مقاله پژوهشی

روش‌هایی برای براورد پارامتر توزیع نمایی در داده‌های تحت سانسور بازه‌ای؛ بررسی مقایسه‌ای

نادر اسدیان،
استادیار، گروه آموزشی آمار دانشکده علوم، دانشگاه لرستان، خرم‌آباد، ایران. asadian.n@lu.ac.ir

محمدحسین پورسعید،
(نویسنده‌ی عهده‌دار مکاتبات)، دانشیار، گروه آموزشی آمار دانشکده علوم، دانشگاه لرستان، خرم‌آباد، ایران. poursaeed.m@lu.ac.ir

چکیده: در این پژوهش برای نرخ خطر داده‌های تحت سانسور بازه‌ای که دارای توزیع نمایی هستند، براوردگرهایی جدید معرفی و بر پایه مطالعه‌های شبیه‌سازی با بعضی از روش‌های موجود مقایسه می‌شوند. با توجه به این‌که در آزمون‌های طول عمر افزایش دقت و همچنین کاهش زمان و هزینه انجام آزمایش‌ها معیارهای مهمی هستند، توزیع مجانبی یکی از روش‌های معرفی شده که نسبت به سایر روش‌های مورد بررسی عملکرد بهتری دارد، به دست آورده می‌شود.

واژه‌های کلیدی: تابع شبه درستنمایی، توزیع نمایی، روش دلتا، زمان نظارت بهینه، سانسور بازه‌ای.

۱. مقدمه

در تجزیه و تحلیل پدیده‌های تصادفی مرتبط با طول عمر، از مباحث نظریه قابلیت اعتماد استفاده می‌شود که در آن، زمان خرابی یک واحد صنعتی و یا زمان مرگ یک واحد بیولوژیک مورد بررسی قرار می‌گیرد. نرخ خطر یکی از شاخص‌های مورد مطالعه در نظریه قابلیت اعتماد است که در تحلیل بقا و بررسی عملکرد سیستم‌ها به آن پرداخته می‌شود. یکی از توزیع‌های طول عمر، توزیع نمایی است که به‌طور گسترده در تجزیه و تحلیل داده‌های صنعتی و پزشکی کاربرد دارد و خاصیت عدم حافظه و همچنین ثابت بودن نرخ خطر از ویژگی‌های جالب این توزیع به شمار می‌رود.

فرض کنید $X_n, \dots, X_2, X_1$ یک نمونه تصادفی n تایی از واحدهای آزمایشی با توزیع نمایی، نرخ خطر θ و تابع چگالی

$$f(x;\theta) = \theta \exp(-\theta x);\ \ x>0, \theta>0$$

است. برای تخمین پارامتر θ ممکن است که آزمایش‌گر به دلیل محدودیت‌های زمانی و برای کاهش هزینه، از طرح‌های مختلف سانسور استفاده کند. برای آشنایی با انواع طرح‌ها و داده‌های سانسور شده می‌توان به کوهن [۱] مراجعه کرد.

یکی از این طرح‌ها سانسور نوع اول است که در آن، آزمایش در زمان از پیش تعیین شده‌ای مانند t پایان می‌یابد و فقط زمان شکست اعضایی از نمونه که قبل از زمان t خراب می‌شوند، مشاهده و ثبت می‌گردد. در این طرح، براورد حداکثر درستنمایی θ به صورت

$$\hat{\theta} = \frac{y}{\sum_{i=1}^{y} x_i + (n-y)t}$$

است که در آن n اندازه نمونه، y تعداد خرابی‌ها و $x_y, \dots, x_1$ زمان‌های خرابی مشاهده شده تا زمان t هستند [۲].

در بعضی موارد ممکن است که واحدهای آزمایشی به‌طور مداوم تحت نظارت نباشند که در این حالت، داده‌های بدست آمده مرتبط با سانسور بازه‌ای هستند که در آن، n واحد آزمایشی در k زمان از پیش تعیین شده مانند $0< t_1 < \cdots < t_k$ مورد بازرسی و نظارت قرار می‌گیرند و تعداد خرابی‌های رخ داده در بین زمان‌های بازرسی ثبت می‌شوند. در این طرح، آزمایش‌گر اطلاعی از زمان دقیق خرابی واحدها ندارد و فقط تعداد خرابی‌های رخ داده در بین زمان‌های بازرسی را می‌داند.

اگر $t_0 \equiv 0$، زمان شروع آزمایش و $(0,t_1], (t_1,t_2], \dots, (t_{k-1},t_k]$ فاصله‌های بین زمان‌های نظارت و y_i؛ $i=1,2,\dots,k$ نیز تعداد واحدهای خراب شده در فاصله زمانی $(t_{i-1},t_i]$ باشد آن‌گاه $(Y_1,Y_2,\dots,Y_k)$ دارای توزیع چندجمله‌ای و تابع درستنمایی به صورت رابطه ۱ خواهد بود

Corresponding author: poursaeed.m@lu.ac.ir
https://doi.org/10.48313/jqem.2023.194410

تاریخ دریافت: ۱۴۰۲/۰۹/۲۱ تاریخ پذیرش: ۱۴۰۳/۰۱/۰۴
دوره ۱۳/ شماره ۲
صفحات ۱۶۵-۱۸۰

$$L(\theta|y_1,y_2,\dots,y_k) \propto (F(t_1;\theta))^{y_1}(F(t_2;\theta)-F(t_1;\theta))^{y_2} \times \dots \times (F(t_k;\theta)-F(t_{k-1};\theta))^{y_k} \times (1-F(t_k;\theta))^{n-\sum_{i=1}^{k} y_i} \quad (1)$$

خواهد بود. در این حالت، براوردگر حداکثر درستنمایی پارامتر θ فُرم بسته‌ای ندارد و برای محاسبه آن بایستی از روش‌های عددی استفاده کرد. با توجه به این‌که در (۱)، t_i ها زمان‌های نظارت هستند و آزمایش‌گر از زمان شکست واحدهای آزمایشی نیز اطلاعی ندارد، نمایشی دیگر از تابع درستنمایی که به تابع شبه درستنمایی موسوم است به صورت

$$L(\theta|y_1,y_2,\dots,y_k) \propto [t_1 f(t_1^*;\theta)]^{y_1}[(t_2-t_1)f(t_2^*;\theta)]^{y_2} \times \dots \times [(t_k-t_{k-1})f(t_k^*;\theta)]^{y_k} \times [1-F(t_k;\theta)]^{n-\sum_{j=1}^{k} y_j};$$
$$t_i^* \in \left\{t_{i-1}, t_i, \frac{t_{i-1}+t_i}{2}\right\}, i=1,\dots,k$$

در نظر گرفته می‌شود. بنا بر این در روش مبتنی بر تابع شبه درستنمایی که موسوم به روش جانشانی است، به جای زمان‌های شکست در بازه i -ام، t_i^* یعنی نقاط کرانی و یا نقطه میانی آن بازه در نظر گرفته شده و پارامتر توسط

$$\hat{\theta}_{(a)} = \frac{\sum_{i=1}^{k} y_i}{\sum_{i=1}^{k} t_i^* + (n-\sum_{i=1}^{k} y_i)t_k};$$
$$t_i^* \in \left\{t_{i-1}, t_i, \frac{t_{i-1}+t_i}{2}\right\}, i=1,\dots,k$$

براورد می‌شود. میکر [۳] با در نظر گرفتن نقطه میانی بازه‌ها، توزیع تقریبی براوردگر پارامتر را به دست آورده است. یکی از روش‌های براورد پارامتر در حالتی خاص از داده‌های تحت سانسور بازه‌ای نیز توسط ژانگ و همکاران [۴] و به صورت

$$\hat{\theta} = \frac{-\ln(1-\frac{y}{n})}{t}$$

معرفی شده که در آن، y تعداد واحدهای خراب شده تا تنها زمان نظارت t است. از این رو

$$\hat{\theta}_{(b)} = \frac{-\ln(1-\sum_{i=1}^{k} y_i/n)}{t_k}$$

نسخه مرتبط با داده‌های تحت سانسور بازه‌ای با k زمان نظارت خواهد بود که مشابه روش براورد پارامتر در داده‌های تحت سانسور نوع اول، $\sum_{i=1}^{k} y_i$ یعنی تعداد خرابی‌ها تا زمان t_k در نظر گرفته می‌شود ولی در این روش همان‌طور که انتظار می‌رود، اطلاعی از زمان خرابی واحدها در دسترس نیست.

چن و همکاران [۵] بر پایه روش نمودار احتمال،

$$\hat{\theta}_{(c)} = \frac{1}{k}\sum_{j=1}^{k} \frac{-\ln(1-\sum_{r=1}^{j} y_r/n)}{t_j}$$

را برای براورد پارامتر معرفی نموده‌اند. همچنین برای بهبود براوردگرهای $\hat{\theta}_{(a)}$، $\hat{\theta}_{(b)}$ و $\hat{\theta}_{(c)}$، روش‌هایی معرفی شده است که مبتنی بر باز براورد پارامتر و به‌کارگیری روش جانشانی هستند [۶].

یکی از موضوع‌های اساسی در مطالعه داده‌های تحت سانسور بازه‌ای، تعیین بهترین زمان‌های بازرسی است. ساده‌ترین طرح، طرحی است که در آن تمام فاصله‌های زمانی برابر باشند که توسط ناناسی [۷] بررسی شده است. الشراوی و همکاران [۸] دستورالعمل‌هایی برای انتخاب زمان‌های بازرسی کارامد معرفی نموده‌اند و اخیراً نیز روش‌هایی برای انتخاب زمان‌های نظارت بهینه با فاصله‌های زمانی برابر و نابرابر، توسط مالویچ و همکاران [۹] معرفی شده است.

در این مقاله، از سه براوردگر $\hat{\theta}_{(a)}$، $\hat{\theta}_{(b)}$ و $\hat{\theta}_{(c)}$ به عنوان روش‌های موجود و مورد استفاده پژوهشگران نام برده شده است. با این حال براوردگر $\hat{\theta}_{(a)}$ که مبتنی بر روش جانشانی است، عملکردی ضعیف دارد و به‌کارگیری آن منجر به کم‌براوردی نرخ خطر می‌شود و پژوهشگران از این روش کم‌تر استفاده می‌کنند [۱۰]. همچنین به دلیل آن‌که براوردگر حداکثر درستنمایی فرم بسته‌ای ندارد و توزیع آن مشخص نشده است، پژوهشگران در انجام استنباط‌های آماری از براوردگرهای $\hat{\theta}_{(b)}$ و $\hat{\theta}_{(c)}$ استفاده می‌کنند. از طرفی، توزیع روش‌های معرفی شده توسط سجادی و همکاران [۶] نیز تعیین نشده است. از این رو در این مقاله سه براوردگر وزنی معرفی می‌گردد که با $\hat{\theta}_{(1)}$، $\hat{\theta}_{(2)}$ و $\hat{\theta}_{(3)}$ نشان داده می‌شوند و عملکرد این سه روش جدید با دو روش موجود $\hat{\theta}_{(b)}$ و $\hat{\theta}_{(c)}$ که عملکردی بهتر از $\hat{\theta}_{(a)}$ دارند، مقایسه می‌شود. همچنین، توزیع مجانبی یکی از روش‌های جدید که نسبت به سایر روش‌ها عملکرد بهتری دارد، بدست آورده می‌شود.

در بخش دوم، سه براوردگر جدید معرفی و عملکرد آن‌ها در بخش سوم بررسی می‌شود. در بخش چهارم توزیع مجانبی براوردگر با عملکرد بهتر، تعیین می‌گردد و بحث و نتیجه‌گیری نیز در بخش پنجم آورده می‌شود.

