مجله علوم آماری، پاییز و زمستان ۱۳۹۹
جلد ۱۴، شماره ۲، ص ۵۳۵ – ۵۴۸
DOI: 10.29252/jss.14.2.535

برآوردهای میانگین جامعه با به‌کارگیری نمونه‌گیری پساطبقه‌ای قضاوتی در نمونه‌گیری طبقه‌بندی‌شده

علی نجفی مجیدآبادی و نادر نعمت‌الهی

گروه آمار، دانشگاه علامه طباطبائی

تاریخ دریافت: ۱۳۹۸/۰۴/۱۲ تاریخ آخرین بازنگری: ۱۳۹۹/۰۳/۰۱

چکیده: روش نمونه‌گیری پساطبقه‌ای قضاوتی روش استفاده از اطلاعات اضافی رتبه‌بندی در نمونه تصادفی ساده به‌منظور افزایش کارایی برآوردگرهای پارامترهای جامعه است. در این مقاله، روش نمونه‌گیری پساطبقه‌ای قضاوتی را در نمونه‌گیری طبقه‌بندی شده به جای نمونه‌گیری تصادفی ساده در داخل طبقه‌ها به‌کار برده و برآوردگرهای جدیدی برای میانگین جامعه ارائه می‌شود. در نهایت با یک مطالعه شبیه‌سازی، برآوردگرهای پیشنهادی با برآوردگر میانگین نمونه‌گیری تصادفی طبقه‌بندی‌شده مورد مقایسه قرار می‌گیرند. نتایج شبیه‌سازی نشان می‌دهد که برآوردگرهای پیشنهادی عمل‌کرد بهتری در بیشتر حالت‌ها نسبت به برآوردگر میانگین نمونه‌گیری تصادفی طبقه‌بندی‌شده دارند.

واژه‌های کلیدی: روش پساطبقه‌ای قضاوتی، کارایی، نمونه‌گیری تصادفی ساده، نمونه‌گیری طبقه‌بندی‌شده.

۱ مقدمه

هدف اصلی نمونه‌گیری ارائهٔ روش‌هایی است که با صرف هزینهٔ اندک بتوان برآوردگرهایی با دقت و کارایی بالا به دست آورد. یکی از این روش‌ها که توسط پژوهش‌گران ارائه شده، روش نمونه‌گیری از مجموعهٔ رتبه‌ای[1] (RSS) است، که اولین بار توسط مک‌اینتایر (۱۹۵۲) پیشنهاد شد. در شرایطی که اندازه‌گیری واحدهای

آدرس الکترونیکی نویسنده مسئول مقاله: نادر نعمت‌الهی، nematollahi@atu.ac.ir
کد موضوع‌بندی ریاضی (۲۰۱۰): 62D05, 62F07, 62F05.

[1]Ranked Set Sampling

جامعه مشکل یا پرهزینه باشد، اما بتوان واحدهای جامعه را به سادگی و با کم‌ترین هزینه رتبه‌بندی کرد، روش نمونه‌گیری از مجموعهٔ رتبه‌ای مطرح می‌شود، که بسیار کاراتر از روش نمونه‌گیری تصادفی ساده با اندازهٔ نمونهٔ برابر است و این افزایش کارایی در نتیجهٔ رتبه‌بندی واحدها که با هزینهٔ کمی صورت می‌گیرد، حاصل می‌شود. هالز و دل (۱۹۶۶) با این روش نمونه‌گیری میانگین محصول علوفه و میانگین ارتفاع درختان یک جنگل را برآورد کردند. تاکاهاسی واکی موتو (۱۹۶۸) میانگین مشاهده‌های حاصل از نمونهٔ مجموعهٔ رتبه‌ای را به‌عنوان برآوردگری نااریب برای میانگین جامعه ارائه کردند و نشان دادند که این برآوردگر کاراتر از برآوردگر میانگین جامعه در نمونه‌گیری تصادفی ساده[1] (SRS) با اندازهٔ نمونه برابر است. همچنین زمان‌زاده و احمدی (۱۳۹۰) بازه اطمینان ناپارامتری با ضریب اطمینان دقیق برای چندک‌های جامعه را بر اساس نمونه‌گیری از مجموعهٔ رتبه‌ای بیان کردند.

برای انتخاب یک نمونهٔ mتایی از جامعه مورد مطالعه بر اساس نمونه‌گیری از مجموعهٔ رتبه‌ای، m نمونه هر یک به اندازهٔ m از جامعه انتخاب می‌شود. سپس هر نمونه به کمک چشم یا متغیر کمکی که با متغیر مورد بررسی همبستگی بالایی دارد به ترتیب افزایشی رتبه‌بندی می‌شود. حال به منظور انتخاب یک نمونهٔ mتایی، از نمونهٔ mتایی اول، واحد دارای کوچک‌ترین رتبه، از نمونهٔ mتایی دوم، واحد داری دومین رتبه و به همین ترتیب از نمونهٔ mتایی mام، واحد دارای بزرگ‌ترین رتبه اندازه‌گیری می‌شوند. به این ترتیب یک نمونهٔ mتایی از جامعهٔ مورد مطالعه به دست می‌آید. توجه شود که این روش نمونه‌گیری منجر به گردآوری $m^۲$ واحد آماری می‌شود اما در نهایت فقط m واحد آماری مورد اندازه‌گیری قرار می‌گیرند. در روش نمونه‌گیری بیان شده وقتی رتبه‌بندی با کمک چشم یا یک متغیر همبسته انجام شود، برای مقدارهای بزرگ m گاهی رتبه‌بندی با خطا همراه است. برای افزایش اندازهٔ نمونه در این روش نمونه‌گیری، افزایش تعداد تکرارها را به جای افزایش m مورد نظر قرار می‌دهند. در حالت کلی، اگر نمونهٔ nتایی مورد نظر به‌صورت $n = n'm$ باشد، می‌توان روش نمونه‌گیری از مجموعهٔ رتبه‌ای را برای m واحد به کار برد و این فرایند را n' بار تکرار کرد. در این صورت نیاز به گردآوری $n'm^۲$ واحد آماری است که از آن‌ها فقط $n'm$ واحد مورد اندازه‌گیری قرار می‌گیرند.

