مجله‌ی بررسی‌های آمار رسمی ایران
سال ۳۱، شماره‌ی ۲، پاییز و زمستان ۱۳۹۹، صص ۵۳۱ –۵۴۷

تأثیر بی‌پاسخی واحدهای نمونه‌گیری اولیه بر برآورد واریانس تغییرات به روش جک‌نایف (مطالعه موردی: داده‌های آمارگیری نیروی کار سال‌های ۱۳۸۸ و ۱۳۸۹)

روشنک علی‌اکبری صبا[†،*]، مریم زنگنه[‡]، لیدا کلهری ندرآبادی[†] و مرجان نورینی[‡]

[†] پژوهشکده‌ی آمار

[‡] مرکز آمار ایران

چکیده. در این مقاله با توجه به اهمیت ارائه برآورد تغییرات شاخص‌های کلیدی آمارگیری نیروی کار به همراه واریانس آن، استفاده از روش جک‌نایف در برآورد واریانس تغییرات مورد بررسی قرار گرفته است. سپس به تأثیر بی‌پاسخی واحدهای نمونه‌گیری اولیه بر برآورد واریانس تغییرات به روش جک‌نایف پرداخته شده است و با انجام یک شبیه‌سازی بر داده‌های آمارگیری نیروی کار سال‌های ۱۳۸۸ و ۱۳۸۹ به بررسی تأثیر بی‌پاسخی بر دقت برآورد واریانس تغییرات شاخص‌ها و نماگرهای کلیدی آمارگیری پرداخته شده است. نتایج این بررسی نشان می‌دهد اگر یک شبه‌حوزه نمونه از یک نقطه زمانی به بعد در دوره‌های آمارگیری متوالی، بی‌پاسخ باشد بر برآورد واریانس تغییرات شاخص‌ها و نماگرهای کلیدی آمارگیری تأثیرگذار خواهد بود هرچند ممکن است تأثیر آن بر برخی از شاخص‌ها یا نماگرها اندک باشد.

واژگان کلیدی: آمارگیری نیروی کار، برآورد واریانس تغییرات، بی‌پاسخی، روش جک‌نایف.

---

* نویسنده‌ی عهده‌دار مکاتبات

دریافت: ۱۴۰۱/۱/۲۰، پذیرش: ۱۴۰۱/۵/۸.

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]

۱- مقدمه

آمارگیری نیروی کار یکی از مهم‌ترین آمارگیری‌های مرکز آمار ایران است که از سال ۱۳۸۴ اطلاعات پایه‌ای نیروی کار کشور را تولید می‌کند. این آمارگیری با هدف دستیابی به برآوردهای فصلی و سالانه شاخص‌های نیروی کار و تغییرات فصلی و سالانه آن‌ها طراحی شده است. به عبارت دیگر در این آمارگیری، علاوه بر برآورد سطوح، برآورد تغییرات شاخص‌های کلیدی نیز مورد نظر است. در این آمارگیری برای به دست آوردن برآورد سطوح و تغییرات شاخص‌ها با دقت مطلوب، از روش نمونه‌گیری چرخشی استفاده می‌شود. نمونه‌گیری چرخشی با ثابت نگه داشتن بخشی از واحدهای نمونه بین دو دوره آمارگیری و تغییر بقیه واحدها، به بهترین نحو، امکان برآورد سطوح و تغییرات را فراهم می‌کند [۳]. از این‌رو امکان برآورد تغییرات پارامترهای مورد نظر با دقت مطلوب در طول زمان با محاسبه واریانس تغییرات فراهم است. بدیهی است مقدار پایین واریانس تغییرات بر دقت بالای برآورد تغییرات دلالت دارد.

واریانس تغییرات شاخص‌ها و نماگرهای نیروی کار به روش‌های گوناگونی در کشورهای مختلف برآورد می‌شود. انجام محاسبات به کمک این روش‌ها در مواردی نیاز به اطلاعاتی دارد که در طول زمان و با بررسی نتیجه‌های حاصل از تعداد زیادی از اجراهای طرح حاصل می‌شود (از جمله برآورد ضریب همبستگی بین برآوردهای دوره‌های متوالی اجرای طرح) یا در مواردی تنها بر اساس برآورد خطای استاندارد برآوردهای مقطعی برآورد می‌شود. کنتی و دیویسون [۴] در سال ۱۹۹۹ در مقاله خود تحت عنوان «برآورد واریانس تحت روش‌های بازنمونه‌گیری برای آمارگیری‌های نیروی کار» به مرور روش‌های بازنمونه‌گیری از جمله روش جک‌نایف، روش تکرار مکرر متعادل و روش خودگردان برای برآورد واریانس تغییرات در آمارگیری‌های نیروی کار پرداختند. علاوه بر این روشی را معرفی کردند که برآوردی از واریانس خطی برآوردگر تغییرات تعداد افراد دارای ویژگی‌های مورد نظر را بین دو دوره‌ی متوالی آمارگیری ارائه می‌داد.

یکی از پرکاربردترین روش‌های بازنمونه‌گیری روش جک‌نایف است که کاربرد گسترده‌ای در برآورد پارامترهای غیر خطی و واریانس برآوردگرها دارد. این روش برای نخستین بار توسط کوئنولی [۶] در سال ۱۹۴۹ برای کاهش اریبی برآوردگرهای غیر خطی معرفی شد. پس از آن توکی [۸] در سال ۱۹۵۸ استفاده از روش جک‌نایف را برای محاسبه واریانس

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]

برآوردگرها پیشنهاد داد. از مزایای روش جک‌نایف، قابلیت استفاده از آن برای هر طرح نمونه‌گیری و هر نوع برآوردگری ازجمله برآوردگرهای غیر خطی مانند نرخ در جامعه است. با وجود مزایای شناخته‌شده روش جک‌نایف در برآورد واریانس تغییرات، کارایی نتایج آن می‌تواند به شرایط اجرای آمارگیری و عوامل تأثیرگذار بر کیفیت داده‌ها از جمله بی‌پاسخی نمونه‌ها بستگی داشته باشد. از این‌رو در این مقاله به تأثیر بی‌پاسخی واحدهای نمونه‌گیری اولیه بر برآورد واریانس تغییرات به روش جک‌نایف پرداخته می‌شود و با انجام یک شبیه‌سازی بر داده‌های آمارگیری نیروی کار سال‌های ۱۳۸۸ و ۱۳۸۹ به بررسی تأثیر بی‌پاسخی بر دقت برآورد واریانس تغییرات شاخص‌ها و نماگرهای کلیدی آمارگیری پرداخته می‌شود. به این منظور ابتدا در بخش دو کلیاتی از آمارگیری نیروی کار ایران ارائه می‌شود. سپس در بخش سه به ضرورت محاسبه واریانس تغییرات در آمارگیری‌های مکرر از جمله آمارگیری نیروی کار مرکز آمار ایران اشاره می‌شود. استفاده از روش جک‌نایف در برآورد واریانس تغییرات در بخش چهار به اختصار شرح داده می‌شود. در نهایت در بخش پنج با انجام یک شبیه‌سازی، تأثیر بی‌پاسخی واحدهای نمونه‌گیری اولیه بر برآورد واریانس تغییرات به روش جک‌نایف مورد بررسی قرار می‌گیرد و نتایج بررسی در بخش شش جمع‌بندی می‌شود.

