

نمونه‌گیری

روشها و کاربردها

پل اس. لهوی

استنلی لمی شو

مترجم

گیتی مختاری امیرمجدی



پژوهشکده آمار

لوی، پل	Levy, Paul S.
نمونه‌گیری: روشها و کاربردها / پل اس. لهوی، استنلی لمی شو؛ مترجم گیتی مختاری امیرمجدی	
-- تهران: مرکز آمار ایران، پژوهشکده آمار، ۱۳۸۱.	
۵۸۵ ص. : جدول، نمودار.	
فهرست‌نویسی بر اساس اطلاعات فیفا.	ISBN 964-365-150-9 ریال : ۳۰۰۰۰
عنوان اصلی: Sampling of Populations: Methods and Applications.	
کتاب‌نامه.	
نمایه	
۱. جمعیت - روشهای آماری. ۲. آمارگیری نمونه‌ای. الف. لمشو، استنلی Lemeshow, Stanley	
ب. مختاری امیرمجدی، گیتی، ۱۳۳۳- ، مترجم. ج. مرکز آمار ایران، پژوهشکده آمار. د. عنوان.	
ان ۹/۴۹/۴۹ HB8۴۹/۴۹/۴۹	۳۰۴/۶۰۱۵۱۹۵۲
۱۳۸۱	
کتابخانه ملی ایران	۴۰۶۴۷-۸۱ م

- مدیریت تولید : گروه پژوهشی طرحهای فنی و روشهای آماری
- ویراستار علمی و ادبی : دکتر علی عمیدی
- حروف‌نگاری، نمونه‌خوانی، صفحه‌بندی، صفحه‌آرایی : محبوبه کاظمی
- ویراستار هنری : فرشید خان‌زاده
- طراحی جلد : نیما دانش‌پرور
- مدیر فنی : علی اصغر حائری مهریزی
- امور فنی و چاپ : مؤسسه انتشارات ستایش

© ۱۳۸۱ پژوهشکده آمار

شماره ۵۲، خیابان شهید فکوری، خیابان باباطاهر، خیابان دکتر فاطمی
تهران ۱۴۱۳۷۱۷۹۱۱، ایران



URL: <http://www.src.ac.ir>

e-mail: src@src.ac.ir

تلفن: ۸۹۵۹۰۲۹ دورنگار: ۸۰۰۷۹۸۹

همه حقوق این اثر برای پژوهشکده آمار محفوظ است. هیچ بخشی از این کتاب را نمی‌توان بدون اجازه کتبی از ناشرش تکثیر یا به هر شکلی و با هر وسیله‌ای ذخیره کرد. استفاده یا تقلید از طرح جلد، ممنوع است.

حروف‌نگاری شده با قلم‌های فارسی لوتوس و تیترو میترا، و قلم لاتین Times New Roman.

چاپ و صحافی شده در ایران.

چاپ یکم

شمارگان: ۶۰۰

پیشنهاد برای نحوه نقل مطلب، جدول یا نمودار از این کتاب، به صورت زیر است:

مختاری امیرمجدی، گیتی (۱۳۸۱). نمونه‌گیری: روشها و کاربردها. لهوی، پل اس.؛ لمی شو، استنلی. ترجمه از انگلیسی به فارسی. تهران: پژوهشکده آمار.

شابک ۹۶۴ - ۳۶۵ - ۱۵۰ - ۹

ISBN 964-365-150-9

بها: سی و پنج هزار ریال

فصل ۱۴

موضوعهای منتخب در طرح نمونه‌های و روش‌شناسی برآورد کردن

در این فصل به بحث در مورد برخی فنون نمونه‌گیری و برآورد کردن می‌پردازیم که برای مقاصد خاص ابداع شده‌اند. به صورتی مشخصتر، ابتدا به بحث دربارهٔ برخی روشهای نمونه‌گیری می‌پردازیم که برای استفاده در آمارگیریهای بهداشتی که در کشورهای در حال توسعه اجرا می‌شوند ابداع شده‌اند و سپس فنونی را بررسی می‌کنیم که در کشورهای توسعه یافته برای تأمین نیازهایی خاص بسط یافته‌اند.

۱.۱۴ آمارگیریهای برنامه گسترش یافته ایمن‌سازی سازمان بهداشت جهانی: تعدیلی

از نمونه‌گیری با احتمال متناسب با اندازه برای استفاده در کشورهای در حال توسعه سازمان بهداشت جهانی (WHO)^۱، از طریق برنامه گسترش یافته ایمن‌سازی (EPI)^۲ خود، دسترسی همهٔ کودکان جهان به ایمن‌سازی را هدف قرار داده است. این سازمان به عنوان بخشی از این برنامه، روش‌شناسی را ابداع کرده است که با استفاده از آمارگیریهای نمونه‌ای نسبتاً سریع و کم‌خرج بتواند

^۱ World Health Organization

^۲ Expanded Program on Immunization

سطوح ایمن‌سازی کودکان را برآورد کند و در صورت پایین بودن این سطوح، مناطقی را به منظور اجرای برنامه‌های ویژه مصون‌سازی هدف قرار دهد. در این روش، تعدیلی از نمونه‌گیری با احتمال متناسب با اندازه (PPS)^۱ به کار می‌رود که در اصل در آمریکا [۱] بسط داده شد و بعدها برای استفاده در برنامه ریشه‌کنی آبله در افریقای غربی [۲] مورد تعدیل قرار گرفت.

با این که آمارگیری برنامه گسترش یافته ایمن‌سازی می‌تواند برای اهداف دیگری نیز به کار رود، هدف عمده از اجرای آن برآورد کردن پوشش ایمن‌سازی (یعنی نسبت کودکانی که همه ایمن‌سازیهای مورد نیاز را دریافت می‌کنند) در یک ناحیه مشخص هدف است (که می‌تواند یک روستا، شهر، و غیره باشد). طی سالها، کاربستی در این آمارگیریها تکامل یافت که طی آن ۳۰ خوشه نمونه با ۷ کودک از هر خوشه انتخاب و اندازه نمونه از ۲۱۰ کودک تشکیل می‌شد. در واقع، این آمارگیریها به آمارگیریهای «۳۰ × ۷» موسوم شدند.

با این که تاکنون صدها آمارگیری «۳۰ × ۷» اجرا شده است، منطق انتخاب اندازه نمونه، نسبتاً غیرمعمول است. با فرض این که ۵۰٪ جامعه هدف تحت پوشش ایمن‌سازی قرار می‌گیرد، برای ۹۵٪ اطمینان از برآورد کردن نسبت مزبور در محدوده نقاط ۱۰ درصدی مقدار واقعی آن به یک نمونه تصادفی ساده با اندازه ۹۶ کودک نیاز است. ولی نمونه‌گیری تصادفی ساده در محیطی که آمارگیریهای برنامه گسترش یافته ایمن‌سازی اجرا می‌شوند مقرون به صرفه نیست و به همین دلیل نوعی فن نمونه‌گیری خوشه‌ای باید به صورت معمول به کار گرفته شود. اگر اثر طرح ۲ فرض شود، اندازه نمونه مورد نیاز دو برابر خواهد شد و به ۱۹۲ نفر (یعنی دو برابر ۹۶ نفر) برای نمونه نیاز است. کسانی که روش‌شناسی برنامه گسترش یافته ایمن‌سازی را ابداع کرده بودند تصمیم داشتند از ۳۰ خوشه استفاده کنند. بنابراین، نمونه‌ای متشکل از ۳۰ خوشه با ۷ نفر از هر خوشه موجب به دست آمدن نمونه‌ای می‌شد که اندازه آن اندکی بزرگتر از نمونه مطلوب بود. این موضوع با جزئیات بیشتر توسط لمی شو و استرا [۳] شرح داده شده است.

خوشه‌ها در آمارگیریهای برنامه گسترش یافته ایمن‌سازی به جامعه ویژه مورد آمارگیری بستگی دارند و معمولاً از روستاها، شهرها، یا نواحی خدمات بهداشتی تشکیل شده‌اند که اطلاعات جمعیتی مربوط به آنها موجود است. در مرحله اول، یک نمونه با احتمال متناسب با اندازه (به صورتی که در فصل ۱۱ شرح داده شد) انتخاب می‌شود. ولی در مرحله دوم به دلایل بودجه‌ای و لوژستیکی، فرایند انتخاب هفت نفر با نمونه‌گیری خوشه‌ای مرسوم تفاوتی بارز دارد. در روش‌شناسی مربوط به برنامه گسترش یافته ایمن‌سازی، به جای انتخاب تصادفی هفت آزمودنی از میان همه آزمودنیهای موجود،

^۱ Probability Proportional to Size

یک خانوار به تصادف برای شروع انتخاب می‌شود، همه اطلاعات مربوط به آزمودنی‌های واجد شرایط موجود در آن خانوار جمع‌آوری می‌شود، و سپس به خانواری مراجعه می‌شود که درب ورودی منزل او از لحاظ فیزیکی به خانوار قبلی از همه نزدیکتر است. این فرایند مراجعه به نزدیکترین خانوار بعدی و جمع‌آوری اطلاعات درباره همه افراد واجد شرایط در آن خانوارها ادامه می‌یابد تا هفت آزمودنی مورد نیاز بررسی شوند.

شیوه‌هایی خاص که برای انتخاب تصادفی اولین خانوار در مرحله دوم نمونه‌گیری به کار می‌روند به ماهیت خوشه بستگی دارند. مثلاً در یک ناحیه روستایی، ممکن است به مصاحبه‌گر آموزش داد که به مکانی که در داخل خوشه دارای مرکزیت است (از قبیل کلیسا یا بازار) مراجعه کند، به طور تصادفی یک جهت را برای طی طریق انتخاب کند (یعنی شمال، جنوب، شرق، یا غرب) و تعداد K خانوار موجود در آن جهت را از نقطه شروع تا مرز شهر بشمارد. سپس یک عدد تصادفی را بین ۱ و K انتخاب کند تا اولین خانوار تعیین شود. طبق قاعده عملکرد، داده‌های مربوط به همه کودکان واجد شرایط در خانواری که هفتمین کودک در آن است جمع‌آوری می‌شوند، حتی اگر این کار موجب شود که بیش از ۷ کودک در یک خوشه گنجانده شود.

دریافته‌اند که روش انتخاب کودکان به شرح بالا، در داخل خوشه‌های نمونه برای آمارگیریه‌های برنامه گسترش یافته ایمن‌سازی از لحاظ اجرای میدانی، نسبتاً آسان است. هزینه‌های یک انتخاب واقعاً تصادفی در مرحله دوم با توجه به منابعی که نوعاً به این آمارگیریه‌های EPI اختصاص داده می‌شود بازدارنده خواهد بود. در تلاشی برای ارزشیابی اثر این انتخاب غیرنمونه‌ای افراد در داخل خوشه‌ها بر برآوردهای مربوط به پوشش ایمن‌سازی، یک مدل شبیه‌سازی رایانه‌ای ابداع شده است [۴]. این مدل، با استفاده از جامعه‌هایی که با مشخصه‌های ویژه به طور مصنوعی ایجاد شده‌اند، نتایج حاصل از روش انتخاب برنامه گسترش یافته ایمن‌سازی را با نتایج به دست آمده از روشهای سنتی تر مقایسه کرده است. معلوم شده است که با وجود ضعف عمل روش برنامه گسترش یافته ایمن‌سازی در داخل خوشه‌های ویژه‌ای که در آنها پنهانکاریهایی درباره افراد ایمن‌سازی نشده وجود داشته است، نتایج حاصل از همه خوشه‌ها در مجموع، گرایش به دقیق بودن تا حدود ۱۰ درصد سطوح واقعی جامعه داشته است. این در محدوده اهدافی است که به وسیله برنامه گسترش یافته ایمن‌سازی تعیین شده و می‌تواند به این واقعیت نسبت داده شود که تعداد خوشه‌های انتخاب شده زیاد است و اریبیهایی که در داخل خوشه‌ها روی می‌دهند در مجموعه ۳۰ تایی گرایش به سربه‌سر شدن دارند. به این ترتیب، به نظر می‌رسد این روش، هنگام استفاده در سراسر مناطق هدف، کاملاً سودمند باشد ولی در صورت

استفاده در خوشه‌هایی خاص یا برای زیرگروهها می‌تواند برآوردهایی فراهم کند که به کلی غیر قابل قبول باشند.

۲.۱۴ نمونه‌گیری تضمین کیفیت

طرح آمارگیری برنامه گسترش یافته ایمن‌سازی که در بخش ۱.۱۴ شرح داده شد دارای مزایای عمده سادگی و کم‌هزینه بودن است که برای برآورد کردن پارامترهایی از قبیل پوشش مایه‌کوبی در کشورها یا نواحی موردنظر، گزینه‌ای جذاب تلقی می‌شود. ولی روش‌شناسی برنامه گسترش یافته ایمن‌سازی، به این علت که یک فن آمارگیری خوشه‌ای است، نمی‌تواند اطلاعات مربوط به واحدهای جامعه‌ای کوچک یا نواحی بهداشت را فراهم آورد. مدیران برنامه بهداشت برای تمرکز بر فعالیتهای نظارتی به این اطلاعات نیاز دارند. زیرا برخی واحدهای بهداشتی ممکن است در اجرای وظایف خاص خود نسبت به دیگر واحدها با شایستگی کمتری عمل کنند. مدیران می‌توانند با شناسایی واحدهای بهداشتی که به اهداف اظهار شده خود دست نیافته‌اند (مثلاً پوشش ایمن‌سازی در سطح عالی) تلاشهای مستقیم خود را در آن مناطقی هدایت کنند که به بیشترین بهبود نیاز دارند. شیوه نمونه‌گیری تضمین کیفیت (QAS)^۱ به عنوان جایگزین نمونه‌گیری برنامه گسترش یافته ایمن‌سازی (EPI) برای استفاده به صورت ابزار نظارتی پیشنهاد شده و مورد توجه قابل ملاحظه‌ای قرار گرفته است.