در برخی بررسی‌ها یک نمونهٔ تصادفی سادهٔ nتایی از جامعه گردآوری و تلاش می‌شود که با افزودن اطلاعات اضافی به این نمونه، دقت آن افزایش یابد. برای این منظور روش پساطبقه‌ای قضاوتی[2] (JPS) توسط مک‌ایچرن و همکاران (۲۰۰۴) پیشنهاد شد. این روش در صورتی نسبت به نمونه‌گیری تصادفی ساده کارایی بیش‌تری دارد که رتبه‌بندی واحدها با دقت بالایی انجام شود. فری و ازتورک (۲۰۱۱) نشان

[1] Simple Random Sample
[2] Judgment Post Stratification

دادند وقتی رتبه‌بندی‌ها با استفاده از متغیرهای کمکی انجام می‌شوند برآوردگرهای میانگین در روش پساطبقه‌ای قضاوتی و روش نمونه‌گیری از مجموعه رتبه‌ای تحت تابع زیان توان دوم خطا ناپذیرفتنی هستند. فری و فیمن (۲۰۱۲) برآوردگری بهبودیافته برای میانگین روش پساطبقه‌ای قضاوتی ارائه و نشان دادند کارایی بیش‌تری نسبت به برآوردگرهای موجود دارد. دست برآورده و همکاران (۲۰۱۶) نیز مطالعاتی در زمینهٔ نمونه‌گیری پساطبقه‌ای قضاوتی انجام دادند.

برای انتخاب نمونه پساطبقه‌ای قضاوتی، فرض کنید نمونه تصادفی ساده $V_1, \cdots, V_n$ از جامعه انتخاب و خصوصیت مورد علاقه یعنی $Y_1, \cdots, Y_n$ ثبت شده است. حال به ازای هر Y_i، یک مجموعهٔ تصادفی شامل $m-1$ عضو جامعه $V_{i1}, \cdots, V_{i,m-1}$ در نظر گرفته می‌شود و بدون اندازه‌گیری خصوصیت Y در آن‌ها، رتبهٔ Y_i در میان mتایی $(Y_{i1}, \cdots, Y_{i,m-1}, Y_i)$ به‌صورت قضاوتی تعیین می‌شود. بنابراین، کل نمونهٔ پساطبقه‌ای قضاوتی به‌صورت

$$\begin{array}{cccc} Y_1 & Y_2 & \cdots & Y_n \\ Y_{11} & Y_{21} & \cdots & Y_{n1} \\ \vdots & \vdots & & \vdots \\ Y_{1,m-1} & Y_{2,m-1} & \cdots & Y_{n,m-1} \end{array} \tag{1}$$

است، که در آن Y_{ir} نمونهٔ کمکی rام، برای رتبه‌بندی Y_i است. برای به دست آوردن نمونهٔ پساطبقه‌ای قضاوتی، هر ستون که دارای m واحد است به‌صورت افزایشی (به وسیله چشم یا با استفاده از یک متغیر کمکی) مرتب می‌شود و رتبهٔ Y_i را بین $(Y_{i1}, \ldots, Y_{i,m-1}, Y_i)$ برای $i = 1, \cdots, n$ تعیین کرده و با R_i نشان داده می‌شود. بدین ترتیب نمونهٔ پساطبقه‌ای قضاوتی $(Y_1, R_1), \ldots, (Y_n, R_n)$ حاصل می‌شود. توجه شود که R_iها دارای توزیع یکنواخت گسسته در مجموعهٔ $\{1, \cdots, m\}$ هستند. اگر پیدا کردن رتبهٔ Y_iها بدون خطا انجام شود، آنگاه $Y_i | R_i = r$ دارای توزیع rامین آماره ترتیبی در نمونه تصادفی ساده به اندازه m است. شیوهٔ معمول در طبقه‌بندی قضاوتی این نمونه به این صورت است که Y_iهایی که دارای رتبهٔ یکسان هستند را در یک طبقه قرار می‌دهند بنابراین m طبقه خواهند شد، که لزوماً همهٔ آن‌ها ناتهی نیستند، به عبارت دیگر ممکن است در یک یا چند طبقه مشاهده‌ای وجود نداشته باشد. توجه شود در روش JPS از همان نمونهٔ تصادفی سادهٔ اولیه $(Y_1, \cdots, Y_n)$ استفاده شده است و استفاده از رتبه‌ها به منظور افزایش دقت این نمونه است. این روش در بسیاری از شرایط نسبت به نمونه‌گیری تصادفی ساده از کارایی بیشتری برخوردار است (مک‌ایچرن و همکاران، ۲۰۰۴). چون رتبه‌بندی‌ها در عمل

به‌صورت قضاوت شخصی یا به کمک یک متغیر کمکی که با متغیر اصلی همبستگی دارد انجام می‌شود، ممکن است رتبه‌بندی با خطا همراه شود. در این صورت نمونهٔ پساطبقه‌ای قضاوتی برای $i = ۱, \cdots, n$ و $r = ۱, \cdots, m$ به‌صورت زوج‌های $(Y_{[r]i}, R_i = r)$ خواهد بود، که در آن $Y_{[r]i}$ واحد iام نمونه است که در پساطبقهٔ rام قرار می‌گیرد، یعنی دارای رتبهٔ قضاوتی r است.

نمونه‌گیری طبقه‌بندی‌شده از روش‌های نمونه‌گیری متداولی است که با تقسیم واحدهای جامعه به دو یا چند طبقه و انتخاب نمونه‌ها به‌طور مستقل در هر یک از طبقه‌ها انجام می‌پذیرد. نمونه‌گیری طبقه‌بندی‌شده امکان استفاده از طرح‌ها و روش‌های مختلف برای انتخاب نمونه‌ها در طبقه‌های مختلف فراهم می‌آورد. هنگام استفاده از نمونه‌گیری طبقه‌بندی‌شده، در انتخاب نمونه در طبقه‌ها، می‌توان روش‌های SRS ، RSS و JPS را به کار برد. در شرایطی که خواسته شود از نمونهٔ SRS در طبقات استفاده کرده و در عین حال دقت نمونه‌گیری طبقه‌بندی‌شده افزایش داده شود، می‌توان از روش JPS به جای SRS در انتخاب نمونه در طبقات استفاده کرد. روش متداول در نمونه‌گیری طبقه‌بندی‌شده استفاده از روش SRS در انتخاب واحدهای نمونه‌ای طبقات است که با نماد[1] (STSRS) نمایش داده می‌شود.

