

روش‌شناسی برآورد واریانس تغییرات شاخص‌های اصلی آمارگیری نیروی کار ایران

مجری:

لیدا کلهری ندرآبادی

همکاران:

محمد رضا ریحانی

روشنک علی‌اکبری صبا

مشاور:

فرهاد مهران



پژوهشکده‌ی آمار

گروه پژوهشی طرح‌های فنی و روش‌های آماری

زمستان ۱۴۰۰

بسم الله الرحمن الرحيم

کد شناسه: RP-0006

عنوان: روش‌شناسی برآورد واریانس تغییرات شاخص‌های اصلی آمارگیری نیروی کار ایران

مجری: لیدا کلهری ندرآبادی

همکاران: محمدرضا ریحانی، روشنک علی‌اکبری صبا

مشاور: فرهاد مهران

گروه پژوهشی: طرح‌های فنی و روش‌های آماری

تاریخ انتشار: زمستان ۱۴۰۰

نوبت انتشار: اول

طراح جلد: ساناز مهندسی

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

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

آمارگیری‌هایی که در سازمان‌های آماری برای گردآوری اطلاعات در طول زمان انجام می‌گیرند اغلب، آمارگیری‌های مکرری هستند که اجرای آن‌ها در دوره‌های زمانی مشخص مانند ماه، سال یا فصل به صورت منظم تکرار می‌شود. دلیل عمده‌ی انجام این آمارگیری‌ها، نیاز به محاسبه‌ی برآوردی از تغییرات پارامترهای مورد نظر در دوره‌های زمانی متوالی است. به عبارت دیگر استنباط در خصوص تغییر پارامتر مورد نظر از یک دوره‌ی آمارگیری به دوره‌ی آمارگیری بعد و قضاوت در مورد معنی‌دار بودن تغییر مشاهده‌شده در آمارگیری‌های مکرر از اهمیت بالایی برخوردار است. واحدهای نمونه در آمارگیری‌های مکرر لزوماً یکسان نیستند، اما در صورت وجود واحدهای نمونه مشترک در دوره‌های مختلف آمارگیری، دقت برآورد تغییرات در طول زمان به طور مستقیم به میزان اشتراک نمونه‌ها در دوره‌های زمانی مختلف بستگی دارد.

با توجه به اهمیت برآورد تغییرات شاخص‌های نیروی کار در طول زمان، آمارگیری نیروی کار مرکز آمار ایران بر اساس یک طرح نمونه‌گیری چرخشی اجرا می‌شود. در این آمارگیری که از سال ۱۳۸۴ به صورت فصلی اجرا می‌شود، برآورد تغییرات فصلی و تغییرات سالانه‌ی شاخص‌های نیروی کار، بر اساس تفاضل برآوردهای حاصل از نمونه‌های دو سال یا دو فصل اجرای آمارگیری محاسبه می‌شوند. بخشی از این نمونه‌ها بر اساس الگوی چرخش طرح نمونه‌گیری در دو دوره‌ی زمانی مورد بررسی، مشترک هستند.

هدف از انجام پژوهش پیش‌رو ارائه‌ی روشی برای برآورد واریانس تغییرات بین دو دوره‌ی آمارگیری نیروی کار مرکز آمار ایران است. در فصل اول این گزارش ضمن توضیح ضرورت محاسبه‌ی واریانس تغییرات در آمارگیری‌های مکرر، روش‌شناسی برآورد واریانس تغییرات و تجربیات برخی از کشورها درباره این موضوع ارائه شده است. در فصل دوم کلیاتی در خصوص آمارگیری نیروی کار ایران و مطالعات پیشین برای برآورد واریانس تغییرات شاخص‌های آمارگیری نیروی کار ارائه شده است. در فصل سوم، برآورد واریانس تغییرات در آمارگیری نیروی کار با استفاده از چهار روش ارائه شده است. در نهایت در فصل چهارم بر اساس روش‌های به کارگرفته شده و نتایج به دست آمده از آن‌ها راهکار عملیاتی برای مرکز آمار ایران در ارائه برآورد واریانس تغییرات ارائه شده است که مرکز آمار ایران می‌تواند به عنوان الگوی اطلاع‌رسانی، همراه نتایج جاری آمارگیری نیروی کار از آن استفاده نماید. برآورد به دست آمده برای واریانس تغییرات از تمام روش‌های مورد مطالعه قابل قبول هستند و تفاوت اساسی با یکدیگر ندارند. این موضوع نشان‌گر این نکته است که مرکز آمار ایران اگر خطای تغییرات را

ارائه کند قویاً می‌تواند از تغییرات شاخص‌های نیروی کار دفاع کند و دامنه خطای قابل قبول، خود پشتیبان‌های برای برآورد تغییرات می‌شود.

این طرح پژوهشی در گروه پژوهشی طرح‌های فنی و روش‌های آماری پژوهشکده‌ی آمار با همکاری سرکار خانم دکتر لیدا کلهری به‌عنوان مجری طرح، آقای دکتر فرهاد مهران به‌عنوان مشاور طرح، و آقای محمدرضا ریحانی سردهائی و خانم دکتر روشنک علی اکبری صبا به‌عنوان همکاران اصلی طرح پژوهشی به انجام رسیده است، که بدین‌وسیله از ایشان صمیمانه تشکر و قدردانی می‌شود.

کروه پژوهشی طرح‌های فنی و روش‌های آماری

پژوهشکده‌ی آمار

فهرست مطالب

۱	روش‌شناسی و تجربه سایر کشورها.....
۲	۱-۱- تجربه سایر کشورها.....
۲	۱-۱-۱- راهکار برآورد واریانس تغییرات برگر و پرایم (۲۰۱۳)، آووز-آلپر و برگر (۲۰۱۵).....
۱۲	۲-۱- رویکرد ایتالیا در برآورد واریانس تغییرات شاخص‌های نیروی کار.....
۱۳	۱-۲-۱- طرح نمونه‌گیری نیروی کار ایتالیا.....
۱۴	۲-۲-۱- برآورد واریانس تغییرات خالص.....
۱۵	۳-۱- رویکرد اتحادیه اروپا در برآورد واریانس تغییرات شاخص‌های نیروی کار.....
۱۶	۱-۳-۱- برآورد واریانس تغییرات خالص.....
۲۰	۲-۳-۱- برآورد واریانس تغییرات ناخالص.....
۲۲	۴-۱- روش‌شناسی آمارگیری نیروی کار کانادا در برآورد واریانس.....
۲۳	۱-۴-۱- خودگردان راثو-وو.....
۲۴	۲-۴-۱- نمونه‌های خودگردان طرح آمارگیری نیروی کار.....
۲۹	۳-۴-۱- وزن‌های خودگردان طرح آمارگیری نیروی کار.....
۳۰	۴-۴-۱- برآورد واریانس.....
۳۳	آمارگیری نیروی کار ایران و مطالعات پیشین.....
۳۴	۱-۲- آمارگیری نیروی کار ایران.....
۳۶	۲-۲- پیشینه برآورد واریانس تغییرات شاخص‌های نیروی کار در ایران.....
۴۵	بررسی برآوردهای واریانس تغییرات در آمارگیری نیروی کار ایران.....
۴۶	۱-۳- چهار روش مورد بررسی.....
۴۸	۲-۳- نتایج مقایسه دو فصل متوالی.....
۴۸	۱-۲-۳- واریانس تغییرات در دو فصل متوالی.....
۵۴	۲-۲-۳- ضریب همبستگی بین دو فصل متوالی.....
۵۹	۳-۳- نتایج مقایسه هر فصل با فصل مشابه سال قبل.....
۵۹	۱-۳-۳- واریانس تغییرات هر فصل نسبت به فصل مشابه سال قبل.....
۶۵	۲-۳-۳- ضریب همبستگی بین هر فصل با فصل مشابه سال قبل.....
۷۱	۴-۳- نتایج مقایسه دو سال متوالی.....
۷۱	۱-۴-۳- واریانس سالانه.....

۷۵	۲-۴-۳- واریانس تغییرات دو سال متوالی.....
۷۸	۳-۴-۳- ضریب همبستگی دو سال متوالی.....
۸۳	۳-۵- کارایی روش خودگردان.....
۸۵	نتیجه‌گیری و پیشنهادها.....
۸۶	بازه اطمینان و معنی‌داری تغییرات.....
۹۱	مرجع‌ها.....
۹۵	پیوست‌ها.....
۹۷	واژه‌نامه.....