۲. معرفی براوردگرهای وزنی

فرض کنید که n واحد آزمایشی تحت سانسور بازه‌ای دارای توزیع نمایی با تابع توزیع

$$F(x;\theta) = 1-\exp(-\theta x);\ x>0, \theta>0$$

هستند و واحدهای آزمایشی در k زمان از پیش تعیین شده $t_1 < \dots < t_k$، مورد بازرسی و نظارت قرار می‌گیرند. اگر $(0,t_1], (t_1,t_2], \dots, (t_{k-1},t_k]$ فاصله‌های بین زمان‌های نظارت

و $y_i;\ i = 1,2,\dots,k$ نیز تعداد واحدهایی باشند که در فاصله زمانی $[t_{i-1}, t_i)$ خراب می‌شوند آن‌گاه با توجه به رابطه بین پارامتر و تابع توزیع نمایی،

$$\hat{\theta}_j = \frac{-\ln\left(1 - \frac{\sum_{i=1}^{j} Y_i}{n}\right)}{t_j};\ j = 1,\dots,k$$

براوردگرهایی مختلف برای پارامتر θ هستند و همان‌طور که نشان داده خواهد شد، به‌طور مجانبی نااریب هستند و واریانس‌هایی متفاوت دارند. همچنین، براوردگرهای موجود $\hat{\theta}_{(b)}$ و $\hat{\theta}_{(c)}$ نیز توابعی از این براوردگرها به شمار می‌روند. براوردگر $\hat{\theta}_{(b)}$، در واقع همان $\hat{\theta}_k$ و $\hat{\theta}_{(c)}$ نیز میانگین $\hat{\theta}_j$ ها است که در واقع حالتی خاص از براوردگر وزنی $\sum_{j=1}^{k} \alpha_j \hat{\theta}_j$ به شمار می‌رود. در روش نمودار احتمال، برای $\hat{\theta}_j$ ها که واریانس‌های متفاوت دارند، وزن‌هایی یکسان در نظر گرفته شده که چندان موجه به نظر نمی‌رسد. از این رو در روش‌های جدید که در این مقاله معرفی می‌شوند، برای $\hat{\theta}_j$ ها وزن‌های متفاوتی در نظر گرفته می‌شود.

به عنوان اولین روش وزنی، ترکیبی خطی از $\hat{\theta}_j$ ها را به صورت $\hat{\theta}_{(1)} = \sum_{j=1}^{k} \alpha_j \hat{\theta}_j$ در نظر می‌گیریم به‌طوری که مجانباً نااریب و دارای کمترین واریانس باشد. با اعمال شرط نااریبی مجانبی و مشتق‌گیری از واریانس $\hat{\theta}_{(1)}$ نسبت به α_j ها می‌توان ضرایب را تعیین نمود. همان‌طور که در بخش چهارم نشان داده می‌شود به ازای n های بزرگ، $\hat{\theta}_j$ ها دارای توزیع نرمال با میانگین θ و

$$Cov(\theta_i, \theta_j) = \frac{F(t_{Min\{i,j\}}; \theta)}{n t_i t_j R(t_{Max\{i,j\}}; \theta)};\quad i,j = 1,\dots,k$$

هستند که در آن

$$R(x;\theta) = 1 - F(x;\theta) = \exp(-\theta x).$$

با اعمال شرط نااریبی مجانبی براوردگر، $\sum_{r=1}^{k} \alpha_r = 1$ و با مشتق‌گیری و انجام محاسبات،

$$\alpha_r = \frac{1}{Var(\hat{\theta}_k - \hat{\theta}_r)} \times \left[Cov\left(\hat{\theta}_k, \hat{\theta}_k - \hat{\theta}_r\right) - \sum_{j=1, j\neq r}^{k-1} \alpha_j Cov\left(\hat{\theta}_k - \hat{\theta}_r, \hat{\theta}_k - \hat{\theta}_j\right)\right];\quad r = 1,\dots,k-1$$

را خواهیم داشت. با توجه به رابطه به دست آمده و حل دستگاه مربوطه، ضرایب وزنی در $\hat{\theta}_{(1)}$ به دست می‌آیند که به صورت توابعی از پارامتر مجهول θ هستند و لذا از براوردگرهای آن‌ها که متغیرهایی تصادفی با فرم‌های پیچیده هستند، استفاده می‌شود. برای مثال، در حالت $k = ۲$، $\alpha_1 = \frac{Cov(\hat{\theta}_2, \hat{\theta}_2 - \hat{\theta}_1)}{Var(\hat{\theta}_2 - \hat{\theta}_1)}$ و براورد آن نیز به صورت $\frac{\frac{y_2}{t_2^2(1-\frac{y_2}{n})} - \frac{y_1}{t_1 t_2(1-\frac{y_2}{n})}}{\frac{y_2}{t_2^2(1-\frac{y_2}{n})} - \frac{2y_1}{t_1 t_2(1-\frac{y_2}{n})} + \frac{y_1}{t_1^2(1-\frac{y_1}{n})}}$ است.

برای معرفی دومین براوردگر وزنی، وابستگی $\hat{\theta}_j$ ها را نادیده می‌گیریم تا ضرایب وزنی فرم ساده‌تری داشته باشند. از این رو از معکوس واریانس مجانبی $\hat{\theta}_j$ ها (و در واقع، معکوس براوردگر آن ها) به عنوان ضرایب وزنی استفاده کرده و دومین براوردگر وزنی به صورت

$$\hat{\theta}_{(2)} = \frac{\sum_{j=1}^{k} \frac{1}{\hat{V}ar(\hat{\theta}_j)} \hat{\theta}_j}{\sum_{j=1}^{k} \frac{1}{\hat{V}ar(\hat{\theta}_j)}} = \frac{\sum_{j=1}^{k} \frac{t_j^2 (1 - \sum_{i=1}^{j} Y_i/n)}{\sum_{i=1}^{j} Y_i} \hat{\theta}_j}{\sum_{j=1}^{k} \frac{t_j^2 (1 - \sum_{i=1}^{j} Y_i/n)}{\sum_{i=1}^{j} Y_i}}$$

معرفی می‌شود که به دلیل به کارگیری براورد ضرایب وزنی، این روش به صورت مجموع براوردگرهایی وابسته خواهد بود. در این روش، ضرایب وزنی فرم ساده‌تری دارند ولی اگر در بازه‌های نظارت اولیه هیچ خرابی مشاهده نشود، قابل استفاده نیستند. حال با توجه به این که اگر n عددی بزرگ و آخرین زمان نظارت نیز برابر با $\frac{\pi}{۲\theta}$ باشد آنگاه $\hat{\theta}_k$ کمترین واریانس ممکن را خواهد داشت [۶]، از این رو می‌توان نتیجه گرفت که اگر زمان‌های نظارت $t_1 < \cdots < t_k$ در فاصله $(0, \frac{\pi}{۲\theta})$ انتخاب شده باشند، آن‌گاه

$$Var(\hat{\theta}_1) > Var(\hat{\theta}_2) > \cdots > Var(\hat{\theta}_k)$$

برقرار است. بنا بر این، سومین براوردگر وزنی را به صورت

$$\hat{\theta}_{(3)} = \frac{\sum_{j=1}^{k} \sum_{r=1}^{j} Y_r \hat{\theta}_j}{\sum_{j=1}^{k} \sum_{r=1}^{j} Y_r}$$

در نظر می‌گیریم تا ضرایب وزنی کوچک‌تر به $\hat{\theta}_j$ های با واریانس بزرگ‌تر نسبت داده شود و هم این‌که فرم ساده‌تری نسبت به دو روش قبلی داشته باشد.

۳. مطالعه شبیه‌سازی

در این بخش، یک مطالعه شبیه‌سازی به منظور مقایسه و بررسی عملکرد دو براوردگر سابق $\hat{\theta}_{(b)}$ و $\hat{\theta}_{(c)}$ و همچنین سه براوردگر وزنی جدید ارائه می‌گردد.

ملاحظه ۳-۱. اگر تا زمان t_k تمام خرابی‌ها رخ داده باشند آن‌گاه هیچ‌یک از براوردگرها قابل تعریف نخواهند بود. از این رو به جای $\hat{\theta}_j = \frac{-ln(1-\sum_{i=1}^{j} Y_i/n)}{t_j}$، از فرم اصلاح شده $\hat{\theta}_j = \frac{-ln(1-\sum_{i=1}^{j} Y_i/n+1)}{t_j}$ استفاده خواهیم کرد.

ملاحظه ۳-۲. در انجام شبیه‌سازی، برای پرهیز از حجیم شدن مقایسات، مقادیر معدودی برای پارامتر و تعداد دفعات نظارت به صورت $\theta = ۰/۰۳, ۰/۰۵, ۰/۰۷$ و $k = ۲,۳,۴$ در نظر

می‌گیریم. برای اندازه نمونه نیز مقادیر نسبتاً بزرگ $n =$ ۳۰،۴۰،۵۰،۶۰ در نظر گرفته می‌شود زیرا در مقاله به توزیع و رفتار مجانبی براوردگرهای مورد بررسی اشاره و استناد شده است.

ملاحظه ۳-۳. اگر $k =$ ۲ باشد و برای تعیین اولین زمان نظارت، میانه توزیع در نظر گرفته شود آن‌گاه به ازای سه مقدار پارامتر، مقادیر مختلف

$$F^{-1}(0/5;0/03) = \frac{-Ln(1-0/5)}{0/03} \simeq 23/1,$$
$$F^{-1}(0/5;0/05) \simeq 13/86\,,\ F^{-1}(0/5;0/07) \simeq 9/9$$

را خواهیم داشت. همان‌طور که نتایج شبیه‌سازی نشان می‌دهند، زمان‌های نظارت و مقدار پارامتر توأماً بر عملکرد براوردگرها تأثیرگذار هستند. از این رو برای هر سه مقدار پارامتر، میانگین میانه‌ها به عنوان اولین زمان نظارت و میانگین صدک‌های ۸۵ را به عنوان دومین زمان نظارت در نظر می‌گیریم.

بنا بر این برای $k =$ ۲،

$$t_1 = \frac{F^{-1}(0/5;0/03) + F^{-1}(0/5;0/05) + F^{-1}(0/5;0/07)}{3} \simeq 16$$
$$t_2 = \frac{F^{-1}(0/85;0/03) + F^{-1}(0/85;0/05) + F^{-1}(0/85;0/07)}{3} \simeq 42$$

اولین و دومین زمان نظارت خواهند بود که نتایج شبیه‌سازی نیز در جدول‌های ۱ تا ۳ آمده است. لازم به ذکر است که (۱۶،۴۲) برای $\theta =$ ۰/۰۷، ۰/۰۵، ۰/۰۳ به ترتیب برابر با صدک‌های (۳۸،۷۲)، (۵۵،۸۸) و (۶۷،۹۵) است. زمان‌های نظارت برای $k =$ ۳ و $k =$ ۴ نیز به ترتیب به صورت

$$t_1 = \frac{F^{-1}(0/4;0/03)+F^{-1}(0/4;0/05)+F^{-1}(0/4;0/07)}{3} \simeq 12$$
$$t_2 = \frac{F^{-1}(0/6;0/03) + F^{-1}(0/6;0/05) + F^{-1}(0/6;0/07)}{3} \simeq 21$$
$$t_3 = \frac{F^{-1}(0/8;0/03) + F^{-1}(0/8;0/05) + F^{-1}(0/8;0/07)}{3} \simeq 36$$

و

$$t_1 = \frac{F^{-1}(0/2;0/03) + F^{-1}(0/2;0/05) + F^{-1}(0/2;0/07)}{3} \simeq 5$$
$$t_2 = \frac{F^{-1}(0/35;0/03) + F^{-1}(0/35;0/05) + F^{-1}(0/35;0/07)}{3} \simeq 10$$
$$t_3 = \frac{F^{-1}(0/65;0/03) + F^{-1}(0/65;0/05) + F^{-1}(0/65;0/07)}{3} \simeq 24$$
$$t_4 = \frac{F^{-1}(0/80;0/03) + F^{-1}(0/80;0/05) + F^{-1}(0/80;0/07)}{3} \simeq 36$$

در نظر گرفته می‌شوند تا بررسی عملکرد براوردگرها تحت تأثیر زمان نظارت مشخصی و یا مقدار خاصی از پارامتر قرار نگیرد و مقایسه آن‌ها به درستی انجام شود.

بر پایه انجام شبیه‌سازی با $N =$ ۱۰۰۰۰ بار تکرار، $Mean$ میانگین براوردگرها و $MSE(\hat{\theta})$ میانگین مربعات خطای آن‌ها و همچنین، $E(\hat{\theta}_{(c)},\hat{\theta})$ کارایی براوردگرها نسبت به $\hat{\theta}_{(c)}$ محاسبه شده که نتایج در جدول‌های زیر آمده است. ضمناً جهت بررسی تأثیر مقدار پارامتر و زمان‌های نظارت بر عملکرد براوردگرها، $\frac{\pi}{۲\theta}$ و $\bar{p}_1,\dots,\bar{p}_4$ محاسبه شده که در آن $\bar{p}_i$ میانگین نسبت خرابی‌ها تا $i-$امین زمان نظارت در ۱۰۰۰۰ بار تکرار است.