در این مقاله برای نخستین‌بار در برآورد میانگین جامعه از روش نمونه‌گیری طبقه‌بندی‌شده که انتخاب واحدهای نمونه در طبقات با استفاده از روش پساطبقه‌ای قضاوتی است، استفاده کرده و سه برآوردگر جدید برای برآورد میانگین جامعه ارائه می‌شود. این روش نمونه‌گیری را روش نمونه‌گیری پساطبقه‌ای قضاوتی طبقه‌بندی‌شده[2] (STJPS) نامیده و با روش STSRS مقایسه می‌شود. به این منظور، در بخش ۲ برآوردگرهای میانگین جامعه با استفاده از روش نمونه‌گیری پساطبقه‌ای قضاوتی و نمونه‌گیری طبقه‌بندی‌شده بیان می‌شود. در بخش ۳ برآوردگر میانگین جامعه با استفاده از روش نمونه‌گیری پساطبقه‌ای قضاوتی طبقه‌بندی‌شده ارائه می‌شود و دو برآودرگر جدید برای برآورد میانگین جامعه براساس نمونه‌گیری پساطبقه‌ای قضاوتی طبقه‌بندی‌شده ارائه می‌شود. در بخش ۵ با استفاده از شبیه‌سازی کارایی برآوردگرهای مطرح شده را با یکدیگر مقایسه کرده و نشان داده می‌شود که برآوردگر میانگین در نمونه‌گیری پساطبقه‌ای قضاوتی طبقه‌بندی‌شده کارایی بیش‌تری نسبت به نمونه‌گیری تصادفی طبقه‌بندی‌شده متداول دارد.

[1] Stratified Simple Random Sampling
[2] Stratified Judgment Post Stratification

۲ برآوردگرهای میانگین با نمونه‌گیری‌های پساطبقه‌ای قضاوتی و تصادفی طبقه‌بندی‌شده

فرض کنید برای $i=۱,\cdots,n$ و m, $r=۱,\cdots,m$، $(Y_{[r]i},R_i)$ یک نمونه پساطبقه‌ای قضاوتی باشد. برای تعریف برآوردگر میانگین JPS استاندارد، فرض کنید n_r تعداد اعضای مجموعهٔ $\{i;R_i=r\}$، یعنی تعداد مشاهده‌ها با رتبهٔ قضاوتی r باشد. در این صورت $(n_۱,\cdots,n_m)$ دارای توزیع چندجمله‌ای با تعداد آزمایش n و احتمال‌های برابر $\frac{۱}{m}$ است. بنابراین، به ازای برخی مقدارهای r ممکن است $n_r=۰$ شود. با قرار دادن

$$I_r=\begin{cases} ۱ & n_r>۰ \\ ۰ & n_r=۰\end{cases}, \quad I_{ri}=\begin{cases} ۱ & R_i=r \\ ۰ & o.w\end{cases}, \quad J_r=\begin{cases} \frac{۱}{n_r} & n_r>۰ \\ ۰ & o.w.\end{cases},$$

$h=\sum_{r=۱}^{m}I_r$ و $n_r=\sum_{i=۱}^{n}I_{ri}$، دست برآورده و همکاران (۲۰۱۶) برآوردگر میانگین جامعه براساس نمونه‌گیری پساطبقه‌ای قضاوتی را به‌صورت

$$\hat{\mu}_{JPS}=\frac{۱}{h}\sum_{r=۱}^{m}\bar{Y}_{[r].}I_r\,, \tag{۲}$$

ارائه کردند، که در آن

$$\bar{Y}_{[r].}=\begin{cases} ۰ & n_r=۰ \\ \frac{۱}{n_r}\sum_{i=۱}^{n}Y_{[r]i}I_{ri} & n_r>۰\end{cases}.$$

قضیه ۱. *(دست برآورده و همکاران، ۲۰۱۶) اگر برای r، $i=۱,\ldots,r$، $\mu_{[r]}=\mathrm{E}(Y_{[r]i}|R=r)$ و $\sigma^۲_{[r]}=\mathrm{Var}(Y_{[r]i}|R=r)$، به ترتیب میانگین و واریانس پساطبقهٔ rام باشند، آنگاه میانگین و*

واریانس برآوردگر میانگین پساطبقه‌ای قضاوتی (۲) *بترتیب عبارتند از:*

$$\mathrm{E}(\hat{\mu}_{JPS}) = \frac{1}{m}\sum_{r=1}^{m}\mu_{[r]} = \mu,$$
$$\mathrm{Var}(\hat{\mu}_{JPS}) = \mathrm{E}(\frac{J_1 I_1^2}{h^2})\sum_{r=1}^{m}\sigma_{[r]}^2 + \mathrm{Var}(\frac{I_1}{h})\frac{m}{m-1}\sum_{r=1}^{m}(\mu_{[r]}-\mu)^2.$$

برای معرفی برآوردگر میانگین در نمونه‌گیری طبقه‌بندی‌شده، فرض کنید جامعهٔ مورد بررسی شامل N واحد نمونه‌گیری باشد که به L طبقهٔ مجزا تقسیم شده باشند به‌طوری که d امین طبقه $(d = 1, \dots, L)$ دارای N_d واحد باشد و $\sum_{d=1}^{L} N_d = N$. همچنین فرض کنید Y_{jd} بر مقدار صفت مورد نظر برای jامین واحد $(j = 1, \dots, N_d)$ در dامین طبقه دلالت نماید و μ_d و σ_d^2 به ترتیب میانگین و واریانس مقادیر صفت مورد نظر در dامین طبقه باشند. برای برآورد میانگین جامعه، μ با استفاده از نمونه‌گیری طبقه‌بندی‌شده، نمونه‌ای به اندازهٔ n_d از طبقهٔ dام انتخاب می‌شود و این نمونه‌ها در طبقه‌های مختلف از یکدیگر مستقلند. در این صورت $\sum_{d=1}^{L} n_d = n$ اندازهٔ نمونه در کل طبقه‌ها است. فرض کنید $W_d = \frac{N_d}{N}$ وزن طبقهٔ dام باشد که $\sum_{L=1}^{d} W_d\mu_d = \mu$. صرف‌نظر از نوع روش نمونه‌گیری داخل هر یک از طبقه‌ها اگر $\hat{\mu}_d$ برآوردی نااریب از μ_d باشد، آن‌گاه $\hat{\mu}_{ST} = \sum_{d=1}^{L} W_d\hat{\mu}_d$، برآوردی نااریب از میانگین جامعه μ است (ککران، ۱۹۷۷). به عبارت دیگر میانگین موزون برآورد نااریب میانگین طبقه‌ها، برآوردی نااریب از میانگین جامعه است. با توجه به استقلال طبقه‌ها، واریانس $\hat{\mu}_{ST}$ به‌صورت

$$\mathrm{Var}(\hat{\mu}_{ST}) = \sum_{d=1}^{L} W_d^2 \mathrm{Var}(\hat{\mu}_d)\,.$$

