



پژوهش‌های آمار



مرکز آمار ایران

# به‌کارگیری پاراداداه در آمارگیری نیروی کار ایران

مجری:

لیدا کلهری ندرآبادی

همکاران:

محمد رضا ریحانی

ندا مصطفوی

مشاور:

فرهاد مهران

گروه پژوهشی طرح‌های فنی و روش‌های آماری

پاییز ۱۳۹۹



# بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

کد شناسه: RP-9930

عنوان: به‌کارگیری پاراداده در آمارگیری نیروی کار ایران

مجری: لیدا کلهری ندرآبادی

همکاران: محمدرضا ریحانی

مشاور: ندا مصطفوی

گروه پژوهشی: طرح‌های فنی و روش‌های آماری

تاریخ انتشار: پاییز ۱۳۹۹

نوبت انتشار: اول

طراح جلد: علیرضا رنجبر

صفحه‌آرا: طاهره امینی

❖ حق مالکیت معنوی این طرح پژوهشی متعلق به پژوهشکده‌ی آمار است و نقل مطالب فقط با ذکر مأخذ مجاز است.



پژوهشکده‌ی آمار



مرکز آمار ایران

تهران، خیابان دکتر فاطمی، خیابان باباطاهر، خیابان سرتیپ فکوری، شماره‌ی

۱۴۵

۰۲۱ ۸۸۶۳۰۴۴۰ -۳

www.srtc.ac.ir



به نام خداوند جان و خرد

## پیش‌گفتار

همزمان با افزایش هزینه‌های آمارگیری، نظارت بر جمع‌آوری داده‌ها و بهره‌وری آمارگیری و در عین حال، سعی در حفظ کیفیت داده‌ها در قالب یک بودجه معقول، اهمیت بیشتری پیدا می‌کند. این موضوع، به ویژه برای مصاحبه‌کنندگان، وضعیت‌هایی را ایجاد می‌کند که در آن منابع کمیاب می‌توانند در جای دیگری در فرایند تولید به کار گرفته شوند. یک طرح واکنش‌گرا، می‌تواند برای مدیران عملیاتی، دقیقاً همان اطلاعاتی را فراهم کند که آن‌ها در هنگام جمع‌آوری مستمر داده‌ها برای کارآمدتر کردن استفاده از منابع، به آن نیاز دارند.

این مطالعه با هدف به‌کارگیری پاراداده در آمارگیری نیروی کار ایران به‌عنوان یکی از مهم‌ترین آمارگیری‌های مرکز آمار ایران در دستور کار پژوهشکده‌ی آمار قرار گرفت. در این مطالعه رایج‌ترین انواع پاراداده شامل داده‌های ثبت اطلاعات تماس، مشخصات آمارگیر، تعامل و اقدامات متقابل در بدو آمارگیری، دلایل عدم همکاری، مانند مسیرهای ممیزی و فایل‌های خلاصه وقایع، نظارت صامت، ارزیابی آمارگیر از کیفیت داده‌ها، ارزیابی پاسخ‌گو از کیفیت پرسش‌نامه و اطلاعات تماس با پاسخگو معرفی شد و به کاربردهای پاراداده به صورت کلی اشاره شد. علاوه بر آن نتایج مطالعات انجام شده در خصوص بررسی تمایلات پاسخگویی در طول دوره‌ی جمع‌آوری داده‌ها و همچنین بررسی نحوه تغییر توزیع این تمایلات پس از هر تلاش برای برقراری تماس ارائه شده است. این کار، یک روش پیش‌بینی را برای شناسایی گروه‌هایی که انجام مصاحبه با آن‌ها مشکل است، در زمانی که جمع‌آوری داده‌ها هنوز ادامه دارد، ارائه می‌کند. همچنین کاربرد پاراداده در زمینه بررسی رصد رفتار مصاحبه‌کننده در چند آمارگیری بین‌المللی و چندملیتی ارائه شده است.

در این طرح پژوهشی به‌ویژه پاراداده‌ی موجود در طرح نیروی کار ایران مورد بررسی قرار گرفته است. بررسی ارتباط میان ویژگی‌های آمارگیر و نرخ بی‌پاسخی به ازای آمارگیر با استفاده از مدل رگرسیون بتاافزوده، زمان تکمیل پرسش‌نامه و اتاق کنترل از مواردی است که به تفصیل مورد بررسی قرار گرفته‌اند. بر این اساس در فصل آخر نیز به برخی موارد پاراداده که گردآوری آن برای نظارت بر آمارگیری نیروی کار قابل استفاده است اشاره شده است. بدیهی است گردآوری پاراداده ممکن است نیازمند تغییرات در روش نمونه‌گیری یا سیستم‌های مدیریت میدانی باشد. با توجه به تغییر شیوه آمارگیری تأمین امکانات مورد نیاز برای ثبت رکوردهای تماس‌های الکترونیک از اهمیت ویژه‌ای برخوردار است که نیازمند به سرمایه‌گذاری در سیستم‌های موجود یا توسعه یک سیستم جدید برای حفظ اطلاعات تماس است.