روشهای نمونه‌گیری تضمین کیفیت در اصل برای نمونه‌گیری و بازرسی محصول تولید شده در مواردی به کار برده می‌شدند که ضرورت ایجاد می‌کرد که هزینه‌های نیروی کار و سایر هزینه‌های نمونه‌گیری در سطوح حداقل حفظ شوند. یک نوع نمونه‌گیری تضمین کیفیت، یعنی نمونه‌گیری تضمین کیفیت دسته‌ای^۲، اساساً نمونه‌گیری طبقه‌بندی شده است، ولی اندازه‌های نمونه کوچکتر از آنند که آنچه را تأمین کنند که معمولاً بازه‌های اطمینان باریک در حد قابل قبول برای برآوردها در داخل یک طبقه معین تلقی می‌شوند (و معمولاً به «دسته» یا «بسته» موسوم‌اند). در عوض، براساس این احتمال که تعداد اقلام معیوب یک نمونه مربوط به دسته یا بسته خاص موردنظر، از عدد تعیین شده‌ای کمتر یا برابر با آن باشد، در مورد کیفیت آن دسته تصمیم‌گیری می‌شود. نتایج حاصل از نمونه‌های انتخاب شده از همه بسته‌های دوه‌دو ناسازگار و فراگیر را می‌توان ترکیب کرد تا برآورد کل دقیقی از متوسط کیفیت کل محصول به دست آید.

متوسط کیفیت محصول غالباً به طور مداوم توسط تولید کننده، مورد نظارت قرار می‌گیرد تا (۱) جایی را که می‌توان در فرایند تولید بهبودی ایجاد کرد شناسایی نمود و (۲) طرح نمونه‌ای را وقتی

^۱ Quality Assurance Sampling

^۲ Lot Quality Assurance Sampling (LQAS)

کیفیت متوسط محصول تغییر می‌کند تعدیل کرد. به طور کلی، یک دسته واحدی است که «از نظر عملیاتی سودمند» است. مثلاً در یک کاربرد صنعتی، اگر چند ماشین یک قطعه بخصوص را تولید می‌کنند می‌توان دسته‌هایی را که یک ماشین تولید می‌کند انتخاب کرد، به خصوص اگر بیشتر احتمال داشته باشد که هرگونه تغییر در قطعات تولیدی ناشی از طرز کار دستگاه باشد تا نهاده عملگر. بازه نمونه‌گیری تولید کننده باید به اندازه کافی کوتاه باشد تا بتوان هرگونه انحراف در اندازه‌گیریها را، پیش از آن که از حدود تحمل خارج شوند، شناسایی کرد. برای این نوع کاربرد، ارزش دارد که دنباله اندازه‌گیریها برای شناسایی سریع گرایش به انحراف تحت نظارت قرار گیرند.

برای کاربردهای مربوط به بهداشت عمومی، یک مدیر ملی می‌تواند دسته‌ها را به صورت دریافت‌کنندگان خدمات از یک واحد عملیاتی تک، از قبیل یک تیم ویژه ایمن‌سازی، در یک دوره زمانی مشخص تعریف کند. فاصله زمانی بین نمونه‌گیریها می‌تواند به بازه‌های بین فصلهای «وقوع زیاد» بیماریهای قابل پیشگیری وابسته باشد، ولی احتمالاً به همین اندازه نیز به مدت زمان و مقدار هزینه مربوط به نمونه‌گیری ارتباط خواهد داشت تا به هرگونه ملاحظات دیگر.

در کاربردهای مربوط به بهداشت عمومی، اگر چنین قضاوت شود که جامعه به طرز مناسبی پوشش داده شده است («یعنی دسته پذیرفته شود»)، در حالی که واقعاً چنین نبوده است، خطای فاحشی روی خواهد داد. برای کنترل این امکان، شیوه‌ای به صورت یک آزمون آماری یک‌طرفه ایجاد شده است. فرض کنید d تعداد افراد مایه‌کوبی نشده در نمونه متشکل از n آزمودنی باشد، و P نسبت واقعی افراد مایه‌کوبی نشده در جامعه‌ای با اندازه N باشد. فرض بر این است که N به نسبت n بسیار بزرگ است. (اگر اتفاقاً N به نسبت n بزرگ نباشد، خواننده می‌تواند به متونی از قبیل متن درسی تألیف براونلی [۲۹، بخش ۱۵.۳] مراجعه کند که چگونگی استفاده از توزیع فوق هندسی را برای ارزشیابی شیوه نمونه‌گیری تضمین کیفیت دسته‌ای نشان می‌دهد.)

فرض صفر چنین است

$$H_0: P \geq P_0 \quad *(\text{یعنی، } \geq 0.50 \text{ نسبت کودکان مایه‌کوبی نشده})$$

در مقابل

$$H_a: P < P_0 \quad (\text{یعنی، } < 0.50 \text{ نسبت کودکان مایه‌کوبی نشده})$$

جدول چهار خانه‌ای زیر، پیامدهای شیوه آزمون را توصیف می‌کند:

* سطح ۰.۵۰٪ در اینجا برای مثال انتخاب شده است. در واقع هر سطحی را می‌توان انتخاب کرد.

پیامدهای آزمون فرض در شیوه نمونه‌گیری تضمین کیفیت دسته‌ای

جامعه واقعی

	به طرز مناسبی مایه‌کوبی نشده	به طرز مناسبی مایه‌کوبی شده	
رد دسته →	مخاطره تأمین‌کننده β نرخ مثبت غلط	آزمون، فقدان پوشش مناسب را تشخیص می‌دهد یا نسبت به آن حساس است $1 - \alpha$ حساسیت	عدم موفقیت در رد H_0 : پوشش نامناسب
پذیرش دسته →	آزمون، پوشش مناسب را تشخیص می‌دهد $1 - \beta$ ویژگی	مخاطره مصرف‌کننده α نرخ منفی غلط	رد H_0 : پوشش مناسب

توجه کنید که در این جدول، چون آزمون به صورت یک‌طرفه تهیه شده و چون فرض بر این است که جامعه به طرز مناسبی پوشش داده نشده است، خطای نوع I، پذیرش دسته هنگامی که معیوب است (منفی غلط)، یعنی احتمالی که می‌توانیم تحت کنترل درآوریم، جدیترین خطاست، مگر این که H_0 را رد کنیم. یعنی، با استفاده از مثال ایمن‌سازی، اگر جامعه‌ای (دسته‌ای) از کودکان دارای نسبت ایمن‌سازی شده قابل قبولی تصور شود درحالی که در حقیقت چنین نیست، تعداد بیشتر کودکان آسیب‌پذیر موجود در جامعه، خطر انتقال بیماری را هنگام بروز آن در دسته افزایش خواهد داد. از این رو است که «هزینه» اعلام این را که جامعه به طرز مناسبی مایه‌کوبی شده است، وقتی حقیقتاً چنین نیست، سنگین در نظر می‌گیریم. از سوی دیگر، خطای نوع II - یعنی رد دسته قابل قبول - کمتر جدی دآوری می‌شود زیرا نتیجه یک تصمیم مثبت غلط آن است که منابع برنامه روی جامعه‌ای متمرکز شود که هم‌اکنون نیز به طرز مناسب ایمن‌سازی شده است.

مشکل اساسی در نمونه‌گیری تضمین کیفیت دسته‌ای بیش از آن که صرفاً تعیین اندازه نمونه باشد انتخاب توازنی مناسب بین اندازه نمونه و ناحیه بحرانی است. محاسبه احتمال خطای نوع II، β ، در همه موارد به مقدار واقعی P ، وقتی فرض می‌شود با P تفاوت دارد، بستگی خواهد داشت. روش محاسبه احتمالها و تعیین اندازه‌های نمونه با استفاده از توزیع دو جمله‌ای انجام خواهد شد و به تفصیل در جاهای دیگر شرح داده شده است [۳۰] و [۳۱].

روش نمونه‌گیری تضمین کیفیت که تا اینجا شرح داده شد به «نمونه‌گیری تک‌مرحله‌ای» موسوم است، زیرا فقط یک نمونه برای تصمیم‌گیری درباره‌ی وضع دسته انتخاب می‌شود. تعدیلی از شیوه‌ی نمونه‌گیری تضمین کیفیت دسته‌ای شامل راهبرد «نمونه‌گیری دوگانه» است و می‌تواند تحت شرایط میدانی معین مفید واقع شود.

یک نمونه‌گیری دوگانه یا برنامه‌ی نمونه‌گیری دو مرحله‌ای را می‌توان برای پایین آوردن هزینه‌های نمونه‌گیری یک آمارگیری مورد استفاده قرار داد. در وضعیت حاضر، دو مقدار بحرانی، d_1^* و d_2^* ، تعیین شده‌اند که $d_1^* \leq d_2^*$ و دو اندازه‌ی نمونه‌ای، n_1 و n_2 ، مشخص شده‌اند. در مرحله‌ی اول، n_1 نفر مورد مطالعه قرار می‌گیرند. اگر تعداد مشاهده‌ی افراد مایه‌کوبی نشده کمتر از d_1^* یا برابر با آن باشد نتیجه می‌گیریم که نسبت واقعی افراد مایه‌کوبی نشده در جامعه به صورتی معنی‌دار کمتر از P است. اگر تعداد مشاهده‌ی افراد مایه‌کوبی نشده در مرحله‌ی اول بیشتر از d_1^* ولی کمتر از d_2^* یا برابر با آن باشد، در آن صورت به مرحله‌ی دوم می‌رویم و نمونه‌گیری را ادامه می‌دهیم تا یا $d_2^* + 1$ نفر مایه‌کوبی نشده را مشاهده کنیم که حاکی از سطح پایین پوشش مایه‌کوبی است، یا مجموعاً $n_1 + n_2$ نفر نمونه‌گیری شوند بدون این که از d_2^* بیشتر شوند و این حاکی از سطح بالای پوشش مایه‌کوبی است. به جای این تصور که نمونه‌گیری دوگانه یک برنامه‌ی نمونه‌گیری تک‌مرحله‌ای است که یک مرحله نمونه‌گیری اضافی به دنبال دارد باید دانست که کل اندازه‌ی نمونه و مقدار بحرانی هر دو مرحله، با برنامه‌ی نمونه‌گیری تک‌مرحله‌ای که قبلاً توصیف شد مطابقت دارد در حالی که مرحله‌ی اول معرف یک نمونه «تحویل یافته» اولیه است. منشأ صرفه‌جویی بیشتر با نمونه‌گیری دوگانه نیز در همین امر نهفته است. هرگاه نتایج حاصل در مرحله‌ی اول کرانگین باشند، برنامه‌ریز بهداشت می‌تواند بر این اساس نتیجه‌گیری کند که آزمودنی‌های نمونه‌گیری شده کمتر از آزمودنی‌های نمونه‌گیری شده با استفاده از برنامه‌ی نمونه‌گیری یک مرحله‌ای است. در غیر این صورت، برنامه‌ریز بهداشت با علم به این که ماکسیمم تعداد نمونه‌گیری شده برابر با اندازه‌ی نمونه حاصل از برنامه‌ی یک مرحله‌ای خواهد بود به نمونه‌گیری ادامه خواهد داد. به عبارت دیگر، استفاده از برنامه‌ی نمونه‌گیری دوگانه تضمین می‌کند که تعداد افراد نمونه‌گیری شده کمتر از تعداد حاصل از برنامه‌ی نمونه‌گیری یک مرحله‌ای بوده یا برابر با آن باشد.

نمونه‌گیری دوگانه در داخل متن طرح نمونه‌گیری تضمین کیفیت دسته‌ای، در عین حالی که می‌تواند تحت شرایطی خاص، صرفه‌جویی در آمارگیری بهداشت را افزایش دهد، ممکن است در شرایطی دیگر اصلاً امکان‌پذیر نباشد. مثلاً اگر تحلیل نتایج آزمون پزشکی در محل امکان‌پذیر نباشد، آن‌گاه ممکن است برنامه‌ی نمونه‌گیری دوگانه مستلزم آن باشد که اگر نتایج آماری حاصل از نمونه‌ی اول کرانگین یا قانع‌کننده نباشند یک تیم پزشکی برای نمونه‌گیری بیشتر به محله‌ی موردنظر بازگردند. هزینه

زمانی و مراجعه به محل ممکن است با صرفه اقتصادی بالقوه حاصل از نمونه‌گیری دوگانه سربه‌سر شود. ولی اگر هزینه اجرای آزمایش پزشکی به ازای هر آزمودنی زیاد باشد ممکن است برنامه‌ریز بهداشت احساس کند که سود بالقوه حاصل از ناگزیر بودن احتمالی اجرای فقط مرحله اول نمونه‌گیری، موجب ترجیح نمونه‌گیری دوگانه شود.

