بهبود براوردهای آمارگیری نیروی کار
با استفاده از چرخشی بودن طرح

مشاور طرح: فرهاد مهران

زهره فلاح محسن خانی

فاطمه هرندی

شیرین گلچی

آرمان بیدار بخت نیا

گروه پژوهشی طرح های فنی و روش های آماری

پژوهشکده ی آمار

بهار ۱۳۸۸

به نام خداوند جان و خرد

پیشگفتار

امروزه آمارگیری نمونه‌ای به صورت گستره برای تولید آمار رسمی کشور مورد استفاده قرار می‌گیرد. معمولاً نمونه‌گیری‌های مهم در فاصله‌های زمانی منظم تکرار می‌شوند. این نمونه‌گیری‌ها تحت عنوان نمونه‌گیری‌های مکرر شناخته می‌شوند. یکی از انواع نمونه‌گیری‌های مکرر که در آمارگیری‌های ملی به کار گرفته می‌شود، نمونه‌گیری چرخشی می‌باشد. روش نمونه‌گیری چرخشی بر اساس چرخش نمونه بین دوره‌ها است. منظور از چرخش نمونه، جایگزین شدن تعدادی از واحدهای نمونه با واحدهای جدید از یک دوره به دوره‌ی دیگر است. چرخشی بودن طرح آمارگیری ذاتاً موجب کم شدن واریانس تغییرات می‌شود ولی اگر از اطلاعات دوره‌های گذشته نیز استفاده شود در براوردهای سطح نیز کاهش واریانس صورت می‌گیرد که بالتبع باعث کاهش بیش از پیش واریانس براورد تغییرات نیز می‌گردد.

آمارگیری نیروی کار مرکز آمار ایران که یکی از مهم‌ترین آمارگیری‌های ملی در ایران به شمار می‌رود در انتخاب نمونه از روش نمونه‌گیری چرخشی استفاده می‌نماید. این آمارگیری با هدف دستیابی به شاخص‌های فصلی و سالانه نیروی کار و تغییرات آن در سطح کل کشور و استان‌ها از بهار ۱۳۸۴ اجرا شده است. در حال حاضر برای ارائه براوردهای سطح تنها از اطلاعات دوره جاری این آمارگیری استفاده می‌شود در صورتی که به دلیل چرخشی بودن نمونه‌ها قابلیت بهبود براوردهای دوره جاری با استفاده از اطلاعات دوره‌های قبل نیز وجود دارد.

پژوهشکده‌ی آمار با توجه به رسالت خود در زمینه معرفی روش‌های نوین و مناسب آماری، طرح پژوهشی بهبود براوردهای آمارگیری نیروی کار با استفاده از چرخشی بودن طرح را در دستور کار خود قرار داد. مجموعه‌ی حاضر با هدف معرفی روشی به منظور بهبود براوردهای آمارگیری نیروی کار که دارای واریانس کمتری نسبت به روش فعلی می‌باشد و کاربرد این روش در ارائه

براوردهای نیروی کار، گامی است در این راه که امید می‌رود مورد توجه و استفاده مسئولین ذی‌ربط قرار گیرد.

اجرای این طرح با مدیریت سرکار خانم زهره فلاح محسن‌خانی و همکاری سرکار خانم فاطمه هرندی، سرکار خانم شیرین گلچی و جناب آقای آرمان بیداربخت‌نیا انجام شده است که در این‌جا از یکایک این عزیزان تشکر و قدردانی می‌شود.

مشاور محترم طرح، جناب آقای دکتر فرهاد مهران نیز علاوه بر ارائه راهنمایی‌های بسیار ارزشمند، با انجام بخشی از مراحل اصلی کار، گروه مطالعاتی را بیش از پیش یاری نمودند که بدین وسیله از همکاری صمیمانه ایشان تشکر و قدردانی می‌شود.

در تهیه‌ی برنامه‌های رایانه‌ای این طرح، سرکار خانم مریم جوادی همکاری صمیمانه‌ای داشتند، که بدین‌وسیله از ایشان تشکر و قدردانی می‌شود.

از سرکار خانم مهرنوش میرمحمد که زحمت تایپ این گزارش را به عهده داشته و با حوصله فراوان آن را به انجام رساندند نیز بسیار تشکر می‌شود.

