مجله‌ی بررسی‌های آمار رسمی ایران
سال ۲۲، شماره‌ی ۱، بهار و تابستان ۱۳۹۰، صص ۷۳ – ۹۵

کاربردی از براورد واریانس براوردگر مرکب تعمیم‌یافته در نمونه‌گیری چرخشی یک‌سطحی متعادل دوسویه

ابراهیم تالانه* و حمیدرضا نواب‌پور

دانشگاه علامه طباطبایی

چکیده: در بیش‌تر آمارگیری‌ها با مشخصه‌هایی سر و کار داریم که در طول زمان تغییر می‌کنند. در این آمارگیری‌ها ممکن است علاقه‌مند باشیم در فاصله‌های زمانی منظم پارامترهای جامعه را براورد کرده و یا تغییرات آن‌ها را از یک دوره به دوره‌ی بعد مطالعه کنیم. با توجه به گسترش استفاده از نمونه‌گیری چرخشی در واحدهای ملی آمار، انجام یک بررسی جامع درباره‌ی براورد واریانس در طرح‌های نمونه‌گیری چرخشی امری ضروری به نظر می‌رسد. نمونه‌گیری چرخشی یکی از روش‌های مورد استفاده در آمارگیری‌های مکرر است. در این مقاله ما با تأکید بر طرح‌های چرخشی یک‌سطحی متعادل دوسویه، از داده‌های طرح آمارگیری هزینه‌ی مصرف‌کننده‌ی ایالت‌های متحده برای مقایسه‌ی براورد واریانس‌های براورد مرکب تعمیم‌یافته و براورد ساده‌ی میانگین هزینه‌ی مصرف‌کننده استفاده کرده‌ایم. نتیجه‌های به دست آمده از این مطالعه نشان‌دهنده‌ی دقت بیش‌تر براورد مرکب تعمیم‌یافته‌ی میانگین نسبت به براورد ساده‌ی میانگین است.

واژگان کلیدی: آمارگیری مکرر؛ نمونه‌گیری چرخشی؛ براوردگر مرکب تعمیم‌یافته؛ طرح متعادل دوسویه؛ براورد واریانس.

۱- مقدمه

بسیاری از خصیصه‌های جامعه از جمله نرخ بیکاری، نرخ اشتغال و ...، مرتب از دوره‌ای به دوره‌ی بعد در حال تغییر هستند. برای دست‌یابی به روند تغییرات این خصیصه‌ها در طول زمان نیاز به آمارگیری‌هایی داریم که به‌طور منظم در طول زمان تکرار شوند. با

* نویسنده‌ی عهده‌دار مکاتبات

دریافت: ۱۳۸۹/۱۱/۱۷، پذیرش: ۱۳۹۰/۶/۱.

استفاده از داده‌هایی که در نمونه‌گیری‌های مکرر از یک جامعه در طول زمان به دست می‌آیند می‌توان به براورد پارامترهای مورد نظر پرداخت و تغییرات آن‌ها را از دوره‌ای به دوره‌ی بعد براورد کرد. یکی از مسائل مطرح در آمارگیری‌های مکرر، الگوی ترکیب در دوره‌های زمانی مختلف است. به‌طور کلی نمونه‌گیری‌های مکرر را به سه دسته تقسیم می‌کنند:

آ) نمونه‌گیری پانلی:
نمونه شامل همان مجموعه واحدهایی است که در دوره‌ی قبل، نمونه‌گیری شده‌اند.

ب) نمونه‌گیری مقطعی:
در این روش که ساده‌ترین نوع آمارگیری در طول زمان است، نمونه شامل مجموعه واحدهایی کاملاً متفاوت با واحدهای نمونه‌گیری‌شده در دوره‌ی قبل است.

ج) نمونه‌گیری چرخشی:
در این شیوه‌ی آمارگیری، نمونه شامل مجموعه‌ای از واحدهای نمونه‌گیری‌شده در دوره‌ی قبل و مجموعه‌ای از واحدهایی که در دوره‌ی قبل نمونه‌گیری نشده‌اند، است. واحدهای نمونه‌ای، طبق الگوی مشخص و از پیش تعیین‌شده‌ای در تعدادی از دوره‌ها، به نمونه وارد شده، سپس با گذشت چند دوره از نمونه خارج می‌شوند. در حقیقت نمونه‌گیری چرخشی، فرایند حذف برخی از عناصر قدیمی از نمونه و اضافه کردن عناصر جدید به آن است. این شیوه‌ی نمونه‌گیری توسط یتس [۷] و پاترسون [۶] "نمونه‌گیری روی فاصله‌های متوالی با جایگزینی جزئی عناصر" و توسط هَنسن و همکاران [۳] "نمونه‌گیری برای سری‌های زمانی" نامیده شده است. اگر رابطه‌ای بین مقدار یک واحد از جامعه در زمان t و مقدار تغییر کرده‌ی همان واحد در زمان متعاقب آن یعنی $t + \Delta t$ وجود داشته باشد، استفاده از اطلاعات قبلی این واحد، براورد جاری پارامتر جامعه را بهبود می‌بخشد. بزرگ‌ترین مزیت استفاده از نمونه‌گیری چرخشی این است که هنگامی که مقدارهای براورد شده‌ی یک دوره با دوره‌ی بعد همبستگی داشته باشند براورد تفاضل دارای واریانس کوچک‌تری است. واریانس تفاضل میانگین‌ها بین دو دوره‌ی زمانی با واحدهای نمونه‌ای متداخل را می‌توان توسط رابطه‌ی زیر نشان داد:

$$\operatorname{var}(\bar{x}_t - \bar{x}_{t-1}) = \operatorname{var}(\bar{x}_t) + \operatorname{var}(\bar{x}_{t-1}) - 2\operatorname{cov}(\bar{x}_t, \bar{x}_{t-1}).$$

$\overline{x_t}$ و $\overline{x_{t-1}}$ به‌ترتیب بیان‌گر میانگین واحدهای نمونه‌ای در دوره‌های زمانی آمارگیری tام و $t-1$ام هستند. کوواریانس در این عبارت به‌علت وجود تداخل میان واحدهای نمونه‌ای در دو دوره‌ی آمارگیری متوالی، مقداری غیر صفر و معمولاً مثبت است که با افزایش میزان تداخل بین دوره‌های آمارگیری، این مقدار افزایش یافته و در نتیجه واریانس تفاضل میانگین‌ها کاهش می‌یابد.

مطالعه‌های اولیه در زمینه‌ی نمونه‌گیری چرخشی به‌عنوان مسئله‌ای در گرداوری داده‌های کشاورزی در دو مقطع زمانی و با مقداری تداخل برای اولین بار توسط جِسِن در سال ۱۹۴۲ صورت گرفت [۴]. هَنسن و همکاران [۳] و پاترسون [۶] مطالعه‌های منسجمی را در ارتباط با براوردگرهای بهین در نمونه‌گیری‌های متداخل انجام دادند. برو و ارنست [۱] براوردگری با عنوان براوردگر مرکب تعمیم‌یافته ارایه دادند و کانت‌وِل [۲] فرمولی برای واریانس براوردگر مرکب تعمیم‌یافته در حالت کلی، ارایه کرد. پارک و همکاران [۵]، طرح نمونه‌گیری چرخشی یک سطحی متعادل دو سویه را ارایه کردند سپس واریانس و اریبی این براوردگر را برای طرح متعادل دوسویه به دست آوردند. آن‌ها همچنین ضریب‌های بهین را با استفاده از مینیمم کردن واریانس و همچنین میانگین توان‌های دوم خطا برای براوردگر مرکب تعمیم‌یافته به دست آوردند.

۲- نمونه‌گیری چرخشی یک‌سطحی

یکی از کاربردی‌ترین طرح‌های نمونه‌گیری چرخشی، نمونه‌گیری چرخشی یک‌سطحی است. در این طرح تنها اطلاعات مربوط به دوره‌ی آمارگیری جاری گرداوری می‌شود. در بیش‌تر طرح‌های آمارگیری نمونه‌ای چرخشی، جامعه به تعداد مناسبی از گروه‌های چرخش تقسیم می‌شود. تعداد گروه‌های چرخش در هر دوره، برابر با تعداد دوره‌های آمارگیری است که در هر مقطع مورد بررسی قرار می‌گیرد [۵].

