

## تعیین اندازه‌ی نمونه در طرح نیروی کار مرکز آمار ایران

نعیمه آبی<sup>۱</sup>\* و حمیدرضا نواب‌پور<sup>۲</sup>

<sup>۱</sup> مرکز آمار ایران

<sup>۲</sup> دانشگاه علامه طباطبایی

چکیده: بیش‌تر مراکز ملی آمار، به‌منظور صرفه‌جویی در منابع موجود برای اجرای طرح‌های آمارگیری نمونه‌ای، به ناچار این طرح‌ها را در سطح ملی بهینه می‌کنند؛ ولی نیاز به اطلاعات جزئی‌تر در سطح جامعه‌های کوچک‌تر (مثل استان‌ها، شهرستان‌ها و یا صفت‌هایی مانند جنسیت یا گروه‌های سنی) همواره احساس می‌شود. با توجه به این رویکرد، باید اندازه‌ی نمونه را به گونه‌ای به دست آورد که با کمک آن، برآوردها هم در سطح ملی و هم در سطح هر یک از زیر جامعه‌ها بهینه شوند. از آن‌جا که آمارگیری نمونه‌ای چرخشی یکی از مهم‌ترین روش‌های آمارگیری نمونه‌ای در طرح‌های مراکز ملی آمار به حساب می‌آید؛ در این مقاله ابتدا اندازه‌ی نمونه‌ی بهینه برای آمارگیری نیروی کار- که مرکز آمار آن را به‌صورت فصلی و با استفاده از نمونه‌گیری چرخشی اجرا می‌کند- را به دست آورده، سپس به کمک روش‌های تعیین اندازه‌ی نمونه برای برآورد در ناحیه‌های کوچک، اندازه‌ی نمونه‌ی تخصیص داده‌شده به هر یک از زیرجامعه‌ها (استان‌ها) به دست می‌آید.

### ۱- مقدمه

تعیین اندازه‌ی نمونه یکی از گام‌های اساسی و مؤثر - و معمولاً دشوار- در هر طرح آمارگیری است. اندازه‌ی نمونه باید به حدی باشد که بتوان آماره‌های قابل اطمینانی از هر آمارگیری را به دست آورد. اندازه‌ی نمونه‌ی کوچک و یا بزرگ، تأثیر ناخوشایندی روی

---

واژگان کلیدی: آمارگیری نمونه‌ای چرخشی؛ آمارگیری نیروی کار؛ برآوردگر مرکب تعمیم یافته؛ برآوردگر مرکب  $AK$ ؛ برآورد برای ناحیه‌های کوچک.

دریافت: ۱۳۸۸/۲/۲۲، پذیرش: ۱۳۸۸/۹/۱۱

\* نویسنده‌دار مکاتبات

کیفیت و دقت برآوردهای حاصل از آمارگیری خواهد گذاشت. اهمیت اندازه‌ی نمونه از جنبه‌های اقتصادی نیز قابل بررسی است. اگر اندازه‌ی نمونه کم‌تر از آن‌چه لازم است در نظر گرفته شود، ممکن است بسیاری از منابع که اطلاعات مفیدی را در بردارند از بین بروند و نتوان به‌درستی به هدف‌های مورد نظر دست یافت؛ و چنان‌چه اندازه‌ی نمونه بیش از حد بزرگ در نظر گرفته شود، اجرای طرح، با نیاز مالی بیش‌تری نسبت به آن‌چه مورد احتیاج بوده، مواجه خواهد شد. لذا آمارشناس با مشکل‌هایی از قبیل افزایش هزینه و افزایش خطای غیر نمونه‌گیری روبرو می‌شود. با پیچیده‌تر شدن طرح‌های آمارگیری نمونه‌ای، این مسئله از اهمیت بالاتری برخوردار می‌شود. چنان‌که آمارشناسان همواره به دنبال راه‌هایی برای رسیدن به اندازه‌ی نمونه‌ی بهینه در این آمارگیری‌های نمونه‌ای هستند. در بخش دوم این مقاله، آمارگیری نمونه‌ای چرخشی و برآوردهای مرتبط با آن توضیح داده شده و در بخش سوم، اثر طرح، به‌عنوان راهنما برای تعدیل اندازه‌ی نمونه در این نوع آمارگیری در نظر گرفته می‌شود. در بخش چهارم، اندازه‌ی نمونه برای برآورد در ناحیه‌های کوچک معرفی می‌شود و در بخش پنجم اندازه‌ی نمونه در طرح نیروی کار مرکز آمار ایران که یک آمارگیری نمونه‌ای چرخشی است، در سطح ملی و استانی تعیین می‌شود.

## ۲- انواع آمارگیری‌های نمونه‌ای چرخشی

در چند دهه‌ی اخیر، آمارگیری نمونه‌ای چرخشی به شدت مورد توجه آمارشناسان و مراکز آماري کشورهای مختلف قرار گرفته است. در این شاخه از آمارگیری‌های نمونه‌ای مکرر، فرض می‌شود نمونه در هر دوره شامل  $n$  واحد است، که  $nP$  ( $0 \leq P < 1$ ) واحد از دوره‌ی زمانی  $(t-1)$  ام در دوره‌ی زمانی  $t$  ام باقی می‌مانند - به اصطلاح به آن‌ها واحدهای جور شده (matched units) گفته می‌شود، و  $n(1-P)$  واحد دیگر از نمونه خارج شده و واحدهای جدید جایگزین آن‌ها می‌شوند- واحدهای جدید را واحدهای ناچور (unmatched units) گویند. واحدهای نمونه‌ای طبق نظم مشخص و از پیش تعیین‌شده‌ای به نمونه وارد شده، سپس با گذشت چند دوره‌ی زمانی از نمونه خارج می‌شوند. در حقیقت منظور از لفظ چرخش در این آمارگیری، جای‌گزین شدن همه یا تعدادی از واحدهای نمونه با واحدهای جدید از یک دوره‌ی زمانی به دوره‌ی زمانی دیگر است. چگونگی ورود، خروج و مدت زمان حضور هر واحد آماری در داخل نمونه را به

اصطلاح " الگوی چرخش (rotation pattern)" طرح نمونه‌گیری چرخشی گویند. الگوهای چرخش متفاوت، منجر به تفاوت‌هایی در طرح‌های نمونه‌گیری چرخشی می‌شوند. به‌طور کلی طرح‌های نمونه‌گیری چرخشی به سه گروه نمونه‌گیری چرخشی یک سطحی (one-level rotation sampling)، نمونه‌گیری چرخشی نیم سطحی (semi-level rotation sampling) و نمونه‌گیری چرخشی چندسطحی (multi-level rotation sampling) تقسیم‌بندی می‌شوند. در دو حالت اول، فقط اطلاعات مرتبط با همان دوره‌ی زمانی مراجعه به واحدهای نمونه‌ای - که به اصطلاح به آن دوره‌ی جاری (current period) می‌گویند-، جمع‌آوری می‌شود ولی در حالت سوم، اطلاعات چند دوره‌ی گذشته نیز مورد پرسش قرار می‌گیرند، لذا پاسخ‌گویان ملزم به یادآوری رخدادها در زمان‌های قبل می‌باشند.

نمادگذاری در طرح آمارگیری نمونه‌ای چرخشی یک‌سطحی به صورت  $r_{11} - r_{21} - \dots - r_{2,m-1} - r_{1m}$  است.  $r_{1i}$  و  $r_{2i}$ ، به ترتیب تعداد دوره‌های زمانی که یک واحد نمونه‌ای داخل و خارج نمونه است را نشان می‌دهند. فرایند ورود و خروج به تعداد  $m$  مرتبه قبل از خروج نهایی واحد نمونه‌گیری از نمونه انجام می‌پذیرد. به‌طور کلی اجرای طرح آمارگیری شامل  $M$  دوره‌ی زمانی است. در هر دوره‌ی زمانی طرح آمارگیری نمونه‌ای چرخشی یک‌سطحی متعادل،  $G = \sum_{i=1}^m r_i$  گروه چرخش مورد مطالعه قرار می‌گیرد.

