

برآورد واریانس توزیع نرمال با استفاده از نمونه‌گیری مجموعه رتبه‌دار

نرگس مهدوی‌منش، نصراله ایران پناه، احسان زمان‌زاده؛

دانشگاه اصفهان، دانشکده علوم، گروه ریاضی

پذیرش ۹۷/۰۶/۲۶

دریافت ۹۷/۰۲/۲۱

چکیده

نمونه‌گیری یکی از مهم‌ترین بخش‌های علم آمار است. در هر تحقیق، پژوهشگر در پی یافتن روش مناسب برای جمع‌آوری نمونه و اطلاعات مربوط به آن است که کارا و کم هزینه باشد. در شرایطی که اندازه‌گیری واحدهای جامعه مشکل یا پرهزینه باشد، اما بتوان واحدهای جامعه را به‌سادگی و با کم‌ترین هزینه رتبه‌بندی کرد، روش نمونه‌گیری مجموعه رتبه‌دار مورد استفاده قرار می‌گیرد. در این مقاله ابتدا روش نمونه‌گیری مجموعه رتبه‌دار معرفی می‌شود. سپس چند روش برآورد واریانس توزیع نرمال با ترکیب برآوردهای بین‌گروهی و درون‌گروهی نارایب ارائه می‌شود. در نهایت برآوردهای ارائه شده با استفاده از پژوهش‌های شبیه‌سازی با یکدیگر مقایسه می‌شوند.

واژه‌های کلیدی: توزیع نرمال، نمونه‌گیری مجموعه رتبه‌دار، نمونه‌گیری تصادفی ساده، کارایی، شبیه‌سازی مونت کارلو.

مقدمه

در برخی از تحقیقات کاربردی، اندازه‌گیری دقیق مشخصه تحت پژوهش پرهزینه یا دشوار بوده است؛ اما می‌توان واحدها را به آسانی انتخاب و رتبه‌بندی کرد. در این شرایط می‌توان از نمونه‌گیری مجموعه رتبه‌دار^۱ (RSS) استفاده کرد. به‌طور مثال برای تعیین دامنه نرمال سطح بیلیروبین در خون نوزادان تازه متولد شده باید نمونه خون از نوزادان انتخاب شده، گرفته و در آزمایشگاه میزان بیلیروبین آزمایش شود. از طرف دیگر سطح بیلیروبین در تعداد کمی نوزاد، با مشاهده صورت، قفسه سینه، قسمت‌های تحتانی بدن و نقاط انتهایی کل بدن که زرد است، قابل رتبه‌بندی است. روش RSS دارای قابلیت اجرایی در آزمایش‌های بالینی است. معمولاً هزینه آزمایش‌های بالینی برای بیمار زیاد است، اما برای هر بیمار آزمایش‌های بالینی را می‌توان با روش RSS بر اساس اطلاعاتی مانند سن، وزن، قد، فشار خون و سوابق سلامت انتخاب کرد که به‌طور معناداری باعث کم‌تر شدن هزینه آزمایش‌های بالینی می‌شود. روش RSS را اولین بار مک‌این‌تایر [۴] به‌منظور برآورد میانگین محصولات یونجه معرفی کرد، اما وی نتوانست کارایی این روش نمونه‌گیری را نشان دهد. هالز و دل [۲] با استفاده از این روش نمونه‌گیری، میانگین ارتفاع درختان را برآورد کردند. تاکاهاشی و واکیموتو [۹] نشان دادند، برآوردگر میانگین جامعه در روش RSS نارایب و کاراتر از برآوردگر میانگین جامعه در نمونه‌گیری تصادفی ساده^۲ (SRS) است. دل و کلاتر [۱] از روش RSS بدون در نظر گرفتن این‌که

رتبه‌بندی واحدها کامل است یا ناقص، نشان دادند برآوردگر میانگین جامعه در روش RSS برای میانگین جامعه نارایب است. استوکس [۷] روش RSS را بر اساس رتبه‌بندی متغیر کمکی به جای متغیر اصلی بررسی کرد. استوکس [۸] در روش RSS برآوردگری برای واریانس جامعه معرفی کرد که به‌طور مجانبی نارایب و کارایی مجانبی آن بیش‌تر از برآوردگر واریانس جامعه در روش SRS است. سینه‌ها و همکاران [۶] مسئله برآورد پارامترهای توزیع‌های نرمال و نمایی را در روش RSS بررسی کردند. یو و همکاران [۱۰] برآوردگر واریانس توزیع نرمال را در روش RSS در هر دو حالت متعادل و نامتعادل ارائه دادند. مک‌ایچرن و همکاران [۳] برآوردگر نارایبی برای واریانس جامعه در روش RSS ارائه کردند که کاراتر از برآوردگر واریانس جامعه در روش SRS و هم‌چنین برآوردگر واریانس استوکس [۸] است. پرون و سینه‌ها [۵] برآورد واریانس را به‌روش ناپارامتری بررسی کرده و هم‌چنین رده‌ای از برآوردگرهای نارایب برای واریانس جامعه در روش RSS معرفی کردند. فری و فیمین [۲] با استفاده از نمونه قضاوتی برآوردی برای واریانس جامعه ارائه کردند. زمان‌زاده و واک [۱۲] برآورد ناپارامتری واریانس با استفاده از متغیر کمکی را بررسی کردند.

در این مقاله، ابتدا به معرفی الگوریتم روش RSS پرداخته می‌شود. سپس برآوردگرهایی نارایب برای واریانس توزیع نرمال با استفاده از روش RSS در حالت یک چرخه و بیش از یک چرخه ارائه می‌شود. در نهایت، با استفاده از شبیه‌سازی، این برآوردگرها با یک‌دیگر و با برآوردگر ارائه شده در روش SRS مقایسه می‌شوند. برنامه‌های مورد نیاز برای پژوهش‌های شبیه‌سازی با استفاده از نرم‌افزار آماری R-3.3.1 انجام شده‌اند.