از نمادگذاری $r_{11} - r_{21} - r_{12} - \cdots - r_{2,m-1} - r_{1m}$، برای نشان دادن الگوی چرخش در نمونه‌گیری چرخشی استفاده می‌شود. نماد r_{1i} نمایانگر تعداد دوره‌های آمارگیری است که یک واحد نمونه‌ای داخل نمونه و r_{2i} نمایانگر تعداد دوره‌های زمانی است که یک واحد نمونه‌ای خارج از نمونه است. فرایند ورود و خروج به تعداد m مرتبه قبل از خروج نهایی واحد نمونه‌ای از نمونه انجام می‌شود. به‌طور کلی در طرح نمونه‌گیری چرخشی، در هر دوره‌ی آمارگیری، $G = \sum_{i=1}^{m} r_{1i}$ گروه چرخش مورد مطالعه قرار خواهند گرفت. میزان

تداخل در مقطع‌های زمانی مختلف، بستگی به اهداف آمارگیری دارد؛ چنانچه ساختاری بهین قادر است پیشنهاد کند که یك واحد چه مدت در نمونه قرار گیرد و چه موقع از روند چرخش خارج شود.

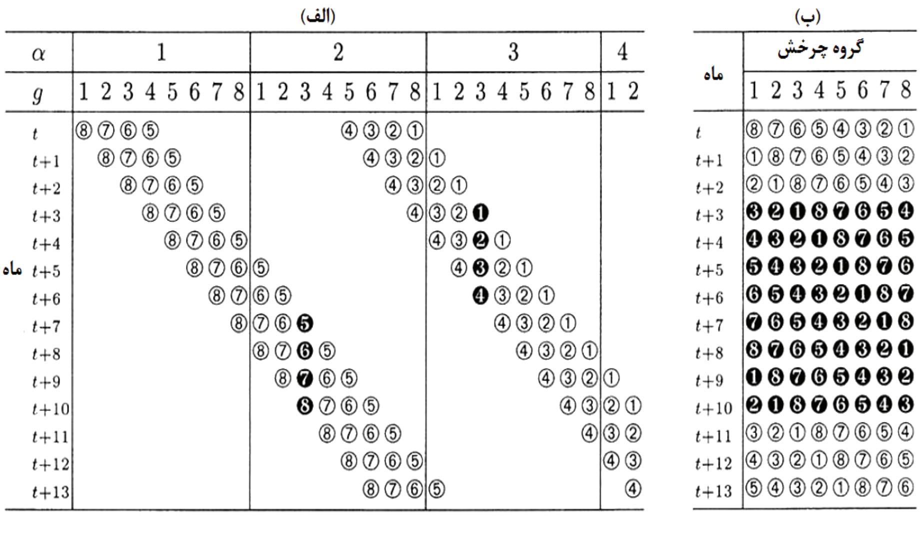
۳- نمونه‌گیری چرخشی یک‌سطحی متعادل دوسویه

برای داشتن براورد سطح کارا، نیاز به نمونه‌هایی است که در دوره‌های مختلف ناهمبسته باشند و برای داشتن براوردهای تغییرات کارا نیاز به نمونه‌هایی است که در دوره‌های مختلف همبستگی بالایی داشته باشند. به این دلیل است که برای داشتن براوردهای مطلوبی از براوردهای مقطع و تغییرات از نمونه‌های چرخشی استفاده می‌شود. هر گروه چرخش یک زیر واحد را در نمونه‌ی مقطعی شرکت می‌دهد. بنا بر این، هر نمونه‌ی مقطعی به تعداد گروه‌های چرخش زیرواحد نمونه‌ای دارد. این گروه‌بندی به شکلی خواهد بود که زیرواحدهای درون هر گروه چرخش همبسته و زیرواحدهای گروه‌های چرخش مختلف ناهمبسته باشند. به وسیله‌ی الگوی چرخش، تعدادی از زیرواحدها در نمونه‌ی ماه بعد می‌مانند و بقیه با زیرواحدهای هم‌گروه جایگزین می‌شوند. در چنین طرح نمونه‌گیری دو نوع همبستگی ممکن است رخ دهد. اول، همبستگی مرتبه‌ی اول است و هنگامی رخ می‌دهد که یک زیرواحد نمونه‌ای در دو دوره‌ی پیاپی در نمونه حاضر باشد. دوم، همبستگی مرتبه‌ی دوم، مربوط به دو زیرواحد مختلف از یک گروه چرخش یکسان است. بهترین راه برای به کارگیری این دو نوع همبستگی متعادل‌سازی دوسویه‌ی الگوی چرخش است. مهم‌ترین براوردی که در نمونه‌گیری چرخشی به آن پرداخته می‌شود براورد تغییرات است. طرح متعادل دوسویه براورد بهتری را برای تغییرات مقطعی فراهم می‌کند زیرا متعادل‌سازی دوسویه همبستگی مثبت اندازه‌های نمونه‌ای بین مقطع‌ها را افزایش داده و در نتیجه واریانس کوچک‌تری ارایه می‌دهد. برای کاهش بار پاسخ‌گویی بر دوش پاسخ‌گویان، می‌توان زیرواحدهای نمونه‌ای را با دیگر زیرواحدها جایگزین کرد. ساده‌ترین روش برای چنین جایگزینی استفاده از یک الگوی چرخش است که فقط بستگی به تعداد نوبت‌هایی داشته باشد که از یک زیرواحد نمونه‌ای آمارگیری به عمل می‌آید. از آن‌جایی که طرح متعادل دوسویه از این نوع الگوی چرخش پیروی می‌کند، اریبی ناشی از انتخاب بیش از معمول گروه‌های چرخش و دوره‌های مشخص را از بین می‌برد. در اینجا متعادل‌سازی دوسویه را با استفاده از طرح ۴-۸-۴ طرح جمعیت جاری آمریکا شرح

خواهیم داد. در این طرح، تمام زیرواحدهای نمونه‌ای در هشت گروه چرخش، گروه‌بندی می‌شوند. هر زیرواحد که نماینده‌ی گروه چرخش خود در نمونه است، به مدت چهار ماه در نمونه می‌ماند و به مدت هشت ماه پیاپی از نمونه خارج می‌شود، سپس به مدت چهار ماه دیگر به نمونه باز می‌گردد.

در شکل ۱ (الف)، (α, g) نشان‌دهنده‌ی زیرواحد αام از گروه چرخش gام است. دایره‌های دور اعداد نشان می‌دهند کدام یک از زیرواحدهای نمونه‌ای از هر گروه چرخش آمارگیری شده‌اند و اعداد درون دایره‌ها نشان‌گر تعداد نوبت‌هایی است که از این زیرواحد نمونه‌ای آمارگیری به‌عمل آمده است.

شکل ۱ (ب)، متعادل‌سازی طولی را برای تعداد نوبت‌های آمارگیری نمایش می‌دهد. هر گروه شامل هشت نوبت آمارگیری در طول هشت دوره‌ی آمارگیری است. این متعادل‌سازی طولی هنگامی امکان‌پذیر است که تمامی هشت گروه چرخش در هر دوره‌ی آمارگیری حاضر باشند. شمول تمام هشت گروه چرخش در هر دوره‌ی آمارگیری این امکان را می‌دهد که جایگزینی زیرواحدها داخل گروه‌های چرخش یکسان به‌راحتی انجام شود. به‌عنوان مثال در این شکل، زیرواحدهای نمونه‌ای (۱,۱) و (۲,۵) در ماه آمارگیری tام به‌وسیله‌ی زیرواحدهای (۳,۱) و (۱,۵) در ماه آمارگیری $t+1$ام جایگزین می‌شوند.



شکل ۱- الگوی چرخش برای طرح ۴-۸-۴

هر دوره‌ی آمارگیری، شامل همه‌ی هشت نوبت آمارگیری است؛ یک زیرواحد برای نوبت اول، یکی برای نوبت دوم، ... و بالاخره آخری برای نوبت هشتم آمارگیری می‌شوند. بنا بر این، این نظام از نظر تعداد نوبت‌های آمارگیری در نمونه و از نظر گروه چرخش متعادل است. با استفاده از این ویژگی متعادل‌سازی دوسویه، هشت زیرواحد نمونه‌ای جدید، یکی از هر گروه چرخش، در هر ماه در نمونه شرکت داده می‌شوند [۵].

۱-۳- طرح متعادل دوسویه

طرح چرخشی $r_1 - r_2 - r_1 - \cdots - r_2 - r_1$ را در نظر بگیرید. در این طرح یک زیرواحد از هر گروه چرخش برای r_1 دوره‌ی اول آمارگیری در نمونه حاضر است و برای r_2 دوره‌ی پیاپی از نمونه کنارگذاشته می‌شود. در این الگوی تناوبی r_1 به مدت m مرتبه و r_2 به مدت$m-1$ مرتبه تکرار می‌شود. به‌طور کلی این الگوی چرخش را به صورت $r_1^m - r_2^{m-1}$ نیز نمایش می‌دهند. به‌عنوان مثال، طرح ۴-۸-۴ را به‌صورت $4^2 - 8^1$ نمایش می‌دهند [۵].