## ۱-۲- برآوردگر مرکب تعمیم‌یافته (Generalized Composite Estimator) در طرح آمارگیری نمونه‌ای چرخشی یک سطحی متعادل

این برآوردگر، کامل‌ترین برآوردگر طرح‌های آمارگیری نمونه‌ای چرخشی است، به طوری که سایر برآوردگرها، حالت خاصی از این برآوردگر هستند. برآوردگر مرکب تعمیم‌یافته‌ی معرفی‌شده توسط برو و ارنست [۱] برای برآورد پارامتر مورد نظر در دوره‌ی زمانی  $t$  ام عبارت است از:

$$(1) \quad y_t = \sum_{i=1}^M a_i x_{t,i} - w \sum_{i=1}^M b_i x_{t-1,i} + w y_{t-1}$$

که  $y_t$  برآورد مرکب تعمیم‌یافته‌ی پارامتر مورد نظر در دوره‌ی زمانی  $t$  ام؛  $x_{t,i}$  برآورد ساده‌ی پارامتر مورد نظر، مربوط به گروه چرخشی که در دوره‌ی زمانی  $t$  ام برای  $i$  امین مرتبه وارد نمونه شده است؛  $x_{t-1,i}$  برآورد ساده‌ی پارامتر مورد نظر، مربوط به گروه چرخشی که در دوره‌ی زمانی  $(t-1)$  ام برای  $i$  امین مرتبه وارد نمونه شده است. ثابت‌های  $w$ ،  $a_i$  ها و  $b_i$  ها توسط تیم مدیریت طرح تعیین می‌شوند و هر کدام می‌توانند مقدارهایی را که در شرط‌های  $0 \leq w < 1$ ،  $\sum_{i=1}^M a_i = 1$  و  $\sum_{i=1}^M b_i = 1$  صدق می‌کنند اختیار نمایند.

### ۱-۱-۲- واریانس برآوردگر مرکب تعمیم‌یافته

یکی از روش‌های مؤثر برای تعدیل اندازه‌ی نمونه در طرح‌های آمارگیری نمونه‌ای پیچیده، استفاده از اثر طرح (یعنی میزان تأثیر طرح جدید بر روی واریانس برآوردگر) است که به صورت زیر تعریف می‌شود:

$$\text{Deff} = \frac{V_{\text{complex}}(\hat{\theta})}{V_{\text{SRS}}(\hat{\theta})}$$

$V_{\text{complex}}(\hat{\theta})$  در این رابطه، واریانس برآوردگر در طرح پیچیده و  $V_{\text{SRS}}(\hat{\theta})$  واریانس برآوردگر در روش نمونه‌گیری تصادفی ساده است. پس برای محاسبه‌ی اثر طرح در آمارگیری نمونه‌ای چرخشی، باید واریانس برآوردگر در این طرح به دست آورده شود. یکی از نکته‌های قابل توجه و تأثیرگذار در به دست آوردن واریانس برآوردگر در آمارگیری نمونه‌ای چرخشی، ساختار کوواریانس گروه‌های یکسان در دوره‌های زمانی مختلف است. در این جا فرض می‌شود، ساختار کوواریانس در طول زمان مانا است. با این پیش فرض، ساختار کوواریانس به صورت زیر در نظر گرفته می‌شود:

الف) برای تمام دوره‌های زمانی  $(t)$  و گروه‌های چرخش  $(i)$ ، واریانس برآورد ساده‌ی پارامتر مورد نظر در هر یک از گروه‌های چرخش عبارت است از:

$$\text{var}(x_{t,i}) = \sigma^2$$

ب) در هر دوره‌ی زمانی، گروه‌های چرخش متفاوت، ناهمبسته هستند. یعنی:

$$\text{cov}(x_{t,i}, x_{t,i}) = 0 \quad i \neq j$$

پ) اگر  $x$  بیان‌گر یک گروه چرخش در دو دوره‌ی زمانی با فاصله‌ی  $|t-s|$  باشد، آن‌گاه:

$$\text{cov}(x_{t,i}, x_{t,i}) = 0 \quad \text{و در غیر این صورت} \quad \text{cov}(x_{t,i}, x_{s,j}) = \rho_{|t-s|} \sigma^2$$

همان‌طور که گفته شد، هر گروه چرخش که برای اولین بار وارد نمونه می‌شود از الگوی چرخش مختص به طرح پیروی می‌کند؛ بر این اساس،  $T_0$  نشان‌دهنده‌ی دوره‌های زمانی که گروه چرخش مد نظر خارج از نمونه قرار دارد می‌باشد. چون طرح متعادل است،  $T_0$  برای تمام گروه‌های چرخش یکسان است. پس از مشخص شدن  $T_0$ ، بردارهای  $a_{M \times 1}$  و  $b_{M \times 1}$  با مولفه‌های زیر تعریف می‌شوند:

$$a_i^* = \begin{cases} 0 & i \in T_0 \\ a_i & \text{سایر جاها} \end{cases} \quad b_i^* = \begin{cases} 0 & i \in T_0 \\ b_i & \text{سایر جاها} \end{cases}$$

کانتول [۲] برای دستیابی به فرمول واریانس برآوردگر، ماتریس  $Q_{M \times M}$  را به صورت زیر معرفی کرد:

$$Q_{ij} = \begin{cases} w^{i-j} p_{i-j} & 1 \leq j < i \leq M \\ 0 & \text{سایر جاها} \end{cases}$$

اگر برآوردگر مرکب تعمیم‌یافته همانند آنچه در (۱) معرفی شده است، به کار گرفته شود و ساختار کوواریانس (الف)، (ب) و (پ) در آن برقرار باشد. آن‌گاه:

$$(۲) \quad V(y_t) = \sigma^2 \frac{a'a + w^2 b'b + 2(a-w^2 b)' Q (a-b)}{(1-w^2)}$$

۲-۲- ارتباط میان برآوردگر مرکب تعمیم‌یافته با برآوردگر مرکب  $AK$  همان‌طور که ملاحظه شد، برای استفاده از فرمول واریانس برآوردگر مرکب تعمیم‌یافته، لازم است مقدارهای مناسب  $a$ ،  $b$  و  $w$  تعیین شوند. فرمول بسته و مشخص شده‌ای برای به دست آوردن این ضریب‌ها وجود ندارد. لذا با معرفی برآوردگر مرکب  $AK$  و بیان ارتباط میان آن با برآوردگر مرکب تعمیم‌یافته، ابتدا، تعداد ضریب‌های نامعلوم را کاهش داده و سپس واریانس معرفی شده به کار گرفته می‌شود.

برآوردگر مرکب  $AK$  (AK composite estimator)

چنانچه  $\bar{x}_t$ ، میانگین  $x_{t,i}$  ( $i = 1, \dots, G$ ) ها باشد با فرض آنکه  $\bar{x}_{m,t}$  و  $\bar{x}_{m,t-1}$  به ترتیب نشان‌گر میانگین گروه‌های چرخش جور شده در دوره‌ی زمانی  $t$  ام و  $(t-1)$  ام و  $\bar{x}_{B,t}$  نیز نشان‌دهنده‌ی میانگین گروه‌های ناجور در دوره‌ی زمانی  $t$  ام باشند. برآوردگر  $AK$  به صورت زیر به دست می‌آید:

$$y_{AK,t} = \frac{G + m(K-1) - Am}{G} \bar{x}_{m,t} + \frac{(1-K+A)m}{G} \bar{x}_{B,t} - K \bar{x}_{m,t-1} + K y_{AK,t-1}$$

که در آن  $A$  و  $K$ ، ثابت‌هایی هستند که با توجه به شرط  $0 \leq A < 1$  و  $0 \leq K < 1$  به دست می‌آیند.  $y_{AK,t-1}$  نیز برآوردگر پارامتر مورد نظر در دوره‌ی زمانی  $t-1$  ام است [۵]. حال، مجموعه‌های  $B$  و  $D$ ، که به ترتیب معرف گروه‌های چرخش ناجور در دوره‌ی زمانی  $t$  ام و  $(t-1)$  ام هستند، به صورت زیر بیان می‌شوند:

$$B = \left\{ i : i = 1, r_1 + 1, \dots, \sum_{k=0}^{m-1} r_k + 1 \right\}$$

که  $r_0 = 0$  و

$$D = \left\{ i : i = r_1, r_1 + r_2, \dots, \sum_{k=1}^m r_k \right\}$$

با توجه به دو مجموعه‌ی تعریف شده در بالا، میانگین گروه‌های چرخش جور شده و ناجور در دوره‌های زمانی  $t$  ام و  $(t-1)$  ام به ترتیب به صورت زیر به دست می‌آیند:

$$\begin{aligned} \bar{x}_{B,t} &= \frac{1}{m} \sum_{i \in B} x_{t,i}, \\ \bar{x}_{m,t} &= \frac{1}{G-m} \sum_{i \notin B} x_{t,i}, \\ \bar{x}_{m,t-1} &= \frac{1}{G-m} \sum_{i \notin D} x_{t-1,i} \end{aligned}$$