روش نمونه‌گیری مجموعه رتبه‌دار

نمونه‌گیری مجموعه رتبه‌دار از روش‌های نمونه‌گیری شناخته شده‌ای است که استفاده از آن برای برآورد پارامترهای نامعلوم جامعه اغلب موجب افزایش کارایی و دقت برآوردگر می‌شود.

مراحل روش RSS به‌صورت زیر است:

مرحله ۱. ابتدا n^2 نمونه تصادفی از جامعه بررسی شده انتخاب و آن را به تصادف به n گروه سطری به حجم n بدین‌صورت تقسیم می‌کنیم:

$$\begin{array}{cccc} X_{11} & X_{21} & \dots & X_{n1} \\ X_{12} & X_{22} & \dots & X_{n2} \\ \vdots & \vdots & & \vdots \\ X_{1n} & X_{2n} & \dots & X_{nn} \end{array}$$

که در آن متغیر تصادفی X_{ij} معرف شاخص مورد نظر برای واحد i -ام در مجموعه j -ام است.

مرحله ۲. بدون آگاهی از متغیر مورد نظر و بر اساس قضاوت شخصی و یا تجهیزات بصری و در کل هر روش بدون هزینه، n واحد درون هر گروه را بدین‌صورت رتبه‌بندی می‌کنیم:

مرحله ۳. از مجموعه r -ام $(r = 1, \dots, n)$ ، r -امین واحد رتبه‌بندی بدین‌صورت انتخاب و مقادیر آن‌ها اندازه‌گیری می‌شود:

که برای سادگی بدین‌صورت نمایش داده می‌شوند:

از آن‌جاکه با بزرگ شدن n امکان دارد رتبه‌بندی با خطای بیش‌تری صورت گیرد، معمولاً n را کوچک و بسته به توانایی محقق در رتبه‌بندی واحدهای نمونه بین ۳ تا ۱۰ انتخاب می‌کنند و برای رسیدن به حجم نمونه دلخواه، این مراحل را m بار، که به هر مرتبه یک چرخه می‌گویند، تکرار می‌کنند و به حجم نمونه $N = nm$ دست می‌یابند. رتبه‌ها برای هر واحد در هر مجموعه ممکن است مطابق با آماره ترتیبی متناظر با آن واحد نباشد، اگر رتبه‌های واحدها مطابق با آماره ترتیبی متناظر با آن واحد باشد رتبه‌بندی را کامل گویند و آن‌ها را با $X_{(r)i}$ نشان می‌دهند. در این مقاله فرض بر این است که نمونه‌گیری کامل است و $X_{[r]}$ ها توزیعی یک‌سان با توزیع r -امین آماره ترتیبی در یک نمونه تصادفی به حجم n دارند و با توجه به کامل بودن نمونه‌گیری آن‌ها را با $X_{(r)}$ نشان می‌دهیم.

برآورد واریانس برای یک چرخه

در این بخش برآوردگرهای نارایب برای σ^2 بر اساس یک چرخه در روش RSS معرفی می‌گردند. ابتدا برآوردگر استوکس [۸] را به‌صورت (۱) در نظر می‌گیریم:

$$\hat{\sigma}_{St}^2 = \frac{\sum_{r=1}^n (X_{(r)} - \hat{\mu})^2}{(n-1)} \quad (1)$$

که در آن $\hat{\mu} = \frac{\sum_{r=1}^n X_{(r)}}{n}$ است. و داریم:

$$E(\hat{\sigma}_{St}^2) = \sigma^2 + \frac{\sum_{r=1}^n (\mu_{(r)} - \mu)^2}{n(n+1)}$$

با قرار دادن $\sigma v_{(r)}$ به جای $(\mu_{(r)} - \mu)$ در رابطه مذکور می‌توان نوشت:

$$E(\hat{\sigma}_{St}^2) = \sigma^2 \left(1 + \frac{\sum_{r=1}^n v_{(r)}^2}{n(n+1)} \right)$$

در نتیجه داریم:

$$E \left(\frac{\hat{\sigma}_{St}^2}{1 + \frac{\sum_{r=1}^n v_{(r)}^2}{n(n+1)}} \right) = E \left(\frac{\sum_{r=1}^n (X_{(r)} - \hat{\mu})^2}{n-1 + \frac{1}{n} \sum_{r=1}^n v_{(r)}^2} \right) = \sigma^2$$

بنابراین برآوردگر (۲) را که تعمیمی از $\hat{\sigma}_{St}^2$ است، می‌توان به‌عنوان برآوردگر نارایب σ^2 در توزیع نرمال در نظر گرفت:

$$\hat{\sigma}_{SM}^2 = \frac{\sum_{r=1}^n (X_{(r)} - \hat{\mu})^2}{n-1 + \frac{1}{n} \sum_{r=1}^n v_{(r)}^2}. \quad (2)$$

یو و همکاران [۱۰] برآوردگر نارایب (۳) را برای واریانس توزیع نرمال با استفاده از روش RSS ارائه کردند:

$$\hat{\sigma}_b^2 = a_n \sum_{r=1}^n \frac{(X_{(r)} - \hat{\mu}_{blue})^2}{v_{2(r)}} \quad (۳)$$

که در آن با توجه به سینه‌ها و همکاران [۶]، $\hat{\mu}_{blue} = \frac{\sum_{r=1}^n (X_{(r)} / v_{2(r)})}{\sum_{r=1}^n (1 / v_{2(r)})}$ بهترین برآوردگر نارایب خطی میانگین با $v_{2(r)} = \frac{\mu_{2(r)}}{\sigma^2}$ است و a_n که به صورتی به دست آمده که $\hat{\sigma}_b^2$ برای σ^2 نارایب باشد، بدین صورت محاسبه می‌شود:

$$a_n = \frac{1}{n - 1 + \sum_{r=1}^n \frac{v_{2(r)}^2}{v_{2(r)}}}.$$

برآورد واریانس برای بیش از یک چرخه

در این بخش برآوردگر واریانس توزیع نرمال با استفاده از روش RSS برای بیش از یک چرخه و n نمونه در هر چرخه ارائه می‌شود.