۲- کلیات آمارگیری نیروی کار

آمارگیری نیروی کار یکی از مهم‌ترین آمارگیری‌های مرکز آمار ایران است که از سال ۱۳۸۴ با اجراهای فصلی، اطلاعات پایه‌ای مورد نیاز در حوزه نیروی کار کشور را تولید می‌کند. طراحی این آمارگیری به‌گونه‌ای انجام شده است که علاوه بر برآورد سطوح، برآورد تغییرات شاخص‌ها و نماگرهای کلیدی نیز قابل ارائه باشد. برای به دست آوردن هر دو نوع برآورد سطوح و تغییرات با دقت مطلوب، از روش نمونه‌گیری چرخشی استفاده می‌شود تا با ثابت نگه‌داشتن بخشی از واحدهای نمونه بین دو دوره آمارگیری و تغییر بقیه واحدها، امکان برآورد سطوح و تغییرات فراهم شود. برای اجرای آمارگیری نیروی کار، خانوارهای نمونه در هر فصل به‌گونه‌ای انتخاب می‌شوند که نیمی از نمونه‌ها در دو فصل متوالی و نیز دو فصل یکسان از دو سال متوالی مشترک باشند. بر این اساس از هر خانوار نمونه در طی فصل‌های مختلف، حداکثر چهار بار آمارگیری به‌عمل می‌آید.

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]

جامعه هدف این آمارگیری را مجموعه افرادی تشکیل می‌دهند که طبق تعریف، عضو خانوارهای معمولی ساکن یا گروهی در نقاط شهری یا روستایی کشور هستند. زمان آمارگیری از واحدهای نمونه، در هر فصل، سومین هفته تقویمی ماه میانی و هفته بعد از آن یعنی هفته پس از هفته مرجع است. در این آمارگیری، اطلاعات مربوط به خانوار از طریق مصاحبه رودرروی مأمور آمارگیر با مطلع‌ترین عضو خانوار و تکمیل پرسشنامه خانوار گردآوری می‌شود. اطلاعات هر یک از اعضای خانوار نیز از طریق مصاحبه رودرروی مأمور آمارگیر با هر یک از اعضای ده ساله و بیشتر خانوار و تکمیل پرسشنامه فردی گردآوری می‌شود.

۲-۱- چارچوب نمونه‌گیری

چارچوب مرحله اول آمارگیری، نمونه پایه آمارگیری‌های خانواری است که شامل واحدهای نمونه‌گیری مرحله‌ی اول (شبه‌حوزه‌ها) است. نمونه پایه، نمونه‌ای است که می‌توان از آن برای تأمین نیازهای چند آمارگیری یا چند دوره از یک آمارگیری، زیرنمونه‌هایی انتخاب کرد. هر شبه‌حوزه شامل یک حوزه، بخشی از یک حوزه بزرگ یا در مواردی مجموعه‌ای از چند حوزه کوچک است که بر اساس اطلاعات آخرین سرشماری نفوس و مسکن سال و تقسیمات کشوری سال به دست آمده است. چارچوب ثانویه طرح آمارگیری، فهرستی از نام و نشانی خانوارهای معمولی ساکن و گروهی در شبه‌حوزه‌های انتخاب‌شده از نمونه پایه است که از آن برای انتخاب خانوارهای نمونه مربوط به هر فصل آمارگیری و مراجعه به آن‌ها، استفاده می‌شود. برای ساخت چارچوب ثانویه طرح، از فایل آخرین سرشماری نفوس و مسکن استفاده شده است [۳].

۲-۲- روش نمونه‌گیری

واحدهای نمونه این طرح طی دو مرحله و با استفاده از روش‌های نمونه‌گیری احتمالی از فایل نمونه پایه انتخاب می‌شوند، به این ترتیب که در مرحله‌ی اول، شبه‌حوزه‌های نمونه در هر یک از طبقه‌ها به روش تصادفی ساده انتخاب می‌شوند، سپس با فرض این‌که ترتیب قرار گرفتن خانوارها در شبه‌حوزه‌ها تصادفی است و محدود کردن نمونه‌ها به ۱۵۰ خانوار اول شبه‌حوزه منجر به ایجاد اریبی در برآوردها نمی‌شود، ۱۵۰ خانوار اول شبه‌حوزه برای

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]

ساخت گروه‌های چرخش سه خانواری در نظر گرفته می‌شوند. در مرحله دوم، چهار گروه چرخش برای آمارگیری در هر فصل تعیین می‌شوند. شایان ذکر است در این آمارگیری، اطلاعات از خانوارهای ساکن در محل سکونت خانوارهای نمونه گردآوری می‌شود، به این معنا که درصورت نقل مکان یک خانوار نمونه، از خانوار جایگزین در همان مکان، آمارگیری به عمل می‌آید. [۳]

۲-۳- الگوی چرخش

الگوی چرخش انتخابی برای این آمارگیری، یک «الگوی ۲-۲-۲» است، یعنی در طول زمان از هر خانوار نمونه، چهار بار آمارگیری به عمل می‌آید، به این ترتیب که خانوار، دو فصل متوالی در نمونه است، سپس به طور موقت برای دو فصل متوالی از نمونه خارج می‌شود، بعد مجدداً برای دو فصل متوالی به نمونه باز می‌گردد و پس از آن برای همیشه از نمونه خارج می‌شود. به این ترتیب میزان تداخل نمونه بین دو فصل متوالی و دو فصل یکسان از دو سال متوالی، ۵۰ درصد و بین دو سال متوالی، ۵۵ درصد است که این تداخل اهداف آمارگیری را تأمین می‌کند [۳].