جدول ۱. مقایسه عملکرد براوردگرها به ازای $t_1 = 16\,, t_2 = 42\,, k = 2\,, \frac{\pi}{۲\theta} = 52\,, \theta = 0/03$

$\hat{\theta}$	n	$(\bar{p}_1\,,\bar{p}_2)$	$Mean$	$MSE(\hat{\theta})$	$E(\hat{\theta}_{(c)}\,,\hat{\theta})$
$\hat{\theta}_{(c)}$	۳۰	(۰/۳۸،۰/۷۲)	۰/۰۲۹۲۱۱	۴/۵۲۶۱e − ۰۵	۱
$\hat{\theta}_{(b)}$	۳۰	(۰/۳۸،۰/۷۲)	۰/۰۲۹۰۶۵	۴/۴۴۷۹e − ۰۵	۱/۰۱۸
$\hat{\theta}_{(1)}$	۳۰	(۰/۳۸،۰/۷۲)	۰/۰۲۷۵۷۹	۴/۶۴۲۲e − ۰۵	۰/۹۷۵
$\hat{\theta}_{(2)}$	۳۰	(۰/۳۸،۰/۷۲)	۰/۰۲۸۴۸۰	۴/۲۴۲۱e − ۰۵	۱/۰۶۷
$\hat{\theta}_{(3)}$	۳۰	(۰/۳۸،۰/۷۲)	۰/۰۲۹۵۱۸	۴/۱۷۸۱e − ۰۵	۱/۰۸۳
$\hat{\theta}_{(c)}$	۴۰	/۳۸،۰/۷۲)	۰/۰۲۹۵۰۹	۳/۵۰۲۷e − ۰۵	۱
$\hat{\theta}_{(b)}$	۴۰	(۰/۳۸،۰/۷۲)	۰/۰۲۹۳۳۶	۳/۴۶۰۹e − ۰۵	۱/۰۱۲
$\hat{\theta}_{(1)}$	۴۰	(۰/۳۸،۰/۷۲)	۰/۰۲۸۲۳۷	۳/۵۲۰۰e − ۰۵	۰/۹۹۲
$\hat{\theta}_{(2)}$	۴۰	(۰/۳۸،۰/۷۲)	۰/۰۲۸۹۴۱	۳/۲۷۶۹e − ۰۵	۱/۰۶۹
$\hat{\theta}_{(3)}$	۴۰	(۰/۳۸،۰/۷۲)	۰/۰۲۹۷۲۱	۳/۲۴۵۲e − ۰۵	۱/۰۷۹
$\hat{\theta}_{(c)}$	۵۰	(۰/۳۸،۰/۷۲)	۰/۰۲۹۵۴۴	۲/۷۶۶۴e − ۰۵	۱
$\hat{\theta}_{(b)}$	۵۰	(۰/۳۸،۰/۷۲)	۰/۰۲۹۴۲۷	۲/۷۹۵۴e − ۰۵	۰/۹۹۰
$\hat{\theta}_{(1)}$	۵۰	(۰/۳۸،۰/۷۲)	۰/۰۲۸۵۴۳	۲/۷۶۷۰e − ۰۵	۱
$\hat{\theta}_{(2)}$	۵۰	(۰/۳۸،۰/۷۲)	۰/۰۲۹۰۹۷	۲/۶۱۰۲e − ۰۵	۱/۰۶۰
$\hat{\theta}_{(3)}$	۵۰	(۰/۳۸،۰/۷۲)	۰/۰۲۹۷۱۶	۲/۵۷۵۳e − ۰۵	۱/۰۷۴
$\hat{\theta}_{(c)}$	۶۰	(۰/۳۸،۰/۷۲)	۰/۰۲۹۵۹۸	۲/۳۱۲۳e − ۰۵	۱
$\hat{\theta}_{(b)}$	۶۰	(۰/۳۸،۰/۷۲)	۰/۰۲۹۴۸۸	۲/۳۲۴۸e − ۰۵	۰/۹۹۵
$\hat{\theta}_{(1)}$	۶۰	(۰/۳۸،۰/۷۲)	۰/۰۲۸۷۵۱	۲/۲۶۹۳e − ۰۵	۱/۰۱۹
$\hat{\theta}_{(2)}$	۶۰	(۰/۳۸،۰/۷۲)	۰/۰۲۹۲۲۰	۲/۱۷۰۵e − ۰۵	۱/۰۶۵
$\hat{\theta}_{(3)}$	۶۰	(۰/۳۸،۰/۷۲)	۰/۰۲۹۷۴۰	۲/۱۴۹۱e − ۰۵	۱/۰۷۶

جدول ۲. مقایسه عملکرد براوردگرها به ازای $t_1 = 16\ , t_2 = 42\ , k = 2\ , \frac{\pi}{۲\theta} = 31\ , \theta = 0/05$

$\hat{\theta}$	n	$(\bar{p}_1\ , \bar{p}_2)$	$Mean$	$MSE(\hat{\theta})$	$E(\ \hat{\theta}_{(c)}\ , \hat{\theta}\ )$
$\hat{\theta}_{(c)}$	۳۰	(۰/۵۵,۰/۸۸)	۰/۰۴۷۹۶۹	۹/۴۲۶۶e − ۰۵	۱
$\hat{\theta}_{(b)}$	۳۰	(۰/۵۵,۰/۸۸)	۰/۰۴۷۰۹۵	۱/۱۵۸۹e − ۰۴	۰/۸۱۳
$\hat{\theta}_{(1)}$	۳۰	(۰/۵۵,۰/۸۸)	۰/۰۴۵۱۷۵	۱/۰۹۹۱۴e − ۰۴	۰/۸۵۸
$\hat{\theta}_{(2)}$	۳۰	(۰/۵۵,۰/۸۸)	۰/۰۴۶۵۰۵	۹/۴۲۲۶e − ۰۵	۱/۰۰۱
$\hat{\theta}_{(3)}$	۳۰	(۰/۵۵,۰/۸۸)	۰/۰۴۸۱۶۸	۹/۲۲۶۳e − ۰۵	۱/۰۱۱
$\hat{\theta}_{(c)}$	۴۰	(۰/۵۵,۰/۸۸)	۰/۰۴۸۳۲۲	۷/۲۳۷۹e − ۰۵	۱
$\hat{\theta}_{(b)}$	۴۰	(۰/۵۵,۰/۸۸)	۰/۰۴۷۶۷۶	۸/۸۹۷۱e − ۰۵	۰/۸۱۴
$\hat{\theta}_{(1)}$	۴۰	(۰/۵۵,۰/۸۸)	۰/۰۴۶۲۱۰	۸/۰۵۰۴e − ۰۵	۰/۸۹۹
$\hat{\theta}_{(2)}$	۴۰	(۰/۵۵,۰/۸۸)	۰/۰۴۷۱۹۶	۷/۱۹۲۲e − ۰۵	۱/۰۰۶
$\hat{\theta}_{(3)}$	۴۰	(۰/۵۵,۰/۸۸)	۰/۰۴۸۴۷۹	۷/۰۸۲۴e − ۰۵	۱/۰۲۲
$\hat{\theta}_{(c)}$	۵۰	(۰/۵۵,۰/۸۸)	۰/۰۴۸۷۰۵	۵/۷۷۵۸۲e − ۰۵	۱
$\hat{\theta}_{(b)}$	۵۰	(۰/۵۵,۰/۸۸)	۰/۰۴۸۱۷۷	۷/۵۰۹۲۲e − ۰۵	۰/۷۶۹
$\hat{\theta}_{(1)}$	۵۰	(۰/۵۵,۰/۸۸)	۰/۰۴۶۹۷۷	۶/۳۸۴۰۶e − ۰۵	۰/۹۰۵
$\hat{\theta}_{(2)}$	۵۰	(۰/۵۵,۰/۸۸)	۰/۰۴۷۷۷۰	۵/۷۶۲۹۳e − ۰۵	۱/۰۰۲
$\hat{\theta}_{(3)}$	۵۰	(۰/۵۵,۰/۸۸)	۰/۰۴۸۸۲۴	۵/۷۲۱۹۵e − ۰۵	۱/۰۱۰
$\hat{\theta}_{(c)}$	۶۰	(۰/۵۵,۰/۸۸)	۰/۰۴۹۰۹۰	۵/۰۲۷۹۳e − ۰۵	۱
$\hat{\theta}_{(b)}$	۶۰	(۰/۵۵,۰/۸۸)	۰/۰۴۸۶۱۴	۶/۳۵۰۶۲e − ۰۵	۰/۷۹۲
$\hat{\theta}_{(1)}$	۶۰	(۰/۵۵,۰/۸۸)	۰/۰۴۷۶۳۸	۵/۳۱۰۲e − ۰۵	۰/۹۴۷
$\hat{\theta}_{(2)}$	۶۰	(۰/۵۵,۰/۸۸)	۰/۰۴۸۲۹۸	۴/۹۴۱۹e − ۰۵	۱/۰۱۷
$\hat{\theta}_{(3)}$	۶۰	(۰/۵۵,۰/۸۸)	۰/۰۴۹۱۸۵	۴/۹۶۶۸e − ۰۵	۱/۰۱۲

جدول ۳. مقایسه عملکرد براوردگرها به ازای $t_1 = 16\ , t_2 = 42\ , k = 2\ , \frac{\pi}{۲\theta} = 22\ , \theta = 0/07$

$\hat{\theta}$	n	$(\bar{p}_1\ , \bar{p}_2)$	$Mean$	$MSE(\hat{\theta})$	$E(\ \hat{\theta}_{(c)}\ , \hat{\theta}\ )$
$\hat{\theta}_{(c)}$	۳۰	(۰/۶۷,۰/۹۵)	۰/۰۶۴۸۰۶	۱/۵۴۱۴e − ۰۴	۱
$\hat{\theta}_{(b)}$	۳۰	(۰/۶۷,۰/۹۵)	۰/۰۶۲۰۱۰	۲/۰۷۷۶e − ۰۴	۰/۷۴۲
$\hat{\theta}_{(1)}$	۳۰	(۰/۶۷,۰/۹۵)	۰/۰۶۱۰۳۹	۲/۲۰۳۸e − ۰۴	۰/۶۹۹
$\hat{\theta}_{(2)}$	۳۰	(۰/۶۷,۰/۹۵)	۰/۰۶۲۷۴۴	۱/۷۷۴۸e − ۰۴	۰/۸۶۹
$\hat{\theta}_{(3)}$	۳۰	(۰/۶۷,۰/۹۵)	۰/۰۶۴۷۵۹	۱/۵۱۶۵e − ۰۴	۱/۰۱۶
$\hat{\theta}_{(c)}$	۴۰	(۰/۶۷,۰/۹۵)	۰/۰۶۵۹۹۹	۱/۲۵۰۳e − ۰۴	۱
$\hat{\theta}_{(b)}$	۴۰	(۰/۶۷,۰/۹۵)	۰/۰۶۳۹۷۰	۱/۷۶۰۳e − ۰۴	۰/۷۱۰
$\hat{\theta}_{(1)}$	۴۰	(۰/۶۷,۰/۹۵)	۰/۰۶۳۱۰۵	۱/۶۲۲۵e − ۰۴	۰/۷۷۱
$\hat{\theta}_{(2)}$	۴۰	(۰/۶۷,۰/۹۵)	۰/۰۶۴۲۷۹	۱/۳۸۶۵e − ۰۴	۰/۹۰۲
$\hat{\theta}_{(3)}$	۴۰	(۰/۶۷,۰/۹۵)	۰/۰۶۵۹۸۰	۱/۲۴۷۴e − ۰۴	۱/۰۰۲
$\hat{\theta}_{(c)}$	۵۰	(۰/۶۷,۰/۹۵)	۰/۰۶۷۱۰۸	۱/۰۴۸۴e − ۰۴	۱
$\hat{\theta}_{(b)}$	۵۰	(۰/۶۷,۰/۹۵)	۰/۰۶۵۳۹۸	۱/۵۴۱۴e − ۰۴	۰/۶۸۰
$\hat{\theta}_{(1)}$	۵۰	(۰/۶۷,۰/۹۵)	۰/۰۶۴۷۰۳	۱/۲۷۵۹e − ۰۴	۰/۸۲۲
$\hat{\theta}_{(2)}$	۵۰	(۰/۶۷,۰/۹۵)	۰/۰۶۵۶۳۲	۱/۱۲۸۰e − ۰۴	۰/۹۲۹
$\hat{\theta}_{(3)}$	۵۰	(۰/۶۷,۰/۹۵)	۰/۰۶۷۰۸۴	۱/۰۵۷۷e − ۰۴	۰/۹۹۱
$\hat{\theta}_{(c)}$	۶۰	(۰/۶۷,۰/۹۵)	۰/۰۶۷۵۶۵	۸/۹۴۰۴e − ۰۵	۱
$\hat{\theta}_{(b)}$	۶۰	(۰/۶۷,۰/۹۵)	۰/۰۶۶۲۱۸	۱/۳۷۵۹e − ۰۴	۰/۶۵۰
$\hat{\theta}_{(1)}$	۶۰	(۰/۶۷,۰/۹۵)	۰/۰۶۵۵۵۲	۱/۰۴۴۷e − ۰۴	۰/۸۵۹
$\hat{\theta}_{(2)}$	۶۰	(۰/۶۷,۰/۹۵)	۰/۰۶۶۲۸۷	۹/۴۰۳۵e − ۰۵	۰/۹۵۱
$\hat{\theta}_{(3)}$	۶۰	(۰/۶۷,۰/۹۵)	۰/۰۶۷۵۵۰	۹/۱۲۹۴e − ۰۵	۰/۹۷۹