این طرح پژوهشی به سفارش مرکز آمار ایران و در گروه پژوهشی طرح‌های فنی و روش‌های آماری پژوهشکده‌ی آمار با همکاری خانم دکتر لیدا کلهری به‌عنوان مجری طرح، آقای دکتر فرهاد مهران به‌عنوان مشاور طرح، آقای محمدرضا ریحانی و خانم ندا مصطفوی به‌عنوان همکاران اصلی طرح پژوهشی به انجام رسیده است، که بدین وسیله از ایشان صمیمانه تشکر و قدردانی می‌شود.

گروه پژوهشی طرح‌های فنی و روش‌های آماری

پژوهشکده‌ی آمار

# فهرست مطالب

مقدمه	۱
۱-۱- مقدمه	۱
۲-۱- انواع پاراداده	۲
۱-۲-۱- مشخصات آمارگیر	۲
۲-۲-۱- داده‌های ثبت اطلاعات تماس	۳
۳-۲-۱- مشاهدات آمارگیر درباره ناحیه و محل سکونت و مشخصات افراد در واحد نمونه‌گیری	۳
۴-۲-۱- تعامل و اقدامات متقابل در بدو آمارگیری	۴
۵-۲-۱- مسیرهای ممیزی	۴
۶-۲-۱- ضبط صدا	۴
۳-۱- برخی کاربردهای پاراداده	۵
۴-۱- معرفی ابزار تاریخچه‌ی تماس CHI	۷
۱-۴-۱- معرفی CHI	۷
۲-۴-۱- فرایند CHI	۷
۵-۱- ارزیابی کیفیت پاراداده‌ی تاریخچه‌ی تماس در تعدادی از آمارگیری‌های آمریکا	۹
۱-۵-۱- پیشینه	۱۰
۲-۵-۱- شرح تلاش‌های برقراری تماس	۱۲
۳-۵-۱- ویژگی‌های تلاش برای برقراری تماس	۱۳
۴-۵-۱- نماگرهای کیفیت داده‌ی CHI	۱۶
۵-۵-۱- گزارش اشتباه در CHI	۲۰
۶-۱- خلاصه و نتیجه‌گیری	۲۱
پاراداده در آمارگیری نیروی کار نروژ	۲۵
۱-۲- مقدمه	۲۵
۲-۲- طرح واکنش‌گرا- یک حالت از طرح آمارگیری سازوار	۲۷
۳-۲- ساختار داده	۲۸
۴-۲- تعداد تلاش برای برقراری تماس	۲۹

- ۲-۵- تمایلات پاسخگویی ..... ۳۱
- ۲-۶- بررسی زمان‌بندی برای برقراری اولین تماس ..... ۳۳
- ۲-۶-۱- اولین تلاش برای برقراری تماس در افراد موج اول ..... ۳۴
- ۲-۶-۲- بهترین روز برای برقراری تماس ..... ۳۵
- ۲-۶-۳- دومین تلاش برای برقراری تماس در افراد موج اول ..... ۳۶
- ۲-۷- ظرفیت فاز و تعویض مد در LFS نروژ ..... ۳۶
- ۲-۷-۱- معرف بودن و شاخص‌های R ..... ۳۷
- ۲-۷-۲- مدل محاسبه شاخص‌های R ..... ۳۸
- ۲-۸- پیشنهاداتی برای یک طرح واکنش‌گرا ..... ۴۰
- ۲-۹- بحث و نتیجه‌گیری ..... ۴۱
- ۲-۴۳- مطالعات موردی در زمینه رصد رفتار مصاحبه‌کننده در آمارگیری‌های بین‌المللی و چندملیتی ..... ۴۳
- ۳-۱- مقدمه ..... ۴۳
- ۳-۲- آمارگیری هرم‌های مصرف‌کننده در هند ..... ۴۶
- ۳-۳- آمارگیری ملی سلامت روان عربستان ..... ۴۹
- ۳-۴- آمارگیری اجتماعی اروپا ..... ۵۶
- ۳-۵- آمارگیری سلامتی، سالمندی و بازنشستگی اروپا ..... ۶۲
- بررسی پاراداده‌ی نیروی کار ایران ..... ۶۷
- ۴-۱- مقدمه ..... ۶۷
- ۴-۲- فرایند فعلی اتاق کنترل ..... ۶۸
- ۴-۲-۱- روش استفاده از اتاق کنترل ..... ۶۹
- ۴-۳- بررسی اطلاعات مدت زمان تکمیل پرسش‌نامه‌ها در طرح نیروی کار ..... ۷۰
- ۴-۴- بررسی رابطه بی‌پاسخی و ویژگی‌های خانوار ..... ۷۳
- ۴-۴-۱- تعیین مشخصات خانوار بی‌پاسخ بر اساس اطلاعات نوبت‌های قبلی در آمارگیری نیروی کار ایران ..... ۷۶
- ۴-۵- بررسی رابطه ویژگی‌های مأمور آمارگیر و نرخ پاسخگویی ..... ۷۸
- ۴-۵-۱- نتایج تحلیل توصیفی ..... ۷۹
- ۴-۵-۲- مدل رگرسیون آمیخته خطی بتا افزوده برای مدل‌بندی نرخ بی‌پاسخی مأمور آمارگیر ..... ۸۱
- ۴-۶- بحث و نتیجه‌گیری ..... ۸۸
- پیشنهاد استفاده از پاراداده برای نظارت بر آمارگیری نیروی کار ..... ۹۱
- ۵-۱- مقدمه ..... ۹۱
- ۵-۲- پاراداده‌ی مورد نیاز برای نظارت بر آمارگیری نیروی کار ..... ۹۲
- ۵-۳- بحث و نتیجه‌گیری ..... ۹۸
- مرجع‌ها ..... ۱۰۱
- پیوست‌ها ..... ۱۰۹
- پیوست آ- شمائی از صفحات و فلوچارت فرایند کار CHI ..... ۱۱۰