۲-۴- روش برآورد

برای دستیابی به برآوردهای مورد نظر، ابتدا داده‌های مربوط به هر یک از خانوارهای نمونه و اعضای آن‌ها وزن‌دهی و سپس از فرمول‌های برآورد استفاده می‌شود. وزن‌دهی در سه مرحله «اعمال وزن پایه»، «تعدیل وزن برای بی‌پاسخی کامل» و «تعدیل وزن بر اساس پیش‌بینی‌های جمعیتی» انجام می‌گیرد. وزنی که پس از اعمال این مراحل وزن‌دهی به دست می‌آید، تعیین می‌کند که هر فرد در نمونه، نماینده چند نفر در جامعه است. پس از تعدیل وزن و محاسبه وزن نهایی، برای برآورد تعداد در هر سطح جغرافیایی، از مجموع وزن واحدهای آماری در آن سطح و برای برآورد نرخ، از حاصل تقسیم برآورد تعداد در دو زیرمجموعه متناظر با صورت و مخرج نرخ استفاده می‌شود. برآورد تغییرات نرخ‌ها در دو دوره از تفاضل برآورد نرخ‌های آن دوره‌ها به دست می‌آید [۳].

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]

۳- ضرورت محاسبه واریانس تغییرات در آمارگیری‌های مکرر

همانگونه که اشاره شد، مهمترین مزیت استفاده از آمارگیری‌های چرخشی، فراهم آمدن امکان برآورد تغییرات برآوردهای مورد نظر با دقت مطلوب در طول زمان است. برای بررسی تغییرات برآوردها در طول زمان، محاسبه برآوردی از واریانس تغییرات مورد نیاز است. فرض کنید تغییر بین مقادیر مشاهده‌شده در دو دوره‌ی آمارگیری با فاصله زمانی s از رابطه

$$\Delta_{y_t}^{(s)} = y_t - y_{t-s}$$

به دست آید. ساده‌ترین تحلیل برای بررسی تغییرات پارامترها، برآورد تغییرات دو دوره متوالی آمارگیری به‌صورت $y_t - y_{t-1}$ است ($s = 1$). واریانس برآورد تغییرات بین دو دوره‌ی آمارگیری با فاصله s دوره از رابطه زیر به دست می‌آید:

$$Var(\Delta_{y_t}^{(s)}) = Var(y_t) + Var(y_{t-s}) - 2\sqrt{Var(y_t)}\sqrt{Var(y_{t-s})}Corr(y_t, y_{t-s})$$

در این رابطه $Corr(y_t, y_{t-s})$ ضریب همبستگی بین برآوردهای دو دوره‌ی آمارگیری t و $t-s$ است. از آن‌جایی که صفت مورد بررسی برای یک واحد آماری در دوره‌های مختلف اجرای آمارگیری اغلب دارای همبستگی مثبتی است، وجود نمونه‌های مشترک بین دوره‌های آمارگیری t و $t-s$ در آمارگیری‌های پانلی یا چرخشی، با افزایش مقدار ضریب همبستگی، منجر به کاهش واریانس برآورد تغییرات نسبت به آمارگیرهای مستقل و افزایش دقت برآورد تغییرات در این آمارگیری‌ها می‌شود. از آنجایی که محاسبه ضریب همبستگی بین برآوردهای دو دوره آمارگیری به ویژه در آمارگیری‌هایی با طرح‌های نمونه‌گیری پیچیده به‌راحتی امکان‌پذیر نیست روش‌های دیگری برای برآورد واریانس تغییرات توسط محققین معرفی و در آمارگیری‌های سازمان‌های آماری سایر کشورها به‌کار گرفته شده است. به‌عنوان مثال در سال‌های قبل در آمارگیری نیروی کار اداره آمار کانادا که به‌صورت ماهانه اجرا می‌شود برای برآورد ضریب تغییرات از رابطه زیر استفاده شده است [۷]:

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]

$$CV\ (y_2 - y_1) = \sqrt{1-\rho} \times \frac{\sqrt{y_1^2 cv^2(y_1) + y_2^2 cv^2(y_2)}}{(y_2 - y_1)}$$

که در آن y_1 و y_2 برآوردهای دو دوره و ρ ضریب همبستگی بین y_1 و y_2 است. بر این اساس واریانس تغییرات $y_2 - y_1$ از رابطه زیر قابل برآورد است

$$Var\left(y_2 - y_1\right) = \left(1-\rho\right)\left[Var\left(y_1\right) + Var\left(y_2\right)\right].$$

اداره‌ی آمار کانادا از نسبت تداخل نمونه برای تقریب ضریب همبستگی استفاده می‌کند، که با توجه به ماهیت آمارگیری چرخشی کانادا برآورد تغییرات ماه به ماه می‌توان فرض کرد $\rho = \frac{5}{6}$. مطالعات تجربی در کانادا نشان داده است که در سطح استان، این مقدار ρ تقریب خوبی برای ضریب همبستگی برآورد تعداد شاغلان ارائه می‌کند اما برای ضریب همبستگی برآورد تعداد بیکاران، مقدار ρ برابر با $\frac{9}{20}$ تقریب بهتری برای تغییرات ماه به ماه نتیجه می‌دهد.

در مطالعات پیشین داخلی، توحیدی و نمازی‌راد [۱] از روش مستقیم و مدل خطی جمع‌پذیر با خطاهای وابسته استفاده کرده و براورد نرخ بیکاری و واریانس تغییرات نرخ بیکاری بین دو فصل بهار و تابستان سال ۱۳۸۵ آمارگیری نیروی کار استان مرکزی را ارائه داده‌اند. هدف اصلی توحیدی و نمازی‌راد [۱]، ارائه برآوردگر مناسب برای پارامترهای اصلی و واریانس آن‌ها بوده است و به‌صورت خاص بر برآوردگر واریانس تغییرات تمرکز نداشته‌اند. با این وجود آن‌ها برآورد واریانس تغییرات را با استفاده از رابطه

$$Var(r - r') = Var(r) + Var(r') - 2\hat{\rho}P\sqrt{Var(r)Var(r')}$$

ارائه کرده‌اند که در آن r و r' برآورد پارامترهای جامعه، $\hat{\rho}$ ضریب همبستگی نمونه‌ای بین داده‌ها در دو فصل متوالی نمونه‌گیری و P میزان تداخل نمونه بین دو فصل متوالی است.