از خوانندگان محترم تقاضا می‌شود، نظرات اصلاحی خود در ارتباط با محتوای مجموعه‌ی حاضر را به گروه پژوهشی طرح‌های فنی و روش‌های آماری منعکس نمایند.

گروه پژوهشی طرح‌های فنی و روش‌های آماری

فهرست مطالب

عنوان **صفحه**

مقدمه ۱

فصل اول: مبانی نظری ۷

۱-۱ براورد مستقیم سطح ۱۰

۱-۲ براورد مستقیم تغییرات ۱۱

۱-۲-۱ براورد مستقیم تغییرات در نمونه‌های مستقل ۱۱

۱-۲-۲ براورد مستقیم تغییرات در نمونه‌های مشترک ۱۱

۱-۲-۳ براورد مستقیم تغییرات در نمونه‌های چرخشی ۱۲

۱-۳ براورد غیر مستقیم سطح ۱۴

۱-۳-۱ پیش‌رو ۱۴

۱-۳-۲ پس‌رو ۱۵

۱-۴ براورد مرکب سطح ۱۷

۱-۴-۱ Kی بهینه ۱۸

۱-۴-۲ δی بهینه ۱۹

۱-۴-۳ واریانس براوردگر مرکب سطح ۲۰

۱-۴-۴ بهره حاصل از K و δهای مختلف ۲۲

۱-۴-۴-۱ بهره حاصل از K و δهای بهینه ۲۲

۱-۴-۴-۲ بهره حاصل از Kی بهینه و $\delta=\frac{۱}{۲}$ ۲۴

۱-۵ براورد مرکب تغییرات ۲۶

۱-۶ براورد کالیبره مرکب ۲۷

۱-۷-جمع‌بندی ۳۱

فصل دوم: مروری بر تجربیات کشورها ۳۵

۲-۱ CPS آمریکا ۳۸

۲-۱-۱ تاریخچه‌ای از CPS ۳۸

۲-۱-۲ وزن‌دهی در CPS ۳۹

۲-۱-۳ براورد مرکب ۴۲

۲-۱-۳-۱ براورد مرکب در CPS ۴۳

۲-۱-۳-۲ محاسبه وزن‌های مرکب ۴۴

۲-۱-۳-۳ انتخاب ضرایب A و K ۴۶

۲-۲ LFS کانادا ۴۸

۲-۲-۱ تاریخچه‌ای از LFS ۴۸

۲-۲-۲ وزن‌دهی در LFS ۴۹

۲-۲-۳ کالیبره مرکب ۵۱

۲-۲-۴ انتخاب مقدار α ۵۳

۲-۳ LFS استرلیا ۵۶

۲-۳-۱ تاریخچه‌ی LFS ۵۶

۲-۳-۲ وزن‌دهی در LFS ۵۶

۲-۳-۳ براوردگر وزن‌دهی مرکب BLUEB1 ..... ۵۸

۲-۴ آمارگیری نیروی کار ایران ۶۱

۲-۴-۱ تاریخچه آمارگیری نیروی کار ایران ۶۱

۲-۴-۲ روش نمونه‌گیری و مراحل وزن‌دهی آمارگیری نیروی کار ۶۱

۲-۵ مقایسه تطبیقی ۶۴

فصل سوم: روش پیشنهادی براورد مرکب آمارگیری نیروی کار ایران ۶۵

۳-۱ متغیرهای اصلی ۶۹

۳-۱-۱ براورد مرکب متغیرهای اصلی ۶۹

۳-۱-۲ میزان کاهش واریانس در نتیجه‌ی استفاده از براورد مرکب ۷۱

۳-۱-۲-۱ مثال عددی ۷۲

۳-۲ سایر متغیرها ۷۳

۳-۲-۱ براورد مرکب سایر متغیرها ۷۳

۳-۳ ضریب همبستگی متغیرهای دو حالتی ۷۵

۳-۳-۱ مثال عددی ۷۶

۳-۴ مقدار K ۷۹

۳-۴-۱ مثال عددی ۷۹

۳-۵ براورد واریانس(بر اساس مجموع‌های خوشه‌های نهایی) ۸۰

۳-۵-۱ براورد واریانس نسبت‌ها ۸۲

۳-۵-۲ براورد واریانس تفاضل نسبت‌ها ۸۳

۳-۵-۳ مثال عددی ۸۴