نمونه‌گیری تضمین کیفیت دوگانه دسته‌ای، یک فن نمونه‌گیری طبقه‌بندی شده است. این نمونه‌گیری به جای ساختن برآوردهای بازه‌های اطمینان پارامترهای نامعلوم، مسئله را به یک سری آزمونه‌های فرض تبدیل می‌کند. ولی اطلاعاتی که فراهم می‌کند از نمونه‌گیری تصادفی طبقه‌بندی شده متداول بیشتر نیست زیرا با استفاده از فن اخیر، اگر اندازه‌های نمونه‌ای به قدر کافی برای تأمین بازه‌های اطمینان سودمند بزرگ باشند، می‌توان بازه‌های اطمینان را برای هر طبقه (دسته) تعیین و درباره مقادیر پوشش داده شده با هر یک از این قبیل بازه‌ها تصمیم‌گیری کرد.

اگرچه اندازه‌های نمونه‌ای برای هر طبقه در یک طرح آمارگیری تضمین کیفیت دسته‌ای، نوعاً کوچکتر از آنند که بازه‌های اطمینان سودمندی برای برآوردهای مربوط به هر طبقه فراهم کنند، ولی یک طرح آمارگیری تضمین کیفیت دسته‌ای که به صورتی مناسب طراحی شده باشد می‌تواند ابزاری برای آزمون پیوسته طبقه‌ها و رده‌بندی آنها به عنوان «قابل قبول» یا «غیر قابل قبول» از لحاظ یک برآمد ویژه فراهم کند. چون نمونه‌های مربوط به نمونه‌گیری تضمین کیفیت دسته‌ای نسبتاً کوچک‌اند احتمال اجرای این نمونه‌گیری بسیار بیشتر است. شاید بتوان نمونه‌ها را همزمان با انجام سایر وظایفی که کارکنان را به میدان کار می‌کشاند انتخاب کرد. چون قواعد ساده‌ای برای اجرا در نظر گرفته شده‌اند مأمور آمارگیری یا رده‌بندی به حداقل آموزش نیاز دارد. نتایج مربوط به طبقه‌ها را می‌توان ترکیب کرد تا برآوردهایی مناسب برای گروههایی از طبقات از قبیل بخشها، ناحیه‌ها، یا کل کشور فراهم شود.

هر چند که بازه‌های اطمینان همواره بیشتر از یک تصمیم دوحالتی ساده اطلاعات فراهم می‌کنند، ولی اندازه‌های نمونه‌ای مورد نیاز در به دست آوردن برآوردهایی با هر سطحی از دقت مفید، برای طبقه‌های نسبتاً کوچک ممکن است عملی نباشد. در این گونه موارد، یک طرح نمونه‌گیری تضمین کیفیت مناسب می‌تواند جایگزینی باشد که ارزش بررسی دارد.

۳.۱۴ اندازه‌های نمونه برای بررسیهای طولی

یک طرح با اهمیت که در مطالعات مربوط به همه‌گیری شناختی مورد استفاده واقع می‌شود بررسی همگروهی است. در این نوع بررسی، افراد را شناسایی و در زمان شناسایی اندازه‌گیری می‌کنند و گاهی سالهای متمادی آنها را برای تعیین متغیرهای برآمد (مثلاً وقوع یک بیماری خاص

موردنظر یا مرگ و میر ناشی از آن) پیگیری می‌نمایند. هدفهای عمده این قبیل مطالعات عبارت از بررسی روابط بین این متغیرهای برآمد و مشخصه‌های افراد حاضر در زمان شناسایی آغازین است. مثالهای مربوط به این قبیل مطالعات شامل بررسی بیماری رگهای قلبی توسط فرامینگام [۳۲]، بررسی فعالیتهای جسمانی توسط رالف پافنبارگر [۳۳]، و بررسی طولی سالخوردگی [۳۴] است.

این قبیل مطالعات غالباً به اندازه همگروهی بزرگ نیاز دارند که براساس فرضیهایی تعیین می‌شود که باید از نظر روابط بین وقوع متغیرهای برآمد و مشخصه‌های موجود در زمان تعیین همگروهها مورد آزمون قرار بگیرند. روشهای مربوط به تعیین این قبیل اندازه‌های همگروهی مورد نیاز، در نوشتگان مربوط به همه‌گیری شناختی شرح داده شده‌اند (مانند اثر روتمن و بویس [۳۵]). ولی به محض این که اندازه همگروهی مورد نیاز تعیین شد، شناسایی تعداد کافی از افراد برای تأمین ملاکهای منظور شدن در گروه ضرورت پیدا می‌کند.

غالباً شناسایی اعضای گروه به وسیله یک آمارگیری انجام می‌پذیرد که در آن نمونه‌ای از افراد انتخاب می‌شوند و افراد نمونه به منظور تأمین ملاکهای مربوط به گنج‌انیده شدن در گروه غربالگری می‌شوند. سپس کسانی که ملاکهای موردنظر درباره آنها مصداق دارد برای برسیهای بعدی در گروه جای داده می‌شوند. اندازه نمونه‌ای که باید برای این غربالگری انتخاب شود در زیر برای دو طرح نمونه‌ای بیان می‌شود: نمونه‌گیری تصادفی ساده از افراد، و نمونه‌گیری خوشه‌ای یک‌مرحله‌ای ساده.

۱.۳.۱۴ نمونه‌گیری تصادفی ساده

برای کسب اطمینان $(1-\alpha) \times 100$ درصد برای به دست آوردن حداقل n^* آزمودنی دارای ملاکی خاص، تعداد تقریبی n آزمودنی مورد نیاز برای نمونه‌گیری از فرمول زیر به دست می‌آید

$$n = \left[A + \sqrt{A^2 + \frac{n^*}{P}} \right]^2 \quad (1.14)$$

که در آن

$$A = \frac{|z_\alpha|}{2} \times \sqrt{\frac{1-P}{P}}$$

z_α = صدک توزیع نرمال $100 \times \alpha$

P = نسبت افراد در جامعه که ملاکهای منظور شدن در بررسی همگروهی را تأمین می‌کنند

(یعنی ورود افراد واجد شرایط برای منظور شدن در گروه)

اندازه نمونه‌ای که از رابطه (۱.۱۴) به دست می‌آید بر مبنای اندازه جامعه است که به مراتب بزرگتر از اندازه نمونه است و مبتنی بر تقریب توزیع نرمال برای توزیع دو جمله‌ای است.

مثال تشریحی: فرض کنید برای یک بررسی مقدماتی پیگیر بلندمدت درباره تحرک شغلی به ۲۰۰ کارمند اجرایی عالی‌رتبه بیمارستانها نیاز است. این افراد برای این که در بررسی گنجانیده شوند باید بین ۳۰ تا ۳۹ سال سن داشته باشند. روش وارد کردن این افراد در بررسی به شرح زیر است. از روی پرونده‌ای که شامل اسامی کارمندان اجرایی عالی‌رتبه تقریباً ۶۰۰۰ بیمارستان محلی در سراسر امریکاست یک نمونه تصادفی ساده گرفته می‌شود. به هر یک از این کارمندان نمونه تلفن کرده، در مورد سن از او سؤال می‌شود. اگر سن او بین ۳۰ تا ۳۹ سال بود، آنگاه کوششی به عمل می‌آید تا مشارکت وی در این بررسی جلب شود. حدس بر این است که تقریباً ۳۰٪ از همه کارمندان اجرایی عالی‌رتبه در دامنه سنی هدف قرار دارند. اگر بخواهیم ۹۵٪ مطمئن باشیم که حداقل ۲۰۰ کارمند اجرایی عالی‌رتبه را به این ترتیب به دست می‌آوریم (با فرض مشارکت ۸۵٪ از میان همه افراد نمونه واجد شرایط)، چند کارمند اجرایی باید نمونه‌گیری شوند؟

با استفاده از رابطه (۱.۱۴) با $P = 0.30$ ، $n^* = 200$ و $z_\alpha = -1.645$

داریم

$$A = \frac{1.645}{2} \times \sqrt{\frac{1-0.30}{0.30}} = 1.256$$

و

$$n = \left[1.256 + \sqrt{(1.256)^2 + \left(\frac{200}{0.30} \right)} \right]^2 \approx 735$$

و با در نظر گرفتن ۸۵٪ موفقیت در جلب همکاری، n نهایی مورد نیاز به این صورت خواهد بود

$$n = \frac{735}{0.85} \approx 865$$

به این ترتیب باید ۸۶۵ کارمند اجرایی عالی‌رتبه نمونه‌گیری شوند تا ۹۵٪ مطمئن باشیم که حداقل ۲۰۰ نفر را در گروه سنی هدف انتخاب خواهیم کرد.

□

فرمول‌بندی که در بالا شرح داده شد برای نمونه‌گیری تصادفی ساده است. ولی یک طرح نمونه‌گیری که احتمال استفاده از آن برای شناسایی افراد در عمل بیشتر است نمونه‌گیری خوشه‌ای

است که در آن ممکن است خانوار یک خوشه تلقی شود و افراد موجود در هر خانوار نمونه، از نظر داشتن شرایط برای گنجاینده شدن در گروه غربالگری شوند. تعیین ویژگی مشابه برای اندازه نمونه آن است که از خوشه‌ها به تعداد کافی، m تا، نمونه‌گیری شود به طوری که تعداد مورد نیاز n^* از افراد با $(1-\alpha) \times 100\%$ اطمینان برای گنجاینده شدن در گروه شناسایی شوند. با این ویژگی و تحت طرح نمونه‌گیری خوشه‌ای یک مرحله‌ای ساده می‌توان رابطه (۱.۱۴) را با تعدیلهای زیر به کار برد.

۲.۳.۱۴ نمونه‌گیری خوشه‌ای یک مرحله‌ای ساده

$$m = \left[A + \sqrt{A^2 + \frac{n^*}{\bar{N}}} \right]^2 \quad (2.14)$$

که در آن

$$A = |z_\alpha| \times \frac{V_N}{\sqrt{2}}$$

$$V_N = \frac{\sigma_N}{\bar{N}}$$

$$\sigma_N = \sqrt{\sum_{i=1}^M (N_i - \bar{N})^2 / M}$$

M = تعداد خوشه‌ها در جامعه

N_i = تعداد افراد واجد شرایط در خوشه i ، $i = 1, \dots, M$

$$\bar{N} = \sum_{i=1}^M N_i / M \quad (\text{میانگین تعداد افراد واجد شرایط به ازای خوشه})$$

باید توجه داشت که \bar{N} ، میانگین تعداد افراد واجد شرایط به ازای خوشه، مشابه P ، یعنی نسبت افراد واجد شرایط جامعه در تعیین ویژگیها برای اندازه نمونه تحت نمونه‌گیری تصادفی ساده است.

مثال تشریحی: فرض کنید می‌خواهیم نمونه‌ای از خانوارها در یک شهر بزرگ بگیریم و یکایک اعضای خانوار را به منظور شناسایی ۵۰۰ زن مجرد ۲۵ تا ۳۹ سال سن که شاغل تمام وقت‌اند و حداقل یک فرزند مدرسه‌رو دارند غربالگری کنیم. باز فرض کنید که در شهر تقریباً ۳۰۰۰۰۰ زن با این شرایط وجود دارند و شهر دارای تقریباً ۳۰۰۰۰۰ خانوار است و $\sigma_N \approx \sqrt{\bar{N}}$. چند خانوار باید نمونه‌گیری شوند تا با اطمینان ۹۵٪ حداقل ۵۰۰ زن با ملاکهای بالا شناسایی شوند؟

$$\begin{aligned} \bar{N} &= \frac{30000}{30000} = 0/1 && \text{(میانگین تعداد زنان مجرد به ازای خانوار)} \\ \sigma_N &= \sqrt{0/1} = 0/316 \\ V_N &= \frac{0/316}{0/1} = 3/16 \\ A &= 1/645 \times \frac{3/16}{2} = 2/60 \\ m &= \left[2/60 + \sqrt{(2/60)^2 + \left(\frac{500}{0/1}\right)} \right]^2 \approx 5382 \end{aligned}$$

به این ترتیب باید ۵۳۸۲ خانوار نمونه‌گیری شوند تا ۹۵٪ مطمئن باشیم که حداقل ۵۰۰ زن واجد شرایط شناسایی می‌شوند (با فرض مشارکت ۱۰۰٪ در بین زنان واجد شرایط).