فلاح [۲] از روش برآورد واریانس تقریبی پیشنهاد شده توسط ورما [۹] و روش جک‌نایف (افرون و تیبشیرانی، [۵]) استفاده کرده است و برآورد واریانس متغیرهای جمعیت ۱۰ ساله و بیشتر، جمعیت فعال، جمعیت شاغل، جمعیت بیکار، جمعیت

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]

غیر فعال، نرخ مشارکت اقتصادی، نسبت اشتغال و نرخ بیکاری را برای فصل بهار سال ۱۳۸۴ و نیز کل سال ۱۳۸۴ محاسبه کرده است. فلاح [۲] همچنین برآورد واریانس تغییرات نرخ مشارکت اقتصادی، نسبت اشتغال و نرخ بیکاری در دو فصل بهار و تابستان سال ۱۳۸۴ را مورد بررسی قرار داده است.

از جمله روش‌های برآورد واریانس تغییرات که نیاز به برآورد ضریب همبستگی بین برآوردهای دو دوره آمارگیری و مشکلات محاسبه این برآورد در آمارگیری‌های پیچیده را رفع می‌کند روش‌های بازنمونه‌گیری از جمله روش جک‌نایف، روش تکرار مکرر متعادل و روش خودگردان است. در بخش بعد به اختصار به روش جک‌نایف در برآورد واریانس تغییرات اشاره می‌شود. این روش به‌دلیل سادگی انجام محاسبات، در برخی از کشورها برای برآورد واریانس تغییرات شاخص‌ها و نماگرهای کلیدی آمارگیری نیروی کار کاربرد دارد.

۴- روش جک‌نایف در برآورد واریانس تغییرات

روش جک‌نایف یکی از روش‌های زیرنمونه‌گیری است که کاربرد وسیعی در برآورد پارامترهای غیر خطی و واریانس برآوردگرها دارد. گام اول در محاسبه واریانس به روش جک‌نایف، تولید زیرنمونه‌های تکراری از داده‌های آمارگیری است. برای این منظور ابتدا در هر طبقه نمونه‌گیری، یک واحد نمونه‌گیری اولیه انتخاب می‌شود. این واحد نمونه‌گیری از نمونه حذف و وزن سایر واحدهای نمونه‌گیری آن طبقه به نحوی تعدیل می‌شود که حذف این واحد نمونه‌گیری را جبران کند. سپس برآورد مورد نظر بر اساس نمونه جدید، یعنی نمونه‌ای که پس از حذف واحد نمونه‌گیری اولیه به دست آمده است، با وزن‌های تعدیل‌شده مجدداً محاسبه می‌شود. با تکرار این فرایند برای هر واحد نمونه‌گیری اولیه در نمونه، برآورد مورد نظر محاسبه می‌شود و به تعداد واحدهای نمونه‌گیری اولیه، برآورد به دست می‌آید. برای محاسبه واریانس برآورد مورد نظر از تغییرپذیری این برآورد در نمونه‌های تکراری استفاده می‌شود. در صورتی که نمونه‌گیری با احتمال نابرابر و بدون جایگذاری انجام گیرد برآورد واریانسی که به روش جک‌نایف محاسبه می‌شود برآوردی محافظه‌کارانه (اریب به سمت بالا) است. برای جبران این اریبی گاهی از ضریب تصحیح جامعه متناهی در محاسبه برآورد واریانس جک‌نایف زیرنمونه‌ها استفاده می‌شود. در این بررسی با فرض

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]

بزرگ بودن اندازه نمونه به‌ویژه در آمارگیری‌های مکرر، از ضریب تصحیح جامعه متناهی صرف‌نظر شده است.

فرض کنید برآورد مورد نظر، تغییر پارامتر مورد بررسی در دو مقطع زمانی باشد و برای انتخاب نمونه در مقطع‌های زمانی مختلف از نمونه‌گیری چندمرحله‌ای طبقه‌بندی‌شده، استفاده شده باشد. همچنین فرض کنید طبقه نمونه‌گیری hام دارای j_h واحد نمونه‌گیری اولیه باشد. در این صورت تعداد کل نمونه‌های تکرارشده در روش جک‌نایف برابر خواهد بود با

$$J = \sum_{h=1}^{H} j_h$$

که در آن H تعداد طبقه‌های نمونه‌گیری است.

اگر برآورد پارامتر مورد بررسی را در مقطع زمانی iام با نماد $\hat{t}^{(i)}$ و برآورد جک‌نایف پارامتر مورد بررسی را در مقطع زمانی iام پس از حذف واحد نمونه‌گیری اولیه aام $(a = 1, 2, \dots, j_h)$ از طبقه hام با نماد $\hat{t}_{ha}^{(i)}$ $(h = 1, 2, \dots, H)$ نشان دهیم، اختلاف بین دو مقدار برآورد نهایی و اختلاف بین دو مقدار برآورد جک‌نایف برای پارامتر مورد بررسی در دو مقطع زمانی با استفاده از روابط زیر محاسبه می‌شود:

$$\widehat{D} = \hat{t}^{(2)} - \hat{t}^{(1)} \ , \qquad \widehat{D}_{ha} = \hat{t}_{ha}^{(2)} - \hat{t}_{ha}^{(1)}$$

با تعریف روابط قبل، برآورد واریانس تغییرات بر اساس روش جک‌نایف از رابطه زیر به دست می‌آید:

$$\hat{Var}(\widehat{D}) = \sum_{h=1}^{H} \frac{(j_h - 1)}{j_h} \sum_{a=1}^{j_h} (\widehat{D}_{ha} - \widehat{D})^2$$

برای بررسی عملکرد روش جک‌نایف در برآورد واریانس تغییرات، مقدار برآورد سطوح در فصل‌های متوالی، برآورد تغییرات آن‌ها و واریانس این برآوردها به‌همراه کران بالا و پایین حدود اطمینان هر یک از برآوردهای تغییرات، محاسبه و ارائه شده است. بررسی و مقایسه نتیجه حاصل از به‌کارگیری این روش‌ها بر اساس داده‌های آمارگیری نیروی کار ایران در سه فصل زمستان ۱۳۸۸، بهار ۱۳۸۹ و تابستان ۱۳۸۹ انجام شده است. صفت مورد بررسی بسته به ماهیت روش برآورد واریانس تغییرات، تعداد بیکار، تعداد شاغل، تعداد فعال یا

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]

نرخ بیکاری در نظر گرفته شده است. برای کسب نتایج بهتر، ابتدا نرخ بی‌پاسخی استان‌ها در سطح واحدهای نمونه‌گیری اولیه بررسی شده و بر اساس نتیجه‌ی حاصل، استان خراسان رضوی به‌عنوان جامعه آماری این بررسی انتخاب شده است. انتخاب این استان به‌دلیل نداشتن واحد نمونه‌گیری اولیه بی‌پاسخ در فصل‌های مورد بررسی است. کران بالا و پایین برآورد تغییرات نیز با فرض نرمال بودن توزیع برآورد تغییرات محاسبه شده است. نتایج در جداول ۱ و ۲ ارائه شده است.