از آن‌جایی که تمام mr_1 گروه چرخش در هر دوره‌ی آمارگیری حضور دارند، جایگزینی زیرواحدها داخل هر گروه انجام می‌شود. به علاوه، رابطه‌ی زیر در طرح $r_1^m - r_2^{m-1}$ بر قرار است:

$$r_2 = lr_1, \qquad l \text{ به ازای مقدارهای صحیح نامنفی}$$

فرض کنید $g_t(i)$ گروه چرخشی باشد که در ماه tام برای نوبت iام آمارگیری شده باشد. طبق شکل ۱ برای گروه‌های چرخش، رابطه‌ی زیر بین ماه t و $t+1$ برقرار است:

$$g_t(i) = g_{t+1}(i+1), \qquad i = 1, 2, \ldots, 7$$

و $g_t(8) = g_{t+1}(1)$.

در زیر یک قاعده‌ی کلی برای چرخش ارائه می‌دهیم:

به ازای مقدارهای معلوم r_1، r_2، m و مقدار مناسبی از عدد صحیح، مثبت و یکتای m^* که، $1 \le m^* \le m$، طرح $r_1^m - r_2^{m-1}$ از الگوی چرخش زیر پیروی می‌کند:

$$(۱)\quad g_t(i)=\begin{cases} g_{t+۱}\left(\operatorname{mod}_m\ \left[m-m^*+\dfrac{i}{r_۱}\right]r_۱+۱\right) & i=r_۱,۲r_۱,...,mr_۱ \\ g_{t+۱}(i+۱) & i=۱,۲,...,mr_۱\ \ (\text{به جز}\ \ i=r_۱,۲r_۱,...,mr_۱) \end{cases}$$

که در آن $\operatorname{mod}_m\ [\cdot]$ باقی‌مانده‌ی تقسیم $[\cdot]$ بر m است و همه‌ی $g_t(۱)$، $g_t(۲)$، ... و $g_t(mr_۱)$ ها وجود دارند.

توجه کنید m زیرواحد از m گروه چرخش $g_t(۱)$، $g_t(۲)$، ...، $g_t(m^*r_۱)$، ... و $g_t(mr_۱)$، به‌وسیله‌ی m زیرواحد دیگر که به‌ترتیب برای نوبت $(m-m^*+۱)\ r_۱+۱$، $(m-m^*+۲)\ r_۱+۱$، ...، ۱، ... و $(۲m-m^*)\ r_۱+۱$ در ماه $t+۱$ام آمارگیری می‌شوند، جایگزین خواهند شد. تعداد $m(r_۱-۱)$ گروه باقیمانده از همان زیرواحدهای نمونه‌ای حاضر در دوره‌ی t برای دوره‌ی $t+۱$ استفاده می‌کنند. یعنی الگوی چرخش (۱) فقط به تعداد نوبت‌هایی که $mr_۱$ زیرواحد در دوره‌ی معین آمارگیری شده‌اند، بستگی دارد.

تعریف ۱. تعادل از نظر تعداد نوبت‌های حضور در نمونه و در گروه چرخش طرح $r_۱^m-r_۲^{m-۱}$ متعادل از نظر تعداد نوبت‌های حضور در نمونه و از نظر تعداد نوبت‌های حضور در گروه چرخش است هرگاه در شرط‌های زیر صدق کند:

۱- در هر دوره‌ی آمارگیری، از هر گروه چرخش یک زیرواحد نمونه‌ای در نمونه باشد؛ یک زیرواحد از یک گروه برای نوبت اول، یکی از دومین گروه برای نوبت دوم و به همین ترتیب یکی از $mr_۱$ امین گروه برای نوبت $mr_۱$ ام آمارگیری شود.

۲- برای هر یک از $mr_۱$ دوره‌ی آمارگیری، هر یک از $mr_۱$ گروه چرخش $mr_۱$ نوبت در نمونه حاضر باشد؛ برای نوبت اول در یک دوره، نوبت دوم در دوره‌ای دیگر و به همین ترتیب [۵].

لم ۱. فرض کنید طرح $r_۱^m-r_۲^{m-۱}$ از الگوی چرخش (۱) تبعیت می‌کند یا متعادل دوسویه باشد. در این صورت عدد صحیح مثبت و یکتای m^* وجود دارد که در رابطه‌ی زیر صدق کند:

$$\operatorname{mod}_m\ [m^*(l+۱)-l\ ]=۰$$

که در آن $۱\le m^*\le m$ و $l=r_۲/r_۱$، [۵].

وقتی $m=۱$ است بدون توجه به مقدار l، $\bmod_m\ [m^*(l+۱)-l\ ]=۰$. این بدین معنی است که تمام طرح‌های $r_۱^۱-۰$ به‌طور خودکار در $\bmod_m\ [m^*(l+۱)-l\ ]=۰$ با $m^*=۱$ صدق می‌کنند.

با توجه به لم ۱، برخی از طرح‌های $r_۱^m-r_۲^{m-۱}$، از الگوی چرخش (۱) برخوردار نیستند و ویژگی طرح‌های متعادل دوسویه را ندارند. به‌عنوان مثال، طرح‌آمارگیری نیروی کار مرکز آمار ایران با الگوی چرخش $۲^۱-۲^۲$ و طرح $۹^۱-۳^۲$ متعلق به این کلاس از طرح‌های متعادل دوسویه نیستند. زیرا در این طرح‌ها عدد مثبت، صحیح و یکتای m^* وجود ندارد که در رابطه‌ی $\bmod_m\ [m^*(l+۱)-l\ ]=۰$ صدق کنند؛ $۱\le m^*\le m$.

۲-۳- براوردگر مرکب تعمیم‌یافته (GCE)[1]

فرض کنید $\mathbf{x}_t=(x_{t,۱},x_{t,۲},\ldots,x_{t,mr_۱})'$ برداری از متغیرها برای نمونه‌ی ماه tام باشد که در آن $x_{t,i}$ اندازه‌ی نمونه‌ای مشخصه‌ی یک زیرواحد که برای نوبت iام در ماه tام آمارگیری شده است. براوردگر مرکب تعمیم‌یافته ([۱]و [۲]) به‌صورت زیر است:

$$\begin{aligned} y_t &= \sum_{i=۱}^{mr_۱} a_i x_{t,i}-\omega\sum_{i=۱}^{mr_۱} b_i x_{t-۱,i}+\omega y_{t-۱} \\ &= \mathbf{a}'\,\mathbf{x}_t-\omega\,\mathbf{b}'\,\mathbf{x}_{t-۱}+\omega y_{t-۱} \end{aligned} \qquad (۲)$$

که در آن: $y_{t-۱}$، براورد مرکب تعمیم‌یافته‌ی پارامتر مورد نظر در دوره‌ی آمارگیری $t-۱$ام؛ $\mathbf{a}'$ و $\mathbf{b}'$ بردارهای ضریب‌ها هستند: $\mathbf{a}'=(a_۱,a_۲,\ldots,a_{mr_۱})$ و $\mathbf{b}'=(b_۱,b_۲,\ldots,b_{mr_۱})$.
در رابطه‌ی بالا، ω عددی است در بازه‌ی $۰\le\omega<۱$ و بردارهای $\mathbf{a}'$ و $\mathbf{b}'$ محدود به قیدهای $\mathbf{a}'\mathbf{1}=\mathbf{b}'\mathbf{1}=۱$ هستند که در آن $\mathbf{1}$ برداری از یک‌ها است. اگر در رابطه‌ی (۲)، $\omega=۰$ قرار داده شود براوردگر مرکب تعمیم‌یافته به براوردگر ساده تبدیل می‌شود.