با جای‌گذاری رابطه‌های به دست آمده در فرمول برآوردگر مرکب  $AK$ ، برآوردگر  $AK$  به صورت زیر تبدیل می‌شود:

$$y_{AK,t} = \frac{(1-K+A)m}{G} \cdot \frac{1}{m} \sum_{i \in B} x_{t,i} + \frac{G+m(K-1)-Am}{G} \cdot \frac{1}{G-m} \sum_{i \notin B} x_{t,i} - \frac{K}{G-m} \sum_{i \notin D} x_{t-1,i} + K y_{AK,t-1}$$

پس از مقایسه‌ی رابطه‌ی بالا با برآوردگر مرکب تعمیم‌یافته‌ی بیان شده در (۱)، به وضوح نتیجه می‌شود که

$$w = K, \\ b_i = \begin{cases} \frac{1}{G-m} & i \notin D \\ 0 & i \in D \end{cases}$$

و

$$a_i = \begin{cases} \frac{G+m(K-1)-Am}{G(G-m)} & i \notin B \\ \frac{1-K+A}{G} & i \in B. \end{cases}$$

پس مشاهده می‌شود برآوردگر مرکب  $AK$  حالت خاصی از برآوردگر مرکب تعمیم‌یافته است. لذا با به کارگیری ضریب‌های بالا، می‌توان واریانس برآوردگر مرکب  $AK$  را نیز به دست آورد. البته باید مقدارهای بهینه‌ی  $A$  و  $K$  مشخص شوند. کومر و لی [۳]، بیان می‌کنند:  $(A, K)$  ی بهینه، مقدارهایی هستند که در میان تمام مقدارهای ممکن، کمترین واریانس برآوردگر مرکب  $AK$  را ایجاد نمایند.

### ۳- اثر طرح و اندازه‌ی نمونه‌ی بهینه

همان‌طور که گفته شد، با استفاده از اثر طرح می‌توان اندازه‌ی نمونه در آمارگیری‌های پیچیده را تعدیل کرد. پس از محاسبه‌ی اثر طرح، چنان‌چه  $Deff < 1$  به دست آید، برای رسیدن به برآورد در سطح دقت نمونه‌گیری تصادفی ساده، باید اندازه‌ی نمونه در طرح آمارگیری نمونه‌ای پیچیده کاسته شود و چنان‌چه  $Deff > 1$  شود، با افزایش اندازه‌ی نمونه

در طرح آمارگیری نمونه‌ای پیچیده، می‌توان به برآورد پارامترها در سطح دقت نمونه‌گیری تصادفی ساده دست یافت.

#### ۴- اندازه‌ی نمونه برای برآورد در ناحیه‌های کوچک

پس از آن‌که استفاده از روش‌های برآورد برای ناحیه‌ی کوچک در مراکز ملی آمار به‌عنوان یک راهکار معتبر و علمی مورد قبول آمارشناسان قرار گرفت، آن‌ها با رویکرد جدیدی مواجه شدند. در این رویکرد، محدوده‌ی ناحیه‌های کوچک قبل از آغاز آمارگیری مشخص است. این رویکرد کاهش هزینه‌ها و خطای نمونه‌گیری را به دنبال دارد. در این روش آمارشناسان با یک نمونه‌گیری با برنامه روبه‌رو هستند. لذا قادرند با تعیین بهینه‌ی اندازه‌ی نمونه به برآوردهایی دست یابند که هم برای پارامترهای ملی و هم برای هر یک از پارامترهای زیر جامعه‌ها به طور همزمان بهینه باشند.

فرض می‌شود طرح آمارگیری نمونه‌ای در  $D$  زیرناحیه‌ی جامعه قابل اجرا بوده و  $\theta_d$  ( $d = 1, 2, \dots, D$ ) نشان‌دهنده‌ی کمیت مورد نظر در زیرجامعه‌ی  $d$  ام است که با  $\hat{\theta}_d$  برآورد می‌شود. برآوردگر  $\hat{\theta}_d$ ، اگر تنها تابعی از عنصرهای زیرجامعه‌ی  $d$  ام باشد، برآوردگر مستقیم (direct estimator) نامیده می‌شود. در این مقاله فرض می‌شود که  $\hat{\theta}_d$  برآوردگری مستقیم و ناریب است. میانگین توان‌های دوم خطای (MSE) زیرجامعه‌ی  $d$  ام که تابعی از اندازه‌ی آن زیرجامعه ( $n_d$ ) است با  $v_d$  نشان داده می‌شود ( $v_d = v_d(n_d)$ ). نماد  $n$  نشان‌دهنده‌ی اندازه‌ی نمونه‌ی کل است که ثابت فرض می‌شود. اندازه‌ی جامعه با  $N$  و اندازه‌ی زیرجامعه‌ی  $d$  ام با  $N_d$  نمایش داده می‌شود.

اساس کار برای رسیدن به اندازه‌ی نمونه‌ی بهینه در هر ناحیه، مینیمم کردن مجموع واریانس‌های موزون برآوردگرها است. وزن‌های در نظر گرفته‌شده که با  $P_d$  نشان داده می‌شوند، اولویت‌های استنباطی هر ناحیه را منعکس می‌کنند. اولویت استنباطی بالا در یک ناحیه، نشان‌دهنده‌ی اهمیت دقت برآوردها در آن ناحیه است، لذا برای دستیابی به برآورد بهینه باید از میزان پراکندگی در آن ناحیه کاسته شود. ضریب‌های  $P_d$ ، وزن‌هایی هستند که با توجه به اطلاعات کمکی موجود در مورد هر زیرجامعه و میزان اهمیت به آن زیرجامعه تعیین می‌شوند. تعیین مجموعه‌ای مناسب از این ضریب‌ها یکی از دشواری‌ها برای دستیابی به اندازه‌ی نمونه‌ی مورد نیاز است. در بسیاری از آمارگیری‌ها، آمارشناس‌ها

با توجه به میزان تراکم هر زیرجامعه، این ضریب را به دست می‌آورند. در این مقاله، خط فقر هر استان به‌عنوان ضریب استنباطی آن در نظر گرفته شده است. از آن‌جا که لازم است طرح هم در سطح ملی و هم در سطح هر یک از استان‌ها بهینه شود، لذا ضریب دیگری ( $J$ ) نیز در نظر گرفته می‌شود. هرچه این ضریب بزرگ‌تر در نظر گرفته شود، به این مفهوم است که برآورد در سطح ملی از اهمیت بیش‌تری نسبت به برآورد در سطح زیرجامعه‌ها برخوردار است. مقدار مورد نیاز  $J$  در هر طرح آمارگیری، توسط تیم مدیریت طرح مشخص می‌شود.

لانگفورد [۴] پس از معرفی ضریب‌های استنباطی  $P_d$  و  $J$ ، نشان داد برای دستیابی به اندازه‌ی نمونه‌ی بهینه در هر یک از زیرجامعه‌ها، کفایت تابع هدف زیر با توجه به اندازه‌ی نمونه‌ی کل که ثابت فرض شده است مینیمم شود.

$$\sum_{d=1}^D P_d v_d(n_d) + JP_+ v(n)$$

در این رابطه،  $v = \text{var}(\hat{\theta})$ ، واریانس برآورد ملی و  $P_+ = \sum_{d=1}^D P_d$  است، وقتی که  $P'_+$  بردار ضریب‌های استنباطی  $P_d$  و  $\mathbf{1}_D$  یک بردار  $D \times 1$  بُعدی با مؤلفه‌های ۱ باشد. هرگاه نمونه‌گیری تصادفی ساده در هر یک از ناحیه‌ها به کار رود و برآورد طبقه‌بندی مورد استفاده قرار گیرد، با فرض ثابت بودن هزینه‌های آمارگیری در تمام ناحیه‌ها، لانگفورد [۴] نشان داد که اندازه‌ی نمونه‌ی بهینه برای هر یک از ناحیه‌ها عبارت خواهد بود از

$$(۳) \quad n_d = n \frac{\sigma_d \sqrt{P'_d}}{\sigma_1 \sqrt{P'_1} + \sigma_2 \sqrt{P'_2} + \dots + \sigma_D \sqrt{P'_D}}$$

که  $\sigma_d$ ، انحراف استاندارد برآوردگر در زیرجامعه‌ی  $d$  ام و  $(JP_+ N_d / N)$  و  $P'_d = P_d + (JP_+ N_d / N)$  است، که در آن  $N_d$  و  $N$  به‌ترتیب اندازه‌ی زیرجامعه‌ی  $d$  ام و اندازه‌ی کل جامعه هستند.