- پیوست ب- متغیرهای بررسی شده برای زمان بندی اولین تماس..... ۱۱۶
- پیوست پ- متغیرهای اضافی بررسی شده برای زمان بندی دومین تلاش برای برقراری تماس..... ۱۱۷
- پیوست ت- متغیرهای اضافی بررسی شده برای تمایلات و شاخص های پاسخگویی..... ۱۱۸
- واژه نامه..... ۱۱۹



# فهرست جدول‌ها

- جدول ۱-۲-۱- متوسط تعداد تلاش‌ها تا به سرانجام رساندن یک مورد به تفکیک فصل در سال ۲۰۱۳ ..... ۳۱
- جدول ۲-۲-۲- تعداد افرادی که بر اساس اطلاعات فصل‌های دوم، سوم و چهارم سال ۲۰۱۳، برای اولین بار با آن‌ها تماس گرفته شده است و مقادیر آماره آزمون توکی به تفکیک بازه‌های زمانی ..... ۳۴
- جدول ۲-۳- میزان پاسخگویی، تعداد افرادی که مورد تماس قرار گرفته‌اند و درصد مصاحبه‌هایی که به صورت دستی شروع شده ..... ۳۵
- جدول ۱-۳-۱- شاخص‌های کنترل کیفیت در آمارگیری ملی سلامت روان عربستان به تفکیک منابع خطا ..... ۵۴
- جدول ۲-۳-۲- مثال از کاربرگی که به صورت خلاصه وضعیت علامت‌گذاری را (۱ علامت‌گذاری شده، ۰ بدون علامت) به تفکیک هر آمارگیر برای شاخص‌های منفرد نشان می‌دهد ..... ۵۴
- جدول ۳-۳-۳- میانگین همبستگی بین مصاحبه‌گرها در ۴۸ آمارگیری برای ۳۶ کشور در ۶ دوره اجرای آمارگیری اجتماعی اروپا ..... ۶۱
- جدول ۳-۴- گزارش کنترل کیفیت عملکرد آمارگیر ..... ۶۵
- جدول ۱-۴-۱- فایل ۱ اتاق کنترل- گزارش وضعیت ..... ۶۸
- جدول ۲-۴-۲- فایل ۲ اتاق کنترل: حجم کار روزانه، جمعیت‌ها و نرخ‌های اصلی نیروی کار به تفکیک استان‌ها ..... ۶۹
- جدول ۳-۴-۳- وضعیت فایل مدت زمان تکمیل در فصل بهار ۱۳۹۸ ..... ۷۱
- جدول ۴-۴-۴- برآورد زمان مورد نیاز برای تکمیل پرسشنامه با توجه به وضع فعالیت افراد ..... ۷۲
- جدول ۵-۴-۵- متغیرهای کمکی مرتبط با همکاری پیمایشی در پژوهش‌های قبلی ..... ۷۵
- جدول ۶-۴-۶- نرخ بی‌پاسخی خانوارها بر اساس بعد خانوار ..... ۷۷
- جدول ۷-۴-۷- چارک‌های سن، تجربه کار (فصل) و حجم کار مأموران آمارگیر ..... ۸۰
- جدول ۸-۴-۸- متوسط نرخ بی‌پاسخی به تفکیک جنس، سطح تحصیلات، تجربه کار، سن و حجم کار آمارگیر ..... ۸۰
- جدول ۱-۵- برنامه‌ریزی مصاحبه مجدد از لحاظ مسئول، حجم مصاحبه مجدد، زمان و اقلام مورد بررسی ..... ۹۵
- جدول ۲-۵-۲- شاخص‌های منفرد و تجمعی پیشنهادی برای ارزیابی آمارگیر ..... ۹۷
- جدول ۳-۵-۳- فرم پیشنهادی ارزیابی آمارگیر ..... ۹۸