۳-۳- واریانس براوردگر مرکب تعمیم‌یافته

برای به دست آوردن واریانس براوردگر مرکب تعمیم‌یافته نیاز به تعریف ساختار کوواریانسی زیر داریم. کوواریانس $x_{t,i}$ و $x_{t',i'}$ برای ماه $t,t'=۰,۱,\ldots$ و زمان

آمارگیری $i,i'=1,2,\ldots,mr_1$ را به‌صورت زیر تعریف می‌کنیم:

$$\operatorname{cov}(x_{t,i},x_{t',i'})=\begin{cases}\sigma_i^2 & i=i', t=t'\\ \rho_{1|t-t'|}\sigma_i\sigma_{i'} & t\neq t',\ \alpha(i)=\alpha(i')\ \text{و}\ g(i)=g(i')\\ \rho_{2|t-t'|}\sigma_i\sigma_{i'} & t\neq t',\ \alpha(i)\neq\alpha(i')\ \text{و}\ g(i)=g(i')\\ 0 & \text{در غیر این صورت}\end{cases} \qquad (3)$$

که در آن σ_i^2 ، واریانس $x_{t,i}$ها یعنی واحدهای نمونه‌ای است که در مقطع t برای نوبت iام آمارگیری می‌شوند. نماد $\rho_{1|t-t'|}$ ضریب همبستگی مرتبه‌ی اول مربوط به اندازه‌ی یک واحد نمونه‌ای در دو مقطع t و اندازه‌ی همان واحد نمونه‌ای در مقطع t' است. نماد $\rho_{2|t-t'|}$ ضریب همبستگی مرتبه‌ی دوم مربوط به دو واحد نمونه‌ای متفاوت از گروه چرخش یکسان است که در مقطع‌های t و t' آمارگیری شده یا خواهند شد.

ما دو نوع براوردگر مرکب سطح مقطعی y_t و تفاضل دو سطح مقطعی $y_t-y_{t-t_0}$ برای $t_0\geq 1$ در نظر می‌گیریم. برای به‌دست آوردن واریانس این براوردگرها نیاز به یافتن $\operatorname{cov}(\mathbf{x}_t,\mathbf{x}_{t+t'})$ تحت ساختار کوواریانسی (۳) داریم.

برای ساده‌نویسی فرمول واریانس براوردگرها، ماتریس‌های زیر را تعریف می‌کنیم:
اولین ماتریس، ماتریس $\mathbf{L}_s$ ، $(mr_1+(m-1)r_2)\times(mr_1+(m-1)r_2)$ است به‌طوری که:

$$(\mathbf{L}_s)_{i,j}=\begin{cases}1 & j=i+1\\ 0 & \text{در غیر این صورت}\end{cases}$$

ماتریس $\mathbf{L}_s$ یک ماتریس انتقال پیش‌رو است با ویژگی $\mathbf{L}_s^{t'}=\mathbf{L}_s^{t'-1}\cdot\mathbf{L}_s$ که $t'\geq 1$ و $\mathbf{L}_s^{0}=\mathbf{I}$.

ماتریس دوم، ماتریس $\mathbf{L}$ ، $mr_1\times mr_1$ است به‌طوری که:

$$(L)_{i,j} = \begin{cases} ۱ & j = \mathrm{mod}_m\left(m - m^* + \frac{i}{r_۱}\right)r_۱ + ۱, \quad i = r_۱, ۲r_۱, \ldots, mr \\ & \text{یا اگر } j = i + ۱, \quad i \neq r_۱, ۲r_۱, \ldots, mr_۱ \\ ۰ & \text{در غیر این صورت} \end{cases}$$

ماتریس جایگشت $\mathbf{L}$ بیان دیگری از الگوی چرخش (۱) با ویژگی $\mathbf{L}^{t'} = \mathbf{L}^{t'-۱} \cdot \mathbf{L}$ است که در آن: $t' \geq ۱$ و $\mathbf{L}^{۰} = \mathbf{I}$. ماتریس سوم ماتریس $\mathbf{L}_۱^{t'}$، $mr_۱ \times mr_۱$ است و از حذف سطر و ستون iام ماتریس $\mathbf{L}_s^{t'}$ به دست می‌آید که در آن: $i = kr_۱ + (k-۱)r_۲ + ۱, kr_۱ + (k-۱)r_۲ + ۲, \ldots, k(r_۱ + r_۲)$ و $k = ۱, ۲, \ldots, m-۱$.
با استفاده از دو ماتریس $\mathbf{L}_۱^{t'}$ و $\mathbf{L}_۲^{t'}$ مقدار کوواریانس به‌صورت زیر به دست می‌آید:

لم ۲. فرض کنید طرح $r_۱^m - r_۲^{m-۱}$ متعادل دوسویه باشد. تحت ساختار کوواریانسی (۳) داریم:

$$\mathrm{cov}(\mathbf{x}_t, \mathbf{x}_{t+t'}) = \rho_{۱t'}\mathbf{V}\mathbf{L}_۱^{t'}\mathbf{V} + \rho_{۲t'}\mathbf{V}\mathbf{L}_۲^{t'}\mathbf{V}$$

که در آن $t_۰ \geq ۰$، $\mathbf{V} = \mathrm{diag}(\sigma_۱, \sigma_۲, \ldots, \sigma_{mr_۱})$ و $\mathbf{L}_۲^{t'} = \mathbf{L}^{t'} - \mathbf{L}_۱^{t'}$ که $\mathbf{L}_۱^{۰} = \mathbf{I}$ و $\mathbf{L}_۲^{۰} = ۰$، [۵].

دو ماتریس $\mathbf{L}_۱^{t'}$ و $\mathbf{L}_۲^{t'}$ به‌طور کامل وضعیت دو اندازه‌گیری $\mathbf{x}_t$ و $\mathbf{x}_{t+t'}$ را تعیین می‌کنند. به‌طوری که $x_{t,i}$ و $x_{t+t',j}$ از یک زیرواحد یکسان هستند اگر $(\mathbf{L}_۱^{t'})_{i,j} = ۱$ و $(\mathbf{L}_۲^{t'})_{i,j} = ۰$ و از دو زیرواحد متفاوت اما گروه چرخش یکسانند اگر $(\mathbf{L}_۱^{t'})_{i,j} = ۰$ و $(\mathbf{L}_۲^{t'})_{i,j} = ۱$.
چون می‌توان براوردگر مرکب تعمیم‌یافته را به‌صورت زیر نیز نوشت:

$$y_t = \mathbf{a}'\mathbf{x}_t - \omega\mathbf{b}'\mathbf{x}_{t-۱} + \omega y_{t-۱} = \mathbf{a}'\mathbf{x}_t + (\mathbf{a}' - \mathbf{b}')\sum_{i=۱}^{\infty}\omega^i\mathbf{x}_{t-i}$$

پس واریانس آن به شکل سری بینهایت در می‌آید. برای رفع این مشکل، ماتریس زیر را تعریف می‌کنیم

$$\mathbf{Q}_{t_\circ,k} = \sum_{j=\circ}^{T-t_\circ+k} \omega^{j+k} \rho_{\text{۱},t_\circ+j-k} \mathbf{L}_\text{۱}^{t_\circ+j-k} + \sum_{j=\circ}^{\infty} \omega^{j+k} \rho_{\text{۲},t_\circ+j-k} \mathbf{L}_\text{۲}^{t_\circ+j-k}$$

که در آن $t_\circ \geq k$، $T \equiv mr_\text{۱} + (m-\text{۱})r_\text{۲} - \text{۱}$ و قسمت اول سمت راست تساوی برابر صفر است هرگاه $T - t_\circ + k < \circ$.

با استفاده از لم ۲ و نمادگذاری‌های بالا، قضیه‌ی زیر را داریم.

قضیه‌ی ۱. *تحت فرض مشابه لم ۲ واریانس‌های y_t و $y_t - y_{t-t_\circ}$ به‌صورت زیر هستند* [۵]:

$$(\text{۱}-\omega^\text{۲})\mathrm{var}(y_t) = \mathbf{a}'\mathbf{V}(\mathbf{I}+\text{۲}\omega\mathbf{Q}_{\text{۱},\circ})\mathbf{V}\mathbf{a} + \omega^\text{۲}\mathbf{b}'\mathbf{V}(\mathbf{I}+\text{۲}\mathbf{Q}_{\text{۱},\circ})\mathbf{V}\mathbf{b} - \text{۲}\omega\mathbf{b}'\mathbf{V}(\mathbf{Q}_{\text{۱},\circ}+\mathbf{Q}'_{\text{۱},\text{۱}})\mathbf{V}\mathbf{a}$$

و

$$\frac{\text{۱}}{\text{۲}}\mathrm{var}(y_t - y_{t-t_\circ}) = (\text{۱}-\omega^{t_\circ})\mathrm{var}(y_t) + \sum_{k=\circ}^{t_\circ-\text{۱}}(\mathbf{a}'\mathbf{V}\mathbf{Q}_{t_\circ,k}\mathbf{V}\mathbf{a} + \omega^\text{۲}\mathbf{b}'\mathbf{V}\mathbf{Q}_{t_\circ,k}\mathbf{V}\mathbf{b} - \mathbf{b}'\mathbf{V}(\mathbf{Q}'_{t_\circ,k+l} + \omega^\text{۲}\,\mathbf{Q}_{t_\circ,k-\text{۱}})\mathbf{V}\mathbf{a}).$$

برای $\rho_{\text{۲}t} = \circ$، کانتوِل به فرمول واریانس برای طرح‌های چرخشی یک‌سطحی دست یافت [۲]. او فقط متعادل‌سازی مقطعی را با طرح‌های چرخشی کلی‌تری مورد بررسی قرار داد. قضیه‌ی ۲ حالت خاصی از رابطه‌ی واریانسی است که در سال ۱۹۹۰ توسط کانتوِل ارایه شد.