##### ۵- تعیین اندازه‌ی نمونه در طرح نیروی کار مرکز آمار ایران

طرح آمارگیری نیروی کار یکی از طرح‌های مهم مرکز آمار ایران است که اطلاعات پایه‌ای نیروی کار کشور را تولید می‌کند. این طرح که از روش نمونه‌گیری چرخشی بهره می‌گیرد، به منظور محاسبه‌ی برآوردهای فصلی و سالانه‌ی نیروی کار و همچنین در نظر گرفتن

تغییرهای آن در سال‌ها و فصل‌های مختلف طراحی شده است. الگوی چرخش انتخاب‌شده برای این طرح یک الگوی «۲-۲-۲» است، یعنی از واحدهای نمونه‌ای، چهار بار آمارگیری به عمل می‌آید، به این ترتیب که هر واحد، دو فصل متوالی در نمونه است، سپس به‌طور موقت برای دو فصل متوالی از نمونه خارج می‌شود، بعد مجدداً برای دو فصل متوالی به نمونه بازگشته و پس از آن برای همیشه از نمونه خارج می‌شود.

در این طرح، با هدف برآورد نسبت بیکاران فصلی در سطح کل کشور و سپس در سطح هر یک از استان‌ها، ابتدا اندازه‌ی نمونه در سطح ملی را به دست آورده، سپس به کمک روش‌های تعیین اندازه‌ی نمونه برای برآورد در ناحیه‌های کوچک، اندازه‌ی نمونه در هر یک از استان‌ها تعیین می‌شود.

همان‌طور که بیان شد، با کمک اثر طرح می‌توان اندازه‌ی نمونه در آمارگیری نمونه‌ای چرخشی را تعدیل کرد. برای محاسبه‌ی اثر طرح، باید واریانس برآوردگر تحت طرح نمونه‌گیری تصادفی ساده نیز محاسبه شود. به علت عدم دسترسی به نمونه‌ی پایه‌ی طرح نیروی کار مرکز آمار، ۴۴۵۶۸ خانوار نمونه‌ای انتخاب شده توسط این مرکز در فصل پاییز سال ۱۳۸۵، به‌عنوان واحدهای چارچوب نمونه‌گیری در نظر گرفته شده و از میان آن‌ها  $n_0 = 22284$  خانوار به‌عنوان اندازه‌ی نمونه برای طرح نمونه‌گیری چرخشی برگزیده شدند. به بیان دیگر، برای به دست آوردن اثر طرح، نمونه‌ای به اندازه‌ی  $n_0 = 22284$  خانوار، یک مرتبه به شیوه‌ی نمونه‌گیری تصادفی ساده و یک مرتبه به شیوه‌ی نمونه‌گیری چرخشی- با حفظ الگوی چرخش طرح نیروی کار مرکز آمار- مورد بررسی قرار گرفتند. انتظار می‌رود با انتخاب این نمونه‌ی ۵۰ درصدی از واحدهای چارچوب، بتوان نتیجه‌های به دست آمده را به چارچوب مورد استفاده‌ی مرکز آمار ایران تعمیم داد.

#### ۱-۵- واریانس نمونه‌ی گزینش شده در روش نمونه‌گیری تصادفی ساده

پس از گزینش واحدهای نمونه‌ای به روش نمونه‌گیری تصادفی ساده، انحراف استاندارد برای برآورد پارامتر مورد نظر (نسبت بیکاران) برابر با  $0.00556$  به دست آمد.

#### ۲-۵- واریانس نمونه‌ی گزینش شده در روش نمونه‌گیری چرخشی

واریانس برآورد نسبت بیکاران در هر یک از گروه‌های چرخش در رابطه‌ی (۱) با  $\sigma^2$

نشان داده شده است. با فرض برابری واریانس در هر چهار گروه چرخش، برآورد واریانس برآورد نسبت بیکاران در هر گروه برابر با  $s^2 = 2(0/001056)^2$  به دست آمد. بردارهای  $a$  و  $b$  به کار رفته در فرمول (۱)، طبق تعریف‌های مطرح شده در بخش ۲-۱-۱ عبارت‌اند از:

$$a = (a, a, \dots, a, a) \quad b = (b, b, \dots, b, b)$$

برای به دست آوردن برآورد واریانس برآورد نسبت بیکاران فصل سوم طرح آمارگیری نیروی کار مرکز آمار در سال ۱۳۸۵، باید مقدار ضریب  $w$  و  $\lambda$  مؤلفه‌ی بردارهای  $a$  و  $b$  مشخص شوند. انتخاب بهینه‌ی این ضریب‌ها به‌طور همزمان دشوار است، لذا از ارتباط میان برآوردگر مرکب تعمیم‌یافته با برآوردگر مرکب  $AK$  کمک گرفته و تعداد ضریب‌ها به دو ضریب  $A$  و  $K$  تقلیل داده می‌شوند. به همین منظور، مجموعه‌های  $B$  و  $D$  همانند آن‌چه در بخش ۲-۲ بیان شد، به قرار زیر معرفی می‌شوند:

$$B = \left\{ i : i = 1, r_1 + 1, \dots, \sum_{k=0}^{r-1} r_{1k} + 1 \right\} = \{1, 3\}$$

$$D = \left\{ i : i = r_1, r_1 + r_2, \dots, \sum_{k=1}^r r_{1k} \right\} = \{2, 4\}.$$

با توجه به مجموعه‌های تعریف شده در بالا، مؤلفه‌های بردارهای  $a$  و  $b$  عبارت‌اند از:

$$a = \left( \frac{1-K+A}{G}, \frac{G+m(K-1)-Am}{G(G-m)}, \dots, \frac{1-K+A}{G}, \frac{G+m(K-1)-Am}{G(G-m)} \right)'$$

$$b = \left( \frac{1}{G-m}, \dots, \dots, \frac{1}{G-m} \right)'$$

بنا بر این تنها ضریب‌های نامعلوم ضریب‌های  $A$  و  $K$  هستند. قابل ذکر است،  $A$  و  $K$  ای که منجر به کوچک‌ترین مقدار واریانس برآوردگر شوند، مقدارهای  $A$  و  $K$  ی بهینه خواهند بود.

برای دستیابی به واریانس برآوردگر، باید ماتریس  $Q$  نیز مشخص شود. با توجه به فرم تعریف شده‌ی این ماتریس در بخش ۲-۱-۱، ماتریس  $Q$  عبارت است از:

$$Q = \begin{bmatrix} \circ & \circ & \circ & \circ & \circ & \circ \\ w^1 \rho_1 & \circ & \circ & \circ & \circ & \circ \\ w^2 \rho_2 & w^1 \rho_1 & \circ & \circ & \circ & \circ \\ w^3 \rho_3 & w^2 \rho_2 & w^1 \rho_1 & \circ & \circ & \circ \\ w^4 \rho_4 & w^3 \rho_3 & w^2 \rho_2 & w^1 \rho_1 & \circ & \circ \\ w^5 \rho_5 & w^4 \rho_4 & w^3 \rho_3 & w^2 \rho_2 & w^1 \rho_1 & \circ \end{bmatrix}$$

$\rho_i$  بیان‌گر میزان همبستگی گروه‌های چرخش مشترک در فصل‌های با اختلاف  $i$  است. چنانچه میزان همبستگی گروه‌های چرخش مشترک میان دو فصل متوالی  $\rho$  در نظر گرفته شود، با فرض برابری همبستگی گروه‌های چرخش مشترک،  $\rho_i$  ها در طرح ۲-۲-۲ به قرار زیر به دست می‌آیند:

- دو فصلی که فاصله‌ی میان آن‌ها، یک (یا چهار) دوره‌ی زمانی است، شامل دو گروه چرخش مشترک می‌باشند، لذا،

$$\rho_1 = \rho_4 = \rho$$

- دو فصلی که دو دوره‌ی زمانی با یکدیگر فاصله دارند، شامل گروه‌های چرخش مشترک نمی‌باشند، لذا،

$$\rho_2 = \circ$$

- دو فصلی که فاصله‌ی میان آن‌ها، سه (یا پنج) دوره‌ی زمانی است، شامل یک گروه چرخش مشترک می‌باشند، لذا،

$$\rho_3 = \rho_5 = \circ/5\rho$$

برآورد همبستگی گروه‌های چرخش مشترک در فصل تابستان و پاییز سال ۱۳۸۵ طرح نیروی کار مرکز آمار ایران- که یک دوره‌ی زمانی با یکدیگر فاصله دارند- برابر  $r \cong \circ/32$  به دست آمد.