به دست می‌آید، که نشان می‌دهد وقتی طبقه‌بندی واحدها در جامعه به گونه‌ای باشد که پراکندگی مقدار صفت مورد نظر در داخل طبقات کم باشد، استفاده از نمونه‌گیری طبقه‌بندی‌شده برآوردی با دقت مطلوب برای میانگین جامعه ارائه می‌دهد. اگر برای انتخاب نمونه‌ها در داخل طبقه‌ها از نمونه‌گیری تصادفی ساده استفاده شود، روش نمونه‌گیری حاصل را نمونه‌گیری تصادفی طبقه‌بندی‌شده می‌نامند و برآوردگر نااریب میانگین جامعه به صورت $\hat{\mu}_{STSRS} = \sum_{d=1}^{L}\frac{W_d}{n_d}\sum_{j=1}^{n_d} Y_{jd}$، به دست می‌آید. در صورتی که از نمونه‌گیری

تصادفی ساده بدون جایگذاری در داخل طبقه‌ها استفاده شود واریانس $\hat{\mu}_{STSRS}$ به‌صورت زیر خواهد بود

$$\mathrm{Var}(\hat{\mu}_{STSRS}) = \sum_{d=1}^{L} W_d^2 (\frac{N_d - n_d}{N_d - 1}) \frac{\sigma_d^2}{n_d} .$$

۳ نمونه‌گیری پساطبقه‌ای قضاوتی طبقه‌بندی‌شده

فرض کنید برای انتخاب نمونه‌ها در داخل هر یک از طبقه‌های جامعه از نمونه‌گیری پساطبقه‌ای قضاوتی استفاده شود. به عبارت دیگر فرض کنید نمونه‌گیری به روش نمونه‌گیری پساطبقه‌ای قضاوتی طبقه‌بندی‌شده (STJPS) انجام شود. مطابق نماد گذاری بخش ۲، فرض کنید m_d به تعداد پساطبقه‌ها و n_d به اندازهٔ نمونه در dامین طبقه دلالت کند و $Y_{d[r]i}$ برای $i = 1, \ldots, n_d$ و $r = 1, \ldots, m_d$ مقدار صفت مورد نظر برای عناصری از پساطبقهٔ iام با رتبهٔ قضاوتی rام از طبقهٔ dام باشد. همچنین فرض کنید n_{dr} تعداد $\{i; R_{di} = r\}$ باشد، یعنی تعداد مشاهده‌ها با رتبهٔ قضاوتی r در طبقهٔ d باشد. در این صورت $(n_{d1}, \cdots, n_{dm_d})$ دارای توزیع چندجمله‌ای با تعداد آزمایش n_d و احتمال‌های برابر $\frac{1}{m_d}$ است. بنابراین، به ازای برخی مقدارهای r ممکن است $n_{dr} = 0$ شود. با قرار دادن

$$I_{dr} = \begin{cases} 1 & n_{dr} > 0 \\ 0 & n_{dr} = 0 \end{cases}, \quad I_{dri} = \begin{cases} 1 & R_{di} = r \\ 0 & o.w \end{cases}, \quad J_{dr} = \begin{cases} \frac{1}{n_{dr}} & n_{dr} > 0 \\ 0 & o.w. \end{cases}$$

$h_d = \sum_{r=1}^{m_d} I_{dr}$ و $n_{dr} = \sum_{i=1}^{n_d} I_{dri}$، برآوردگر میانگین جامعه بر اساس نمونه‌گیری پساطبقه‌ای قضاوتی طبقه‌بندی‌شده $\hat{\mu}_{STJPS} = \sum_{d=1}^{L} W_d \hat{\mu}_{dJPS}$، است، که در آن $\hat{\mu}_{dJPS} = \frac{1}{h_d} \sum_{r=1}^{m_d} \bar{Y}_{d[r].} I_{dr}$ و

$$\bar{Y}_{d[r].} = \begin{cases} 0 & n_{dr} = 0 \\ \frac{1}{n_{dr}} \sum_{i=1}^{n_d} Y_{d[r]i} I_{dri} & n_{dr} > 0 \end{cases} .$$

قضیه ۲. برآوردگر میانگین پساطبقه‌ای قضاوتی طبقه‌بندی‌شده برای μ نااریب با واریانس

$$\begin{aligned}\mathrm{Var}(\hat{\mu}_{STJPS}) &= \sum_{d=1}^{L} W_d^2 \{\mathrm{E}(\frac{J_{d1} I_{d1}^2}{h_d^2}) \sum_{r=1}^{m_d} \sigma_{d[r]}^2 \\ &+ \mathrm{Var}(\frac{I_{d1}}{h_d}) \frac{m_d}{m_d - 1} \sum_{r=1}^{m_d} (\mu_{d[r]} - \mu_d)^2\},\end{aligned}$$

است، که در آن $\mu_{d[r]} = \mathrm{E}(Y_{d[r]i} | R = r)$ و $\sigma_{d[r]}^2 = \mathrm{Var}(Y_{d[r]i} | R = r)$، $i = 1, \ldots, n_d$ و μ_d، $\mu_{d[r]}$ و $\sigma_{d[r]}^2$ به ترتیب میانگین طبقهٔ dام، واریانس rامین آماره ترتیبی در طبقهٔ dام و میانگین rامین آمارهٔ ترتیبی در طبقهٔ dام هستند.

برهان: با استفاده از قضیهٔ ۱، اثبات سرراست است.

