای پاسخ ندادن به یک پرسش، کلید خاصی را فشار دهد. در پرسش‌نامه‌های پستی نیز می‌توان به پاسخگو اطمینان داد که اگر پاسخی را واقعاً نمی‌داند، مجاز به عدم پاسخ‌گویی است.

پس از به‌کارگیری نکات فوق در طراحی پرسش‌نامه، و پیش از اجرای نهایی، باید آن را مورد آزمایش قرار دهیم. برای این کار، دو راه پیشنهاد می‌شود: پیش‌آزمون (pretest) و استفاده از نمونه‌ی مقدماتی (pilot).

۱-۲-۴- پیش‌آزمون: در این روش، تمام مراحل پرسش و پاسخ با استفاده از مصاحبه‌ی دقیقی که با افراد یا گروه‌ها صورت می‌گیرد آزمایش می‌شود و مشکلات پرسش‌نامه و چگونگی برداشت پاسخگو از پرسش‌ها مورد بررسی قرار می‌گیرد. پیش‌آزمون معمولاً قبل از نهایی شدن پرسش‌نامه، با دعوت از تعداد کمی از پاسخگویان و مصاحبه با آنان صورت می‌گیرد.

۲-۲-۴- نمونه‌ی مقدماتی: از این روش در آمارگیری‌های با مقیاس بزرگ استفاده می‌شود، که شامل نمونه‌گیری، مراجعه به پاسخگو، گردآوری داده‌ها، کدگذاری، و ویرایش داده‌ها می‌باشد و در واقع، تمامی مراحل یک آمارگیری را مورد آزمایش قرار می‌دهد؛ در حالی که در روش پیش‌آزمون فقط روند پرسش و پاسخ آزمایش می‌شود.

### ۳-۴- پاسخگو

دلایل مختلفی از قبیل موضوع پرسش (مانند مسائل حساس اجتماعی)، ساختار پرسش (باز یا بسته بودن)، دشوار بودن پرسش، بروز مشکل در فرایند پرسش و پاسخ (مثلاً، پاسخگو متوجه پرسش نشود)، یا ضعف انگیزه‌ی پاسخ‌گویی از جانب فرد، ممکن است باعث بی‌پاسخی در اقلام اطلاعاتی از جانب پاسخگو شود. از آن‌جا که عدم پاسخ‌گویی ارتباط زیادی با میزان سواد و سن پاسخگو دارد [۱] و در افراد مسن و کم‌سواد، بی‌پاسخی بیش‌تری مشاهده می‌شود، این نوع بی‌پاسخی، کاملاً تصادفی نیست و حذف مقادیر بی‌پاسخ، موجب اریبی در نتایج خواهد شد. اما چند روش برای کمینه کردن این نوع بی‌پاسخی پیشنهاد می‌شود، که عبارت‌اند از:

- توجه به قابل درک بودن پرسش‌های پرسش‌نامه، ویرایش آن‌ها، مرتبط بودن پرسش با گزینه‌های پاسخ، و برخی موارد دیگر که در بخش مربوط به پرسش‌نامه ذکر شد.

- استفاده از رایانه در مصاحبه‌ها. این کار می‌تواند باعث کاهش خطا از جانب پاسخگو شود؛ زیرا در این صورت، خود پاسخگو است که آهنگ و سرعت پاسخ‌گویی را تنظیم می‌کند و لذا برای فهمیدن و پاسخ دادن به پرسش‌ها وقت کافی را در اختیار خواهد داشت و در پاسخ‌گویی به پرسش‌های حساس نیز صداقت بیش‌تری از خود نشان خواهد داد.
- یک معارفه‌ی جذاب و دقیق از جانب مصاحبه‌گر تعلیم‌دیده، یا یک مقدمه‌ی مناسب برای پرسش‌نامه، که می‌تواند به تولید پاسخ صحیح و هدایت پاسخگو در فرایند پرسش و پاسخ کمک کند.
- استفاده از روش‌های پاسخ تصادفی (لنس‌فلت-مول‌درس و دیگران [۴])، که می‌تواند راهی برای ایجاد اطمینان در فرد و دریافت پاسخ صحیح از وی باشد. به‌خصوص در مورد پرسش‌های حساس، علاوه بر پرسش‌نامه، روش گردآوری اطلاعات نیز باید به‌گونه‌ای باشد که اطمینان و احساس امنیت را در پاسخگو به وجود آورد، که روش پاسخ تصادفی، یکی از راه‌های دست‌یابی به این هدف می‌باشد.