پس از مشخص شدن ماتریس  $Q$ ، با کمک برنامه‌نویسی در R برای مقدارهای مختلف  $A$  و  $K$ ، واریانس و اثر طرح برآوردگر محاسبه شد. اثرهای طرح متناسب با هر مقدار  $A$  و  $K$ ، در جدول ۱، نشان داده شده است.

با توجه به نتیجه‌های نشان داده شده در جدول، با  $A = K = \circ/1$  کم‌ترین مقدار برآورد واریانس حاصل می‌شود.

## ۳-۵- تعدیل اندازه‌ی نمونه

پس از محاسبه‌ی واریانس در هر یک از دو روش نمونه‌گیری تصادفی ساده و نمونه‌گیری چرخشی، می‌توان مقدار  $Deff$  را به دست آورد.

$$Deff \cong 0/9107048$$

مقدار اثر طرح به دست آمده، بیان‌گر آن است که با تعداد نمونه‌ی کم‌تری می‌توان به سطح دقت مورد نظر در نمونه‌گیری تصادفی ساده دست یافت. تعداد نمونه‌ی تعدیل‌شده به کمک اثر طرح، عبارت است از:

$$n = n_0 \cdot Deff \\ = 22284 \times 0/9107048 \cong 20294$$

## ۴-۵- اندازه‌ی نمونه در سطح استان‌های کشور

در طرح نیروی کار مرکز آمار ایران، پس از آن‌که برآوردها در سطح ملی به دست آورده شدند، لازم است برآورد پارامترهای مورد نظر در سطح هر یک از استان‌های کشور نیز گزارش شوند؛ به‌همین منظور برآوردهای ناحیه‌ی کوچک مورد استفاده قرار می‌گیرد، و چون از ابتدای شروع به اجرای طرح، به دست آوردن برآوردها در سطح استان‌ها جزء هدف‌های طرح است، لذا می‌توان با انتخاب اندازه‌ی نمونه‌ی مناسب در سطح هر یک از استان‌ها به گونه‌ای عمل کرد که برآوردها، هم در سطح ملی و هم در سطح هر یک از استان‌ها بهینه باشند.

همان‌طور که در بخش ۳ بیان شد، برای دستیابی به اندازه‌ی نمونه‌ی بهینه در هر یک از استان‌ها، لازم است ضریب‌هایی به‌عنوان ضریب اولویت استنباطی به هر یک از ناحیه‌های کوچک اختصاص داده شده و بر اساس آن در مورد اندازه‌ی نمونه‌ی بهینه‌ی هر یک از استان‌ها تصمیم‌گیری شود. از آن‌جا که برآورد نسبت بیکاران یکی از شاخص‌های مهم اقتصادی در تصمیم‌گیری‌های کلان جامعه است، لذا برآوردها در استان‌های فقیرتر از اهمیت بسزایی برخوردار هستند. به‌همین منظور خط فقر هر استان در سال ۱۳۸۵ به‌عنوان ضریب اولویت استنباطی در تعیین اندازه‌ی نمونه‌ی بهینه مورد استفاده قرار گرفته است. به بیان دیگر، هر چه خط فقر در استانی بالاتر باشد (درصد تعداد افراد زیر خط فقر بیشتر باشد)، ضریب اولویت استنباطی آن بزرگ‌تر در نظر گرفته می‌شود. پس وزن‌های

$P_d$  به صورت زیر محاسبه می‌شوند:

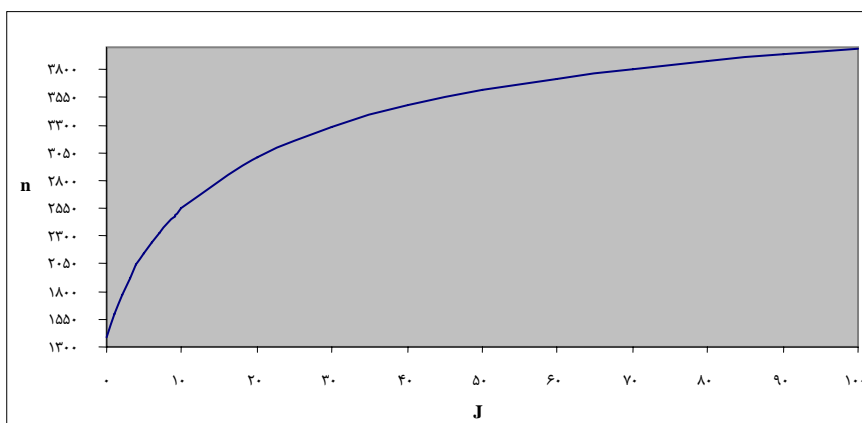
$$p_d = \frac{\text{خط فقر در استان } d \text{ ام}}{\text{مجموع خط فقر استان‌ها}} \text{ (اولویت استنباطی استان } d \text{ ام)}$$

پس از مشخص شدن ضریب‌های اولویت استنباطی، با کمک رابطه‌ی (۳) اندازه‌ی نمونه‌ی بهینه در هر استان مشخص خواهد شد.

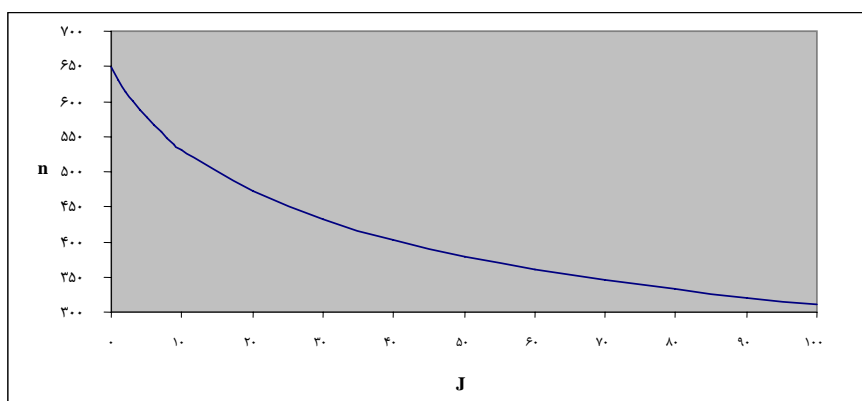
ضریب  $J$ ، عامل دیگری است که در تعیین اندازه‌ی نمونه در سطح هر یک از استان‌ها تأثیرگذار است. این ضریب، میزان اهمیت برآوردها در سطح ملی را نشان می‌دهد، پس هر چه مقدار بزرگتری برای آن در نظر گرفته شود، برآورد در سطح ملی اهمیت بیشتری نسبت به برآورد در سطح هر یک از استان‌ها خواهد داشت. همان‌طور که در بخش چهارم گفته شد، تصمیم‌گیری در مورد مقدار این ضریب بسته به هدف‌های طرح بر عهده‌ی تیم مدیریت طرح می‌باشد. جدول (۲) در بردارنده‌ی اندازه‌ی نمونه به دست آمده برای مقدارهای مختلف  $J$  است. نسبت نمونه‌ی تخصیص داده شده به هر استان نیز در جدول (۳) نشان داده شده است.

همان‌طور که در جدول (۳) مشاهده می‌شود، هر چه مقدار  $J$  افزایش یابد، سهم نمونه‌ی تخصیص داده شده به استان تهران نیز افزایش می‌یابد. در حقیقت، هر چه  $J$  بزرگ‌تر در نظر گرفته شود (برآوردها در سطح ملی از اهمیت بالاتری برخوردار شوند)، تعداد نمونه‌ی تخصیص داده شده به پرجمعیت‌ترین استان، بسیار بیشتر از سایر استان‌ها می‌شود. با کوچک‌تر انتخاب کردن مقدار  $J$ ، این مسئله تعدیل می‌یابد. پس از تهران، بیش‌ترین اندازه‌ی نمونه‌ی به استان اصفهان تخصیص داده شده است (برای  $J = 5$ ، حدود ۵ درصد از کل نمونه‌ی گزینش شده).