۴-۳- ضریب‌های بهین برای براوردگر مرکب تعمیم‌یافته

اکنون دو مجموعه از ضریب‌های بهین را ارایه می‌دهیم. یک مجموعه به‌وسیله‌ی مینیمم کردن واریانس براوردگر مرکب تعمیم‌یافته و دیگری به‌وسیله‌ی مینیمم کردن میانگین توان دوم خطا.

تا اینجا دو نوع براوردگر مرکب تعمیم‌یافته معرفی کردیم: y_t و $y_t - y_{t-t_{\circ}}$. با در نظر گرفتن مقدارهای خاصی از $t_{\circ}$ و β، فرض می‌کنیم n تا از این براوردگرهای مرکب داشته باشیم که آن‌ها را با z_{ht}، $h = ۱,۲,...,n$، نشان می‌دهیم. توجه کنید که z_{ht} و $z_{h't}$، $h \neq h'$، ممکن است دو نوع براوردگر تعمیم‌یافته‌ی یکسان از دو مشخصه‌ی مختلف باشند. به‌عنوان مثال، نماد y_t برای افراد شاغل با نماد y_t برای افراد بیکار یکسان است. با استفاده از مینیمم کردن دو تابع هدف زیر ضریب‌های بهین را برای براوردگر مرکب به دست می‌آوریم:

$$O_۱ = \sum_{h=۱}^{n} \delta_h \operatorname{var}(z_{ht}) - \lambda_{۱,۱}(\mathbf{۱}'\mathbf{a}_۱ - ۱) - \lambda_{۲,۱}(\mathbf{۱}'\mathbf{b}_۱ - ۱)$$
$$O_۲ = \sum_{h=۱}^{n} \delta_h \operatorname{MSE}(z_{ht}) - \lambda_{۱,۲}(\mathbf{۱}'\mathbf{a}_۲ - ۱) - \lambda_{۲,۲}(\mathbf{۱}'\mathbf{b}_۲ - ۱)$$

که در آن λ ها ضریب‌های لاگرانژ و δ ها اهمیت نسبی هر یک از براوردگرها هستند. برای مثال اگر هر کدام از این n براوردگر مرکب تعمیم‌یافته اهمیت یکسانی داشته باشند $\delta_h = ۱/n$، به ازای همه‌ی δ_h ها.

با استفاده از قضیه‌ی ۱ و تعریف مناسبی از ماتریس‌های $\mathbf{P}_{۱,h}$ ، $\mathbf{P}_{۲,h}$ و $\mathbf{P}_{۳,h}$ می‌توان واریانس را به‌صورت زیر نوشت ([۵]):

$$\operatorname{var}(z_{ht}) = \mathbf{a}_۱'\mathbf{P}_{۱,h}\mathbf{a}_۱ + \mathbf{b}_۱'\mathbf{P}_{۲,h}\mathbf{a}_۱ + \mathbf{b}_۱'\mathbf{P}_{۳,h}\mathbf{b}_۱$$

که در آن $\mathbf{P}_{۳,h} = \omega^۲\mathbf{P}_{۱,h}$. به‌عنوان مثال وقتی $y_t = z_{ht}$، ماتریس‌های بالا به‌صورت زیر به دست می‌آیند:

$$\mathbf{P}_{۱,h} = \frac{۱}{(۱-\omega^۲)}\mathbf{V}(\mathbf{I} + ۲\omega\mathbf{Q}_{۱,\circ})\mathbf{V}$$
$$\mathbf{P}_{۲,h} = \frac{-۲\omega}{(۱-\omega^۲)}\mathbf{V}(\mathbf{Q}_{۱,\circ} + \mathbf{Q}_{۱,۱}')\mathbf{V}$$

که در آن $\mathbf{Q}_{۱,۱}'$ ترانهاده‌ی ماتریس $\mathbf{Q}_{۱,۱}$ است. به‌طور مشابه، با تعریف مناسبی از ماتریس‌های $\mathbf{C}_{۱,h}$ ، $\mathbf{C}_{۲,h}$ و $\mathbf{C}_{۳,h}$، توان دوم اریبی را می‌توان به‌صورت زیر نوشت:

$$\text{bias}^{۲}(z_{ht}) \ = \ \mathbf{a}'_{۲}\mathbf{C}_{۱,h}\mathbf{a}_{۲} + \mathbf{b}'_{۲}\mathbf{C}_{۲,h}\mathbf{a}_{۲} + \mathbf{b}_{۲}\mathbf{C}_{۳,h}\mathbf{b}_{۲}$$

اکنون ماتریس‌های زیر را تعریف می‌کنیم:

$$\mathbf{G}_{۱}^{-۱} = \sum_{h=۱}^{n} \delta_h (\mathbf{P}_{۱,h} + \mathbf{P}'_{۱,h})$$

$$\mathbf{G}_{۲۱}^{-۱} = \sum_{h=۱}^{n} \delta_h (\mathbf{P}_{۱,h} + \mathbf{P}'_{۱,h} + \mathbf{C}_{۱,h} + \mathbf{C}'_{۱,h})$$

$$\mathbf{G}_{۲۲}^{-۱} = \sum_{h=۱}^{n} \delta_h (\mathbf{P}_{۳,h} + \mathbf{P}'_{۳,h} + \mathbf{C}_{۳,h} + \mathbf{C}'_{۳,h}).$$

فرض کنید $\mathbf{A}_{۱} = \mathbf{G}_{۱}((\mathbf{۱}'\mathbf{G}_{۱}\mathbf{۱})^{-۱}\mathbf{J}\mathbf{G}_{۱} - \mathbf{I})$ ، $\mathbf{A}_{۲j} = \mathbf{G}_{۲j}((\mathbf{۱}'\mathbf{G}_{۲j}\mathbf{۱})^{-۱}\mathbf{J}\mathbf{G}_{۲j} - \mathbf{I})$ که در آن $\mathbf{I}$ ماتریس همانی است و $j = ۱,۲$. همچنین فرض کنید $\mathbf{B}_{۱} = (\mathbf{۱}'\mathbf{G}_{۱}\mathbf{۱})^{-۱}\mathbf{G}_{۱}\mathbf{۱}$ و $\mathbf{B}_{۲j} = (\mathbf{۱}'\mathbf{G}_{۲j}\mathbf{۱})^{-۱}\mathbf{G}_{۲j}\mathbf{۱}$ ، $j = ۱,۲$ ([۵]).

اکنون با استفاده از بهینه‌سازی دو تابع هدف $O_{۱}$ و $O_{۲}$ به ضریب‌های بهین دست خواهیم یافت. این ضریب‌ها را ضریب‌های بینابینی می‌نامند. لم زیر نحوه‌ی براورد این ضریب‌ها را نشان می‌دهد.