به‌طور کلی، (هنگامی که در مورد مسائل حساس مصاحبه می‌شود، نه‌تنها در پرسش‌نامه)، بلکه در تمامی فرایند پرسش و پاسخ باید جانب احتیاط را نگه داشت. مثلاً، اطلاع‌رسانی در مورد چگونگی استفاده از نتایج (پس از گردآوری)، یا چگونگی استفاده از نتایج حاصل و دلیل طرح موضوع می‌تواند در پاسخگو ایجاد اطمینان کند. لذا در این موارد برای این که بی‌پاسخی به جواب واقعی پرسش ارتباط پیدا نکند و با بی‌پاسخی ناتصادفی مواجه نشویم، باید پرسش‌های پرسش‌نامه را به‌شدت کنترل کرد. برای دست‌یابی به این هدف، معمولاً باید پرسش‌ها را به شکل بی‌طرفانه و غیر انتقادی طرح کرد. همچنین دامنه‌ی گزینه‌های پاسخ نیز تا حد ممکن باید وسیع باشد.

در برخی از مصاحبه‌ها ممکن است در مورد اتفاقاتی که در گذشته رخ داده‌اند، پرسش شود. در این صورت، بهتر است برای کمک کردن به حافظه‌ی پاسخگو، از پرسش‌های کمکی استفاده شود.

#### ۴-۴- مصاحبه‌گر

مصاحبه‌گران عامل مهمی در کاهش بی‌پاسخی هستند، اما با این حال می‌توانند منبع بروز خطا نیز باشند. یک مصاحبه‌گر ممکن است در پرسیدن سؤال، دریافت اطلاعات حساس، ثبت پاسخ، و موارد مشابه دچار خطا گردد، که هر یک ممکن است به دو صورت رخ دهد:

- مصاحبه‌گر به شکل غیر عمدی و بر حسب تصادف، مرتکب اشتباه شود (مانند از قلم انداختن سؤال)؛
- مصاحبه‌گر به دلیل تعجیل در کار یا دلایل دیگر، به‌طور مغرضانه مرتکب خطا شود (مثلاً، از برخی پرسش‌ها چشم‌پوشی کند یا توجه پاسخگو را به گزینه‌هایی مثل «نمی‌دانم» جلب کند)، که با نظارت و بازرسی دقیق می‌توان مانع این امر شد.

نقش مهم مصاحبه‌گر در کاهش بی‌پاسخی‌های ناتصادفی را نباید فراموش کرد. لذا آموزش صحیح مصاحبه‌گر و آگاه کردن وی از ساختار پرسش‌نامه می‌تواند در کاهش این نوع بی‌پاسخی مفید باشد.

#### ۴-۵- پردازش داده‌ها

حتی اگر محققان، تمامی تلاش خود را برای جلوگیری از بروز خطا در فرایند پرسش و پاسخ به کار گیرند، باز هم در عملیات مربوط به وارد کردن داده‌ها، کدگذاری، و ویرایش ممکن است خطاهایی رخ دهد. استفاده از رایانه برای گردآوری داده‌ها در آمارگیری‌ها، ابزاری مفید برای پیشگیری از بروز این خطاهاست. همچنین کنترل کیفیت در حین کدگذاری و تحلیل نیز می‌تواند مفید باشد.

#### ۵- روش‌های مقابله با بی‌پاسخی در ارقام اطلاعاتی

اطلاع از مکانیسم بی‌پاسخی می‌تواند راهی مفید برای پیشگیری از بی‌پاسخی، یا انتخاب یک روش آماری مناسب برای کاهش آثار آن باشد. بررسی الگوی بی‌پاسخی، اولین گامی