با توجه به این‌که کم براوردی نرخ خطر یکی از ضعف‌های یک براوردگر به شمار می‌رود و کم براوردی نرخ خطر در توزیع نمایی باعث بیش براوردی میانگین می‌شود، بایستی علاوه بر کارایی براوردگرها به این شاخص نیز توجه ویژه‌ای داشت. همان‌طور که

نتایج شبیه‌سازی در جدول‌های ۱ تا ۳ نشان می‌دهند، عملکرد $\hat{\theta}_{(b)}$ و $\hat{\theta}_{(c)}$ بستگی زیادی به مقدار زمان‌های نظارت t_1 و t_2 دارد و حتی در بعضی موارد عملکرد $\hat{\theta}_{(b)}$ بهتر است. ضمناً برخلاف انتظار، $\hat{\theta}_{(1)}$ عملکرد خوبی ندارد زیرا به جای ضرایب وزنی از براورد آن‌ها استفاده می‌شود. همچنین $\hat{\theta}_{(3)}$ نسبت به اولین و دومین روش وزنی $\hat{\theta}_{(1)}$ و $\hat{\theta}_{(2)}$ عملکرد بهتری دارد، به ویژه اگر زمان‌های نظارت در فاصله $(0,\frac{\pi}{۲\theta})$ انتخاب شده باشند.

بنا بر این با توجه به نتایج بدست آمده و این‌که تعیین ضرایب وزنی و تعیین توزیع مجانبی $\hat{\theta}_{(1)}$ به آسانی میسر نیست، به نظر می‌رسد که ادامه بررسی عملکرد $\hat{\theta}_{(1)}$ و $\hat{\theta}_{(2)}$ ضروری نیست. از این رو در بررسی‌های بعدی و در جدول‌های ۴ تا ۹، دو روش $\hat{\theta}_{(b)}$ و $\hat{\theta}_{(c)}$ فقط با سومین روش وزنی که عملکرد بهتری داشته است، مقایسه می‌شود

.

جدول ۴. مقایسه عملکرد براوردگرها به ازای $t_1 = 12\,, t_2 = 21\,, t_3 = 36\,, k = 3\,, \frac{\pi}{۲\theta} = 52\,, \theta = 0/03$

$\hat{\theta}$	n	$(\bar{p}_1, \bar{p}_2, \bar{p}_3)$	$Mean$	$MSE(\hat{\theta})$	$E(\hat{\theta}_{(c)}, \hat{\theta})$
$\hat{\theta}_{(c)}$	۳۰	(۰/۳۰,۰/۴۷,۰/۶۶)	۰/۰۲۹۲۹۶	۵/۰۹۴۷e − ۰۵	۱
$\hat{\theta}_{(b)}$	۳۰	(۰/۳۰,۰/۴۷,۰/۶۶)	۰/۰۲۹۱۳۴	۴/۶۹۹۶e − ۰۵	۱/۰۸۴
$\hat{\theta}_{(3)}$	۳۰	(۰/۳۰,۰/۴۷,۰/۶۶)	۰/۰۲۹۶۸۶	۴/۶۵۱۵e − ۰۵	۱/۰۹۵
$\hat{\theta}_{(c)}$	۴۰	(۰/۳۰,۰/۴۷,۰/۶۶)	۰/۰۲۹۳۸۲	۳/۸۲۹۰e − ۰۵	۱
$\hat{\theta}_{(b)}$	۴۰	(۰/۳۰,۰/۴۷,۰/۶۶)	۰/۰۲۹۲۵۴	۳/۵۶۴۷e − ۰۵	۱/۰۷۴
$\hat{\theta}_{(3)}$	۴۰	(۰/۳۰,۰/۴۷,۰/۶۶)	۰/۰۲۹۶۸۲	۳/۴۹۱۷e − ۰۵	۱/۰۹۷
$\hat{\theta}_{(c)}$	۵۰	(۰/۳۰,۰/۴۷,۰/۶۶)	۰/۰۲۹۶۰۴	۳/۰۵۴۴e − ۰۵	۱
$\hat{\theta}_{(b)}$	۵۰	(۰/۳۰,۰/۴۷,۰/۶۶)	۰/۰۲۹۵۳۰	۲/۹۴۸۳e − ۰۵	۱/۰۳۶
$\hat{\theta}_{(3)}$	۵۰	(۰/۳۰,۰/۴۷,۰/۶۶)	۰/۰۲۹۸۴۷	۲/۸۰۶۱e − ۰۵	۱/۰۸۹
$\hat{\theta}_{(c)}$	۶۰	(۰/۳۰,۰/۴۷,۰/۶۶)	۰/۰۲۹۵۷۱	۲/۶۱۹۵e − ۰۵	۱
$\hat{\theta}_{(b)}$	۶۰	(۰/۳۰,۰/۴۷,۰/۶۶)	۰/۰۲۹۵۳۸	۲/۴۷۷۹e − ۰۵	۱/۰۵۷
$\hat{\theta}_{(3)}$	۶۰	(۰/۳۰,۰/۴۷,۰/۶۶)	۰/۰۲۹۷۸۳	۲/۳۷۷۷e − ۰۵	۱/۱۰۲

جدول ۵. مقایسه عملکرد براوردگرها به ازای $t_1 = 12\,, t_2 = 21\,, t_3 = 36\,, k = 3\,, \frac{\pi}{۲\theta} = 31\,, \theta = 0/05$

$\hat{\theta}$	n	$(\bar{p}_1, \bar{p}_2, \bar{p}_3)$	$Mean$	$MSE(\hat{\theta})$	$E(\hat{\theta}_{(c)}, \hat{\theta})$
$\hat{\theta}_{(c)}$	۳۰	(۰/۴۵,۰/۶۵,۰/۸۴)	۰/۰۴۸۳۵۸	۹/۴۲۶۶e − ۰۵	۱
$\hat{\theta}_{(b)}$	۳۰	(۰/۴۵,۰/۶۵,۰/۸۴)	۰/۰۴۷۷۰۷	۱/۱۵۸۹e − ۰۴	۰/۸۱۳
$\hat{\theta}_{(3)}$	۳۰	(۰/۴۵,۰/۶۵,۰/۸۴)	۰/۰۴۸۷۲۳	۹/۲۲۶۳e − ۰۵	۱/۰۲۲
$\hat{\theta}_{(c)}$	۴۰	(۰/۴۵,۰/۶۵,۰/۸۴)	۰/۰۴۸۸۰۷	۷/۷۰۹۱e − ۰۵	۱
$\hat{\theta}_{(b)}$	۴۰	(۰/۴۵,۰/۶۵,۰/۸۴)	۰/۰۴۸۲۸۷	۹/۱۳۸۱e − ۰۵	۰/۸۴۴
$\hat{\theta}_{(3)}$	۴۰	(۰/۴۵,۰/۶۵,۰/۸۴)	۰/۰۴۹۰۸۴	۷/۴۱۷۹e − ۰۵	۱/۰۳۹
$\hat{\theta}_{(c)}$	۵۰	(۰/۴۵,۰/۶۵,۰/۸۴)	۰/۰۴۹۰۲۳	۶/۲۱۱۳۴e − ۰۵	۱
$\hat{\theta}_{(b)}$	۵۰	(۰/۴۵,۰/۶۵,۰/۸۴)	۰/۰۴۸۵۹۹	۷/۲۸۱۱۱e − ۰۵	۰/۸۵۳
$\hat{\theta}_{(3)}$	۵۰	(۰/۴۵,۰/۶۵,۰/۸۴)	۰/۰۴۹۲۳۷	۵/۹۴۵۵۷e − ۰۵	۱/۰۴۵
$\hat{\theta}_{(c)}$	۶۰	(۰/۴۵,۰/۶۵,۰/۸۴)	۰/۰۴۹۱۹۶	۵/۲۸۴۳۶e − ۰۵	۱
$\hat{\theta}_{(b)}$	۶۰	(۰/۴۵,۰/۶۵,۰/۸۴)	۰/۰۴۸۸۹۸	۶/۲۶۲۰۸e − ۰۵	۰/۸۴۴
$\hat{\theta}_{(3)}$	۶۰	(۰/۴۵,۰/۶۵,۰/۸۴)	۰/۰۴۹۳۹۶	۵/۰۶۰۱۵e − ۰۵	۱/۰۴۴

جدول ۶. مقایسه عملکرد براوردگرها به ازای $\theta = 0/07$, $\frac{\pi}{۲\theta} = 22$, $t_1 = 12\,, t_2 = 21\,, t_3 = 36\,, k = 3$

$\hat{\theta}$	n	$(\bar{p}_1, \bar{p}_2, \bar{p}_3)$	$Mean$	$MSE(\hat{\theta})$	$E(\hat{\theta}_{(c)}, \hat{\theta})$
$\hat{\theta}_{(c)}$	۳۰	(۰/۵۷,۰/۷۷,۰/۹۲)	۰/۰۶۶۶۲۱	۱/۶۲۳۸e − ۰۴	۱
$\hat{\theta}_{(b)}$	۳۰	(۰/۵۷,۰/۷۷,۰/۹۲)	۰/۰۶۴۲۶۷	۲/۱۱۸۹e − ۰۴	۰/۷۶۶
$\hat{\theta}_{(3)}$	۳۰	(۰/۵۷,۰/۷۷,۰/۹۲)	۰/۰۶۶۸۴۱	۱/۵۷۷۸e − ۰۴	۱/۰۲۹
$\hat{\theta}_{(c)}$	۴۰	(۰/۵۷,۰/۷۷,۰/۹۲)	۰/۰۶۷۳۶۴	۱/۳۲۳۴e − ۰۴	۱
$\hat{\theta}_{(b)}$	۴۰	(۰/۵۷,۰/۷۷,۰/۹۲)	۰/۰۶۵۴۲۳	۱/۸۲۷۵e − ۰۴	۰/۷۲۴
$\hat{\theta}_{(3)}$	۴۰	(۰/۵۷,۰/۷۷,۰/۹۲)	۰/۰۶۷۵۱۰	۱/۳۰۵۶e − ۰۴	۱/۰۱۴
$\hat{\theta}_{(c)}$	۵۰	(۰/۵۷,۰/۷۷,۰/۹۲)	۰/۰۶۸۰۲۵	۱/۰۳۹۰e − ۰۴	۱
$\hat{\theta}_{(b)}$	۵۰	(۰/۵۷,۰/۷۷,۰/۹۲)	۰/۰۶۶۷۶۱	۱/۵۵۹۲e − ۰۴	۰/۶۶۶
$\hat{\theta}_{(3)}$	۵۰	(۰/۵۷,۰/۷۷,۰/۹۲)	۰/۰۶۸۱۸۱	۱/۰۳۳۸e − ۰۴	۱/۰۰۵
$\hat{\theta}_{(c)}$	۶۰	(۰/۵۷,۰/۷۷,۰/۹۲)	۰/۰۶۸۴۴۱	۸/۸۱۴۴e − ۰۵	۱
$\hat{\theta}_{(b)}$	۶۰	(۰/۵۷,۰/۷۷,۰/۹۲)	۰/۰۶۷۲۶۳	۱/۲۸۱۸۸e − ۰۴	۰/۶۸۸
$\hat{\theta}_{(3)}$	۶۰	(۰/۵۷,۰/۷۷,۰/۹۲)	۰/۰۶۸۵۵۲	۸/۷۴۵۶۷e − ۰۵	۱/۰۰۸