۴ برآوردگرهای جدید با نمونه‌گیری پساطبقه‌ای قضاوتی طبقه‌بندی‌شده

لم ۱. *(وانگ و همکاران، ۲۰۰۸) فرض کنید $Y_{[1]}, \ldots, Y_{[m]}$ به‌طور تصادفی مرتب شده‌اند. برای هر y، نابرابری $F_{[1]}(y) \geq \ldots \geq F_{[m]}(y)$ برقرار است، که در آن $F_{[r]}(y)$ تابع توزیع تجمعی درون پساطبقه rام تعریف شده است.*

اگر دو توزیع به‌طور تصادفی بصورت $(F(x) \geq G(x))$ مرتب شده باشند، آن‌گاه برای هر تابع غیر نزولی Φ نابرابری $E_G(\Phi(X)) \geq E_F(\Phi(X))$ برقرار است. لذا وانگ و همکاران (۲۰۰۸) با قرار دادن $\Phi(X) = X$ نشان دادند $\mu_{[1]} \leq \ldots \leq \mu_{[m]}$. اما چون میانگین‌های نمونه $\bar{Y}_{[r]}$ ممکن است در عمل این نابرابری را نقض کنند، آنها صورت همنوا شده آن، یعنی $\bar{Y}_{[r]}^{iso}$، را در $\hat{\mu} = \frac{1}{h} \sum_{r=1}^{m} \bar{Y}_{[r]} I_r$ جایگزین کرده و برآوردگر میانگین هم‌توان[1] در روش نمونه‌گیری پساطبقه‌ای قضاوتی را به‌صورت $\hat{\mu}_w = \frac{1}{h} \sum_{r=1}^{m} \bar{Y}_{[r]}^{iso} I_r$، ارائه کردند، که در آن $\bar{Y}_{[r]}^{iso} = \max_{t \leq r} \min_{s \geq r} \sum_{g=t}^{s} \frac{n_g \bar{Y}_{[g]}}{\sum_{u=t}^{s} n_u}$، به ازای $n_r > 0$ است. برای حالتی که $n_r = 0$ باشد، پیشنهاد دادند $\bar{Y}_{[r]}^{iso}$ با میانگین آمیخته یک یا دو رده کناری آن برآورد شود. به دلیل پیچیدگی محاسبات اثباتی از نااریبی آن ارائه نکردند. قاسمی و همکاران

---
[1] Isotonic

(۱۳۹۵) با توجه به ایده‌ای که بیان شد به جای $\{(Y_i, R_i), i = ۱, \dots, n\}$ از نمونهٔ مرتب‌شدهٔ آن استفاده کردند. آن‌ها $\{(Y_{(i)}, R_{[i]}), i = ۱, \dots, n\}$ را نمونهٔ مرتب‌شدهٔ طرح JPS در نظر گرفتند که در آن $\mathrm{Y}_{(۱)} \leq \dots \leq \mathrm{Y}_{(m)}$ آماره‌های مرتب‌شدهٔ $\mathrm{Y}_۱, \dots, \mathrm{Y}_m$ هستند و $\mathrm{R}_{[j]}$ متغیر R_j متناظر با $Y_{(j)}$ است. در این حالت ممکن است رابطهٔ $\mathrm{R}_{[۱]} \leq \dots \leq \mathrm{R}_{[m]}$ برقرار نباشد. ایدهٔ معرفی برآوردگر جدید توسط قاسمی و همکاران (۱۳۹۵) این است که از تصادفی قرار گرفتن هر یک از مشاهدات درون طبقات چشم‌پوشی کرده و مشاهدات را به‌صورت $\{(Y_{(i)}, R_{(i)}), i = ۱, \dots, n\}$ در نظر گرفته‌اند، که در آن $Y_{(۱)} \leq \dots \leq Y_{(n)}$ و $R_{(۱)} \leq \dots \leq R_{(n)}$ است. در نتیجه برآوردگر پیشنهادی آنها به‌صورت $\hat{\mu}_g = \frac{۱}{m}\sum_{r=۱}^{m} \bar{Y}^*_{(r)}$ است، که در آن $\bar{Y}^*_{(r)} = \frac{۱}{n_r}\sum_{i=۱}^{n} Y_{(i)} I_{ri}$ و

$$I_{ri} = \begin{cases} ۱ & R_{[i]} = r \\ ۰ & o.w \end{cases}, \quad n_r = \sum_{i=۱}^{n} I_{ri}.$$

اگر $n_r > ۰$ باشد و چنانچه $n_r = ۰$ باشد، در این صورت $\bar{Y}^*_{(r)}$ به‌صورت میانگین آمیخته رده یا رده‌های مجاور آن تعریف می‌شود. از آنجا که محاسبهٔ میانگین توان دوم خطا و اریبی برآوردگر $\hat{\mu}_g$ به آسانی امکان‌پذیر نیست در نتیجه با استفاده از یک مطالعه شبیه‌سازی نشان دادند که این برآوردگر نسبت به برآوردگر میانگین استاندارد پساطبقه‌ای قضاوتی بهتر عمل می‌کند.

فرض کنید برای انتخاب نمونه‌ها در داخل هر یک از طبقه‌های جامعه از نمونه‌گیری پساطبقه‌ای قضاوتی استفاده شود. مطابق نمادگذاری بخش ۲ فرض کنید m_d به تعداد پساطبقه‌ها و n_d به اندازهٔ نمونه در dامین طبقه دلالت نماید. در این صورت برآوردگر جدید با استفاده از برآوردگر پیشنهادی وانگ و همکاران (۲۰۰۸) به‌صورت $\hat{\mu}_{new_۱} = \sum_{d=۱}^{L} W_d \hat{\mu}_{w_d}$، معرفی می‌شود، که در آن $\hat{\mu}_{w_d} = \frac{۱}{h_d}\sum_{r=۱}^{m_d} \bar{Y}^{iso}_{d[r]} I_r$ است، اندیس d در $\hat{\mu}_{w_d}$ یعنی برآورد میانگین در طبقهٔ dام است.

فرض کنید برای انتخاب نمونه‌ها در داخل هر یک از طبقه‌های جامعه از نمونه‌گیری پساطبقه‌ای قضاوتی استفاده شود. مطابق نمادگذاری بخش ۲ فرض کنید m_d به تعداد پساطبقه‌ها و n_d به اندازه نمونه در dامین طبقه دلالت نماید. در این صورت برآوردگر جدید به‌صورت $\hat{\mu}_{new_۲} = \sum_{d=۱}^{L} W_d \hat{\mu}_{g_d}$ معرفی می‌شود، که در آن $\hat{\mu}_{dg}$ همان برآوردگر میانگین پساطبقه‌ای قضاوتی در طبقه d است. با توجه به این که قاسمی و همکاران (۱۳۹۵) با شبیه‌سازی نشان داده‌اند که $\hat{\mu}_d$ عملکرد بهتری نسبت به $\hat{\mu}_{JPS}$

دارد، لذا به‌صورت شهودی انتظار می‌رود که برآوردگر پیشنهادی نسبت به برآوردگر میانگین پساطبقه‌ای قضاوتی طبقه‌بندی‌شده و همچنین برآوردگر میانگین تصادفی طبقه‌بندی‌شده عملکرد بهتری داشته باشد. چون محاسبه میانگین توان دوم خطا و اریبی این برآوردگر به آسانی امکان‌پذیر نیست برای مقایسه کارایی برآوردگرها از شبیه‌سازی مونت‌کارلو استفاده می‌شود.