فهرست جدول‌ها

- جدول ۱-۱- مشخصه‌های اصلی طرح نمونه‌گیری نیروی کار ایتالیا..... ۱۳
- جدول ۱-۳- شاخص‌های مکانی نسبت برآورد خودگردان به برآورد خودگردان ساده در برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه دو فصل پاییز و زمستان ۱۳۹۹ در سطح استانی..... ۵۳
- جدول ۲-۳- شاخص‌های مکانی نسبت برآورد خودگردان ساده به برآورد اول دستورالعمل اروپا در برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه دو فصل پاییز و زمستان ۱۳۹۹ در سطح استانی..... ۵۳
- جدول ۳-۳- شاخص‌های مکانی نسبت برآورد اول دستورالعمل اروپا به برآورد دوم دستورالعمل اروپا در برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه دو فصل پاییز و زمستان ۱۳۹۹ در سطح استانی..... ۵۳
- جدول ۴-۳- شاخص‌های مکانی نسبت برآورد خودگردان به برآورد خودگردان ساده در برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه دو فصل زمستان ۱۳۹۹ و زمستان ۱۳۹۸ در سطح استانی..... ۶۴
- جدول ۵-۳- شاخص‌های مکانی نسبت برآورد خودگردان ساده به برآورد اول دستورالعمل اروپا در برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه دو فصل زمستان ۱۳۹۹ و زمستان ۱۳۹۸ در سطح استانی..... ۶۴
- جدول ۶-۳- شاخص‌های مکانی نسبت برآورد اول دستورالعمل اروپا به برآورد دوم دستورالعمل اروپا در برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه دو فصل زمستان ۱۳۹۹ و زمستان ۱۳۹۸ در سطح استانی..... ۶۴
- جدول ۷-۳- برآورد واریانس سالانه پارامترهای اصلی آمارگیری نیروی کار ایران برای سال‌های ۱۳۹۷ تا ۱۳۹۹ به دو روش در سطح کل کشور..... ۷۲
- جدول ۸-۳- برآورد واریانس سالانه پارامترهای اصلی آمارگیری نیروی کار ایران برای سال‌های ۱۳۹۷ تا ۱۳۹۹ به دو روش در سطح مناطق شهری..... ۷۲
- جدول ۹-۳- برآورد واریانس سالانه پارامترهای اصلی آمارگیری نیروی کار ایران برای سال‌های ۱۳۹۷ تا ۱۳۹۹ به دو روش در سطح مناطق روستایی..... ۷۳
- جدول ۱۰-۳- شاخص‌های مکانی نسبت برآورد خودگردان به برآورد خودگردان ساده در برآورد واریانس سالانه پارامترهای اصلی آمارگیری نیروی کار ایران برای سال ۱۳۹۹ در سطح استانی..... ۷۵
- جدول ۱۱-۳- برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه دو سال متوالی به دو روش در سطح کل کشور..... ۷۵
- جدول ۱۲-۳- برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه دو سال متوالی به دو روش در سطح مناطق شهری..... ۷۶

- جدول ۳-۱۴- شاخص‌های مکانی نسبت برآورد خودگردان به برآورد خودگردان ساده در برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه سال ۱۳۹۹ با سال ۱۳۹۸ در سطح استانی..... ۷۸
- جدول ۳-۱۵- برآورد ضریب همبستگی بین پارامترهای اصلی یکسان آمارگیری نیروی کار ایران در دو سال متوالی به دو روش در سطح کل کشور..... ۷۹
- جدول ۳-۱۶- برآورد ضریب همبستگی بین پارامترهای اصلی یکسان آمارگیری نیروی کار ایران در دو سال متوالی به دو روش در سطح مناطق شهری..... ۷۹
- جدول ۳-۱۷- برآورد ضریب همبستگی بین پارامترهای اصلی یکسان آمارگیری نیروی کار ایران در دو سال متوالی به دو روش در سطح مناطق روستایی..... ۸۰
- جدول ۴-۱- خطای معیار تغییرات و بازه اطمینان ۹۵ درصدی مقایسه سالانه تعداد شاغلان، تعداد بیکاران، تعداد فعالان و نرخ بیکاری در آمارگیری نیروی کار مرکز آمار ایران ۱۳۹۸-۱۳۹۹..... ۸۸
- جدول ۴-۲- برآورد خطای معیار تغییرات سالانه و فصلی شاخص‌های اصلی نیروی کار مرکز آمار ایران..... ۹۰

فهرست شکل‌ها

- شکل ۳-۱- برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه دو فصل متوالی به چهار روش در سطح کل کشور..... ۴۹
- شکل ۳-۲- برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه دو فصل متوالی به چهار روش در سطح مناطق شهری..... ۵۰
- شکل ۳-۳- برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه دو فصل متوالی به چهار روش در سطح مناطق روستایی..... ۵۱
- شکل ۳-۴- برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه دو فصل پاییز و زمستان ۱۳۹۹ به چهار روش در سطح استانی..... ۵۲
- شکل ۳-۵- برآورد ضریب همبستگی بین پارامترهای اصلی یکسان آمارگیری نیروی کار ایران در دو فصل متوالی به چهار روش در سطح کل کشور..... ۵۴
- شکل ۳-۶- برآورد ضریب همبستگی بین پارامترهای اصلی یکسان آمارگیری نیروی کار ایران در دو فصل متوالی به چهار روش در سطح مناطق شهری..... ۵۵
- شکل ۳-۷- برآورد ضریب همبستگی بین پارامترهای اصلی یکسان آمارگیری نیروی کار ایران در دو فصل متوالی به چهار روش در سطح مناطق روستایی..... ۵۶
- شکل ۳-۸- برآورد ضریب همبستگی بین پارامترهای اصلی یکسان آمارگیری نیروی کار ایران در دو فصل پاییز و زمستان ۱۳۹۹ به چهار روش در سطح استانی..... ۵۷
- شکل ۳-۹- برآورد ضریب همبستگی بین پارامترهای اصلی یکسان آمارگیری نیروی کار ایران در دو فصل تابستان و پاییز ۱۳۹۹ به چهار روش در سطح استانی..... ۵۸
- شکل ۳-۱۰- برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه هر فصل با فصل مشابه سال قبل به چهار روش در سطح کل کشور..... ۶۰
- شکل ۳-۱۱- برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه هر فصل با فصل مشابه سال قبل به چهار روش در سطح مناطق شهری..... ۶۱
- شکل ۳-۱۲- برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه هر فصل با فصل مشابه سال قبل به چهار روش در سطح مناطق روستایی..... ۶۲
- شکل ۳-۱۳- برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه دو فصل زمستان ۱۳۹۹ و زمستان ۱۳۹۸ به چهار روش در سطح استانی..... ۶۳

- شکل ۳-۱۴- برآورد ضریب همبستگی بین پارامترهای اصلی یکسان آمارگیری نیروی کار ایران در هر فصل با فصل مشابه سال قبل به چهار روش در سطح کل کشور..... ۶۶
- شکل ۳-۱۵- برآورد ضریب همبستگی بین پارامترهای اصلی یکسان آمارگیری نیروی کار ایران در هر فصل با فصل مشابه سال قبل به چهار روش در سطح مناطق شهری..... ۶۷
- شکل ۳-۱۶- برآورد ضریب همبستگی بین پارامترهای اصلی یکسان آمارگیری نیروی کار ایران در هر فصل با فصل مشابه سال قبل به چهار روش در سطح مناطق روستایی..... ۶۸
- شکل ۳-۱۷- برآورد ضریب همبستگی بین پارامترهای اصلی یکسان آمارگیری نیروی کار ایران در دو فصل زمستان ۱۳۹۹ و زمستان ۱۳۹۸ به چهار روش در سطح استانی..... ۶۹
- شکل ۳-۱۸- برآورد ضریب همبستگی بین پارامترهای اصلی یکسان آمارگیری نیروی کار ایران در دو فصل پاییز ۱۳۹۹ و پاییز ۱۳۹۸ به چهار روش در سطح استانی..... ۷۰
- شکل ۳-۱۹- برآورد واریانس سالانه پارامترهای اصلی آمارگیری نیروی کار ایران برای سال ۱۳۹۹ به دو روش در سطح استانی..... ۷۴
- شکل ۳-۲۰- برآورد واریانس تغییرات پارامترهای اصلی آمارگیری نیروی کار ایران برای مقایسه سال ۱۳۹۹ با سال ۱۳۹۸ به دو روش در سطح استانی..... ۷۷
- شکل ۳-۲۱- برآورد ضریب همبستگی بین پارامترهای اصلی یکسان آمارگیری نیروی کار ایران در دو سال ۱۳۹۹ و ۱۳۹۸ به دو روش در سطح استانی..... ۸۱
- شکل ۳-۲۲- برآورد ضریب همبستگی بین پارامترهای اصلی یکسان آمارگیری نیروی کار ایران در دو سال ۱۳۹۸ و ۱۳۹۷ به دو روش در سطح استانی..... ۸۲
- شکل ۳-۲۳- برآورد واریانس‌های فصلی برای پارامترهای اصلی آمارگیری نیروی کار ایران به سه روش در سطح کل کشور..... ۸۴