ابتدا فرض کنید نمونه‌ای به حجم mn با n رده هر کدام با اندازه m به صورت مجموعه

$$\{X_{(r)i}; r = 1, \dots, n, i = 1, \dots, m\}$$

داشته باشیم، برآوردگر (۴) را که تعمیم برآوردگر استوکس [۸] بر اساس m چرخه است و مانند برآوردگر $\hat{\sigma}_{SM}^2$ در حالت یک چرخه به دست می‌آید، به عنوان برآوردی نارایب از σ^2 در نظر می‌گیریم:

$$\hat{\sigma}_{MSM}^2 = \frac{\sum_{i=1}^m \sum_{r=1}^n (X_{(r)i} - \hat{\mu})^2}{mn - 1 + \frac{1}{n} \sum_{r=1}^n v_{2(r)}^2} \quad (۴)$$

برآوردگرهای دیگری برای σ^2 با استفاده از این روش‌های به دست می‌آیند:

۱. در این روش، هر چرخه را یک نمونه جداگانه به صورت $\{X_{(1)i}, \dots, X_{(n)i}\}$ در نظر گرفته و برآوردگر σ^2 برای چرخه i -ام که با $\hat{\sigma}_{(i)}^2$ نشان داده می‌شود، محاسبه می‌شود. سپس از آن جاکه $\hat{\sigma}_{(i)}^2$ ‌ها از یکدیگر مستقل هستند، برآوردگر بر اساس m چرخه به صورت $\frac{1}{m} \sum_{i=1}^m \hat{\sigma}_{(i)}^2$ به دست می‌آید. با جای گذاری دو برآوردگر $\hat{\sigma}_{SM}^2$ و $\hat{\sigma}_b^2$ که در روابط (۲) و (۳) ارائه شده‌اند، به جای $\hat{\sigma}_{(i)}^2$ ، برآوردگرهای حاصل از این روش به ترتیب با $\hat{\sigma}_{CSM}^2$ و $\hat{\sigma}_{Cb}^2$ نشان داده می‌شوند.

۲. برای ارائه برآوردگر با استفاده از این روش RSS، mn نمونه، n رده هر کدام با اندازه m در نظر گرفته می‌شود. به منظور ارائه برآوردگر ابتدا $\hat{\sigma}_{MSM}^2$ بدین صورت بازنویسی می‌شود:

$$\hat{\sigma}_{MSM}^2 = C \left\{ (m-1) \sum_{r=1}^n V_r + m \sum_{r=1}^n (\bar{X}_{(r)} - \hat{\mu})^2 \right\}$$

$$\text{که در آن } \bar{X}_{(r)} = \frac{1}{m} \sum_{i=1}^m X_{(r)i} \text{ و } C = \left(mn - 1 + \frac{1}{n} \sum_{r=1}^n v_{2(r)}^2 \right)^{-1}$$

$$V_r = \frac{1}{m-1} \sum_{i=1}^m (X_{(r)i} - \bar{X}_{(r)})^2$$

چنان که مشاهده می‌شود $\hat{\sigma}_{MSM}^2$ را می‌توان به دو برآوردگر درون گروهی و بین گروهی تجزیه کرد. با توجه به یو و همکاران [۱۰]، $E(V_r) = \sigma^2 v_{2(r)}$ ، پس یک برآوردگر ناریب برای σ^2 بر اساس V_r بدین صورت است:

$$\hat{\sigma}_W^2 = \sum_{r=1}^n d_r V_r ; d_r = \frac{\frac{v_{2(r)}}{\eta_{(r)}}}{\sum_{k=1}^n \frac{v_{2(k)}}{\eta_{(k)}}} \quad (5)$$

علاوه بر این، V_r ها مستقل با واریانس $Var(V_r) = \sigma^4 \eta_{(r)}$ هستند که در آن $\eta_{(r)}$ بدین صورت محاسبه می‌شود:

$$\eta_{(r)} = \frac{v_{4(r)} + v_{2(r)}^2}{m} + \frac{2}{m(m-1)} v_{2(r)}^2$$

بنابر این

$$Var(\hat{\sigma}_W^2) = \sum_{r=1}^n d_r^2 Var(V_r) = \frac{\sigma^4}{\sum_{r=1}^n \frac{v_{2(r)}^2}{\eta_{(r)}}}$$

یو و همکاران [۱۰]، سپس قسمت بین گروهی را مدنظر قرار داده و دو برآوردگر برای σ^2 بر اساس میانگین‌های n رده یعنی $\{\bar{X}_{(r)}; r=1, \dots, n\}$ ارائه کردند. دو برآوردگر مطرح شده بدین صورت است:

الف) برآوردگر بین گروهی بر اساس $\hat{\sigma}_{SM}^2$

$$\hat{\sigma}_{B1}^2 = \frac{\sum_{r=1}^n (\bar{X}_{(r)} - \hat{\mu})^2}{\frac{n-1}{m} + \frac{nm-n+1}{nm} \sum_{r=1}^n v_{2(r)}^2} \quad (6)$$