۵ مطالعه شبیه‌سازی

برای بررسی کارایی سه برآوردگر ارائه شده بر اساس نمونه‌گیری پساطبقه‌ای قضاوتی طبقه‌بندی‌شده نسبت به نمونه‌گیری تصادفی طبقه‌بندی‌شده بدون جایگذاری، از داده‌های گل‌خانه‌های ایران استفاده می‌شود که در سرشماری عمومی کشاورزی توسط مرکز آمار ایران (۱۳۸۲) جمع‌آوری شده‌اند. برای سهولت بیشتر و جلوگیری از تأثیر اندازهٔ جامعهٔ بزرگ بر نتایج شبیه‌سازی، در این مطالعه از اطلاعات استان تهران استفاده می‌شود. فرض کنید قرار باشد میانگین ارزش محصولات گل‌خانه‌های استان تهران برآورد شود. بدین منظور از برآوردگرهای $\hat{\mu}_{STJPS}$، $\hat{\mu}_{new_۲}$، $\hat{\mu}_{new_۱}$ و $\hat{\mu}_{STSRS}$ استفاده می‌شوند و دقت چهار برآوردگر با تکرار نمونه‌گیری محاسبه می‌شوند. در محاسبهٔ برآوردها نیاز به رتبه‌بندی واحدهای نمونه است. برای رتبه‌بندی ناقص از مدل دل و کلاتر (۱۹۷۲) استفاده می‌شود. در این مدل برای رتبه‌بندی از متغیر کمکی

$$\mathrm{X} = \rho(\frac{Y-\mu}{\sigma}) + \sqrt{۱-\rho^۲}Z\,. \qquad (۳)$$

استفاده می‌شود، که در آن μ و σ میانگین و انحراف معیار متغیر Y (در طبقه مربوطه)، ρ ضریب همبستگی X و Y، و Z متغیر تصادفی نرمال استاندارد است. هر چه مقدار ρ به یک نزدیک‌تر باشد رتبه‌بندی دقیق‌تر خواهد بود. لذا در شبیه‌سازی، دو مقدار $\rho = ۰/۹۹$ برای رتبه‌بندی با خطای اندک و $\rho = ۰/۵۰$ برای رتبه‌بندی ضعیف در نظر گرفته می‌شود. با تولید داده Z از توزیع نرمال استاندارد و قرار دادن در (۳) مقادیر متغیر کمکی X به دست می‌آید و رتبه‌بندی براساس مقادیر X انجام می‌شود. با محدود کردن اطلاعات موجود به گل‌خانه‌های استان تهران $N = ۲۷۵$ واحد در جامعه وجود دارد. برای طبقه‌بندی این واحدها از روش فراوانی ریشه تجمعی (دالینوس ، ۲۰۱۶) استفاده می‌شود.

الگوریتم ۱. الگوریتم طبقه‌بندی فراوانی ریشهٔ تجمعی :

گام ۱- مرتب کردن واحدهای جامعه براساس متغیر X به‌صورت صعودی،

گام ۲- گروه‌بندی X به J گروه مجزا،
گام ۳- تعیین فراوانی هر یک از گروه‌ها، f_i $(i = ۱, \dots, J)$،
گام ۴ - محاسبهٔ ریشهٔ توان دوم فراوانی هر گروه، $\sqrt{f_i}$ ،
گام ۵- محاسبهٔ جمع تجمعی ریشهٔ فراوانی‌ها $\sum_{i=۱}^{J} \sqrt{f_i}$ ،
گام ۶ - تقسیم مجموع ریشهٔ توان دوم فراوانی‌ها بر تعداد طبقه‌های فرض شده $Q = \frac{۱}{L} \sum_{i=۱}^{J} \sqrt{f_i}$،
گام ۷ - تعیین مرز بالای هر یک از طبقه‌ها برای مقادیر X به‌صورت $Q, ۲Q, \dots, (L-۱)Q, LQ$.

با الگوریتم ۱ واحدهای موجود در جامعه به $L = ۲$ (با اندازه طبقه‌های $N_۱ = ۱۹۱$ و $N_۲ = ۸۴$ واحد) و $L = ۳$ طبقه (با اندازه طبقه‌های $N_۱ = ۱۹۱$، $N_۲ = ۴۷$ و $N_۳ = ۳۷$ واحد) تقسیم‌بندی شد. برای برآورد میانگین جامعه، در هر تکرار نمونه‌گیری میانگین نمونه‌های حاصل از چهار روش محاسبه می‌شود. انتخاب نمونه‌ها و محاسبهٔ برآوردگرها با نرم‌افزار R با ۵۰۰۰۰ بار تکرار انجام می‌گیرد. مقدار دقیق میانگین جامعه که از چارچوب قابل محاسبه است برابر $\mu = ۲۹/۸۲$ است. بنابراین می‌توان برآوردی از میانگین توان دوم خطای هر یک از برآوردگرهای $\hat{\mu}_k$ را به‌صورت

$$M\hat{S}E(\hat{\mu}_k) = \frac{۱}{۵۰۰۰۰} \sum_{d=۱}^{۵۰۰۰۰} (\hat{\mu}_{dk} - \mu)^۲, \quad k = ۱, ۲, ۳, ۴,$$

به دست آورد، که در آن $\hat{\mu}_k$ برای $k = ۱, ۲, ۳, ۴$ به ترتیب برآوردگرهای $\hat{\mu}_{new_۱}$، $\hat{\mu}_{new_۲}$، $\hat{\mu}_{STSRS}$ و $\hat{\mu}_{STSRS}$ هستند و همچنین $\hat{\mu}_{d۱}$، $\hat{\mu}_{d۲}$، $\hat{\mu}_{d۳}$ و $\hat{\mu}_{d۴}$ به ترتیب برآوردگرهای $\hat{\mu}_{new_۱}$، $\hat{\mu}_{new_۲}$، $\hat{\mu}_{STSRS}$ و $\hat{\mu}_{STSRS}$ در d امین تکرار نمونه‌گیری هستند. برای مقایسهٔ کارایی هر چهار برآوردگر از کارایی نسبی برآوردگرهای حاصل استفاده می‌شود. کارایی نسبی $\hat{\mu}_۱$ نسبت به $\hat{\mu}_۲$ به‌صورت

$$RE(\hat{\mu}_۱, \hat{\mu}_۲) = \frac{M\hat{S}E(\hat{\mu}_۲)}{M\hat{S}E(\hat{\mu}_۱)}. \qquad (۴)$$