□

۳.۳.۱۴ نمونه‌گیری خوشه‌ای با بیش از یک حوزه

غالباً وضعیت طوری است که می‌خواهیم n^* فرد را در هر یک از چندین حوزه معین برای بررسی آینده شناسایی کنیم. برای مثال، ممکن است بخواهیم ۱۵۰۰ نفر را در هر یک از چهار رسته نژادی - جنسی شناسایی کنیم (مرد سفیدپوست، مرد سیاهپوست، زن سفیدپوست، زن سیاهپوست). اگر طرح نمونه‌ای تعیین شده برای شناسایی این افراد، نمونه‌گیری خوشه‌ای یک‌مرحله‌ای ساده باشد، در آن صورت، تعداد خوشه‌های مورد نیاز که باید نمونه‌گیری شوند تا حداقل n^* نفر در هر یک از این حوزه‌ها به دست آید با فرمول زیر محاسبه می‌شود

$$m = \max(m_1, \dots, m_H) \tag{۳.۱۴}$$

که در آن

$$m_h = \left[A_h + \sqrt{A_h^2 + \frac{n^*}{N_h}} \right]^2, \quad h = 1, \dots, H \tag{۴.۱۴}$$

A_h و \bar{N}_h همتهای ویژه حوزه‌ای پارامترهایی هستند که برای رابطه (۲.۱۴) تعریف شدند و $H =$ تعداد حوزه‌های موردنظر.

می‌توان بدون از دست دادن کلیت، فرض کرد که $m_1 \leq m_2 \leq \dots \leq m_H$. پس، با نمونه‌گیری خوشه‌ای یک‌مرحله‌ای ساده، این احتمال برای همه حوزه‌ها بجز حوزه H وجود دارد که بیش از تعداد

مورد نیاز برای تأمین ویژگیهای موردنظر، دارای آزمودنیهای واجد شرایط باشند. مطلب مذکور، به دلیل این واقعیت است که در نمونه‌گیری خوشه‌ای یک‌مرحله‌ای معمولی، همین که یک خوشه در نمونه انتخاب شد دیگر نمونه‌گیری فرعی از افراد انجام نمی‌شود. لهوی و همکاران [۵]، [۶] نوعی نمونه‌گیری خوشه‌ای یک‌مرحله‌ای ساده را ابداع کرده‌اند که در آن خوشه‌های نمونه به طور تصادفی در یکی از گروه‌های H قرار داده می‌شوند. آنها که در گروه ۱ قرار دارند می‌توانند افرادی را از همه حوزدهای H بگیرند. آنهایی که در گروه ۲ قرار دارند می‌توانند افرادی را از حوزه ۲ تا H بگیرند ولی از حوزه ۱ نمی‌توانند بگیرند. همین طور تا خوشه‌های گروه h که فقط می‌توانند افرادی را از حوزه H بگیرند. این تعدیل نمونه‌گیری خوشه‌ای یک‌مرحله‌ای ساده، به نمونه‌گیری خوشه‌ای یک‌مرحله‌ای تلسکوپی موسوم است و به این نتیجه منتهی می‌شود که ویژگیهای موردنظر برای هر حوزه دقیقاً تأمین می‌شود و از مشکل گرفتن تعداد افرادی بیش از حد لازم در بیشتر حوزه‌ها اجتناب می‌شود. این روش در یک آمارگیری بزرگ از اشخاص سالخورده در شانگهای، چین، با موفقیت به کار رفته است [۵]، [۶].

۴.۱۴ برآورد کردن شیوع بیماریها از روی بررسیهای غربالگری

غالباً مایل‌اند شیوع یک بیماری یا شرطی در یک جامعه را براساس آزمون غربالگری برآورد کنند که حساسیت و ویژگی آن چندان کامل نباشد. این دو اصطلاح به معنای زیرند:

حساسیت، احتمال آن است که شخصی که دچار یک بیماری یا شرط خاص است آزمون غربالگری مثبت داشته باشد.

ویژگی، احتمال آن است که شخصی که دچار یک بیماری یا شرط خاص نیست دارای آزمون غربالگری مثبت نباشد.

آزمون غربالگری معمولاً آزمونی است که در مقایسه با آزمون تشخیصی دقیقتری که در عین اینکه قابل دسترس است برای استفاده در وضعیت آمارگیری، شدنی نیست، برای استفاده میدانی کم‌هزینه و شدنی باشد. مثلاً آزمایش وضعیت ذهنی (MMSE)^۱ شامل چند سؤال است که برای غربالگری افراد از لحاظ اختلالات شناختی که می‌تواند نشانه زوال عقل (اختلال مشاعر) باشد یا نباشد [۶] مورد استفاده قرار می‌گیرد در حالی که برای تشخیص این اختلال مشاعر، با هر میزان معقولی از درستی، به یک سری آزمایشات پیچیده روانی و عصب‌شناختی نیاز خواهد بود.

^۱ Mini Mental State Examination

اهداف عمده برنامه‌های غربالگری معمولاً شناسایی افراد دچار شرط یا بیماری موردنظر برای مداخله بعدی است، و مسایل آماری این برنامه‌ها مستلزم ارزیابی احتمال این است که فردی که به عنوان مثبت غربالگری شده واقعاً دچار آن بیماری است و به همین قیاس احتمال این است که فردی که به عنوان منفی غربالگری شده واقعاً دچار آن بیماری نیست (که به ترتیب مقدار پیشگویی آزمون مثبت و مقدار پیشگویی آزمون منفی نامیده می‌شوند). مباحث مربوط به این مطالب در نوشتگان مربوط به همه‌گیری شناختی فراوانند [۷] و ما در عوض، بحث خود را بر مسئله برآورد کردن شیوع بیماری از طریق آزمون غربالگری متمرکز می‌کنیم.

روش‌شناسی پایه‌ای، مستلزم انتخاب نمونه‌ای n فردی از جامعه‌ای متشکل از N نفر و اجرای آزمون غربالگری درباره هر یک از افراد نمونه است. اگر k از n نفر افراد نمونه به عنوان مثبت غربالگری شوند، لهوی و کاس [۸] نشان داده‌اند که در آن صورت $\hat{\pi}$ ، برآورد درست‌نمایی ماکسیمم π ، شیوع نامعلوم برای افراد دچار آن بیماری از فرمول زیر به دست می‌آید.

$$\hat{\pi} = \frac{\hat{p} + S_p - 1}{S_e + S_p - 1} \quad (5.14)$$

که در آن، $\hat{p} = \frac{k}{n}$ = نسبت دارندگان آزمون مثبت، S_e = حساسیت آزمون و S_p = ویژگی آزمون. از رابطه (۵.۱۴) می‌توان دید که در وضعیتهایی که مجموع ویژگی و حساسیت کمتر از یک است یا هنگامی که مجموع شیوع مربوط به آزمون و ویژگی آن کمتر از یک است برآورد شیوع $\hat{\pi}$ منفی خواهد بود. گاستویرت [۹] و لیو و لهوی [۱۰] تعدیلهایی از معادله (۵.۱۴) را پیشنهاد کرده‌اند که امکان برآوردهای منفی را برای شیوع از میان می‌برند. در عمل، یک آزمون غربالگری که حساسیت و ویژگی آن به اندازه‌ای کم است که احتمال برآوردهای منفی برای شیوع را به وجود می‌آورد احتمالاً مطلوبیت چندانی به عنوان یک غربال برای بیماری یا شرطی خاص نخواهد داشت.

خطای معیار برآورد شیوع، $\hat{\pi}$ ، از فرمول زیر به دست می‌آید

$$SE(\hat{\pi}) = \frac{SE(\hat{p})}{(S_e + S_p - 1)} \quad (6.14)$$

که در آن خطای معیار \hat{p} ، $SE(\hat{p})$ ، به طرح نمونه‌ای خاص و شیوه‌های برآورد کردن مورد استفاده بستگی خواهد داشت.

مثال تشریحی: فرض کنید یک نمونه تصادفی ساده متشکل از ۱۵۰ عضو از میان اعضای اتحادیه‌ای گرفته شده است که نماینده معلمان شاغل در هفت ناحیه آموزشی در بخش بزرگی از ایالت ایلی‌نوی

است. این هفت ناحیه آموزشی ۲۵۶۰ آموزگار را در استخدام خود دارند. همه معلمهای نمونه‌گیری شده از لحاظ بیماری آب سیاه تحت آزمون غربالگری قرار گرفتند که معلوم بود دارای ۹۶٪ حساسیت و ۸۹٪ ویژگی است. از میان ۱۵۰ معلم مورد آزمون، ۲۳ نفر مثبت بودند. از روی این داده‌ها مطلوب است برآورد نسبت معلمان دچار آب سیاه در اتحادیه و بازه‌های اطمینان ۹۵٪. برای این نسبت از دانسته‌های خود درباره نمونه‌گیری تصادفی ساده (فصل ۳)، داریم

$$\hat{p} = \frac{23}{150} = 0.153$$

و

$$\begin{aligned} \hat{SE}(\hat{p}) &= \left(\frac{N-n}{N} \right)^{\frac{1}{2}} \left(\frac{\hat{p}(1-\hat{p})}{n-1} \right)^{\frac{1}{2}} \\ &= \left(\frac{2560-150}{2560} \right)^{\frac{1}{2}} \left(\frac{0.153 \times 0.847}{149} \right)^{\frac{1}{2}} \\ &= 0.029 \end{aligned}$$

و از رابطه‌های (۵.۱۴) و (۶.۱۴)، داریم

$$\hat{\pi} = \frac{0.153 + 0.89 - 1}{0.96 + 0.89 - 1} = \frac{0.043}{0.85} = 0.051$$

و

$$\hat{SE}(\hat{\pi}) = \left(\frac{0.029}{0.85} \right) = 0.034$$

سرانجام، بازه‌های اطمینان ۹۵ درصدی برای π ، شیوع نامعلوم آب سیاه در میان همه معلمان در هفت ناحیه آموزشی به صورت زیر است

$$0.051 - 1.96 \times 0.034 \leq \pi \leq 0.051 + 1.96 \times 0.034$$

یا

$$0 \leq \pi \leq 0.118$$

توجه داشته باشید که حد پایین به صورت -0.016 محاسبه شده و برابر با صفر گرفته شده است، زیرا نسبت منفی مفهومی ندارد.

□

روشی که در بالا ارائه شد فرض را بر این می‌گیرد که حساسیت و ویژگی آزمون غربالگری معلوم‌اند. اگر معلوم نباشند می‌توان از یک طرح نمونه‌گیری دوگانه برای برآورد شیوع نامعلوم، π ،

استفاده کرد. در این طرح، زیرنمونه‌ای از آن دسته از افرادی که مثبت غربالگری شده‌اند و زیرنمونه دیگری از آن دسته که منفی غربالگری شده‌اند انتخاب می‌شود. افراد هر دو نمونه مورد آزمایش تشخیصی دقیقتری برای آن بیماری یا شرط خاص قرار می‌گیرند که به عنوان «استاندارد طلایی» برای اثبات یا رد نتایج آزمون غربالگری اولیه مورد استفاده قرار خواهد گرفت. در این نوع طرح، برآورد شیوع، $\hat{\pi}$ ، از فرمول زیر به دست می‌آید

$$\hat{\pi} = \frac{\frac{k_1}{m_1} \times m + \frac{k_2}{m_2} \times (n - m)}{n} \quad (7.14)$$

که در آن

n = تعداد غربالگری شده از جامعه متشکل از N عنصر

m = تعداد مثبت برای آزمون غربالگری

$n-m$ = تعداد منفی برای آزمون غربالگری

m_1 = تعداد نمونه‌گیری شده از گروه متشکل از m فردی که آزمون غربالگری آنها ابتدا مثبت بوده است

k_1 = تعدادی که مثبت بودن آنها از میان m_1 آزمودنی که ابتدا مثبت بوده و مجدداً نمونه‌گیری شده‌اند تأیید شده است

m_2 = تعداد نمونه‌گیری شده از گروه متشکل از $n-m$ فردی که آزمون غربالگری آنها ابتدا منفی بوده است

k_2 = تعداد آزمودنیهایی که از میان m_2 آزمودنی که ابتدا منفی بوده و مجدداً نمونه‌گیری شده‌اند مثبت تشخیص داده شده‌اند

برآوردی از واریانس این برآورد (که با استفاده از واریانس شرطی و نادیده گرفتن تصحیح جامعه

متناهی به دست آمده است) با فرمول زیر محاسبه می‌شود

$$\begin{aligned} \widehat{Var}(\hat{\pi}) = & \left(\frac{k_1}{m_1} - \frac{k_2}{m_2} \right)^2 \times \frac{m \left(1 - \frac{m}{n} \right)}{n} + \left(\frac{m}{n} \right)^2 \frac{\frac{k_1}{m_1} \left(1 - \frac{k_1}{m_1} \right)}{m_1} \\ & + \left(1 - \frac{m}{n} \right)^2 \frac{\frac{k_2}{m_2} \left(1 - \frac{k_2}{m_2} \right)}{m_2} \end{aligned} \quad (8.14)$$

مثال تشریحی: یک نمونه تصادفی ساده متشکل از ۱۲۰ نفر از میان ۴۲۳ نفری که در باجه‌های اخذ

عوارض در سیستم بزرگراههای یک ایالت بزرگ کار می‌کنند گرفته شده است. هر کارمند نمونه تحت