۳-۶ براورد واریانس (با استفاده از اثر طرح) ۸۴

۳-۶-۱ مثال عددی ۸۵

۳-۷ خلاصه مراحل محاسبه براوردهای مرکب با روش پیشنهادی ۸۷

فصل چهارم: براورد مرکب فصل دوم سال ۱۳۸۴ آمارگیری نیروی کار ۸۹

۴-۱ مقادیر ضریب همبستگی و K ۹۲

۴-۲ براورد مرکب متغیرهای اصلی ۹۳

۴-۳ واریانس براورد مرکب ۹۷

۴-۳-۱ واریانس براوردهای مرکب با استفاده از روش ورما ۹۷

۴-۳-۲ واریانس براوردهای مرکب با استفاده از روش $Deft^{2}$ ۹۸

۴-۴ بهره‌ی براوردگر مرکب ۱۰۰

۴-۵ نتیجه‌گیری ۱۰۱

منابع ۱۰۳

واژه‌نامه ۱۰۹

پیوست‌ها ۱۱۵

پیوست ۱: نحوه محاسبه‌ی Kی بهینه برای بیش از ۲ نمونه (۲ تکرار)

با الگوی چرخش ۲-۲-۲ ..... ۱۱۷

پیوست ۲: متغیرهای کمکی مرکب تعریف شده در سطح ایالت برای

LFS کانادا ۱۲۲

پیوست ۳- وزن‌دهی مرکب BLUE B1 استرالیا در دو مرحله‌ی براورد

مرکب و وزن‌دهی ۱۲۵

پیوست ۴- برنامه‌های رایانه‌ای ۱۲۷

مقدمه

بسیاری از نمونه‌گیری‌های مهم در فاصله‌های زمانی معین تکرار می‌شوند. می‌توان از داده‌هایی که در نمونه‌گیری‌های مکرر از یک جامعه در دوره‌های مختلف زمانی به دست می‌آید، به منظور براورد پارامترهای مربوط به ویژگی‌های مورد مطالعه برای دوره زمانی جاری، براورد تغییرات پارامترها از دوره قبل تا دوره زمانی جاری، و براورد پارامترهایی مانند مجموع کل یا میانگین برای چند دوره زمانی استفاده کرد. در نمونه‌گیری‌های مکرر، بررسی تغییرات از اهمیت خاصی برخوردار است. به منظور بالا بردن دقت براورد تغییرات بین دوره‌های زمانی مختلف همراه با ارائه براوردهای مطلوب از دوره زمانی جاری، وجود تداخل بین نمونه‌های دوره‌های مختلف، لازم است. یعنی در هر دوره باید نمونه شامل مجموعه‌ای از واحدهای نمونه‌گیری شده در دوره قبلی و مجموعه‌ای از واحدهایی که در دوره قبل نمونه‌گیری نشده‌اند، باشد. نمونه‌گیری مکرر با این نوع ترکیب نمونه، نمونه‌گیری چرخشی نامیده می‌شود (کوتز و همکاران، ۲۰۰۵).

در نمونه‌گیری چرخشی ممکن است برخی واحدها برای چند دوره زمانی متوالی در نمونه باقی بمانند و یا برخی واحدها پس از خروج موقت از نمونه، دوباره در دوره‌های بعدی وارد نمونه شوند. منظور از چرخش نمونه، جایگزین شدن تعدادی از واحدهای نمونه با واحدهای جدید از یک دوره به دوره دیگر است (فلاح‌محسن‌خانی، ۱۳۸۵).