تعریف می‌شود. کارایی نسبی برآوردگرهای $\hat{\mu}_{new_۱}$، $\hat{\mu}_{new_۲}$ و $\hat{\mu}_{STJPS}$ نسبت به برآوردگر $\hat{\mu}_{STSRS}$ برای $n_d = ۵, ۱۰$ و $m_d = ۲, ۴, ۶, ۸, ۱۰$ در جدول‌های ۱ و ۲ ارائه شده است. همان‌طور که ملاحظه می‌شود وقتی رتبه‌بندی دارای خطای اندک است دقت برآورد حاصل از نمونه‌گیری STJPS در بیشتر موارد، بیش از دقت برآورد حاصل از نمونه‌گیری تصادفی طبقه‌بندی‌شده است. مقایسه مقادیر کارایی نسبی، نشان

جدول ۱. کارایی $\hat{\mu}_{STJPS}$، $\hat{\mu}_{new_1}$ و $\hat{\mu}_{new_2}$ نسبت به $\hat{\mu}_{STSRS}$ برای $\rho = ۰٫۹۹$

	$L = ۲$						$L = ۳$					
	$n_d = ۵$			$n_d = ۱۰$			$n_d = ۵$			$n_d = ۱۰$		
m_d	$\hat{\mu}_{STJPS}$	$\hat{\mu}_{new_1}$	$\hat{\mu}_{new_2}$	$\hat{\mu}_{STJPS}$	$\hat{\mu}_{new_1}$	$\hat{\mu}_{new_2}$	$\hat{\mu}_{STJPS}$	$\hat{\mu}_{new_1}$	$\hat{\mu}_{new_2}$	$\hat{\mu}_{STJPS}$	$\hat{\mu}_{new_1}$	$\hat{\mu}_{new_2}$
۲	۱/۰۰	۱/۰۱	۱/۰۳	۱/۰۶	۱/۰۷	۱/۰۸	۱/۰۰	۱/۰۳	۱/۰۶	۱/۰۵	۱/۰۶	۱/۰۷
۴	۱/۰۶	۱/۳۸	۱/۴۸	۱/۱۷	۱/۳۸	۱/۴۷	۱/۰۹	۱/۴۲	۱/۵۴	۱/۲۰	۱/۴۵	۱/۵۶
۶	۱/۰۸	۱/۷۳	۱/۸۵	۱/۲۰	۱/۷۴	۱/۸۸	۱/۰۶	۱/۷۲	۱/۸۷	۱/۲۰	۱/۸۶	۲/۰۲
۸	۱/۰۶	۱/۹۶	۲/۰۸	۱/۲۰	۲/۱۱	۲/۲۷	۱/۰۵	۱/۹۲	۲/۰۸	۱/۱۷	۲/۲۶	۲/۴۷
۱۰	۱/۰۴	۲/۰۵	۲/۱۹	۱/۱۶	۲/۴۰	۲/۵۹	۱/۰۴	۱/۹۴	۲/۰۹	۱/۱۴	۲/۵۷	۲/۸۰

جدول ۲. کارایی $\hat{\mu}_{STJPS}$، $\hat{\mu}_{new_1}$ و $\hat{\mu}_{new_2}$ نسبت به $\hat{\mu}_{STSRS}$ برای $\rho = ۰٫۵۰$

	$L = ۲$						$L = ۳$					
	$n_d = ۵$			$n_d = ۱۰$			$n_d = ۵$			$n_d = ۱۰$		
m_d	$\hat{\mu}_{STJPS}$	$\hat{\mu}_{new_1}$	$\hat{\mu}_{new_2}$	$\hat{\mu}_{STJPS}$	$\hat{\mu}_{new_1}$	$\hat{\mu}_{new_2}$	$\hat{\mu}_{STJPS}$	$\hat{\mu}_{new_1}$	$\hat{\mu}_{new_2}$	$\hat{\mu}_{STJPS}$	$\hat{\mu}_{new_1}$	$\hat{\mu}_{new_2}$
۲	۰/۸۴	۰/۸۸	۰/۸۴	۰/۸۵	۰/۸۸	۰/۷۷	۰/۸۴	۰/۸۹	۰/۸۸	۰/۸۲	۰/۸۵	۰/۷۴
۴	۰/۸۹	۰/۹۸	۱/۰۰	۰/۷۹	۰/۸۹	۰/۸۲	۰/۸۸	۰/۹۶	۰/۹۷	۰/۷۶	۰/۸۷	۰/۷۹
۶	۰/۹۱	۱/۰۳	۱/۰۵	۰/۸۱	۰/۹۳	۰/۸۶	۰/۸۹	۰/۹۸	۱/۰۰	۰/۷۶	۰/۸۸	۰/۸۲
۸	۰/۹۲	۱/۰۶	۱/۰۹	۰/۸۳	۰/۹۶	۰/۹۰	۰/۹۱	۰/۹۹	۱/۰۲	۰/۸۰	۰/۹۲	۰/۸۷
۱۰	۰/۹۲	۱/۰۸	۱/۱۱	۰/۸۴	۰/۹۸	۰/۹۳	۰/۹۱	۱/۰۰	۱/۰۳	۰/۸۱	۰/۹۴	۰/۸۸

می‌دهد با افزایش m_d کارایی $\hat{\mu}_{STJPS}$ همواره افزایشی نیست. کارایی برآوردگرهای $\hat{\mu}_{new_1}$ و $\hat{\mu}_{new_2}$ نسبت به برآوردگرهای $\hat{\mu}_{STJPS}$ و $\hat{\mu}_{STSRS}$ زمانی که رتبه‌بندی با خطای اندک است بیشتر است. همچنین با افزایش m_d کارایی بروادرگرهای $\hat{\mu}_{new_1}$ و $\hat{\mu}_{new_2}$ افزایش می‌یابد. جدول ۳ نشان می‌دهد