روش‌شناسی و تجربه سایر کشورها

نظارت بر تغییرات یا روند شاخص‌ها در طول زمان در بسیاری از زمینه‌های علوم اقتصادی و اجتماعی از اهمیت شایانی برخوردار است. شاخص‌های بازار کار بر اساس برآوردهای حاصل از نمونه‌گیری به دست می‌آیند و نمایانگر وضعیت بازار کار هستند، بنابراین بررسی تغییرات آن‌ها از اهمیت ویژه‌ای برخوردار است. با توجه به استفاده از نمونه‌گیری چرخشی در آمارگیری نیروی کار اکثر کشورها، همپوشانی بین واحدهای نمونه فصلی و سالانه باعث می‌شود تفسیر تفاوت بین برآوردهای مقطع گمراه‌کننده باشد، مگر اینکه علاوه بر برآوردهای نقطه‌ای شاخص‌ها، برآورد واریانس تغییرات نیز ارائه شود تا امکان بررسی معنی‌دار بودن اختلاف‌ها فراهم شود. اگر برآورد مقطعی بر اساس نمونه‌های مستقل باشد، برآورد واریانس برای اندازه‌گیری تغییرات نسبتاً ساده خواهد بود. در این صورت، فقط باید واریانس هر برآورد مقطعی را محاسبه کنیم.

در ایران و بعضی از کشورهای عضو اتحادیه اروپا از طرح نمونه‌گیری چرخشی ۲-۲-۲ برای آمارگیری نیروی کار استفاده می‌شود، که در آن واحدهای نمونه‌گیری شده برای دو فصل متوالی در نمونه هستند، سپس برای دو فصل بعدی از نمونه خارج می‌شوند و بعد از آن دوباره دو فصل دیگر به نمونه وارد می‌شوند و پس از آن برای همیشه از نمونه خارج می‌شوند. الگوهای چرخشی دیگری نیز مورد استفاده قرار می‌گیرند، به‌عنوان مثال در آمارگیری نیروی کار کانادا در هر دوره‌ی آمارگیری یک‌ششم نمونه‌ها جدید هستند.

بنابراین نمی‌توان فرض کرد که برآوردهای فصلی یا سالانه نیروی کار مستقل از زمان بوده و لازم است همبستگی زمانی بین برآوردهای مقطع نیز در محاسبات واریانس لحاظ شود. چندین برآوردهای واریانس برای برآورد تغییرات بین دو مقطع زمانی در مقالات مختلف مانند کیش (۱۹۶۵)، مونیچ و زینز (۲۰۱۱)، نوردبرگ (۲۰۰۰)، کوالیته (۲۰۰۹)، کوالیته و تیله (۲۰۰۸)، تام (۱۹۸۴) و وود (۲۰۰۸) ارائه شده است.

اویسر و رایمون (۲۰۱۵) اظهار داشتند هیچ یک از این برآوردها به اندازه کافی دقیق، کاربردی و انعطاف‌پذیر نیستند تا بتوانند در چارچوب تعهدی در سطح اتحادیه اروپا اجرا شوند. اویسر و رایمون (۲۰۱۵) بیان کردند

برای محاسبه واریانس تغییرات نیروی کار در اتحادیه اروپا در حالت ایده آل، روش مورد استفاده برای برآورد واریانس تغییرات خالص سالانه باید:

- همپوشانی بین داده‌های فصلی و سالانه را لحاظ کند (جایی که داده‌های سالانه از میانگین داده‌های فصلی حاصل می‌شود).
- بیشتر ویژگی‌های پیچیده در طراحی نمونه مانند طبقه‌بندی، احتمالات نابرابر انتخاب، انتخاب چندمرحله‌ای، تعدیل وزن برای بی‌پاسخی واحد و کالیبراسیون براساس منابع داده خارجی را در نظر بگیرد.
- قادر به حل و فصل کردن تغییرات سطح‌ها، نسبت‌ها و نرخ‌ها باشد.

با توجه به آنچه گفته شد، در ادامه این فصل ابتدا برای آشنایی بیشتر تجربه سایر کشورها در برآورد واریانس تغییرات شاخص‌های آمارگیری‌های خانواری به‌ویژه آمارگیری نیروی کار^۱ (LFS) ارائه می‌شود. در فصل دوم پیشینه مختصری از مطالعات روش‌شناسی برآورد واریانس تغییرات شاخص‌های نیروی کار ایران ارائه می‌شود. در فصل سوم، روش پیشنهادی این پژوهش برای برآورد واریانس تغییرات شاخص‌های نیروی کار ایران به همراه یافته‌های طرح آمده است. در فصل چهارم نتیجه‌گیری و پیشنهادات کاربردی برای برآورد واریانس تغییرات شاخص‌های نیروی کار ایران ارائه می‌شود.

۱-۱- تجربه سایر کشورها

در سال‌های اخیر با توجه به اهمیت برآورد تغییرات شاخص‌های مورد نظر در آمارگیری‌های در طول زمان، روش‌های مختلفی برای برآورد واریانس تغییرات شاخص‌ها بین مقاطع زمانی مختلف ارائه شده است تا بر اساس آن امکان اعتبارسنجی و اطمینان از درستی برآوردهای محاسبه‌شده برای تغییرات شاخص‌ها در طول زمان فراهم شود. در این بخش بر سه رویکرد مرسوم برآورد واریانس تغییرات تمرکز کرده و جزئیاتی از چگونگی به‌کارگیری این روش‌ها برای داده‌های آمارگیری‌های خانواری به‌ویژه آمارگیری نیروی کار ارائه می‌شود.

۱-۱-۱- راهکار برآورد واریانس تغییرات برگر و پرایم (۲۰۱۳)، آووز-آلپر و برگر (۲۰۱۵)

در این بخش، برآوردگر ارائه شده توسط برگر و پرایم (۲۰۱۳) و آووز-آلپر و برگر (۲۰۱۵) ارائه شده است. این برآوردگر همچنین برای تجزیه و تحلیل آمارهای اتحادیه اروپا در خصوص درآمد و شرایط زندگی^۲، EU-SILC و محاسبه شاخص‌های تغییرات فقر، نابرابری و محرومیت اجتماعی توصیه شده است (آمار اروپا، ۲۰۱۳).

^۱ Labour Force Survey

^۲ The European Union Statistics on Income and Living Conditions

برآوردگر پیشنهادی چندین مزیت دارد. از نظر تئوری معتبر است، برای اکثر طرح‌های نمونه‌گیری نیروی کار در کشورهای عضو اتحادیه اروپا قابل استفاده است و با استفاده از نرم‌افزارهای SAS، SPSS، Stata یا R قابل محاسبه و پیاده‌سازی است.

در بخش بعدی، رویکرد برآورد واریانس در مورد شاخص‌های خطی (به‌عنوان مثال کل جمعیت، میانگین یا نسبت‌ها) و پس از آن شاخص‌های غیر خطی مانند شاخص‌های نسبت، ارائه شده است. همچنین رویکرد گسترش این روش برای برآورد واریانس تغییرات سالانه شاخص‌های نیروی کار آمده است.

۱-۱-۱-۱- برآورد واریانس تغییرات شاخص‌های خطی

فرض کنید می‌خواهیم تغییر Δ را بین دو مجموع $\tau_1 = \sum_{i \in U_1} y_{1i}$ و $\tau_2 = \sum_{i \in U_2} y_{2i}$ از دوره اول و دوم آمارگیری برآورد کنیم، که در آن y_{1i} و y_{2i} به ترتیب مقادیر متغیر مورد نظر را در دوره آمارگیری اول و دوم نشان می‌دهند. دوره‌های آمارگیری می‌توانند مربوط به فصل‌ها یا سال‌ها باشند.

به‌عنوان مثال، τ_1 ممکن است تعداد کل افراد شاغل (به‌طور مشابه بیکار) در دوره آمارگیری اول و τ_2 تعداد کل افراد شاغل (به‌طور مشابه بیکار) در دوره آمارگیری دوم باشد. در این حالت، $y_{1i} = 1$ برای فرد i که در دوره آمارگیری اول شاغل است و در غیر این صورت $y_{1i} = 0$ ، برای دوره آمارگیری دوم نیز به همین ترتیب y_{2i} تعریف می‌شود.