جدول ۱- تغییرات بهار ۱۳۸۹ نسبت به زمستان ۱۳۸۸ در استان خراسان رضوی با روش جک‌نایف

صفت مورد نظر	برآورد فصل اول	برآورد فصل دوم	برآورد تغییرات	واریانس برآورد تغییرات	کران پایین فاصله اطمینان ۹۵ درصد	کران بالای فاصله اطمینان ۹۵ درصد
تعداد بیکاران	۲۶۸۸۴۸	۲۶۰۷۴۶	-۸۱۰۲	۲۵۸۰۱۷۱۶۰۲	-۱۰۷۶۶۱	۹۱۴۵۷
تعداد شاغلان*	۱۶۰۷۲۷۹	۱۸۹۴۸۱۳	۲۸۷۵۳۴	۵۲۳۷۱۹۱۲۶۴	۱۴۵۶۹۲	۴۲۹۳۷۶
تعداد فعالان*	۱۸۷۶۱۲۷	۲۱۵۵۵۵۹	۲۷۹۴۳۲	۸۸۳۴۹۵۳۷۷۶	۹۵۲۰۳	۴۶۳۶۶۱
نرخ بیکاری	۱۴٫۳	۱۲٫۱	-۲٫۲	۴٫۷	-۶٫۴	۲

* معنی‌دار بودن تغییر فصل دوم نسبت به فصل اول

جدول ۲- تغییرات تابستان ۱۳۸۹ نسبت به بهار ۱۳۸۹ در استان خراسان رضوی با روش جک‌نایف

صفت مورد نظر	برآورد فصل اول	برآورد فصل دوم	برآورد تغییرات	واریانس برآورد تغییرات	کران پایین فاصله اطمینان ۹۵درصد	کران بالای فاصله اطمینان ۹۵ درصد
تعداد بیکاران	۲۶۰۷۴۶	۱۹۲۲۵۷	-۶۸۴۸۹	۱۴۱۱۳۸۶۸۱۶	-۱۴۲۱۲۳	۵۱۴۵
تعداد شاغلان	۱۸۹۴۸۱۳	۱۸۹۰۴۳۴	-۴۳۷۹	۷۶۲۳۴۵۷۴۰۱	-۱۷۵۵۱۱	۱۶۶۷۵۳
تعداد فعالان	۲۱۵۵۵۵۹	۲۰۸۲۶۹۱	-۷۲۸۶۸	۷۳۳۰۶۶۸۱۴۵	-۲۴۰۶۸۲	۹۴۹۴۶
نرخ بیکاری	۱۲٫۱	۹٫۲	-۲٫۹	۲٫۸	-۶٫۲	۰٫۴

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]

۵-تأثیر بی‌پاسخی واحدهای نمونه‌گیری اولیه بر برآورد واریانس تغییرات به روش جک‌نایف

عمده‌ترین بی‌پاسخی‌های آمارگیری نیروی کار، بی‌پاسخی شبه‌حوزه (واحد نمونه‌گیری اولیه) و بی‌پاسخی خانوار در شبه‌حوزه است. برای کاهش تأثیر بی‌پاسخی بر نتایج این آمارگیری، پیش از محاسبه برآوردها از تعدیل بی‌پاسخی (بی‌پاسخی شبه‌حوزه و بی‌پاسخی خانوار) برای اصلاح وزن‌های نمونه‌گیری استفاده می‌شود.

در صورت استفاده از روش جک‌نایف در برآورد واریانس تغییرات، تأثیر بی‌پاسخی خانوار در شبه‌حوزه‌ها با تعدیل وزن اصلاح می‌شود اما در صورت وجود شبه‌حوزه بی‌پاسخ، تعدیل آن در روش جک‌نایف نیاز به بررسی بیش‌تری دارد. اگر بی‌پاسخی به دلایلی مانند عدم دسترسی به واحدهای نمونه‌گیری اولیه در دوره‌های متوالی آمارگیری رخ داده باشد، استفاده از روش معمول جک‌نایف امکان‌پذیر است، اما اگر بی‌پاسخی واحدهای نمونه‌گیری اولیه تنها در یک دوره آمارگیری رخ داده باشد برای محاسبه واریانس تغییرات، ایجاد رابطه‌ای یک به یک بین همه واحدهای نمونه‌گیری اولیه در دو دوره آمارگیری امکان‌پذیر نیست. به‌عنوان یکی از راه‌های رفع این مشکل می‌توان واحدهای نمونه‌گیری اولیه بی‌پاسخ در یک فصل را در فصل بعد نیز بی‌پاسخ در نظر گرفت. در این صورت اگر برآورد تغییرات با کل واحدهای نمونه‌گیری اولیه آمارگیری‌شده در دو فصل محاسبه شود لازم است در نتایج خروجی به این نکته اشاره شود که برای برآورد واریانس از واحدهای نمونه‌گیری کمتری استفاده شده است. به‌عبارت دیگر تنها از اطلاعات واحدهای نمونه‌گیری اولیه‌ای که در هر دو فصل مورد بررسی، آمارگیری شده‌اند برای محاسبه واریانس تغییرات استفاده شده است.