# فهرست شکل‌ها

- شکل ۱-۲- تعداد تلاش‌های انجام گرفته برای قراری تماس قبیل از این‌که یک مورد به سرانجام رسیده باشد به تفکیک فصل‌های دوم تا چهارم سال ۲۰۱۳ در طرح LFS نروژ..... ۳۰
- شکل ۲-۲- تمایلات پاسخگویی مربوط به ۵ تلاش اول برای برقراری تماس در فصل‌های سوم و چهارم ۲۰۱۳..... ۳۲
- شکل ۳-۲- تمایلات پاسخگویی مربوط به ۱۵ تلاش اول برای برقراری تماس در فصل‌های سوم و چهارم ۲۰۱۳..... ۳۳
- شکل ۴-۲- شاخص‌های R برای فصل‌های سوم و چهارم سال ۲۰۱۳. برآورد برای نرخ اشتغال بدون وزن‌دهی..... ۳۸
- شکل ۵-۲- شاخص‌های جزئی R برای فصل‌های سوم و چهارم سال ۲۰۱۳..... ۳۹
- شکل ۱-۳- ارتباطات بین دفتر مرکزی، مرکز بازبینی، تیم نظارت، شعب عملیات میدانی و تیم‌های عملیات میدانی. خانوار با HHH دفتر مرکزی با HO، مرکز توسعه منابع انسانی با HDR و مرکز اعتبارسنجی همزمان آفلاین با ORV .. ۴۹
- شکل ۱-۴- نمودار جعبه‌ای نرخ بی‌پاسخی به تفکیک سن، جنس و سطح تحصیلات بر اساس اطلاعات مأموران آمارگیر در همه استان‌های مورد مطالعه ..... ۸۱
- شکل ۲-۴- نمودار اثر نمونه‌های تولیدشده پارامترهای مدل ..... ۸۵
- شکل ۳-۴- نمودار خودهمبستگی نمونه‌های تولیدشده پارامترهای مدل ..... ۸۶





## مقدمه

### ۱-۱- مقدمه

کیفیت داده‌های به‌دست آمده از آمارگیری‌ها و روش‌های بهبود کیفیت از موضوعات پرچالشی هستند که پرداختن به آن‌ها از اهمیت بالایی برخوردار است. از این رو آمارشناسان همواره به دنبال روش‌ها و ابزارهایی هستند که با استفاده از آن‌ها بتوان خطاهای آمارگیری را کاهش و کیفیت نتایج را افزایش داد. پاراداده<sup>۱</sup> از ابزارهایی است که برای کشف منابع خطا در آمارگیری‌ها به کار می‌رود. کوپر (۱۹۹۸) اولین شخصی بود که اصطلاح «پاراداده» را در حوزه روش‌شناسی آمارگیری معرفی کرد. او ابتدا این اصطلاح را برای اشاره به داده‌های فرایندهای خودکار، ارائه کرد. از آن به بعد، این اصطلاح به همه انواع داده‌های مرتبط با فرایند گردآوری داده‌ها مانند ضبط مکالمه آمارگیر، مدت زمان مصاحبه و مشخصات آمارگیر توسعه یافت. کروتز (۲۰۱۳) بیان می‌کند پاراداده، داده‌های بیشتری است که در طول فرایند اجرای آمارگیری‌ها می‌توان گردآوری کرد. او همچنین پاراداده را نوعی داده کمکی معرفی می‌کند که از منابع داخلی آمارگیری به‌دست می‌آید. گردآوری پاراداده‌ی مربوط به آمارگیری، موضوع جدیدی نیست اما دامنه و جزئیات پاراداده‌یی که گردآوری می‌شود، در سال‌های اخیر به‌طور اساسی افزایش یافته است، که البته بیشتر به دلیل کامپیوتری شدن تمام فرایند گردآوری داده در آمارگیری است و تنها به بخش گردآوری داده‌های پرسشنامه محدود نمی‌شود. به‌طور کلی، آگاهی در مورد فواید بالقوه استفاده از پاراداده برای از میان برداشتن برخی چالش‌ها از قبیل، کاهش نرخ پاسخگویی، افزایش مخاطره اریبی ناشی از بی‌پاسخی و خطای اندازه‌گیری و هزینه‌های فزاینده گردآوری داده، در میان تولیدکنندگان و استفاده‌کنندگان داده رو به افزایش است. پاراداده قابلیت گردآوری در هر نوع آمارگیری را دارد، گرچه مقدار، نوع و سطح جزئیاتی که با پاراداده به‌دست می‌آید بسته به این که داده‌ها به‌صورت دستی یا به‌صورت خودکار با استفاده از نرم‌افزارهای

---

<sup>۱</sup> Paradata

کامپیوتری ثبت شده باشند، متفاوت است. در سال‌های اخیر برای ارزیابی کیفیت داده‌های گردآوری شده در بررسی‌های روش‌شناختی آمارگیری به‌طور فزاینده‌ای از پاراداده استفاده شده است.