جدول ۳. کارایی $\hat{\mu}_{new_2}$ نسبت به $\hat{\mu}_{new_1}$

	$\rho = ۰٫۹۹$				$\rho = ۰٫۵۰$			
	$L = ۲$		$L = ۳$		$L = ۲$		$L = ۳$	
m_d	$n_d = ۵$	$n_d = ۱۰$	$n_d = ۵$	$n_d = ۱۰$	$n_d = ۵$	$n_d = ۱۰$	$n_d = ۵$	$n_d = ۱۰$
۲	۱/۰۱۹	۱/۰۰۹	۱/۰۲۹	۱/۰۰۹	۰/۹۵	۰/۸۷۵	۰/۹۸۸	۰/۸۷۰
۴	۱/۰۷۲	۱/۰۶۵	۱/۰۸۴	۱/۰۷۶	۱/۰۲	۰/۹۲۱	۱/۰۱۰	۰/۹۰
۶	۱/۰۶۹	۱/۰۸۰	۱/۰۷۸	۱/۰۸۶	۱/۰۱۹	۰/۹۲۴	۱/۰۲۰	۰/۹۳۱
۸	۱/۰۶۱	۱/۰۷۵	۱/۰۸۳	۱/۰۹۲	۱/۰۲۸	۰/۹۳۷	۱/۰۳۰	۰/۹۴۵
۱۰	۱/۰۶۸	۱/۰۷۹	۱/۰۷۷	۱/۰۸۹	۱/۰۲۷	۰/۹۴۸	۱/۰۳	۰/۹۳۶

که با افزایش ρ مقدار کارایی $\hat{\mu}_{new_2}$ نسبت به $\hat{\mu}_{new_1}$ افزایشی است. همچنین با افزایش m_d و L بهبود کارایی همواره افزایشی نیست. وقتی که رتبه‌بندی دارای خطای اندک است $\hat{\mu}_{new_2}$ برآوردگر بهتری نسبت به $\hat{\mu}_{new_1}$ است. با توجه به مقایسه کارایی‌ها در مواردی که برای رتبه‌بندی واحدها از یک متغیر کمکی که دارای همبستگی بالایی با متغیر اصلی است استفاده می‌شود، برآوردگرهای پساطبقه‌ای قضاوتی طبقه‌بندی‌شده به‌جای استفاده از برآوردگر طبقه‌ای متداول پیشنهاد می‌شود.

بحث و نتیجه‌گیری

معالعه شبیه‌سازی در بخش ۵ که بر اساس داده‌های واقعی انجام شد نشان می‌دهد که وقتی رتبه‌بندی دارای خطای اندک است برآوردگر $\hat{\mu}_{STJPS}$ برآوردگر بهتری نسبت به $\hat{\mu}_{STSRS}$ است. همچنین برآوردگرهای $\hat{\mu}_{new_1}$ و $\hat{\mu}_{new_2}$ در مقایسه با برآوردگرهای $\hat{\mu}_{STJPS}$ و $\hat{\mu}_{STSRS}$ در بیشتر حالت‌ها برآوردگر بهتری هستند.

تشکر و قدردانی

نویسندگان از داوران و پراستار محترم که با پیشنهادات ارزنده خود باعث بهبود مقاله شدند، کمال تشکر و قدردانی را دارند.

مراجع

زمان‌زاده، الف. و احمدی، ج. (۱۳۹۰)، بازه اطمینان ناپارامتری با ضریب اطمینان دقیق برای چندک‌های جامعه براساس نمونه‌گیری از مجموعهٔ رتبه‌ای، *مجله علوم آماری*، ۵(۱)، ۲۳–۳۹.

محمدقاسمی، ح.، زمان‌زاده، الف. و محمدی، م. (۱۳۹۵)، برآوردگر جدید میانگین در طرح نمونه‌گیری طبقه‌ای قضاوتی با مرتب کردن مشاهدات درون طبقات، *مجله علوم آماری*، ۱۰(۱)، ۱۲۹–۱۳۷.

مرکز آمار ایران، (۱۳۸۲)، نتایج سرشماری عمومی کشاورزی.

Cochran, W. G. (1977), *Sampling Techniques*, 3rd Edition. Wiley, New York.

Dalenius, T. (1950), The Problem of Optimum Stratification, *Skand. Aktuarie Tidskr*, **33**, 203-213.

Dastbaravarde, A., Arghami, N. R. and Sarmad, M. (2016), Some Theoretical Results Concerning Non-parametric Estimation by Using a Judgment Poststratification Sample, *Communications in Statistics -Theory and Methods*, **45**, 2181-2203.

Dell, T. R. and Clutter, J. L. (1972), Ranked Set Sampling Theory with Order Statistics Background, *Biometrics*, 545–555.

Frey, J. and Feeman, T. G. (2012), An Improved Mean Estimator for Judgment Post-Sratification, *Computational Statistics and Data Analysis*, **56**, 418–426.

Frey, J. and Ozturk, O. (2011), Constrained Estimation Using Judgment Post-Stratification, *Annals of the Institute of Statistical Mathematics*, **63**, 769–789.

Halls, L. K. and Dell, T. R. (1966), Trial of Ranked-Set Sampling for Forage Yields, *Forest Science*, **12**, 22–26.

MacEachern, S. N., Stasny, E. A. and Wolfe, D. A. (2004), Judgment Post-Stratification with Imprecise Rankings, *Biometrics*, **60**, 207–215.

McIntyre, G. A. (1952), A Method For Unbiased Selective Sampling, Using Ranked Sets, *Australian J. Agricultural Research*, **3**, 385–390.

Takahasi, K. and Wakimoto, K. (1968), On Unbiased Estimates of the Population Mean Based on the Sample Stratified by Means of Ordering, *Annals of the Institute of Statistical Mathematics*, **20**, 1–31.

Wang, X., Lim, J. and Stokes, L. (2008), A Nonparametric Mean Estimator for Judgment Poststratified Data, *Biometrics*, **64**, 355–363.

Journal of Statistical Sciences, Autumn and Winter, 2020
Vol. 14, No. 2, pp 535-548
DOI: 10.29252/jss.13.2.535

The Population Mean Estimators by using Judgment Post Stratification in Stratified Sampling

Najafi Majid Abadi، A.، Nematollahi، N.
Department of Statistics, Allameh Tabataba'i University, Tehran, Iran.

Abstract: Judgment post-stratification is a method of using additional information of ranking in the simple random sampling, to increase the efficiency of the estimators of population parameters. In this paper, we use judgment post-stratification instead of simple random sampling in stratums of stratified sampling, and present new estimators for population mean. Then, we compare the proposed estimators with random stratified mean estimator by using a simulation study. The simulation results show that the proposed estimators perform better than the random stratified mean estimator in most of the cases.

Keywords: Judgment post-stratification method, Efficiency, Simple random sampling, Stratified sampling.

Mathematics Subject Classification (2010): 62D05, 62F07, 62F05.