در این تحقیق برای بررسی تأثیر بی‌پاسخی بر دقت برآورد واریانس تغییرات، از یک مطالعه شبیه‌سازی بر اساس داده‌های استان خراسان رضوی استفاده شده است. در طراحی نمونه‌گیری این استان شبه‌حوزه‌ها به ۳۴ طبقه تقسیم شده‌اند و در مجموع از کل طبقه‌ها ۱۴۱ شبه‌حوزه در مرحله اول نمونه‌گیری انتخاب شده است. هیچ یک از شبه‌حوزه‌های نمونه‌ی این استان در سه فصل زمستان ۱۳۸۸، بهار ۱۳۸۹ و تابستان ۱۳۸۹ بی‌پاسخ نبوده‌اند. برای انجام این مطالعه ابتدا در هر تکرار یک واحد از واحدهای نمونه‌گیری اولیه (شبه‌حوزه‌ها) بی‌پاسخ فرض شده و سپس نتایج حاصل با حالت نبود بی‌پاسخی در

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

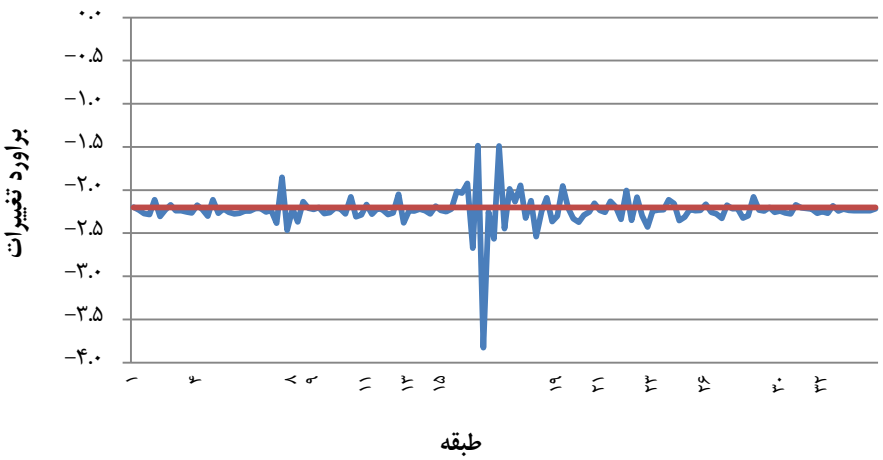
[DOR: 20.1001.1.25385798.1399.31.2.9.5]

واحدهای نمونه‌گیری اولیه مقایسه شده است. به این منظور طبقه‌های نمونه‌گیری در هر فصل به نحوی ادغام شده‌اند که حداقل سه شبه‌حوزه در هر طبقه وجود داشته باشد تا بی‌پاسخ فرض کردن واحدهای نمونه‌گیری خللی در انجام محاسبات واریانس ایجاد نکند. پس از ادغام طبقه‌ها، با بی‌پاسخ فرض کردن شبه‌حوزه‌ها در تکرارهای مختلف، ابتدا مقایسه‌ای بین مقدار این برآوردها و مقدار برآورد حاصل از طرح نمونه‌گیری، بدون وجود شبه‌حوزه بی‌پاسخ، انجام گرفته است. شکل‌های ۱ تا ۴ میزان تأثیر بی‌پاسخی بر برآوردهای کلیدی آمارگیری را نشان می‌دهند. در این شکل‌ها خطوط قرمزرنگ نشانگر مقدار برآورد تغییرات حاصل از طرح نمونه‌گیری، بدون وجود شبه‌حوزه بی‌پاسخ در نمونه است و خطوط آبی‌رنگ مقدار برآورد تغییرات را با فرض بی‌پاسخ بودن شبه‌حوزه‌های نمونه برای هر یک از شبه‌حوزه‌ها در همه طبقه‌ها (۳۴ طبقه) نشان می‌دهد.

شکل ۱ میزان تأثیر بی‌پاسخی شبه‌حوزه‌های نمونه را بر برآورد تغییرات نرخ بیکاری، به‌عنوان اصلی‌ترین متغیر مورد مطالعه در آمارگیری نیروی کار ایران، نشان می‌دهد. تعداد شبه‌حوزه‌ها در طبقه‌های نمونه‌گیری آمارگیری نیروی کار در استان خراسان رضوی بین ۳ تا ۱۲ تغییر می‌کند که بیشترین مقدار آن (۱۲) به طبقه ۱۶ نمونه‌گیری تعلق دارد. بیشترین تعداد شبه‌حوزه‌ها در سایر طبقه‌ها ۶ شبه‌حوزه است. همانطور که در این شکل مشاهده می‌شود تأثیر بی‌پاسخی شبه‌حوزه‌ها بر برآورد تغییرات نرخ بیکاری در طبقه‌هایی که تعداد شبه‌حوزه نمونه بیشتری دارند بیش از سایر طبقه‌ها است. احتمالاً یکی از دلایل این امر تغییر وزن تعداد بیشتری از شبه‌حوزه‌ها در این طبقه‌ها به دلیل تعدیل وزن برای بی‌پاسخی است که برآورد تغییرات متغیرهای مورد بررسی را تحت تأثیر قرار می‌دهد. بیشترین تأثیر بی‌پاسخی در طبقه‌ها مربوط به شبه‌حوزه‌ای است که بی‌پاسخی آن مقدار مطلق برآورد تغییرات نرخ بیکاری استان خراسان رضوی را در بهار ۱۳۸۹ نسبت به زمستان ۱۳۸۸ از عدد ۲/۲ به ۳/۸ افزایش می‌دهد.

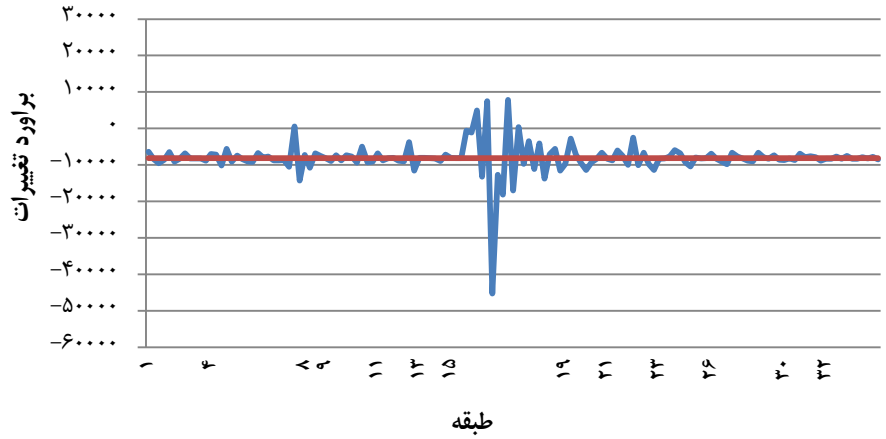
[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]



شکل ۱- برآورد تغییرات نرخ بیکاری بهار ۱۳۸۹ نسبت به زمستان ۱۳۸۸ با بی‌پاسخی شبه‌حوزه‌ها در طبقه‌ها