در نمونه‌گیری‌های مکرر که واحدهای انتخاب شده در هر دوره مستقل از دوره‌های دیگر هستند، براوردهای سطح تنها بر اساس اطلاعات همان دوره ارائه می‌شوند در صورتی که وجود واحدهای نمونه‌گیری شده در دوره قبلی در نمونه‌گیری چرخشی ارائه براوردهایی دقیق‌تر با نام براوردهای مرکب را امکان‌پذیر می‌سازد. به دلیل وجود این واحدها، نتایج حاصل از دوره‌های مختلف دارای همبستگی مثبت می‌باشند. براورد مرکب با استفاده از این همبستگی‌های مثبت که در طول زمان بین اطلاعات واحدهای یکسان وجود دارد، دقت براوردها به خصوص براوردهای تغییرات را بهبود می‌دهد. تئوری این روش آن‌گونه که در آمارگیری‌های ماهانه ایالات متحده به کار می‌رود توسط

وودراف در سال ۱۹۶۳ شرح داده شده که عبارت است از محاسبه‌ی براوردهای جاری به صورت مجموع موزون دو مؤلفه a و b که a براوردی است که مستقیماً از دوره جاری به دست آمده است و b براورد غیر مستقیم دوره جاری است به این ترتیب که براورد دوره قبل، برای تغییرات مشاهده شده بین دوره قبل و دوره جاری برای واحدهای تکراری که در هر دو دوره اندازه‌گیری شده‌اند، تعدیل شده است. تغییرات مشاهده شده ممکن است به صورت اختلاف مقادیر دو دوره یا به صورت نسبت آن‌ها نشان داده شود. وزن‌های مؤلفه‌ها به اندازه همبستگی‌های دوره به دوره بستگی خواهند داشت. هر چه ضرایب همبستگی بزرگ‌تر باشد، وزن‌های مربوط به مؤلفه b بزرگ‌تر خواهند بود. در آمارگیری نیروی کار متغیرهایی مانند بیکاری، همبستگی‌های دوره به دوره کوچک‌تری نسبت به متغیرهایی مثل اشتغال دارند (مهران، و ورما، ۱۹۹۲).

هدف این پژوهش، مطالعه، بررسی و کاربرد روش‌های بهبود براوردهای حاصل از آمارگیری نیروی کار به منظور دست یافتن به براوردهای دقیق‌تر در آمارگیری مزبور می‌باشد.

در فصل اول گزارش، مبانی نظری براوردگر مرکب بیان شده است که شامل براورد مستقیم سطح و تغییرات، براورد غیر مستقیم سطح و تغییرات، براورد مرکب سطح و تغییرات و براورد کالیبره مرکب سطح و تغییرات می‌باشد.

فصل دوم گزارش در ۵ قسمت تدوین شده است. در سه قسمت اول، ویژگی‌های آمارگیری نیروی کار کشورهای آمریکا، کانادا و استرالیا ارائه شده است به گونه‌ای که بعد از بیان تاریخچه‌ای مختصر از آمارگیری مزبور در این کشورها، مراحل براورد این آمارگیری‌ها مورد بحث و بررسی قرار گرفته است، در قسمت چهارم مختصری از تاریخچه و روش براورد در آمارگیری نیروی کار ایران ارائه شده است و در قسمت آخر مشخصه‌های اصلی طرح نمونه‌گیری و خطای نسبی شاخص‌های عمده آمارگیری نیروی کار در چهار کشور آمریکا، کانادا، استرالیا و ایران مقایسه شده است.

فصل سوم گزارش با نگاهی توأم به مبانی نظری و تجربیات کشورهای آمریکا، کانادا و استرالیا روشی برای بهبود براوردهای آمارگیری نیروی کار ایران ارائه کرده است. همچنین با ارائه واریانس براوردها، نشان داده شده است که روش معرفی شده براوردهای کاراتری را ارائه می‌نماید.

در فصل چهارم با استفاده از داده‌های فصل بهار و تابستان آمارگیری نیروی کار سال ۱۳۸۴ و به کمک روش معرفی شده در فصل سوم، براوردهای بهبود یافته فصل تابستان سال ۱۳۸۴ آمارگیری نیروی کار محاسبه شده است. همچنین میزان دقت این براوردها نسبت به براوردهای مستقیم نیز ارائه شده است. در پایان این فصل نتیجه‌گیری این طرح پژوهشی ارائه شده است.

فصل اول

مبانی نظری

هدف از ارائه‌ی مطالب این فصل علاوه بر آشنا نمودن متخصصان نمونه‌گیری‌های چرخشی با اهم روش‌هایی که برای بهبود براوردهای آمارگیری‌های چرخشی استفاده می‌شود، آگاه نمودن از نحوه‌ی محاسبه دقت براوردها می‌باشد. از آن‌جا که در این فصل مبانی نظری براوردها ارائه شده است، به منظور درک بهتر، مطالب این قسمت با نگاه کتاب درسی تدوین شده است.