جدول۷. مقایسه عملکرد براوردگرها به ازای $\theta = 0/03$, $\frac{\pi}{۲\theta} = 52$, $t_1 = 5, t_2 = 10, t_3 = 24, t_4 = 36, k = 4,$

$\hat{\theta}$	n	$(\bar{p}_1, \bar{p}_2, \bar{p}_3, \bar{p}_4)$	$Mean$	$MSE(\hat{\theta})$	$E(\hat{\theta}_{(c)}, \hat{\theta})$
$\hat{\theta}_{(c)}$	۳۰	(۰/۱۴,۰/۲۶,۰/۵۱,۰/۶۶)	۰/۰۲۹۲۸۵	۶/۴۲۴۲e − ۰۵	۱
$\hat{\theta}_{(b)}$	۳۰	(۰/۱۴,۰/۲۶,۰/۵۱,۰/۶۶)	۰/۰۲۹۱۲۱	۴/۷۵۲۱e − ۰۵	۱/۳۵۲
$\hat{\theta}_{(3)}$	۳۰	(۰/۱۴,۰/۲۶,۰/۵۱,۰/۶۶)	۰/۰۳۰۰۵۶	۴/۸۱۶۰e − ۰۵	۱/۳۳۴
$\hat{\theta}_{(c)}$	۴۰	(۰/۱۴,۰/۲۶,۰/۵۱,۰/۶۶)	۰/۰۲۹۴۸۵	۴/۷۱۵۸e − ۰۵	۱
$\hat{\theta}_{(b)}$	۴۰	(۰/۱۴,۰/۲۶,۰/۵۱,۰/۶۶)	۰/۰۲۹۲۵۷	۳/۵۹۸۵e − ۰۵	۱/۳۱۰
$\hat{\theta}_{(3)}$	۴۰	(۰/۱۴,۰/۲۶,۰/۵۱,۰/۶۶)	۰/۰۳۰۰۰۷	۳/۵۵۸۲e − ۰۵	۱/۳۲۵
$\hat{\theta}_{(c)}$	۵۰	(۰/۱۴,۰/۲۶,۰/۵۱,۰/۶۶)	۰/۰۲۹۵۶۸	۳/۸۳۸۱e − ۰۵	۱
$\hat{\theta}_{(b)}$	۵۰	(۰/۱۴,۰/۲۶,۰/۵۱,۰/۶۶)	۰/۰۲۹۳۵۵	۲/۹۳۸۷e − ۰۵	۱/۳۰۶
$\hat{\theta}_{(3)}$	۵۰	(۰/۱۴,۰/۲۶,۰/۵۱,۰/۶۶)	۰/۰۲۹۹۸۵	۲/۹۰۵۴e − ۰۵	۱/۳۲۱
$\hat{\theta}_{(c)}$	۶۰	(۰/۱۴,۰/۲۶,۰/۵۱,۰/۶۶)	۰/۰۲۹۶۱۳	۳/۲۳۴۶e − ۰۵	۱
$\hat{\theta}_{(b)}$	۶۰	(۰/۱۴,۰/۲۶,۰/۵۱,۰/۶۶)	۰/۰۲۹۵۸۰	۲/۴۱۱۴e − ۰۵	۱/۳۴۱
$\hat{\theta}_{(3)}$	۶۰	(۰/۱۴,۰/۲۶,۰/۵۱,۰/۶۶)	۰/۰۳۰۰۳۲	۲/۴۰۵۷e − ۰۵	۱/۳۴۵

جدول ۸. مقایسه عملکرد براوردگرها به ازای $t_1 = 5, t_2 = 10, t_3 = 24, t_4 = 36, k = 4, \frac{\pi}{۲\theta} = 31, \theta = 0/05$

$\hat{\theta}$	n	$(\bar{p}_1, \bar{p}_2, \bar{p}_3, \bar{p}_4)$	$Mean$	$MSE(\hat{\theta})$	$E(\hat{\theta}_{(c)}, \hat{\theta})$
$\hat{\theta}_{(c)}$	۳۰	(۰/۲۲،۰/۳۹،۰/۷۰،۰/۸۴)	۰/۰۴۸۶۵۲	۱/۱۸۹۶e − ۰۴	۱
$\hat{\theta}_{(b)}$	۳۰	(۰/۲۲،۰/۳۹،۰/۷۰،۰/۸۴)	۰/۰۴۷۷۳۸	۱/۱۹۷۰e − ۰۴	۰/۹۹۴
$\hat{\theta}_{(3)}$	۳۰	(۰/۲۲،۰/۳۹،۰/۷۰،۰/۸۴)	۰/۰۴۹۳۳۶	۹/۹۷۱۴e − ۰۵	۱/۱۹۳
$\hat{\theta}_{(c)}$	۴۰	(۰/۲۲،۰/۳۹،۰/۷۰،۰/۸۴)	۰/۰۴۸۷۹۱	۸/۷۸۲۵e − ۰۵	۱
$\hat{\theta}_{(b)}$	۴۰	(۰/۲۲،۰/۳۹،۰/۷۰،۰/۸۴)	۰/۰۴۷۹۶۰	۸/۷۴۴۸e − ۰۵	۱/۰۰۴
$\hat{\theta}_{(3)}$	۴۰	(۰/۲۲،۰/۳۹،۰/۷۰،۰/۸۴)	۰/۰۴۹۲۵۹	۷/۲۴۱۲e − ۰۵	۱/۲۱۳
$\hat{\theta}_{(c)}$	۵۰	(۰/۲۲،۰/۳۹،۰/۷۰،۰/۸۴)	۰/۰۴۹۱۵۴	۷/۳۴۳۸e − ۰۵	۱
$\hat{\theta}_{(b)}$	۵۰	(۰/۲۲،۰/۳۹،۰/۷۰،۰/۸۴)	۰/۰۴۸۵۷۳	۷/۵۵۲۴e − ۰۵	۰/۹۷۲
$\hat{\theta}_{(3)}$	۵۰	(۰/۲۲،۰/۳۹،۰/۷۰،۰/۸۴)	۰/۰۴۹۵۴۸	۶/۱۵۸۱e − ۰۵	۱/۱۹۳
$\hat{\theta}_{(c)}$	۶۰	(۰/۲۲،۰/۳۹،۰/۷۰،۰/۸۴)	۰/۰۴۹۱۹۴	۶/۰۶۳۱e − ۰۵	۱
$\hat{\theta}_{(b)}$	۶۰	(۰/۲۲،۰/۳۹،۰/۷۰،۰/۸۴)	۰/۰۴۸۸۲۲	۶/۳۱۷۰e − ۰۵	۰/۹۶۰
$\hat{\theta}_{(3)}$	۶۰	(۰/۲۲،۰/۳۹،۰/۷۰،۰/۸۴)	۰/۰۴۹۵۷۹	۵/۰۷۹۱e − ۰۵	۱/۱۹۴

جدول ۹. مقایسه عملکرد براوردگرها به ازای $t_1 = 5, t_2 = 10, t_3 = 24, t_4 = 36, k = 4, \frac{\pi}{۲\theta} = 22, \theta = 0/07$

$\hat{\theta}$	n	$(\bar{p}_1, \bar{p}_2, \bar{p}_3, \bar{p}_4)$	$Mean$	$MSE(\hat{\theta})$	$E(\hat{\theta}_{(c)}, \hat{\theta})$
$\hat{\theta}_{(c)}$	۳۰	(۰/۳۰،۰/۵۰،۰/۸۱،۰/۹۲)	۰/۰۶۶۹۹۳	۱/۷۸۳۴ − ۰۴	۱
$\hat{\theta}_{(b)}$	۳۰	(۰/۳۰،۰/۵۰،۰/۸۱،۰/۹۲)	۰/۰۶۴۱۰۷	۲/۱۳۴۲e − ۰۴	۰/۸۳۵
$\hat{\theta}_{(3)}$	۳۰	(۰/۳۰،۰/۵۰،۰/۸۱،۰/۹۲)	۰/۰۶۷۴۲۸	۱/۵۶۳۸e − ۰۴	۱/۱۴۰
$\hat{\theta}_{(c)}$	۴۰	(۰/۳۰،۰/۵۰،۰/۸۱،۰/۹۲)	۰/۰۶۷۹۲۴	۱/۳۹۱۱e − ۰۴	۱
$\hat{\theta}_{(b)}$	۴۰	(۰/۳۰،۰/۵۰،۰/۸۱،۰/۹۲)	۰/۰۶۵۷۴۰	۱/۸۲۳۷e − ۰۴	۰/۷۶۳
$\hat{\theta}_{(3)}$	۴۰	(۰/۳۰،۰/۵۰،۰/۸۱،۰/۹۲)	۰/۰۶۸۲۲۰	۱/۲۵۲۱e − ۰۴	۱/۱۱۱
$\hat{\theta}_{(c)}$	۵۰	(۰/۳۰،۰/۵۰،۰/۸۱،۰/۹۲)	۰/۰۶۸۰۶۰	۱/۱۳۲۹e − ۰۴	۱
$\hat{\theta}_{(b)}$	۵۰	(۰/۳۰،۰/۵۰،۰/۸۱،۰/۹۲)	۰/۰۶۶۴۳۸	۱/۵۳۲۷e − ۰۴	۰/۷۳۹
$\hat{\theta}_{(3)}$	۵۰	(۰/۳۰،۰/۵۰،۰/۸۱،۰/۹۲)	۰/۰۶۸۳۵۳	۱/۰۲۱۸e − ۰۴	۱/۱۰۹
$\hat{\theta}_{(c)}$	۶۰	(۰/۳۰،۰/۵۰،۰/۸۱،۰/۹۲)	۰/۰۶۸۴۶۶	۹/۴۷۲۸e − ۰۵	۱
$\hat{\theta}_{(b)}$	۶۰	(۰/۳۰،۰/۵۰،۰/۸۱،۰/۹۲)	۰/۰۶۷۲۱۷	۱/۳۰۷۸ − ۰۴	۰/۷۲۴
$\hat{\theta}_{(3)}$	۶۰	(۰/۳۰،۰/۵۰،۰/۸۱،۰/۹۲)	۰/۰۶۸۷۲۹	۸/۵۶۰۰e − ۰۵	۱/۱۰۷

۴. تعیین توزیع مجانبی

با توجه به نتایج بدست آمده در بخش قبلی، مقایسه عملکرد $\hat{\theta}_{(c)}\hat{\theta}_{(b)}$، و $\hat{\theta}_{(3)}$ و همچنین با در نظر گرفتن کم براوردی نرخ خطر و کارایی براوردگرها، ملاحظه می‌شود که $\hat{\theta}_{(3)}$ عملکرد بهتری دارد. از این رو در این بخش توزیع مجانبی $\hat{\theta}_{(3)}$ به دست آورده می‌شود. برای سهولت در انجام محاسبات، از این پس به جای $F(t;\theta)$ و $R(t;\theta)$ به ترتیب از نمادهای مختصر $F(t)$ و $R(t)$ استفاده می‌شود.

لم ۱.۴. *اگر $W_i = \sum_{r=1}^{i} Y_r$، آن‌گاه بردار تصادفی $\boldsymbol{W} = (W_i)_{k\times1}; i = 1, \dots, k$ به ازای n های بزرگ دارای توزیع نرمال k متغیره است که بردار میانگین و ماتریس کوواریانس آن به ترتیب به صورت*

$$E(\mathbf{W}) \equiv \boldsymbol{\mu} = (\mu_i)_{k\times1};\;\; \mu_i = nF(t_i)$$

$$Cov(\mathbf{W}) \equiv \mathbf{\Sigma} = (\sigma_{ij})_{k\times k};$$
$$\sigma_{ij} = nF(t_{Min\{i,j\}})R(t_{Max\{i,j\}})$$

می‌باشد.