است که در راه مقابله با این مشکل باید برداشته شود؛ زیرا این الگو ممکن است حاوی اطلاعات مفیدی باشد. به عنوان مثال، اگر مشاهده شود که اغلب بی‌پاسخی‌ها فقط مربوط به یک متغیرند که تأثیر زیادی بر تحلیل ما ندارد و یک متغیر مرکزی محسوب نمی‌شود، پیش از آن که با نگر داشتن این متغیر در مدل، دچار مشکل شویم یا مجبور به انجام جانهی‌های زیاد باشیم، باید این متغیر را از مدل حذف کنیم. این موضوع در مورد یک واحد اطلاعاتی که بی‌پاسخی زیادی تولید کرده باشد نیز صدق می‌کند. به طور کلی، اگر الگوی خاصی در توزیع بی‌پاسخی در ماتریس داده‌ها مشاهده شود، می‌توان در مدل تحلیلی یا انتخاب روش آماری مناسب از آن استفاده کرد. نرم‌افزارهایی برای بررسی الگوهای بی‌پاسخی در ماتریس داده‌ها موجود است؛ از جمله SPSS که در بخش MVA مجهز به این تکنیک می‌باشد، و همچنین برنامه‌ی SOLAS که امکان بررسی الگوهای بی‌پاسخی را برای کاربر ایجاد می‌کند.

ممکن است بررسی الگوی بی‌پاسخی نتواند در تشخیص این که آیا بی‌پاسخی در قلم اطلاعاتی، با پاسخ واقعی آن پرسش مرتبط بوده است یا خیر، به ما کمکی کند. لذا برای آزمون تصادفی بودن به اطلاعات کمکی نیاز است، که این اطلاعات کمکی می‌تواند اطلاعات جانبی حاصل از منابع دیگر باشد، یا از طریق مراجعه به پاسخگو (پیش‌آزمون و استفاده از نمونه‌ی مقدماتی) حاصل شود.

اگر در بررسی مقدماتی داده‌ها متوجه بی‌پاسخی زیاد در چند متغیر شویم، با اجرای یک پس‌آزمون (post-test) می‌توانیم اطلاعات کمکی تولید کنیم. پس‌آزمون، عبارت است از مصاحبه‌ی دقیق و هدفمند با تعداد کمی از افرادی که خصوصیت‌های مشابهی با افراد تولیدکننده‌ی بی‌پاسخی در آمارگیری دارند. راه دیگر برای گردآوری اطلاعات کمکی، مراجعه به افرادی است که تولید بی‌پاسخی کرده‌اند. اما هر دو روش فوق، مستلزم صرف وقت و هزینه‌ی بالایی هستند. یک روش کم‌هزینه برای گردآوری این اطلاعات مفید، استفاده از مشاهدات مصاحبه‌گران است. گزارش هوشمندانه‌ی مصاحبه‌گر از عملیات پرسش و پاسخ و مشکلاتی که با آن مواجه بوده است، اطلاعات مفیدی برای دستیابی به دلایل بی‌پاسخی در ارقام اطلاعاتی به دست می‌دهد.

از آن جا که هیچ‌گاه نمی‌توانیم مطمئن باشیم که به‌طور کامل مانع از بروز بی‌پاسخی

شده‌ایم، پس از این‌که همه‌ی تلاش خود را برای کاهش بی‌پاسخی به کار بستیم، در صدد انتخاب روشی مناسب برای مقابله با آن هستیم، به‌طوری‌که اثر اریبی حاصل از بی‌پاسخی را به بهترین نحو کاهش دهد. نرم‌افزارهای موجود آماری معمولاً به‌طور خودکار، روش‌های حذف (listwise, pairwise) و برخی روش‌های بسیار ساده‌ی جانپهی (جانپهی مقدار میانگین کل، جانپهی به‌روش رگرسیون) را در برخورد با بی‌پاسخ‌ها اتخاذ می‌کنند. تمامی این روش‌ها بر اساس فرضیه‌ی کاملاً تصادفی بودن بی‌پاسخ‌هاست، که در موارد بسیار کمی با این حالت مواجه خواهیم بود. لذا روش‌های نام‌برده، در صورت تصادفی بودن یا ناتصادفی بودن بی‌پاسخ‌ها، همگی سبب مشکلاتی از قبیل اریبی و کاهش مصنوعی پی‌مقدار ( $p$ -value) می‌شوند. در انتخاب روش آماری مناسب باید به دو نکته توجه نمود: (۱) استفاده از تمام اطلاعات موجود برای دستیابی به الگوی بی‌پاسخی در ماتریس داده‌ها، و (۲) استفاده از روش تحلیلی مناسب برای مجموعه مقادیر باپاسخ، به‌طوری‌که اثر مقادیر بی‌پاسخ را به‌شکل صحیح و شایسته‌ای تعدیل نماید.