دو نمودار در شکل‌های ۱ و ۲ روند تغییر تعداد نمونه در پرجمعیت‌ترین و کم جمعیت‌ترین استان کشور را با توجه به مقدارهای مختلف  $J$  نشان می‌دهند.



شکل ۱- روند تغییر اندازه‌ی نمونه در استان تهران



شکل ۲- روند تغییر اندازه‌ی نمونه در استان ایلام

همان‌گونه که در شکل‌ها مشاهده می‌شود، با افزایش مقدار  $J$ ، اندازه‌ی نمونه در پرجمعیت‌ترین استان کشور (تهران) روندی صعودی و در کم جمعیت‌ترین استان کشور (ایلام) روندی نزولی پیدا می‌کند. در حقیقت هرگاه برآورد در سطح ملی نسبت به برآورد در سطح استان‌ها از اهمیت بالاتری برخوردار باشد، باید از ناحیه‌های بزرگ‌تر، تعداد نمونه‌ی بیش‌تری گزینش شود و بالعکس.

جدول ۱- اثر طرح برآوردگر نسبت بیکاران، متناسب با هر یک از مقادیرهای  $A$  و  $K$ 

$A$	$K$	Deff	$A$	$K$	Deff
۰/۱	۰/۱	۰/۹۱۰۷۰۴۸	۰/۵	۰/۶	۱/۳۰۸۷۲۷
۰/۱	۰/۲	۰/۹۴۰۵۵۳۶	۰/۵	۰/۷	۱/۵۷۹۹۶۴
۰/۱	۰/۳	۱/۰۰۲۹۳۸	۰/۵	۰/۸	۲/۱۳۵۶۶۶
۰/۱	۰/۴	۱/۱۰۶۲۸۵	۰/۵	۰/۹	۳/۷۹۴۱۳۸
۰/۱	۰/۵	۱/۲۶۸۷۳۶	۰/۶	۰/۱	۱/۱۴۵۶۰۱
۰/۱	۰/۶	۱/۵۲۸۴۷۷	۰/۶	۰/۲	۱/۱۰۱۲۰۵
۰/۱	۰/۷	۱/۹۷۵۴۷۶	۰/۶	۰/۳	۱/۰۹۰۲۹۰
۰/۱	۰/۸	۲/۸۷۹۴۹۴	۰/۶	۰/۴	۱/۱۱۴۵۳۰
۰/۱	۰/۹	۵/۵۸۸۶۴۱	۰/۶	۰/۵	۱/۱۸۱۸۸۴
۰/۲	۰/۱	۰/۹۲۲۴۲۲۶	۰/۶	۰/۶	۱/۳۱۱۹۰۱
۰/۲	۰/۲	۰/۹۳۷۵۳۵	۰/۶	۰/۷	۱/۵۵۲۷۶۲
۰/۲	۰/۳	۰/۹۸۴۵۸۹۸	۰/۶	۰/۸	۲/۰۵۰۷۳۸
۰/۲	۰/۴	۱/۰۷۰۴۲۷	۰/۶	۰/۹	۳/۵۳۸۹۶۶
۰/۲	۰/۵	۱/۲۱۰۶۵۳	۰/۷	۰/۱	۱/۲۴۵۴۷۲
۰/۲	۰/۶	۱/۴۳۸۶۷۳	۰/۷	۰/۲	۱/۱۸۶۰۵۸
۰/۲	۰/۷	۱/۸۳۳۵۹۳	۰/۷	۰/۳	۱/۱۶۱۴۸۸
۰/۲	۰/۸	۲/۶۳۲۹۱۹	۰/۷	۰/۴	۱/۱۷۲۴۴۰
۰/۲	۰/۹	۵/۰۲۳۹۴۳	۰/۷	۰/۵	۱/۲۲۵۵۸۳
۰/۳	۰/۱	۰/۹۵۱۷۷۱۱	۰/۷	۰/۶	۱/۳۳۸۳۱۹
۰/۳	۰/۲	۰/۹۵۲۰۹۰۸	۰/۷	۰/۷	۱/۵۵۴۲۲۹
۰/۳	۰/۳	۰/۹۸۴۱۵۰۷	۰/۷	۰/۸	۲/۰۰۶۲۲۲
۰/۳	۰/۴	۱/۰۵۳۳۲۲	۰/۷	۰/۹	۳/۳۶۱۱۷۶
۰/۳	۰/۵	۱/۱۷۲۹۲۶	۰/۸	۰/۱	۱/۳۶۲۹۷۴
۰/۳	۰/۶	۱/۳۷۲۱۱۴	۰/۸	۰/۲	۱/۲۸۸۴۸۶
۰/۳	۰/۷	۱/۷۲۰۳۸	۰/۸	۰/۳	۱/۲۵۰۵۹۶
۰/۳	۰/۸	۲/۴۲۶۷۵۶	۰/۸	۰/۴	۱/۲۴۹۱۰۴
۰/۳	۰/۹	۴/۵۳۶۶۲۷	۰/۸	۰/۵	۱/۲۸۹۶۳۸
۰/۴	۰/۱	۰/۹۹۸۱۷۵۰۴	۰/۸	۰/۶	۱/۳۸۱۷۹۸۱
۰/۴	۰/۲	۰/۹۸۴۲۲۱	۰/۸	۰/۷	۱/۵۸۴۳۶۶
۰/۴	۰/۳	۱/۰۰۱۶۲۱	۰/۸	۰/۸	۲/۰۰۲۱۱۸
۰/۴	۰/۴	۱/۰۵۴۹۷۱	۰/۸	۰/۹	۳/۲۶۰۷۶۷
۰/۴	۰/۵	۱/۱۵۵۵۵۶	۰/۹	۰/۱	۱/۴۹۸۱۰۷
۰/۴	۰/۶	۱/۳۲۸۷۹۸	۰/۹	۰/۲	۱/۴۰۸۴۸۹
۰/۴	۰/۷	۱/۶۳۵۸۳۷	۰/۹	۰/۳	۱/۳۵۷۶۱۳
۰/۴	۰/۸	۲/۲۶۱۰۰۵	۰/۹	۰/۴	۱/۳۴۴۵۲۱
۰/۴	۰/۹	۴/۱۲۶۶۹۲	۰/۹	۰/۵	۱/۳۷۴۰۵۰
۰/۵	۰/۱	۱/۰۶۳۳۶۰	۰/۹	۰/۶	۱/۴۶۰۸۸۷

جدول ۱- اثر طرح برآورده‌گر نسبت بیکاران، متناسب با هر یک از مقدارهای  $A$  و  $K$ 

$A$	$K$	Deff	$A$	$K$	Deff
۰/۵	۰/۲	۱/۰۳۳۹۲۶	۰/۹	۰/۷	۱/۶۴۳۱۷۴
۰/۵	۰/۳	۱/۰۳۷۰۰۱	۰/۹	۰/۸	۲/۰۳۸۴۲۶
۰/۵	۰/۴	۱/۰۷۵۳۷۴	۰/۹	۰/۹	۳/۲۳۷۷۴۰
۰/۵	۰/۵	۱/۱۵۸۵۴۲			

جدول ۲- اندازه‌ی نمونه‌ی تخصیص داده‌شده به هر استان با توجه به مقدارهای مختلف برای  $J$ 