ب) برآوردگر بین گروهی بر اساس $\hat{\sigma}_b^2$

$$\hat{\sigma}_{B2}^2 = a_{n,m} \sum_{r=1}^n \frac{(\bar{X}_{(r)} - \tilde{\mu}_{blue})^2}{v_{2(r)}/m} \quad (7)$$

$$\tilde{\mu}_{blue} = \frac{\sum_{r=1}^n (\bar{X}_{(r)}/v_{2(r)})}{\sum_{r=1}^n (1/v_{2(r)})} \text{ و } a_{n,m} = \left(n-1 + m \sum_{r=1}^n \frac{v_{2(r)}^2}{v_{2(r)}} \right)^{-1}$$

که در آن

۳. در این روش رده‌ای از برآوردگرهای ناریب بدین صورت معرفی می‌شود:

$$\hat{\sigma}_{WB}^2(\varepsilon) = \varepsilon \hat{\sigma}_W^2 + (1-\varepsilon) \hat{\sigma}_B^2 ; \varepsilon \in [0,1] \quad (8)$$

که در آن $\hat{\sigma}_W^2$ برآوردگر ناریب درون گروهی ارائه شده در (۵) است و $\hat{\sigma}_B^2$ یکی از دو برآوردگر بین گروهی (۶) یا (۷) است. با قرار دادن هر کدام از دو برآوردگر (۶) و (۷) در رابطه (۸) برآوردگرهایی حاصل می‌شوند که به ترتیب با $\hat{\sigma}_{WB1}^2$ و $\hat{\sigma}_{WB2}^2$ نشان داده می‌شوند.

قضیه ۱. مقدار بهینه ε که واریانس $\hat{\sigma}_{WB}^2$ را به کم‌ترین مقدار می‌رساند بدین صورت است:

$$\varepsilon^* = \frac{Var(\hat{\sigma}_B^2) - Cov(\hat{\sigma}_W^2, \hat{\sigma}_B^2)}{Var(\hat{\sigma}_W^2) + Var(\hat{\sigma}_B^2) - 2Cov(\hat{\sigma}_W^2, \hat{\sigma}_B^2)}$$

اثبات.

$$Var(\hat{\sigma}_{WB}^2) = \varepsilon^2 Var(\hat{\sigma}_W^2) + (1 - \varepsilon)^2 Var(\hat{\sigma}_B^2) + 2\varepsilon(1 - \varepsilon)Cov(\hat{\sigma}_W^2, \hat{\sigma}_B^2) \quad (9)$$

با مشتق‌گیری از رابطه (۹) نسبت به ε و برابر صفر قرار دادن آن داریم:

$$\varepsilon = \frac{Var(\hat{\sigma}_B^2) - Cov(\hat{\sigma}_W^2, \hat{\sigma}_B^2)}{Var(\hat{\sigma}_W^2) + Var(\hat{\sigma}_B^2) - 2Cov(\hat{\sigma}_W^2, \hat{\sigma}_B^2)} \quad (10)$$

از طرفی مشتق دوم $Var(\hat{\sigma}_{WB}^2)$ نسبت به ε بدین صورت است:

$$\begin{aligned} \frac{\partial^2 Var(\hat{\sigma}_{WB}^2)}{\partial \varepsilon^2} &= 2Var(\hat{\sigma}_W^2) + 2Var(\hat{\sigma}_B^2) - 4Cov(\hat{\sigma}_W^2, \hat{\sigma}_B^2) \\ &= 2Var(\hat{\sigma}_W^2 - \hat{\sigma}_B^2). \end{aligned}$$

چنان‌که مشاهده می‌شود مشتق دوم یک عبارت مثبت است و در نتیجه رابطه (۱۰) باعث کمینه شدن (۹)

می‌شود.

پژوهش‌های شبیه‌سازی

در این بخش از شبیه‌سازی مونت کارلو برای مقایسه برآوردهای ارائه شده در دو بخش قبل با یکدیگر و مقایسه آن‌ها با برآوردهای واریانس در روش SRS (S^2) و $\hat{\sigma}_{St}^2$ در رابطه (۱) استفاده شده است. تمامی نمونه‌های این پژوهش از توزیع نرمال استاندارد تولید شده‌اند.

در جدول ۱ مقادیر کارایی برآوردهای $\hat{\sigma}_b^2$ نسبت به $\hat{\sigma}_{SM}^2$ و S^2 برای ۱۰۰۰۰ بار تکرار مونت کارلو با $n = 3, \dots, 10$ و $m = 1$ ارائه شده است. با توجه به این جدول مشاهده می‌شود که کارایی $\hat{\sigma}_b^2$ از هر دو برآوردهای $\hat{\sigma}_{SM}^2$ و S^2 بیش‌تر است و با افزایش n بیش‌تر می‌شود.

جدول ۱. مقادیر کارایی $\hat{\sigma}_b^2$ نسبت به دو برآوردهای $\hat{\sigma}_{SM}^2$ و S^2

n	۳	۴	۵	۶	۷	۸	۹	۱۰
$e(\hat{\sigma}_b^2, S^2)$	۱/۲۷	۱/۴۱	۱/۵۷	۱/۷۰	۱/۸۴	۱/۹۵	۲/۰۸	۲/۲۴
$e(\hat{\sigma}_b^2, \hat{\sigma}_{SM}^2)$	۱/۰۳	۱/۰۴	۱/۰۵	۱/۰۸	۱/۱۱	۱/۱۲	۱/۱۳	۱/۱۵

جدول ۲ شامل مقادیر کارایی برآوردهای $\hat{\sigma}_W^2$ ، $\hat{\sigma}_{B1}^2$ و $\hat{\sigma}_{B2}^2$ نسبت به S^2 و $\hat{\sigma}_{MSM}^2$ می‌باشد. هم‌چنین جدول ۳ مقادیر کارایی $\hat{\sigma}_{WB1}^2$ و $\hat{\sigma}_{WB2}^2$ را نسبت به این دو برآوردهای نشان می‌دهد که در فرمول این دو برآوردهای جهت شبیه‌سازی، از برآورد ε^* استفاده شده است. در پژوهش مربوط به این دو جدول تعداد تکرار مونت کارلو ۱۰۰۰۰۰ با $n=2, \dots, 5$ و $m=2, \dots, 6$ است. چنان‌که از جدول ۲ مشاهده می‌شود برآوردهای $\hat{\sigma}_W^2$ ، $\hat{\sigma}_{B1}^2$ و $\hat{\sigma}_{B2}^2$ عملکرد ضعیفی نسبت به برآوردهای روش SRS و هم‌چنین برآوردهای تعمیم یافته استوکس دارند.