یکی از روش‌های مناسب برای حالتی که بی‌پاسخی تصادفی رخ داده باشد، تحلیل مقادیر باپاسخ به‌شرط مشاهده‌ی متغیرهای کمکی است، که اثر زیادی بر کاهش اریبی ناشی از بی‌پاسخی دارد. رایین و دیگران [۷] اشاره کرده‌اند که در آمارگیری‌های گسترده که تماس به‌طور کامل برقرار شده باشد، برای شروع تحلیل باید فرض تصادفی بودن بی‌پاسخی را در نظر گرفت. در صورتی که بی‌پاسخی از نوع تصادفی فرض شود، دو روش را برای تحلیل می‌توان به کار گرفت: برآورد مستقیم و روش‌های جانپهی.

در روش برآورد مستقیم، مقادیر باپاسخ با استفاده از روش بیشینه‌ی درست‌نمایی، به‌طور کامل مورد تحلیل قرار می‌گیرند. برخی از برنامه‌های کامپیوتری که در مدل‌سازی معادلات ساختاری می‌توانند با ماتریس ناقص داده‌ها کار کنند، عبارت‌اند از AMOS و Mplus و Lisrel 8.5، که البته دست‌یابی به آن‌ها کار آسانی نیست. این در حالی است که نرم‌افزارهای استاندارد و رایج آماری، روش‌های جانپهی را به‌طور خودکار (در یک مرحله) انجام می‌دهند. در این روش‌ها مکان‌های بی‌پاسخ در ماتریس داده‌ها با مقادیر مناسبی جایگزین می‌شوند. پس از آن می‌توان (بر خلاف روش برآورد مستقیم) ماتریس کامل داده‌ها را در تحلیل شرکت داد. همان‌گونه که پیش‌تر گفته شد، تمامی روش‌های

ساده‌ی به کار گرفته شده در نرم‌افزارهای آماری، برآوردهایی اریب ارائه می‌کنند. لذا روش بی‌درنگ (hot deck) به عنوان روشی مناسب بر پایه‌ی فرض بی‌پاسخی تصادفی برای دستیابی به نتایج ناریب پیشنهاد می‌گردد. در این روش، پس از آن که گروه‌هایی بر اساس متغیرهای کمکی (که با توجه به الگوی بی‌پاسخی انتخاب می‌شوند) تشکیل شد، واحدهای اطلاعاتی در این گروه‌ها به گونه‌ای مرتب می‌شوند که ترتیب ثابت واحدهای نمونه، متناظر با ساختار طرح نمونه باشد. سپس مکان‌های بی‌پاسخ هر گروه، توسط ثبت‌کننده و با استفاده از یک مقدار تصادفی از گروه مربوط، بی‌درنگ جایگزین می‌شوند. دو مشکل اساسی ممکن است با عملیات جانهی همراه باشد:

- (آ) استفاده از اطلاعات و مقادیر مشاهده شده برای پیش‌بینی مقادیر بی‌پاسخ، منجر به یک ساختار در مجموعه‌ی کامل شده‌ی داده‌ها می‌گردد؛ و
- (ب) از آن‌جا که در تحلیل مجموعه داده‌های کامل شده، به‌طور مصنوعی از تعداد زیادی مشاهده استفاده می‌شود، این کار منجر به اریبی در آزمون‌های معناداری می‌گردد.

برای برطرف کردن این مشکلات، رابین [۶] روش جانهی چندگانه را پیشنهاد نمود، که در آن، یک مقدار بی‌پاسخ با دو یا چند (M) مقدار برآورد شده جایگزین می‌گردد. لذا یک مجموعه‌ی M تایی از ماتریس‌های کامل شده خواهیم داشت، که در این جا مقادیر برآورد شده باید شامل یک جمله‌ی خطا و از یک توزیع مناسب باشند. این کار از ایجاد یک ساختار پس از کامل شدن ماتریس داده‌ها جلوگیری می‌کند. پس از جانهی چندگانه باید M ماتریس کامل شده را تحلیل و سپس برآوردها را ترکیب نماییم، که این عمل نیز مشکل اریبی مذکور را برطرف می‌کند.

به‌طور کلی، جانهی چندگانه ممکن است به سه روش انجام شود: روش رگرسیونی پارامتری، روش رگرسیونی ناپارامتری (propensity score)، روش زنجیره‌ی مارکوف مونت‌کارلویی (MCMC). روش رگرسیونی فقط برای الگوی یکنوا کاربرد دارد، ولی روش MCMC برای همه‌ی الگوهای بی‌پاسخی به کار گرفته می‌شود.