فرضیاتی که با توجه به آن‌ها براوردگرها مورد مطالعه قرار گرفته‌اند، به این ترتیب بوده است که در زمان‌های ۱ و ۲ نمونه‌های s_1 و s_2 با اندازه‌ی یکسان از جامعه‌ی ثابت و بسته به روش تصادفی ساده انتخاب می‌شوند، فرض بر این است که در این دو تکرار بی‌پاسخی وجود ندارد.

در این فصل با در نظر گرفتن گزینه‌های مختلف برای انتخاب نمونه‌های یک آمارگیری مکرر با دو تکرار، براوردگرهای مستقیم، غیر مستقیم و مرکب ابتدا برای سطح و سپس برای تغییرات ارائه شده است. براوردگر مستقیم براوردگری است که در آن فقط از اطلاعات یک مقطع استفاده می‌شود. در براوردگر غیر مستقیم از اطلاعات هر دو مقطع استفاده می‌شود. هدف از ارائه‌ی براوردگر غیر مستقیم، شناخت بهتر براوردگر مرکب بوده است زیرا براوردگر مرکب ترکیبی از براوردگر مستقیم و غیر مستقیم است. بعد از ارائه‌ی این براوردگرها، فرمول واریانس آن‌ها نیز ارائه شده است. در ادامه براوردگر کالیبره‌ی مرکب که روش عملی دیگری برای تصحیح براوردگرهای حاصل از نمونه‌گیری‌های چرخشی است، معرفی و کارایی آن بررسی شده است. در پایان جمع‌بندی این فصل ارائه شده است.

۱-۱ براورد مستقیم سطح

فرض کنید $U=\{۱,\ldots,k,\ldots N\}$ مجموعه واحدهای جامعه مورد بررسی است. y_k، مقدار متغیر مورد بررسی برای kامین واحد جامعه است و t مجموع y_kهای جامعه است. از این جامعه با طرح نمونه‌گیری تصادفی ساده، نمونه‌ی احتمالی s با اندازه‌ی n انتخاب می‌شود. به عبارتی $s=\{y_۱,\ldots,y_n\}$ و $\pi_k=P_r(k\in s)=\frac{n}{N}$ احتمال انتخاب عضو kام جامعه در نمونه‌ی s می‌باشد. هدف، براورد t همراه با ارائه‌ی دقت آن می‌باشد.

با توجه به فرمول براوردگر هورویتز- تامپسون (HT)، براوردگر مستقیم مجموع y به صورت زیر خواهد بود.

$$\hat{t}^D=\sum_{k\in s}d_k y_k=\sum_{k\in s}\frac{y_k}{\pi_k}=\sum_{k\in s}y_k.\frac{N}{n}=N\bar{y} \qquad , \qquad d_k=\frac{۱}{\pi_K}$$

همچنین واریانس براوردگر $\hat{t}^D$ وقتی $\mathrm{var}(y)=\sigma^۲$ عبارت است از:

$$\mathrm{var}(\hat{t}^D)=\mathrm{var}(N\bar{y})=N^۲(۱-\frac{n}{N})\frac{\sigma^۲}{n}$$

با فرض کوچک بودن کسر نمونه‌گیری داریم:

$$\mathrm{var}(\hat{t}^D)\simeq N^۲\frac{\sigma^۲}{n} \qquad (۱)$$

۱-۲ براورد مستقیم تغییرات

فرض کنید در دو زمان ۱ و ۲ نمونه‌های s_1 و s_2 با اندازه‌ی یکسان انتخاب شده است. هدف براورد $(t_2 - t_1)$ به همراه واریانس آن است، از آن‌جا که در یک آمارگیری مکرر با دو تکرار، نمونه‌ها در زمان‌های ۱ و ۲ ممکن است کاملاً مستقل باشند، کاملاً مشترک باشند (داده‌های پانلی) یا قسمتی از آن‌ها مشترک باشد (نمونه‌گیری چرخشی)، با در نظر گرفتن سه گزینه‌ی مذکور برای نمونه‌ها براوردگر مورد نظر و واریانس آن برای هر گزینه به طور جداگانه ارائه می‌شود.