تغییر Δ به‌صورت زیر برآورد می‌شود:

$$\hat{\Delta} = \hat{\tau}_2 - \hat{\tau}_1$$

که $\hat{\tau}_2 = \sum_{i \in S_2} \omega_{2i} y_{2i}$ و $\hat{\tau}_1 = \sum_{i \in S_1} \omega_{1i} y_{1i}$ برآوردهای τ_2 و τ_1 براساس نمونه‌های مقطع S_2 و S_1 در دوره آمارگیری اول و دوم، و ω_{1i} و ω_{2i} به ترتیب وزن نمونه‌گیری در دوره آمارگیری اول و دوم را نشان می‌دهند. واریانس $\hat{\Delta}$ به‌صورت زیر محاسبه می‌شود:

$$\begin{aligned} \text{var}(\hat{\Delta}) &= \text{var}(\hat{\tau}_2 - \hat{\tau}_1) \\ &= \text{var}(\hat{\tau}_2) + \text{var}(\hat{\tau}_1) - 2\text{cov}(\hat{\tau}_2, \hat{\tau}_1) \\ &= \text{var}(\hat{\tau}_2) + \text{var}(\hat{\tau}_1) - 2\sqrt{\text{var}(\hat{\tau}_2)}\sqrt{\text{var}(\hat{\tau}_1)}\rho(\hat{\tau}_2, \hat{\tau}_1) \end{aligned}$$

واریانس‌های مقطعی $\text{var}(\hat{\tau}_2)$ و $\text{var}(\hat{\tau}_1)$ ممکن است با استفاده از هر برآوردگر واریانس مبتنی بر فرض استقلال نمونه‌ها محاسبه شود. به‌عنوان مثال، ممکن است از برآوردگر پیشنهاد شده Net-SILC، (اداره آمار اروپا، ۲۰۱۳)، که به تقریب «خوشه نهایی» متکی است، استفاده شود:

$$\text{var}(\hat{\tau}_l) = \sum_{h=1}^H \frac{n_h}{n_h-1} \sum_{i=1}^{n_h} (y_{lhi} - \bar{y}_{lh..})^2 \quad l \in \{1, 2\}$$

که

$$\bar{y}_{lh..} = n_h^{-1} (\sum_{j=1}^{n_h} y_{lhi.}) \text{ و } y_{lhi.} = \sum_{j=1}^{m_{lhi}} \omega_{lhij} \cdot y_{lhij}$$

- علامت h برچسب طبقه و H تعداد طبقات است. اگر هیچ طبقه‌بندی وجود نداشته باشد، می‌توان کل جمعیت هدف را به‌عنوان یک طبقه واحد در نظر گرفت ($H = 1$).
 - علامت i برچسب واحد نمونه‌گیری اولیه (PSU)^۳ است. در هر طبقه به تعداد n_h ، واحد نمونه‌گیری اولیه وجود دارد.
 - نماد j برچسب خانوار در i مین PSU در طبقه h ، با مجموع m_{lhi} خانوار است. برای طرح‌های نمونه‌گیری تک‌مرحله‌ای، می‌توان هر خانوار را به‌عنوان یک واحد PSU در نظر گرفت.
 - y_{lhij} مقدار متغیر مورد مطالعه y_l برای j امین خانوار در i مین PSU در طبقه h است.
 - ω_{lhij} وزن نمونه‌گیری برای j امین خانوار در i مین PSU در طبقه h است.
- محاسبه ضریب همبستگی $\rho(\hat{t}_1, \hat{t}_2)$ بین برآوردهای مقطعی τ_1 و τ_2 قطعاً دشوارترین بخش برای برآورد است و در اغلب موارد به دلیل تداخل دوره‌های آمارگیری نمی‌توان آن را نادیده گرفت. نادیده گرفتن ضریب همبستگی منجر به بیش‌برآوردی واریانس $\hat{\Delta}$ و سپس تفسیر نادرست تفاوت‌های مشاهده شده می‌شود. کیش (۱۹۶۵) پیشنهاد می‌کند که همبستگی با محاسبه نمونه مشترک در موج‌های اول و دوم برآورد شود. برآوردگر کیش (۱۹۶۵) اریب است زیرا همبستگی در بین واحدهای نمونه مشترک قوی‌تر است. در نتیجه، واریانس $\hat{\Delta}$ ممکن است کم تخمین زده شود و سپس اختلافات مشاهده‌شده بیش از حد تفسیر شود.
- برآوردگر پیشنهادی برگر و پرایم (۲۰۱۳) و آووز-آلپر و برگر (۲۰۱۵) براساس ماتریس مانده‌های مدل رگرسیون چندمتغیره است. این مدل شامل متغیرهای کمکی است که مؤلفه‌های طبقه‌بندی را مشخص می‌کند، همچنین اثر متقابل چرخش طرح‌های نمونه‌گیری را مشخص می‌کند:

$$\begin{pmatrix} y_{1i} \\ y_{2i} \end{pmatrix} = \sum_{h=1}^H \begin{pmatrix} \beta_h^{(1)} z_{1h,i} + \beta_h^{(2)} z_{2h,i} + \beta_h^{(12)} z_{1h,i} z_{2h,i} \\ \gamma_h^{(1)} z_{1h,i} + \gamma_h^{(2)} z_{2h,i} + \gamma_h^{(12)} z_{1h,i} z_{2h,i} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix}$$

که $\varepsilon_i = \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix}$ و $i \in S = S_1 \cup S_2$ از توزیع نرمال دومتغیره با میانگین صفر و ماتریس واریانس کوواریانس نامعلوم پیروی می‌کند. $z_{1h,i}$ و $z_{2h,i}$ متغیرهای طراحی (ظاهری) هستند که طبقه‌بندی و دوره‌ی آمارگیری را مشخص می‌کنند:

$$z_{1h,i} = \begin{cases} 1 & \text{اگر } i \in S_{1h} \\ 0 & \text{در غیر این صورت} \end{cases} \quad z_{2h,i} = \begin{cases} 1 & \text{اگر } i \in S_{2h} \\ 0 & \text{در غیر این صورت} \end{cases}$$

$z_{1h,i}$ در صورتی مقدار ۱ اختیار می‌کند که واحد نمونه i ام در دوره‌ی اول آمارگیری در طبقه h باشد، حوزه مقادیر $z_{2h,i}$ برای دوره‌ی دوم آمارگیری به‌صورت مشابه تعریف می‌شود.

$$\beta_h^{(1)}, \beta_h^{(2)}, \gamma_h^{(1)}, \gamma_h^{(2)}, \gamma_h^{(12)}, \beta_h^{(12)}$$

³ Primary Sampling Unit

با فرض اینکه طرح نمونه‌گیری دارای یک آنتروپی بزرگ است، همبستگی بین مانده‌های رگرسیون برآوردشده $\varepsilon_{1i}, \varepsilon_{2i}$ برآورد همبستگی $\rho(\hat{\varepsilon}_1, \hat{\varepsilon}_2)$ بین $\hat{\varepsilon}_1$ و $\hat{\varepsilon}_2$ را به دست می‌دهد. برگر و پرایم (۲۰۱۳) نشان دادند که این برآوردگر نیز سازگار با طراحی^۵ است. مطالعات شبیه‌سازی^۶ نشان داده‌اند که برآوردگر پیشنهادی از کیش (۱۹۶۵)، تام (۱۹۸۴) و کوالیت و تیل (۲۰۰۸) بهتر از برآوردگرهای سنتی است. علاوه بر این، با استفاده از تقریب «خوشه نهایی» می‌توان به راحتی این رویکرد را برای کنار آمدن با طرح‌های نمونه‌گیری چندمرحله‌ای گسترش داد (به عنوان مثال سارندال و همکاران، ۱۹۹۲). با این حال، این تقریب با این فرض است که کسرهای نمونه‌گیری قابل اغماض هستند، هرچند ممکن است در عمل چنین نباشد.