جدول ۲. مقادیر کارایی $\hat{\sigma}_{B1}^2$ و $\hat{\sigma}_{B2}^2$ نسبت به S^2 و $\hat{\sigma}_{MSM}^2$

$e(\hat{\sigma}_{B2}^2, \hat{\sigma}_{MSM}^2)$	$e(\hat{\sigma}_{B1}^2, \hat{\sigma}_{MSM}^2)$	$e(\hat{\sigma}_W^2, \hat{\sigma}_{MSM}^2)$	$e(\hat{\sigma}_{B2}^2, S^2)$	$e(\hat{\sigma}_{B1}^2, S^2)$	$e(\hat{\sigma}_W^2, S^2)$	m	n
۰/۵۲	۰/۵۲	۰/۶۰	۰/۵۵	۰/۵۶	۰/۶۵	۲	۲
۰/۴۱	۰/۴۱	۰/۷۵	۰/۴۳	۰/۴۳	۰/۷۹	۳	
۰/۳۶	۰/۳۶	۰/۸۰	۰/۳۷	۰/۳۷	۰/۸۳	۴	
۰/۳۳	۰/۳۳	۰/۸۵	۰/۳۴	۰/۳۴	۰/۸۷	۵	
۰/۳۱	۰/۳۱	۰/۸۷	۰/۳۲	۰/۳۲	۰/۸۸	۶	
۰/۶۷	۰/۶۶	۰/۵۰	۰/۷۷	۰/۷۶	۰/۵۸	۲	۳
۰/۵۸	۰/۵۷	۰/۶۴	۰/۶۵	۰/۶۴	۰/۷۲	۳	
۰/۵۲	۰/۵۲	۰/۷۱	۰/۵۸	۰/۵۷	۰/۷۹	۴	
۰/۵۰	۰/۴۹	۰/۷۴	۰/۵۶	۰/۵۵	۰/۸۳	۵	
۰/۴۸	۰/۴۸	۰/۷۷	۰/۵۴	۰/۵۳	۰/۸۵	۶	
۰/۷۶	۰/۷۳	۰/۴۴	۰/۹۷	۰/۹۴	۰/۵۶	۲	۴
۰/۶۷	۰/۶۵	۰/۵۷	۰/۸۲	۰/۸۰	۰/۷۰	۳	
۰/۶۳	۰/۶۲	۰/۶۴	۰/۷۷	۰/۷۵	۰/۷۸	۴	
۰/۶۱	۰/۶۰	۰/۶۶	۰/۷۴	۰/۷۲	۰/۸۰	۵	
۰/۶۰	۰/۵۸	۰/۷۰	۰/۷۲	۰/۷۱	۰/۸۵	۶	
۰/۸۲	۰/۷۹	۰/۴۰	۱/۱۱	۱/۰۶	۰/۵۴	۲	۵
۰/۷۵	۰/۷۲	۰/۵۱	۰/۹۹	۰/۹۵	۰/۶۸	۳	
۰/۷۱	۰/۶۹	۰/۵۷	۰/۹۳	۰/۹۰	۰/۷۵	۴	
۰/۶۹	۰/۶۷	۰/۶۲	۰/۹۰	۰/۸۷	۰/۸۱	۵	
۰/۶۸	۰/۶۶	۰/۶۴	۰/۸۷	۰/۸۵	۰/۸۳	۶	

جدول ۳. مقادیر کارایی $\hat{\sigma}_{WB2}^2$ و $\hat{\sigma}_{WB1}^2$ نسبت به S^2 و $\hat{\sigma}_{MSM}^2$