لم ۳. برای مقدارهای معلوم وزن‌های δ_h ، $h = ۱,۲,...,n$ ، ضریب‌های بینابینی $\mathbf{a}_{۱}$ و $\mathbf{b}_{۱}$ با استفاده از مینیمم‌سازی واریانس موزون، یعنی $\sum_{h=۱}^{n}\delta_k \operatorname{var}(z_{ht})$ ، و ضریب‌های بینابینی $\mathbf{a}_{۲}$ و $\mathbf{b}_{۲}$ با استفاده از مینیمم‌سازی میانگین توان دوم خطای موزون، یعنی $\sum_{h=۱}^{n}\delta_k \operatorname{MSE}(z_{ht})$ ، به‌صورت زیر هستند [۵]:

$$\hat{\mathbf{a}}_{۱} = [\mathbf{I} - \frac{۱}{\omega^{۲}}\mathbf{A}_{۱}F'_{۱}\mathbf{A}_{۱}F_{۱}]^{-۱}[\mathbf{I} + \mathbf{A}_{۱}F'_{۱}]\mathbf{B}_{۱}$$

$$\hat{\mathbf{b}}_{۱} = \frac{۱}{\omega^{۲}}\mathbf{A}_{۱}F'_{۱}\hat{\mathbf{a}}_{۱} + \mathbf{B}_{۱} \qquad (۴)$$

و

$$\hat{\mathbf{a}}_{۲} = [\mathbf{I} - \mathbf{A}_{۲۱}F'_{۲}\mathbf{A}_{۲۲}F_{۲}]^{-۱}\ [\mathbf{B}_{۲۱} + \mathbf{A}_{۲۱}F'_{۲}\mathbf{B}_{۲۲}]$$

$$\hat{\mathbf{b}}_2 = \mathbf{A}_{22} F'_{22} \hat{\mathbf{a}}_2 + \mathbf{B}_{22}$$

که در آن $F_1 = \sum_{h=1}^{n} \delta_h \mathbf{P}_{2,h}$ و $F_2 = \sum_{h=1}^{n} \delta_h (\mathbf{P}_{2,h} + \mathbf{C}_{2,h})$ ، و F'_1 و F'_2 ترانهاده‌ی ماتریس‌های F_1 و F_2 هستند.

۴- یک مطالعه‌ی موردی

دفتر آمار کار ایالت‌های متحده (BLS)[2]، از داده‌های مربوط به هزینه‌های خانواری برای ساختن وزن‌ها برای شاخص قیمت مصرف‌کننده استفاده می‌کند. طرح آمارگیری هزینه‌ی مصرف‌کننده جریانی از داده‌های مربوط به عادت‌های خرید مصرف‌کننده‌های آمریکایی را ارایه می‌دهد. این داده‌ها به‌طور گسترده‌ای در پژوهش‌ها و تحلیل‌های اقتصادی به کار برده می‌شوند. در این بخش از داده‌های طرح آمارگیری هزینه‌ی مصرف‌کننده‌ی ایالت‌های متحده (۱۹۸۵- ۱۹۸۹) برای براورد واریانس استفاده می‌کنیم. این آمارگیری به‌صورت چرخشی انجام می‌شود و از یک الگوی چرخش برای ورود و خروج واحدهای نمونه‌ای استفاده می‌کند. الگوی چرخش این آمارگیری ۰-۵ است که بر اساس آن هر واحد نمونه‌ای پنج فصل متوالی در نمونه حضور دارد و سپس به‌طور کامل از نمونه خارج می‌شود. بیست درصد از نمونه‌ی هر فصل شامل خانوارهای جدیدی هستند که برای اولین بار به نمونه وارد می‌شوند. این واحدهای نمونه‌ای جایگزین واحدهایی می‌شوند که پنجمین آمارگیریِ آن‌ها تکمیل شده و باید از نمونه خارج شوند.

نخستین آمارگیری از هر واحد مصرف‌کننده اطلاعات جمعیتی و مشخصه‌های خانوادگی را روی یک کارت کنترلی گرداوری می‌کند. این اطلاعات شامل سن، جنس، نژاد، وضعیت تأهل و تحصیلات است. این اطلاعات در هر مقطع آمارگیری به‌روز می‌شوند. در این آمارگیری اطلاعات مربوط به هزینه و درامد گرداوری نمی‌شوند، از این رو در براورد میانگین هزینه‌ی مصرف‌کننده و میانگین درامد قبل و بعد از پرداخت مالیات و همچنین براورد واریانس این براوردها از اطلاعات آمارگیری اول استفاده نمی‌شود. آمارگیری‌های دوم تا پنجم دارای یک پرسشنامه‌ی یک‌دست و یکسان برای گرداوری اطلاعات مربوط به هزینه‌ی سه ماه گذشته است. اطلاعات درامدی دوازده ماه گذشته‌ی مربوط به هر واحد از قبیل دست‌مزد، حقوق، درامد تجاری، درامد حاصل از فعالیت‌های کشاورزی، درامد حاصل از تأمین اجتماعی، غرامت بیکاری، بیمه‌ی کارگری،

حقوق بازنشستگی، سود و سود سهام، درامدهای مربوط به کمک‌های دولت به نیازمندان و معلولین، درامدهای حاصل از اتاق‌های اجاره‌ای و دیگر درامدها، فقط در نوبت‌های دوم و پنجم آمارگیری انجام می‌شود. این اطلاعات از آمارگیری نوبت دوم به آمارگیری‌نوبت‌های سوم و چهارم منتقل می‌شوند. در نوبت پنجم آمارگیری این اطلاعات به‌روز می‌شوند. در این طرح، به‌دلیل این که با چهار مقطع آمارگیری سر و کار داریم، الگوی چرخش ۰-۵ همانند الگوی ۰-۴ عمل می‌کند. در هر مقطع آمارگیری، چهار گروه چرخش مورد مطالعه قرار می‌گیرند که یکی از این گروه‌های چرخش برای نوبت دوم، یکی برای نوبت سوم، دیگری برای نوبت چهارم و بالاخره آخری برای نوبت پنجم آمارگیری می‌شوند. پس این طرح یک طرح نمونه‌گیری چرخشی متعادل دوسویه است.

۱-۴- براورد مرکب تعمیم‌یافته برای میانگین هزینه‌ی مصرف‌کننده

همان‌طور که گفتیم در این مطالعه چون از اطلاعات نوبت‌های آمارگیری دوم تا پنجم استفاده می‌کنیم، الگوی چرخش آن شبیه الگوی چرخش ۰-۴ عمل می‌کند. ابتدا ماتریس $\mathbf{L}$ را به‌صورت زیر تعریف می‌کنیم: درایه‌ی (i,j) آن برابر یک است هرگاه $j=i+1$ و سایر درایه‌های آن برابر صفر است. ماتریس $\mathbf{L}$ بیان دیگری از الگوی چرخش (۱) است. در نتیجه ماتریس $\mathbf{L}$ به‌صورت زیر به دست می‌آید:

$$\mathbf{L}=\begin{bmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0\end{bmatrix}\qquad \mathbf{L}^2=\begin{bmatrix}0&0&1&0\\0&0&0&1\\0&0&0&0\\0&0&0&0\end{bmatrix}$$

$$\mathbf{L}^3=\begin{bmatrix}0&0&0&1\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}\qquad \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}=\mathbf{L}^4=\mathbf{L}^5=\cdots$$

جدول ۱- ضریب‌های بینابینی برای براوردگر مرکب تعمیم‌یافته

	$a_{۱}$	$a_{۲}$	$a_{۳}$	$a_{۴}$
$\rho =$ ۰/۵	۰/۲۲۴۵۸۹۴	۰/۲۸۵۶۳۰۱	۰/۲۵۵۹۲۷۱	۰/۲۳۳۸۵۳۴
$\rho =$ ۰/۶	۰/۲۰۸۸۶۶۵	۰/۲۹۸۷۸۴۱	۰/۲۷۱۳۲۳۴	۰/۲۲۱۰۲۶
	$b_{۱}$	$b_{۲}$	$b_{۳}$	$b_{۴}$
$\rho =$ ۰/۵	۰/۰۰۴۴۲۳۱	۰/۲۶۰۰۳۱۱	۰/۳۶۷۰۴۴۳	۰/۳۶۸۵۰۱۵
$\rho =$ ۰/۶	۰/۰۰۰۰۰۰۰	۰/۲۵۰۴۵۶۲	۰/۳۷۵۹۸۹۷	۰/۳۷۳۵۵۴۱

با توجه به بازگشتی بودن براوردگر مرکب تعمیم‌یافته، این براوردگر از اطلاعات دوره‌های قبل آمارگیری استفاده می‌کند. در نتیجه زمانی کاراست که اندازه‌گیری‌های مکرر روی یک واحد نمونه‌ای دارای همبستگی باشند. این براوردگر میانگین موزونی است با ضریب‌های نامعلوم ω، a و b. در عمل باید این ضریب‌ها را براورد کرد.