استان	اولویت	$J=۰$	$J=۱$	$J=۲$	$J=۳$	$J=۴$	$J=۵$	$J=۶$
تهران		۱۳۸۶	۱۵۹۹	۱۷۷۱	۱۹۱۴	۲۰۲۸	۲۱۴۷	۲۲۴۴
اردبیل		۵۷۶	۵۶۳	۵۵۱	۵۴۲	۵۳۳	۵۲۵	۵۱۸
اصفهان		۹۲۴	۹۴۴	۹۶۲	۹۸۰	۹۹۶	۱۰۱۱	۱۰۲۶
ایلام		۶۴۸	۶۳۰	۶۱۴	۶۰۱	۵۸۸	۵۷۷	۵۶۶
آذربایجان شرقی		۴۴۰	۴۴۸	۴۵۶	۴۶۳	۴۷۰	۴۷۶	۴۸۲
آذربایجان غربی		۵۱۳	۵۱۵	۵۱۷	۵۲۰	۵۲۳	۵۲۵	۵۲۸
بوشهر		۶۱۵	۵۹۹	۵۸۵	۵۷۲	۵۶۱	۵۵۱	۵۴۲
چهارمحال بختیاری		۶۵۰	۶۳۳	۶۱۸	۶۰۵	۵۹۳	۵۸۳	۵۷۳
خراسان جنوبی		۶۰۱	۵۸۵	۵۷۱	۵۵۸	۵۴۷	۵۳۷	۵۲۸
خراسان رضوی		۵۶۶	۶۱۵	۶۵۷	۶۹۴	۷۲۶	۷۵۴	۷۸۰
خراسان شمالی		۴۴۸	۴۳۶	۴۲۷	۴۱۸	۴۱۰	۴۰۳	۳۹۷
خوزستان		۷۴۱	۷۴۵	۷۵۰	۷۵۵	۷۶۰	۷۶۵	۷۷۰
زنجان		۷۲۲	۷۰۳	۶۸۷	۶۷۳	۶۶۰	۶۴۸	۶۳۸
سمنان		۸۲۳	۸۰۰	۷۸۱	۷۶۳	۷۴۷	۷۳۳	۷۲۰
سیستان و بلوچستان		۴۲۳	۴۲۳	۴۲۴	۴۲۵	۴۲۶	۴۲۸	۴۲۹
فارس		۶۵۸	۷۰۲	۷۱۸	۷۳۳	۷۴۶	۷۵۹	۷۷۱
قزوین		۵۸۱	۵۶۷	۵۵۶	۵۴۵	۵۳۶	۵۲۸	۵۲۰
قم		۸۶۰	۸۳۷	۸۱۸	۸۰۱	۷۸۵	۷۷۱	۷۵۸
کردستان		۵۷۸	۵۶۵	۵۵۴	۵۴۴	۵۳۶	۵۲۸	۵۲۲
کرمان		۶۲۰	۶۲۳	۶۲۶	۶۲۹	۶۳۳	۶۳۷	۶۴۰
کرمانشاه		۶۴۷	۶۳۹	۶۳۳	۶۲۸	۶۲۴	۶۲۰	۶۱۷
کهگیلویه و بویراحمد		۷۵۹	۷۳۸	۷۲۰	۷۰۳	۶۸۹	۶۷۵	۶۶۳
گلستان		۵۷۷	۵۶۴	۵۵۴	۵۴۵	۵۳۷	۵۳۰	۵۲۴
گیلان		۸۷۳	۸۶۳	۸۵۵	۸۴۸	۸۴۲	۸۳۸	۸۳۴
لرستان		۷۰۴	۶۹۰	۶۷۸	۶۶۸	۶۶۰	۶۵۲	۶۴۵
مازندران		۸۰۶	۷۹۸	۷۹۱	۷۸۶	۷۸۳	۷۷۹	۷۷۷
مرکزی		۷۲۹	۷۱۳	۶۹۹	۶۸۷	۶۷۶	۶۶۷	۶۵۸
هرمزگان		۴۸۵	۴۷۳	۴۶۴	۴۵۵	۴۴۷	۴۴۰	۴۳۴
همدان		۶۶۶	۶۵۴	۶۴۵	۶۳۷	۶۳۰	۶۲۴	۶۱۸
یزد		۶۴۵	۶۲۸	۶۱۴	۶۰۱	۵۹۰	۵۸۰	۵۷۱

گزیده‌مطالب آماری، سال ۲۰، شماره‌ی ۲، پاییز و زمستان ۱۳۸۸، صص ۲۲۹-۲۴۹

ادامه‌ی جدول ۲- اندازه‌ی نمونه‌ی تخصیص داده‌شده به هر استان با توجه به مقدارهای مختلف برای  $J$ 

استان	اولویت	$J=7$	$J=8$	$J=9$	$J=10$	$J=20$	$J=30$
تهران		۲۳۳۰	۲۴۰۹	۲۴۸۱	۲۵۴۸	۳۰۱۱	۳۲۸۶
اردبیل		۵۱۱	۵۰۵	۵۰۰	۴۹۴	۴۵۷	۴۳۳
اصفهان		۱۰۳۹	۱۰۵۲	۱۰۶۴	۱۰۷۶	۱۱۶۴	۱۲۲۲
ایلام		۵۵۷	۵۴۸	۵۳۹	۵۳۱	۴۷۲	۴۳۲
آذربایجان شرقی		۴۸۸	۴۹۴	۴۹۹	۵۰۴	۵۴۲	۵۶۷
آذربایجان غربی		۵۳۱	۵۳۴	۵۳۷	۵۳۹	۵۶۲	۵۷۸
بوشهر		۵۳۳	۵۲۵	۵۱۸	۵۱۱	۴۵۸	۴۲۳
چهارمحال بختیاری		۵۶۴	۵۵۶	۵۴۸	۵۴۱	۴۸۷	۴۵۱
خراسان جنوبی		۵۱۹	۵۱۱	۵۰۳	۴۹۶	۴۴۳	۴۰۹
خراسان رضوی		۸۰۴	۸۲۶	۸۴۵	۸۶۴	۹۹۶	۱۰۷۶
خراسان شمالی		۳۹۱	۳۸۵	۳۸۰	۳۷۵	۳۳۹	۳۱۶
خوزستان		۷۷۵	۷۸۰	۷۸۴	۷۸۹	۸۲۶	۸۵۳
زنجان		۶۲۸	۶۱۹	۶۱۰	۶۰۳	۵۴۳	۵۰۵
سمنان		۷۰۷	۶۹۶	۶۸۵	۶۷۵	۵۹۹	۵۴۸
سیستان و بلوچستان		۴۳۱	۴۳۲	۴۳۴	۴۳۵	۴۴۹	۴۶۰
فارس		۷۸۳	۷۹۳	۸۰۳	۸۱۳	۸۸۵	۹۳۲
قزوین		۵۱۳	۵۰۷	۵۰۱	۴۹۶	۴۵۵	۴۳۰
قم		۷۴۶	۷۳۵	۷۲۵	۷۱۶	۶۴۳	۵۹۶
کردستان		۵۱۵	۵۱۰	۵۰۴	۴۹۹	۴۶۴	۴۴۲
کرمان		۶۴۴	۶۴۷	۶۵۱	۶۵۴	۶۸۲	۷۰۳
کرمانشاه		۶۱۵	۶۱۳	۶۱۱	۶۰۹	۶۰۰	۵۹۸
کهگیلویه و بویراحمد		۶۵۱	۶۴۱	۶۳۱	۶۲۲	۵۵۰	۵۰۳
گلستان		۵۱۸	۵۱۳	۵۰۸	۵۰۳	۴۷۱	۴۵۱
گیلان		۸۳۱	۸۲۸	۸۲۵	۸۲۳	۸۱۱	۸۰۸
لرستان		۶۳۸	۶۳۳	۶۲۷	۶۲۲	۵۸۸	۵۶۸
مازندران		۷۷۵	۷۷۳	۷۷۱	۷۷۰	۷۶۵	۷۶۶
مرکزی		۶۵۱	۶۴۳	۶۳۷	۶۳۱	۵۸۶	۵۵۹
هرمزگان		۴۲۸	۴۲۳	۴۱۸	۴۱۳	۳۷۹	۳۵۸
همدان		۶۱۴	۶۰۹	۶۰۵	۶۰۲	۵۷۸	۵۶۵
یزد		۵۶۲	۵۵۴	۵۴۶	۵۳۹	۴۸۷	۴۵۳

جدول ۳- نسبت نمونه‌ی تخصیص داده شده به هر استان با توجه به مقادیرهای مختلف برای  $J$ 