یک آزمون سریع غربالگری قرار می‌گیرد که نقص شنوایی را ارزیابی می‌کند. از میان ۱۲۰ کارمندی که به این ترتیب غربالگری شده‌اند ۶۳ نفر علائمی از کاهش شنوایی نشان داده‌اند. یک نمونه تصادفی ساده ۳۰ نفری از میان این ۶۳ نفری که مثبت غربالگری شده‌اند انتخاب شده و مورد آزمایش دقیقتری برای سنجش کاهش شنوایی قرار گرفته‌اند. از این ۳۰ نفر در ۲۳ نفر کاهش شنوایی تأیید شده است. به همین ترتیب، نمونه‌ای ۳۰ نفری از میان ۵۷ نفری که در ابتدا منفی غربالگری شده‌اند گرفته شده و ۱۴ نفر از آنها نیز هنگامی که مورد آزمایش دقیقتر قرار گرفتند نقص شنوایی داشتند. از معادله (۷.۱۴) داریم

$$\begin{aligned} n &= 120 & m &= 63 & n - m &= 57 \\ m_1 &= m_2 = 30 & k_1 &= 23 & k_2 &= 14 \\ \hat{\pi} &= \frac{\frac{23}{30} \times 63 + \frac{14}{30} \times 57}{120} \end{aligned}$$

از رابطه (۸.۱۴)، خطای معیار $\hat{\pi}$ به صورت زیر برآورد می‌شود

$$\begin{aligned} \hat{SE}(\hat{\pi}) &= \left\{ \left(\frac{23}{30} - \frac{14}{30} \right)^2 \times \frac{0.525 \times 0.475}{120} \right. \\ &\quad \left. + (0.525)^2 \times \frac{0.767 \times 0.233}{30} + (0.475)^2 \times \frac{0.467 \times 0.533}{30} \right\}^{1/2} = 0.061 \end{aligned}$$

به این ترتیب، بازه‌های اطمینان ۹۵ درصدی برای نسبت π ، مبتلایان به بیماری یا شرطی خاص که به منظور آن غربالگری شده‌اند به صورت زیر است

$$0.62 - 1.96 \times 0.061 \leq \pi \leq 0.62 + 1.96 \times 0.061$$

یا

$$0.50 \leq \pi \leq 0.74$$

□

مثال بالا اهمیت نمونه‌گیری دوگانه را هنگامی که حساسیت و ویژگی آزمون غربالگری نامعلوم‌اند نشان می‌دهد. اگر نسبت مثبت غربالگری آغازین به کار رفته بود، برآورد به دست آمده، یعنی $\frac{63}{120} = 0.525$ ، به صورتی قابل ملاحظه کمتر از برآوردی می‌شد که در رابطه (۷.۱۴) ارائه شده است.

در بررسیهای غربالگری غالباً نه تنها برآورد شیوع بیماری یا شرطی خاص، بلکه برآورد حساسیت و ویژگی آزمون غربالگری خاصی نیز موردنظر است. این کار در صورتی قابل اجراست که نمونه‌ها هم از آنهایی که در ابتدا منفی غربالگری شده‌اند و هم از آنها که ابتدا مثبت غربالگری شده‌اند انتخاب شوند. برآوردهای مناسب، \hat{S}_e و \hat{S}_p ، برای حساسیت و ویژگی به صورت زیر به دست می‌آیند

$$\hat{S}_e = \frac{m \frac{k_1}{m_1}}{m \left(\frac{k_1}{m_1} - \frac{k_2}{m_2} \right) + n \frac{k_2}{m_2}} \quad (9.14)$$

و

$$\hat{S}_p = \frac{(n-m) \left(1 - \frac{k_2}{m_2} \right)}{m \left(\frac{k_2}{m_2} - \frac{k_1}{m_1} \right) + n \left(1 - \frac{k_2}{m_2} \right)} \quad (10.14)$$

۵.۱۴ برآورد کردن پیشامدهای نادر: نمونه‌گیری شبکه‌ای

نشان دادیم که تعدیل برنامه نمونه‌گیری تصادفی ساده پایه، بدون تغییر در شیوه برآورد کردن (مثلاً نمونه‌گیری تصادفی طبقه‌بندی شده با تخصیص متناسب) گاهی می‌تواند به برآوردهایی منجر شود که خطای نمونه‌گیری آنها کمتر از برآوردهایی است که از طریق نمونه‌گیری تصادفی ساده به دست آمده‌اند. همچنین بحث کردیم که چگونه یک برنامه نمونه‌گیری تصادفی ساده که از شیوه برآورد کردن تعدیل شده استفاده می‌کند (مثلاً برآورد نسبتی) می‌تواند به برآوردهایی با خطای نمونه‌گیری کمتر منتهی شود. به راهی مشابه، در این بخش، نشان خواهیم داد که چگونه تعدیل قاعده شمارش می‌تواند به برآوردهایی منجر شود که خطاهای نمونه‌گیری آنها کمتر از آن است که از طریق قواعد شمارش مرسوم به دست می‌آید.

قاعده شمارش، الگوریتمی است که واحدهای شمارش (یا واحدهای فهرست‌برداری) را به واحدهای اولیه ربط می‌دهد. مثلاً در یک آمارگیری نمونه‌ای از خانوارها که برای برآورد کردن کل نوزادانی اجرا می‌شود که در یک جامعه طی دوره زمانی مشخصی متولد می‌شوند، قاعده واضح شمارش آن است که گزارش کردن تولد به وسیله خانوار والدین نوزاد میسر باشد. چنین قاعده‌ای برای شمارش، هر عنصر (مثلاً تولد) را به فقط و فقط یک واحد شمارش (مثلاً خانوار) ربط می‌دهد. ولی یک قاعده شمارش دیگر این امکان را فراهم می‌سازد که گزارش تولد نوزادان به وسیله خانوار

پدربزرگ و مادربزرگ نوزاد نیز مانند خانوار والدین میسر باشد. در آن صورت این قاعده شمارش اجازه می‌دهد که یک عنصر به بیش از یک واحد فهرست‌برداری ربط داده شود.

قاعده شمارشی که ارتباط یک عنصر را به تنها یک واحد شمارش میسر می‌سازد قاعده شمارش مرسوم نامیده می‌شود. قاعده شمارشی که ارتباط یک عنصر را به بیش از یک واحد شمارش میسر می‌سازد قاعده شمارش چندبارگی نامیده می‌شود. طرح‌های نمونه‌ای که از قواعد شمارش چندبارگی استفاده می‌کنند نمونه‌های شبکه‌ای نامیده می‌شوند. نمونه‌های شبکه‌ای طی دو دهه گذشته مورد توجه قابل ملاحظه‌ای قرار گرفته‌اند، به خصوص در علوم رفتاری و بهداشتی، و به خصوص در وضعیت‌هایی که متضمن پیشامدهای نادر یا مشخصه‌های کمیاب‌اند. این طرح را با مثال زیر بررسی می‌کنیم.

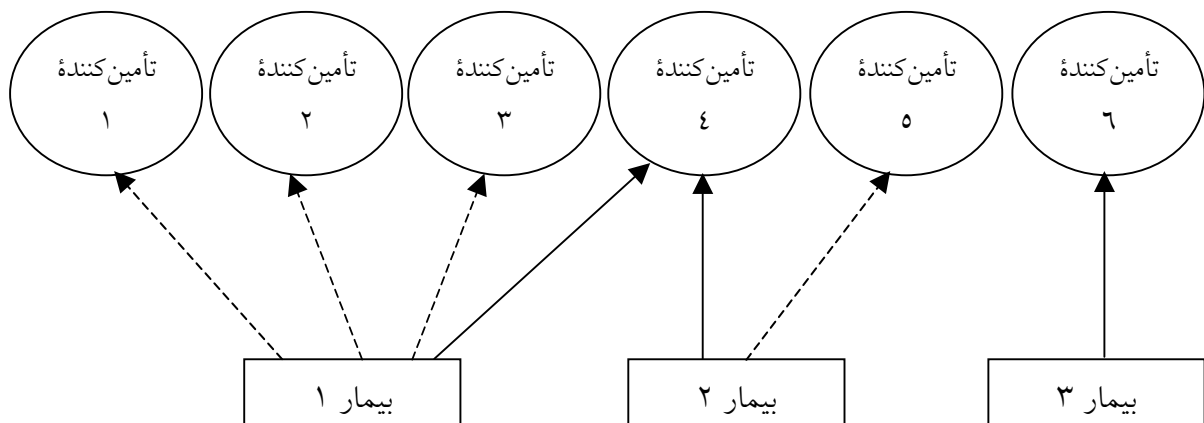
مثال تشریحی: بخشی را در نظر می‌گیریم که دارای ۱۰۰ تأمین‌کننده شناسایی شده مراقبت‌های اولیه بهداشتی است. فرض کنید می‌خواهیم یک آمارگیری نمونه‌ای با ده تأمین‌کننده مراقبت‌های بهداشتی اجرا کنیم تا کل تعداد اشخاصی را که طی یک دوره زمانی خاص تحت درمان سرطان پوست بوده‌اند برآورد کنیم. اگر هر بیمار فقط و فقط توسط یک تأمین‌کننده مراقبت‌های بهداشتی تحت درمان قرار گرفته بود، فرایند مزبور فرایند ساده‌ای بود. ولی، بیماری که سرطان پوست دارد توسط چند تأمین‌کننده مراقبت‌های بهداشتی درمان می‌شود (مانند متخصص داخلی، جراح، پرتوشناس، تن‌درمانگر). به این ترتیب، در طراحی آمارگیری نمونه‌ای باید قاعده‌ای برای شمارش تعیین کنیم که ارتباط عناصر را (یعنی اشخاصی را که تحت درمان سرطان پوست بوده‌اند) با بیش از یک واحد شمارش (یعنی تأمین‌کننده مراقبت‌های بهداشتی) میسر سازد.

فرض می‌کنیم که در طول سال سه نفر تحت درمان سرطان پوست قرار گرفته‌اند. نفر ۱ با چهار پزشک (پزشکان ۱، ۲، ۳ و ۴)، نفر ۲ با دو پزشک (۴ و ۵) و نفر ۳ با یک پزشک (پزشک ۶) درمان می‌شده است. به این ترتیب شبکه‌ای داریم که تأمین‌کنندگان مراقبت‌های بهداشتی را با بیماران مبتلا به سرطان پوست ربط می‌دهد. این شبکه در شکل ۱.۱۴ نشان داده شده است. خط‌های توپر این شکل، اولین تأمین‌کننده را (از نظر ترتیب زمانی) نشان می‌دهند که بیمار را تحت درمان قرار داده است.

اگر از یک قاعده شمارش مرسوم استفاده کنیم - مثل ربط دادن بیماران به اولین تأمین‌کننده‌ای که آنان را تحت درمان گرفته است - در آن صورت تأمین‌کننده ۴، بیماران ۱ و ۲ را گزارش خواهد داد و تأمین‌کننده ۶ بیمار ۳ را گزارش خواهد کرد. ۹۸ تأمین‌کننده دیگر هیچ بیماری را گزارش نخواهند کرد. اگر X_i را تعداد بیماران گزارش شده توسط i امین تأمین‌کننده براساس این قاعده شمارش در نظر بگیریم، در آن صورت $X_4 = 2$ ، $X_6 = 1$ و بقیه $X_i = 0$ اند. برای یک نمونه تصادفی ساده متشکل

از $n = 10$ تأمین‌کننده مراقبتهای بهداشتی، می‌بینیم که برآوردگر $x' = \left(\frac{100}{10}\right)x$ با شمارش همه نمونه‌ها دارای توزیع زیر است

فرآوانی نسبی	x'
۰/۸۰۹۱	۰
۰/۰۹۰۹	۱۰
۰/۰۹۰۹	۲۰
۰/۰۰۹۱	۳۰



شکل ۱.۱۴ نمونه شبکه‌ای برای تأمین‌کنندگان مراقبتهای بهداشتی و بیماران مبتلا به سرطان پوست

میانگین توزیع x' یعنی $E(x')$ ، برابر است با ۳ (مجموع واقعی جامعه) و خطای معیار، $SE(x')$ ، برابر است با $\sqrt{6/81}$. توجه کنید که در بیش از ۸۰٪ نمونه‌ها هیچ بیماری شناسایی نخواهد شد و برآورد، x' ، برابر با صفر خواهد بود.

□

حالا، با مراجعه به مثال بالا فرض کنید از یک قاعده شمارش استفاده کنیم که گزارش یک بیمار را با هر تأمین‌کننده بهداشتی که آن بیمار را درمان می‌کرده است میسر سازد. این یک قاعده چندبارگی است، زیرا گزارش بیمار توسط بیش از یک تأمین‌کننده مقدور می‌شود. از آنجا که بیمارانی که با بیش از یک تأمین‌کننده تحت درمان قرار گرفته‌اند بیشتر از آنهایی که تنها تحت درمان یک تأمین‌کننده بوده‌اند شانس انتخاب شدن دارند، باید از یک شیوه برآورد کردن تعدیل شده برای به دست

آوردن یک برآوردگر ناریب از مجموع استفاده شود. این کار با تعریف متغیر زیر، x_i^* ، برای هر واحد شمارش (مثلاً تأمین کننده مراقبت‌های بهداشتی) انجام می‌پذیرد که از رابطه زیر به دست می‌آید

$$x_i^* = \sum_{j=1}^m \frac{\delta_{ij}}{s_j}$$

و در آن

m = کل تعداد عناصر (بیماران) شناسایی شده در نمونه

۱، اگر i امین واحد شمارش (تأمین کننده i) با قاعده شمردن به j امین عنصر ربط داده شده باشد
 ۰، در غیر این صورت } = δ_{ij}

s_j = تعداد واحدهای شمارش ربط داده شده به عنصر j (که چندبارگی عنصر نامیده می‌شود)

s_j نشان‌دهنده اطلاعات اضافی است که باید (در صورت امکان) از واحدهای شمارش نمونه یا، در برخی موارد، از بعضی منابع دیگر مانند خود عناصر به دست آید. تذکر می‌دهیم که این فرایند به دست آوردن چندبارگیها ممکن است به صورتی قابل ملاحظه هزینه آمارگیری را افزایش دهد.