با توجه به این‌که $(Y_1, Y_2, \dots, Y_k)$ دارای توزیع چندجمله‌ای است،

$$P(Y_1 = y_1, \dots, Y_k = y_k) = \frac{n!}{y_1!\dots y_k!(n-\sum_{r=1}^{k} y_r)!} \times \vartheta_1^{y_1} \dots \vartheta_k^{y_k}(1 - \sum_{r=1}^{k} \vartheta_r)^{n-\sum_{r=1}^{k} y_r}$$

که در آن

$$0 \le y_1, \dots, y_k \le n, \vartheta_r = F(t_r) - F(t_{r-1}); \quad r = 1, \dots, k.$$

بنا بر این $W_i = \sum_{r=1}^{i} Y_r$ تعداد واحدهای خراب شده تا t_i و دارای توزیع دوجمله‌ای با پارامترهای n و $F(t_i)$ خواهد بود و توزیع مجانبی آن نیز به صورت نرمال با میانگین $nF(t_i)$ و واریانس $nF(t_i)R(t_i)$ است. همچنین برای $i < j$:

$$Cov(W_i, W_j) = Cov\left(\sum_{r=1}^{i} Y_r, \sum_{r=1}^{j} Y_r\right) = Cov\left(\sum_{r=1}^{i} Y_r, \sum_{r=1}^{i} Y_r + \sum_{r=i+1}^{j} Y_r\right) = Var(\sum_{r=1}^{i} Y_r) + Cov(\sum_{r=1}^{i} Y_r, \sum_{r=i+1}^{j} Y_r)$$

با توجه به این‌که $\sum_{r=i+1}^{j} Y_r$ دارای توزیع دوجمله‌ای با پارامترهای n و $F(t_j) - F(t_i)$ است و با توجه به مقدار کوواریانس مؤلفه‌ها در توزیع چندجمله‌ای، برای $i < j$ خواهیم داشت:

$$Cov(W_i, W_j) = nF(t_i)R(t_i) - nF(t_i)\left(F(t_j) - F(t_i)\right) = nF(t_i)R(t_j)$$

از این رو

$$Cov(W_i, W_j) = nF(t_{Min\{i,j\}})R(t_{Max\{i,j\}}); \quad i, j = 1, \dots, k$$

و همچنین نتیجه مطلوب برقرار است.

قضیه ۲.۴. *برای n های بزرگ*

$$\sqrt{n}(\hat{\theta}_{(3)} - \theta) \mathrel{\dot\sim} N(0, \sigma^2)$$

که در آن

$$\sigma^2 = \frac{1}{(\sum_{r=1}^{k} F(t_r))^2}$$

$$\times \sum_{i=1}^{k}\sum_{j=1}^{k} \frac{F^2(t_{Min\{i,j\}})F(t_{Max\{i,j\}})}{t_i t_r R(t_{Min\{i,r\}})}$$

اثبات: *$\hat{\theta}_{(3)}$ را می‌توان به صورت*

$$\hat{\theta}_{(3)} = \frac{\sum_{j=1}^{k}\sum_{r=1}^{j} Y_r \hat{\theta}_j}{\sum_{j=1}^{k}\sum_{r=1}^{j} Y_r} = \frac{\sum_{j=1}^{k} W_j \hat{\theta}_j}{\sum_{j=1}^{k} W_j} = \frac{\sum_{j=1}^{k} B_j \hat{\theta}_j}{\sum_{j=1}^{k} B_j} = \frac{\sum_{j=1}^{k} B_j (\frac{-Ln(1-B_j)}{t_j})}{\sum_{j=1}^{k} B_j} \quad (۲)$$

نوشت که در آن

$$\mathbf{B} = (B_i)_{k\times 1} = \frac{1}{n}\mathbf{W} \text{ و } B_i = \frac{W_i}{n}; i = 1, \dots, k.$$

از این رو با توجه به لم ۱، توزیع مجانبی بردار $\mathbf{B}$ نیز به صورت نرمال k متغیره با بردار میانگین و ماتریس کوواریانس

$$E(\mathbf{B}) \equiv \boldsymbol{\mu} = (\mu_i)_{k\times 1}; \; \mu_i = F(t_i) \quad (۳)$$

$$Cov(\mathbf{B}) \equiv \mathbf{\Sigma} = (\sigma_{ij})_{k\times k}; \quad \sigma_{ij} = \frac{F\left(t_{Min\{i,j\}}\right)R\left(t_{Max\{i,j\}}\right)}{n} \quad (۴)$$

است. با توجه به (۲)، $\hat{\theta}_{(3)}$ به صورت $\hat{\theta}_{(3)} \equiv g(\mathbf{B})$، یعنی تابعی از بردار تصادفی $\mathbf{B}$ است. بنا بر این با توجه به (۳)، (۴) و به کارگیری روش دلتا در حالت چندمتغیره، $\hat{\theta}_{(3)}$ دارای توزیع مجانبی نرمال است که میانگین آن $g(E(\mathbf{B})) = g((F(t_i))_{k\times 1}) = \theta$ و واریانس آن نیز به صورت $(\nabla g(E(\mathbf{B})))^T Cov(\mathbf{B}) \nabla g(E(\mathbf{B}))$ می‌باشد که در آن $\nabla g(E(\mathbf{B}))$ ژاکوبین تابع g است [۱۱]. از این رو

$$\nabla g(\mathbf{B}) = (\frac{\partial}{\partial B_i}\{g(\mathbf{B})\})_{k\times 1};$$
$$\frac{\partial}{\partial B_i}\{g(\mathbf{B})\} = \frac{C}{(\sum_{i=1}^{k} B_i)^2}$$

و

$$C = (\textstyle\sum_{i=1}^{k} B_i) \times \left(-\frac{Ln(1-B_i)}{t_i} + B_i \frac{\partial}{\partial B_i}\{-Ln(1-B_i)\}\right) - \sum_{i=1}^{k} B_i \left(\frac{-Ln(1-B_i)}{t_i}\right)$$

که با جایگذاری $E(\mathbf{B})$ به جای $\mathbf{B}$،

$$\frac{\partial}{\partial B_i}\{g(E(\mathbf{B}))\} = \frac{\left(\theta + \frac{F(t_i)}{t_i R(t_i)} - \theta\right)}{\sum_{i=1}^{k} F(t_i)} = \frac{F(t_i)}{t_i R(t_i) \sum_{i=1}^{k} F(t_i)}$$

و

$$(\nabla g(E(\mathbf{B})))^T Cov(\mathbf{B}) \nabla g(E(\mathbf{B})) = \frac{1}{(\sum_{i=1}^{k} F(t_i))^2} \times \mathbf{A}\Sigma\mathbf{A}^T$$

را خواهیم داشت که در آن

$$\mathbf{A} = (a_j)_{1\times k} = (\frac{F(t_j)}{t_j R(t_j)})_{1\times k},$$

$$\Sigma = (\sigma_{ij})_{k\times k} = \left(\frac{F(t_{Min\{i,j\}})R(t_{Max\{i,j\}})}{n}\right)_{k\times k}.$$

بنا بر این

$$\Sigma\mathbf{A}^T \equiv \mathbf{D} = (d_i)_{k\times 1};$$

$$d_i = \sum_{r=1}^{k} \frac{F(t_{Min\{i,r\}})R(t_{Max\{i,r\}})}{n} \frac{F(t_r)}{t_r R(t_r)}$$

و

$$\mathbf{A}\Sigma\mathbf{A}^T = \sum_{i=1}^{k} \frac{F(t_i)}{t_i R(t_i)} \times \sum_{r=1}^{k} \frac{F(t_{Min\{i,r\}})R(t_{Max\{i,r\}})}{n} \frac{F(t_r)}{t_r R(t_r)}$$
$$= \frac{1}{n}\sum_{i=1}^{k} \frac{F(t_i)}{t_i R(t_i)} \times \sum_{r=1}^{i} \frac{F(t_{Min\{i,r\}})R(t_{Max\{i,r\}})F(t_r)}{t_r R(t_r)} + \frac{1}{n}\sum_{i=1}^{k} \frac{F(t_i)}{t_i R(t_i)} \times \sum_{r=i+1}^{k} \frac{F(t_{Min\{i,r\}})R(t_{Max\{i,r\}})F(t_r)}{t_r R(t_r)}$$
$$= \frac{1}{n}\sum_{i=1}^{k} \frac{F(t_i)}{t_i R(t_i)} \sum_{r=1}^{i} \frac{F(t_r)R(t_i)F(t_r)}{t_r R(t_r)} + \frac{1}{n}\sum_{i=1}^{k} \frac{F(t_i)}{t_i R(t_i)} \sum_{r=i+1}^{k} \frac{F(t_i)R(t_r)F(t_r)}{t_r R(t_r)}$$
$$= \frac{1}{n}\sum_{i=1}^{k}\sum_{r=1}^{i} \frac{F(t_i)F^2(t_r)}{t_i t_r R(t_r)} + \frac{1}{n}\sum_{i=1}^{k}\sum_{r=i+1}^{k} \frac{F^2(t_i)F(t_r)}{t_i t_r R(t_i)}$$
$$= \frac{1}{n}\sum_{i=1}^{k}\sum_{r=1}^{i} \frac{F^2(t_{Min\{i,r\}})F(t_{Max\{i,r\}})}{t_i t_r R(t_{Min\{i,r\}})} + \frac{1}{n}\sum_{i=1}^{k}\sum_{r=i+1}^{k} \frac{F^2(t_{Min\{i,r\}})F(t_{Max\{i,r\}})}{t_i t_r R(t_{Min\{i,r\}})}$$
$$= \frac{1}{n}\sum_{i=1}^{k}\sum_{r=1}^{k} \frac{F^2(t_{Min\{i,r\}})F(t_{Max\{i,r\}})}{t_i t_r R(t_{Min\{i,r\}})}$$

نتیجه ۳.۴. *با توجه به لم ۱، به کارگیری روش دلتا و نیز انجام محاسبات می‌توان نشان داد*

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \theta\mathbf{1}) \sim N_k(\mathbf{0}, \Sigma)$$

که در آن

$$\hat{\boldsymbol{\theta}} = (\hat{\theta}_i)_{k\times 1},$$

$$\Sigma = (\sigma_{ij})_{k\times k} = \left(\frac{F\left(t_{Min\{i,j\}}\right)}{t_{Min\{i,j\}} t_{Max\{i,j\}} R\left(t_{Min\{i,j\}}\right)}\right)_{k\times k}$$

نتیجه ۴.۴. *هرگاه تعداد دفعات نظارت $k=2$ باشد آنگاه با توجه به قضیه ۲*

$$\sqrt{n}\left(\frac{Y_1\hat{\theta}_1 + (Y_1+Y_2)\hat{\theta}_2}{2Y_1+Y_2} - \theta\right) \sim N(0,\sigma^2),$$

که در آن $\sigma^2 = \dfrac{\frac{F^3(t_1)}{t_1^2 R(t_1)} + 2\frac{F^2(t_1)F(t_2)}{t_1 t_2 R(t_1)} + \frac{F^3(t_2)}{t_2^2 R(t_2)}}{n(F(t_1)+F(t_2))^2}$.

نتیجه ۵.۴. *هرگاه تعداد دفعات نظارت $k=2$ باشد، یک فاصله اطمینان مجانبی $100(1-\alpha)\%$ برای θ به صورت*

$$\hat{\theta}_{(3)} \pm \frac{z_{\frac{\alpha}{2}}}{n(2y_1+y_2)}\sqrt{\frac{y_1^3}{t_1^2(1-\frac{y_1}{n+1})} + \frac{2y_1^2(y_1+y_2)}{t_1 t_2(1-\frac{y_1}{n+1})} + \frac{(y_1+y_2)^3}{t_2^2(1-\frac{y_1+y_2}{n+1})}}$$

است که در آن $\hat{\theta}_{(3)} = \frac{y_1\hat{\theta}_1 + (y_1+y_2)\hat{\theta}_2}{2y_1+y_2}$.

۵. بحث و نتیجه‌گیری

در این پژوهش، برای براورد نرخ خطر توزیع نمایی در داده‌های تحت سانسور بازه‌ای، سه روش $\hat{\theta}_{(1)}$، $\hat{\theta}_{(2)}$ و $\hat{\theta}_{(3)}$ معرفی و بر پایه مطالعات شبیه‌سازی و در نظر گرفتن شاخص‌های کم براوردی نرخ خطر و کارایی با دو روش $\hat{\theta}_{(b)}$ و $\hat{\theta}_{(c)}$ که مورد استفاده پژوهشگران هستند، مقایسه گردید. با توجه به این‌که در بررسی‌های مقدماتی، $\hat{\theta}_{(3)}$ در مقایسه با دو روش معرفی شده دیگر عملکرد بهتری از خود نشان داد، در بررسی‌های بعدی فقط این روش با دو روش $\hat{\theta}_{(b)}$ و $\hat{\theta}_{(c)}$ مقایسه گردید تا از حجیم شدن مقایسه‌ها پرهیز شود. همان‌طور که نتایج بدست آمده نشان می‌دهند، کم براوردی $\hat{\theta}_{(3)}$ نسبت به $\hat{\theta}_{(b)}$ و $\hat{\theta}_{(c)}$ کم‌تر است و هم این‌که با افزایش تعداد دفعات نظارت، میزان کم براوردی آن نیز کاهش می‌یابد. با توجه به این‌که کم براوردی نرخ خطر از نقاط ضعف یک براوردگر به شمار می‌رود و کم براوردی نرخ خطر در توزیع نمایی نیز منجر به بیش براوردی میانگین توزیع می‌شود، عملکرد $\hat{\theta}_{(3)}$ را می‌توان بهتر از سایر روش‌ها دانست. علاوه بر آن، همان‌طور که نتایج شبیه‌سازی نشان می‌دهند، کارایی این روش نیز نسبت به سایر روش‌ها بیش‌تر است که این دو ویژگی بیانگر عملکرد خوب و مزیت‌های این روش می‌باشد و لذا توزیع مجانبی آن نیز به دست آورده شد.