نرم‌افزارهای SAS 8.2 و SOLAS برای اجرای عملیات جانهی چندگانه مناسب‌اند. در SAS، دو برنامه‌ی PROC MI و PROC MIANALYZE به ترتیب

برای جانمایی و ترکیب برآوردها به کار گرفته می‌شوند.

## ۶- نتیجه‌گیری

با بررسی دقیق مراحل عملیات، می‌توان میزان بی‌پاسخی در ارقام اطلاعاتی را کاهش داد. به‌طور کلی عواملی مانند آموزش مناسب مصاحبه‌گر، روش‌های مناسب برای گردآوری داده‌ها، و طراحی خوب پرسش‌نامه، همگی به کاهش بی‌پاسخی در ارقام اطلاعاتی کمک می‌کنند. پیش‌آزمون پرسش‌نامه برای رفع مشکلاتی نظیر جمله‌بندی و همچنین ترتیب پرسش‌ها بسیار مؤثر است. در آمارگیری‌های گسترده، استفاده از نمونه‌ی مقدماتی می‌تواند محققان را در بازرسی نهایی مقادیر بی‌پاسخ یاری دهد، به‌طوری که با تحلیل آماری داده‌های به دست آمده از نمونه‌ی مقدماتی می‌توان به الگوی بی‌پاسخی پی برد، که در طراحی نهایی پرسش‌نامه بسیار مؤثر است.

اطلاع از این موضوع که بی‌پاسخ‌ها چگونه و با چه الگویی در ماتریس داده‌ها رخ داده‌اند، نه تنها در توسعه و بهبود روش‌های گردآوری داده‌ها، بلکه در انتخاب یک روش آماری مناسب برای برخورد با بی‌پاسخی در ارقام اطلاعاتی، به محقق کمک می‌کند. پیشگیری کامل از بی‌پاسخی در ارقام اطلاعاتی امکان‌پذیر نیست. بر این اساس، لازم است در تحلیل آماری، روش‌هایی را برای مقابله با بی‌پاسخی به کار ببریم. لذا اطلاع بیش‌تر محقق از مکانیسم بی‌پاسخی، وی را در تعدیل هرچه بهتر نتایج یاری می‌دهد. پس به‌طور کلی اولین راهبرد، پیشگیری از بروز بی‌پاسخی تا حد ممکن، و گردآوری اطلاعات کمکی (مثلاً، با استفاده از پیش‌آزمون یا اطلاعات مصاحبه‌گر) می‌باشد. در گام دوم باید مراحل دوگانه‌ی تحلیل را دنبال کرد؛ یعنی، (۱) استفاده از همه‌ی اطلاعات کمکی در تشخیص الگوی بی‌پاسخی، و (۲) تحلیل ماتریس داده‌ها به‌طوری که تعدیل لازم برای مقادیر بی‌پاسخ صورت گرفته باشد.

## توضیح

<sup>۱</sup> در این آمارگیری‌ها که در مراحل مختلف انجام می‌شوند، در هر مرحله از آمارگیری، بخشی از نمونه متعلق به مرحله‌ی قبل است.

## مرجع‌ها

- Colsher, P.L.; Wallace, R.B. (1989). Data quality and age: health and psycho behavioral correlates of inconsistent responses. *J. Gerontology*. **44**, 45-52. [۱]
- Engel, U.; Reinecke, J. (1994). *Panelanalyse: Grundlagen, Techniken, Beispiele*. Walter de Gruyter, Berlin [In German]. [۲]
- de Leeuw, E.D. (1992). *Data Quality in Mail, Telephone, and Face to Face Surveys*. TT-publikaties, Amsterdam. [۳]
- Lensvelt-Mulders, G.J.L.M.; Hox, J.J.; van der Heijden, P.G.M.; Mass, C.J.M. (2005). Meta-analysis of randomized response research. *Sociol. Methods Res.* **33**, 319-348. [۴]
- van de Pol, F.J.R. (1989). *Issues of Design and Analysis of Panels*. Sociometric Research Foundation, Amsterdam. [۵]
- Rubin, D.B. (1987). *Multiple Imputation for Nonresponse in Surveys*. Wiley, New York. [۶]
- Rubin, D.B.; Stern, H.S.; Vehovar, V. (1995). Handling “don’t know” survey responses: the case of the Slovenian Plebiscite. *J. Amer. Statist. Assoc.* **91**, 822-828. [۷]