۱-۲-۱ براورد مستقیم تغییرات در نمونه‌های مستقل

براوردگر مستقیم تغییرات در دو نمونه که واحدهای آن‌ها کاملاً مستقل هستند عبارت است از

$$\hat{d}_{.} = \hat{t}_2^D - \hat{t}_1^D = N(\bar{y}_2 - \bar{y}_1)$$

$$\begin{aligned}\operatorname{var}(\hat{d}_{.}) &= N^2 \operatorname{var}(\bar{y}_2 - \bar{y}_1) = N^2(\operatorname{var}(\bar{y}_1) + \operatorname{var}(\bar{y}_2)) \\ &= N^2((1-\frac{n}{N})\frac{\sigma^2}{n} + (1-\frac{n}{N})\frac{\sigma^2}{n}) = 2N^2(1-\frac{n}{N})\frac{\sigma^2}{n}\end{aligned}$$

با فرض کوچک بودن کسر نمونه‌گیری، واریانس براورد مستقیم تغییرات عبارت است از:

$$\operatorname{var}(\hat{d}_{.}) \simeq 2N^2\frac{\sigma^2}{n} \qquad (2)$$

۱-۲-۲ براورد مستقیم تغییرات در نمونه‌های مشترک

براوردگر مستقیم تغییرات در دو نمونه که واحدهای آن‌ها کاملاً یکسان هستند به صورت زیر به دست می‌آید:

$$\hat{d}_1 = \hat{t}_2^D - \hat{t}_1^D = N(\bar{y}_2 - \bar{y}_1)$$

$$\begin{aligned} Var(\hat{d}_1) &= Var(N(\bar{y}_2 - \bar{y}_1)) = N^2 Var(\bar{y}_2 - \bar{y}_1) \\ &= N^2(Var(\bar{y}_1) + Var(\bar{y}_2) - 2Cov(\bar{y}_1, \bar{y}_2)) \\ &= N^2((1-\frac{n}{N})\frac{\sigma^2}{n} + (1-\frac{n}{N})\frac{\sigma^2}{n} - 2(1-\frac{n}{N})\frac{\sigma^2}{n}\rho) \\ &= N^2((1-\frac{n}{N})2\sigma^2\frac{(1-\rho)}{n}) \end{aligned}$$

با فرض کوچک بودن کسر نمونه‌گیری، واریانس براوردگر تغییرات به صورت ذیل به دست می‌آید:

$$\text{var}(\hat{d}_1) \simeq \frac{2N^2\sigma^2}{n}(1-\rho) \qquad (3)$$

۱-۲-۳ براورد مستقیم تغییرات در نمونه‌های چرخشی

فرض می‌شود δ درصد از واحدهای نمونه‌ی s_2، عضو نمونه‌ی s_1 نیز بوده‌اند. به عبارتی داریم:

$$\begin{aligned} s_1 &= s_1^{\delta} \cup s_1^{1-\delta} \\ s_2 &= s_2^{\delta} \cup s_2^{1-\delta} \\ s_1^{\delta} &= s_2^{\delta} \end{aligned}$$

بنابراین $\hat{t}_1^D$ و $\hat{t}_2^D$ عبارتند از:

$$\begin{aligned} \hat{t}_1^D &= N\bar{y}_1 = N(\delta\bar{y}_{1\delta} + (1-\delta)\bar{y}_{1(1-\delta)}) \\ \hat{t}_2^D &= N\bar{y}_2 = N(\delta\bar{y}_{2\delta} + (1-\delta)\bar{y}_{2(1-\delta)}) \end{aligned}$$

تفاضل دو عبارت $\hat{t}_1^D$ و $\hat{t}_2^D$، براورد مستقیم تغییرات در نمونه‌گیری چرخشی است. به عبارتی:

$$\hat{d}_\delta = (\hat{t}_2^D - \hat{t}_1^D) = N(\delta(\bar{y}_{2\delta} - \bar{y}_{1\delta}) + (1-\delta)(\bar{y}_{2(1-\delta)} - \bar{y}_{1(1-\delta)}))$$

با توجه به روابط ذیل واریانس $\hat{\Delta}_\delta$ نیز به دست می‌آید.