همان‌طور که در جدول‌ها ملاحظه می‌شود، سومین براوردگر وزنی

عملکرد بهتری دارد هرگاه زمان‌های نظارت در فاصله $(0,\frac{\pi}{۲\theta})$ باشند. با توجه به این‌که در سانسور بازه‌ای بایستی زمان‌های نظارت مقادیری از پیش تعیین شده و مستقل از اطلاعات به‌دست آمده نمونه‌ای باشند، نمی‌توان زمان‌های نظارت را کم‌تر از کران بالای این فاصله و یا براورد آن در نظر گرفت. با این حال، اگر زمان‌های نظارت کم‌تر از $\frac{\pi}{۲\theta}$ باشند، $F(t_k;\theta) \le F(\frac{\pi}{۲\theta};\theta) \simeq 0.8$ را خواهیم داشت و انتظار می‌رود که تا آخرین زمان نظارت تقریباً 80% از واحدهای مورد آزمایش خراب شوند. در جدول‌ها مقادیر $\bar{p}_1,\dots,\bar{p}_4$ محاسبه شده‌اند و همان‌طور که انتظار می‌رود، این مقادیر تقریباً برابر با $F(t_1;\theta),\dots,F(t_4;\theta)$ هستند. با بررسی مجدد جدول‌ها نیز ملاحظه می‌شود که هرگاه تعداد دفعات نظارت زیاد و هم این‌که متوسط نسبت خرابی‌ها تا آخرین زمان نظارت تقریباً کم‌تر از 80% باشد، سومین براوردگر وزنی عملکرد بسیار بهتری از خود نشان می‌دهد. برای مثال همان‌طور که در جدول شماره ۷ ملاحظه می‌شود، آخرین زمان نظارت برابر با صدک ۶۶ -ام است و میانگین نسبت واحدهای خراب شده تا این زمان نیز برابر با ۰/۶۶ است و کارایی $\hat{\theta}_{(3)}$ نسبت به $\hat{\theta}_{(c)}$ بیش‌تر از ۱/۳۲ و همچنین کم براوردی آن نسبت به دو روش دیگر کم‌تر می‌باشد. از این رو می‌توان نتیجه گرفت که در کارهای عملی، اگر تعداد دفعات نظارت زیاد و تا آخرین زمان نظارت نیز ۸۰ − ۷۰ درصد از واحدهای آزمایشی خراب شده باشد آن‌گاه انتظار می‌رود که زمان‌های نظارت تعیین شده در فاصله $(0,\frac{\pi}{۲\theta})$ باشند. در این صورت، زمینه ارتقا عملکرد سومین براوردگر وزنی بیش‌تر فراهم می‌شود و صرفه‌جویی در هزینه و مدت زمان انجام آزمایش که یکی از اهداف به کارگیری انواع طرح‌های سانسور است، تحقق می‌یابد.

۶. اعلام تعارض منافع

نویسندگان اعلام می‌کنند که هیچ نوع تعارض منافعی وجود ندارد.

۷. مراجع

[1] Cohen, A.C. (1991), Truncated and Censored Samples: Theory and Application, Marcel Dekker. New York.

[2] Rausand, M. and Høyland, A. (2004), System Reliability Theory: Models, Statistical Methods and Applications, Wiley-Interscience. Hoboken, NJ.

[3] Meeker, W.Q. (1986), Planning life tests in which units are inspected for failure. *IEEE Transactions on Reliability*, 35, 571–578.

[4] Zhang, C.W., Zhang, T., Xu, D. and Xie, M. (2013), Analyzing highly censored reliability data without exact failure times: An efficient tool for practitioners. *Quality Engineering*, 25, 392–400.

[5] Chen, D.G. and Lio, Y.L. (2010), Parameter estimations for generalized exponential distribution under progressive type-I interval censoring. *Computational Statistics & Data Analysis*, 54, 1581–1591.

[6] Sajadi, F., Poursaeed, M. (2021), Improving the parameter estimation under interval-censored exponential lifetimes. *Journal of Statistical Modelling: Theory and Applications*, 2, 63-78.

[7] Nanasi, T. (2014), Interval censored data analysis with Weibull and exponential distribution. *Applied Mechanics and Materials*, 693, 74-79.

[8] EL-Shaarawi, A.H. and Naderi, A. (1991), Statistical inference from multiply censored environmental data. *Environmental Monitoring and Assessment*, 17, 339–347.

[9] Malevich, N. and Müller, C.H. (2019), Optimal design of inspection times for interval censoring. *Statistical Papers*, 60, 449–464.

[10] Hewett, P., Ganser GH. (2007), A comparison of several methods for analyzing censored data. *Ann Occup Hyg.*, 51, 611-32.

[11] Casella, G., Berger, R. L. (2002), Statistical inference. 2nd ed. , Pacific Grove, Australia .

Methods for Estimating Parameter of Exponential Distribution in Interval Censored Data; Comparison Study

Nader Asadian
Assistant Professor, Lorestan University. Khorramabad. Iran

Mohammad Hossein Poursaeed[1]
Associate Professor, Lorestan University. Khorramabad. Iran. poursaeed.m@lu.ac.ir

Abstract: In this research, new estimators for the hazard rate of exponential distribution in interval censored data are compared with some existing methods based on simulation studies. Then the asymptotic distribution of one of the introduced methods that has better performance is obtained.
Key words: Exponential distribution, Delta method, Interval censoring, Pseudo-likelihood function

1. Introduction

Hazard rate is one of the indicators which is studied in the theory of reliability, survival analysis and system performance review, and one of the lifetime distributions which is widely used in industrial and medical data analysis is exponential distribution.
Let $X_1, X_2, \dots, X_n$ be a random sample of n experimental units with exponential distribution, hazard rate θ and density function $f(x;\theta) = \theta exp(-\theta x);\ x > 0, \theta > 0$. To estimate the parameter θ, the experimenter may use different censoring schemes to limit the experiment time and reduce costs. For more information about kinds of censored data refer to Cohen [1].
In some cases, where experimental units are not continuously monitored, the obtained data is related to interval censoring, in which n experimental units are inspected and monitored at k predetermined times such as $0 < t_1 < \cdots < t_k$ and the number of failures occurring between inspection times are recorded. In this scheme, the experimenter does not know the exact time of failures and only knows the number of failures which occurred between the inspection times.
Let $t_0 \equiv 0$, $(0, t_1], (t_1, t_2], \dots, (t_{k-1}, t_k]$ are the intervals between monitoring times, and $y_i;\ i = 1,2, \dots, k$ are the number of failures in the time interval $(t_{i-1}, t_i]$, then the maximum likelihood estimator of θ does not have a closed form. In the method based on the pseudo-likelihood function, which is called substitution method, the estimator of parameter is given by

$$\hat{\theta}_{(a)} = \frac{\sum_{i=1}^{k} y_i}{\sum_{i=1}^{k} t_i^* + (n - \sum_{i=1}^{k} y_i) t_k}; \qquad t_i^* \in \left\{t_{i-1}, t_i, \frac{t_{i-1}+t_i}{2}\right\}, i = 1, \dots, k.$$

Meeker [4] has obtained the asymptotic distribution of the parameter estimator by considering the midpoint of the intervals. One of the parameter estimation methods in a special case of interval censored data is

$$\hat{\theta} = \frac{-\ln(1 - \frac{y}{n})}{t}$$

which is introduced by Zhang et al. [6], where y is the number of failures until one monitoring time t. Therefore,

$$\hat{\theta}_{(b)} = \frac{-\ln(1 - \sum_{i=1}^{k} y_i / n)}{t_k}$$

can be considered as a version which is related to k monitoring times. Based on the probability plot method,

---

[1] Corresponding Author :poursaeed.m@lu.ac.ir

$$\hat{\theta}_{(c)} = \frac{1}{k}\sum_{j=1}^{k} \frac{-\ln(1 - \sum_{r=1}^{j} y_r/n)}{t_j}$$

has been introduced for parameter estimation by Chen et al. [2]. Also, to improve the estimators $\hat{\theta}_{(a)}$, $\hat{\theta}_{(b)}$ and $\hat{\theta}_{(c)}$, some methods have been introduced that use parameter re-estimation and substitution method [5].

In this study, three estimators $\hat{\theta}_{(a)}$, $\hat{\theta}_{(b)}$ and $\hat{\theta}_{(c)}$ are considered as the existing methods used by researchers. Since, using $\hat{\theta}_{(a)}$ leads to an underestimation of the hazard rate, the researchers use this method rarely [3]. In addition, the maximum likelihood estimator does not have a closed form and its distribution is not determined. Therefore, researchers use the estimators $\hat{\theta}_{(b)}$ and $\hat{\theta}_{(c)}$ in making statistical inferences.

Since, the two existing methods $\hat{\theta}_{(b)}$ and $\hat{\theta}_{(c)}$, have better performance than $\hat{\theta}_{(a)}$, therefore, those are compared with the three new ones, $\hat{\theta}_{(1)}$, $\hat{\theta}_{(2)}$ and $\hat{\theta}_{(3)}$. Moreover, the asymptotic distribution of one of the new methods that performs better than others will be obtained.

2- Research Method

Suppose that n experimental units under interval censoring have exponential distribution with distribution function

$$F(x;\theta) = 1 - \exp(-\theta x); \quad x > 0, \theta > 0,$$

And experimental units are inspected and monitored at k predetermined times $t_1 < \cdots < t_k$. If $(0, t_1], (t_1, t_2], \ldots, (t_{k-1}, t_k]$ are the intervals between monitoring times and $y_i;\ i = 1,2, \ldots, k$ are the number of failures in the time interval $(t_{i-1}, t_i]$, then in view of relationship between the parameter and the exponential distribution function,

$$\hat{\theta}_j = \frac{-\ln\left(1 - \frac{\sum_{i=1}^{j} Y_i}{n}\right)}{t_j}; \ j = 1, \ldots, k,$$

are different estimators for the parameter θ. It can be shown that they are asymptotically unbiased and have different variances. In addition, the existing estimators $\hat{\theta}_{(b)}$ and $\hat{\theta}_{(c)}$ are also functions of these estimators. The method introduced by Chen et al. [2] is actually the same as $\hat{\theta}_k$ and probability plot method is also the average of $\hat{\theta}_j$. In fact, probability plot method is a special case of the weighted estimator $\sum_{j=1}^{k} \alpha_j \hat{\theta}_j$ which considers the same coefficients for components ($\hat{\theta}_j$ s) with different variances, which cannot be reasonable.