برای نمونه‌های تصادفی ساده از تأمین کنندگان مراقبت‌های بهداشتی، از جامعه‌ای متشکل از N تأمین کننده از این نوع، برآورد مجموع، x'_{mult} ، به صورت زیر داده می‌شود

$$x'_{mult} = \left(\frac{N}{n} \right) x^*$$

که در آن x^* مجموع نمونه‌ای x_i^* است. به مثال خود برمی‌گردیم.

مثال تشریحی: با مراجعه به مثال قبل، از یک قاعده شمارش استفاده می‌کنیم که گزارش بیمار را با هر تأمین کننده مراقبت‌های بهداشتی که بیمار را درمان می‌کرده است میسر می‌سازد. پس، برای نمونه‌ای که در شکل ۱.۱۴ نشان داده شده است، داریم

$$\begin{array}{cccccc} s_1 = 4 & s_2 = 2 & s_3 = 1 & & & \\ x_1^* = \frac{1}{4} & x_2^* = \frac{1}{4} & x_3^* = \frac{1}{4} & x_4^* = \frac{1}{4} + \frac{1}{2} = \frac{3}{4} & x_5^* = \frac{1}{2} & x_6^* = 1 \end{array}$$

جدول ۱.۱۴ توزیع x'_{mult} برای نمونه‌های تصادفی ساده $n=10$ از $N=100$ تأمین کننده مراقبت‌های بهداشتی را در این مثال نشان می‌دهد.

میانگین، $E(x'_{mult})$ ، و خطای معیار، $SE(x'_{mult})$ ، که از روی توزیعی که در جدول ۱.۱۴ نشان داده شده (و با استفاده از عبارتهای ۲.۲ ارائه شده در تابلوی ۴.۲) محاسبه شده‌اند چنین‌اند

$$SE(x'_{mult}) = ۴/۱۳ \quad \text{و} \quad E(x'_{mult}) = ۳ = X$$

به این ترتیب می‌بینیم که برای این مثال، برآورد مجموع، x'_{mult} ، نه تنها برآوردگری ناریب برای مجموع جامعه‌ای X است بلکه در این مورد دارای خطای معیاری به مراتب کمتر از x' ، برآورد مجموع به دست آمده از قاعده شمارش مرسوم است.

□

نظریه نمونه‌گیری شبکه‌ای با جزئیات بیشتر در مقاله‌های سیرکن [۱۱]، [۱۲] و سودمان، سیرن و کووان [۱۳] مورد بحث قرار گرفته است. نشان داده شده است که این نظریه به خصوص در برآورد وقوع یا شیوع صفات کیفی که در کمتر از ۳٪ جامعه روی می‌دهند (یعنی «پیشامدهای نادر») سودمند است. مثلاً برای برآورد تعداد تولد و ازدواج [۱۴] و قربانیان جرایم [۱۵] که در یک دوره زمانی معین روی داده‌اند و نیز شیوع بیماریهایی از قبیل فیبروز کیستی [۱۲]، سرطان [۱۶] و دیابت [۱۷] به کار رفته است. نظریه مزبور به طبقه‌بندی [۱۲]، [۱۸]، برآورد نسبی [۱۹]، و نمونه‌گیری خوشه‌ای [۲۰] نیز گسترش یافته است.

قواعد شمارش چندبارگی می‌توانند قابلیت اعتماد برآورد را بهبود بخشند زیرا غالباً عملکرد نمونه را افزایش می‌دهند. مثلاً در مثالی که در بالا مورد استفاده قرار گرفت، تقریباً ۸۱٪ همه نمونه‌های دارای

جدول ۱.۱۴ توزیع x'_{mult}

فرآوانی نسبی (الف)	x'_{mult}
۰/۵۲۲۳	۰/۰
۰/۱۸۴۳	۲/۵
۰/۰۸۰۷	۵/۰
۰/۰۸۱۳۲	۷/۵
۰/۰۸۲۵۱	۱۰/۰
۰/۰۳۰۷۶	۱۲/۵
۰/۰۱۰۰۳	۱۵/۰
۰/۰۰۸۳۹	۱۷/۵
۰/۰۰۱۹۲	۲۰/۰
۰/۰۰۰۷۴	۲۲/۵
۰/۰۰۰۱۹	۲۵/۰
۰/۰۰۰۰۱	۲۷/۵
۰/۰۰۰۰۰۳	۳۰/۰

الف. به دست آمده از نظریه ترکیبانی

اندازه $n=10$ در صورت استفاده از قاعده شمارش مرسوم، هیچ بیمار دچار سرطان پوست را نشان نمی‌دادند در حالی که تحت قاعده چندبارگی، فقط در حدود ۵۲٪ از همه نمونه‌ها ممکن است هیچ بیمار دچار سرطان پوست را نشان ندهند. تحت قاعده چندبارگی، شبکه تأمین کنندگان مراقبت‌های بهداشتی که واجد شرایط برای گزارش بیماران مبتلا به سرطان پوست بوده‌اند از ۲ به ۶ افزایش یافته‌اند (ن. ک. شکل ۱.۱۴).

۶.۱۴ برآورد کردن پیشامدهای نادر: نمونه‌گیری دوگانه

راهبرد دیگری که می‌تواند گاهی برای برآورد فراوانی نسبی پیشامدهای نادر یا اعضای جوامع نادر به کار رود استفاده از دو آمارگیری مستقل یا ترکیبی از یک آمارگیری و سیستم ثبتي است. برآورد کردن از این نوع سیستم دوگانه به شرح زیر است. N را تعداد نامعلوم پیشامدهای نادر در جامعه در نظر می‌گیریم، $x_1 =$ کل تعداد پیشامدهای نادر شناسایی شده در اولین آمارگیری، $x_2 =$ کل تعداد پیشامدهای نادر شناسایی شده در دومین آمارگیری، و $x_{12} =$ تعداد پیشامدهای نادر شناسایی شده در هر دو آمارگیری. اینها را می‌توان به صورت زیر نشان داد.

$$\begin{array}{c}
 \text{آمارگیری ۲} \\
 \text{آری} \quad \text{نه} \\
 \begin{array}{|c|} \hline x_{12} \\ \hline \end{array} \\
 \begin{array}{c} x_1 \quad \text{آری} \\ \quad \quad \text{نه} \end{array} \\
 \hat{N} \quad x_2
 \end{array}
 \quad \text{آمارگیری ۱}$$

اگر دو آمارگیری مستقل باشند، در آن صورت x_{12} تعداد پیشامدهای نادر موجود در هر دو آمارگیری تقریباً برابر با $\frac{x_1 x_2}{N}$ خواهد بود و N می‌تواند با استفاده از \hat{N} برآورد شود که از فرمول زیر به دست می‌آید

$$\hat{N} = \frac{x_1 x_2}{x_{12}} \quad (9.14)$$

این برآورد در اصل به وسیله چاندرا سکار و دمینگ [۲۱] ابداع شد که از آن برای برآورد وقایع حیاتی در کشورهای در حال توسعه استفاده کردند. در این کاربرد از برآورد مزبور، دو مأخذ گزارش دهنده وقایع حیاتی از یک سیستم ثبتي و یک آمارگیری تشکیل شده است. اخیراً، کووان و مالک،

روش مذکور را به وضعیتهایی پیچیده‌تر گسترش داده‌اند که در آنها وقایعی که در هیچ یک از دو سیستم گنجانیده نشده‌اند خوشه‌بندی می‌شوند [۲۲].

مثال تشریحی: فرض کنید سیستم ثبتی وقایع حیاتی یک کشور در حال توسعه، ۲۰۵۴۳ مورد فوت را طی یک سال معین شمارش کرده است و یک آمارگیری نمونه‌ای متشکل از یکی از هر ۱۰۰ خانوار، تعداد ۱۸۵ مورد فوت را شمارش کرده که ۱۴۶ مورد آن با مرگ و میر فهرست شده در آمارهای ثبتی مطابقت دارد.

با استفاده از برآوردگری که در رابطه (۹.۱۴) نشان داده شده است، داریم

$$x_1 = 20543$$

$$x_2 = 100 \times 185 = 18500$$

$$x_{12} = 100 \times 146 = 14600$$

$$\hat{N} = \frac{20543 \times 18500}{14600} \approx 26031$$

به این ترتیب، برآورد می‌شود که در طی مدت موردنظر ۲۶۰۳۱ مورد مرگ و میر در آن کشور روی داده است.

۷.۱۴ برآورد کردن مشخصه‌هایی برای مناطق محلی: برآورد کردن ترکیبی

برنامه‌ریزی خدمات اجتماعی و بهداشتی در آمریکا غالباً بر مبنای جغرافیایی نیمه سراسری انجام می‌پذیرد که طی آن مناطقی از قبیل گروههایی از ایالتها، ایالت‌های تکی، بخشها، و شهرداریها به عنوان «واحدهای تحلیل» عملی به خدمت گرفته می‌شوند. به این دلیل، بسیار مهم است که برآوردهایی معتبر و قابل اعتماد از مشخصه‌های اجتماعی و بهداشتی این مناطق محلی موجود باشد تا تصمیمهای مربوط به برنامه‌ریزی براساس اطلاعات صحیح گرفته شوند.

از سوی دیگر، این قبیل برآوردها برای مناطق محلی غالباً به آسانی به دست نمی‌آیند. اگرچه برآوردهایی از تولد، وفات، ازدواج و طلاق را می‌توان برای مناطق محلی از طریق نظام ثبتی آمارهای حیاتی به دست آورد، برآوردهای مربوط به ناخوشی، از کار افتادگی، و استفاده از خدمات معمولاً براساس محل موجود نیستند. علت اصلی موجود نبودن داده‌های بهداشتی در سطح مناطق محلی در این واقعیت نهفته است که اکثر داده‌های بهداشتی از یک نظام آمارگیری سراسری جمع‌آوری می‌شوند که مرکز ملی آمارهای بهداشتی^۱ آن را اجرا می‌کند. این نظام آمارهای بهداشتی می‌تواند برآوردهایی معتبر و قابل اعتماد برای سراسر آمریکا و نیز برای زیرمنطقه‌های جغرافیایی بزرگ

^۱ National Center for Health Statistics (NCHS)

(یعنی شمال شرق، شمال مرکزی، جنوب، غرب) فراهم کند. ولی محدودیتهای مربوط به اندازه نمونه و طراحی این آمارگیریه‌ها به طور کلی آنها را در تولید برآوردهایی دقیق برای نواحی نسبتاً کوچک که غالباً به عنوان واحدهای برنامه‌ریزی به کار می‌روند نامناسب می‌سازد.

مشخصه اصلی طرحهای این آمارگیریه‌های سراسری که باعث می‌شود آمارگیریه‌های مزبور برای تولید برآوردهای نواحی کوچک مناسب نباشند راه ساختن طبقه‌هاست. مثلاً در آمارگیری مصاحبه‌ای بهداشت ملی^۱ که یک آمارگیری مصاحبه‌ای سراسری از خانوار است که برآوردهایی برای ناخوشی، از کارافتادگی، و استفاده از خدمات بهداشتی به دست می‌دهد، طبقه‌ها از یک یا چند بخش یا نواحی آماری کلانشهری استاندارد^۲ تشکیل شده‌اند که براساس مشخصه‌های جمعیت‌شناختی مشابه گروه‌بندی شده‌اند. مثلاً، یک طبقه از آمارگیری مصاحبه‌ای بهداشت ملی می‌تواند شامل سه بخش باشد: یکی از ایالت آیووا، یکی از نبراسکا و یکی از کانزاس. در نمونه، تنها یکی از این سه بخش انتخاب خواهد شد تا معرف کل طبقه باشد. برآوردهای مربوط به آمریکا با انبوه‌سازی این قبیل برآوردهای طبقه‌ای از همه طبقه‌ها به دست می‌آید. ولی، کنار هم گذاشتن برآوردهای مربوط به مناطقی از قبیل ایالتها، شهرها، یا نواحی برنامه‌ریزی بهداشتی که عموماً ترکیباتی از طبقه‌های آمارگیری مصاحبه‌ای بهداشت ملی نیستند در واقع امکان‌پذیر نیست.

یک روش استفاده از داده‌های حاصل از آمارگیریه‌های سراسری بهداشت برای مقاصد به دست آوردن برآوردهایی برای نواحی کوچک که عمدتاً - صرفاً به علت سادگی و جاذبه شهودی آن - مورد قبول واقع شده است برآورد کردن ترکیبی نامیده می‌شود. با این شیوه، برآوردهای مشخصه‌های نواحی کوچک از ترکیب برآوردهای سراسری مشخصه‌های ویژه گروه‌های جامعه‌ای با برآوردهای توزیع متناسب درون جامعه ناحیه کوچک در همان گروه‌های جامعه‌ای به دست می‌آیند. گروه‌ها هم براساس ارتباط با مشخصه مورد برآورد و هم براساس موجود بودن داده‌های جامعه‌ای ناحیه کوچک ویژه آن گروه انتخاب می‌شوند.