طرح پژوهشی «پاراداده و کاربرد آن در بهبود نتایج آمارگیری‌ها» توسط کلهری و همکاران (۱۳۹۵) انجام شد در این مطالعه متداول‌ترین انواع پاراداده شامل داده‌های ثبت اطلاعات تماس، مشخصات آمارگیر، تعامل و اقدامات متقابل در بدو آمارگیری، دلایل عدم همکاری، مسیرهای ممیزی و فایل‌های خلاصه وقایع، نظارت صامت، ارزیابی آمارگیر از کیفیت داده‌ها، ارزیابی پاسخ‌گو از کیفیت پرسشنامه و اطلاعات تماس با پاسخگو معرفی و به کاربردهای آن اشاره شده است. در این فصل، ابتدا برخی از انواع پاراداده به اختصار مرور می‌شوند برای مطالعه بیشتر می‌توانید به کلهری و همکاران (۱۳۹۵) مراجعه کنید. باید توجه شود که امکان گردآوری همه‌ی انواع پاراداده در هر آمارگیری وجود ندارد، بلکه نوع آمارگیری در تعیین پاراداده‌ی قابل حصول، مؤثر است. به‌عنوان مثال، مشاهدات آمارگیر درباره ناحیه، مربوط به آمارگیری‌هایی است که با مراجعه حضوری انجام می‌شوند و در آمارگیری‌های تلفنی یا اینترنتی امکان دستیابی به آن‌ها فراهم نیست. در ادامه‌ی فصل برخی کارکردهای پاراداده و نحوه استفاده از پاراداده در اداره سرشماری امریکا بیان می‌شود. معیارهای ارزیابی کیفیت پاراداده گسترش نیافته است و در مطالعه مربوط به اداره سرشماری امریکا، پاراداده‌ی به‌دست آمده از سه آمارگیری به‌منظور ارزیابی کیفیت مورد بررسی قرار می‌گیرند. در فصل دوم از مطالعه پیش‌رو نحوه به‌کارگیری پاراداده در طرح نیروی کار نروژ بیان می‌شود. در فصل سوم نحوه استفاده از داده و پاراداده به‌منظور نظارت بر فرایند آمارگیری و کنترل کیفیت آن در چهار آمارگیری که عبارت‌اند از آمارگیری هرم‌های مصرف‌کننده در هند، آمارگیری ملی سلامت روان عربستان، آمارگیری اجتماعی اروپا و آمارگیری سلامتی، سالمندی و بازنشستگی اروپا ارائه می‌شود. در فصل چهارم به مطالعه پاراداده‌یی که درباره اطلاعات تماس و ویژگی‌های مأمور آمارگیر در طرح آمارگیری نیروی کار ایران به‌دست می‌آیند، خواهیم پرداخت. بررسی‌های فصل چهارم نشان می‌دهد که علی‌رغم استفاده از تبلت در جمع‌آوری داده‌های طرح نیروی کار، پاراداده‌ی کاملی از تبلت به‌دست نمی‌آید. بنا بر این فصل پنجم به ارائه راهنموده‌هایی برای تولید پاراداده‌ی مفید با هدف نظارت بر فرایند آمارگیری و کنترل و مدیریت آن ارائه شده است.

## ۱-۲- انواع پاراداده

در این بخش برخی از انواع پاراداده شامل مشخصات آمارگیر، داده‌های ثبت اطلاعات تماس، مشاهدات آمارگیر درباره ناحیه، محل سکونت و مشخصات افراد در واحد نمونه‌گیری، تعامل و اقدامات متقابل در شروع آمارگیری مرور می‌شوند.

### ۱-۲-۱- مشخصات آمارگیر

در آمارگیری‌هایی که اطلاعات توسط آمارگیر جمع‌آوری می‌شود، آمارگیر نقش مهمی در افزایش میزان پاسخ به آمارگیری و کاهش خطای اندازه‌گیری دارد. مطالعات گذشته نشان داده‌اند که نرخ پاسخگویی قویاً با نگرش مثبت آمارگیر نسبت به متقاعد ساختن پاسخگو همبستگی دارد (لهتونن ۱۹۹۶، و گروز و کوپر ۱۹۹۸). همچنین بخشی از تفاوت بین پاسخ ارائه‌شده توسط پاسخگو و آنچه پژوهشگر واقعاً به دنبال آن است، می‌تواند



در اثر خصوصیات فردی آمارگیر ایجاد شود. مهارت آمارگیران در مدیریت مصاحبه با یکدیگر تفاوت دارد، از طرف دیگر مهارت آمارگیران در درک مفهوم سؤالات و انتقال صحیح آن به پاسخ‌گویان یکسان نیست. علاوه بر آن، برخی از ویژگی‌های فردی آمارگیر مانند جنس، سن، نژاد یا زبان در آمارگیری‌هایی که سؤالات حساس دارند، در کسب اطلاعات در برخی مناطق با فرهنگ یا زبان خاص اثر دارد. بنا بر این کسب برخی اطلاعات در مورد ویژگی‌های آمارگیر و تحلیل آن به فهم بهتر عوامل مؤثر بر این تغییرات در بین آمارگیران کمک می‌کند، که می‌تواند برای بهبود فرایند جذب، آموزش و نظارت بر آمارگیران مفید باشد.

### ۱-۲-۲- داده‌های ثبت اطلاعات تماس

داده‌های ثبت اطلاعات تماس درباره مراجعه یا تماس با خانوار است که می‌تواند نشان‌دهنده تلاش انجام‌شده توسط آمارگیر برای تکمیل پرسشنامه واحد نمونه‌ای باشد. این اطلاعات شامل زمان مراجعه، تعداد مراجعات، نتیجه مراجعات و نام و نام خانوادگی یا کد شناسایی آمارگیر یا آمارگیرانی است که به خانوار مراجعه کرده‌اند. هدف اصلی از ثبت اطلاعات تماس، رصد کردن نحوه تصمیم‌گیری‌های آمارگیر در عملیات میدانی و تأثیر آن بر میزان مشارکت افراد است که برای مدیریت آمارگیری‌ها با هدف افزایش نرخ پاسخ‌گویی مورد استفاده قرار می‌گیرد. به‌عنوان مثال، نشان می‌دهد که در چه زمانی از روز افراد تمایل بیشتری به پاسخ‌گویی دارند، که ممکن است با توجه به نوع آمارگیری و پاسخ‌گویان هدف، متفاوت باشد. ثبت اطلاعات تماس برای تمام واحدهای نمونه، اعم از پاسخگو و بی‌پاسخ، امکان‌پذیر است و اطلاعات تماس مربوط به واحدهای بی‌پاسخ برای تحلیل بی‌پاسخی مفیدند، بنا بر این اطلاعات آن‌ها به اندازه اطلاعات پاسخگویان ارزشمند است.