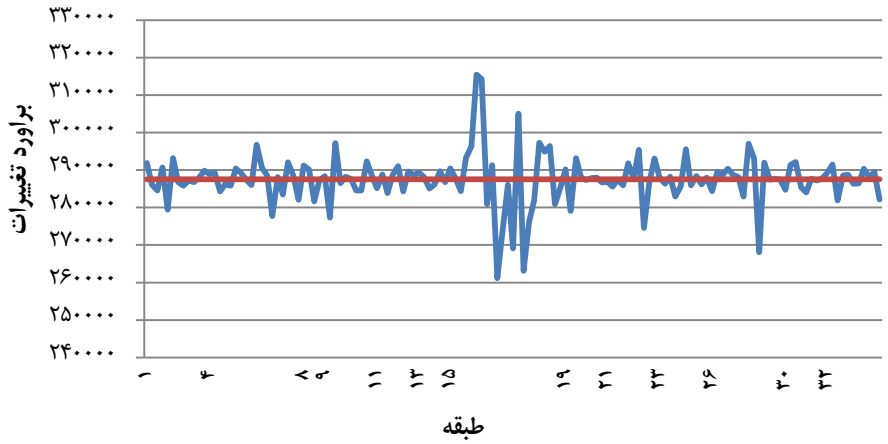
شکل‌های ۲ تا ۴ نشان می‌دهند تأثیر بی‌پاسخی شبه‌حوزه‌ها بر برآورد تغییرات متغیرهای تعداد افراد بیکار، تعداد افراد فعال و تعداد افراد شاغل نیز در طبقه‌هایی که تعداد شبه‌حوزه نمونه بیشتری دارند بیش از سایر طبقه‌ها است.



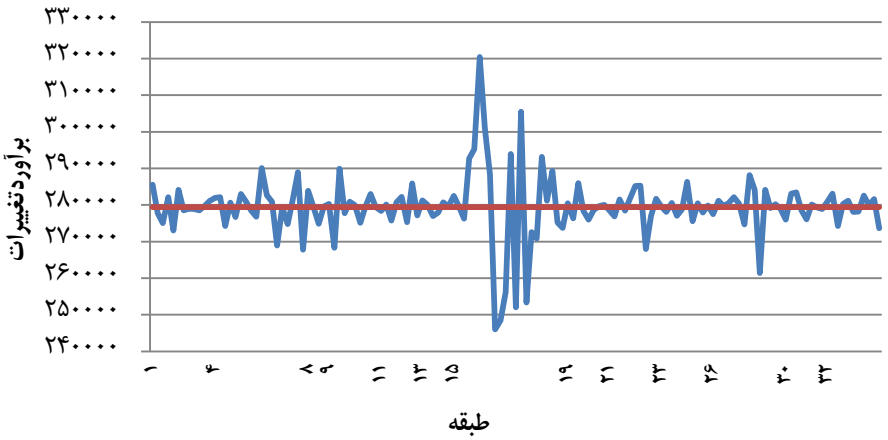
شکل ۲- برآورد تغییرات تعداد افراد بیکار بهار ۱۳۸۹ نسبت به زمستان ۱۳۸۸ با بی‌پاسخی شبه‌حوزه‌ها در طبقه‌ها

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]



شکل ۳- برآورد تغییرات تعداد افراد شاغل بهار ۱۳۸۹ نسبت به زمستان ۱۳۸۸ با بی‌پاسخی شبه‌حوزه‌ها در طبقه‌ها



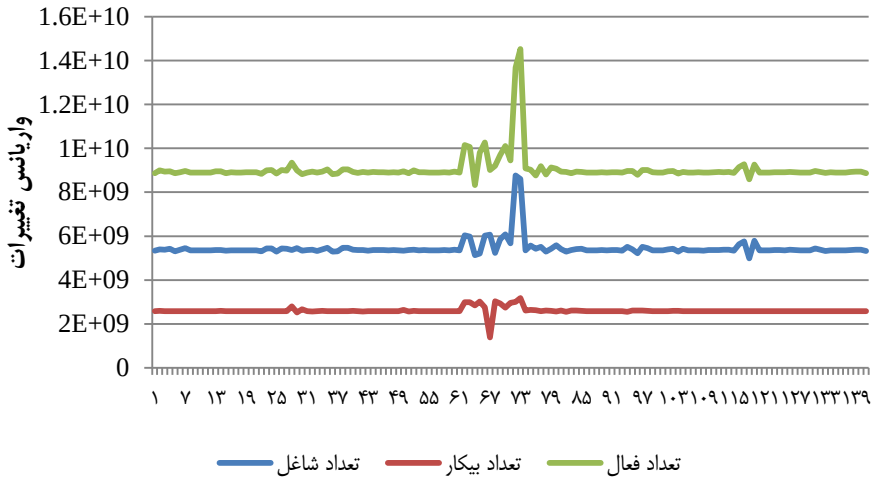
شکل ۴- برآورد تغییرات تعداد افراد فعال بهار ۱۳۸۹ نسبت به زمستان ۱۳۸۸ با بی‌پاسخی شبه‌حوزه‌ها در طبقه‌ها

در ادامه، برای بررسی تأثیر بی‌پاسخی واحدهای نمونه‌گیری اولیه بر برآورد واریانس تغییرات، از مقایسه‌ی واریانس تغییرات متغیرهای کلیدی تعداد افراد فعال، تعداد افراد بیکار، تعداد افراد شاغل و نرخ بیکای در بهار ۱۳۸۹ نسبت به زمستان ۱۳۸۸ با

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]

حضور شبه‌حوزه بی‌پاسخ در نمونه‌ها و در حالت نبود شبه‌حوزه بی‌پاسخ استفاده شده است. برای سادگی، نتیجه این بررسی ابتدا برای سه متغیر تعداد افراد بیکار، تعداد افراد شاغل و تعداد افراد فعال و سپس برای نرخ بیکاری ارائه می‌شود. تعداد دفعات محاسبه برآوردها برابر با تعداد کل واحدهای نمونه‌گیری اولیه در نمونه (۱۴۱) است. برآورد واریانس تغییرات تعداد افراد فعال، تعداد افراد بیکار و تعداد افراد شاغل به روش جک‌نایف با وجود یک شبه‌حوزه بی‌پاسخ در نمونه‌ها، در شکل ۵ نشان داده شده است. این شکل به وضوح میزان تأثیر بی‌پاسخی واحدهای نمونه‌گیری اولیه را بر برآورد واریانس تغییرات متغیرهای مورد بررسی به‌ویژه در طبقه‌های بزرگ نشان می‌دهد. شایان ذکر است در این شکل شبه‌حوزه‌های میانی (شماره‌ی ۶۲ تا ۷۳) متعلق به بزرگترین طبقه نمونه‌گیری (طبقه ۱۶) است.