به صورت رسمیت، برآورد ترکیبی \tilde{x}_a از سطح میانگین \bar{X}_a ی مشخصه X برای ناحیه a از فرمول زیر به دست می‌آید

$$\tilde{x}_a = \sum_{k=1}^K \hat{P}_{ak} \bar{x}_k \quad (۱۰.۱۴)$$

که در آن \bar{x}_k برای $k = 1, \dots, K$ ، برآورد سراسری سطح میانگین مشخصه X برای افراد گروه k است که از آمارگیری سراسری به دست آمده است، و \hat{P}_{ak} برآورد نسبت همه افراد ناحیه a است که در گروه k

^۱ National Health Interview Survey (NHIS)

^۲ Standard Metropolitan Statistical Areas (SMSAs)

هستند. \hat{P}_{ak} ها برآوردهایی از نسبت‌های جامعه‌ای محلی هستند که عموماً از یک سرشماری آمریکا یا از نمایندگیهای محلی به دست آمده‌اند؛ K کل تعداد گروه‌های جامعه‌ای را نشان می‌دهد که برای بررسی انتخاب شده‌اند.

می‌بینیم که برآورد ترکیبی \tilde{x}_a برخی مشخصه‌هایی را دارد که بسیار شبیه به برآوردهایی هستند که در بحث پس طبقه‌بندی در فصل ۶ بسط داده شدند (هر چند تفاوت‌های مهمی نیز وجود دارند). شباهت به پس طبقه‌بندی در این واقعیت نهفته است که هر دو روش میانگینهای ویژه گروه‌هایی را که طبقه نیستند با نسبت‌های جامعه‌ای مناسب برای آن گروه‌ها ترکیب می‌کنند. تفاوت عمده در این واقعیت نهفته است که میانگینهای ویژه گروهی \bar{x}_k که در ساخت برآورد ترکیبی مورد استفاده قرار می‌گیرند خود از روی داده‌هایی ساخته می‌شوند که در بیشتر موارد از افرادی خارج از ناحیه کوچک به دست آمده‌اند.

همان طور که قبلاً گفته شد، برآورد ترکیبی تا حدودی مورد قبول واقع شده است زیرا از لحاظ منابع و هزینه‌ها جایگزینی عملیتر برای اجرای آمارگیری در ناحیه کوچکی است که برآوردهایی برای آن موردنظر است. ویژگیهای آماری برآوردهای ترکیبی توسط گونزالز و هوزا [۲۳] و لهوی و فرنچ [۲۴] بررسی شده‌اند. این برآوردها/ریب‌اند و اریبی آنها از فرمول زیر به دست می‌آید

$$B(\tilde{x}) = \sum_{k=1}^K \hat{P}_{ak} (\bar{X}_k - \bar{X}_{ak}) \quad (11.14)$$

که در آن، \bar{X}_k برای $k = 1, 2, \dots, K$ سطح میانگین مشخصه x در سطح کشور برای گروه k و \bar{X}_{ak} سطح میانگین مشخصه x برای افراد ناحیه a است که در گروه k قرار دارند.

عبارتی برای واریانس مناسب یک برآورد ترکیبی به صورت زیر به دست می‌آید

$$Var(\tilde{x}) \approx \sum_{k=1}^K P_{ak}^2 Var(\bar{x}_k) + 2 \sum_{k < r} P_{ak} P_{ar} Cov(\bar{x}_k, \bar{x}_r) \quad (12.14)$$

که در آن

$$Var(\bar{x}_k) = \text{واریانس برآورد سطح میانگین } x \text{ در گروه } k \text{ (سراسری)}$$

$$Cov(\bar{x}_k, \bar{x}_r) = \text{کوواریانس بین برآورد سطوح میانگین مشخصه } x \text{ در گروه‌های } k \text{ و } r \text{ (سراسری)}$$

چون برآوردهای ترکیبی عموماً بر مبنای نمونه‌های بزرگ‌اند، خطاهای نمونه‌گیری آنها غالباً کم است و درستی آنها تا حدود زیادی به بزرگی اریبی آنها بستگی دارد. توجه داشته باشید که اریبی در برآورد ترکیبی، متوسطی موزون از تفاوت‌های موجود بین میانگین مربوط به یگ گروه در سطح سراسری و میانگین مربوط به همان گروه در ناحیه کوچک است.

اگر یک برآورد نارایب \bar{x}'_a برای ناحیه کوچک وجود داشته و از نظر آماری مستقل از برآورد ترکیبی باشد، در آن صورت میانگین توان دوم خطای برآورد ترکیبی را می‌توان به صورت زیر برآورد کرد

$$MSE(\tilde{x}_a) = (\tilde{x}_a - \bar{x}'_a)^2 - Var(\bar{x}'_a) \quad (۱۳.۱۴)$$

که در آن، $Var(\bar{x}'_a)$ برآوردی نارایب از واریانس \bar{x}'_a است. به طور کلی، واریانس برآورد نارایب موجود، \bar{x}'_a ، برای معادله (۱۳.۱۴) ناپایدارتر از آن است که مفید فایده‌ای باشد (در غیر این صورت باید به جای برآورد ترکیبی، \bar{x}'_a برآورد منتخب می‌بود). ولی، این برآوردهای میانگین توان دوم خطای برآوردهای ترکیبی تکی می‌توانند با تعدادی عبارتهای مشابه در میان چندین ناحیه کوچک متوسط‌گیری شوند تا برآورد پایدارتری به دست آید که دارای تعبیر متوسط میانگین توان دوم خطای مجموعه‌ای از برآوردهای ترکیبی (مانند ایالتها در داخل کشور آمریکا) است.

مثال تشریحی: فرض کنید می‌خواهیم میانگین تعداد روزهای کاری تلف شده به علت بیماری را برای ایالتی بزرگ برآورد کنیم و فرض کنید که برآوردهای سراسری زیر برای هر یک از چهار حوزه نژادی - جنسی همراه با توزیع جمعیت ایالت بر حسب چهار گروه نژادی - جنسی موجودند.

گروه نژادی - جنسی	نسبت جمعیت ایالت در گروه	برآورد روزهای کاری تلف شده در سطح کشور به ازای هر نفر در سال (\bar{x}_k)
سفیدپوست مرد	۰/۳۹	۵/۲
سفیدپوست زن	۰/۴۳	۵/۶
همه مردان دیگر	۰/۰۸	۵/۹
همه زنان دیگر	۰/۱۰	۶/۳

از رابطه (۱۳.۱۴)، برآورد ترکیبی ایالت به صورت زیر به دست می‌آید

$$\tilde{x}_a = ۰/۳۹ \times ۵/۲ + ۰/۴۳ \times ۵/۶ + ۰/۰۸ \times ۵/۹ + ۰/۱۰ \times ۶/۳ = ۵/۵۴ \quad \text{روز به ازای هر نفر در سال}$$

فرض کنید برآوردی نارایب برای ایالت موجود و برابر با $۶/۸۵$ با برآورد واریانس $۷۶/۲۹$ است. پس برآورد میانگین توان دوم خطای برآوردهای ترکیبی به صورت زیر به دست می‌آید

$$MSE(\tilde{x}_a) = (۶/۸۵ - ۵/۵۴)^2 + ۷۶/۲۹ = ۷۸/۰۱$$

و برآورد ریشه میانگین توان دوم خطا $\sqrt{۷۸/۰۱} = ۸/۸۳$ خواهد بود.

□

برآورد کردن ترکیبی یکی از روشهای متعدد برای به دست آوردن برآوردهایی از مشخصه‌های نواحی کوچک است. از مدل‌های رگرسیونی به صورتی روزافزون استفاده شده است. این روش عموماً به وسیله معادله‌های رگرسیونی که روابط بین برآورد یک متغیر برآمد موردنظر و مجموعه‌ای از متغیرهای نشانگر موجود برای نواحی محلی را بیان می‌کنند توسعه می‌یابد. سپس این معادله‌های رگرسیونی در تهیه برآوردی بهبود یافته از متغیر برآمد موردنظر برای ناحیه محلی بر مبنای متغیرهای نشانگر موجود برای ناحیه محلی مورد استفاده قرار می‌گیرند. بحث کاملتر درباره این نوع روش‌شناسی و سایر روشهای به دست آوردن برآوردهای ناحیه کوچک را می‌توان در مروری توسط له‌وی [۲۵] یافت.

۸.۱۴ استخراج اطلاعات حساس: فنون پاسخ تصادفی شده

در بسیاری از آمارگیریهای نمونه‌ای باید اطلاعاتی را که دارای ماهیت حساس‌اند از افرادی که در نمونه انتخاب شده‌اند استخراج کرد. مثلاً در بررسیهای مربوط به کاربستهای برنامه تنظیم خانواده ممکن است پرسیدن سؤالاتی درباره استفاده از داروها، کاربستهای جلوگیری از بارداری، یا تاریخچه سقط جنین ضروری باشد. این قبیل سؤالات ممکن است برای بعضی از افراد ترساننده یا عذاب‌آور باشد. برای اجتناب از بی‌پاسخی بیش از حد یا پاسخهای گمراه کننده، روشی به نام پاسخ تصادفی شده در اصل توسط وارنر [۲۶] ابداع شد که با قدری موفقیت در بسیاری از آمارگیریهای مورد استفاده قرار گرفته است.

فن پاسخ تصادفی شده در اساسی‌ترین شکل آن دو سؤال را برای پاسخگو مطرح می‌کند: یک سؤال حساس و یک سؤال بی‌گزند. سپس یک ابزار تصادفی‌سازی از قبیل کیسه‌ای محتوی مهره‌های قرمز و سفید به پاسخگو داده می‌شود و از او درخواست می‌شود تا یک مهره از کیسه خارج کند بدون این که مصاحبه‌گر آن را ببیند. اگر مهره قرمز انتخاب شد به پاسخگو گفته می‌شود که به سؤال بی‌گزند با «آری» یا «نه» پاسخ دهد. اگر مهره سفید انتخاب شد پاسخگو باید به سؤال حساس با «آری» یا «نه» پاسخ دهد. سپس مصاحبه‌گر جواب پاسخگو را یادداشت می‌کند بدون این که بداند پاسخ، مربوط به کدام سؤال بوده است.

منطقی که در پس این روش نهفته آن است که اگر پاسخگو احساس کند که مصاحبه‌گر نمی‌داند که پاسخ مربوط به سؤال حساس یا سؤال بی‌گزند است علاقمند خواهد بود که به سؤال حساس پاسخ دهد. برای تشریح مطلب، ممکن است کارت زیر در اختیار پاسخگو قرار داده شود:

۱. آیا هرگز به عنوان تفریح کوکابین مصرف کرده‌اید؟

۲. آیا آخرین رقم شماره تأمین اجتماعی شما یک عدد فرد است (۱، ۳، ۵، ۷، ۹)؟

پس از این که به پاسخگو گفته شد در صورت انتخاب مهره سفید به سؤال ۱ و در صورت انتخاب مهره قرمز به سؤال ۲ جواب دهد، پاسخگو با «آری» یا «نه» پاسخ می‌دهد بدون این که به مصاحبه‌گر بگوید که جواب کدام سؤال را داده است.

اگر هم ترکیب ابزار تصادفی‌سازی، و هم احتمال نظری پاسخ «آری» به سؤال بی‌گزند معلوم باشند می‌توان برآوردی از نسبت افراد جامعه را که دارای صفت کیفی تعیین شده در سؤال حساس بوده‌اند از روی کل نسبت پاسخگوییانی که برای آنها پاسخ «آری» ثبت شده است به دست آورد. این برآورد از فرمول زیر به دست می‌آید

$$\hat{P}_1 = \frac{P^* - P_1(1-\theta)}{\theta} \quad (14.14)$$

که در آن \hat{P}_1 برآورد نسبت کسانی در جامعه است که صفت تعیین شده در سؤال حساس را دارند؛ P^* نسبت پاسخهای «آری» حاصل از آمارگیری است؛ P_1 احتمال نظری پاسخ «آری» به سؤال بی‌گزند است؛ و θ احتمال پاسخگویی به سؤال حساس (مثلاً نسبت مهره‌های قرمز در کیسه) است.

مثال تشریحی: در مثالی که در بالا ارائه شد، فرض کنید که کیسه حاوی ۷۵٪ مهره‌های سفید است و احتمال فرد بودن آخرین رقم شماره تأمین اجتماعی هر شخص برابر با احتمال زوج بودن آن است. پس θ و P_1 به صورت زیر به دست می‌آیند

$$\theta = \frac{3}{4} = 0.75 \quad \text{و} \quad P_1 = \frac{1}{4} = 0.25$$

فرض کنید ۴۰٪ از همه پاسخگویان جواب «آری» داده‌اند. به عبارت دیگر، $P^* = 0.40$. پس \hat{P}_1 ، برآورد نسبت اشخاصی که به عنوان تفریح کوکابین مصرف می‌کنند به صورت زیر به دست می‌آید

$$\hat{P}_1 = \frac{0.40 - 0.25 \times (1 - 0.75)}{0.75} = 0.367$$

به این ترتیب، برآورد می‌کنیم که ۳۶.۷٪ از همه افراد موجود در این جامعه برای تفریح کوکابین مصرف کرده‌اند.