### ۱-۲-۳- مشاهدات آمارگیر درباره ناحیه و محل سکونت و مشخصات افراد در واحد نمونه‌گیری

مشاهدات ناحیه برای ارزیابی تأثیر شرایط محله بر پیامدهای مهم وضعیت اجتماعی و اقتصادی گردآوری می‌شوند. در دسترس بودن شواهد برای تمام واحدهای نمونه، فرصت را برای کشف اینکه آیا موارد بی‌پاسخ به‌صورت معنی‌داری با موارد پاسخگو متفاوت هستند، به وجود می‌آورد. برای مؤثر بودن این کار، مشاهدات باید قویاً با احتمال پاسخگویی واحد نمونه و متغیرهای مورد نظر آمارگیری مرتبط باشند.

مشاهدات واحد مسکونی معمولاً شامل ارزیابی کیفیت واحد مسکونی در مقایسه با سایر واحدها در یک محله است. به‌عنوان مثال این‌که، واحد مسکونی در یک ساختمان چند واحدی قرار دارد، آیا موانع دسترسی وجود دارد و این‌که شرایط آن ایمن است یا خیر. شرایط واحد مسکونی در مقایسه با سایر واحدها در منطقه، به‌طور قابل توجهی با نرخ غیبت خانوار و امتناع از پاسخگویی مرتبط است (دورانت و استیل، ۲۰۰۹).

مشاهدات آمارگیر از اعضای ساکن در واحدهای مسکونی منتخب، به دو رده تقسیم می‌شوند: مشاهدات در مورد ویژگی‌های جمعیتی افراد و مشاهداتی که نماینده متغیرهای کلیدی آمارگیری هستند. خصوصیات جمعیت‌شناسی که در آمارگیری‌ها ثبت می‌شوند، اطلاعاتی در خصوص افراد و زیر مجموعه‌هایی که حضورشان در نمونه اهمیت دارد فراهم می‌کنند. به‌عنوان مثال، گروه سنی، نژاد یا زبان افراد می‌تواند بر پاسخگویی آن‌ها

اثرگذار باشد. باید توجه داشت که شناسایی شواهدی که به‌صورت تنگاتنگ با شانس پاسخگویی و متغیرهای مورد نظر آمارگیری مرتبط باشند و همچنین اندازه‌گیری آن‌ها حائز اهمیت است.

### ۱-۲-۴- تعامل و اقدامات متقابل در بدو آمارگیری

گام بسیار مهم در جلب همکاری اعضای نمونه، تعامل متقابل بین آمارگیر و عضو نمونه کاندید پاسخگویی است. تا به حال تلاش‌هایی برای گردآوری اطلاعات خلاصه‌ای از این تعامل‌ها صورت گرفته است. این اطلاعات، برای ایجاد تمایز بین مشخصه‌های واحدهایی که در آمارگیری شرکت کرده‌اند و نمونه‌هایی که در نهایت در آمارگیری شرکت نکرده‌اند، و برای پیش‌گویی اینکه کدام مخاطب از بین واحدهای نمونه حتی با پیگیری‌های بیشتر شرکت در آمارگیری را رد خواهد کرد، به‌کار گرفته می‌شود. در حال حاضر به‌عنوان مثال در کشور انگلستان در تعداد کمی از مطالعات، داده‌هایی درباره تعامل آمارگیر با واحد نمونه گردآوری می‌شود. آمارگیران در این آمارگیری‌ها ارزیابی خود درباره میزان مقاومت اولیه پاسخ‌گویان در همکاری را تحت عناوینی مانند هیچ، خیلی کم، متوسط، قوی، ارائه می‌کنند.

### ۱-۲-۵- مسیرهای ممیزی

فایل‌های مسیرهای ممیزی حاوی اطلاعاتی هستند که با استفاده از آن‌ها فرایند جمع‌آوری داده‌ها قابل تمیز دادن است. در حال حاضر، در برخی از آمارگیری‌های بزرگ از سیستم‌های نرم‌افزاری مصاحبه همراه با کامپیوتر (CAI<sup>۲</sup>) استفاده می‌شود؛ که این سیستم در کنار ثبت داده‌های پرسشنامه، دامنه گسترده‌ای پاراداده هم تولید می‌کند. این فایل‌ها طیف وسیعی از اطلاعات را راجع به تمام اقلام نمونه‌ای فراهم می‌کنند که می‌تواند ابزارهای تحقیق، عملکرد آمارگیر و کیفیت داده را ارزیابی کند. به‌عنوان مثال، اگر در یک حوزه آمارگیری بیشتر از حد معمول از دکمه Help استفاده شود، یعنی مشکلی در رابطه با آن سؤال یا آموزش آن وجود دارد.