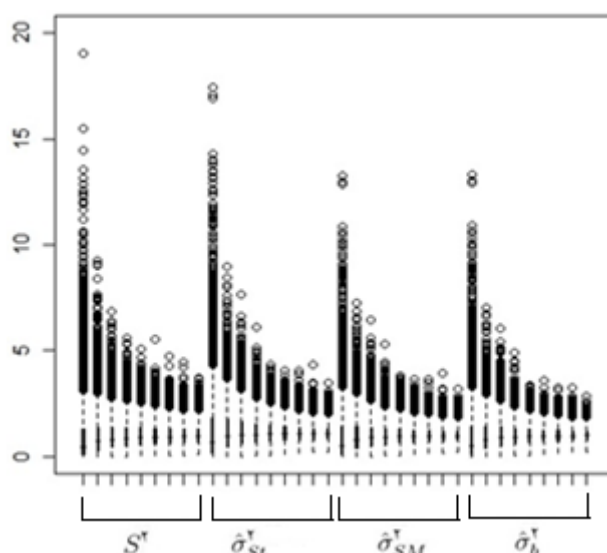
$e(\hat{\sigma}_{WB2}^2, \hat{\sigma}_{MSM}^2)$	$e(\hat{\sigma}_{WB1}^2, \hat{\sigma}_{MSM}^2)$	$e(\hat{\sigma}_{WB2}^2, S^2)$	$e(\hat{\sigma}_{WB1}^2, S^2)$	m	n
۱/۰۶	۱/۰۶	۱/۱۳	۱/۱۳	۲	۲
۱/۰۸	۱/۰۸	۱/۱۴	۱/۱۴	۳	
۱/۰۸	۱/۰۸	۱/۱۲	۱/۱۲	۴	
۱/۱۰	۱/۱۰	۱/۱۲	۱/۱۲	۵	
۱/۱۰	۱/۱۰	۱/۱۲	۱/۱۲	۶	
۱/۰۸	۱/۰۷	۱/۲۵	۱/۲۴	۲	۳
۱/۱۱	۱/۱۰	۱/۲۵	۱/۲۴	۳	
۱/۱۱	۱/۱۱	۱/۲۴	۱/۲۴	۴	
۱/۱۳	۱/۱۲	۱/۲۶	۱/۲۶	۵	
۱/۱۳	۱/۱۳	۱/۲۶	۱/۲۶	۶	
۱/۱۰	۱/۰۸	۱/۴۰	۱/۳۷	۲	۴
۱/۱۲	۱/۱۱	۱/۳۸	۱/۳۶	۳	
۱/۱۴	۱/۱۳	۱/۴۰	۱/۳۸	۴	
۱/۱۳	۱/۱۲	۱/۳۷	۱/۳۶	۵	
۱/۱۶	۱/۱۵	۱/۴۰	۱/۳۸	۶	
۱/۱۱	۱/۰۸	۱/۵۱	۱/۴۶	۲	۵
۱/۱۴	۱/۱۱	۱/۵۱	۱/۴۷	۳	
۱/۱۴	۱/۱۳	۱/۵۰	۱/۴۷	۴	
۱/۱۶	۱/۱۲	۱/۵۱	۱/۴۸	۵	
۱/۱۷	۱/۱۵	۱/۵۱	۱/۴۸	۶	

پس از ادغام $\hat{\sigma}_W^2$ با هر یک از دو برآوردگر $\hat{\sigma}_{B1}^2$ و $\hat{\sigma}_{B2}^2$ به ترتیب برآوردگرهای $\hat{\sigma}_{WB1}^2$ و $\hat{\sigma}_{WB2}^2$ به دست می‌آیند که با توجه به جدول ۳ مشاهده می‌شود کارایی این دو برآوردگر از S^2 و $\hat{\sigma}_{MSM}^2$ بیش‌تر است و دو برآوردگر ادغامی عملکرد تقریباً یکسانی دارند. همچنین مشاهده می‌شود که کارایی این دو برآوردگر نسبت به S^2 با افزایش m تقریباً ثابت می‌ماند اما با افزایش n افزایش می‌یابد و نسبت به $\hat{\sigma}_{MSM}^2$ با افزایش n و m با شدت بسیار کمی افزایش می‌یابد و می‌توان گفت با تغییر n و m کارایی $\hat{\sigma}_{WB1}^2$ و $\hat{\sigma}_{WB2}^2$ نسبت به $\hat{\sigma}_{MSM}^2$ تقریباً ثابت می‌ماند.

جدول ۴. مقادیر MSE برآوردگرهای ارائه شده برای واریانس توزیع نرمال

$\hat{\sigma}_{WB2}^2$	$\hat{\sigma}_{WB1}^2$	$\hat{\sigma}_{B2}^2$	$\hat{\sigma}_{B1}^2$	$\hat{\sigma}_W^2$	$\hat{\sigma}_{Cb}^2$	$\hat{\sigma}_{CSM}^2$	n	m
۰/۵۸	۰/۵۸	۱/۱۵۸	۱/۱۵	۱/۰۵	۰/۷۹	۰/۸۰	۲	۲
۰/۳۲	۰/۳۳	۰/۵۳	۰/۵۴	۰/۶۸	۰/۳۹	۰/۳۹	۳	
۰/۲۱	۰/۲۱	۰/۳۱	۰/۳۱	۰/۵۲	۰/۲۴	۰/۲۵	۴	
۰/۳۶	۰/۳۶	۰/۹۴	۰/۹۴	۰/۵۳	۰/۵۶	۰/۵۶	۲	۳
۰/۱۹	۰/۲۰	۰/۳۸	۰/۳۹	۰/۳۴	۰/۲۵	۰/۲۶	۳	
۰/۱۳	۰/۱۳	۰/۲۱	۰/۲۱	۰/۲۷	۰/۱۶	۰/۱۷	۴	
۰/۲۵	۰/۲۵	۰/۸۰	۰/۷۷	۰/۳۴	۰/۴۲	۰/۴۲	۲	۴
۰/۱۵	۰/۱۵	۰/۳۱	۰/۳۱	۰/۲۴	۰/۱۹	۰/۲۰	۳	
۰/۰۹	۰/۱۰	۰/۱۸	۰/۱۸	۰/۱۸	۰/۱۲	۰/۱۲	۴	