شکل ۵ – برآورد واریانس تغییرات بهار ۱۳۸۹ نسبت به زمستان ۱۳۸۸ به روش جک‌نایف با حضور شبه‌حوزه‌ی بی‌پاسخ (بر حسب میلیون)

برای خلاصه کردن نتایج بررسی، از میانگین واریانس تغییرات متغیرهای تعداد افراد فعال، تعداد افراد بیکار و تعداد افراد شاغل بهار ۱۳۸۹ نسبت به زمستان ۱۳۸۸ در دو حالت وجود شبه‌حوزه بی‌پاسخ در نمونه‌ها و نبود آن استفاده شد که نتیجه این بررسی در

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]

جدول ۳ ارائه شده است. این جدول نشان می‌دهد که اگر یک شبه‌حوزه نمونه از یک نقطه زمانی به بعد در دوره‌های آمارگیری متوالی، بی‌پاسخ باشد (یا بی‌پاسخ فرض شود) حتی با تعدیل وزن‌های نمونه‌گیری، برآورد واریانس تغییرات تعداد افراد فعال، تعداد افراد بیکار و تعداد افراد شاغل افزایش اندکی خواهد داشت. بی‌پاسخی شبه‌حوزه‌ها کمترین تأثیر را بر برآورد تغییرات تعداد بیکاران دارد که مقدار آن کمتر از تعداد افراد فعال و تعداد افراد شاغل است.

با انجام مقایسه‌ای مشابه برای متغیر نرخ بیکاری نتیجه گرفته می‌شود که برآورد واریانس تغییرات نرخ بیکاری بهار ۱۳۸۹ نسبت به زمستان ۱۳۸۸ با حضور شبه‌حوزه بی‌پاسخ در نمونه‌ها به‌طور متوسط از عدد ۴/۷ (بدون بی‌پاسخی) به ۴/۰۳ (در حضور بی‌پاسخی) می‌رسد.

جدول ۳ – میانگین واریانس تغییرات متغیرهای کلیدی بهار ۱۳۸۹ نسبت به زمستان ۱۳۸۸ با وجود شبه‌حوزه بی‌پاسخ و بدون وجود شبه‌حوزه بی‌پاسخ در نمونه‌ها (بر حسب میلیون)

پارامتر	با وجود شبه‌حوزه‌ی بی‌پاسخ	بدون وجود شبه‌حوزه‌ی بی‌پاسخ
تعداد شاغل	۵/۴۵۰	۵/۲۳۷
تعداد بیکار	۲/۶۰۵	۲/۵۸۰
تعداد فعال	۹/۰۵۲	۸/۸۳۵

۶- جمع‌بندی و نتیجه‌گیری

شبیه‌سازی انجام‌شده در بخش قبل نشان داد تأثیر بی‌پاسخی شبه‌حوزه‌ها بر برآورد تغییرات شاخص‌های کلیدی آمارگیری نیروی کار شامل نرخ بیکاری، تعداد افراد بیکار، تعداد افراد فعال و تعداد افراد شاغل به‌خصوص در طبقه‌هایی که تعداد شبه‌حوزه‌های نمونه بیشتری دارند ممکن است قابل توجه باشد. این بررسی همچنین نشان داد اگر یک شبه‌حوزه نمونه از یک نقطه زمانی به بعد در دوره‌های آمارگیری متوالی، بی‌پاسخ باشد با تعدیل وزن‌های نمونه‌گیری، بر برآورد واریانس تغییرات شاخص‌های کلیدی آمارگیری تأثیرگذار خواهد بود هرچند ممکن است تأثیر اندکی باشد. بی‌پاسخی شبه‌حوزه‌ها در بین شاخص‌های تعداد مورد نظر، کمترین تأثیر را بر برآورد تغییرات تعداد بیکاران دارد که مقدار آن کمتر از تعداد افراد فعال و تعداد افراد شاغل است.

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]

مرجع‌ها

[۱] توحیدی، مینا؛ نمازی‌راد، محمدرضا (۱۳۸۷). برآورد واریانس در نمونه‌گیری چرخشی، پژوهشکده آمار.

[۲] فلاح محسن‌خانی، زهره (۱۳۸۹). واریانس برآوردگرها در آمارگیری از نیروی کار. مجله بررسی‌های آمار رسمی ایران، ۲۱، ۱-۱۲.

[۳] مرکز آمار ایران (۱۳۹۰). دستورالعمل طرح نمونه‌گیری آمارگیری نیروی کار ۱۳۹۰.

[4] Canty, A.J. and Davison, A.C. (1999). Resampling–based variance estimation for labour force surveys. Journal of the Royal Statistical Society: Series D (The Statistician), **48**, 379-391.

[5] Efron, B. and Tibshirani, R. (1993). An Introduction to the Bootstrap, Chapman & Hall/CRC, Florida.

[6] Quenouille, M.H. (1949). Problems in plane sampling. The Annals of Mathematical Statistics, 355-375.

[7] Statistics Canada (2010). Guide to the Labour Force Survey 2010.

[8] Tukey, J. (1958). Bias and confidence in not quite large samples. Ann. Math. Statist., **29**, 614.

[9] Verma, V. (2002). Sampling Methods. Training Handbook, Revised, SIAP, Tokyo.

روشنک علی‌اکبری صبا
دکتری آمار
تهران، خیابان فاطمی، خیابان باباطاهر، خیابان فکوری، پلاک ۱۴۵، پژوهشکده‌ی آمار.
رایانشانی: r_saba@srtc.ac.ir

مریم زنگنه
فوق‌لیسانس آمار
تهران، خیابان فاطمی، خیابان رهی‌معیری، مرکز آمار ایران.
رایانشانی: zangeneh_m9@yahoo.com

لیدا کلهری ندرآبادی
دکتری آمار
تهران، خیابان فاطمی، خیابان باباطاهر، خیابان فکوری، پلاک ۱۴۵، پژوهشکده‌ی آمار.
رایانشانی: kalhori@srtc.ac.ir

مرجان نورینی
فوق‌لیسانس آمار
تهران، خیابان فاطمی، خیابان رهی‌معیری، مرکز آمار ایران.
رایانشانی: m_noorini@sci.org.i

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[DOR: 20.1001.1.25385798.1399.31.2.9.5]