برای براورد کردن این ضریب‌ها با استفاده از رابطه‌ی (۴)، و با استفاده از داده‌ها ابتدا $\sigma_{\bar{X}_۲}$، $\sigma_{\bar{X}_۳}$، $\sigma_{\bar{X}_۴}$ و $\sigma_{\bar{X}_۵}$ را به‌صورت ۱۴۱، ۱۳۵، ۱۳۲ و ۱۳۲ دلار براورد کردیم که در آن $\bar{x}_i$ میانگین هزینه‌ی واحدهای مصرف‌کننده‌ای است که برای نوبت iام آمارگیری شده‌اند $(i = ۲,۳,۴,۵)$. همچنین با استفاده از داده‌ها، $\rho_{۱,۱}$، $\rho_{۱,۲}$ و $\rho_{۱,۳}$ به‌ترتیب به‌صورت ۰/۵۸، ۰/۵۲ و ۰/۶۲ براورد شدند. این آماره‌ها نشان‌دهنده‌ی تغییراتی اندک در طول دوره‌ی آمارگیری هستند. بنا بر این، ما مقدار ρ را در این کاربرد ۰/۵ و ۰/۶ فرض می‌کنیم. از طرفی با توجه به این که هدف مینیمم کردن واریانس براوردگر تعمیم‌یافته است، با اجراهای متعدد برنامه به‌وسیله‌ی SAS/IML، بهترین مقدار برای ω را ۰/۵ به ازای $\rho = ۰/۵$ و ۰/۶ به ازای $\rho = ۰/۶$ به دست آوردیم. با توجه به براوردهای به دست آمده، ضریب‌های بینابینی به‌صورت جدول ۱ به دست آمدند.

با در نظر گرفتن جدول ۳ (پیوست) فرض کنید $\bar{\mathbf{x}}_t = (\bar{x}_{t,۲}, \bar{x}_{t,۳}, \bar{x}_{t,۴}, \bar{x}_{t,۵})$ برداری از میانگین‌های نوبت‌های مختلف آمارگیری در مقطع t باشد که در آن $\bar{x}_{t,i}$ میانگین هزینه برای واحدهای مصرف‌کننده‌ای است که در مقطع t برای نوبت iام آمارگیری شده‌اند. در نتیجه با استفاده از این ضریب‌های بینابینی، براورد مرکب

تعمیم‌یافته‌ی میانگین هزینه‌ی مصرف‌کننده برای فصل اول سال ۱۹۸۹ با استفاده از $\rho=۰/۵$ و $\omega=۰/۵$ به‌صورت زیر به‌دست آمد:

$$\begin{aligned}\hat{y}_t &= \hat{\mathbf{a}}'\mathbf{x}_t - \omega\hat{\mathbf{b}}'\mathbf{x}_{t-۱} + \omega\hat{y}_{t-۱} \\ &= \hat{\mathbf{a}}'\mathbf{x}_t + (\hat{\mathbf{a}}' - \hat{\mathbf{b}}')\sum_{i=۱}^{\infty}\omega^i\,\mathbf{x}_{t-i} = ۶۵۵۰/۸۳\end{aligned}$$

و با استفاده از $\rho=۰/۶$ و $\omega=۰/۶$ به‌صورت زیر حاصل شد:

$$\begin{aligned}\hat{y}_t &= \hat{\mathbf{a}}'\mathbf{x}_t - \omega\hat{\mathbf{b}}'\mathbf{x}_{t-۱} + \omega\hat{y}_{t-۱} \\ &= \hat{\mathbf{a}}'\mathbf{x}_t + (\hat{\mathbf{a}}' - \hat{\mathbf{b}}')\sum_{i=۱}^{\infty}\omega^i\,\mathbf{x}_{t-i} = ۶۵۳۲/۵.\end{aligned}$$

براورد مرکب تعمیم‌یافته برای تغییرات فصلی میانگین هزینه‌ی مصرف‌کننده یعنی $\mu_t - \mu_{t-۱}$ به‌صورت زیر به دست آمد:

$$\widehat{\mu_t - \mu_{t-۱}} = \hat{y}_t - \hat{y}_{t-۱} = ۳۴۷/۴۴.$$

۲-۴- براورد واریانس براورد مرکب تعمیم‌یافته

براورد واریانس براوردگر مرکب تعمیم‌یافته‌ی میانگین هزینه‌ی مصرف‌کننده و براورد واریانس براوردگر مرکب تعمیم‌یافته برای براورد تغییرات فصلی هزینه‌ی مصرف‌کننده با استفاده از $\rho=۰/۵$ و $\omega=۰/۵$ به‌صورت زیر محاسبه شدند:

$$\begin{aligned}\widehat{\mathrm{var}}(y_t) = \{&\mathbf{a}'\,\mathbf{V}(\mathbf{I}+۲\omega\mathbf{Q}_{۱,۰})\mathbf{V}\mathbf{a} + \omega^۲\mathbf{b}'\,\mathbf{V}(\mathbf{I}+۲\mathbf{Q}_{۱,۰})\mathbf{V}\mathbf{b} \\ &-۲\omega\mathbf{b}'\,\mathbf{V}(\mathbf{Q}_{۱,۰}+\mathbf{Q}'_{۱,۱})\mathbf{V}\mathbf{a}\}\big/(۱-\omega^۲) = ۵۷۲۹/۱۱\end{aligned}$$

و

$$\begin{aligned}\widehat{\mathrm{var}}(y_t - y_{t-۱}) = ۲\{&(۱-\omega)\widehat{\mathrm{var}}(y_t) \\ &+\sum_{k=۰}^{t_۰-۱}(\mathbf{a}'\,\mathbf{V}\mathbf{Q}_{t_۰,k}\mathbf{V}\mathbf{a} + \omega^۲\mathbf{b}'\,\mathbf{V}\mathbf{Q}_{t_۰,k}\mathbf{V}\mathbf{b} \\ &-\mathbf{b}'\,\mathbf{V}\mathbf{Q}'_{۱,۱}\mathbf{V}\mathbf{a})\} = ۴۴۴۶/۴۱\end{aligned}$$

و با استفاده از $\rho=۰/۶$ و $\omega=۰/۶$ به‌صورت زیر به‌دست آمدند:

$$\widehat{\mathrm{var}}(y_t) = \{\mathbf{a}'\mathbf{V}(\mathbf{I}+۲\omega\mathbf{Q}_{۱,۰})\mathbf{Va}+\omega^۲\mathbf{b}'\mathbf{V}(\mathbf{I}+۲\mathbf{Q}_{۱,۰})\mathbf{Vb} \\ -۲\omega\mathbf{b}'\mathbf{V}(\mathbf{Q}_{۱,۰}+\mathbf{Q}'_{۱,۱})\mathbf{Va}\}\big/(۱-\omega^۲) = ۶۶۹۲/۳۴$$

و

$$\widehat{\mathrm{var}}(y_t - y_{t-۱}) = ۲\{(۱-\omega)\widehat{\mathrm{var}}(y_t) \\ +\sum_{k=۰}^{t_۰-۱}(\mathbf{a}'\mathbf{VQ}_{t_۰,k}\mathbf{Va}+\omega^۲\mathbf{b}'\mathbf{VQ}_{t_۰,k}\mathbf{Vb} \\ -\mathbf{b}'\mathbf{VQ}'_{۱,۱}\mathbf{Va})\} = ۳۳۷۱/۰۸$$

براورد ساده‌ی میانگین هزینه‌ی مصرف‌کننده برای فصل اول سال ۱۹۸۹ به‌صورت زیر محاسبه شد:

$$\bar{x}_t = \frac{۱}{n}\sum_{i=۱}^{n} x_i = ۶۵۲۳/۸۵$$

واریانس آن نیز از رابطه‌ی زیر براورد می‌شود:

$$s^۲ = \frac{۱}{n-۱}\sum_{i=۱}^{n}(x_i-\bar{x})^۲ = ۳۱۷۵۴۳۱۴/۰۱$$

$$\widehat{\mathrm{var}}(\bar{x}_t) = \frac{s^۲}{n} = ۶۱۶۲/۳$$

همچنین براورد تغییرات فصلی به‌صورت زیر است:

$$\widehat{\mu_t - \mu_{t-۱}} = \bar{x}_t - \bar{x}_{t-۱} = ۲۴۶/۹۶$$

براورد واریانس براورد تغییرات فصلی با $\omega = ۰/۵$ عبارت است از:

$$\widehat{\mathrm{var}}(\bar{x}_t - \bar{x}_{t-۱}) = \widehat{\mathrm{var}}(\bar{x}_t)+\widehat{\mathrm{var}}(\bar{x}_{t-۱})-۲\widehat{\mathrm{cov}}(\bar{x}_t,\bar{x}_{t-۱}) = ۵۹۶۶/۹۳$$