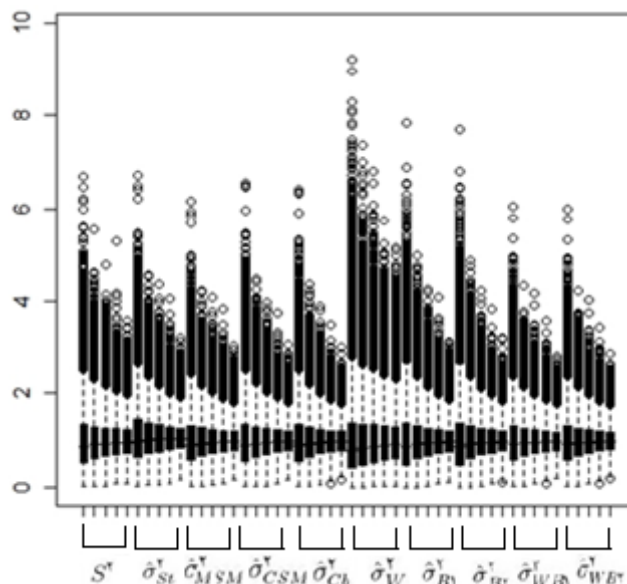
در جدول ۴ مقادیر شبیه‌سازی شده MSE برای برآوردگرهای ارائه شده برای واریانس توزیع نرمال با اندازه‌های $m=2,3,4$ و $n=2,3,4$ برای ۱۰۰۰۰ بار تکرار مونت کارلو ارائه شده است. با توجه به این جدول مشاهده می‌شود که در بین برآوردگرهای ارائه شده برای واریانس توزیع نرمال، دو برآوردگر $\hat{\sigma}_{WB1}^2$ و $\hat{\sigma}_{WB2}^2$ دارای MSE کم‌تری نسبت به بقیه هستند. همچنین با افزایش n و m برای همه برآوردگرها MSE کاهش می‌یابد.



شکل ۱. نمودار جعبه‌ای مقایسه برآوردگرهای S^2 ، $\hat{\sigma}_{St}^2$ ، $\hat{\sigma}_{SM}^2$ و $\hat{\sigma}_b^2$

در شکل ۱ نمودار جعبه‌ای برآوردگرهای $\hat{\sigma}_{SM}^2$ و $\hat{\sigma}_b^2$ در (۲) و (۳) و برآوردگرهای S^2 و $\hat{\sigma}_{St}^2$ برای $n=2, \dots, 10$ و $m=1$ با ۱۰۰۰۰ تکرار مونت کارلو نشان داده شده است. چنان‌که مشاهده می‌شود با افزایش n دامنه تغییرات و دامنه

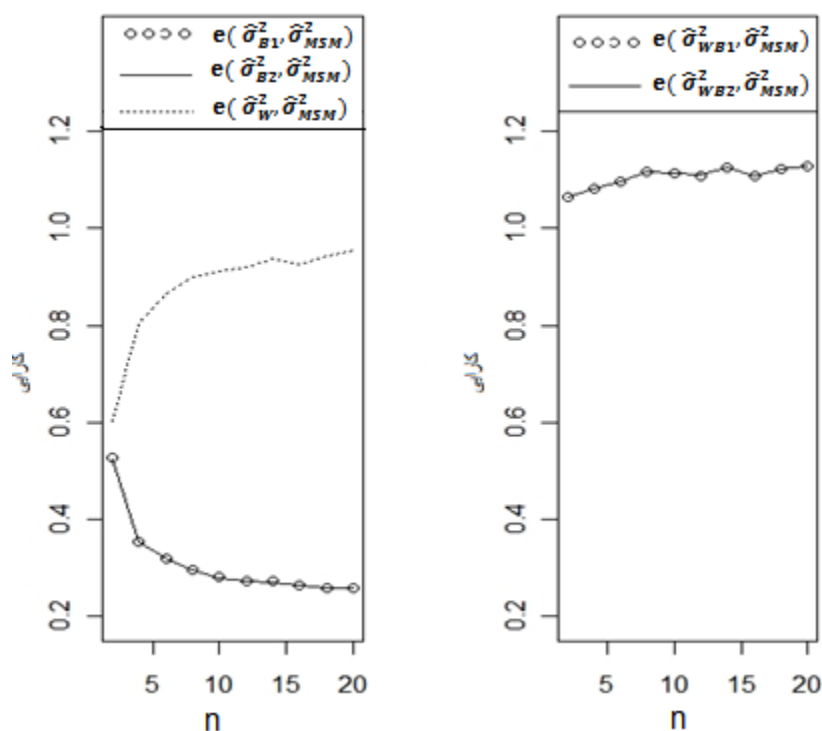
میان چارکی برای هر یک از ۴ برآوردگر کاهش می‌یابد. همچنین برآوردگرهای به‌دست آمده برای توزیع نرمال دارای پراکندگی کمتر و در نتیجه MSE کم‌تری نسبت به برآوردگر استوکس و برآوردگر به‌دست آمده از روش SRS هستند. بین دو برآوردگر مربوط به توزیع نرمال، $\hat{\sigma}_B^2$ پراکندگی و MSE کم‌تری نسبت به $\hat{\sigma}_{SM}^2$ دارد.



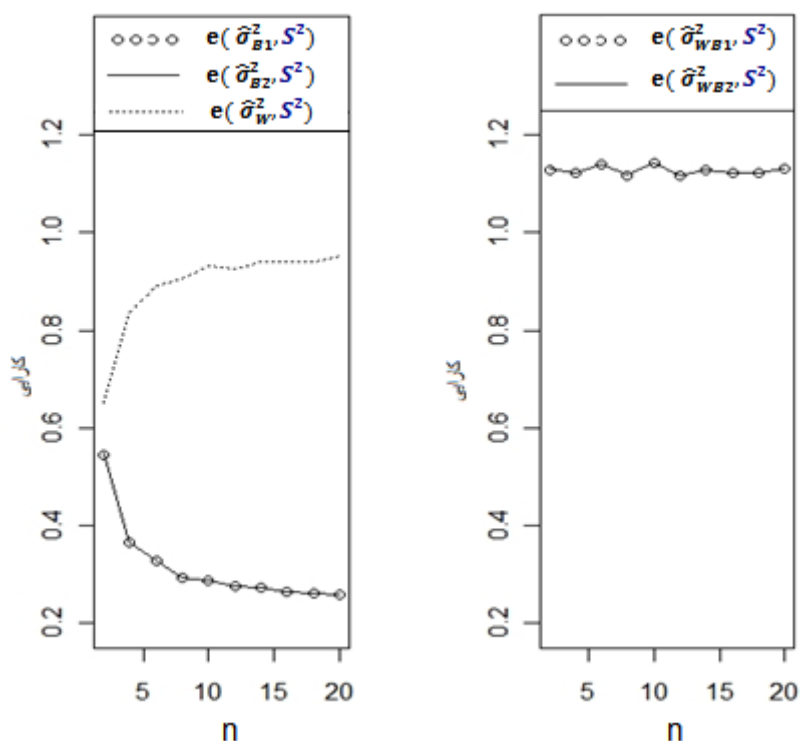
شکل ۲. نمودار جعبه‌ای مقایسه‌ی برآوردگرهای ارائه شده برای واریانس با $m=2$