### ۱-۲-۶- ضبط صدا

سیستم‌های نرم‌افزاری پیشرفته CAI همچنین امکان ضبط صدا در طول مصاحبه را فراهم می‌کنند که به آن مصاحبه ضبط‌شده به وسیله کامپیوتر (CARI<sup>۳</sup>) اطلاق می‌شود. با رضایت مخاطبین، می‌توان کل مصاحبه، بخشی از مصاحبه یا یک بخش تصادفی از مصاحبه را ضبط کرد. در آمارگیری‌های تلفنی، سرپرستان آمارگیری می‌توانند به‌صورت زنده مصاحبه آمارگیر را گوش کنند (نظارت صامت<sup>۴</sup>). CARI این فرصت را برای مدیران و روش‌شناسان آمارگیری فراهم می‌کند که به تعامل بین آمارگیر و مخاطب گوش دهند و اطلاعات مورد نیاز برای ارزیابی ابزارها، آمارگیر و کیفیت داده‌ها را به‌دست آورند.

<sup>۲</sup> Computer Assisted Interview

<sup>۳</sup> Computer Assisted Recorded Interview

<sup>۴</sup> Silent Monitoring

### ۱-۳- برخی کاربردهای پاراداده

پاراداده می‌تواند توسط روش‌شناسان آمارگیری در بحث قابلیت اطمینان، پاسخگو بودن به سؤال‌ها و قابل تعمیم بودن دانش راه و روش آمارگیری استفاده شود. هدف کلی استفاده از این دانش بهبود فرایند آمارگیری و کمینه کردن خطای آمارگیری است. در ابتدا تمرکز اصلی پژوهش‌های روش‌شناسی که از پاراداده بهره می‌بردند، بر روی خطای بی‌پاسخی بوده است. برای مثال، تاریخچه طولانی از تحلیل ثبت اطلاعات تماس‌ها برای دستیابی به دانش الگوی مناسب تماس در آمارگیری‌های چهره به چهره وجود دارد (دورانت و استیل ۲۰۰۹، کامپانلی و دیگران ۱۹۹۷، سویرس و درک ۱۹۹۲). روش‌شناسان آمارگیری همچنین به ارزش استفاده از پاراداده برای مطالعه اریبی ناشی از بی‌پاسخی پی‌برده‌اند. به‌ویژه پاراداده‌یی که برای همه موارد نمونه‌گیری شده در دسترس باشند و با مشارکت پاسخ‌گویان و متغیرهای مورد نظر آمارگیری، همبستگی داشته باشد. پاراداده‌ی حاصل از ثبت اطلاعات تماس‌ها که معمولاً شامل زمان تماس (روز و زمان) و نیز نتیجه تماس (عدم برقراری تماس، امتناع از پاسخگویی، غیر واجد شرایط بودن، تکمیل مصاحبه، تعیین وقت برای تماس بعدی و غیره) هستند، اغلب برای ارزیابی و کاهش خطای بی‌پاسخی مورد استفاده قرار می‌گیرند (هیروه و دیگران، ۲۰۰۷، بلوم و دیگران، ۲۰۰۸). داده‌های ثبت اطلاعات تماس برای پاسخ‌گویان و غیر پاسخ‌گویان در یک آمارگیری موجود هستند و در نتیجه، کاندیداهای اولیه برای مطالعه اریبی بی‌پاسخی به شمار می‌روند.

روش‌شناسان آمارگیری، از پاراداده برای کاوش در مورد منابع ایجاد خطای اندازه‌گیری نیز استفاده می‌کنند (باسیلی و اسکات ۱۹۹۶، کسپر و کوپر ۱۹۹۷، هیروی ۲۰۰۳، درایسما و دیجکسترا ۲۰۰۴، مک گی ۲۰۰۷، یان و تورنگیو ۲۰۰۸). برای مثال، مدت زمان تکمیل پرسشنامه، فایل مسیرهای ممیزی و ضبط صداها می‌تواند مشکلات پاسخگو در هنگام پاسخ به سؤالات را مشخص کند. پاراداده همچنین برای آزمودن اثر شخصیت آمارگیر در اندازه‌گیری، مورد استفاده قرار می‌گیرد. علاوه بر خطای بی‌پاسخی و خطای اندازه‌گیری در صورت امکان این پتانسیل وجود دارد که از پاراداده برای ارزیابی سایر منابع خطا از جمله خطای پردازش نیز استفاده شود.

پاراداده علاوه بر آن که برای کاهش خطای آمارگیری از طریق درک بهتر دلایل بروز خطاها و بهبود مدیریت در گردآوری داده‌ها قابل استفاده است، بعد از گردآوری داده‌ها برای تعدیل خطای برآوردهای آمارگیری نیز به کار می‌آید.