Therefore, in the new methods introduced in this paper, different weights are considered for $\hat{\theta}_j$s. As the first weighted estimator, we consider a linear combination of $\hat{\theta}_j$s as $\hat{\theta}_{(1)} = \sum_{j=1}^{k} \alpha_j \hat{\theta}_j$ so that it is asymptotically unbiased and has the lowest variance. It can be shown that for large n, $\hat{\theta}_j$s have a normal distribution with mean θ and

$$Cov(\theta_i, \theta_j) = \frac{F(t_{Min\{i,j\}};\theta)}{n t_i t_j R(t_{Max\{i,j\}};\theta)}; \quad i, j = 1, \ldots, k,$$

where $R(x;\theta) = 1 - F(x;\theta) = \exp(-\theta x)$. By applying the asymptotic unbiased condition of the estimator and some calculations, we have $\sum_{r=1}^{k} \alpha_r = 1$, and

$$\alpha_r = \frac{1}{Var(\hat{\theta}_k - \hat{\theta}_r)} \times [Cov(\hat{\theta}_k, \hat{\theta}_k - \hat{\theta}_r) - \sum_{j=1, j\neq r}^{k-1} \alpha_j Cov(\hat{\theta}_k - \hat{\theta}_r, \hat{\theta}_k - \hat{\theta}_j)]; \quad r = 1, \ldots, k-1.$$

Since the weighted coefficients are functions of the unknown parameter, their estimators which have complex forms, are used. For example, in the case of $k = 2$, $\alpha_1 = \frac{Cov(\hat{\theta}_2, \hat{\theta}_2 - \hat{\theta}_1)}{Var(\hat{\theta}_2 - \hat{\theta}_1)}$, its estimator is

$$\frac{\frac{y_2}{t_2^2(1-\frac{y_2}{n})} - \frac{y_1}{t_1 t_2(1-\frac{y_2}{n})}}{\frac{y_2}{t_2^2(1-\frac{y_2}{n})} - \frac{2y_1}{t_1 t_2(1-\frac{y_2}{n})} + \frac{y_1}{t_1^2(1-\frac{y_1}{n})}}.$$

The second weighted estimator is introduced as follows

$$\hat{\theta}_{(2)} = \frac{\sum_{j=1}^{k} \frac{1}{\hat{V}ar(\hat{\theta}_j)} \hat{\theta}_j}{\sum_{j=1}^{k} \frac{1}{\hat{V}ar(\hat{\theta}_j)}} = \frac{\sum_{j=1}^{k} \frac{t_j^2 (1-\sum_{i=1}^{j} Y_i/n)}{\sum_{i=1}^{j} Y_i} \hat{\theta}_j}{\sum_{j=1}^{k} \frac{t_j^2 (1-\sum_{i=1}^{j} Y_i/n)}{\sum_{i=1}^{j} Y_i}},$$

with simpler form of coefficients than the first method.

Moreover, the third weighted estimator is established by $\hat{\theta}_{(3)} = \frac{\sum_{j=1}^{k} \sum_{r=1}^{j} Y_r \hat{\theta}_j}{\sum_{j=1}^{k} \sum_{r=1}^{j} Y_r}$, so that smaller weight coefficients are attributed to $\hat{\theta}_j$s with larger variance. In addition, it has a simpler form than the forms of the previous two weighted methods.

In order to compare the two former estimators $\hat{\theta}_{(b)}$ and $\hat{\theta}_{(c)}$ as well as three new weighted estimators, a simulation study is presented. If all failures have occurred by time t_k, then none of the estimators will be definable. Therefore, instead of $\hat{\theta}_j = \frac{-ln(1-\sum_{i=1}^{j} Y_i/n)}{t_j}$, the modified form $\hat{\theta}_j = \frac{-ln(1-\sum_{i=1}^{j} Y_i/n+1)}{t_j}$ will be used. In simulation, to avoid making the comparisons bulky, we consider few values for the parameter and the number of monitoring times as $\theta = 0.03,0.05,0.07$ and $k = 2,3,4$. For the sample size, relatively large values of $n = 30,40,50,60$ are considered, because the distribution and asymptotic behavior of the estimators under review are mentioned and cited in this study.

Based on the simulation with N=10,000 repetitions, the mean of the estimators, the mean square of their error, as well as $E(\hat{\theta}_{(c)}, \hat{\theta})$ (i.e. the efficiency of the estimators) compared to $\hat{\theta}_{(c)}$ are calculated. In addition, in order to investigate the effect of the parameter value and monitoring times on the estimators, $\frac{\pi}{2\theta}$ and $\bar{p}_1, \dots, \bar{p}_4$ are calculated, where $\bar{p}_i; i = 1, \dots, 4$ is the average of failures ratio up to the $i - th$ monitoring time in 10,000 repeated times.

3. Research findings

Since, the underestimation of the hazard rate is considered one of the weaknesses of an estimator, and the underestimation of the hazard rate in the exponential distribution causes an overestimation of the average, thus in addition to the efficiency of the estimators, special attention should also be paid to this criterion.

As the initial simulation results show, $\hat{\theta}_{(1)}$ does not work well, because their estimates are used instead of weighted coefficients. Given that if n is a large number and the last monitoring time is equal to $\frac{\pi}{2\theta}$, then $\hat{\theta}_k$ will have the lowest possible variance [4]. Hence, the values of the variances is a decreasing function whenever t_k is in $(0, \frac{\pi}{2\theta})$ and it can be concluded that if the monitoring times $t_1 < \cdots < t_k$ are chosen in the interval $(0, \frac{\pi}{2\theta})$, then $Var(\hat{\theta}_1) > Var(\hat{\theta}_2) > \cdots > Var(\hat{\theta}_k)$. Also, $\hat{\theta}_{(3)}$ performs better than $\hat{\theta}_{(1)}$ and $\hat{\theta}_{(2)}$, especially when the monitoring times are selected in $(0, \frac{\pi}{2\theta})$. Next, the two methods $\hat{\theta}_{(b)}$ and $\hat{\theta}_{(c)}$ were compared with $\hat{\theta}_{(3)}$ that has better performance. Based on further investigations and as seen in the table below, $\hat{\theta}_{(3)}$ has a better performance. Hence, the asymptotic distribution of $\hat{\theta}_{(3)}$ is obtained.

Table 1. Comparing the estimators for
$\theta = 0.03, k = 4, \frac{\pi}{2\theta} = 52\,, t_1 = 5, t_2 = 10, t_3 = 24, t_4 = 36$

$\hat{\theta}$	n	$(\bar{p}_1, \bar{p}_2, \bar{p}_3, \bar{p}_4)$	$Mean$	$MSE(\hat{\theta})$	$E(\hat{\theta}_{(c)}, \hat{\theta})$
$\hat{\theta}_{(c)}$	30	(0.14,0.26,0.51,0.66)	0.029285	$6.4242e-05$	1
$\hat{\theta}_{(b)}$	30	(0.14,0.26,0.51,0.66)	0.029121	$4.7521e-05$	1.352
$\hat{\theta}_{(3)}$	30	(0.14,0.26,0.51,0.66)	0.030056	$4.8160e-05$	1.334
$\hat{\theta}_{(c)}$	40	(0.14,0.26,0.51,0.66)	0.029485	$4.7158e-05$	1
$\hat{\theta}_{(b)}$	40	(0.14,0.26,0.51,0.66)	0.029257	$3.5985e-05$	1.310
$\hat{\theta}_{(3)}$	40	(0.14,0.26,0.51,0.66)	0.030007	$3.5582e-05$	1.325
$\hat{\theta}_{(c)}$	50	(0.14,0.26,0.51,0.66)	0.029568	$3.8381e-05$	1
$\hat{\theta}_{(b)}$	50	(0.14,0.26,0.51,0.66)	0.029355	$2.9387e-05$	1.306
$\hat{\theta}_{(3)}$	50	(0.14,0.26,0.51,0.66)	0.029985	$2.9054e-05$	1.321
$\hat{\theta}_{(c)}$	60	(0.14,0.26,0.51,0.66)	0.029613	$3.2346e-05$	1
$\hat{\theta}_{(b)}$	60	(0.14,0.26,0.51,0.66)	0.029580	$2.4114e-05$	1.341
$\hat{\theta}_{(3)}$	60	(0.14,0.26,0.51,0.66)	0.030032	$2.4057e-05$	1.345

Theorem 2.1. For large n, $\sqrt{n}(\hat{\theta}_{(3)} - \theta) \dot{\sim} N(0, \sigma^2)$, where in
$$\sigma^2 = \frac{1}{(\sum_{r=1}^{k} F(t_r))^2} \times \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{F^2(t_{Min\{i,j\}}) F(t_{Max\{i,j\}})}{t_i t_r R(t_{Min\{i,r\}})}$$

Result 2.2. By applying the delta method and some of calculations, it can be shown that
$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \theta\mathbf{1}) \dot{\sim} N_k(\mathbf{0}, \Sigma)$$
where in
$$\hat{\boldsymbol{\theta}} = (\hat{\theta}_i)_{k\times 1}, \Sigma = (\sigma_{ij})_{k\times k} = \left(\frac{F(t_{Min\{i,j\}})}{t_{Min\{i,j\}} t_{Max\{i,j\}} R(t_{Min\{i,j\}})}\right)_{k\times k}.$$

Result 2.3. If the number of monitoring times is k=2, then

$$\sqrt{n}\left(\frac{Y_1\hat{\theta}_1 + (Y_1 + Y_2)\hat{\theta}_2}{2Y_1 + Y_2} - \theta\right) \dot{\sim} N(0, \sigma^2);\;\; \sigma^2 = \frac{\frac{F^3(t_1)}{t_1^2 R(t_1)} + 2\frac{F^2(t_1)F(t_2)}{t_1 t_2 R(t_1)} + \frac{F^3(t_2)}{t_2^2 R(t_2)}}{n(F(t_1) + F(t_2))^2}$$

Result 2.4. when k=2, a $\%100(1-\alpha)$ asymptotic confidence interval for θ is

$$\frac{y_1\hat{\theta}_1 + (y_1 + y_2)\hat{\theta}_2}{2y_1 + y_2} \pm \frac{z_{\frac{\alpha}{2}}}{n(2y_1 + y_2)} \sqrt{\frac{y_1^3}{t_1^2(1 - \frac{y_1}{n+1})} + \frac{2y_1^2(y_1 + y_2)}{t_1 t_2(1 - \frac{y_1}{n+1})} + \frac{(y_1 + y_2)^3}{t_2^2(1 - \frac{y_1 + y_2}{n+1})}}$$

4. Discussion and conclusion

In this research, to estimate the hazard rate of exponential distribution under interval censored data, three methods $\hat{\theta}_{(1)}$, $\hat{\theta}_{(2)}$ and $\hat{\theta}_{(3)}$ are introduced. Based on simulation studies and considering underestimate and efficiency, the new estimators were compared with the two methods $\hat{\theta}_{(b)}$ and $\hat{\theta}_{(c)}$ which are used by the researchers. Preliminary investigations showed the $\hat{\theta}_{(3)}$ has better performance compared to the other two introduced methods. Hence, in the subsequent investigations only this method was compared with the two methods $\hat{\theta}_{(b)}$ and $\hat{\theta}_{(c)}$ to avoid making the comparisons bulky.

As the obtained results show, the underestimation of $\hat{\theta}_{(3)}$ is lower than $\hat{\theta}_{(b)}$ and $\hat{\theta}_{(c)}$, and its underestimation also decreases when the number of monitoring times are increased. In addition, the efficiency of $\hat{\theta}_{(3)}$ is also higher compared to the other methods, therefore its asymptotic distribution has been obtained.
As the tables show, $\hat{\theta}_{(3)}$ performs better when the monitoring times are in $(0, \frac{\pi}{2\theta})$ and the number of monitoring times is large. On the other hand, the monitoring times should be predetermined values and independent of the information obtained, hence the monitoring times cannot be considered lower than $\frac{\pi}{2\theta}$ or than its estimator. However, if the monitoring times are less than $\frac{\pi}{2\theta}$, we have $F(t_k;\theta) \leq F(\frac{\pi}{2\theta};\theta) \simeq 0.8$. Therefore, it can be concluded that in practical works, if the number of monitoring times is large and until the last monitoring time 70-80% of the test units are damaged, then, it is expected that the monitoring times have been chosen in $(0, \frac{\pi}{2\theta})$. In this case, improving the performance of $\hat{\theta}_{(3)}$ will be provided as well as saving the cost and duration of the experiment, which both of are the goals of using various censoring schemes.

References

[1] Cohen, A.C. (1991), Truncated and Censored Samples*: Theory and Application*, Marcel Dekker. New York.

[2] Chen, D.G. and Lio, Y.L. (2010), Parameter estimations for generalized exponential distribution under progressive type-I interval censoring. *Computational Statistics & Data Analysis*, 54, 1581–1591.

[3] Hewett, P., Ganser GH. (2007), A comparison of several methods for analyzing censored data. *Ann Occup Hyg.*, 51, 611-632.

[4] Meeker, W.Q. (1986), Planning life tests in which units are inspected for failure. *IEEE Transactions on Reliability*, 35, 571–578.

[5] Sajadi, F., Poursaeed, M. (2021), Improving the parameter estimation under interval-censored exponential lifetimes. *Journal of Statistical Modelling: Theory and Applications*, 2, 63-78.

[6] Zhang, C.W., Zhang, T., Xu, D. and Xie, M. (2013), Analyzing highly censored reliability data without exact failure times: An efficient tool for practitioners. *Quality Engineering*, 25, 392–400.