جدول ۲- براورد واریانس‌های براورد ساده و براورد مرکب تعمیم‌یافته

براورد واریانس		میانگین فصلی هزینه‌ی مصرف‌کننده	تغییرات فصلی هزینه‌ی مصرف‌کننده
با استفاده از براورد تعیمیم‌یافته‌ی مرکب	$\omega = ۰/۵$ و $\rho = ۰/۵$	۶۶۹۲/۳	۴۴۴۶/۴۱
	$\omega = ۰/۶$ و $\rho = ۰/۶$	۵۷۲۹/۱۱	۳۳۷۱/۰۸
با استفاده از براورد ساده‌ی میانگین	$\rho = ۰/۵$	۶۱۶۲/۳	۵۹۶۶/۹۳
	$\rho = ۰/۶$	۶۱۶۲/۳	۴۷۷۴/۸۷

و با استفاده از $\omega = ۰/۶$

$$\widehat{\mathrm{var}}(\bar{x}_t - \bar{x}_{t-۱}) = \widehat{\mathrm{var}}(\bar{x}_t) + \widehat{\mathrm{var}}(\bar{x}_{t-۱}) - ۲\widehat{\mathrm{cov}}(\bar{x}_t, \bar{x}_{t-۱}) = ۴۷۷۴/۸۷$$

خلاصه‌ی مقایسه‌ی دو روش براورد ساده و براورد مرکب تعمیم‌یافته‌ی میانگین هزینه‌ی مصرف‌کننده در جدول ۲ آمده است.

۵- بحث و نتیجه‌گیری

در بخش‌های قبل برای میانگین فصلی هزینه‌ی مصرف‌کننده، دو براورد ساده و براورد مرکب تعمیم‌یافته و همچنین براورد واریانس این براوردها را ارایه دادیم. در این مطالعه‌ی موردی، با توجه به جدول ۲، که از داده‌های این طرح به دست آمده است، با استفاده از $\rho = ۰/۵$ ، براورد ساده‌ی‌میانگین هزینه‌ی مصرف‌کننده دارای واریانس کوچک‌تری نسبت به براورد مرکب تعمیم‌یافته‌ی میانگین هزینه‌ی مصرف‌کننده است و با استفاده از $\rho = ۰/۶$ ، براورد مرکب تعمیم‌یافته‌ی میانگین هزینه‌ی مصرف‌کننده دارای واریانس کوچک‌تری نسبت به براورد ساده‌ی میانگین‌هزینه‌ی مصرف‌کننده است. این در حالی است که هم با استفاده از $\rho = ۰/۵$ و هم با استفاده از $\rho = ۰/۶$ ، براورد تغییرات فصلی هزینه‌ی مصرف‌کننده با استفاده از براوردگر مرکب تعمیم‌یافته دارای واریانس کوچک‌تری نسبت به براورد ساده‌ی تغییرات فصلی میانگین هزینه‌ی مصرف‌کننده است و در نتیجه

دقت آن نیز بالاتر است. این موضوع به این دلیل است که براوردگر مرکب تعمیم‌یافته از اطلاعات دوره‌های گذشته‌ی آمارگیری در محاسبات استفاده می‌کند و به‌دلیل وجود همبستگی نسبتاً بالا انتظار داریم که با استفاده از این روش به براورد واریانس کوچک‌تری دست پیدا کنیم. این نتیجه‌گیری برای داده‌های این آمارگیری است. اگرچه براوردگر مرکب تعمیم‌یافته از براوردگر ساده بهتر عمل می‌کند ولی نمی‌توان با این مثال حکم قطعی به برتری براوردگر مرکب تعمیم‌یافته داد.

توضیحات

1. Generalized Composite Estimator
2. U. S. Bureau of Labor Statistics

مرجع‌ها

[1] Breau, P. and Ernst, L. (1983). Alternative estimators to the current composite estimators. In Proceedings of the Survey Research Methods Section, *American Statistical Association*. 397–402.

[2] Cantwell, P.J. (1990). Variance formulae for composite estimators in rotation designs. *Survey Methodology*, **16**, 153–163.

[3] Hansen, M.H.; Hurwitz, W.N. and Madow, W.G. (1953). *Sample Survey Methods and Theory*, (2 vols.). John Wiley and Sons, New York

[4] Jessen, R.J. (1942). Statistical investigation of a sample survey for obtaining farm facts. Iowa State College of Agriculture and Mechanic Arts Research Bulletin, **304**, 54–59.

[5] Park, Y.S.; Kim, K.W.and Choi, J.W. (2001). One-level rotation design balanced on time in monthly sample and in rotation group. *Journal of the American Statistical Association*. **96**, 1483–1496.

[6] Patterson, H.D. (1950). Sampling on successive occasions with partial replacement of units. *Journal of the Royal Statistical Society*. Ser. B, **12**, 241–255.

[7] Yates, F. (1949). *Sampling Methods for Censuses and Surveys*. Griffin, 175–182, 233–235, 260–262.

پیوست

جدول ۳- میانگین‌های هزینه‌ی مصرف‌کننده برای هرکدام از نوبت‌های آمارگیری مربوط به هر دوره

	$\bar{x}_{t,۲}$	$\bar{x}_{t,۳}$	$\bar{x}_{t,۴}$	$\bar{x}_{t,۵}$
۸۵-۱	۵۱۵۰/۱	۵۰۳۸/۶۵	۵۲۸۴/۲۵	۵۵۷۳/۵
۸۵-۲	۵۲۲۷/۳۵	۴۹۷۷/۹۹	۴۸۷۱/۰۲	۵۵۲۶/۳
۸۵-۳	۵۳۸۹/۶۵	۵۲۸۵/۵۹	۵۵۳۷/۷۱	۵۳۵۹/۶
۸۵-۴	۵۰۳۴/۵۲	۵۳۲۹/۱۹	۵۳۲۸/۱۱	۵۶۲۳/۶۱
۸۶-۱	۵۴۵۱/۷۴	۵۱۴۶/۳۵	۵۴۶۶/۰۲	۵۸۰۵/۰۷
۸۶-۲	۵۴۳۸/۳	۵۰۸۳/۸۹	۵۰۴۶/۶۸	۵۶۷۶/۷۸
۸۶-۳	۵۷۴۵/۲۸	۵۶۵۵/۱۸	۵۵۱۸/۹۸	۵۷۰۵/۴۸
۸۶-۴	۵۵۴۷/۶۸	۵۷۲۲/۱	۵۶۵۳/۸۹	۵۷۷۴/۰۶
۸۷-۱	۵۶۵۸/۴	۵۵۲۰/۷۹	۵۷۴۷/۸۶	۵۸۹۸/۱۹
۸۷-۲	۵۳۲۶/۶۴	۵۳۹۸/۸۶	۵۱۶۷/۷	۵۸۵۰/۵۱
۸۷-۳	۵۸۶۱/۴۳	۵۶۶۶/۴۸	۵۸۸۵/۹۹	۶۰۰۰/۴۴
۸۷-۴	۵۸۳۰/۲۸	۵۶۰۴/۷۹	۵۶۷۳/۳۸	۶۰۸۱/۰۴
۸۸-۱	۵۷۸۷/۷۳	۶۱۵۰/۰۸	۵۹۳۲/۹۷	۶۳۱۵/۱۲
۸۸-۲	۵۹۴۵/۵۸	۵۹۰۸/۸۱	۶۳۳۱/۹۷	۶۱۰۷/۰۸
۸۸-۳	۶۳۲۴/۸	۶۰۶۹/۷۴	۶۱۰۲/۸۳	۶۶۱۷/۷۱
۸۸-۴	۶۵۳۲/۲۶	۶۲۴۰/۳	۶۰۲۹/۴۷	۶۳۰۴/۹۲
۸۹-۱	۶۵۲۸/۳۵	۶۶۹۱/۹۹	۶۴۱۰/۶۱	۶۴۷۰/۷۴

به‌عنوان مثال در ۸۹-۱، $\bar{x}_{t,۲}$، میانگین هزینه‌ی واحدهای نمونه‌ای که برای نوبت دوم در فصل اول سال ۸۹ آمارگیری شده‌اند، می‌باشد.

ابراهیم تالانه
فوق لیسانس آمار
تهران، خیابان شهید بهشتی، نبش احمد قصیر، دانشکده‌ی اقتصاد دانشگاه علامه طباطبایی، گروه آمار.
رایانشانی: ebrahim.talaneh@gmail.com

حمیدرضا نواب‌پور
دکتری آمار
تهران، خیابان شهید بهشتی، نبش احمد قصیر، دانشکده‌ی اقتصاد دانشگاه علامه طباطبایی، گروه آمار.
رایانشانی: h.navvabpour@srtc.ac.ir