۱-۱-۲- برآورد واریانس تغییرات شاخص‌های غیر خطی

برآوردگر واریانس معرفی شده در بخش قبلی می‌تواند گسترش یابد تا برای محاسبه شاخص‌های پیچیده غیر خطی نیز قابل استفاده باشد. ایده آن، استفاده از روش خطی‌سازی (سارندال و همکاران ۱۹۹۲، دوئل ۱۹۹۹، دنماتی و راثو ۲۰۰۴، ولتر ۲۰۰۷ و اوایسر ۲۰۰۹) به منظور برآورد تقریبی یک آماره پیچیده غیر خطی با یک تابع خطی از مشاهدات، بر اساس خصوصیات مجانبی برآوردگر است. اگر $\hat{\theta}$ یک آماره غیرخطی باشد فرض کنید که $\hat{\theta}$ می‌تواند به عنوان یک تابع هموار از مجموع کل برآوردشده $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_p$ بیان شود.

$$\hat{\varepsilon}_k = \sum_{i \in S} \omega_i \cdot y_{ki}$$

y_{ki} مقدار متغیر مطالعه y_k ، ایده آن، استفاده از روش خطی‌سازی (سارندال و همکاران ۱۹۹۲، دوئل ۱۹۹۹، دنماتی و راثو ۲۰۰۴، ولتر ۲۰۰۷ و اوایسر ۲۰۰۹) به منظور برآورد تقریبی یک آماره پیچیده غیر خطی با یک تابع خطی از مشاهدات، بر اساس خصوصیات مجانبی برآوردگر است. $k = 1, \dots, p$ برای واحد i و ω_i وزن‌های نمونه‌گیری است. تقریب مرتبه اول سری تیلور برای $\hat{\theta}$ به صورت زیر است:

$$\hat{\theta} = f(\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_p) \cong \sum_{i \in S} \omega_i \cdot u_i + R_n$$

که در آن باقیمانده R_n با افزایش تعداد نمونه به سمت صفر میل می‌کند. در نتیجه، واریانس $\hat{\theta}$ به صورت مجانبی به بخش خطی سری تیلور یعنی $\hat{\varepsilon} = \sum_{i \in S} \omega_i u_i$ میل می‌کند.

$$\text{var}(\hat{\varepsilon}) \cong \text{var}(\hat{\theta})$$

u_i را متغیر «خطی‌شده» برای i می‌نامند.^۷

رویکرد خطی‌سازی بر اساس توابع تأثیر (دوئل، ۱۹۹۹) به اندازه کافی عمومیت دارد که می‌تواند با اکثر شاخص‌های معمول، از جمله شاخص‌هایی که نمی‌توان آن‌ها را به عنوان یک تابع هموار از مجموع بیان کرد (به عنوان مثال چندک‌ها و ضریب جینی) مورد استفاده قرار داد.

فرض کنید هدف برآورد کردن تغییر Δ بین دو پارامتر غیر خطی θ_1 و θ_2 از دوره‌های آمارگیری اول و دوم (سالانه یا فصلی) باشد.

^۴ Design Consistent

^۵ آندرسون و همکاران (۲۰۱۱) و موارد قابل تحویل پروژه DACSEIS در <https://www.uni-trier.de/> مراجعه کنید.

^۶ برای مثال به اوایسر ۲۰۰۹ مراجعه کنید.

$$\hat{\Delta} = \hat{\theta}_\gamma - \hat{\theta}_\alpha$$

که $\hat{\theta}_\gamma$ و $\hat{\theta}_\alpha$ برآوردگرهای استاندارد θ_γ و θ_α بر اساس نمونه‌های مقطعی s_γ و s_α از دوره آمارگیری اول و دوم هستند. به عنوان مثال، اگر θ_α نسبت مجموع دو صفت در جامعه هدف (مانند نرخ اشتغال یا بیکاری) در t_α و θ_γ نسبت مشابه در t_γ باشد، آنگاه داریم:

$$\hat{\theta}_\alpha = \frac{\sum_{i \in s_\alpha} \omega_{\alpha i} \cdot y_{\alpha i}}{\sum_{i \in s_\alpha} \omega_{\alpha i} \cdot x_{\alpha i}} \quad \hat{\theta}_\gamma = \frac{\sum_{i \in s_\gamma} \omega_{\gamma i} \cdot y_{\gamma i}}{\sum_{i \in s_\gamma} \omega_{\gamma i} \cdot x_{\gamma i}}$$

فرض کنید $\hat{t}_\gamma = \sum_{i \in s_\gamma} \omega_{\gamma i} u_{\gamma i}$ یک تقریب خطی از $\hat{\theta}_\gamma$ و $\hat{t}_\alpha = \sum_{i \in s_\alpha} \omega_{\alpha i} u_{\alpha i}$ یک تقریب خطی از $\hat{\theta}_\alpha$ است. به عنوان مثال، در مورد شاخص‌های نسبت، فرمول زیر به دست می‌آید:

$$u_{\alpha i} = \frac{1}{\sum_{i \in s_\alpha} \omega_{\alpha i} \cdot x_{\alpha i}} (y_{\alpha i} - \hat{\theta}_\alpha x_{\alpha i}) \quad u_{\gamma i} = \frac{1}{\sum_{i \in s_\gamma} \omega_{\gamma i} \cdot x_{\gamma i}} (y_{\gamma i} - \hat{\theta}_\gamma x_{\gamma i})$$

از این رو، واریانس $\hat{\Delta}$ توسط فرمول زیر ارائه می‌شود:

$$\begin{aligned} \text{var}(\hat{\Delta}) &= \text{var}(\hat{\theta}_\gamma - \hat{\theta}_\alpha) \\ &\cong \text{var}(\hat{t}_\gamma - \hat{t}_\alpha) \\ &= \text{var}(\hat{t}_\gamma) + \text{var}(\hat{t}_\alpha) - 2\text{cov}(\hat{t}_\gamma, \hat{t}_\alpha) \\ &= \text{var}(\hat{t}_\gamma) + \text{var}(\hat{t}_\alpha) - 2\sqrt{\text{var}(\hat{t}_\gamma)}\sqrt{\text{var}(\hat{t}_\alpha)}\rho(\hat{t}_\gamma, \hat{t}_\alpha) \end{aligned}$$

واریانس‌های مقطعی $\text{var}(\hat{t}_\gamma)$ و $\text{var}(\hat{t}_\alpha)$ ممکن است با استفاده از برآوردگر پیشنهاد Net-SILC محاسبه شود. همبستگی $\rho(\hat{t}_\gamma, \hat{t}_\alpha)$ بین $\hat{t}_\gamma$ و $\hat{t}_\alpha$ از طریق مدل رگرسیون چندمتغیره ارائه شده توسط برگر و پرایم (۲۰۱۳) برآورد می‌شود، که متغیرهای مطالعه $y_{\alpha i}$ و $y_{\gamma i}$ با همتایان «خطی‌شده» خود $u_{\alpha i}$ و $u_{\gamma i}$ جایگزین می‌شوند:

$$\begin{pmatrix} u_{\alpha i} \\ u_{\gamma i} \end{pmatrix} = \sum_{h=1}^H \begin{pmatrix} \beta_h^{(1)} z_{\alpha h, i} + \beta_h^{(2)} z_{\gamma h, i} + \beta_h^{(12)} z_{\alpha h, i} z_{\gamma h, i} \\ \gamma_h^{(1)} z_{\alpha h, i} + \gamma_h^{(2)} z_{\gamma h, i} + \gamma_h^{(12)} z_{\alpha h, i} z_{\gamma h, i} \end{pmatrix} + \begin{pmatrix} \varepsilon_{\alpha i} \\ \varepsilon_{\gamma i} \end{pmatrix}$$

$\varepsilon_{\alpha i}$ و $\varepsilon_{\gamma i}$ مانده‌های رگرسیون برآوردشده و $\text{cor}(\hat{t}_\gamma, \hat{t}_\alpha)$ برآورد همبستگی بین $\hat{t}_\gamma$ و $\hat{t}_\alpha$ را نشان می‌دهد.

۱-۱-۳- گسترش روش برای تغییرات سالانه در شاخص‌های نیروی کار

شاخص‌های آمارگیری نیروی کار عمدتاً به صورت تعداد کل یا نرخ، محاسبه و گزارش می‌شوند. در این زیربخش روش بیان‌شده برای برآورد واریانس تغییرات به تفکیک نوع شاخص‌ها (تعداد یا نرخ) ارائه می‌شود.

• تعداد کل افراد شاغل (یا بیکار)

فرض کنید هدف، برآورد واریانس تغییر سالانه تعداد کل افراد شاغل (یا بیکار) بین سال‌های اول و دوم باشد

$$\hat{\Delta} = \hat{t}_2 - \hat{t}_1$$

که $\hat{t}_1$ و $\hat{t}_2$ میانگین برآوردهای فصلی در سال‌های اول و دوم هستند.

$$\hat{t}_1 = \frac{\hat{P}_1^{(1)} + \hat{P}_1^{(2)} + \hat{P}_1^{(3)} + \hat{P}_1^{(4)}}{4}$$

$$\hat{t}_2 = \frac{\hat{P}_2^{(1)} + \hat{P}_2^{(2)} + \hat{P}_2^{(3)} + \hat{P}_2^{(4)}}{4}$$

$$\bullet \quad \hat{P}_1^{(l)} = \sum_{i \in S_{1,l}} \omega_{1i}^{(l)} y_{1i}^{(l)} \quad \text{برآورد تعداد افراد شاغل در فصل } l \text{ سال اول } (l \in \{1, 2, 3, 4\}) \text{ و}$$

$$\bullet \quad \hat{P}_2^{(l)} = \sum_{i \in S_{2,l}} \omega_{2i}^{(l)} y_{2i}^{(l)} \quad \text{برآورد تعداد افراد شاغل در فصل } l \text{ سال دوم است.}$$