استان	اولویت	$J=0$	$J=1$	$J=2$	$J=3$	$J=4$
تهران		۶/۸۳	۷/۸۸	۸/۷۳	۹/۴۳	۱۰/۰۴
اردبیل		۲/۸۴	۲/۷۷	۲/۷۲	۲/۶۷	۲/۶۳
اصفهان		۴/۵۵	۴/۶۵	۴/۷۴	۴/۸۳	۴/۹۱
ایلام		۳/۱۹	۳/۱۰	۳/۰۳	۲/۹۶	۲/۹۰
آذربایجان شرقی		۲/۱۷	۲/۲۱	۲/۲۵	۲/۲۸	۲/۳۲
آذربایجان غربی		۲/۵۳	۲/۵۴	۲/۵۵	۲/۵۶	۲/۵۸
بوشهر		۲/۰۳	۲/۹۵	۲/۸۸	۲/۸۲	۲/۷۷
چهارمحال بختیاری		۳/۲۰	۳/۱۲	۳/۰۵	۲/۹۸	۲/۹۲
خراسان جنوبی		۲/۹۶	۲/۸۸	۲/۸۱	۲/۷۵	۲/۷۰
خراسان رضوی		۲/۷۹	۳/۰۳	۳/۲۴	۳/۴۲	۳/۵۸
خراسان شمالی		۲/۲۱	۲/۱۵	۲/۱۰	۲/۰۶	۲/۰۲
خوزستان		۳/۶۵	۳/۶۷	۳/۶۹	۳/۷۲	۳/۷۵
زنجان		۳/۵۶	۳/۴۷	۳/۳۹	۳/۳۲	۳/۲۵
سمنان		۴/۰۶	۳/۹۴	۳/۸۵	۲/۷۶	۳/۶۸
سیستان و بلوچستان		۲/۰۹	۲/۰۹	۲/۰۹	۲/۰۹	۲/۱۰
فارس		۳/۳۷	۳/۴۶	۳/۵۴	۳/۶۱	۳/۶۸
قزوین		۲/۸۶	۲/۸۰	۲/۷۴	۲/۶۹	۲/۶۴
قم		۴/۲۴	۴/۱۳	۴/۰۳	۳/۹۵	۳/۸۷
کردستان		۲/۸۵	۲/۷۸	۲/۷۳	۲/۶۸	۲/۶۴
کرمان		۳/۰۶	۳/۰۷	۳/۰۸	۳/۱۰	۳/۱۲
کرمانشاه		۳/۱۹	۳/۱۵	۳/۱۲	۳/۰۹	۳/۰۷
کهگیلویه و بویراحمد		۳/۷۴	۳/۶۴	۳/۵۵	۳/۴۷	۳/۳۹
گلستان		۲/۸۴	۲/۷۸	۲/۷۳	۲/۶۹	۲/۶۵
گیلان		۴/۳۰	۴/۲۵	۴/۲۱	۴/۱۸	۴/۱۵
لرستان		۳/۴۷	۳/۴۰	۳/۳۴	۳/۲۹	۳/۲۵
مازندران		۳/۹۷	۳/۹۳	۳/۹۰	۳/۸۸	۳/۸۶
مرکزی		۳/۵۹	۳/۵۱	۳/۴۴	۳/۳۹	۳/۳۳
هرمزگان		۲/۳۹	۲/۳۳	۲/۲۸	۲/۲۴	۲/۲۰
همدان		۳/۲۸	۳/۲۲	۳/۱۸	۳/۱۴	۳/۱۰
یزد		۳/۱۸	۳/۱۰	۳/۰۳	۲/۹۶	۲/۹۱

ادامه‌ی جدول ۳- نسبت نمونه‌ی تخصیص داده شده به هر استان با توجه به مقادیرهای مختلف برای  $J$

استان	اولویت	$J=5$	$J=6$	$J=7$	$J=8$	$J=9$
تهران		۱۰/۵۸	۱۱/۰۶	۱۱/۴۸	۱۱/۸۷	۱۲/۲۳
اردبیل		۲/۵۹	۲/۵۵	۵۲۲	۲/۴۹	۲/۴۶
اصفهان		۴/۹۸	۵/۰۵	۵/۱	۵/۱۹	۵/۲۴
ایلام		۲/۸۴	۲/۷۹	۲/۷۴	۲/۷۰	۲/۶۶
آذربایجان شرقی		۲/۳۵	۲/۳۸	۲/۴۱	۲/۴۳	۲/۴۶
آذربایجان غربی		۲/۵۹	۲/۶۰	۲/۶۲	۲/۶۳	۲/۶۴
بوشهر		۲/۷۲	۲/۶۷	۲/۶۳	۲/۵۹	۲/۵۵
چهارمحال بختیاری		۲/۸۷	۲/۸۲	۲/۷۸	۲/۷۴	۲/۷۰
خراسان جنوبی		۲/۶۵	۲/۶۰	۲/۵۶	۲/۵۲	۲/۴۸
خراسان رضوی		۳/۷۲	۳/۸۵	۳/۹۶	۴/۰۷	۴/۱۷
خراسان شمالی		۱/۹۹	۱/۹۵	۱/۹۲	۱/۹۰	۱/۸۷
خوزستان		۳/۷۷	۳/۸۰	۳/۸۲	۳/۸۴	۳/۸۷
زنجان		۳/۲۰	۳/۱۴	۳/۰۹	۳/۰۵	۳/۰۱
سمنان		۳/۶۱	۳/۵۵	۳/۴۹	۳/۴۳	۳/۳۸
سیستان و بلوچستان		۲/۱۱	۲/۱۲	۲/۱۲	۲/۱۳	۲/۱۴
فارس		۳/۷۴	۳/۸۰	۳/۸۶	۳/۹۱	۳/۹۶
قزوین		۲/۶۰	۲/۵۶	۲/۵۳	۲/۵۰	۲/۴۷
قم		۳/۸۰	۳/۷۴	۳/۶۸	۳/۶۲	۳/۵۷
کردستان		۲/۶۰	۲/۵۷	۲/۵۴	۲/۵۱	۲/۴۸
کرمان		۳/۱۴	۳/۱۵	۳/۱۷	۳/۱۹	۳/۲۱
کرمانشاه		۳/۰۶	۳/۰۴	۳/۰۳	۳/۰۲	۳/۰۱
کهگیلویه و بویراحمد		۳/۳۳	۳/۲۷	۳/۲۱	۳/۱۶	۳/۱۱
گلستان		۲/۶۱	۲/۵۸	۲/۵۵	۲/۵۳	۲/۵۰
گیلان		۴/۱۳	۴/۱۱	۴/۰۹	۴/۰۸	۴/۰۷
لرستان		۳/۲۱	۳/۱۸	۳/۱۵	۳/۱۲	۳/۰۹
مازندران		۳/۸۴	۳/۸۳	۳/۸۲	۳/۸۱	۳/۸۰
مرکزی		۳/۲۹	۳/۲۴	۳/۲۱	۳/۱۷	۳/۱۴
هرمزگان		۲/۱۷	۲/۱۴	۲/۱۱	۲/۰۸	۲/۰۶
همدان		۳/۰۷	۳/۰۵	۳/۰۲	۳/۰۰	۲/۹۸
یزد		۲/۸۶	۲/۸۱	۲/۷۷	۲/۷۳	۲/۶۹

## مرجع‌ها

- [1] Breau, P. and Ernest, L.R. (1983). Alternative estimator to the current composite estimator, Proceeding of the Section on Survey Research Methods, *American Statistical Association*, 397-402.
- [2] Balan, Jorge (1981). *Why People Move*. The UNESCO Press Printed in France.
- [3] Cantwell, P. (1988). Variance Formulae for the Generalized Composite Estimator under a Balanced One-Level Rotation Plan. *Statistical Research Division Research Report Series*, No. 88-26, Bureau of the Census, Washington, D. C.
- [4] Dejong, Gordon F., Kerry Richter and Pimonpan Isarabhakdi (1996). Gender, values, & intentions to move in rural thailand. *International Migration Review*, 30, 748-770.
- [5] Kumer, L. and Lee, H. (1983). Evaluation of Composite Estimation for the Canadian Labour Force Survey. *Statistics Canada*, 403-408.
- [6] Guilamoto, Christopher Z. (1998). Institutions & migration: Short-term versus long-term moves in ruel West Africa. *Population Studies*, 52, 85-103.
- [7] Longford, N.T. (2006). Sample size calculation for small-area estimation. *Survey Methodology*, 32, 87-96.
- [8] Park, Y., Kim, K., and Choi, J.W. (2007). A balanced multi-level rotation sampling design and its efficient composite estimator. *Journal of Statistical Planning and Inference*, 137, 594-610.

## نعیمه آبی

کارشناس ارشد آمار

تهران، خیابان فاطمی، خیابان رهی‌معیری، پلاک ۱، مرکز آمار ایران.

رایانشانی: n\_a\_abi@yahoo.com

## حمیدرضا نواب‌پور

دکتری آمار

تهران، خیابان شهید بهشتی، خیابان احمد قصیر، دانشکده‌ی اقتصاد دانشگاه علامه طباطبائی، گروه آمار.

رایانشانی: h.navvabpour@src.ac.ir