تعدیل بی‌پاسخی برای داده‌های حاصل از اغلب آمارگیری‌هایی که در مقیاس بزرگ اجرا می‌شوند، انجام می‌شود. به‌طور کلی، وزن‌های نمونه‌گیری اغلب بر اساس اطلاعات حاصل از داده‌های کمکی، محاسبه می‌شوند. به‌عنوان مثال، اطلاعات چارچوب نمونه‌گیری، داده‌های سرشماری یا منابع جمعیت‌شناسی معتبر، که منجر به تعدیل اختلاف مشخصه‌های جمعیت‌شناسی بین جمعیت مورد نظر در کل جامعه و نمونه می‌شود، مورد استفاده قرار می‌گیرند. داده‌های آمارگیری اغلب برای تعدیل اثرات بی‌پاسخی، وزن‌دهی مجدد می‌شوند. با این وجود، اگر در تعدیل وزن برای بی‌پاسخی مشخصه‌های جمعیت‌شناسی به کار گرفته شده برای تعدیل وزن با متغیرهای مورد نظر در آمارگیری و میزان تمایل به پاسخگویی مرتبط نباشند، تعدیل وزن اریبی ناشی از بی‌پاسخی را کاهش نمی‌دهد.

با توجه به این محدودیت، در حال حاضر توجه آمارشناسان به سمت استفاده از پاراداده برای تعدیل خطای بی‌پاسخی جلب شده است. شواهد نیز حاکی از آن است که پاراداده می‌تواند نقش مهمی در تعدیل خطای بی‌پاسخی داشته باشد (بلوم ۲۰۰۹). اما چالشی که هم‌اکنون روش‌شناسان آمارگیری با آن رو به رو هستند، شناسایی اقلامی از پاراداده است که با تمایل به مشارکت و متغیرهای مورد نظر آمارگیری، همبستگی داشته باشد. این امر ممکن است نیازمند شناسایی دامنه‌ای از اقلام پاراداده‌ی مورد استفاده برای تعدیل‌های بی‌پاسخی باشد و در نظر گرفتن یک تعدیل برای همه برآوردها امکان‌پذیر نباشد.

به‌کارگیری پاراداده در تعدیل خطای اندازه‌گیری، کمتر از کاربرد آن در تعدیل خطای بی‌پاسخی است. با این حال، شواهدی مبنی بر اینکه پاراداده در سطح اقلام<sup>۵</sup> می‌تواند اطلاعات مفیدی را در مورد فرایند پرسش و پاسخ فراهم کند، وجود دارد. این امید وجود دارد که پاراداده‌ی در سطح قلم به همان خوبی که در مدل‌بندی آماری برای تحلیل داده‌ها مورد استفاده قرار گرفته است بتواند برای تعدیل خطای اندازه‌گیری هم استفاده شود. برای دستیابی به این امر و درک بهتر ارتباط بین پاراداده و خطای اندازه‌گیری لازم است مطالعات بیشتری انجام شود.

همچنین تشخیص داده شده است که کیفیت خوب آمارگیری بر حسب نرخ خطای کوچک برای منابع ویژه خطا، کوچک بودن میانگین مربعات خطا برای برآوردهای منتخب یا سایر ویژگی‌های محصول آمارگیری، به طراحی فرایندهای اصلی که این محصولات را تولید می‌کنند و نیز کنترل این فرایندها، بستگی دارد. سازمان‌های آماری همواره دغدغه کیفیت دارند اما در بسیاری از آن‌ها یک نمای واضح که ارتباط بین فرایند و محصول آن را بیان کند، وجود ندارد. با کنترل و احتمالاً تنظیم این فرایندها، اهداف آمارگیری آسان‌تر قابل حصول خواهند بود و در غیر این صورت برای دستیابی به کیفیت مورد انتظار لازم است با صرف هزینه قابل توجهی فرایندها مجدداً تکرار شوند. این رویکرد برای سنجش کیفیت و بهبود مداوم آن است چرا که تحلیل هر فرایند می‌تواند دلایل ریشه‌ای مشکلاتی را تشخیص دهد که با حذف آن‌ها، فرایند در ویژگی‌های محصول آمارگیری تغییر ایجاد کند. بنا بر این پاراداده در کنترل فرایند آمارگیری نیز قابل استفاده است. از منظر دیگر پاراداده برای پایش و مدیریت آمارگیری‌های در حال انجام نیز قابل استفاده است. در روش‌های بهبود مستمر کیفیت (ایم‌ای، ۱۹۸۶، دمینگ، ۱۹۸۶)، اندازه‌گیری‌ها در مورد یک فرایند طی زمان اجرای آن انجام می‌شود. با الهام گرفتن از این روش‌ها، پژوهشگران استفاده از چنین استراتژی‌هایی برای فرایند عملیات آمارگیری را پیشنهاد می‌کنند. بنا بر این می‌توان منابع خطا را همزمان با اجرای فرایند شناسایی کرد و مداخلات لازم را انجام داد (بیمر و کاسپار، ۱۹۹۴، مورگانشتاین و مارکر، ۱۹۹۷). همانطور که بحث شد، پاراداده می‌تواند نقش مهمی را در به‌کارگیری این نوع استراتژی‌ها، ایفا کند. در ادامه، ابزار گردآوری و تحلیل اطلاعات در رابطه با برقراری تماس با واحدهای نمونه و نحوه ارزیابی کیفیت پاراداده‌ی تاریخچه‌ی تماس در تعدادی از طرح‌های فدرال امریکا ارائه می‌شود.

<sup>5</sup> Item-Level