این برآوردها براساس آمارگیری نیروی کار (اروپا) و نمونه‌های فصلی $S_{1,l}$ و $S_{2,l}$ در فصل اول و دوم است. $\omega_{1i}^{(l)}$ و $\omega_{2i}^{(l)}$ به ترتیب وزن نمونه‌گیری برای نمونه‌های فصلی $S_{1,l}$ و $S_{2,l}$ را نشان می‌دهد و $y_{1i}^{(l)}$ و $y_{2i}^{(l)}$ نشانگر ظاهری^۷ و ۱ برای شاغلان است که اگر فرد i ام در فصل l از سال اول شاغل باشد مقدار ۱ و در غیر این صورت مقدار ۰ اختیار می‌کند. با استفاده از این نشانه‌ها، واریانس تغییر سالانه $\hat{\Delta}$ را می‌توان به صورت زیر نوشت:

$$\text{var}(\hat{\Delta})$$

$$= \text{var}(\hat{t}_2 - \hat{t}_1)$$

$$= \text{var}\left(\frac{\hat{P}_2^{(1)} + \hat{P}_2^{(2)} + \hat{P}_2^{(3)} + \hat{P}_2^{(4)}}{4} - \frac{\hat{P}_1^{(1)} + \hat{P}_1^{(2)} + \hat{P}_1^{(3)} + \hat{P}_1^{(4)}}{4}\right)$$

$$= \text{var}\left(\frac{\hat{P}_2^{(1)} + \hat{P}_2^{(2)} + \hat{P}_2^{(3)} + \hat{P}_2^{(4)}}{4}\right) + \text{var}\left(\frac{\hat{P}_1^{(1)} + \hat{P}_1^{(2)} + \hat{P}_1^{(3)} + \hat{P}_1^{(4)}}{4}\right) - 2$$

$$\times \text{Cov}\left(\frac{\hat{P}_2^{(1)} + \hat{P}_2^{(2)} + \hat{P}_2^{(3)} + \hat{P}_2^{(4)}}{4}, \frac{\hat{P}_1^{(1)} + \hat{P}_1^{(2)} + \hat{P}_1^{(3)} + \hat{P}_1^{(4)}}{4}\right)$$

⁷ Dummy indicator

$$\begin{aligned}
&= \frac{1}{16} \left[\sum_{l=1}^4 \text{var}(\hat{Y}_r^{(l)}) + 2 \sum_{l=1}^4 \sum_{\substack{m=1 \\ m>l}}^4 \text{cov}(\hat{Y}_r^{(l)}, \hat{Y}_r^{(m)}) + \sum_{l=1}^4 \text{var}(\hat{Y}_l^{(l)}) \right. \\
&\quad \left. + 2 \sum_{l=1}^4 \sum_{\substack{m=1 \\ m>l}}^4 \text{cov}(\hat{Y}_l^{(l)}, \hat{Y}_l^{(m)}) - 2 \sum_{l=1}^4 \sum_{\substack{m=1 \\ m>l}}^4 \text{cov}(\hat{Y}_r^{(l)}, \hat{Y}_l^{(m)}) \right] \\
&= \frac{1}{16} \left[\sum_{l=1}^4 \text{var}(\hat{Y}_r^{(l)}) + 2 \sum_{l=1}^4 \sum_{\substack{m=1 \\ m>l}}^4 \sqrt{\text{var}(\hat{Y}_r^{(l)})} \sqrt{\text{var}(\hat{Y}_r^{(m)})} \rho(\hat{Y}_r^{(l)}, \hat{Y}_r^{(m)}) + \sum_{l=1}^4 \text{var}(\hat{Y}_l^{(l)}) \right. \\
&\quad \left. + 2 \sum_{l=1}^4 \sum_{\substack{m=1 \\ m>l}}^4 \sqrt{\text{var}(\hat{Y}_l^{(l)})} \sqrt{\text{var}(\hat{Y}_l^{(m)})} \rho(\hat{Y}_l^{(l)}, \hat{Y}_l^{(m)}) \right. \\
&\quad \left. - 2 \sum_{l=1}^4 \sum_{\substack{m=1 \\ m>l}}^4 \sqrt{\text{var}(\hat{Y}_r^{(l)})} \sqrt{\text{var}(\hat{Y}_l^{(m)})} \rho(\hat{Y}_r^{(l)}, \hat{Y}_l^{(m)}) \right] \\
&= \frac{1}{16} X^T A X
\end{aligned}$$

که ماتریس‌های X و A به ترتیب شامل واریانس و کوواریانس فصلی هستند:

$$X = \begin{pmatrix} \sqrt{\text{var}(\hat{Y}_1^{(1)})} \\ \sqrt{\text{var}(\hat{Y}_1^{(2)})} \\ \sqrt{\text{var}(\hat{Y}_1^{(3)})} \\ \sqrt{\text{var}(\hat{Y}_1^{(4)})} \\ \sqrt{\text{var}(\hat{Y}_r^{(1)})} \\ \sqrt{\text{var}(\hat{Y}_r^{(2)})} \\ \sqrt{\text{var}(\hat{Y}_r^{(3)})} \\ \sqrt{\text{var}(\hat{Y}_r^{(4)})} \end{pmatrix}$$

$$A = \begin{pmatrix} \rho(\hat{Y}_1^{(1)}, \hat{Y}_1^{(1)}) & \rho(\hat{Y}_1^{(1)}, \hat{Y}_1^{(2)}) & \rho(\hat{Y}_1^{(1)}, \hat{Y}_1^{(3)}) & \rho(\hat{Y}_1^{(1)}, \hat{Y}_1^{(4)}) & -\rho(\hat{Y}_1^{(1)}, \hat{Y}_r^{(1)}) & -\rho(\hat{Y}_1^{(1)}, \hat{Y}_r^{(2)}) & -\rho(\hat{Y}_1^{(1)}, \hat{Y}_r^{(3)}) & -\rho(\hat{Y}_1^{(1)}, \hat{Y}_r^{(4)}) \\ \rho(\hat{Y}_1^{(2)}, \hat{Y}_1^{(1)}) & \rho(\hat{Y}_1^{(2)}, \hat{Y}_1^{(2)}) & \rho(\hat{Y}_1^{(2)}, \hat{Y}_1^{(3)}) & \rho(\hat{Y}_1^{(2)}, \hat{Y}_1^{(4)}) & -\rho(\hat{Y}_1^{(2)}, \hat{Y}_r^{(1)}) & -\rho(\hat{Y}_1^{(2)}, \hat{Y}_r^{(2)}) & -\rho(\hat{Y}_1^{(2)}, \hat{Y}_r^{(3)}) & -\rho(\hat{Y}_1^{(2)}, \hat{Y}_r^{(4)}) \\ \rho(\hat{Y}_1^{(3)}, \hat{Y}_1^{(1)}) & \rho(\hat{Y}_1^{(3)}, \hat{Y}_1^{(2)}) & \rho(\hat{Y}_1^{(3)}, \hat{Y}_1^{(3)}) & \rho(\hat{Y}_1^{(3)}, \hat{Y}_1^{(4)}) & -\rho(\hat{Y}_1^{(3)}, \hat{Y}_r^{(1)}) & -\rho(\hat{Y}_1^{(3)}, \hat{Y}_r^{(2)}) & -\rho(\hat{Y}_1^{(3)}, \hat{Y}_r^{(3)}) & -\rho(\hat{Y}_1^{(3)}, \hat{Y}_r^{(4)}) \\ \rho(\hat{Y}_1^{(4)}, \hat{Y}_1^{(1)}) & \rho(\hat{Y}_1^{(4)}, \hat{Y}_1^{(2)}) & \rho(\hat{Y}_1^{(4)}, \hat{Y}_1^{(3)}) & \rho(\hat{Y}_1^{(4)}, \hat{Y}_1^{(4)}) & -\rho(\hat{Y}_1^{(4)}, \hat{Y}_r^{(1)}) & -\rho(\hat{Y}_1^{(4)}, \hat{Y}_r^{(2)}) & -\rho(\hat{Y}_1^{(4)}, \hat{Y}_r^{(3)}) & -\rho(\hat{Y}_1^{(4)}, \hat{Y}_r^{(4)}) \\ -\rho(\hat{Y}_r^{(1)}, \hat{Y}_1^{(1)}) & -\rho(\hat{Y}_r^{(1)}, \hat{Y}_1^{(2)}) & -\rho(\hat{Y}_r^{(1)}, \hat{Y}_1^{(3)}) & -\rho(\hat{Y}_r^{(1)}, \hat{Y}_1^{(4)}) & \rho(\hat{Y}_r^{(1)}, \hat{Y}_r^{(1)}) & \rho(\hat{Y}_r^{(1)}, \hat{Y}_r^{(2)}) & \rho(\hat{Y}_r^{(1)}, \hat{Y}_r^{(3)}) & \rho(\hat{Y}_r^{(1)}, \hat{Y}_r^{(4)}) \\ -\rho(\hat{Y}_r^{(2)}, \hat{Y}_1^{(1)}) & -\rho(\hat{Y}_r^{(2)}, \hat{Y}_1^{(2)}) & -\rho(\hat{Y}_r^{(2)}, \hat{Y}_1^{(3)}) & -\rho(\hat{Y}_r^{(2)}, \hat{Y}_1^{(4)}) & \rho(\hat{Y}_r^{(2)}, \hat{Y}_r^{(1)}) & \rho(\hat{Y}_r^{(2)}, \hat{Y}_r^{(2)}) & \rho(\hat{Y}_r^{(2)}, \hat{Y}_r^{(3)}) & \rho(\hat{Y}_r^{(2)}, \hat{Y}_r^{(4)}) \\ -\rho(\hat{Y}_r^{(3)}, \hat{Y}_1^{(1)}) & -\rho(\hat{Y}_r^{(3)}, \hat{Y}_1^{(2)}) & -\rho(\hat{Y}_r^{(3)}, \hat{Y}_1^{(3)}) & -\rho(\hat{Y}_r^{(3)}, \hat{Y}_1^{(4)}) & \rho(\hat{Y}_r^{(3)}, \hat{Y}_r^{(1)}) & \rho(\hat{Y}_r^{(3)}, \hat{Y}_r^{(2)}) & \rho(\hat{Y}_r^{(3)}, \hat{Y}_r^{(3)}) & \rho(\hat{Y}_r^{(3)}, \hat{Y}_r^{(4)}) \\ -\rho(\hat{Y}_r^{(4)}, \hat{Y}_1^{(1)}) & -\rho(\hat{Y}_r^{(4)}, \hat{Y}_1^{(2)}) & -\rho(\hat{Y}_r^{(4)}, \hat{Y}_1^{(3)}) & -\rho(\hat{Y}_r^{(4)}, \hat{Y}_1^{(4)}) & \rho(\hat{Y}_r^{(4)}, \hat{Y}_r^{(1)}) & \rho(\hat{Y}_r^{(4)}, \hat{Y}_r^{(2)}) & \rho(\hat{Y}_r^{(4)}, \hat{Y}_r^{(3)}) & \rho(\hat{Y}_r^{(4)}, \hat{Y}_r^{(4)}) \end{pmatrix}$$