در شکل ۲ مقادیر شبیه‌سازی شده برآوردگرهای معرفی شده در بخش قبل با تعداد تکرار ۱۰۰۰۰ و اندازه $n=3, \dots, 6$ و $m=2$ نشان داده شده است. چنان‌که مشاهده می‌شود دامنه تغییرات سه برآوردگر $\hat{\sigma}_{B1}^2$ و $\hat{\sigma}_{B2}^2$ و $\hat{\sigma}_W^2$ نسبت به سایر برآوردگرها بیش‌تر است و در نتیجه MSE بیش‌تری نسبت به آن‌ها دارند اما برآوردگرهای به‌دست آمده از ادغام این سه برآوردگر کم‌ترین دامنه تغییرات و در نتیجه کم‌ترین MSE را دارند. همچنین با افزایش n و m دامنه تغییرات همه برآوردگرها کاهش می‌یابد.

در شکل‌های ۳ و ۴ به‌ترتیب نمودار کارایی برآوردگرهای بین‌گروهی، درون‌گروهی و ترکیب آن‌ها نسبت به برآوردگر تعمیم‌یافته استوکس و برآوردگر روش تصادفی ساده رسم شده است. با توجه به این دو شکل مشاهده می‌شود که کارایی برآوردگرهای بین‌گروهی و درون‌گروهی نسبت به دو برآوردگر تعمیم‌یافته استوکس و برآوردگر روش تصادفی ساده کم‌تر از یک است و با ترکیب برآوردگر درون‌گروهی با هر یک از برآوردگرهای بین‌گروهی دو برآوردگر $\hat{\sigma}_{WB1}^2$ و $\hat{\sigma}_{WB2}^2$ به‌دست می‌آیند که کارایی آن‌ها نسبت به هر دو برآوردگر تعمیم‌یافته استوکس و برآوردگر روش تصادفی ساده بیش‌تر از یک است.



شکل ۳. نمودار کارایی برآوردگرهای بین‌گروهی، درون‌گروهی و ترکیب آن‌ها نسبت به برآوردگر تعمیم یافته استوکس



شکل ۴. نمودار کارایی برآوردگرهای بین‌گروهی، درون‌گروهی و ترکیب آن‌ها نسبت به برآوردگر روش تصادفی ساده.

نتیجه‌گیری

در این مقاله ابتدا به معرفی روش RSS پرداختیم. سپس با استفاده از این روش، چند برآوردگر برای واریانس توزیع نرمال ارائه شد. برای مقایسه دقت این برآوردگرها از پژوهش‌های شبیه‌سازی استفاده شد. نتایج شبیه‌سازی نشان داد که MSE این برآوردگرها با افزایش n و کاهش m می‌یابد و در بین برآوردگرهای ارائه شده برای واریانس توزیع نرمال با استفاده از روش RSS، آن‌هایی که از ترکیب برآوردگرهای بین‌گروهی و درون‌گروهی به‌دست آمده‌اند دارای دقت بیش‌تری هستند.

منابع

1. Dell T. R., Clutter J. L. "Ranked set sampling theory with order statistics background", *Biometrics*, 28 (1972) 545-555.
2. Halls L. K., Dell T. R., "Trial of ranked-set sampling for forage yields, *Forest Science*", 12, (1966) 22-26.
3. Frey J., Feeman T. G., "Variance estimation using judgment post-stratification", *Annals of the Institute of Statistical Mathematics* 65 (2013) 551-569.
4. MacEachern S. N., Ozturk O., Wolfe D. A., "A new ranked set sample estimator of variance", *Journal of the Royal Statistical Society* 64 (2002) 177-188.
5. McIntyre G. A. "A method for unbiased selective sampling using ranked sets", *Australian Journal of Agricultural Research* 3 (1952) 385-390.
6. Perron F., Sinha B. K., "Estimation of variance based on a ranked set sample", *Journal of Statistical Planning and Inference* 120 (2004) 21-28.
7. Sinha Bikas k., Purkayastha S., "On some aspects of ranked set sampling for estimation of normal and exponential parameters", *Statistical and Decision*, 14 (1996) 223-240.
8. Stokes S. L., "Ranked set sampling with concomitant variables", *Communications in Statistics-Theory and Methods* 12 (1977) 1207-1211.
9. Stokes S. L., "Estimation of variance using judgment ordered ranked set samples", *Biometrics* 36 (1980) 35-42.
10. Takahasi K., Wakimoto K., "On unbiased estimates of the population mean based on the sample stratified by means of ordering", *Annals of the Institute of Statistical Mathematics* 20(1968) 1-31.
11. Yu P. H. L., Lam K., Sinha B. K., "Estimation of normal variance based on balanced and unbalanced ranked set samples", *Biometrics*, 53 (1999) 1070-1080.

12. Zamanzade E., Vock M., "Variance estimation in ranked set sampling using a concomitant variable", Statistics and Probability Letters, 105 (2015) 1-5.
13. Zamanzade E., Al-Omari A. I., "New ranked set sampling for estimating the population mean and variance", Hacettepe Journal of Mathematics and Statistics (2016) 1891.