□

چون پاسخهای سؤال بی‌گزند به مفهومی هدر رفتن اطلاعات است، فنون پاسخ تصادفی شده عموماً نسبت به آمارگیریهای مرسوم به اندازه‌های نمونه‌ای بزرگتر نیاز دارند تا با معیارهای تعیین شده قابلیت اعتماد مطابقت پیدا کنند. همچنین امکان به دست آوردن برآوردهای منفی با این روش هست.

شرح مفصلتر این روش در مقاله‌های ابرناتی و همکاران [۲۷] ارائه شده و برخی جزئیات این فن توسط چنگ و همکاران [۲۸] مورد بحث قرار گرفته است.

۹.۱۴ خلاصه

در این فصل به بحث در مورد برخی کاربردهای نمونه‌گیری و روشهای برآورد کردن پرداختیم که به مشکلات ویژه در علوم بهداشتی ارتباط پیدا می‌کنند. یکی از این روشها که نوعی نمونه‌گیری با احتمال متناسب با اندازه است به منظور تعیین سطوح ایمن‌سازی در کشورهای در حال توسعه ابداع شده بود. یک روش دیگر - یعنی نمونه‌گیری تلسکوپی - برای یک آمارگیری از اختلال مشاعر در شانگهای، جمهوری خلق چین، ابداع شده بود. این فن در موقعیتهایی سودمند است که نمونه‌گیری فرعی در داخل خوشه‌ها شدنی نیست. سایر موضوعهایی که مورد بحث قرار گرفتند عبارت‌اند از روشهای برآورد کردن میزان شیوع بیماریها یا شرایط خاص از طریق آزمونهای غربالگری که دارای حساسیت و ویژگی ناقص‌اند، برآورد کردن پیشامدهای نادر یا گروههای جامعه‌ای کمیاب، روشهای برآورد کردن برای نواحی کوچک؛ و روشهای به دست آوردن اطلاعات حساس از طریق آمارگیریهای نمونه‌ای.

تمرین

۱.۱۴ در کارخانه‌ای با ۳۵۷۵ کارگر، یک نمونه تصادفی ساده از ۵۲۵ کارگر گرفته شده است. از همه اشخاص نمونه‌گیری شده نوار قلب گرفته شده است که از نظر هرگونه ناهنجاری، توسط دو پزشک، به طور مستقل خوانده شده‌اند. از ۵۲۵ نفر افراد نمونه ۲۵ نفر دارای ناهنجاریهایی بوده‌اند که هر دو پزشک متوجه شده‌اند؛ ۱۵ نفر ناهنجاریهایی داشته‌اند که پزشک A تشخیص داده ولی پزشک B تشخیص نداده است؛ ۳۷ نفر ناهنجاریهایی داشته‌اند که فقط پزشک B تشخیص داده است. دربقیه هیچ‌گونه ناهنجاری توسط هیچ‌یک از دو پزشک تشخیص داده نشده است. براساس این داده‌ها، برآورد تعداد ناهنجاریها در میان ۳۵۷۵ کارگر چقدر است؟

۲.۱۴ همین اشخاص براساس یک آزمایش تکی قند خون ناشتا از لحاظ بیماری قند، غربالگری شده‌اند. این روش غربالگری به خصوص دارای حساسیت معلوم معادل ۸۰٪ و ویژگی معادل ۹۶٪ است. براساس یافته‌های حاصل از ۱۴ آزمایش مثبت، شیوع بیماری قند را در میان کارمندان برآورد کرده و بازه‌های اطمینان ۹۵ درصدی آن را تعیین کنید.

۳.۱۴ همان افراد نمونه با استفاده از یک معاینه با دستگاه استاندارد سنجش فشار خون از لحاظ بیماری فشار خون نیز غربالگری شده‌اند و براساس این معاینه، ۶۳ نفر مثبت تشخیص داده شده‌اند. از میان کسانی که مثبت تشخیص داده شده بودند ۲۵ نفر از نظر فشار خون مورد ارزیابی دقیقتر قرار گرفتند و بیماری فشار خون ۲۱ نفر از آنها تأیید شد. از میان کسانی که ابتدا به عنوان منفی غربالگری شده بودند نمونه‌ای متشکل از ۵۰ نفر گرفته شد و در ارزیابی دقیقتر معلوم شد که ۸ نفر از این افراد فشار خون دارند. براساس این یافته‌ها، برآورد شیوع فشار خون در این جامعه چقدر است؟ بازه‌های اطمینان ۹۵ درصدی را برای این نسبت برآورد شده تعیین کنید.

۴.۱۴ از داده‌های ارائه شده در تمرین قبل، حساسیت و ویژگی معاینه سنجش فشار خون را که در غربالگری برای بیماری فشار خون به کار رفته است برآورد کنید.

۵.۱۴ یک نمونه تصادفی ساده از ۳۰۰ خانوار در جامعه‌ای متشکل از ۳۵۶۲ خانوار گرفته شده است. از پاسخگویان هر خانوار سؤال شد که آیا در طی سال گذشته در آن خانه و یا در هر خانه دیگر بلوک خانه پاسخگو سرقتی اتفاق افتاده است یا نه. شش مورد سرقت زیر گزارش داده شدند:

تعداد خانوارهای واجد شرایط	سرقت
برای گزارش سرقت	
۳	۱
۲	۲
۷	۳
۶	۴
۸	۵
۲	۶

براساس این یافته‌ها، تعداد سرقت‌های اتفاق افتاده طی سال گذشته در آن جامعه را برآورد کنید.

کتابشناسی

The following publications discuss the methodology used in the EPI Surveys.

1. Serfling, R. E., and Sherman, I. L., *Attribute Sampling Methods*, Publication No. 1230, U.S. Department of Health and Human Services, Public Health Service, Washington, D.C., 1975.
2. Henderson, R. H., et al., Assessment of vaccination coverage, vaccination scar rates, and smallpox scarring in five areas of West Africa, *Bulletin of the World Health Organization*, 48: 183, 1973.
3. Lemeshow, S., and Stroh, G., Jr., *Sampling Techniques for Evaluating Health Parameters in Developing Countries*, National Academy Press, Washington, D.C., 1988.
4. Lemeshow, S., et al., A computer simulation of the EPI survey strategy, *International Journal of Epidemiology*, 14: 473, 1985.

The following publications discuss the Shanghai Survey of Alzheimer's Disease and Dementia and single-stage cluster sampling with a telescopic respondent rule.

5. Levy, P. S., Yu, E. S. H., Liu, W. T., Zhang, M. Y., Wang, Z. Y., Wong, S., and Katzman, R., Single stage cluster sampling with a telescopic respondent rule: A variation motivated by a survey of dementia in elderly residents of Shanghai. *Statistics in Medicine*, 8, 1537, 1989.
6. Levy, P. S., Yu, E. S. H., Liu, W. T., Zhang, M., Wang, Z., Wong, S., and Katzman, R., Variation on single stage cluster sampling used in a survey of elderly people in Shanghai. *International Journal of Epidemiology*, 17: 931, 1988.

The following publications discuss estimation of prevalence from screening tests.

7. Weiss, N. S., *Clinical Epidemiology: The Study of Outcome of Disease*, Oxford University Press, New York, 1986.
8. Levy, P. S., and Kass, E. H., A three population model for sequential screening for Bacteriuria. *American Journal of Epidemiology*, 91: 148, 1970.
9. Gastwirth, J., The statistical precision of medical screening procedures: Application to polygraph and AIDS antibodies test data. *Statistical Science*, 2: 213, 1987.
10. Lew, R. A., and Levy, P. S., Estimation of prevalence on the basis of screening tests. *Statistics in Medicine*, 8: 1225, 1989.

The following publications deal with network sampling.

11. Sirken, M. G., Household surveys with multiplicity. *Journal of the American Statistical Association*, 65: 257, 1970.
12. Sirken, M. G., Stratified sample surveys with multiplicity. *Journal of the American Statistical Association*, 67: 224, 1972.
13. Sudman, S., Sirken, M. G., and Cowan, C.D., *Science* 240, 991-996, 1988.
14. Nathan, G., Schmelz, U. O., and Kenvin, J., *Multiplicity Study of Marriages and Births in Israel*, Vital and Health Statistics, Series 2, No. 70, National Center for Health Statistics, Rockville, MD, 1977.
15. Czaja, R., Blair, J., Using network sampling for rare populations: An application to local crime victimization surveys. American Statistical Association, Proceedings of the Survey Research Section, 38-43, 1988.
16. Czaja, R., Snowden, C., and Cassady, R., Reporting bias and sampling errors in a survey of a rare population using multiplicity counting rules. *Journal of the American Statistical Association*, 81: 411, 1986.

17. Sirken, M. G., Inderfurth, G. P., Burnham, C. E., and Danchik, K. M., Household sample survey of diabetes: design effects of counting rules. *Proceedings of the American Statistical Association, Social Statistics Section*, 659, 1975.
18. Levy, P. S., Optimum allocation in stratified random network sampling for estimating the prevalence of attributes in rare populations. *Journal of the American Statistical Association*, 72: 758, 1978.
19. Sirken, M. G., and Levy, P. S., Multiplicity estimation of proportions based on ratios of random variables. *Journal of the American Statistical Association*, 69: 68, 1974.
20. Levy, P. S., Simple cluster sampling with multiplicity. *Proceedings of the American Statistical Association, Social Statistics Section*, 963, 1977.

The following publications develop methodology for estimation of events from dual systems.

21. Chandra Sekar, C., and Deming, W. E., On a method of estimating birth and death rates and the extent of registration. *Journal of the American Statistical Association*, 44: 101, 1949.
22. Cowan, C. D., and Malec, D., Capture-recapture models when both sources have clustered observations. *Journal of the American Statistical Association*, 81: 347, 1986.

The following publications discuss methodology for obtaining estimates for small areas.

23. Gonzalez, M. E., and Hoza, C., Small area estimation with applications to unemployment and housing estimates. *Journal of the American Statistical Association*, 73: 7, 1978.
24. Levy, P. S., and French, D., *Synthetic Estimation of State Health Characteristics Based on the Health Interview Survey*, Vital and Health Statistics, Series 2, No. 75, DHEW publication No. (PHS) 78-1349, U.S. Government Printing Office, Washington, D.C., 1977.
25. Levy, P. S., Small-area estimation-Synthetic and other procedures, 1968-1978, *Proceedings of the Workshop on Synthetic Estimates for Small Areas*, National Institute of Drug Abuse, Research Monography 24. U.S. Government Printing Office, Washington, D.C., 1979.

The following publications discuss randomized response methodology.

26. Warner, S. L., Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60: 63, 1975.
27. Abernathy, J. R., Greenberg, B. G., and Horvitz, D. G., Estimates of induced abortion in urban North Carolina. *Demography*, 7: 19, 1970.
28. I-Cheng, C., Chow, L. P., and Rider, R. V., The randomized response technique as used in the Taiwan outcome of pregnancy study. *Studies in Family Planning (A Publication of the Population Council)*, 3: 265, 1972.

The following publications deal with issues related to acceptance sampling.

29. Brownlee, K. A., *Statistical Theory and Methodology in Science and Engineering*, 2nd ed., Wiley, New York, 1965.
30. Lemeshow, S., and Stroh, G., Jr., *Sampling Techniques for Evaluating Health Parameters in Developing Countries*, National Academy Press, Washington, D.C., 1988.
31. Lemeshow, S., Hosmer, D., Klar, J., and Lwanga, S., *Adequacy of Sample Size in Health Studies*, Wiley, New York, 1990.

The following publications deal with issues related to cohort studies described in Section 14.3.

32. Kannel, W. B., An epidemiological study of cardiovascular disease. In: *Fifth Conference on Cerebral Vascular Diseases*, R. G. Sickert and J. P. Whisnat, eds., Grune and Stratton, New York and London, 1966.
33. Paffenbarger, R. L., Physical activity, all-cause mortality, and longevity of college alumni. *New England Journal of Medicine*, 314: 605, 1986.
34. Kovar, M. G., Fitti, J. E., and Chyba, M. M., The longitudinal study of aging: 1984-1990. *Vital and Health Statistics*, 1: 28, 1992.
35. Rothman, K. J., and Boice, J. D., *Epidemiologic Analysis with a Programmable Calculator*. Epidemiology Resources, Inc., Boston, 1982.

The following expository articles on topics discussed in this chapter have appeared recently in the Encyclopedia of Biostatistics.

36. Cohen, S., Small area estimation. In *The Encyclopedia of Biostatistics*, Armitage, P. A., and Colton, T., Eds., Wiley, Chichester, U.K., 1998.
37. Franklin, L., Randomized response techniques in sample surveys. In *The Encyclopedia of Biostatistics*, Armitage, P. A., and Colton, T., Eds., Wiley, Chichester, U.K., 1998.
38. Krotki, K., Sampling in developing countries. In *The Encyclopedia of Biostatistics*, Armitage, P. A., and Colton, T., Eds., Wiley, Chichester, U.K., 1998.
39. Sirken, M., Network sampling. In *The Encyclopedia of Biostatistics*, Armitage, P. A., and Colton, T., Eds., Wiley, Chichester, U.K., 1998.