تمام مؤلفه‌های کوواریانس در ماتریس A را می‌توان با استفاده از روش پیشنهادی برگر و پرایم (۲۰۱۳) و روش‌های توصیف‌شده در زیربخش‌های قبلی برآورد کرد.

• نرخ اشتغال (یا بیکاری)

نرخ اشتغال (بیکاری) سالانه به‌عنوان نسبت تعداد افراد شاغل (بیکار) به جمعیت کل افراد فعال تعریف می‌شود. با استفاده از نمادگذاری قبل، نرخ اشتغال در سال اول و سال دوم را می‌توان به‌صورت زیر نوشت:

$$\hat{\theta}_1 = \frac{\hat{Y}_1^{(1)} + \hat{Y}_1^{(2)} + \hat{Y}_1^{(3)} + \hat{Y}_1^{(4)}}{\hat{Z}_1^{(1)} + \hat{Z}_1^{(2)} + \hat{Z}_1^{(3)} + \hat{Z}_1^{(4)}}$$

$$\hat{\theta}_y = \frac{\hat{Y}_y^{(1)} + \hat{Y}_y^{(2)} + \hat{Y}_y^{(3)} + \hat{Y}_y^{(4)}}{\hat{Z}_y^{(1)} + \hat{Z}_y^{(2)} + \hat{Z}_y^{(3)} + \hat{Z}_y^{(4)}}$$

که در آن

- $\hat{Y}_y^{(l)} = \sum_{i \in S_{y,l}} \omega_{yi}^{(l)} y_{yi}^{(l)}$ تعداد برآوردشده افراد شاغل در فصل l سال اول است.
- $\hat{Z}_y^{(l)} = \sum_{i \in S_{y,l}} \omega_{yi}^{(l)} z_{yi}^{(l)}$ تعداد برآوردشده افراد فعال در فصل l سال اول است.
- $\hat{Y}_y^{(l)} = \sum_{i \in S_{y,l}} \omega_{yi}^{(l)} y_{yi}^{(l)}$ تعداد برآوردشده افراد شاغل در فصل l سال دوم است.
- $\hat{Z}_y^{(l)} = \sum_{i \in S_{y,l}} \omega_{yi}^{(l)} z_{yi}^{(l)}$ تعداد برآوردشده افراد فعال در فصل l سال دوم است.

فرض کنید $\hat{R}_y^{(l)} = \frac{\hat{Y}_y^{(l)}}{\hat{Z}_y^{(l)}}$ نرخ اشتغال برآوردشده در فصل l سال باشد ($l \in \{1, 2, 3, 4\}$ و $y \in \{1, 2\}$)

با فرض اینکه در همان سال تعداد جمعیت فعال از یک فصل به فصل دیگر بدون تغییر باقی بماند، داریم:

$$\hat{\theta}_1 = \frac{\hat{R}_1^{(1)}\hat{Z}_1^{(1)} + \hat{R}_1^{(2)}\hat{Z}_1^{(2)} + \hat{R}_1^{(3)}\hat{Z}_1^{(3)} + \hat{R}_1^{(4)}\hat{Z}_1^{(4)}}{\hat{Z}_1^{(1)} + \hat{Z}_1^{(2)} + \hat{Z}_1^{(3)} + \hat{Z}_1^{(4)}} \cong \frac{\hat{R}_1^{(1)} + \hat{R}_1^{(2)} + \hat{R}_1^{(3)} + \hat{R}_1^{(4)}}{4}$$

$$\hat{\theta}_2 = \frac{\hat{R}_2^{(1)}\hat{Z}_2^{(1)} + \hat{R}_2^{(2)}\hat{Z}_2^{(2)} + \hat{R}_2^{(3)}\hat{Z}_2^{(3)} + \hat{R}_2^{(4)}\hat{Z}_2^{(4)}}{\hat{Z}_2^{(1)} + \hat{Z}_2^{(2)} + \hat{Z}_2^{(3)} + \hat{Z}_2^{(4)}} \cong \frac{\hat{R}_2^{(1)} + \hat{R}_2^{(2)} + \hat{R}_2^{(3)} + \hat{R}_2^{(4)}}{4}$$

بنابراین، می‌توان رویکردی مشابه قبل را برای برآورد تغییرات سالانه نرخ اشتغال، به کار گرفت، و نرخ اشتغال

فصلی $\hat{R}_y^{(l)}$ را جایگزین تعداد تخمینی افراد شاغل $\hat{Y}_y^{(l)}$ کرد.

فرض کنید می‌خواهیم تغییر نرخ اشتغال بین سال‌های ۱ و ۲ را برآورد کنیم:

$$\hat{\Delta} = \hat{\theta}_2 - \hat{\theta}_1$$

در این صورت به دست می‌آوریم:

$$\text{var}(\hat{\Delta}) = \frac{1}{16} X^T A X$$

که در آن ماتریس‌های X و A به صورت زیر هستند:

$$= \begin{pmatrix} \sqrt{\text{var}(\hat{R}_1^{(1)})} \\ \sqrt{\text{var}(\hat{R}_1^{(2)})} \\ \sqrt{\text{var}(\hat{R}_1^{(3)})} \\ \sqrt{\text{var}(\hat{R}_1^{(4)})} \\ \sqrt{\text{var}(\hat{R}_2^{(1)})} \\ \sqrt{\text{var}(\hat{R}_2^{(2)})} \\ \sqrt{\text{var}(\hat{R}_2^{(3)})} \\ \sqrt{\text{var}(\hat{R}_2^{(4)})} \end{pmatrix}$$

$$A = \begin{pmatrix} \rho(\mathcal{P}_1^{(1)}, \mathcal{P}_1^{(2)}) & \rho(\hat{R}_1^{(1)}, \hat{R}_1^{(2)}) & \rho(\hat{R}_1^{(1)}, \hat{R}_1^{(3)}) & \rho(\hat{R}_1^{(1)}, \hat{R}_1^{(4)}) & -\rho(\hat{R}_1^{(1)}, \hat{R}_2^{(1)}) & -\rho(\hat{R}_1^{(1)}, \hat{R}_2^{(2)}) & -\rho(\hat{R}_1^{(1)}, \hat{R}_2^{(3)}) & -\rho(\hat{R}_1^{(1)}, \hat{R}_2^{(4)}) \\ \rho(\hat{R}_1^{(2)}, \hat{R}_1^{(1)}) & \rho(\mathcal{P}_1^{(1)}, \mathcal{P}_1^{(2)}) & \rho(\hat{R}_1^{(2)}, \hat{R}_1^{(3)}) & \rho(\hat{R}_1^{(2)}, \hat{R}_1^{(4)}) & -\rho(\hat{R}_1^{(2)}, \hat{R}_2^{(1)}) & -\rho(\hat{R}_1^{(2)}, \hat{R}_2^{(2)}) & -\rho(\hat{R}_1^{(2)}, \hat{R}_2^{(3)}) & -\rho(\hat{R}_1^{(2)}, \hat{R}_2^{(4)}) \\ \rho(\hat{R}_1^{(3)}, \hat{R}_1^{(1)}) & \rho(\hat{R}_1^{(3)}, \hat{R}_1^{(2)}) & \rho(\mathcal{P}_1^{(1)}, \mathcal{P}_1^{(3)}) & \rho(\hat{R}_1^{(3)}, \hat{R}_1^{(4)}) & -\rho(\hat{R}_1^{(3)}, \hat{R}_2^{(1)}) & -\rho(\hat{R}_1^{(3)}, \hat{R}_2^{(2)}) & -\rho(\hat{R}_1^{(3)}, \hat{R}_2^{(3)}) & -\rho(\hat{R}_1^{(3)}, \hat{R}_2^{(4)}) \\ \rho(\hat{R}_1^{(4)}, \hat{R}_1^{(1)}) & \rho(\hat{R}_1^{(4)}, \hat{R}_1^{(2)}) & \rho(\hat{R}_1^{(4)}, \hat{R}_1^{(3)}) & \rho(\mathcal{P}_1^{(1)}, \mathcal{P}_1^{(4)}) & -\rho(\hat{R}_1^{(4)}, \hat{R}_2^{(1)}) & -\rho(\hat{R}_1^{(4)}, \hat{R}_2^{(2)}) & -\rho(\hat{R}_1^{(4)}, \hat{R}_2^{(3)}) & -\rho(\hat{R}_1^{(4)}, \hat{R}_2^{(4)}) \\ -\rho(\hat{R}_2^{(1)}, \hat{R}_1^{(1)}) & -\rho(\hat{R}_2^{(1)}, \hat{R}_1^{(2)}) & -\rho(\hat{R}_2^{(1)}, \hat{R}_1^{(3)}) & -\rho(\hat{R}_2^{(1)}, \hat{R}_1^{(4)}) & \rho(\mathcal{P}_2^{(1)}, \mathcal{P}_2^{(1)}) & \rho(\hat{R}_2^{(1)}, \hat{R}_2^{(2)}) & \rho(\hat{R}_2^{(1)}, \hat{R}_2^{(3)}) & \rho(\hat{R}_2^{(1)}, \hat{R}_2^{(4)}) \\ -\rho(\hat{R}_2^{(2)}, \hat{R}_1^{(1)}) & -\rho(\hat{R}_2^{(2)}, \hat{R}_1^{(2)}) & -\rho(\hat{R}_2^{(2)}, \hat{R}_1^{(3)}) & -\rho(\hat{R}_2^{(2)}, \hat{R}_1^{(4)}) & \rho(\hat{R}_2^{(2)}, \hat{R}_2^{(1)}) & \rho(\mathcal{P}_2^{(1)}, \mathcal{P}_2^{(2)}) & \rho(\hat{R}_2^{(2)}, \hat{R}_2^{(3)}) & \rho(\hat{R}_2^{(2)}, \hat{R}_2^{(4)}) \\ -\rho(\hat{R}_2^{(3)}, \hat{R}_1^{(1)}) & -\rho(\hat{R}_2^{(3)}, \hat{R}_1^{(2)}) & -\rho(\hat{R}_2^{(3)}, \hat{R}_1^{(3)}) & -\rho(\hat{R}_2^{(3)}, \hat{R}_1^{(4)}) & \rho(\hat{R}_2^{(3)}, \hat{R}_2^{(1)}) & \rho(\hat{R}_2^{(3)}, \hat{R}_2^{(2)}) & \rho(\mathcal{P}_2^{(1)}, \mathcal{P}_2^{(3)}) & \rho(\hat{R}_2^{(3)}, \hat{R}_2^{(4)}) \\ -\rho(\hat{R}_2^{(4)}, \hat{R}_1^{(1)}) & -\rho(\hat{R}_2^{(4)}, \hat{R}_1^{(2)}) & -\rho(\hat{R}_2^{(4)}, \hat{R}_1^{(3)}) & -\rho(\hat{R}_2^{(4)}, \hat{R}_1^{(4)}) & \rho(\hat{R}_2^{(4)}, \hat{R}_2^{(1)}) & \rho(\hat{R}_2^{(4)}, \hat{R}_2^{(2)}) & \rho(\hat{R}_2^{(4)}, \hat{R}_2^{(3)}) & \rho(\mathcal{P}_2^{(1)}, \mathcal{P}_2^{(4)}) \end{pmatrix}$$

تمام شرایط واریانس و کوواریانس در ماتریس‌های فوق را می‌توان با ترکیب روش خطی‌سازی و برآوردهای مناسب محاسبه کرد. با استفاده از نشانه‌های قبلی، متغیر «خطی‌شده» برای نرخ اشتغال:

$$u_{1i}^{(l)} = \frac{1}{\sum_{i \in S_1} \omega_{1i}^{(l)} z_{1i}^{(l)}} (y_{1i}^{(l)} - \hat{R}_1^{(l)} z_{1i}^{(l)})$$

$$u_{2i}^{(l)} = \frac{1}{\sum_{i \in S_2} \omega_{2i}^{(l)} z_{2i}^{(l)}} (y_{2i}^{(l)} - \hat{R}_2^{(l)} z_{2i}^{(l)})$$

$$l \in \{1, 2, 3, 4\}$$

برآوردگر واریانس پیشنهادی در این بخش برای برآورد تغییرات خالص سالانه برآوردگرهای نیروی کار هم آسان و هم انعطاف‌پذیر است. این رویکرد همچنین قادر به ارائه برآوردگرهای دامنه است و می‌تواند با استفاده از تکنیک‌های آماری استاندارد، در کلاس بزرگی از طرح‌ها و برآوردگرهای نمونه‌گیری استفاده شود. برآورد واریانس را می‌توان با استفاده از نرم‌افزارهای مختلف آماری محاسبه کرد. برای مثال عملیات ماتریسی که در ماژول SAS IML ارائه شده است، این امکان را فراهم می‌کند. با این حال، اجرای آن دشوار است زیرا برای هشت برآورد فصلی باید ماتریس واریانس و کوواریانس برآورد شوند. در این روش چند تقریب و محدودیت وجود دارد:

- برآوردگر واریانس ارائه‌شده توسط برگر و پرایم (۲۰۱۳) بر فرض یک طراحی آنتروپی بالا متکی است، که ممکن است در عمل اینگونه نباشد.
- از آنجا که روش خطی‌سازی بر اساس خصوصیات مجانبی توجیه می‌شود، اگر اندازه نمونه به اندازه کافی بزرگ نباشد، برآورد واریانس ممکن است قابل اعتماد نباشد. این امر انتظار می‌رود با برآورد دامنه کوچک مشکل‌ساز شود.
- در مورد طرح‌های نمونه‌گیری چندمرحله‌ای، باید فرض کرد که کسرهای نمونه‌گیری مرحله اول نمونه‌گیری برای معتبر بودن تقریب «خوشه نهایی» ناچیز است. این فرض ممکن است در عمل برقرار نباشد.

برای پیاده‌سازی این رویکرد، متغیرهای دقیق طراحی نمونه مورد نیاز است:

- در صورت استفاده، کد شناسه برای طبقات، واحدهای اولیه نمونه‌گیری (PSU)، خانوارها و افراد؛
- وزن‌های نمونه‌گیری؛
- هر واحد نمونه‌گیری در پایگاه داده باید یک شناسه منحصر به فرد دریافت کند که بیش از همه، باید در طول زمان و در دوره‌های آمارگیری مختلف ثابت باشد.

بنا بر آنچه گفته شد با توجه به کمبود احتمالی متغیرهای طرح نمونه‌گیری و همچنین مشکلات بالقوه کیفیت طرح نمونه‌گیری در مجموعه داده‌های آمارگیری نیروی کار اتحادیه اروپا، باید هر نتیجه عددی مبتنی بر این رویکرد با احتیاط به کار گرفته شود.

۱-۲- رویکرد ایتالیا در برآورد واریانس تغییرات شاخص‌های نیروی کار

در این بخش بر اساس مقاله سکارلی و همکاران (۲۰۱۷) در خصوص بهبود کیفیت برآورد واریانس در آمارگیری نیروی کار (LFS)، به راهکاری که کشور ایتالیا برای برآورد واریانس تغییرات شاخص‌های کلیدی نیروی کار از آن بهره می‌برد، اشاره می‌شود.